European Journal of Operational Research 233 (2014) 23–33

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Continuous Optimization

On distributional robust probability functions and their computations

Man Hong Wong [a,*], Shuzhong Zhang [b,1]

[a] Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong
[b] Department of Industrial and Systems Engineering, University of Minnesota, 111 Church Street S.E., Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:
Received 21 August 2012
Accepted 26 August 2013
Available online 8 September 2013

Keywords:
Risk management
Distributional robust
Moment bounds
Semidefinite programming (SDP)
Conic programming

ABSTRACT

Consider a random vector, and assume that a set of its moments information is known. Among all possible distributions obeying the given moments constraints, the envelope of the probability distribution functions is introduced in this paper as *distributional robust probability function*. We show that such a function is computable in the bi-variate case under some conditions. Connections to the existing results in the literature and its applications in risk management are discussed as well.

1. Introduction

One of the tasks in risk management is to manage what to do in all scenarios, especially when worst comes to worst. With reference to a risk measure, a common way to describe "worst-case" is through distributional robustness, which refers to any distribution fitting some given moments, say $\mathfrak{m}_0, \mathfrak{m}_1, \ldots, \mathfrak{m}_n$. Consequently, the formulation below is often known as the *moment bound problem*:

$$(GP) \quad \sup_{x\sim(\mathfrak{m}_0,\ldots,\mathfrak{m}_n)} \mathbb{E}[\psi(x)] := \sup \ \mathbb{E}[\psi(x)]$$
$$\text{s.t.} \quad \mathbb{E}[\underbrace{x\odot x\odot\cdots\odot x}_{\#\text{ of } x=i}] = \mathfrak{m}_i, \quad i=0,\ldots,n.$$

A few words about our notations are in order here. For $x\in\mathbf{R}$, "$\odot$" denotes the usual scalar multiplication and $\mathfrak{m}_i\in\mathbf{R}$ for all i, with $\mathfrak{m}_i := \mathbb{E}[x^i]$. For $x\in\mathbf{R}^d$, "$\odot$" is the tensor multiplication (or matrix multiplication in case $d=2$) in the corresponding spaces and $\mathfrak{m}_i\in\mathbf{R}^{\overbrace{d\times\cdots\times d}^{\#\text{ of } d=i}}$. For example, if $x=[x^{(1)}, \ x^{(2)}]^T\in\mathbf{R}^2$, then $x\odot x = xx^T = \begin{pmatrix} (x^{(1)})^2 & x^{(1)}x^{(2)} \\ x^{(1)}x^{(2)} & (x^{(2)})^2 \end{pmatrix}\in\mathbf{R}^{2\times2}$, which is in the same space where $\mathfrak{m}_2$ resides. We will use the lower case letters (e.g. m) for scalars and vectors, capital letters (e.g. M) for matrices, and *fraktur* lower case letters $\mathfrak{m}$ for ambiguous (implied) dimensions.

Scarf (1958) was the first to apply this worst-case analysis in inventory management, where he took $\psi(x) = \min\{x, k\}$ for some constant k and assumed the knowledge of the first two moments. Lo (1987) and Grundy (1991) applied the similar concept for option bounds. As a matter of fact, a sizeable amount of relevant literature can be found (see e.g. Chen, He, & Zhang (2011), Cox (1991), Cox, Lin, Tian, & Zuluaga (2008), Han, Li, Sun, & Sun (2005), He, Zhang, & Zhang (2010), Jansen, Haezendonck, & Goovaerts (1986), Liu & Li (2009), De Schepper & Heijnen (2007, 2010), De Vylder & Goovaerts (1982, 1983), De Vylder (1982), Courtois & Denuit (2007), Zymler, Rustem, & Kuhn (2011), Gulpinar & Rustem (2007), Huang, Zhu, Fabozzi, & Fukushima (2011), Pena, Vera, & Zuluaga (2012)) With the recent computational developments of moment bounds, applications have been introduced in different streams in financial engineering. For example, Bertsimas and Popescu (2005) discussed the moment bounds using semidefinite programming and its relevance in probability theory (see also Bertsimas, Natarajan, & Teo, 2006; Popescu, 2007) considered the mean–covariance solutions for stochastic optimization; Chen et al. (2011) as well as Natarajan and Sim (2010) discussed the moment bounds in the context of robust portfolio selection; Wong and Zhang (2013) discussed the moment bounds in the context of nonlinear risk management; Lasserre, Preito-Rumeau, and Zervos (2006) discussed the pricing of a class of exotic options with moments and SDP relaxation. Regarding the theory underlying the computation of moment bounds, we refer the interested reader to Popescu (2005) and Lasserre (2008).

