D

Distributionally Robust Two-Stage Linear Programs with Wasserstein Distance: Tractable Formulations

Nan Jiang and Weijun Xie
Georgia Institute of Technology, Atlanta, GA, USA

Article Outline

Keywords

Distributionally robust · Two-stage · Stochastic program · Wasserstein · Tractable

MSC Codes

90C11, 90C15, 62J07

Introduction

Consider a distributionally robust two-stage linear programs (DRTSLP) of the form

$$v^* = \min_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} + \mathcal{Z}(\boldsymbol{x}) := \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[Z(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \right] \right\}, \tag{1}$$

where set $\mathcal{X} \in \mathbb{R}^{n_1}$ represents the deterministic feasible set of the here-and-now decisions $\boldsymbol{x}$, $\boldsymbol{c}$ denotes the cost coefficients of the here-and-now objective, and $\mathcal{Z}(\boldsymbol{x})$ is the worst-case expected wait-and-see cost. Function $Z(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$ is known as the recourse function with uncertain parameters $\tilde{\boldsymbol{\xi}} \in \Xi$ and the probability distribution $\mathbb{P}$, which comes from a family of distributions represented by the ambiguity set $\mathcal{P}$.

Following the notation in [1, 3, 8, 9, 11], the recourse function $Z(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})$ in (1) with a realization $\boldsymbol{\xi}$ of $\tilde{\boldsymbol{\xi}}$ comprises the following optimization problem:

$$Z(\boldsymbol{x}, \boldsymbol{\xi}) = \min_{\boldsymbol{y}\in\mathbb{R}^{n_2}} \left\{ (\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{y} \colon \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y} \geq \boldsymbol{h}(\boldsymbol{x}) \right\}, \tag{2}$$

where $\boldsymbol{y}$ denotes the wait-and-see decisions in the second-stage problem; $\boldsymbol{T} : \mathbb{R}^{n_1} \to \mathbb{R}^{\ell\times m_2}$ and $\boldsymbol{h} : \mathbb{R}^{n_1} \to \mathbb{R}^{\ell}$ represent the technology affine mapping and the right-hand mapping, separately; and $\boldsymbol{\xi} = (\boldsymbol{\xi}_q, \boldsymbol{\xi}_T) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, $\boldsymbol{Q} \in \mathbb{R}^{n_2\times m_1}$, $\boldsymbol{q} \in \mathbb{R}^{n_2}$. Following the convention, DRTSLP (1) is termed as

– DRTSLP with *objective uncertainty* when $\Xi = \mathbb{R}^{m_1} \times \{\boldsymbol{\xi}_T\}$
– DRTSLP with *left-hand side uncertainty* when $\Xi = \{\boldsymbol{\xi}_q\} \times \mathbb{R}^{m_2}$

P. M. Pardalos, O. A. Prokopyev (eds.), *Encyclopedia of Optimization*,
https://doi.org/10.1007/978-3-030-54621-2_778-1

- DRTSLP with *right-hand side uncertainty* when $\Xi = \{\boldsymbol{\xi}_q\} \times \mathbb{R}^{m_2}$ and $\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{T}$ (i.e., the constant terms $\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T$ can be moved to right-hand sides of the constraints)

Throughout this paper, the following assumptions are imposed, which are quite standard in the two-stage stochastic program literature [3, 8, 9, 11]:

- (Fixed Recourse) The recourse matrix $\boldsymbol{W} \in \mathbb{R}^{\ell \times n_2}$ is fixed.
- (Sufficiently Expensive Recourse) DRTSLP (1) has sufficiently expensive recourse if for any $\boldsymbol{x} \in \mathcal{X}$, the dual of the second-stage problem (2) is feasible for all $\boldsymbol{\xi} \in \Xi$.
- (Separable Uncertainty) The support $\Xi = \Xi_q \times \Xi_T$, where $\Xi_q \subseteq \mathbb{R}^{\ell}$, $\Xi_T \subseteq \mathbb{R}^{n_2}$.

These assumptions are important to derive the deterministic counterparts of DRTSLP (1).

