Multistage Distributionally Robust Mixed-Integer Programming with Decision-Dependent Moment-based Ambiguity Sets

Siqian Shen

Department of Industrial and Operations Engineering

University of Michigan

Joint work with Xian Yu, PhD student in IOE@UMich

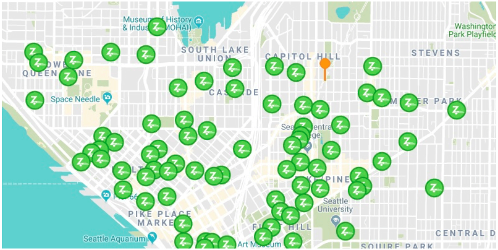
supported by NSF grants #1727618, #1709094, and DoE grant #DE-SC0018018

1. Introduction
2. Multistage DRO with Endogenous Uncertainty
   - Type 1: decision-dependent bounds on moments
   - Type 2: the mean and covariance matrix exactly matching empirical ones
   - Type 3: general decision-dependent moments
3. Computational Studies
4. Application: COVID-19 Test Kit Allocation

Why Consider Endogenous Uncertainty (EU)?

In some applications, uncertain parameters of systems depend on decisions made over time (e.g., rental facilities open sequentially over time).

(a) Zipcar parking lots



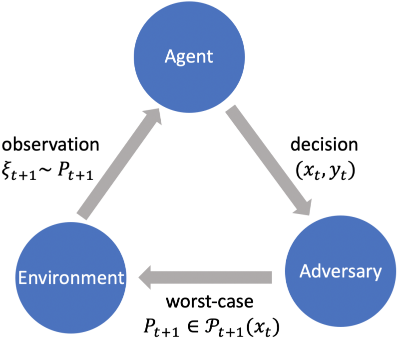
(b) Zipcar in Seattle

A multistage decision making process where each stage is a mixed-integer program involving

- Binary planning variables, e.g., where to locate facilities.
- Continuous/integer operational variables.

Dynamic Process

- $\boldsymbol{x}_t \in \{0,1\}^I$: binary state variable that connects the consecutive two stages t and $t+1$
- $\boldsymbol{y}_t \in \mathbb{R}^{I\times J}$: continuous/integer stage variable that only appears at stage t
- $\boldsymbol{\xi}_t \in \Xi_t \subset \mathbb{R}^J$: random vector at stage t
- $\mathcal{P}_t(\boldsymbol{x}_{t-1})$: ambiguity set at stage t



$$\overbrace{\text{decision } (x_1, y_1)}^{\text{Stage 1}} \rightarrow \overbrace{\text{worst-case } (P_{\xi_2} \in \mathcal{P}_2(x_1)) \rightarrow \text{parameter } (\xi_2) \rightarrow \text{decision } (x_2, y_2)}^{\text{Stage 2}} \rightarrow \cdots$$

$$\rightarrow \underbrace{\text{worst-case } (P_{\xi_T} \in \mathcal{P}_T(x_T)) \rightarrow \text{parameter } (\xi_T) \rightarrow \text{decision } (x_T, y_T)}_{\text{Stage T}}$$

Multistage Distributionally Robust Optimization

DDDR:

$$\min_{(\boldsymbol{x}_1,\boldsymbol{y}_1)\in X_1}\Big\{g_1(\boldsymbol{x}_1,\boldsymbol{y}_1)+\max_{P_2\in\mathcal{P}_2(\boldsymbol{x}_1)}\mathbb{E}_{P_2}\Big[\min_{(\boldsymbol{x}_2,\boldsymbol{y}_2)\in X_2(\boldsymbol{x}_1,\boldsymbol{\xi}_2)}g_2(\boldsymbol{x}_2,\boldsymbol{y}_2)+\cdots$$
$$+\max_{P_T\in\mathcal{P}_T(\boldsymbol{x}_{T-1})}\mathbb{E}_{P_T}\Big[\min_{(\boldsymbol{x}_T,\boldsymbol{y}_T)\in X_T(\boldsymbol{x}_{T-1},\boldsymbol{\xi}_T)}g_T(\boldsymbol{x}_T,\boldsymbol{y}_T)\Big]\Big]\Big\},$$

- $g_t(x_t, y_t)$: objective function in stage t; linear in x_t and y_t.
- ξ_t has a decision independent finite support $\Xi_t := \{\xi_t^k\}_{k=1}^K$ with each realization $k \in [K]$ having probability p_k.

Multistage Distributionally Robust Optimization

DDDR:

$$\min_{(\boldsymbol{x}_1,\boldsymbol{y}_1)\in X_1}\Big\{g_1(\boldsymbol{x}_1,\boldsymbol{y}_1)+\max_{P_2\in\mathcal{P}_2(\boldsymbol{x}_1)}\mathbb{E}_{P_2}\Big[\min_{(\boldsymbol{x}_2,\boldsymbol{y}_2)\in X_2(\boldsymbol{x}_1,\boldsymbol{\xi}_2)} g_2(\boldsymbol{x}_2,\boldsymbol{y}_2)+\cdots$$
$$+\max_{P_T\in\mathcal{P}_T(\boldsymbol{x}_{T-1})}\mathbb{E}_{P_T}\Big[\min_{(\boldsymbol{x}_T,\boldsymbol{y}_T)\in X_T(\boldsymbol{x}_{T-1},\boldsymbol{\xi}_T)} g_T(\boldsymbol{x}_T,\boldsymbol{y}_T)\Big]\Big]\Big\},$$

- $g_t(x_t, y_t)$: objective function in stage t; linear in x_t and y_t.
- ξ_t has a decision independent finite support $\Xi_t := \{\xi_t^k\}_{k=1}^K$ with each realization $k \in [K]$ having probability p_k.

