European Journal of Operational Research 249 (2016) 188–199

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective☆

Tito Homem-de-Mello*, Bernardo K. Pagnoncelli

School of Business, Universidad Adolfo Ibáñez, Diagonal las Torres 2640, Peñalolén, Santiago, Chile

ARTICLE INFO

Article history:
Received 17 April 2014
Accepted 18 May 2015
Available online 1 June 2015

Keywords:
Stochastic programming
Risk aversion
Multistage
Consistency
Pension funds

ABSTRACT

We discuss the incorporation of risk measures into multistage stochastic programs. While much attention has been recently devoted in the literature to this type of model, it appears that there is no consensus on the best way to accomplish that goal. In this paper, we discuss pros and cons of some of the existing approaches. A key notion that must be considered in the analysis is that of consistency, which roughly speaking means that decisions made today should agree with the planning made yesterday for the scenario that actually occurred. Several definitions of consistency have been proposed in the literature, with various levels of rigor; we provide our own definition and give conditions for a multi-period risk measure to be consistent. A popular way to ensure consistency is to nest the one-step risk measures calculated in each stage, but such an approach has drawbacks from the algorithmic viewpoint. We discuss a class of risk measures—which we call expected conditional risk measures—that address those shortcomings. We illustrate the ideas set forth in the paper with numerical results for a pension fund problem in which a company acts as the sponsor of the fund and the participants' plan is defined-benefit.

1. Introduction

The evolution and widespread use of stochastic programming is closely related to the increasing computing power made available since the foundation of the field. The important class of two-stage stochastic programs found immediate use in applications since its general framework of first- and second-stage decisions is suitable for a number of real-world problems, see for instance Wallace and Ziemba (2005). Later, the attention turned to multistage stochastic programs (MSSPs), which are a natural extension of two-stage models. In those problems the sequence of events starts with a decision, followed by a realization of a random vector, and then a decision is made knowing the outcome of the random vector, a new realization occurs, and so on. Randomness is often described by a continuous stochastic process or by a discrete process with a very high number of possible outcomes. A common approach is to build a *scenario tree*, which is a discrete representation that in some sense is close to the original process according to some distance. The generation of scenario trees has received a great deal of attention in the literature: see, for instance, Pflug (2001), Høyland and Wallace (2001), Dupačová, Gröwe-Kuska, and Römisch (2003), Heitsch and Römisch (2009), Pflug and Pichler (2011), Pflug and Pichler (2012), Mehrotra and Papp (2013). Scenario trees are crucial for numerical solution of the problem, as algorithms used in practice to solve the problem are typically rooted in some decomposition principles such as the Nested Decomposition (Donohue & Birge, 2006) or the Stochastic Dual Dynamic Programming scheme (Pereira & Pinto, 1991). MSSPs have been used in a number of areas, including finance, revenue management, energy planning, and natural resources management, among others.

The classical formulation of stochastic programs (in two or more stages) optimizes the expected value of an objective function that depends on the decision variables as well as on the random variables that represent the uncertainty in the problem. Such a formulation assumes that the decision maker is risk-neutral, i.e., he or she will not mind large losses in some scenarios as long as those are offset by large gains in other scenarios. While such an approach is useful in a number of applications, it does not reflect the situation where the decision is very concerned about large losses—in other words, such a decision maker is *risk-averse*. It is natural then to consider risk-averse formulation of stochastic programs.