Wasserstein Ambiguity Set

Following [6, 8, 11, 12], the nominal distribution and the true distribution are assumed to be bounded with respect to the Wasserstein metric, i.e., under the Wasserstein ambiguity set, which is defined as

$$\mathcal{P}_r^W = \left\{\mathbb{P} \colon \mathbb{P}\left\{\tilde{\boldsymbol{\xi}} \in \Xi\right\} = 1, W_r(\mathbb{P}, \mathbb{P}_{\tilde{\zeta}}) \le \theta\right\},$$

where for any $r \in [1, \infty]$, the $r-$Wasserstein distance is defined as

$$W_r(\mathbb{P}_1, \mathbb{P}_2) = \inf\left\{\left[\int_{\Xi\times\Xi} \left\|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\right\|_p^r \mathbb{Q}(d\boldsymbol{\xi}_1, d\boldsymbol{\xi}_2)\right]^{\frac{1}{r}} : \begin{array}{l}\mathbb{Q} \text{ is a joint distribution of } \tilde{\boldsymbol{\xi}}_1 \text{ and } \tilde{\boldsymbol{\xi}}_2 \\ \text{with marginals } \mathbb{P}_1 \text{ and } \mathbb{P}_2, \text{ respectively}\end{array}\right\},$$

and $\mathbb{P}_{\tilde{\zeta}}$ is a discrete empirical reference distribution of random parameters $\tilde{\zeta}$ generated by N i.i.d. samples such that $\mathbb{P}_{\tilde{\zeta}}\{\tilde{\zeta} = \zeta^i\} = 1/N$, i.e., $\mathbb{P}_{\tilde{\zeta}} = 1/N \sum_{i\in[N]} \delta_{\zeta^i}$ and δ_{ζ^i} is the Dirac function that places unit mass on the realization $\tilde{\zeta} = \zeta^i$ for each $i \in [N]$, and $\theta \ge 0$ is the Wasserstein radius. Notice that if $r = \infty$, then the $\infty-$Wasserstein distance reduces to

$$W_\infty(\mathbb{P}_1, \mathbb{P}_2) = \inf\left\{\text{ess.sup}_{\mathbb{Q}} \left\|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\right\|_p : \begin{array}{l}\mathbb{Q} \text{ is a joint distribution of } \tilde{\boldsymbol{\xi}}_1 \text{ and } \tilde{\boldsymbol{\xi}}_2 \\ \text{with marginals } \mathbb{P}_1 \text{ and } \mathbb{P}_2, \text{ respectively}\end{array}\right\}.$$

Tractable Formulations

In general, it has been shown in [8, 11] that solving a DRTSLP with $r-$Wasserstein ambiguity set can be NP-hard. This section reviews the tractable reformulations of DRTSLP (1) with $r-$Wasserstein ambiguity set with $r = 1, r = \infty$ and derives new results for DRTSLP (1) with $r \in (1, \infty)$. Formally, tractability of a convex program is formally defined below.

Definition 1 (Tractability, [2]) Given a compact set $\widehat{\mathcal{X}} \in \mathbb{R}^{n_1}$ with nonempty interior, suppose it is contained in a Euclidean ball with radius R and is containing a Euclidean ball with radius r. Then there exists an efficient algorithm to solve DRTSLP (1) to $\widehat{\varepsilon} > 0$ accuracy, whose running time is polynomial in $n_1, n_2, m_1, m_2, \ell, N, \ln(R/r), \ln(1/\widehat{\varepsilon})$, and the encoding length of DRTSLP (1).

Type 1−Wasserstein Ambiguity Set

There are two known special cases, under which DRTSLP (1) with 1−Wasserstein ambiguity set can be tractable. The first case is DRTSLP with objective uncertainty, which has been studied by [6].