Considered in this talk:

- Decision maker's risk attitude: risk-neutral v.s. risk-averse
- Support Ξ_t: discrete v.s. continuous
- Ambiguity set $\mathcal{P}_t(\boldsymbol{x}_{t-1})$:
  - Type 1 : decision-dependent bounds on moments
  - Type 2 : the mean and covariance matrix exactly matching empirical ones
  - Type 3 : general decision-dependent moment set

Ambiguity Sets with Bounded Moment Functions[1]

Given empirical demand mean $\bar{\mu}_j$, standard deviation $\bar{\sigma}_j$ and tolerance parameters $\epsilon_j^\mu, \underline{\epsilon}^S, \bar{\epsilon}_j^S$, define the ambiguity set in stage $t+1$ as

$$\mathcal{P}_{t+1}^D(\boldsymbol{x}_t) = \Big\{ \boldsymbol{p} \in \mathbb{R}_+^K \mid \sum_{k=1}^K p_k = 1, \tag{1a}$$

$$\mu_j(\boldsymbol{x}_t) - \epsilon_j^\mu \leq \sum_{k=1}^K p_k \xi_{t+1,j}^k \leq \mu_j(\boldsymbol{x}_t) + \epsilon_j^\mu, \ \forall j \in [J], \tag{1b}$$

$$S_j(\boldsymbol{x}_t)\underline{\epsilon}_j^S \leq \sum_{k=1}^K p_k (\xi_{t+1,j}^k)^2 \leq S_j(\boldsymbol{x}_t)\bar{\epsilon}_j^S, \ \forall j \in [J] \Big\}, \tag{1c}$$

where for each $j \in [J]$,
$\mu_j(x_t) = \bar{\mu}_j(1 + \sum_{i\in[I]} \lambda_{ji}^\mu x_{ti}),\ S_j(x_t) = (\bar{\mu}_j^2 + \bar{\sigma}_j^2)(1 + \sum_{i\in[I]} \lambda_{ji}^S x_{ti})$.
$\lambda_{ji}^\mu, \lambda_{ji}^S \in [0, 1]$: effects of building a facility at i on the moments of j

[1] Basciftci, B., Ahmed, S., & Shen, S. (2020). Distributionally robust facility location problem under decision-dependent stochastic demand. Forthcoming in *EJOR*, 2020.
https://doi.org/10.1016/j.ejor.2020.11.002

Ambiguity Sets with Bounded Moment Functions[2]

In general, Type 1 ambiguity set is in the form of

$$\mathcal{P}^D_{t+1}(\boldsymbol{x}_t) := \left\{ \boldsymbol{p} \in \mathbb{R}^K \mid \underline{\boldsymbol{p}}(\boldsymbol{x}_t) \leq \boldsymbol{p} \leq \bar{\boldsymbol{p}}(\boldsymbol{x}_t),\ \boldsymbol{l}(\boldsymbol{x}_t) \leq \sum_{k=1}^K p_k \boldsymbol{f}(\boldsymbol{\xi}_{t+1}) \leq \boldsymbol{u}(\boldsymbol{x}_t) \right\},$$

Bellman Equation:

$$Q_t(x_{t-1}, \xi_t) = \min_{(x_t, y_t) \in X_t(x_{t-1}, \xi_t)} g_t(x_t, y_t) + \max_{P_{t+1} \in \mathcal{P}(x_t)} \mathbb{E}_{P_{t+1}}[Q_{t+1}(x_t, \xi_{t+1})]$$

Theorem 1 (Yu and S. (2020))

If for any feasible $\boldsymbol{x_t} \in X_t$, the ambiguity set is nonempty, then the Bellman equation can be reformulated as a mixed-integer nonlinear program (MINLP):

$$\begin{aligned} Q_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t) = \min_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \underline{\boldsymbol{\gamma}}, \bar{\boldsymbol{\gamma}}, \boldsymbol{x}_t, \boldsymbol{y}_t} \quad & g_t(\boldsymbol{x}_t, \boldsymbol{y}_t) - \boldsymbol{\alpha}^\top \boldsymbol{l}(\boldsymbol{x}_t) + \boldsymbol{\beta}^\top \boldsymbol{u}(\boldsymbol{x}_t) - \underline{\boldsymbol{\gamma}}^\top \underline{\boldsymbol{p}}(\boldsymbol{x}_t) + \bar{\boldsymbol{\gamma}}^\top \bar{\boldsymbol{p}}(\boldsymbol{x}_t) \\ \text{s.t.} \quad & (-\boldsymbol{\alpha} + \boldsymbol{\beta})^\top \boldsymbol{f}(\boldsymbol{\xi}^k_{t+1}) - \underline{\gamma}_k + \bar{\gamma}_k \geq Q_{t+1}(\boldsymbol{x}_t, \boldsymbol{\xi}^k_{t+1}),\ \forall k \in [K], \\ & (\boldsymbol{x}_t, \boldsymbol{y}_t) \in X_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t), \\ & \boldsymbol{\alpha},\ \boldsymbol{\beta},\ \underline{\boldsymbol{\gamma}},\ \bar{\boldsymbol{\gamma}} \geq 0. \end{aligned}$$

[2] Luo, F., & Mehrotra, S. (2020). Distributionally robust optimization with decision dependent ambiguity sets. *Optimization Letters*, 1-30.

Specific Bellman Equation

$$Q_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t) =$$

$$\begin{aligned}
\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{x}_t,\boldsymbol{y}_t} \quad & g_t(\boldsymbol{x}_t, \boldsymbol{y}_t) - \alpha_1 - \sum_{j\in[J]} \alpha_{2j}(\bar{\mu}_j - \epsilon_j^{\mu}) - \sum_{j\in[J]}\sum_{i\in[I]} \lambda_{ji}^{\mu}\bar{\mu}_j \alpha_{2j} x_{ti} \\
& - \sum_{j\in[J]} \alpha_{3j}(\bar{\mu}_j^2 + \bar{\sigma}_j^2)\underline{\epsilon}_j^S - \sum_{j\in[J]}\sum_{i\in[I]} \lambda_{ji}^S \underline{\epsilon}_j^S (\bar{\mu}_j^2 + \bar{\sigma}_j^2)\alpha_{3j} x_{ti} \\
& + \beta_1 + \sum_{j\in[J]} \beta_{2j}(\bar{\mu}_j + \epsilon_j^{\mu}) + \sum_{j\in[J]}\sum_{i\in[I]} \lambda_{ji}^{\mu}\bar{\mu}_j \beta_{2j} x_{ti} \\
& + \sum_{j\in[J]} \beta_{3j}(\bar{\mu}_j^2 + \bar{\sigma}_j^2)\bar{\epsilon}_j^S + \sum_{j\in[J]}\sum_{i\in[I]} \lambda_{ji}^S \bar{\epsilon}_j^S (\bar{\mu}_j^2 + \bar{\sigma}_j^2)\beta_{3j} x_{ti} \\
\text{s.t.} \quad & -\alpha_1 + \beta_1 + \sum_{j\in[J]} \xi_{t+1,j}^k(-\alpha_{2j} + \beta_{2j}) + \sum_{j\in[J]} (\xi_{t+1,j}^k)^2(-\alpha_{3j} + \beta_{3j}) \geq Q_{t+1}^k, \ \forall k \in [K] \\
& (\boldsymbol{x}_t, \boldsymbol{y}_t) \in X_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t), \\
& \boldsymbol{\alpha}, \ \boldsymbol{\beta} \geq 0.
\end{aligned}$$