In the case of two-stage models, the structure of a first-stage deterministic cost plus a random recourse cost in the second stage makes the extension to the risk-averse case immediate from a modeling perspective, in the sense that the natural choice is to

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

* Corresponding author. Tel.: +56 2 2331 1208.
E-mail addresses: tito.hmello@uai.cl (T. Homem-de-Mello), bernardo.pagnoncelli@uai.cl (B.K. Pagnoncelli).

http://dx.doi.org/10.1016/j.ejor.2015.05.048
0377-2217/

replace the expectation of the second stage cost with some other risk measure; see, for instance, Schultz and Tiedemann (2006), Fábián (2008), Shapiro, Dentcheva, and Ruszczyński (2009) and Miller and Ruszczyński (2011). The difficulty associated with the risk-averse model depends on the choice of the risk measure. Ahmed (2006) shows that if the risk measure is the variance, then the resulting problem is NP-hard. Furthermore, monotonicity in the second stage would be lost, and the cost units of first and second stage would be different unless the standard deviation was used, but the problem would likely become intractable in this case. Rockafellar and Uryasev (2000) show that if the Conditional Value-at-Risk is chosen, the sampled version of the continuous problem can be approximated by a linear programming problem. Noyan (2012) proposes two decomposition algorithms to efficiently solve a disaster management problem with the Conditional Value-at-Risk as the risk measure.

For multistage stochastic programming the picture is quite different and several questions arise. When sequential decision is involved, there is no natural or obvious way of measuring risk. Should risk be measured at every stage separately? Should it be applied to the several scenario paths in the tree? Or should risk be measured in a nested way, in the spirit of dynamic programming? What if only the risk at the end of the time horizon is relevant and we do not want to measure risk at the other stages? The difficulty in extending risk measures to the multistage setting has been discussed in several papers and it can be argued that the differences among the approaches are far more significant than the two-stage case.

A number of recent papers have considered the importance of measuring risk in MSSPs (see, for instance, Collado, Papp, & Ruszczyński, 2012; Eichhorn & Römisch, 2005; Guigues & Sagastizábal, 2013; Kozmík & Morton, 2015; Pagnoncelli & Piazza, 2012; Pflug & Römisch, 2007; Pflug & Pichler, 2014b; Philpott & de Matos, 2012; Philpott, de Matos, & Finardi, 2013; Shapiro, 2012a; Shapiro, Tekaya, Soares, & da Costa, 2013). Several of these papers focus on how to adapt existing algorithms from the risk-neutral case to the risk-averse case, often with the Conditional Value-at-Risk as the risk measure. One of the goals of our paper is to address some of the popular ways to measure risk and discuss their advantages and drawbacks. In addition, we revisit and extend a class of multi-period risk measures proposed by Pflug and Ruszczyński (2005) (see also Pflug, 2006 for a more extensive discussion), which we call *expected conditional risk measures* (*ECRMs*), and discuss how the resulting problem can be efficiently solved. ECRMs combine two attractive features: on the one hand, ECRMs can be represented in a nested form, a feature that is desirable and the focus of much of the recent literature, as we shall see later; on the other hand, we show that when ECRMs are applied with the Conditional Value-at-Risk (CVaR) as the underlying risk measure, the resulting MSSP can be represented by a simpler risk-neutral MSSP with additional variables, much in the spirit of the polyhedral risk measures introduced by Eichhorn and Römisch (2005).

As it has been observed in the literature, one very important issue that arises when modeling risk-averse MSSPs is that of *time consistency*. Time consistency in MSSPs has been highlighted by several authors in recent years as a desirable property a problem should have. Informally, time consistency means that if you solve an MSSP today and find solutions for each node of a tree, you should find the same solutions if you re-solve the problem tomorrow given what was observed and decided today. The definitions in the literature differ mainly by their focus: the works of Ruszczyński (2010) and Kovacevic and Pflug (2014) deal with sequences of random variables, while Detlefsen and Scandolo (2005), Cheridito, Delbaen, and Kupper (2006), and Bion-Nadal (2008), define time consistency for continuous-time dynamic models. The definitions in Shapiro (2009), Carpentier, Chancelier, Cohen, De Lara, and Girardeau (2012), Rudloff, Street, and Valladão (2014) and De Lara and Leclère (2014) are centered on optimization and on the stability of decision variables at every stage. Xin, Goldberg, and Shapiro (2013) propose definitions of time-consistency of policies in the context of distributionally robust MSSPs, whereas Pflug and Pichler (2014a) propose a related notion of time-consistent decisions. We propose a new definition of consistency, closer to the optimization-oriented papers. Our definition is suitable for MSSPs that can be represented via scenario trees. Using a simple three-stage inventory problem we show that several natural ways of measuring risk lead to inconsistent formulations, according to our definition. We also show the class of ECRMs we study in this paper is time-consistent.

We illustrate the applicability of ECRMs by using it in a pension fund problem proposed by Haneveld, Streutker, and Van Der Vlerk (2010). This numerical example illustrates two important aspects of ECRMs: first, the simplicity of implementation when the CVaR is used as an ingredient for the ECRM—indeed, we use standard software for risk-neutral multistage programs available in the literature to solve the corresponding risk-averse problem. The second important aspect is the flexibility allowed by the model to represent the change in the degree of risk aversion over time; for example, the decision maker may be more risk-averse about the earlier stages and less risk-averse about the stages farther in the future. We also use the numerical example to propose a (to the best of our knowledge) novel way to compare optimal solutions of MSSPs. The majority of applications only analyzes the first-stage solution since in most cases a rolling-horizon procedure will be implemented in practice and the solutions of other stages will not be implemented. We show in our pension fund example that the solutions of subsequent stages carry important information concerning the quality and robustness of the first-stage solution. By using first- and second-order dominance we show that despite having an attractive first-stage allocation, some solutions exhibit a poor behavior in subsequent stages, such as having a very high probability of needing extra money injection in the fund.

The rest of the paper is organized as follows. Section 2 defines our notion of consistency. In Section 3 we present an inventory problem that illustrates our notion of consistency and discuss several modeling paradigms for risk-averse MSSPs. We prove some results that characterize consistency according to our definition in Section 4. In Section 5 we introduce the notion of ECRMs and study in detail their properties, including consistency and the equivalent risk-neutral formulation of the case with CVaR. The pension fund example that illustrates our approach is presented in Section 6, while Section 7 presents some concluding remarks.

2. Consistency

We start by defining precisely the notation and the class of problems we want to study. Consider a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \mathcal{F}_T$ be sub sigma-algebras of $\mathcal{F}$ such that each $\mathcal{F}_t$ corresponds to the information available up to (and including) stage t, with $\mathcal{F}_1 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. Let $\mathcal{Z}_t$ denote a space of $\mathcal{F}_t$-measurable functions from Ω to $\mathbb{R}$, and let $\mathcal{Z} := \mathcal{Z}_1 \times \cdots \times \mathcal{Z}_T$. We define a *multi-period risk function* $\mathbb{F}$ as a mapping from $\mathcal{Z}$ to $\mathbb{R}$. For example, we may have, for $Z \in \mathcal{Z}$,

$$\mathbb{F}(Z) = \mathbb{F}_1(Z_1) + \cdots + \mathbb{F}_T(Z_T), \tag{2.1}$$

where each $\mathbb{F}_t$ is a *one-period risk function*, i.e., a mapping from $\mathcal{Z}_t$ to $\mathbb{R}$. Note that some assumptions may be required for the existence of the mapping $\mathbb{F}$. For example, consider the additive case (2.1) with each $\mathbb{F}_i$ being the expectation operator; then, each expectation must exist and be finite, which can be ensured for example when Ω is finite. Throughout this paper we assume that $\mathbb{F}$ is well-defined and finite.

Consider now the space $\mathcal{D}_T$ of distributions of T-dimensional random vectors in $\mathcal{Z}$. That is, each element $G \in \mathcal{D}_T$—which is a mapping from $\mathcal{B}_T$ to [0, 1], where $\mathcal{B}_T$ is the Borel sigma-algebra in $\mathbb{R}^T$—can be written as the distribution function G_Z of some $Z = (Z_1, \ldots, Z_T) \in \mathcal{Z}$,