Proposition 1 (DRTSLP with Objective Uncertainty, [6]) *Suppose $\Xi = \mathbb{R}^{m_1} \times \{\boldsymbol{\xi}_T\}$ and for any given $\boldsymbol{x} \in \mathcal{X}$, the feasible region of second-stage problem in DRTSLP (1) (i.e., set $\{\boldsymbol{y} : \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y} \geq \boldsymbol{h}(\boldsymbol{x})\}$) is nonempty and compact. Then DRTSLP (1) can be tractable for any $p \in [1, \infty]$ and admits the following equivalent formulation:*

$$
\begin{aligned}
v^* = \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y},\lambda\geq 0} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \lambda\theta + \frac{1}{N}\sum_{j\in[N]} (\boldsymbol{Q}\boldsymbol{\zeta}_q^j + \boldsymbol{q})^\top \boldsymbol{y}^j, \\
\text{s.t.} \quad & \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y}^j \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N], \\
& \lambda \geq \|\boldsymbol{Q}^\top \boldsymbol{y}^j\|_{p*}, \forall j \in [N], \\
& \boldsymbol{y}^j \in \mathbb{R}^{n_2}, \forall j \in [N],
\end{aligned}
$$

where $p* = \frac{p}{p-1}$.

If there is only constraint uncertainty involved, then DRTSLP (1) can have a tractable representation with the reference distance $\|\cdot\|_1$.

Proposition 2 (DRTSLP with Left-Hand Side Uncertainty, [8]) *Suppose $\Xi = \{\boldsymbol{\xi}_q\} \times \mathbb{R}^{m_2}$ and $p = 1$. Then DRTSLP (1) can be tractable and admits the following equivalent formulation:*

$$
\begin{aligned}
v^* = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y},\\ \boldsymbol{\varphi},\boldsymbol{\psi}}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \lambda\theta + \frac{1}{N}\sum_{j\in[N]} (\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{y}^j, \\
\text{s.t.} \quad & \boldsymbol{\varphi}_i, \boldsymbol{\psi}_i \in \mathbb{R}^{n_2}, \forall i \in [m_2], \\
& \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\zeta}_T^j + \boldsymbol{W}\boldsymbol{y}^j \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N], \\
& (\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{\varphi}_i \leq \lambda, (\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{\psi}_i \leq \lambda, \forall i \in [m_2], \\
& \boldsymbol{T}(\boldsymbol{x})\boldsymbol{e}_i \leq \boldsymbol{W}\boldsymbol{\varphi}_i, -\boldsymbol{T}(\boldsymbol{x})\boldsymbol{e}_i \leq \boldsymbol{W}\boldsymbol{\psi}_i, \forall i \in [m_2], \\
& \lambda \geq 0, \boldsymbol{y}^j \in \mathbb{R}^{n_2}, \forall j \in [N].
\end{aligned}
$$

Note that the result in Proposition 2 also holds for DRTSLP with right-hand side uncertainty, which is summarized below:

Corollary 1 (DRTSLP with Right-Hand Side Uncertainty) *Suppose $\Xi = \{\boldsymbol{\xi}_q\} \times \mathbb{R}^{m_2}$, $p = 1$, and $\boldsymbol{T}(\boldsymbol{x}) := \boldsymbol{T}$. Then DRTSLP (1) can be tractable and admits the following equivalent formulation:*

$$
\begin{aligned}
v^* = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y},\\ \boldsymbol{\varphi},\boldsymbol{\psi}}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \lambda\theta + \frac{1}{N}\sum_{j\in[N]} (\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{y}^j, \\
\text{s.t.} \quad & \boldsymbol{\varphi}_i, \boldsymbol{\psi}_i \in \mathbb{R}^{n_2}, \forall i \in [m_2], \\
& \boldsymbol{T}\boldsymbol{\zeta}_T^j + \boldsymbol{W}\boldsymbol{y}^j \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N], \\
& (\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{\varphi}_i \leq \lambda, (\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{\psi}_i \leq \lambda, \forall i \in [m_2], \\
& |\boldsymbol{T}\boldsymbol{e}_i| \leq \boldsymbol{W}\boldsymbol{\varphi}_i, \forall i \in [m_2], \\
& \lambda \geq 0, \boldsymbol{y}^j \in \mathbb{R}^{n_2}, \forall j \in [N].
\end{aligned}
$$

Type ∞—Wasserstein Ambiguity Set

In [11], the author showed that by exploring the neat representation of the worst-case expected wait-and-seet cost, DRTSLP (1) enjoys more tractable results.