nonlinear

where $Q_{t+1}^k := Q_{t+1}(\boldsymbol{x}_t, \boldsymbol{\xi}_{t+1}^k)$.

- Bilinear terms: McCormick envelopes, Q_{t+1}^k: approximate by linear cuts.

Reformulation of Specific Bellman Equation

$$
\begin{aligned}
\underline{Q}_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t) = \\
\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{x}_t,\boldsymbol{y}_t} \quad & g_t(\boldsymbol{x}_t, \boldsymbol{y}_t) - \alpha_1 - \sum_{j\in[J]} \alpha_{2j}(\bar{\mu}_j - \epsilon^{\mu}_j) - \sum_{j\in[J]}\sum_{i\in[I]} \lambda^{\mu}_{ji}\bar{\mu}_j z^{\alpha_2}_{tji} \\
& - \sum_{j\in[J]} \alpha_{3j}(\bar{\mu}^2_j + \bar{\sigma}^2_j)\underline{\epsilon}^S_j - \sum_{j\in[J]}\sum_{i\in[I]} \lambda^S_{ji}\underline{\epsilon}^S_j(\bar{\mu}^2_j + \bar{\sigma}^2_j) z^{\alpha_3}_{tji} \\
& + \beta_1 + \sum_{j\in[J]} \beta_{2j}(\bar{\mu}_j + \epsilon^{\mu}_j) + \sum_{j\in[J]}\sum_{i\in[I]} \lambda^{\mu}_{ji}\bar{\mu}_j z^{\beta_2}_{tji} \\
& + \sum_{j\in[J]} \beta_{3j}(\bar{\mu}^2_j + \bar{\sigma}^2_j)\bar{\epsilon}^S_j + \sum_{j\in[J]}\sum_{i\in[I]} \lambda^S_{ji}\bar{\epsilon}^S_j(\bar{\mu}^2_j + \bar{\sigma}^2_j) z^{\beta_3}_{tji} \\
\text{s.t.} \quad & -\alpha_1 + \beta_1 + \sum_{j\in[J]} \xi^k_{t+1,j}(-\alpha_{2j} + \beta_{2j}) + \sum_{j\in[J]} (\xi^k_{t+1,j})^2(-\alpha_{3j} + \beta_{3j}) \ge \theta^k_t, \ \forall k \in [K], \\
& \theta^k_t \ge v^{lk}_{t+1} + (\boldsymbol{\pi}^{lk}_{t+1})^{\mathsf{T}} \boldsymbol{x}_t, \ \forall k \in [K], \ l \in [\ell], \\
& (\boldsymbol{x}_t, \boldsymbol{y}_t) \in X_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t), \\
& (z^{\alpha_2}_{tji}, \alpha_{2j}, x_{ti}) \in M^{\alpha_2}_{tji}, \ (z^{\alpha_3}_{tji}, \alpha_{3j}, x_{ti}) \in M^{\alpha_3}_{tji}, \ \forall i \in [I], \ j \in [J], \\
& (z^{\beta_2}_{tji}, \beta_{2j}, x_{ti}) \in M^{\beta_2}_{tji}, \ (z^{\beta_3}_{tji}, \beta_{3j}, x_{ti}) \in M^{\beta_3}_{tji}, \ \forall i \in [I], \ j \in [J], \\
& \boldsymbol{\alpha}, \ \boldsymbol{\beta} \ge 0.
\end{aligned}
$$

Mixed-integer
Linear Program

Ambiguity Set with Exact First and Second Moments

Ambiguity sets defined by exact mean and covariance:

$$\mathcal{P}(\boldsymbol{x_t}) := \Big\{ p \in \mathbb{R}_+^K : \quad \sum_{k=1}^{K} p_k = 1 \tag{2a}$$

$$\mathbb{E}[\boldsymbol{\xi}_{t+1}] = \boldsymbol{\mu}(\boldsymbol{x_t}) \tag{2b}$$

$$\mathbb{E}[(\boldsymbol{\xi}_{t+1} - \boldsymbol{\mu}(\boldsymbol{x_t}))(\boldsymbol{\xi}_{t+1} - \boldsymbol{\mu}(\boldsymbol{x_t}))^{\mathsf{T}}] = \Sigma(\boldsymbol{x_t}) \Big\} \tag{2c}$$

Theorem 2 (Yu and S. (2020))

If for any feasible $\boldsymbol{x_t} \in X_t$, the ambiguity set defined in (2) is nonempty, then the Bellman equation can be reformulated as $Q_t(\boldsymbol{x_{t-1}}, \boldsymbol{\xi_t}) =$

$$\begin{aligned} \min_{\boldsymbol{x_t}, \boldsymbol{y_t}, s, \boldsymbol{u}, \boldsymbol{Y}} \quad & g_t(\boldsymbol{x_t}, \boldsymbol{y_t}) + s + \boldsymbol{u}^{\mathsf{T}} \boldsymbol{\mu}(\boldsymbol{x_t}) + \Sigma(\boldsymbol{x_t}) \bullet \boldsymbol{Y} \\ \text{s.t.} \quad & s + \boldsymbol{u}^{\mathsf{T}} \boldsymbol{\xi}_{t+1}^{k} + (\boldsymbol{\xi}_{t+1}^{k} - \boldsymbol{\mu}(\boldsymbol{x_t}))(\boldsymbol{\xi}_{t+1}^{k} - \boldsymbol{\mu}(\boldsymbol{x_t}))^{\mathsf{T}} \bullet \boldsymbol{Y} \geq Q_{t+1}^{k}, \ \forall k \\ & (\boldsymbol{x_t}, \boldsymbol{y_t}) \in X_t(\boldsymbol{x_{t-1}}, \boldsymbol{\xi_t}) \end{aligned}$$

where $A \bullet B = \text{trace}(A^{\mathsf{T}} B)$.

