Accepted Manuscript

Data-driven risk-averse stochastic optimization with Wasserstein metric

Chaoyue Zhao, Yongpei Guan

PII: S0167-6377(18)30050-6
DOI: https://doi.org/10.1016/j.orl.2018.01.011
Reference: OPERES 6334

To appear in: *Operations Research Letters*

Received date : 25 June 2016
Revised date : 29 January 2018
Accepted date : 29 January 2018

Please cite this article as: C. Zhao, Y. Guan, Data-driven risk-averse stochastic optimization with Wasserstein metric, *Operations Research Letters* (2018), https://doi.org/10.1016/j.orl.2018.01.011

Data-Driven Risk-Averse Stochastic Optimization with Wasserstein Metric

Chaoyue Zhao[†] **and Yongpei Guan**[‡]
[†]School of Industrial Engineering and Management
Oklahoma State University, Stillwater, OK 74074
[‡]Department of Industrial and Systems Engineering
University of Florida, Gainesville, FL 32611
Emails: chaoyue.zhao@okstate.edu; guan@ise.ufl.edu

Abstract

In this paper, we study a data-driven risk-averse stochastic optimization approach with Wasserstein Metric for the general distribution case. By using the Wasserstein Metric, we can successfully reformulate the risk-averse two-stage stochastic optimization problem with distributional ambiguity to a traditional two-stage robust optimization problem. In addition, we derive the worst-case distribution and perform convergence analysis to show that the risk aversion of the proposed formulation vanishes as the size of historical data grows to infinity.

Keywords: stochastic optimization; data-driven decision making; Wasserstein metric

1 Introduction

The traditional two-stage stochastic optimization problem can be described as follows (cf. [3] and [15]):

$$\text{(SP)} \quad \min_{x\in X} \quad c^\top x + \mathbb{E}_{\mathbb{P}}[\mathcal{Q}(x,\xi)],$$

where the first-stage decision variable x is in a compact set X and the second-stage problem is

$$\mathcal{Q}(x,\xi) = \min_{y\in Y}\{d(\xi)^\top y : A(\xi)x + By \geq b(\xi)\}, \tag{1}$$

where $\mathcal{Q}(x,\xi)$ is assumed continuous on ξ (e.g., when $A(\xi)$ and $b(\xi)$ are continuous on ξ) and the uncertain random variable ξ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, in which Ω is a compact convex sample space for ξ, $\mathcal{F}$ is a σ-algebra of Ω, and $\mathbb{P}$ is the associated given true probability distribution. SP has broad applications. However, there are challenges. For instance, in practice, the true distribution of the random parameters is usually unknown and hard to predict accurately and the inaccurate estimation of the true distribution may lead to biased solutions and make the solutions sub-optimal. Meanwhile, there are usually a series of historical data available for the unknown true distribution. To incorporate distribution ambiguity and utilize the historical data, we consider a data-driven risk-averse stochastic optimization formulation:

$$\text{(DD-SP)} \quad \min_{x\in X} \quad c^\top x + \max_{\hat{\mathbb{P}}\in\mathcal{D}} \mathbb{E}_{\hat{\mathbb{P}}}[\mathcal{Q}(x,\xi)],$$

where $\hat{\mathbb{P}}$ represents an unknown distribution in the given confidence set $\mathcal{D}$. This formulation allows distribution ambiguity and introduces a confidence set $\mathcal{D}$ to ensure that the true distribution $\mathbb{P}$ is within this set with a certain confidence level based on nonparametric statistics. In general, if we let $d_{\text{M}}(\mathbb{P}_0, \hat{\mathbb{P}})$ be any distance measure between a reference distribution $\mathbb{P}_0$ and an unknown distribution $\hat{\mathbb{P}}$ and use θ, as a function of the size of historical data, to represent the corresponding distance, then the confidence set $\mathcal{D}$ can be represented as follows:

$$\mathcal{D} = \{\hat{\mathbb{P}} : d_{\text{M}}(\mathbb{P}_0, \hat{\mathbb{P}}) \leq \theta\}.$$