Reformulations for data-driven stochastic optimization problems with structured ambiguity sets

Lotfi M. Chaouach[1], Dimitris Boskos[1], and Tom Oomen[1,2]

[1]Delft Center for Systems and Control, 2628CD, Delft, the Netherlands (e-mail: l.chaouach@tudelft.nl)
[2]Eindhoven University of Technology, the Netherlands

1 Introduction

Hedging against uncertainty in the choice of probability distributions is essential in data-driven stochastic optimization. Such choices need to account for limitations like the amount of available data, the dimensionality of the uncertainty, or the fact that the data might be corrupted, as these factors influence the accuracy of the inferred model and the reliability of the decision. Thus, it is required to derive trustworthy ambiguity sets from the available data, which contain the data-generating distribution with high probability. It is desirable to achieve this under the least possible amount of conservativeness and with tractability guarantees for the optimization problem.

2 Problem formulation

A typical data-driven distributionally robust optimization problem has the form

$$\min_{x\in\mathcal{X}} \max_{P_\xi\in\widehat{\mathcal{P}}^N} \mathbb{E}_{P_\xi}\big[f(x,\xi)\big], \tag{1}$$

where f is the cost function, x is the decision variable, $\xi\in\mathbb{R}^d$ is a random vector with distribution P_ξ, and $\widehat{\mathcal{P}}^N$ is an ambiguity set of distributions that is informed by N samples of ξ. A convenient choice to build $\widehat{\mathcal{P}}^N$ is by grouping the distributions around the empirical distribution P_ξ^N of the samples up to a given distance ε in the Wasserstein metric [2], which yields a ball $\mathcal{B}(P_\xi^N,\varepsilon)$. It is possible to tune the radius of this ball to guarantee that it contains the true distribution with prescribed confidence. However, this becomes very conservative when the dimension d of the uncertainty is high [1]. This curse of dimensionality can be overcome by exploiting independence between the lower-dimensional components of $\xi=(\xi_1,\ldots,\xi_n)$ to build an ambiguity ball for each component and construct a product measure hyperrectangle

$$\mathcal{H}(\boldsymbol{P}_\xi^N,\boldsymbol{\varepsilon}) := \mathcal{B}(\hat{P}_{\xi_1}^N,\varepsilon_1)\otimes\ldots\otimes\mathcal{B}(\hat{P}_{\xi_n}^N,\varepsilon_n), \tag{2}$$

around the product of the empirical distributions of the components $\boldsymbol{P}_\xi^N$. Nevertheless, $\mathcal{H}$ is a non-convex set and tractable reformulations of (1) with $\widehat{\mathcal{P}}^N=\mathcal{H}(\boldsymbol{P}_\xi^N,\boldsymbol{\varepsilon})$ are only available when the cost is the sum or the product of functions that depend on the individual components $\xi_1,\ldots,\xi_n$ [1].

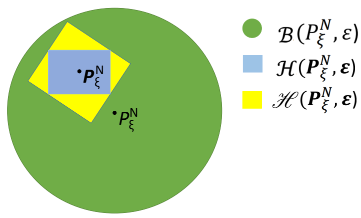

Figure 1: The multi-transport hyperrectangle contains the product measure hyperrectangle while enjoying the same statistical guarantees and similar size-reduction properties compared to the monolithic ambiguity ball.

3 Multi-transport ambiguity hyperrectangles

To obtain tractable reformulations for a broader class of functions, we propose another ambiguity set, which contains $\mathcal{H}$ without being much larger. Given a baseline measure μ and a vector of transport budgets $\boldsymbol{\delta}=(\delta_1,\ldots,\delta_n)\in\mathbb{R}_+^n$, consider the set of transport plans

$$\Pi(\mu,\boldsymbol{\delta}) := \Big\{\pi\in\mathcal{P}_p(\mathbb{R}^d\times\mathbb{R}^d) : \mathrm{pr}_1(\pi)=\mu \\ \int\|\xi_k-\hat{\xi}_k\|^p d\pi(\hat{\xi},\xi)\le\delta_k,\ k=1,\ldots,n\Big\}.$$

where $\mathcal{P}_p(\mathbb{R}^d\times\mathbb{R}^d)$ denotes the class of distributions on $\mathbb{R}^d$ with finite pth momen and $\mathrm{pr}_1(\pi)$ denotes the first marginal of π. We next introduce the *multi-transport* hyperrectangle

$$\mathscr{H}(\boldsymbol{P}_\xi^N,\boldsymbol{\varepsilon}) := \Big\{P_\xi\in\mathcal{P}_p(\mathbb{R}^d)\text{ s.t }\exists\pi\in\Pi(\boldsymbol{P}_\xi^N,\boldsymbol{\varepsilon}) \\ \text{with marginal } P_\xi\Big\}. \tag{3}$$

By exploiting the analysis in [2], we formulate a tractable dual problem for (1) with $\widehat{\mathcal{P}}^N=\mathscr{H}(\boldsymbol{P}_\xi^N,\boldsymbol{\varepsilon})$ and establish strong duality for any upper semi-continuous cost function f. In addition, $\mathscr{H}$ enjoys similar statistical properties to $\mathcal{H}$ with favorable size reduction rates (see Figure 1).

References

[1] Chaouach, L. M., Boskos, D., & Oomen, T., Uncertain uncertainty in data-driven stochastic optimization: Towards structured ambiguity sets. In 61st IEEE Conference on Decision and Control, Cancun, Mexico, 2022.

[2] J. Blanchet and K. Murthy, "Quantifying distributional model risk via optimal transport," Mathematics of Operations Research, vol. 44, no. 2, pp. 565–600, 2019.