McCormick Envelopes

If we assume that the elements in $\boldsymbol{\mu}(\boldsymbol{x_t})$, $\Sigma(\boldsymbol{x_t})$ are affine in $\boldsymbol{x_t}$, i.e.,

$$\mu_j(\boldsymbol{x_t}) = \bar{\mu}_j(1 + \sum_{i \in [I]} \lambda_{ji}^{\mu} x_{ti})$$

$$\Sigma(\boldsymbol{x_t}) = \bar{\Sigma}(1 + \sum_{i \in [I]} \lambda_{i}^{cov} x_{ti})$$

where $\lambda_{ji}^{\mu}, \lambda_i^{cov} \in [0, 1]$ are effects of building a facility at i on the moments, then

$$\Sigma(\boldsymbol{x_t}) \bullet Y = \sum_{i=1}^{d} \sum_{j=1}^{d} \bar{\Sigma}_{ij}(1 + \sum_{s \in [I]} \sigma_s x_{ts}) y_{ij}$$

$$\boldsymbol{\mu}(\boldsymbol{x_t})\boldsymbol{\mu}(\boldsymbol{x_t})^{\mathsf{T}} \bullet Y = \sum_{i=1}^{d} \sum_{j=1}^{d} \bar{\mu}_i \bar{\mu}_j (1 + \sum_{s \in [I]} \lambda_{is} x_{ts})(1 + \sum_{s \in [I]} \lambda_{js} x_{ts}) y_{ij}$$

Bellman equation $\Rightarrow$ mixed-integer linear program (MILP)

Ambiguity Sets with General Decision-dependent Moments

We extend the general moment-based ambiguity set commonly used in the DRO literature [3] into the Type 3 decision-dependent moment set:

$$\mathcal{P}^D_{t+1}(\boldsymbol{x}_t) := \Bigg\{ \boldsymbol{p} \in \mathbb{R}^K \mid \sum_{k=1}^K p_k = 1, \\ \left(\sum_{k=1}^K p_k \boldsymbol{\xi}^k_{t+1} - \boldsymbol{\mu}(\boldsymbol{x}_t)\right)^\mathsf{T} \Sigma(\boldsymbol{x}_t)^{-1} \left(\sum_{k=1}^K p_k \boldsymbol{\xi}^k_{t+1} - \boldsymbol{\mu}(\boldsymbol{x}_t)\right) \leq \gamma, \\ \sum_{k=1}^K p_k (\boldsymbol{\xi}^k_{t+1} - \boldsymbol{\mu}(\boldsymbol{x}_t))(\boldsymbol{\xi}^k_{t+1} - \boldsymbol{\mu}(\boldsymbol{x}_t))^\mathsf{T} \preceq \eta \Sigma(\boldsymbol{x}_t) \Bigg\}.$$

where $\boldsymbol{\mu}(\boldsymbol{x_t})$, $\Sigma(\boldsymbol{x_t})$ are empirical mean vector and covariance matrix:

$$\mu_j(\boldsymbol{x}_t) = \bar{\mu}_j(1 + \sum_{i \in [I]} \lambda^{\mu}_{ji} x_{ti}), \ \forall j \in [J],$$

$$\Sigma(\boldsymbol{x}_t) = \bar{\Sigma}(1 + \sum_{i \in [I]} \lambda^{cov}_{i} x_{ti}).$$

$\lambda^{\mu}_{ji}, \lambda^{cov}_{i} \in [0, 1]$: effects of building a facility at i on the moments.

[3] Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3), 595-612.

Reformulation of Bellman Equations

Theorem 3 (Yu and S. (2020))

Suppose that Slater's constraint qualification conditions are satisfied, i.e., for any feasible $\boldsymbol{x_t}$, there exists a vector $p = [p_1, p_2, \ldots, p_K]$ such that $\sum_{k=1}^{K} p_k = 1$, $(\mathbb{E}[\boldsymbol{\xi_{t+1}}] - \boldsymbol{\mu}(\boldsymbol{x_t}))^{\mathsf{T}}\Sigma(\boldsymbol{x_t})^{-1}(\mathbb{E}[\boldsymbol{\xi_{t+1}}] - \boldsymbol{\mu}(\boldsymbol{x_t})) < \gamma_t$, $\mathbb{E}[(\boldsymbol{\xi_{t+1}} - \boldsymbol{\mu}(\boldsymbol{x_t}))(\boldsymbol{\xi_{t+1}} - \boldsymbol{\mu}(\boldsymbol{x_t}))^{\mathsf{T}}] \prec \eta_t \Sigma(\boldsymbol{x_t})$. Using the Type 3 ambiguity set, the Bellman equation can be recast as $Q_t(\boldsymbol{x_{t-1}}, \boldsymbol{\xi_t}) =$