In particular, when we choose $\psi(x) = \mathbf{1}_{x\in\mathfrak{E}}$ for some event $\mathfrak{E}$ in the sample space $\Omega\subseteq\mathbf{R}^d$ of x, (GP) is the worst-case probability, which can be regarded as an implicit function of the moments, given the event $\mathfrak{E}$:

* Corresponding author.
E-mail addresses: edmhwong@gmail.com (M.H. Wong), zhangs@umn.edu (S. Zhang).
[1] Tel. 1 612 6248406

0377-2217/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2013.08.044

$$\mathbb{F}_{d,n}(\mathfrak{E}) := \sup_{(\mathbf{R}^d \ni) x \sim (\mathfrak{m}_0,\ldots,\mathfrak{m}_n)} \mathbb{P}[x \in \mathfrak{E}] \quad \left(= \sup_{x \sim (\mathfrak{m}_0,\ldots,\mathfrak{m}_n)} \mathbb{E}[\mathbf{1}_{x \in \mathfrak{E}}] \right). \tag{1}$$

As we shall see later, choosing $\mathfrak{E} = \{x \in \mathbf{R} : x \leqslant t\}$, $\mathbb{F}_{1,2}(\mathfrak{E})$ is in fact a probability distribution itself. In general, however, this is not the case; although we always have $0 \leqslant \mathbb{F}_{d,n} \leqslant 1$, it may not satisfy the additivity of joint countable union, namely, for any countable sequence of pairwise disjoint event $\mathfrak{E}_1, \mathfrak{E}_2, \ldots$, we only have

$$\mathbb{F}_{d,n}\left(\bigcup_j^{\infty} \mathfrak{E}_j\right) \leqslant \sum_{j=1}^{\infty} \mathbb{F}_{d,n}(\mathfrak{E}_j).$$

In other words, only subadditivity is guaranteed and the equality holds only when the right hand side is attained by the same extremal distribution of x for all $\mathfrak{E}_j$.

The possibility of deriving an analytical form or devising a simple computational procedure for $\mathbb{F}_{d,n}$ remains open for general n and d. Throughout this paper, we adhere our discussion to the case $n = 2$, unless specified otherwise. When $d = 1$ and $n = 2$, there are nice distributional robust functions in analytical form. In our subsequent discussion we will revisit them while applying the Value-at-Risk in the context of portfolio selection. The formulation is in line with El Ghaoui, Oks, and Oustry (2003), who discuss the worst-case Value-at-Risk knowing the first two moments. To the best of our knowledge, even when $d = 2$ there is no analytical form or method for exact computation. The closest approximation is due to Cox et al. (2008), who use sum-of-squares (SOS) polynomials to approximate $\mathbb{F}_{2,2}(\mathfrak{E})$ for nonnegative random variables, where $\mathfrak{E} = \{x \in \mathbf{R}^2 : x \leqslant t \text{ for some } t \in \mathbf{R}^2_+\}$. In this paper, we propose a computational method, in the realm of semidefinite programming (SDP), to exactly compute $\mathbb{F}_{2,2}(\cdot)$. The methodology is based on the characterization of copositive cones in $\mathbf{R}^{d+1}$, where $d \leqslant 3$, and some results in Luo, Sturm, and Zhang (2004), which state that, given either (i) $x^{(1)} \in [0,1]$ or (ii) $x^{(1)} \in \mathbf{R}_+$, and $x^{(2)} \in \mathbf{R}^m$, the nonnegativity of a bi-quadratic function reduces to LMIs.

Before proceeding, let us formally summarize and highlight the main contributions of this paper.

1. We introduce the moment bound of the probability as a function given their moments, and formally introduce this function as the *distributional robust probability function*.
2. In particular for the moment bound of two joint events, Cox et al. (2008) developed an approximation approach through sum-of-squares polynomials. In contrast, we provide an methodology of the exact moment bound in the form of a semidefinite program.
3. In association with risk management, we give three examples as applications of the distributional robust probability functions. Our approach is in particular useful and powerful when involving the bounded events, i.e. $(l_1 \leqslant x^{(1)} \leqslant u_1, l_2 \leqslant x^{(2)} \leqslant u_2)$.

The rest of this paper is organized as follows. In Section 2, we review $\mathbb{F}_{1,2}$ and its connection to Value-at-Risk in the context of portfolio selection. In Section 3, we derive the LMIs for computing $\mathbb{F}_{2,2}$, where three events are taken into account as our "base cases": $\mathfrak{E}_1 := \{x \in \mathbf{R}^2 : x^{(1)} \leqslant u^{(1)}, x^{(2)} \leqslant u^{(2)}\}$, $\mathfrak{E}_2 := \{x \in \mathbf{R}^2 : l^{(1)} \leqslant x \leqslant u^{(1)}, l^{(2)} \leqslant x \leqslant u^{(2)}\}$ and $\mathfrak{E}_3 := \{x \in \mathbf{R}^2 : x^{(1)} \leqslant u^{(1)}, 1_2 \leqslant x \leqslant u^{(2)}\}$. Model extensions are introduced in Section 4, followed by applications in Section 5, and finally we conclude the paper in Section 7.

