Supplementary material for "On Data-Driven Robust Optimization With Multiple Uncertainty Subsets: Unified Uncertainty Set Representation and Mitigating Conservatism"

Yun Li, Neil Yorke-Smith, Tamas Keviczky

1 Parameters Used for *Case Study 1*

The basic uncertainty set for the uncertainty $\mathbf{v}_t$ in each time period t consists two subsets $\mathcal{V}_{t,1}$ and $\mathcal{V}_{t,2}$ that are defined as

$$\mathcal{V}_{t,1} := \{\mathbf{v}_t | \mathbf{D}_1 \mathbf{v}_t \leq \mathbf{d}_1\}, \quad \mathcal{V}_{t,2} := \{\mathbf{v}_t | \mathbf{D}_2 \mathbf{v}_t \leq \mathbf{d}_2\} \tag{1}$$

where

$$\mathbf{D}_1 = \mathbf{D}_2 = [1, -1]^{\mathbf{T}}, \quad \mathbf{d}_1 = [2, 0]^{\mathrm{T}}, \quad \mathbf{d}_2 = [0, 2]^{\mathrm{T}}. \tag{2}$$

The system matrices defining the building thermal dynamics are

$$\boldsymbol{\Phi} = \begin{bmatrix} 0.0167 & 0.0048 & 0.1245 & 0.409 \\ 0.0005 & 0.0002 & 0.0039 & 0.0044 \\ 0.0253 & 0.0073 & 0.3321 & 0.0617 \\ 0.0244 & 0.0070 & 0.0526 & 0.3456 \end{bmatrix}, \boldsymbol{\Gamma}_u = \begin{bmatrix} 0.0986 \\ 0.0029 \\ 0.0288 \\ 0.0275 \end{bmatrix}, \boldsymbol{\Gamma}_w = \begin{bmatrix} 0.2536 & 0.4596 \\ 0.0070 & 0.9840 \\ 0.4450 & 0.1287 \\ 0.4477 & 0.1225 \end{bmatrix}, \boldsymbol{\Gamma}_v = \begin{bmatrix} 0.2536 \\ 0.0070 \\ 0.4450 \\ 0.4477 \end{bmatrix}.$$

2 Parameters Used for *Case Study 2*

The uncertainty set for the demand uncertainties consists of four uncertainty sets $\mathcal{V}_k := \{\mathbf{v} \mid \mathbf{D}_k \mathbf{v} \leq \mathbf{d}_k\}$ with parameters $(\mathbf{D}_k, \mathbf{d}_k)$ set as

$$\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{D}_3 = \mathbf{D}_4 = [\mathbf{I}_3, -\mathbf{I}_3]^{\mathrm{T}} \tag{3}$$

$$\mathbf{d}_1 = [0.3, 0.3, 0.3, 0, 0, 0]^{\mathrm{T}}, \quad \mathbf{d}_2 = [1.2, 1.2, 1.2, -1, -1, -1]^{\mathrm{T}}, \tag{4}$$

$$\mathbf{d}_3 = [1, 0.3, 0.3, -0.8, 0, 0], \quad \mathbf{d}_4 = [0.3, 1, 0.3, 0, -0.7, 0]^{\mathrm{T}}. \tag{5}$$

3 Parameters Used for *Case Study 3*

The mass balance coefficient matrix is

$$\boldsymbol{\kappa} = \begin{bmatrix} 0.63 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0.58 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.25 & -4 & 0 & -1 & -1.5 & 0 \\ 0 & 0.1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0.24 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2.3 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.93 & 0.77 & -1 \end{bmatrix} \tag{6}$$

The uncertainty set consists of four subsets that are defined as

$$\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{D}_3 = \mathbf{D}_4 = \mathbf{I}_5 \otimes \begin{bmatrix} \mathbf{I}_{\text{in}} & \mathbf{0} \\ -\mathbf{I}_{\text{in}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\text{out}} \\ \mathbf{0} & -\mathbf{I}_{\text{out}} \end{bmatrix} \tag{7}$$

$$\begin{aligned} \mathbf{d}_1 &= \mathbf{e}_5 \otimes [130, 112, -110, -88, 35, 49, -25, -41]^{\mathrm{T}}, & (8) \\ \mathbf{d}_2 &= \mathbf{e}_5 \otimes [76, 68, -64, -56, 40, 37, -30 - 27]^{\mathrm{T}}, & (9) \\ \mathbf{d}_3 &= \mathbf{e}_5 \otimes [65, 35, -55, -25, 23, 14, -17, -6]^{\mathrm{T}}, & (10) \\ \mathbf{d}_4 &= \mathbf{e}_5 \otimes [52, 52, -48, -48, 27, 27, -23, -23]^{\mathrm{T}}. & (11) \end{aligned}$$

4 CCG algorithm for solving Wasserstein metric-based DRO problem

For the DRO problem

$$\min_{\mathbf{x}} \mathbf{c}^{\mathrm{T}}\mathbf{x} + \max_{f(\mathbf{v})\in\mathcal{P}} \min_{\mathbf{y}} \mathbf{b}^{\mathrm{T}}\mathbf{y} \tag{12a}$$

$$\text{s.t. } \mathbf{A}\mathbf{x} \le \mathbf{q}, \tag{12b}$$

$$\mathbf{T}\mathbf{x} + \mathbf{W}\mathbf{y} + \mathbf{M}\mathbf{v} \le \mathbf{h} \tag{12c}$$

where the ambiguity set is defined as

$$D_W = \{\mathbb{Q} \in M(\mathcal{V}) : d_W(\mathbb{Q}, \mathbb{Q}_0) \le \varepsilon\} \tag{13}$$

$$d_W(\mathbb{Q}, \mathbb{Q}_0) = \inf_{\Pi} \left(\int_{\mathcal{V}^2} ||\mathbf{v} - \mathbf{v}^0||\ \Pi(\mathrm{d}\mathbf{v}, \mathrm{d}\mathbf{v}^0) \middle| \begin{array}{l} \Pi \text{ is a joint distribution of} \\ \mathbf{v} \text{ and } \mathbf{v}^0 \text{ with marginals } \mathbb{Q} \text{ and } \mathbb{Q}_0 \end{array} \right) \tag{14}$$

with ε as the Wasserstein distance. According to the results in [1], the above optimization can be reformulated as

$$\min_{\mathbf{x},\lambda,\mathbf{z}_i,s_i} \mathbf{c}^{\mathrm{T}}\mathbf{x} + \lambda\varepsilon + \frac{1}{N}\sum_{i=1}^{N} s_i \tag{15a}$$

$$\text{s.t. } \sup_{\mathbf{v}\in\mathcal{U}} \left[\min_{\mathbf{y}} \mathbf{b}^{\mathrm{T}}\mathbf{y} - \mathbf{z}_i^{\mathrm{T}}\mathbf{v}\right] + \mathbf{z}_i^{\mathrm{T}}\hat{\mathbf{v}}_i \leq s_i,\ i = 1,2,\cdots,N, \tag{15b}$$

$$\mathbf{T}\mathbf{x} + \mathbf{W}\mathbf{y} + \mathbf{M}\mathbf{v} \leq \mathbf{h}, \tag{15c}$$

$$\mathbf{A}\mathbf{x} \leq \mathbf{q}, \tag{15d}$$

$$||\mathbf{z}_i||_* \leq \lambda,\ i = 1,2,\cdots,N. \tag{15e}$$

where N is the number of uncertainty samples $\hat{\mathbf{v}}_i$. Then, based on the above formulation and applying a *column-and-constraint generation* approach, the DRO problem can be solved via Algorithm 1.

$$\textbf{MP:} \qquad \min_{\mathbf{x},\lambda,s_i,\mathbf{z}_{t,i},\mathbf{y}_{t,i}} \mathbf{c}^{\mathrm{T}}\mathbf{x} + \lambda\varepsilon + \frac{1}{N}\sum_{i=1}^{N} s_i \tag{16a}$$

$$\text{s.t. } \mathbf{b}^{\mathrm{T}}\mathbf{y}_{t,i} - \mathbf{z}_i^{\mathrm{T}}\mathbf{v}_{t,i}^* + \mathbf{z}_i^{\mathrm{T}}\hat{\mathbf{v}}_i \leq s_i, \tag{16b}$$

$$\mathbf{T}\mathbf{x} + \mathbf{W}\mathbf{y}_{t,i} + \mathbf{M}\mathbf{v}_{t,i} \leq \mathbf{h}, \tag{16c}$$

$$\mathbf{A}\mathbf{x} \leq \mathbf{q}, \tag{16d}$$

$$i = 1,\cdots,N, \quad t = 1,\cdots,r. \tag{16e}$$

$$\textbf{SP:} \qquad \mathcal{C}(\mathbf{x}^*, \mathbf{z}_i^*) := \max_{\mathbf{v}_i} \left[\min_{\mathbf{y}_i} \mathbf{b}^{\mathrm{T}}\mathbf{y}_i - \mathbf{v}_i^{\mathrm{T}}\mathbf{z}_i^*\right] \tag{17a}$$

$$\text{s.t. } \mathbf{T}\mathbf{x}^* + \mathbf{W}\mathbf{y}_i + \mathbf{M}\mathbf{v}_i \leq \mathbf{h}. \tag{17b}$$

References

[1] Mohajerin Esfahani, P., Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. Mathematical Programming, 171(1), 115-166.

Algorithm 1 *column-and-constraint generation* algorithm for DRO problem (15).

Input: suboptimality gap ϵ

Output: the optimal decision variable $\mathbf{x}^*$ and objective function value $\mathbf{c}^{\mathrm{T}}\mathbf{x}^* + \lambda^*\varepsilon + \frac{1}{N}\sum_{i=1}^{N} s_i^*$

Set $LB = -\infty$, $UB = \infty$, $r = 0$

while $|UB - LB| > \epsilon$ **do**

  Solve **MP** in (16) to derive solutions $\{\mathbf{x}^*, \lambda^*, s_i^*, \mathbf{z}_i^*\}$ and update $LB = \mathbf{c}^{\mathrm{T}}\mathbf{x}^* + \lambda^*\varepsilon + \frac{1}{N}\sum_{i=1}^{N} s_i^*$

  for $i = 1, \cdots, N$ **do**

    Solve **SP** $\mathcal{C}(\mathbf{x}^*, \mathbf{z}_i^*)$ in (17) to derive solutions $\{\mathbf{v}_i^*, \mathbf{y}_i^*\}$

  end for

  Update UB as

$$UB = \min\left\{UB, \mathbf{c}^{\mathrm{T}}\mathbf{x}^* + \lambda^*\varepsilon + \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{b}^{\mathrm{T}}\mathbf{y}_i^* - \mathbf{z}_i^{*\mathrm{T}}\mathbf{v}_i^* + \mathbf{z}_i^{*\mathrm{T}}\hat{\mathbf{v}}_i\right)\right\}$$

  Create decision variables $\{\mathbf{y}_{r,i} \mid i = 1, \cdots, N\}$, set parameters $\mathbf{v}_{r,i} = \mathbf{v}_i^*$ $(i = 1, \cdots, N)$, and add the following constraints to **MP** in (16)

$$\begin{cases} \mathbf{b}^{\mathrm{T}}\mathbf{y}_{r,i} - \mathbf{z}_i^{\mathrm{T}}\mathbf{v}_{r,i} + \mathbf{z}_i^{\mathrm{T}}\hat{\mathbf{v}}_i \leq s_i, \\ \mathbf{T}\mathbf{x} + \mathbf{W}\mathbf{y}_{r,i} + \mathbf{M}\mathbf{v}_{r,i}^* \leq \mathbf{h}. \end{cases}$$

  $r \leftarrow r + 1$

end while

Return: $\mathbf{x}^*$ and $\mathbf{c}^{\mathrm{T}}\mathbf{x}^* + \lambda^*\varepsilon + \frac{1}{N}\sum_{i=1}^{N} s_i^*$