$$
\begin{aligned}
\min_{\boldsymbol{x_t}, \boldsymbol{y_t}, s, \boldsymbol{Z}, \boldsymbol{Y}} \quad & g_t(\boldsymbol{x_t}, \boldsymbol{y_t}) + s + \Sigma(\boldsymbol{x_t}) \bullet \boldsymbol{z}_1 - 2\boldsymbol{\mu}(\boldsymbol{x_t})^{\mathsf{T}}\boldsymbol{z}_2 + \gamma_t z_3 + \eta_t \Sigma(\boldsymbol{x_t}) \bullet \boldsymbol{Y} \\
\text{s.t.} \quad & s - 2{\boldsymbol{z}_2}^{\mathsf{T}}\boldsymbol{\xi}^{\boldsymbol{k}}_{t+1} + (\boldsymbol{\xi}^{\boldsymbol{k}}_{t+1} - \boldsymbol{\mu}(\boldsymbol{x_t}))(\boldsymbol{\xi}^{\boldsymbol{k}}_{t+1} - \boldsymbol{\mu}(\boldsymbol{x_t}))^{\mathsf{T}} \bullet Y \geq Q^k_{t+1}, \ \forall k \\
& \boldsymbol{Z} = \begin{pmatrix} \boldsymbol{z}_1 & \boldsymbol{z}_2 \\ {\boldsymbol{z}_2}^{\mathsf{T}} & z_3 \end{pmatrix} \succeq 0 \\
& \boldsymbol{Y} \succeq 0 \\
& (\boldsymbol{x_t}, \boldsymbol{y_t}) \in X_t(\boldsymbol{x_{t-1}}, \boldsymbol{\xi_t})
\end{aligned}
$$

where $A \bullet B = \text{trace}(A^{\mathsf{T}}B)$.

Solving the MISDP Reformulation

We linearize $\Sigma(\boldsymbol{x_t}) \bullet \boldsymbol{z}_1,\ \mu(\boldsymbol{x_t})^\top \boldsymbol{z}_2,\ \Sigma(\boldsymbol{x_t}) \bullet \boldsymbol{Y},\ \mu(\boldsymbol{x_t})\mu(\boldsymbol{x_t})^\top \bullet \boldsymbol{Y}$, and then the Bellman equation becomes a mixed-integer semidefinite program (MISDP):

$$Q_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t) = \min_{\substack{\boldsymbol{x}_t, \boldsymbol{y}_t, s, \boldsymbol{Z}, \boldsymbol{Y} \\ \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}}} \quad \tilde{g}_t(\boldsymbol{x}_t, \boldsymbol{y}_t, s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}) \tag{3a}$$

$$\text{s.t.} \quad f_t(s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{v}, \boldsymbol{\xi}^k_{t+1}) \geq Q^k_{t+1},\ \forall k \in [K], \tag{3b}$$

$$(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}) \in \tilde{X}_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t), \tag{3c}$$

$$\boldsymbol{Z} \succeq 0,\ \boldsymbol{Y} \succeq 0. \tag{3d}$$

Solving the MISDP Reformulation

We linearize $\Sigma(\boldsymbol{x_t}) \bullet \boldsymbol{z}_1,\ \boldsymbol{\mu}(\boldsymbol{x_t})^\top \boldsymbol{z}_2,\ \Sigma(\boldsymbol{x_t}) \bullet \boldsymbol{Y},\ \boldsymbol{\mu}(\boldsymbol{x_t})\boldsymbol{\mu}(\boldsymbol{x_t})^\top \bullet \boldsymbol{Y}$, and then the Bellman equation becomes a mixed-integer semidefinite program (MISDP):

$$Q_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t) = \min_{\substack{\boldsymbol{x}_t, \boldsymbol{y}_t, s, \boldsymbol{Z}, \boldsymbol{Y} \\ \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}}} \quad \tilde{g}_t(\boldsymbol{x}_t, \boldsymbol{y}_t, s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}) \tag{3a}$$

$$\text{s.t.} \quad f_t(s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{v}, \boldsymbol{\xi}_{t+1}^k) \geq Q_{t+1}^k, \ \forall k \in [K], \tag{3b}$$

$$(\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}) \in \tilde{X}_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t), \tag{3c}$$

$$\boldsymbol{Z} \succeq 0, \ \boldsymbol{Y} \succeq 0. \tag{3d}$$

The MISDP itself is difficult to solve directly, and we are dealing with a multistage stochastic MISDP. Next, we approximate the problem by

- lower bounding (3) via *Relaxed Lagrangian Cuts*;
- upper bounding (3) via inner approximating MISDP by MILPs.

Lower bounding via Relaxed Lagrangian Cuts[4]

We make a copy of the state variable $\boldsymbol{z}_t = \boldsymbol{x}_{t-1}$ and then relax it to get a Lagrangian function. Specifically, we solve the following problem in the backward step to get *Relaxed Lagrangian Cuts* $\{(v_t^{lk}, \pi_t^{lk})\}_{k=1}^K$ where $v_t^{\ell k} = \mathcal{L}_t^k(\pi_t^{\ell k})$:

$$
\begin{aligned}
\mathcal{L}_t^k(\boldsymbol{\pi}_t) = \min_{\substack{\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{z}_t, s, \boldsymbol{Z}, \boldsymbol{Y} \\ \boldsymbol{\theta}_t, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}}} \quad & \tilde{g}_t(\boldsymbol{x}_t, \boldsymbol{y}_t, s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}) - \boldsymbol{\pi}_t^\top \boldsymbol{z}_t \\
\text{s.t.} \quad & f_t(s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{v}, \boldsymbol{\xi}_{t+1}^{k'}) \geq \theta_t^{k'}, \ \forall k' \in [K], \\
& \theta_t^k \geq v_{t+1}^{lk} + (\boldsymbol{\pi}_{t+1}^{lk})^\top \boldsymbol{x}_t, \ \forall k \in [K], \ l \in [\ell], \\
& (\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}) \in \tilde{X}_t(\boldsymbol{z}_t, \boldsymbol{\xi}_t^k), \\
& \boldsymbol{Z} \succeq 0, \ \boldsymbol{Y} \succeq 0.
\end{aligned}
$$

Proposition

The collection of Relaxed Lagrangian Cuts $\{(v_t^{lk}, \pi_t^{lk})\}_{k=1}^K$ is valid because the true value function is bounded from below by these cuts for all $\boldsymbol{x}_{t-1}$, i.e., $Q_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t^k) \geq v_t^{lk} + (\boldsymbol{\pi}_t^{lk})^\top \boldsymbol{x}_{t-1}$ for all $\boldsymbol{x}_{t-1} \in \{0,1\}^I$.

[4] Zou, J., Ahmed, S., & Sun, X. A. (2019). Stochastic dual dynamic integer programming. *Mathematical Programming*, 175(1-2), 461-502.

Upper bounding via inner approximating MISDP by MILPs[5]

Definition

A symmetric matrix A is *diagonally dominant* if $a_{ii} \geq \sum_{j \neq i} |a_{ij}|$ for all i.

$$DD(\boldsymbol{U}) := \{\boldsymbol{M} \in S_n \mid \boldsymbol{M} = \boldsymbol{U}^\top \boldsymbol{Q} \boldsymbol{U} \text{ for some dd matrix } \boldsymbol{Q}\} \subset P_n.$$

At iteration l, we replace the conditions $\boldsymbol{Z} \succeq 0$ and $\boldsymbol{Y} \succeq 0$ by $\boldsymbol{Z} \in DD(\boldsymbol{U})$, $\boldsymbol{Y} \in DD(\boldsymbol{V})$ for some fixed matrices $\boldsymbol{U}, \boldsymbol{V}$:

$$\begin{aligned} \overline{Q}_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t) = \min_{\substack{\boldsymbol{x}_t, \boldsymbol{y}_t, s, \boldsymbol{Z}, \boldsymbol{Y} \\ \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}}} \quad & \tilde{g}_t(\boldsymbol{x}_t, \boldsymbol{y}_t, s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}) \\ \text{s.t.} \quad & f_t(s, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{v}, \boldsymbol{\xi}_{t+1}^k) \geq \overline{Q}_{t+1}(\boldsymbol{x}_t, \boldsymbol{\xi}_{t+1}^k), \ \forall k \in [K], \\ & (\boldsymbol{x}_t, \boldsymbol{y}_t, \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{v}) \in \tilde{X}_t(\boldsymbol{x}_{t-1}, \boldsymbol{\xi}_t), \\ & \boldsymbol{Z} \in DD(\boldsymbol{U}), \ \boldsymbol{Y} \in DD(\boldsymbol{V}). \end{aligned}$$

which gives an upper bound on (3): $\overline{Q}_t \geq Q_t$.

[5] Ahmadi, A. A., & Hall, G. (2015). Sum of squares basis pursuit with linear and second order cone programming. Algebraic and Geometric Methods in Discrete Mathematics, *Contemp. Math*, 685, 27-53.

Facility Location with Endogenous, Stochastic Demand

Parameters:

- Stages $1, \ldots, T$, facility sites $1, \ldots, I$, demand locations $1, \ldots, J$
- c_{tij}: unit transportation cost, R_j: unit revenue, h_{ti}: capacity of facility i, f_{ti}: building cost of facility i, N: total budget for building facilities.
- ξ_{tj}: demand of location j in stage t.

Decision Variables:

- $x_{ti} \in \{0, 1\}$: $x_{ti} = 1$ if a facility is open at location i in stage t, and 0 o.w.
- $y_{tij} \in \mathbb{N}$: flow from facility i to customer j in stage t.

$$\min_{x_t, y_t} \quad \sum_{i \in [I]} \sum_{j \in [J]} c_{tij} y_{tij} - \sum_{j \in [J]} R_j \sum_{i \in [I]} y_{tij} \tag{4a}$$

$$\sum_{i \in [I]} y_{tij} \leq \xi_{tj}, \quad \forall j \in [J], \tag{4b}$$

$$\sum_{j \in [J]} y_{tij} \leq h_{ti} \sum_{\tau=1}^{t} x_{\tau i}, \quad \forall i \in [I], \tag{4c}$$

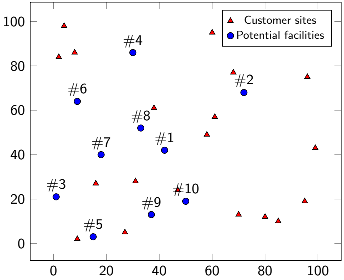
$$\sum_{i \in [I]} f_{ti}(x_{ti} - x_{t-1,i}) \leq N, \tag{4d}$$

$$x_{ti} \geq x_{t-1,i}, \quad \forall i \in [I], \tag{4e}$$

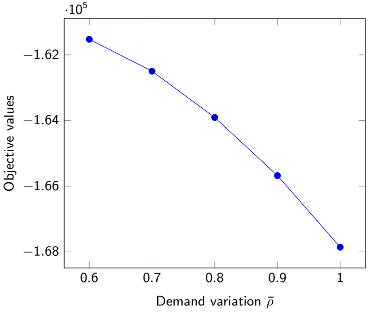
$$x_{ti} \in \{0, 1\}, \ y_{tij} \in \mathbb{Z}_+, \quad \forall i \in [I], \ j \in [J]. \tag{4f}$$

Experimental Setup for Type 1 Ambiguity Set

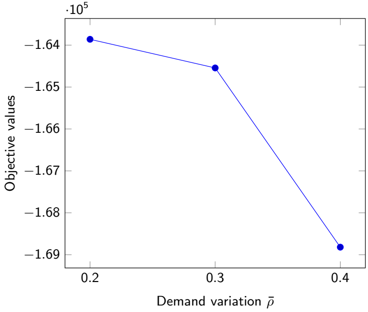
- $T = 3$ stages, $I = 10$ facilities, $J = 20$ customer sites on a 100×100 grid.
- $c_{ij} = \mathrm{dist}(i,j)/4$, $f_{ti} = 100$, $N = 100$, $h_{ti} = 1000$, $R_j = 100$.
- $\epsilon^{\mu} = 25$, $\underline{\epsilon}^S = 0.1$, $\bar{\epsilon}^S = 1.9$.
- $\lambda^{\mu}_{ji} = e^{-\mathrm{dist}(i,j)/25}$, $\lambda^{S}_{ji} = e^{-\mathrm{dist}(i,j)/50}$, and then are normalized.
- Sample K points from $\mathcal{N}(\bar{\mu}_j, \bar{\sigma}_j^2)$ where $\bar{\mu}_j \sim U[20, 40]$ and $\bar{\sigma}_j = \bar{\mu}_j \times \bar{\rho}$.



Optimal Objective Values under Different $\bar{\rho}$ and Demand Distributions



(c) Normal distribution of DDDR



(d) Log-normal distribution of DDDR

- obj(DDDR) $\downarrow$ as demand variation $\bar{\rho} \uparrow$.

Varying Budgets and Fixed Transportation Costs

Budget N	DDDR Obj.	DDDR Sol.	DIDR Obj.	DIDR Sol.
100	$-162,752$	2	$-145,811$	9
300	$-165,461$	[2, 4, 8]	$-145,815$	[7, 8, 9]
500	$-168,114$	[1, 2, 4, 8, 10]	$-145,820$	[4, 5, 7, 8, 10]

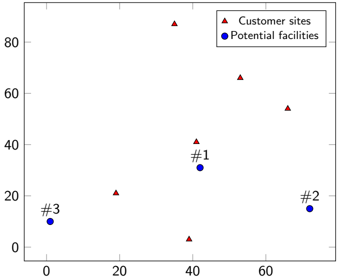
Table: Total impact on the first and second moments of each facility

I	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
$\sum_{j\in[J]}\lambda_{ji}^{\mu}$	2.11	3.47	1.08	2.15	1.43	1.69	1.47	1.86	2.09	2.65
$\sum_{j\in[J]}\lambda_{ji}^{S}$	2.23	2.58	1.48	2.02	1.57	1.80	1.85	2.11	2.06	2.31

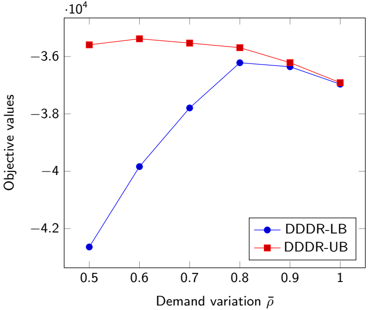
- Facility #2 has the largest total impact on the first and second moments, which is always chosen by DDDR.
- DIDR always chooses facilities having smaller impacts on the first and second moments, and thus obj(DIDR) $\geq$ obj(DDDR).

Experimental Setup for Type 3 Ambiguity Set

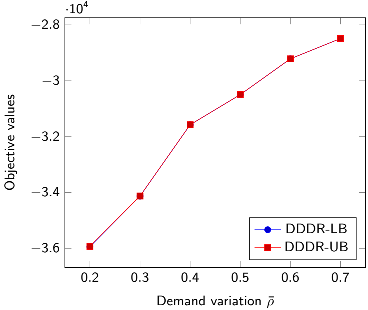
- $T = 3$ stages, $I = 3$ facilities and $J = 6$ customer sites
- $\gamma = 10, \ \eta = 100$
- λ_{ji}^{μ} are the same as before, $\lambda_i^{cov} \sim U[0, 1]$ and then are normalized.
- Sample K data points following $\mathcal{N}(\bar{\mu}_j, \bar{\sigma}_j^2)$ for each $j \in [J]$ and $\bar{\Sigma}$ is set to the sample covariance matrix of the $K \times J$ data points among all customer sites.



Diverse $\bar{\rho}$ and Demand Distributions



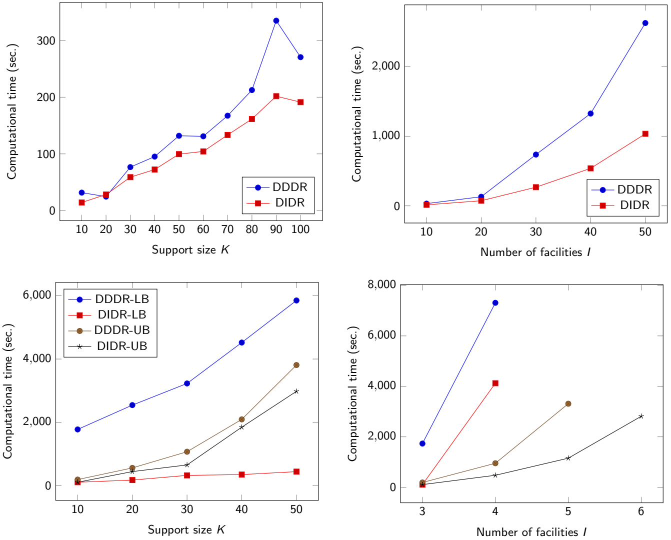
(e) Normal distributions of DDDR



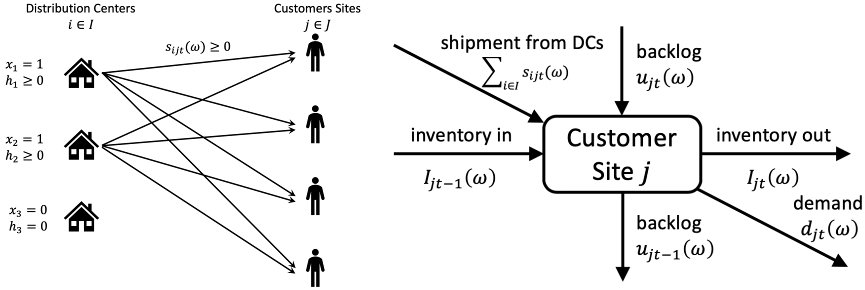
(f) Log-normal distributions of DDDR

- Gaps $\downarrow$ as demand variation $\bar{\rho} \uparrow$, and gap(Log-normal) $\ll$ gap(Normal).

CPU Time Comparison



COVID-19 Test Kit Allocation[6]



We consider a two-stage stochastic mixed-integer linear program (TS-MILP):

- The first-stage involves binary facility-location decisions x_i and capacity-design decisions h_i;
- The second-stage involves continuous decision variables s_{ijt}, u_{jt}, I_{jt} representing shipments from DCs, backlog, and inventory at each customer site depending on the random demand $d_{jt}(\omega)$.

[6] Basciftci, B., Yu, X., & Shen, S., "An optimization approach for location and capacity design of COVID-19 test kit distribution centers and allocation under uncertain outbreaks," working paper, 2020.

