Numerical method for approximately optimal solutions of two-stage distributionally robust optimization with marginal constraints — supplementary material

---

The form of (LSIP^*) inspires the following approach for setting the input $\mathfrak{C}^{(0)}$ in Algorithm 2. For any finitely supported probability measure $\hat{\mu} = \sum_{j=1}^{J} \alpha_j \delta_{\boldsymbol{x}_j} \in \mathcal{P}(\boldsymbol{\mathcal{X}})$ which satisfies $\hat{\mu} \in \Gamma(\nu_1, \ldots, \nu_N)$ with $\nu_i \in [\mu_i]_{\mathcal{G}_i}$ for $i = 1, \ldots, N$, one can solve the following linear programming problem:

$$\begin{aligned} \underset{\boldsymbol{\xi}_{\mathrm{in}}, \boldsymbol{\xi}_{\mathrm{eq}}, (\boldsymbol{\lambda}_j)}{\text{maximize}} \quad & \langle \boldsymbol{q}_{\mathrm{in}}, \boldsymbol{\xi}_{\mathrm{in}} \rangle + \langle \boldsymbol{q}_{\mathrm{eq}}, \boldsymbol{\xi}_{\mathrm{eq}} \rangle + \sum_{j=1}^{J} \langle \alpha_j (\mathbf{W} \boldsymbol{x}_j + \boldsymbol{b}), \boldsymbol{\lambda}_j \rangle \\ \text{subject to} \quad & \mathbf{L}_{\mathrm{in}}^{\mathsf{T}} \boldsymbol{\xi}_{\mathrm{in}} + \mathbf{L}_{\mathrm{eq}}^{\mathsf{T}} \boldsymbol{\xi}_{\mathrm{eq}} - \left(\sum_{j=1}^{J} \alpha_j \mathbf{V}^{\mathsf{T}} \boldsymbol{\lambda}_j \right) = \boldsymbol{c}_1, \\ & \boldsymbol{\lambda}_j \in S_2^* \qquad \forall 1 \le j \le N, \\ & \boldsymbol{\xi}_{\mathrm{in}} \in \mathbb{R}_-^{n_{\mathrm{in}}}, \; \boldsymbol{\xi}_{\mathrm{eq}} \in \mathbb{R}^{n_{\mathrm{eq}}}. \end{aligned}$$

Suppose that this problem is feasible, and that $\hat{\boldsymbol{\xi}}_{\mathrm{in}}, \hat{\boldsymbol{\xi}}_{\mathrm{eq}}, (\hat{\boldsymbol{\lambda}}_j)_{j=1:J}$ is an optimal solution (notice that this problem is bounded from above due to Theorem 3.16), one can then define $\hat{\mu}_{\mathrm{aug}} := \sum_{j=1}^{J} \alpha_j \delta_{(\boldsymbol{x}_j, \hat{\boldsymbol{\lambda}}_j)} \in \mathcal{P}(\boldsymbol{\mathcal{X}} \times S_2^*)$. One may check that $(\hat{\boldsymbol{\xi}}_{\mathrm{in}}, \hat{\boldsymbol{\xi}}_{\mathrm{eq}}, \hat{\mu}_{\mathrm{aug}})$ is feasible for (LSIP^*). Subsequently, one can let $\mathfrak{C}^{(0)} := \left\{ (\boldsymbol{x}_j^{\mathsf{T}}, \hat{\boldsymbol{\lambda}}_j^{\mathsf{T}})^{\mathsf{T}} : 1 \le j \le J \right\}$. Even though there is no theoretical guarantee that ($\mathsf{LSIP}_{\mathsf{relax}}(\mathfrak{C}^{(0)})$) has bounded sublevel sets when $\mathfrak{C}^{(0)}$ is chosen this way, we have not experienced any issue with the convergence of Algorithm 2 when we used this approach for solving the numerical examples in Section 5.