Proposition 3 (DRTSLP (1) with Both Objective and Constraint Uncertainty, [11]) *Suppose $p = \infty$ and either $\boldsymbol{T}(\boldsymbol{x}) \in \mathbb{R}_+^{\ell\times m_2}$ for all $\boldsymbol{x} \in \mathcal{X}$ or $\boldsymbol{T}(\boldsymbol{x}) \in \mathbb{R}_-^{\ell\times m_2}$ for all $\boldsymbol{x} \in \mathcal{X}$. Then DRTSLP (1) can be tractable and admits the following equivalent formulation:*

$$
\begin{aligned}
v^* = \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \frac{1}{N}\sum_{j\in[N]} \left[\left(\boldsymbol{Q}\boldsymbol{\zeta}_q^j + \boldsymbol{q}\right)^\top \boldsymbol{y}^j + \theta \|\boldsymbol{Q}^\top \boldsymbol{y}^j\|_1 \right], \\
\text{s.t.} \quad & \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\zeta}_T^j + \boldsymbol{W}\boldsymbol{y}^j - \theta|\boldsymbol{T}(\boldsymbol{x})|\boldsymbol{e} \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N], \\
& \boldsymbol{y}^j \in \mathbb{R}^{n_2}, \forall j \in [N].
\end{aligned}
$$

For DRTSLP (1) with objective uncertainty, different from Proposition 1, nonemptiness and compactness assumptions are not required to confirm the tractability.

Proposition 4 (DRTSLP with Objective Uncertainty, [11]) *Suppose $\Xi = \mathbb{R}^{m_1} \times \{\boldsymbol{\xi}_T\}$. DRTSLP (1) can be tractable for any $p \in [1, \infty]$ and admits the following equivalent formulation:*

$$
v^* = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\\ \boldsymbol{y}}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} + \frac{1}{N}\sum_{j\in[N]} \left[\left(\boldsymbol{Q}\boldsymbol{\zeta}_q^j + \boldsymbol{q}\right)^\top \boldsymbol{y}^j + \theta \|\boldsymbol{Q}^\top \boldsymbol{y}^j\|_{p^*} \right] : \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y}^j \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N], \boldsymbol{y}^j \in \mathbb{R}^{n_2}, \forall j \in [N] \right\}.
$$

Following the same assumption as Proposition 2, the tractability result for DRTSLP (1) with left-hand side uncertainty under type ∞-Wasserstein ambiguity set is derived below:

Proposition 5 (DRTSLP with Left-Hand Side Uncertainty, [11]) *Suppose $\Xi = \{\boldsymbol{\xi}_q\} \times \mathbb{R}^{m_2}$ and $p = 1$. Then DRTSLP (1) can be tractable and admits the following equivalent formulation:*

$$\begin{aligned} v^* = \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y},\boldsymbol{\eta}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \frac{1}{N}\sum_{j\in[N]} \eta_j, \\ \text{s.t.} \quad & \eta_j \geq \left(\boldsymbol{Q}\boldsymbol{\xi}_q^j + \boldsymbol{q}\right)^\top \boldsymbol{y}^{ijk}, \forall j \in [N], \forall i \in [m_2], \forall k \in \{-1,1\}, \\ & \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\zeta}_T^j + \boldsymbol{W}\boldsymbol{y}^{ijk} - \theta k \boldsymbol{T}(\boldsymbol{x})\boldsymbol{e}_i \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N], \forall i \in [m_2], \forall k \in \{-1,1\}, \\ & \boldsymbol{y}^{ijk} \in \mathbb{R}^{n_2}, \forall j \in [N], \forall i \in [m_2], \forall k \in \{-1,1\}. \end{aligned}$$

Similarly, this tractability can be directly extended to DRTSLP (1) with right-hand side uncertainty.

Corollary 2 (DRTSLP with Right-Hand Side Uncertainty) *Suppose $\Xi = \{\boldsymbol{\xi}_q\} \times \mathbb{R}^{m_2}$, $p = 1$, and $\boldsymbol{T}(\boldsymbol{x}) := \boldsymbol{T}$. Then DRTSLP (1) can be tractable and admits the following equivalent formulation:*

$$\begin{aligned} v^* = \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y},\boldsymbol{\eta}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \frac{1}{N}\sum_{j\in[N]} \eta_j, \\ \text{s.t.} \quad & \eta_j \geq \left(\boldsymbol{Q}\boldsymbol{\xi}_q^j + \boldsymbol{q}\right)^\top \boldsymbol{y}^{ijk}, \forall j \in [N], \forall i \in [m_2], \forall k \in \{-1,1\}, \\ & \boldsymbol{T}\boldsymbol{\zeta}_T^j + \boldsymbol{W}\boldsymbol{y}^{ijk} - \theta k \boldsymbol{T}\boldsymbol{e}_i \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N], \forall i \in [m_2], \forall k \in \{-1,1\}, \\ & \boldsymbol{y}^{ijk} \in \mathbb{R}^{n_2}, \forall j \in [N], \forall i \in [m_2], \forall k \in \{-1,1\}. \end{aligned}$$

Interested readers can check the work [11] for more tractability results under $\infty-$Wasserstein ambiguity set $\mathcal{P}_\infty^W$.

Type r–Wasserstein Ambiguity Set with $r \in (1, \infty)$

There is limited prior work on DRTSLP (1) under $r-$Wasserstein ambiguity set with $r \in (1, \infty)$. It is worthy of mentioning that [8] discussed DRTSLP (1) under type 2$-$Wasserstein ambiguity set, where the authors used a hierarchy of semidefinite programming approximations to approximate the worst-case expected wait-and-see cost. In this subsection, the tractable formulations with general $r-$Wasserstein ambiguity set with only objective uncertainty or only constraint uncertainty are provided using the duality result from theorem 1 in [4] or theorem 1 in [7].

Proposition 6 (DRTSLP with Objective Uncertainty) *Suppose $\Xi = \mathbb{R}^{m_1} \times \{\boldsymbol{\xi}_T\}$ and for any given $\boldsymbol{x} \in \mathcal{X}$, the feasible region of second-stage problem in DRTSLP (1) (i.e., set $\{\boldsymbol{y} : \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y} \geq \boldsymbol{h}(\boldsymbol{x})\}$) is nonempty and compact. Then DRTSLP (1) can be tractable for any $p \in [1, \infty], r \in (1, \infty)$ and admits the following formulation:*

$$
\begin{aligned}
v^* = \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y},\lambda\geq 0} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \lambda\theta^r + \frac{1}{N}\sum_{j\in[N]}\left[\left(\boldsymbol{Q}\boldsymbol{\zeta}_q^j + \boldsymbol{q}\right)^\top \boldsymbol{y}^j + \|\boldsymbol{Q}^\top \boldsymbol{y}^j\|_{p^*}^{\frac{r}{r-1}} \lambda^{-\frac{1}{r-1}} r^{-\frac{r}{r-1}}(r-1)\right],\\
\text{s.t.} \quad & \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y}^j \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N],\\
& \boldsymbol{y}^j \in \mathbb{R}^{n_2}, \forall j \in [N].
\end{aligned}
$$

Proof Let us first consider the reformulation of $\mathcal{Z}(\boldsymbol{x})$ under $r-$Wasserstein ambiguity set. Applying the duality result from theorem 1 in [4] or theorem 1 in [7], $\mathcal{Z}(\boldsymbol{x})$ can be written as

$$
\mathcal{Z}(\boldsymbol{x}) = \min_{\lambda\geq 0} \lambda\theta^r + \frac{1}{N}\sup_{\boldsymbol{\xi}_q}\sum_{j\in[N]}\min_{\boldsymbol{y}^j}\left\{(\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{y}^j - \lambda\left\|\boldsymbol{\xi}_q - \boldsymbol{\zeta}_q^j\right\|_p^r : \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y}^j \geq \boldsymbol{h}(\boldsymbol{x})\right\}.
$$

Since set $\{\boldsymbol{y} : \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y} \geq \boldsymbol{h}(\boldsymbol{x})\}$ is nonempty and compact, according to Sion's minimax theorem [10], one can interchange the supremum and infimum operators, that is,

$$
\begin{aligned}
&\mathcal{Z}(\boldsymbol{x}) =\\
&\min_{\lambda\geq 0,\boldsymbol{y}}\left\{\lambda\theta^r + \frac{1}{N}\sum_{j\in[N]}\sup_{\boldsymbol{\xi}_q}\left[(\boldsymbol{Q}\boldsymbol{\xi}_q + \boldsymbol{q})^\top \boldsymbol{y}^j - \lambda\left\|\boldsymbol{\xi}_q - \boldsymbol{\zeta}_q^j\right\|_p^r\right] : \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}_T + \boldsymbol{W}\boldsymbol{y}^j \geq \boldsymbol{h}(\boldsymbol{x}), \forall j \in [N]\right\}.
\end{aligned}
$$

Letting $\widehat{\boldsymbol{\zeta}}_q^j = \boldsymbol{\xi}_q - \boldsymbol{\zeta}_q^j$ for each $j \in [N]$ and using the Hölder's inequality and the fact that the supremum is attainable, for each $j \in [N]$, the inner supremum becomes

$$
\sup_{\widehat{\boldsymbol{\zeta}}_q^j}\left[\widehat{\boldsymbol{\zeta}}_q^{j\top}(\boldsymbol{Q}^\top \boldsymbol{y}^j) - \lambda\left\|\widehat{\boldsymbol{\zeta}}_q^j\right\|_p^r\right] = \sup_{\widehat{\boldsymbol{\zeta}}_q^j}\left[\left\|\widehat{\boldsymbol{\zeta}}_q^j\right\|_p \|\boldsymbol{Q}^\top \boldsymbol{y}^j\|_{p^*} - \lambda\left\|\widehat{\boldsymbol{\zeta}}_q^j\right\|_p^r\right] = \|\boldsymbol{Q}^\top \boldsymbol{y}^j\|_{p^*}^{\frac{r}{r-1}}\lambda^{-\frac{1}{r-1}} r^{-\frac{r}{r-1}}(r-1).
$$

Since the term $\|\boldsymbol{Q}^\top \boldsymbol{y}\|_{p^*}^{\frac{r}{r-1}}\lambda^{-\frac{1}{r-1}}$ can be written as

$$
\|\boldsymbol{Q}^\top \boldsymbol{y}\|_{p^*}^{\frac{r}{r-1}}\lambda^{-\frac{1}{r-1}} = \lambda\left(\frac{\|\boldsymbol{Q}^\top \boldsymbol{y}\|_{p^*}}{\lambda}\right)^{1+\frac{1}{r-1}},
$$

and together with the fact that $(\|\boldsymbol{Q}^\top \boldsymbol{y}\|_{p^*}/\lambda)^{1+\frac{1}{r-1}}$ is convex in $\boldsymbol{y}$, according to the convexity of perspective function (see, for example, the convexity result in [5]), the objective function of its corresponding DRTSLP is joint convex in $\boldsymbol{x}, \boldsymbol{y}, \lambda$. Hence, in this case, DRTSLP (1) can be tractable. □

The following proposition shows a sufficient condition under which DRTSLP (1) with right-hand side uncertainty can be tractable.

Proposition 7 (DRTSLP with Right-Hand Side Uncertainty) *Suppose $\Xi = \{\boldsymbol{\xi}_q\} \times \mathbb{R}^{m_2}$, $p = 1$, and $\boldsymbol{T}(\boldsymbol{x}) := \boldsymbol{T}$. Assume that the separations over the epigraphs of the inner supremum functions can be done efficiently and DRTSLP (1) can be tractable for any $r \in (1, \infty)$ and admits the following formulation:*

$$v^* = \min_{\substack{x\in\mathcal{X},\\ \lambda\geq 0}} c^\top x + \lambda\theta^r + \frac{1}{N}\sum_{j\in[N]} \max_{i\in[m_2]} \max \left\{ \begin{array}{l} \sup\limits_{\substack{\pi\geq 0, W^\top\pi = Q\xi_q + q,\\ t\geq 0, t\leq (T^\top\pi)_i}} \left[\left(h(x) - T\zeta_T^j \right)^\top \pi + t^{\frac{r}{r-1}} \lambda^{-\frac{1}{r-1}} r^{-\frac{r}{r-1}} (r-1) \right], \\ \sup\limits_{\substack{\pi\geq 0, W^\top\pi = Q\xi_q + q,\\ t\geq 0, t\leq -(T^\top\pi)_i}} \left[\left(h(x) - T\zeta_T^j \right)^\top \pi + t^{\frac{r}{r-1}} \lambda^{-\frac{1}{r-1}} r^{-\frac{r}{r-1}} (r-1) \right] \end{array} \right\}.$$

Proof Consider the reformulation of $\mathcal{Z}(x)$ under $r-$Wasserstein ambiguity set. Applying the duality result from theorem 1 in [4] or theorem 1 in [7], $\mathcal{Z}(x)$ can be written as

$$\mathcal{Z}(x) = \min_{\lambda\geq 0} \lambda\theta^r + \frac{1}{N}\sum_{j\in[N]} \sup_{\xi_T} \sup_{\pi\geq 0, W^\top\pi = Q\xi_q + q} \left[\left(h(x) - T(x)\xi_T \right)^\top \pi - \lambda \left\| \xi_T - \zeta_T^j \right\|_p^r \right].$$

Switching the supremum operators and letting $\widehat{\zeta}_T^j = \xi_T - \zeta_T^j$ for each $j \in [N]$, the inner supremum becomes

$$\sup_{\widehat{\zeta}_T^j} \left(h(x) - T(x)\zeta_T - T(x)\widehat{\zeta}_T^j \right)^\top \pi - \lambda \left\| \widehat{\zeta}_T^j \right\|_p^r = \left(h(x) - T(x)\zeta_T \right)^\top \pi - \inf_{\widehat{\zeta}_T^j} \left[\left(T(x)\widehat{\zeta}_T^j \right)^\top \pi + \lambda \left\| \widehat{\zeta}_T^j \right\|_p^r \right].$$

According to the Hölder's inequality and the fact that the infimum can be attainable, the infimum above becomes

$$\inf_{\widehat{\zeta}_T^j} \left[\left(T(x)\widehat{\zeta}_T^j \right)^\top \pi + \lambda \left\| \widehat{\zeta}_T^j \right\|_p^r \right] = \inf_{\widehat{\zeta}_T^j} \left[- \left\| \widehat{\zeta}_T^j \right\|_p \left\| T(x)^\top \pi \right\|_{p*} + \lambda \left\| \widehat{\zeta}_T^j \right\|_p^r \right] = \left\| T(x)^\top \pi \right\|_{p*}^{\frac{r}{r-1}} \lambda^{-\frac{1}{r-1}} r^{-\frac{r}{r-1}} (1-r).$$

Therefore, together with the presumptions that $p = 1$ and $T(x) := T$, $\mathcal{Z}(x)$ is equivalent to

$$\mathcal{Z}(x) = \min_{\lambda\geq 0} \lambda\theta^r + \frac{1}{N}\sum_{j\in[N]} \sup_{\pi\geq 0, W^\top\pi = Q\xi_q + q} \left[\left(h(x) - T\zeta_T^j \right)^\top \pi + \left\| T^\top\pi \right\|_\infty^{\frac{r}{r-1}} \lambda^{-\frac{1}{r-1}} r^{-\frac{r}{r-1}} (r-1) \right].$$

Since $\|T^\top\pi\|_\infty = \max_{i\in[m_2]} \max\{(T^\top\pi)_i, -(T^\top\pi)_i\}$, one nonnegative slack variable t is introduced to represent the term $\|T^\top\pi\|_\infty$, and then $\mathcal{Z}(x)$ reduces to

$$\mathcal{Z}(\boldsymbol{x}) = \min_{\lambda \geq 0} \lambda\theta^r + \frac{1}{N}\sum_{j\in[N]} \sup_{\substack{\boldsymbol{\pi}\geq \mathbf{0}, \boldsymbol{W}^\top\boldsymbol{\pi}=\boldsymbol{Q}\boldsymbol{\xi}_q+\boldsymbol{q},\\ t\geq 0, t\leq \max_{i\in[m_2]}|(\boldsymbol{T}^\top\boldsymbol{\pi})|_i}} \left[\left(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{T}\boldsymbol{\zeta}_T^j\right)^\top \boldsymbol{\pi} + t^{\frac{r}{r-1}}\lambda^{-\frac{1}{r-1}}r^{-\frac{r}{r-1}}(r-1)\right].$$

Then, DRTSLP (1) can be written as

$$v^* = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\\ \lambda\geq 0}} \boldsymbol{c}^\top\boldsymbol{x} + \lambda\theta^r + \frac{1}{N}\sum_{j\in[N]} \max_{i\in[m_2]} \max \left\{ \begin{array}{l} \sup\limits_{\substack{\boldsymbol{\pi}\geq \mathbf{0}, \boldsymbol{W}^\top\boldsymbol{\pi}=\boldsymbol{Q}\boldsymbol{\xi}_q+\boldsymbol{q},\\ t\geq 0, t\leq (\boldsymbol{T}^\top\boldsymbol{\pi})_i}} \left[\left(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{T}\boldsymbol{\zeta}_T^j\right)^\top \boldsymbol{\pi} + t^{\frac{r}{r-1}}\lambda^{-\frac{1}{r-1}}r^{-\frac{r}{r-1}}(r-1)\right], \\ \sup\limits_{\substack{\boldsymbol{\pi}\geq \mathbf{0}, \boldsymbol{W}^\top\boldsymbol{\pi}=\boldsymbol{Q}\boldsymbol{\xi}_q+\boldsymbol{q},\\ t\geq 0, t\leq -(\boldsymbol{T}^\top\boldsymbol{\pi})_i}} \left[\left(\boldsymbol{h}(\boldsymbol{x}) - \boldsymbol{T}\boldsymbol{\zeta}_T^j\right)^\top \boldsymbol{\pi} + t^{\frac{r}{r-1}}\lambda^{-\frac{1}{r-1}}r^{-\frac{r}{r-1}}(r-1)\right] \end{array} \right\}.$$

where the objective function is jointly convex in $\boldsymbol{x}, \lambda$.

According to the assumption that the separations over the epigraphs of the inner supremum functions can be done efficiently, one can apply the ellipsoid method to solve the reformulation efficiently. □

It is worthy of mentioning that DRTSLP (1) with left-hand side uncertainty can be intractable and the authors are unable to obtain any tractable result for the general type r-Wasserstein ambiguity set when $r \in (1, \infty)$.

Conclusion

This paper surveyed the tractable reformulations of distributionally robust two-stage linear programs under Wasserstein ambiguity set. New tractable reformulations were also provided for objective and right-hand side uncertainties.

See also

- ▸ Distributionally Robust Optimization in Facility Location Problems
- ▸ L-Shaped Method for Two-Stage Stochastic Programs with Recourse
- ▸ Minimax Theorems
- ▸ Robust Linear Programming with Right-Hand–Side Uncertainty, Duality, and Applications
- ▸ Robust Optimization
- ▸ Two-Stage Stochastic Programs with Recourse

Acknowledgments This research was supported by the National Science Foundation Grant 2046426.

References

1. Ahmed S (2010) Two-stage stochastic integer programming: a brief introduction. Wiley encyclopedia of operations research and management science
2. Ben-Tal A, El Ghaoui L, Nemirovski A (2009) Robust optimization. Princeton University Press, Princeton, New Jersey
3. Bertsimas D, Doan XV, Natarajan K, Teo C-P (2010) Models for minimax stochastic linear optimization problems with risk aversion. Math Oper Res 35(3):580–602
4. Blanchet J, Murthy K (2019) Quantifying distributional model risk via optimal transport. Math Oper Res 44(2):565–600
5. Boyd S, Boyd SP, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge, England
6. Esfahani PM, Kuhn D (2018) Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. Math Program 171(1):115–166
7. Gao R, Kleywegt AJ (2016) Distributionally robust stochastic optimization with Wasserstein distance. arXiv preprint arXiv:1604.02199
8. Hanasusanto GA, Kuhn D (2018) Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls. Oper Res 66(3):849–869

9. Shapiro A, Dentcheva D, Ruszczynski A (2021) Lectures on stochastic programming: modeling and theory. SIAM, Philadelphia
10. Sion M (1958) On general minimax theorems. Pac J Math 8(1):171–176
11. Xie W (2020) Tractable reformulations of two-stage distributionally robust linear programs over the type-∞ Wasserstein ball. Oper Res Lett 48(4):513–523
12. Zhao C, Guan Y (2018) Data-driven risk-averse stochastic optimization with Wasserstein metric. Oper Res Lett 46(2):262–267