TS-MILP for Facility Location and Test Kit Distribution

$$\begin{aligned}
\min \quad & \sum_{i\in\mathcal{I}} c_i^o x_i + \sum_{i\in\mathcal{I}} c_i^h |\mathcal{T}| h_i \\
& + \sum_{\omega\in\Omega} p^\omega \left(\sum_{i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T}} c_{ijt}^s s_{ijt}(\omega) + \sum_{j\in\mathcal{J}, t\in\mathcal{T}} \left(c_{jt}^l l_{jt}(\omega) + c_{jt}^u u_{jt}(\omega) \right) \right) \\
\text{s.t.} \quad & \sum_{j\in\mathcal{J}} s_{ijt}(\omega) \le h_i x_i, \ \forall i \in \mathcal{I}, \ t\in\mathcal{T}, \ \omega\in\Omega \\
& \sum_{i\in\mathcal{I}, j\in\mathcal{J}} s_{ijt}(\omega) \le B_t, \ \forall t\in\mathcal{T}, \ \omega\in\Omega \\
& \sum_{i\in\mathcal{I}} s_{ijt}(\omega) + l_{jt-1}(\omega) + u_{jt}(\omega) = d_{jt}(\omega) + l_{jt}(\omega) + u_{jt-1}(\omega), \ \forall j\in\mathcal{J}, \ t\in\mathcal{T}, \ \omega\in\Omega \\
& x_i \in \{0,1\}, \ h_i, \ s_{ijt}(\omega), \ l_{j,t}(\omega), \ u_{jt}(\omega) \ge 0, \ \forall i\in\mathcal{I}, \ j\in\mathcal{J}, \ t\in\mathcal{T}, \ \omega\in\Omega,
\end{aligned}$$

- $h_i x_i$: McCormick envelopes

DDDR Formulation

$$\min \quad \sum_{i\in\mathcal{I}} c_i^o x_i + \sum_{i\in\mathcal{I}} c_i^h |\mathcal{T}| h_i + \max_{\mathbb{P}\in U(h)} \mathbb{E}_{\mathbb{P}}[g(h,d)] \tag{5a}$$

$$\text{s.t.} \quad h_i = \sum_{l\in\mathcal{L}} H_i^l \lambda_i^l, \ \forall i \in \mathcal{I}, \tag{5b}$$

$$\sum_{l\in\mathcal{L}} \lambda_i^l = x_i, \ \forall i \in \mathcal{I}, \tag{5c}$$

$$x_i \in \{0,1\} \ \forall i \in \mathcal{I}, \ \lambda_i^l \in \{0,1\} \ \forall i \in \mathcal{I}, \ \forall l \in \mathcal{L}. \tag{5d}$$

where

$$g(h,d) = \min \quad \sum_{i\in\mathcal{I}, j\in\mathcal{J}, t\in\mathcal{T}} c_{ijt}^s s_{ijt} + \sum_{j\in\mathcal{J}, t\in\mathcal{T}} \left(c_{jt}^l l_{jt} + c_{jt}^u u_{jt} \right) \tag{6a}$$

$$\text{s.t.} \quad \sum_{j\in\mathcal{J}} s_{ijt} \le h_i, \ \forall i \in \mathcal{I}, \ t \in \mathcal{T}, \tag{6b}$$

$$\sum_{i\in\mathcal{I}, j\in\mathcal{J}} s_{ijt} \le B_t, \ \forall t \in \mathcal{T}, \tag{6c}$$

$$\sum_{i\in\mathcal{I}} s_{ijt} + l_{jt-1} + u_{jt} = d_{jt} + l_{jt} + u_{jt-1}, \ \forall j \in \mathcal{J}, \ t \in \mathcal{T}, \tag{6d}$$

$$s_{ijt}, \ l_{j,t}, \ u_{jt} \ge 0, \ \forall i \in \mathcal{I}, \ j \in \mathcal{J}, \ t \in \mathcal{T}. \tag{6e}$$

Type 2 Ambiguity Set

We assume support information of each demand variable is known, and represented in set S. We introduce the ambiguity set $U(h)$ in the following form:

$$U(h) = \left\{ \mathbb{P} \in \mathcal{P}(S) \quad \middle| \quad \int_S d\mathbb{P} = 1, \quad \mathbb{E}_{\mathbb{P}}[\mathrm{d}] = \mu(h), \quad \mathbb{E}_{\mathbb{P}}[(\mathrm{d} - \mu(h))(\mathrm{d} - \mu(h))^\top] = \Sigma(h) \right\},$$

Theorem (Basciftci, Yu, S. (2020))

Problem (5) *can be represented as the following single level problem:*

$$\begin{aligned} \min \quad & \sum_{i \in \mathcal{I}} c_i^o x_i + \sum_{i \in \mathcal{I}} c_i^h |\mathcal{T}| h_i + \beta + \delta^\top \mu(h) + \Sigma(h) \cdot Y \\ s.t. \quad & (5\mathrm{b}) - (5\mathrm{d}) \\ & \beta + \delta^\top \xi^k + (\xi^k - \mu(h))(\xi^k - \mu(h))^\top \cdot Y \geq g(h, \xi^k), \forall k \in S. \end{aligned}$$

Future Work

- Compare the solution performance and objective values between TS-MILP, DRO and DDDR with Type 2 ambiguity set in COVID-19 test kit allocation problem;
- Consider the effect of decision dependency in DRO models with distance-based ambiguity sets, such as Wasserstein metric.

- Yu, X., Shen, S., "Multistage Distributionally Robust Mixed-Integer Programming with Decision-Dependent Moment-based Ambiguity Sets," forthcoming in *Mathematical Programming*. `https://doi.org/10.1007/s10107-020-01580-4`.
- Beste Basciftci, Shabbir Ahmed, Siqian Shen. "Distributionally robust facility location problem under decision-dependent stochastic demand," *European Journal of Operational Research*, 292(2), 548-561, 2021.
- Basciftci, B., Yu, X., Shen, S., "Resource Distribution Under Spatiotemporal Uncertainty of Disease Spread: Stochastic versus Robust Approaches," `https://arxiv.org/abs/2103.04266`.