2. Distributional robust function with a single random variable

Take $\mathfrak{E}_1 = \{x \in \mathbf{R} : x \leqslant t\}$ and let μ_1 and σ^2 be the mean and variance respectively. $\mathbb{F}_{1,2}(\mathfrak{E}_1)$ can be represented as a function of t (Cantelli, 1910; Chebyshev, 1874):

$$\mathbb{F}_{1,2}(\mathfrak{E}_1) := F(t) = \begin{cases} \frac{\sigma^2}{(\mu_1 - t)^2 + \sigma^2}, & t \leqslant \mu_1; \\ 1, & t > \mu_1. \end{cases} \tag{2}$$

The above follows essentially from the Chebyshev–Cantelli inequality. An alternative proof can be found in Chen et al. (2011). It is also well-known that this worse-case probability is achieved by a two-point distribution of x. However, the story is completely different when $F(t)$ is regarded as a distribution function of some random variable ζ, since it now has a smooth and continuous distribution (2), which allows us to compute its moments analytically. In the next paragraph, we review the portfolio selection based on worst-case Value-at-Risk. We will use this $F(t)$ in the proof of Lemma 1 to enable a second-order cone programming formulation of the portfolio selection problem. Meanwhile, it is interesting to note that the first two moments of ζ and x are no longer the same: $\mathbb{E}(\zeta) = \mu_1 - \frac{\pi}{2}\sigma$ versus $\mathbb{E}(x) = \mu_1$; and $\mathbb{E}(\zeta^2) = \infty$ versus $\mathbb{E}(x^2) = \mu_1^2 + \sigma^2$ (see Appendix A). Recall that ζ follows a worst-case distribution and represents an extreme event. This infinite variance of ζ actually meets our intuition that an extreme event has a high "fluctuation".

In risk management, as extreme events are often associated with the Value-at-Risk (VaR), let us apply ζ with this risk measure and consider a portfolio selection problem (Castellacci & Siclari, 2003; Goh, Lim, Sim, & Zhang, 2012; Gotoh & Takeda, 2012; Rossello, 2008). Suppose that $\theta \in \mathbf{R}^p$ is the vector of investment return from p assets with a mean $m \in \mathbf{R}^p$ and second moment matrix $M \in \mathcal{S}^p_+$. Let $w \in \mathbf{R}^p$ be the portfolio weights and $x = w^T\theta$ the portfolio return. Then $\mathbb{E}(x) = w^Tm$ and $\mathbb{E}(x^2) = w^TMw$. Applying $\mathbb{F}_{1,2}(\mathfrak{E}_1)$ in the definition VaR, where we regard $-w^T\theta$ as the loss and choose $t = -\alpha$ in $\mathfrak{E}_1$, we have

$$\mathrm{VaR}_\epsilon(w^T\theta) := \arg\min_\alpha \{\mathbb{F}_{1,2}(-w^T\theta \geqslant \alpha) \leqslant \epsilon\},$$

where $\epsilon \in (0,1)$ is the level of confidence. The higher the α, the higher the risk. Therefore we would like to minimize the risk over the set of admissible portfolio $\mathcal{W}$ (which typically incorporates the target of return, budget constraint and sometimes no short selling constraints) as follows:

$$\begin{aligned} \min \quad & \alpha \\ \text{s.t.} \quad & \mathbb{F}_{1,2}(-w^T\theta \geqslant \alpha) = F(-\alpha) \leqslant \epsilon \\ & w \in \mathcal{W}, \end{aligned} \tag{3}$$

where ϵ is given. Recall that $F(-\alpha) = \sup_{x \sim (w^Tm, w^TMw)} \mathbb{P}(x \geqslant -\alpha)$. Below we show that (3) is a convex optimization model and can be solved efficiently.

Lemma 1 (*See also Theorem 1 of (El Ghaoui et al., 2003)*). *Problem (3) can be reformulated by second-order cone programming (SOCP).*

Proof. The assertion follows from the observation that

$$F(-\alpha) \leqslant \epsilon \Longleftrightarrow \frac{w^TMw}{(w^Tm + \alpha)^2 + w^TMw} \leqslant \epsilon \Longleftrightarrow (1-\epsilon)w^TMw \leqslant \epsilon(w^Tm + \alpha)^2 \Longleftrightarrow \begin{pmatrix} w^Tm + \alpha \\ \frac{1-\epsilon}{\epsilon} M^{\frac{1}{2}} w \end{pmatrix} \in SOC(d+1).$$

Note that we have implicitly assumed $-w^Tm \leqslant \alpha$. Otherwise $\mathbb{F}_{1,2}(-w^T\theta \geqslant \alpha) = 1 > \epsilon$, which contradicts the definition of VaR. The proof in Lemma 1 can be regarded as an alternative approach compared to that of El Ghaoui et al. (2003). The key difference is that we work with the extremal distribution $\sup_{x \sim (\mu_1, \Gamma)} \mathbb{P}(\mathfrak{E}_1) = F(t)$, which makes the proof explicit. (Recall that in El Ghaoui et al. (2003) the strong duality is discussed for $\sup_{x \sim (\mu_1, \Gamma)} \mathbb{P}(\mathfrak{E}_1)$.)

Let us introduce the worst-case probability of the event $\mathfrak{E}_2 := \{x \in \mathbf{R} : l \leqslant x \leqslant u\}$ as: