Robust Decision Models with Estimation and Prediction

Melvyn SIM
National University of Singapore

Stochastic Programming Society (SPS)
Virtual Seminar Series

Robust Decision Models with Estimation and Prediction

Melvyn SIM
National University of Singapore

Joint work with




Zhiyuan Lou
TUM

Zhi Chen
CUHK

Jingui Xie
TUM

Peng Xiong
NUS

Agenda

- Challenges of Optimizing Decisions Under Uncertainty
- A Tour of Robust Decision Models
- Estimating Parameters of Disparity Metric
- Asymmetric Uncertainty Sets
- Predictive Model with Side Information
- Numerical Studies

Challenges of Optimizing Decisions Under Uncertainty

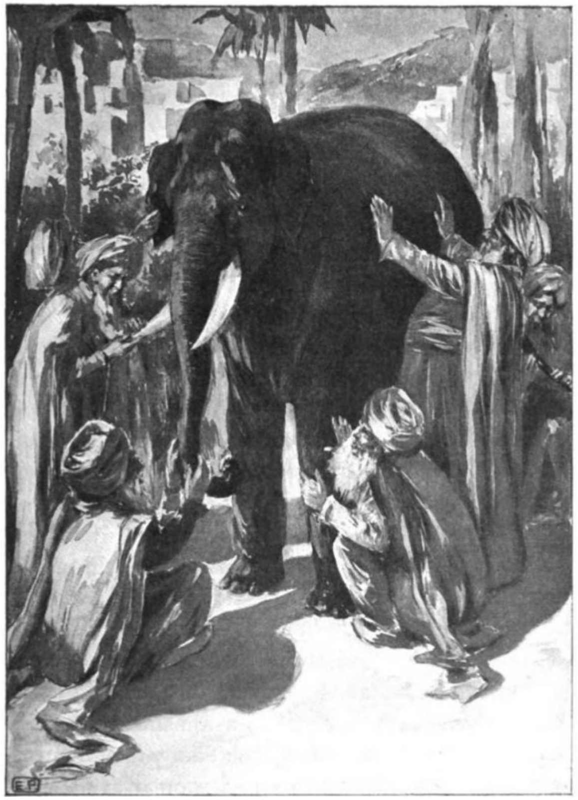
"All Models are Wrong"

– George Box, JASA 1976

Expertise on Decisions Making under Uncertainty

- **Applied Probabilist**
  - Model decision problem as a stochastic process.
  - Focus on formulating stochastic process with analytical tractability.
- **Decision Theorist**
  - Rank various decisions under uncertainty.
  - Focus on rational agents with extensive computational powers.
- **Statistician**
  - Estimate parameters of stochastic process from data.
  - Focus on convergence analysis as data increases.
- **Optimization Expert**
  - Obtain the best solution for the decision problem
  - Focus on efficiency of algorithm and optimality of the solution.

The Blind Men and the Elephant



"The moral of the parable is that humans have a tendency to claim absolute truth based on their limited, subjective experience as they ignore other people's limited, subjective experiences which may be equally true." - Wikipedia

- Illustrator unknown - From *The Heath readers by grades*, D.C. Heath and Company (Boston).

A Portfolio Management Problem

- Investment in N stocks.
- Uncertain future stock returns $\tilde{\boldsymbol{r}}$.
- Portfolio allocation $\boldsymbol{x} \in \mathbb{R}^N$.
- Future portfolio returns: $\tilde{\boldsymbol{r}}^\top \boldsymbol{x}$.

A Portfolio Management Problem

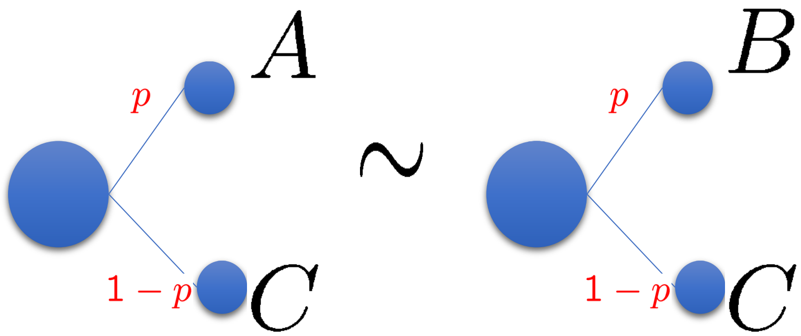
- **Applied Probablist**
  - Model stock returns as a multivariate normal distributed random variable.

$$\tilde{\boldsymbol{r}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
$$\tilde{\boldsymbol{r}}^\top \boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}^\top \boldsymbol{x}, \boldsymbol{x}^\top \boldsymbol{\Sigma} \boldsymbol{x})$$

A Portfolio Management Problem

- **Decision Theorist**
  - Rational agents follows Expected Utility Theory (EUT)
  - Consistent with the following Axioms (von Neumann/Morgenstern)
    - **Complete Ordering**
    - **Transitivity**
    - **Independence**
    - **Continuity**

$$A \sim B$$



A Portfolio Management Problem

- **Decision Theorist**
  - Rank portfolio decisions based on an agent with Constant Absolute Risk Aversion- Exponential Utility.

$$u_\kappa^{-1}\left(\mathbb{E}\left[u_\kappa\left(\tilde{\boldsymbol{r}}^\top \boldsymbol{x}\right)\right]\right) = \boldsymbol{\mu}^\top \boldsymbol{x} - \frac{\boldsymbol{x}^\top \boldsymbol{\Sigma} \boldsymbol{x}}{2\kappa}$$

where

$$u_\kappa(r) = -\exp(-r/\kappa)$$

A Portfolio Management Problem

- **Statistician**

Given empirical distribution:

$$\tilde{\boldsymbol{r}} \sim \widehat{\mathbb{P}}$$

Estimate mean and covariance by maximizing likelihood:

$$\hat{\boldsymbol{\mu}} = \mathbb{E}_{\widehat{\mathbb{P}}}\left[\tilde{\boldsymbol{r}}\right], \hat{\boldsymbol{\Sigma}} = \mathbb{E}_{\widehat{\mathbb{P}}}\left[(\tilde{\boldsymbol{r}} - \hat{\boldsymbol{\mu}})(\tilde{\boldsymbol{r}} - \hat{\boldsymbol{\mu}})^\top\right]$$

A Portfolio Management Problem

- **Optimization Expert**
  - Obtain optimal portfolio decision by solving a tractable convex quadratic optimization problem.

$$\begin{aligned} \max \quad & \hat{\boldsymbol{\mu}}^\top \boldsymbol{x} - \frac{\boldsymbol{x}^\top \widehat{\boldsymbol{\Sigma}} \boldsymbol{x}}{2\kappa} \\ \text{s.t.} \quad & \mathbf{1}^\top \boldsymbol{x} = 1 \end{aligned}$$

A Portfolio Management Problem

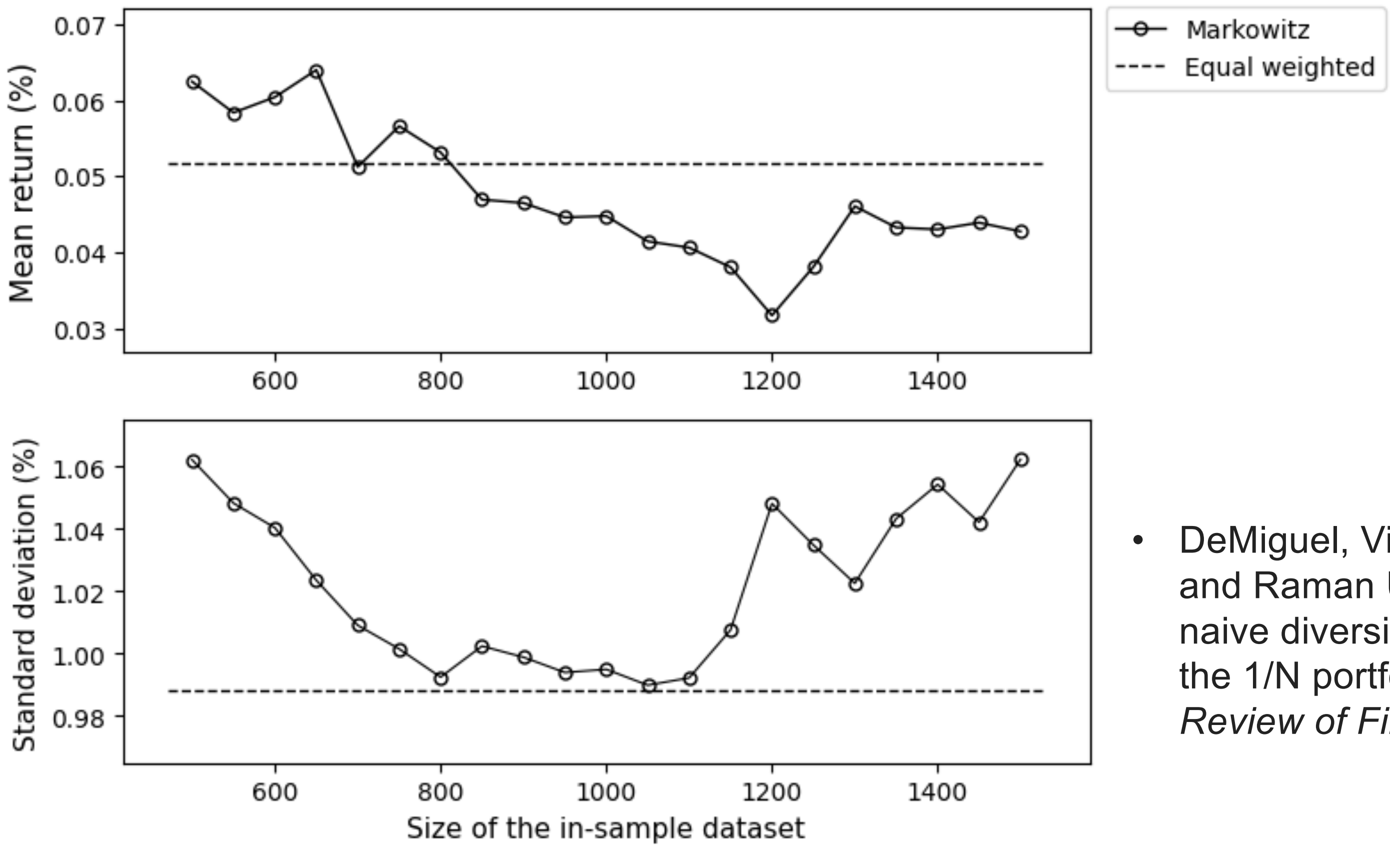
- **Optimization Expert**
  - Constraint on expected return:

$$\begin{aligned} \text{min} \quad & \boldsymbol{x}^\top \widehat{\boldsymbol{\Sigma}} \boldsymbol{x} \\ \text{s.t.} \quad & \widehat{\boldsymbol{\mu}}^\top \boldsymbol{x} = \tau \\ & \mathbf{1}^\top \boldsymbol{x} = 1 \end{aligned}$$

Expected return of equal weighted portfolio:

$$\tau = \frac{1}{N} \widehat{\boldsymbol{\mu}}^\top \mathbf{1}$$

Optimizer’s Curse: Inferior Performance



- DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?." *The Review of Financial Studies* (2009).

Agenda

- Challenges of Optimizing Decisions Under Uncertainty
- A Tour of Robust Decision Models
- Estimating Parameters of Disparity Metric
- Asymmetric Uncertainty Sets
- Predictive Model with Side Information
- Numerical Studies

A Deterministic Decision Model

$$V = \begin{cases} \min & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & f(\boldsymbol{x}, \hat{\boldsymbol{z}}) \leq \tau \\ & \boldsymbol{x} \in \mathcal{X} \end{cases}$$

- f: attribute function.
- $\boldsymbol{x} \in \mathcal{X}$: here-and-now decisions.
- $\hat{\boldsymbol{z}} \in \mathcal{Z}$: Specified input vector, where $\mathcal{Z}$ is the support set.

On Attribute Function

- Often expressed as an optimization problem over wait-and-see or recourse variables.

$$f(\boldsymbol{x}, \boldsymbol{z}) = \begin{cases} \text{min} & \boldsymbol{d}^\top \boldsymbol{y} \\ \text{s.t.} & \mathbf{B}\boldsymbol{y} \geq \mathbf{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{b}(\boldsymbol{z}) \\ & \boldsymbol{y} \in \mathbb{R}^{N_1} \end{cases}$$

- A decision model can have multiple attribute functions

$$V = \begin{cases} \text{min} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & f_m(\boldsymbol{x}, \hat{\boldsymbol{z}}) \leq \tau_m \quad \forall m \in [M] \\ & \boldsymbol{x} \in \mathcal{X} \end{cases}$$

Budgeted Robust Optimization Model

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- $\mathcal{Z}(\gamma)$: uncertainty set: subset of support set that contains the specified input vector

$$\mathcal{Z}(\gamma) = \{\boldsymbol{z} \in \mathcal{Z} \mid \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \leq \gamma\}.$$

- $\Delta(\mathrm{z}, \hat{\mathrm{z}})$:Disparity metric
- γ: Disparity budget parameter

On Disparity Metric

- Evaluate distance from specified input.
- Typically based in L-p norm:

$$\Delta(z, \hat{z}) = \|\mathbf{D}^{-1}(z - \hat{z})\|_p.$$

- $\mathbf{D}$: Deviation Matrix: Determine the shape of the uncertainty set
- The choice of L-p norm has the impact on its tractability.
  - p=1 is associated with the uncertainty set of “The Price of Robustness”.

Budgeted Robust Optimization Model

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- Exact or safe tractable approximations available.
  - Many works on tractable reformulations
  - Facilitated by algebraic modeling language such as RSOME.

RSOME – Robust Stochastic Optimization Made Easy

RSOME

Star 291 | Watch 4 | Fork 56

- Home
- User Guide
- Examples
- About

Download ZIP File | Download TAR Ball | View On GitHub

The current version of RSOME supports deterministic, robust optimization and distributionally robust optimization problems. In the default configuration, RSOME relies on open-source solvers imported from the `scipy.optimize` package to solve linear programming (`linprog()`) and mixed-integer linear programming (`milp`) problems. Besides the default solver, RSOME also provides interfaces for other open-source and commercial solvers. Detailed information of these solver interfaces is presented in the following table.

Solver	License type	Required version	RSOME interface	Second-order cone constraints	Exponential cone constraints	Semidefiniteness constraints
scipy.optimize	Open-source	>= 1.9.0	`lpg_solver`	No	No	No
CyLP	Open-source	>= 0.9.0	`clp_solver`	No	No	No
OR-Tools	Open-source	>= 7.5.7466	`ort_solver`	No	No	No
ECOS	Open-source	>= 2.0.10	`eco_solver`	Yes	Yes	No
Gurobi	Commercial	>= 9.1.0	`grb_solver`	Yes	No	No
Mosek	Commercial	>= 10.0.44	`msk_solver`	Yes	Yes	Yes
CPLEX	Commercial	>= 12.9.0.0	`cpx_solver`	Yes	No	No
COPT	Commercial	>= 7.2.2	`cpt_solver`	Yes	Yes	Yes

Budgeted Robust Optimization Model

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \quad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- Question:
  - How do we specify the disparity budget parameter?
    - Statistical Guarantee?
    - Interpretability?

Relation to Chance Constraint

Suppose $\tilde{z} \sim \mathbb{P}$, then

$$f(x, z) \leq \tau \quad \forall z \in \mathcal{Z}(\gamma)$$
$$\Downarrow$$
$$\mathbb{P}[f(x, \tilde{z}) > \tau] \leq \mathbb{P}[\Delta(\tilde{z}, \hat{z}) > \gamma]$$

- Robust constraint as a tractable approximation of chance constraints
- Caveat:
  - Bound is weak and depends on attribute function
  - Probability distribution is often not known.

Satisficing Decision Model

- Simon (1955, 1959)
  - Human behavior to achieve at least some minimum level of a particular variable, but which does not necessarily maximize its value. Coined the term **satisfice**. (Satisfy + Suffice).
- Lanzilloti (1958)
  - Managers are primarily concerned about target returns on investment.
- Mao (1970)
  - Manager's perception of risk: the prospect of not meeting some target rate of return.

Satisficing Decision Model

P-Model:

$$\begin{aligned} \max \quad & \mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) \leq \tau\right] \\ \text{s.t.} \quad & \boldsymbol{c}^\top \boldsymbol{x} \leq V_1 \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- Charnes and Cooper. Operations Research. 1963. Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints.
- Computationally intractable

Tractable Satisficing Models

- D. Brown and M. Sim. 2009. Satisficing Measures for Analysis of Risky Positions. *Management Science*.
- D. Brown, E. De Giorgi and M. Sim. 2012. Aspirational Preferences and their Representation by Risk Measures. *Management Science*.
- D.Z. Long, M. Sim and M. Zhou 2022. Robust Satisficing. *Operations Research*.
- M Sim, Q Tang, M Zhou, T Zhu. 2024. The Analytics of Robust Satisficing - Predict, Optimize, Satisfice, then Fortify. *Operations Research*.

Budgeted Robust Satisficing Model

$$\begin{aligned} \max \quad & \gamma \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \le \tau \qquad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \\ & \boldsymbol{c}^\top \boldsymbol{x} \le V_1 \\ & \boldsymbol{x} \in \mathcal{X}, \ \ \gamma \ge 0 \end{aligned}$$

- $V_1 \ge V$: the specified target.
  - Target is less abstract than budget parameter
- Maximize the budget parameter, subject to budgeted robust feasibility
  - Quasiconcave maximation problem
  - Binary search on budget parameter

Tolerated Robust Optimization Model

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \le \tau + \kappa \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z} \in \mathcal{Z} \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- $\kappa \in [0, +\infty]$: disparity tolerance parameter.
- Violation of attribute constraint can be tolerated
- Magnitude of violation is proportional to the distance from the specified input vector

Relation to Profiled Chance Constraint

Suppose $\tilde{\boldsymbol{z}} \sim \mathbb{P}$, then

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z} \in \mathcal{Z}$$

$$\Downarrow$$

$$\mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau + \kappa r\right] \leq \mathbb{P}\left[\Delta(\tilde{\boldsymbol{z}}, \hat{\boldsymbol{z}}) > r\right] \quad \forall r > 0$$

- Ensure that greater violation of constraint might occur at lower probability.
- May not provide probabilistic guarantee of constraint feasibility.

Tolerated Robust Satisficing Model

$$
\begin{aligned}
\min \quad & \kappa \\
\text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \le \tau + \kappa \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z} \in \mathcal{Z} \\
& \boldsymbol{c}^\top \boldsymbol{x} \le V_1 \\
& \boldsymbol{x} \in \mathcal{X}, \ \ \kappa \ge 0
\end{aligned}
$$

- $V_1 \ge V$: the specified target.
  - Target is less abstract than tolerance parameter
- Minimize the tolerance parameter, subject to tolerated robust feasibility
  - Limit violation of constraint as much as possible
  - Convex optimization problem

Unified Robust Optimization Framework

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) - \gamma)^+, \quad \forall \boldsymbol{z} \in \mathcal{Z} \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- Dual parameter specifications
  - $\gamma \in [0, +\infty]$: disparity budget.
  - $\kappa \in [0, +\infty]$: disparity tolerance.
- Recovers budgeted robust optimization model if $\kappa = \infty$.
- Recovers tolerated robust optimization model if $\gamma = 0$.

Probabilistic Guarantee

Suppose $\tilde{z} \sim \mathbb{P}$, then

$$f(x, z) \leq \tau + \kappa(\Delta(z, \hat{z}) - \gamma)^{+} \quad \forall z \in \mathcal{Z}$$

$$\Downarrow$$

$$\mathbb{P}\left[f(x, \tilde{z}) > \tau + \kappa r\right] \leq \mathbb{P}\left[\Delta(\tilde{z}, \hat{z}) > \gamma + r\right] \quad \forall r > 0$$

- Ensure that greater violation of constraint might occur at lower probability.
- Also provide probabilistic guarantee of constraint feasibility.

Unified Robust Satisficing Framework

- Specify two targets: $V_2 \geq V_1 \geq V$
- Step 1: Maximize budget as much as possible:

$$\gamma^\star = \begin{cases} \max & \gamma \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau \qquad \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \\ & \boldsymbol{c}^\top \boldsymbol{x} \leq V_1 \\ & \boldsymbol{x} \in \mathcal{X}, \ \gamma \geq 0 \end{cases}$$

- Step 2: Minimize tolerance as much as possible:

$$\begin{array}{ll} \min & \kappa \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) - \gamma^\star)^+, \ \ \forall \boldsymbol{z} \in \mathcal{Z} \\ & \boldsymbol{c}^\top \boldsymbol{x} \leq V_2 \\ & \boldsymbol{x} \in \mathcal{X}, \kappa \geq 0 \end{array}$$

Agenda

- Challenges of Optimizing Decisions Under Uncertainty
- A Tour of Robust Decision Models
- Estimating Parameters of Disparity Metric
- Asymmetric Uncertainty Sets
- Predictive Model with Side Information
- Numerical Studies

Estimating Parameters of Disparity Metric

- Robust models are characterized by the disparity metric:

$$\begin{aligned} \text{min} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & f(\boldsymbol{x}, \boldsymbol{z}) \le \tau + \kappa(\Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) - \gamma)^+, \quad \forall \boldsymbol{z} \in \mathcal{Z} \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- Suppose we restrict disparity to L-p norm:

$$\Delta(z, \hat{z}) = \|\mathbf{D}^{-1}(z - \hat{z})\|_p.$$

- Can we estimate the parameters of the disparity metric from data?
  - What would be the objective function of such estimation procedure?

Probabilistic Guarantee

Suppose $\boldsymbol{x}$ is feasible in:

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\|\mathbf{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}})\|_p - \gamma)^+, \quad \forall \boldsymbol{z} \in \mathcal{Z}$$

and

$$\mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_p)\right] \leq 1.$$

ψ: nonnegative, convex and increasing deviation penalty function.

By Markov inequality,

$$\begin{aligned} \mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau + \kappa r\right] &\leq \mathbb{P}\left[\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_p > r)\right] \\ &\leq \mathbb{P}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_p) > \psi(r))\right] \\ &\leq 1/\psi(r) \quad \forall r \geq 0. \end{aligned}$$

Probabilistic Guarantee

By specifying the deviation penalty function ψ and ensuring that

$$\mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_p)\right] \leq 1,$$

we have:

$$\mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau + \kappa r\right] \leq 1/\psi(r) \quad \forall r \geq 0.$$

Suppose $\psi(r) = r^7$, then

$$\mathbb{P}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{z}}) > \tau + \kappa r\right] \leq r^{-7} \quad \forall r \geq 0.$$

Volume of Confidence Set

Define confidence set:

$$\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma) = \left\{ \boldsymbol{z} \in \mathbb{R}^L \mid \|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|_p \leq \gamma \right\}.$$

Volume of confidence set:

$$\mathsf{Vol}(\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \gamma)) = \det(\mathbf{D})\gamma^L \mathsf{Vol}(\mathcal{B}(\mathbf{0}, \mathbf{I}, 1)).$$

- The volume of the norm ball is related to the size of the uncertainty set.
- The larger the size of the uncertainty set, the more conservative the robust solution might become.

Minimum Volume Confidence Set (MVCS)

$$\begin{aligned} \inf \quad & \det(\mathbf{D}) \\ \text{s.t.} \quad & \mathbb{E}_{\mathbb{P}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \leq 1 \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \mathbf{D} \in \mathbb{S}^L_{++}. \end{aligned}$$

Ensures probabilistic guarantee.

Volume of confidence set.

- Key idea in estimating the parameters of the deviation metric:
  - Minimizing the volume of the confidence set, while providing the probabilistic guarantee.
  - Estimating based on providing the least conservative solution with the desired probabilistic guarantee.

MVCS Estimation Problem

Consider the data-driven setting. We have a set of historical data

$$\mathcal{D} = \{(\omega, z_\omega) \mid \omega \in [\Omega]\},$$

where z_ω is the observed outcome of the input vector in scenario ω.

MVCS Estimation Problem

$$\begin{aligned} \inf \quad & \det(\mathbf{D}) \\ \text{s.t.} \quad & \mathbb{E}_{\widehat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \leq 1 \\ & \mathbf{D} \succeq \text{Diag}(\boldsymbol{r}) \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L, \; \mathbf{D} \in \mathbb{S}^L_{++}. \end{aligned}$$

$$\widehat{\mathbb{P}}[\tilde{\boldsymbol{z}} = \boldsymbol{z}_\omega] = \frac{1}{\Omega} \quad \forall \omega \in \Omega$$

- Regularization vector $\boldsymbol{r}$ to avoid degeneration.
- Without it, $\mathbf{D}$ may degenerate to a singular matrix, when there is enough data, *i.e.* Ω is small,

MVCS Estimation Problem

$$
\begin{aligned}
\inf \quad & \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\|)\right] \leq 1 \\
& \mathbf{D} \succeq \text{Diag}(\boldsymbol{r}) \\
& \hat{\boldsymbol{z}} \in \mathbb{R}^L, \ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{aligned}
$$

Not a Convex Optimization Problem

$\Updownarrow$

$$
\begin{aligned}
\max \quad & \sqrt[L]{\det(\mathbf{Q})} \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q}\|)\right] \leq 1 \\
& \mathbf{Q} \preceq \text{Diag}^{-1}(\boldsymbol{r}) \\
& \boldsymbol{q} \in \mathbb{R}^L, \ \mathbf{Q} \in \mathbb{S}^L_{+}.
\end{aligned}
$$

Convert to Convex Optimization via change of variables:

$$\mathbf{D} \rightarrow \mathbf{Q}^{-1}, \quad \hat{\boldsymbol{z}} \rightarrow \mathbf{D}\boldsymbol{q}$$

- Max root-determinant optimization problem with semidefinite constraints
- Root-determinant is concave and SOCP representable - Supported in RSOME
- Deviation penalty function is conic representable – Power of exponential cone

Deviation Metric Norm and Penalty Function

Deviation metric based on L-p_1 norm:

$$\|\cdot\| = \|\cdot\|_{p_1}.$$

Deviation penalty based on p_2 power function:

$$\psi(r) = r^{p_2}.$$

Estimation problem:

$$\begin{array}{ll} \inf & \det(\mathbf{D}) \\ \text{s.t.} & \mathbb{E}_{\widehat{\mathbb{P}}}\left[\left(\left\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}})\right\|_{p_1}\right)^{p_2}\right] \leq 1 \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L,\ \mathbf{D} \in \mathbb{S}^L_{++}. \end{array}$$

RSOME Code

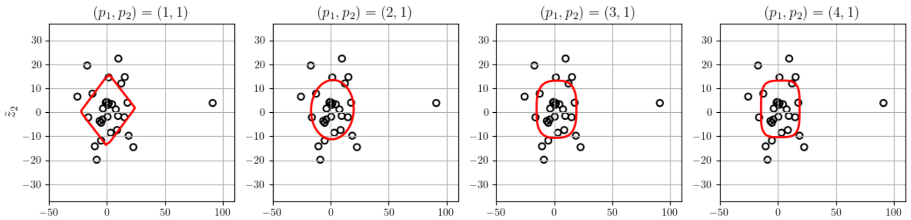
```python
import numpy as np
import rsome as rso
from rsome import ro
from rsome import cpt_solver as cpt
```

```python
N, L = z.shape

model = ro.Model()
q = model.dvar(L)
u = model.dvar(N)
v = model.dvar(N)
Q = model.dvar((L, L))

model.max(rso.rootdet(Q))
model.st((1/N) * u.sum() <= 1)
model.st((1/L) * rso.power(v, p2) <= u)
l_norm = lambda x: rso.norm(x, p1)
for n in range(N):
    model.st(l_norm(Q@z[n] - q) <= v[n])

model.solve(cpt, display=display)

return q.get(), Q.get()
```

$$
\begin{aligned}
\max \quad & \sqrt[L]{\det(\mathbf{Q})} \\
\text{s.t.} \quad & \mathbb{E}_{\widehat{\mathbb{P}}}\left[(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q}\|_{p_1})^{p_2}\right] \leq 1 \\
& \boldsymbol{q} \in \mathbb{R}^L, \ \mathbf{Q} \in \mathbb{S}^L_+.
\end{aligned}
$$

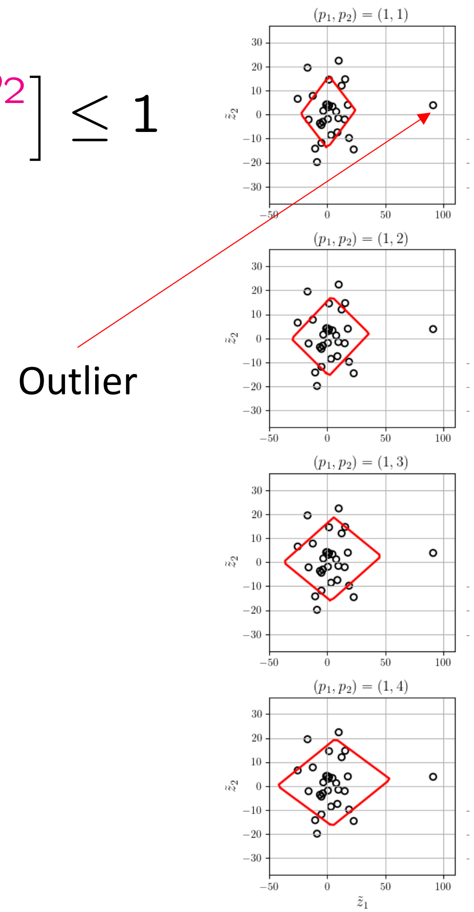
Shape of Uncertainty Set via p_1

$$
\begin{aligned}
\text{inf} \quad & \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\widehat{\mathbb{P}}}\left[\left\|\mathbf{D}^{-1}(\tilde{z} - \hat{z})\right\|_{p_1}\right] \le 1 \\
& \hat{z} \in \mathbb{R}^L,\ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{aligned}
$$



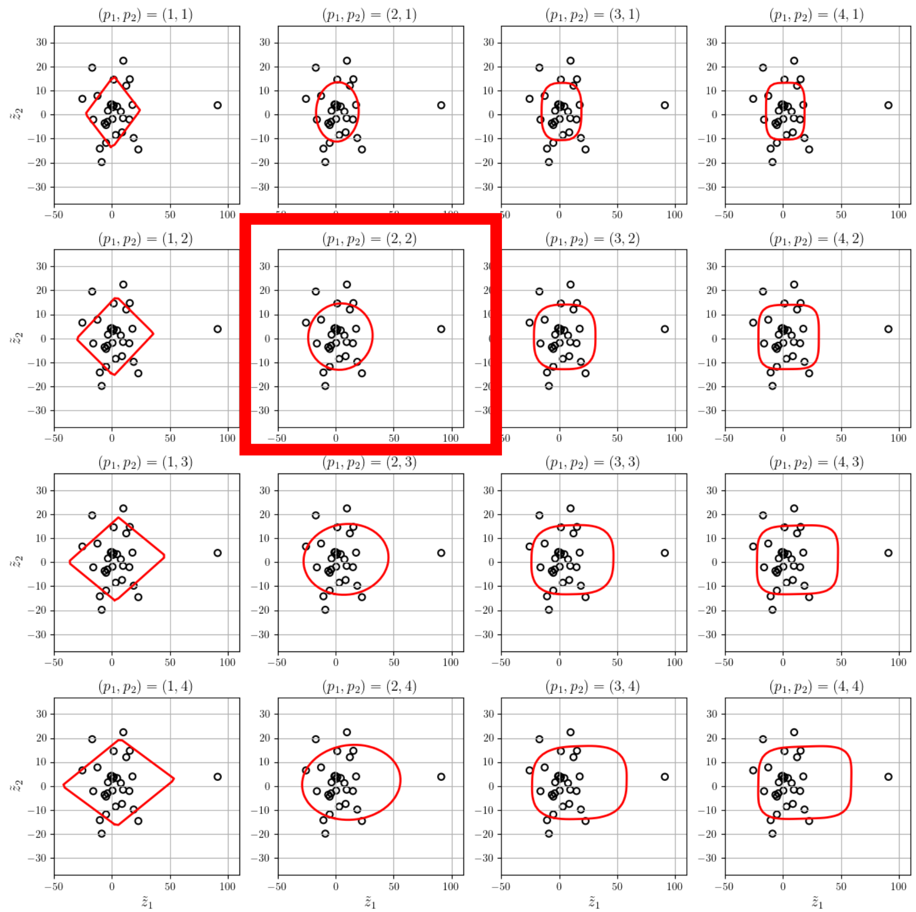
Outlier Sensitivity to Estimation via p_2

$$\begin{aligned} \inf \quad & \det(\mathbf{D}) \\ \text{s.t.} \quad & \mathbb{E}_{\widehat{\mathbb{P}}}\left[\left(\left\|\mathbf{D}^{-1}(\tilde{z} - \hat{z})\right\|_1\right)^{p_2}\right] \leq 1 \\ & \hat{z} \in \mathbb{R}^L,\ \mathbf{D} \in \mathbb{S}^L_{++}. \end{aligned}$$



Uncertainty Set and Outlier Sensitivity

$$
\begin{aligned}
\inf \quad & \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\widehat{\mathbb{P}}}\left[\left(\left\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \widehat{\boldsymbol{z}})\right\|_{p_1}\right)^{p_2}\right] \leq 1 \\
& \widehat{\boldsymbol{z}} \in \mathbb{R}^L, \ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{aligned}
$$



Case of $p_1 = p_2 = 2$

$$\widehat{\boldsymbol{z}}^\star = \mathbb{E}_{\widehat{\mathbb{P}}}[\tilde{\boldsymbol{z}}]$$

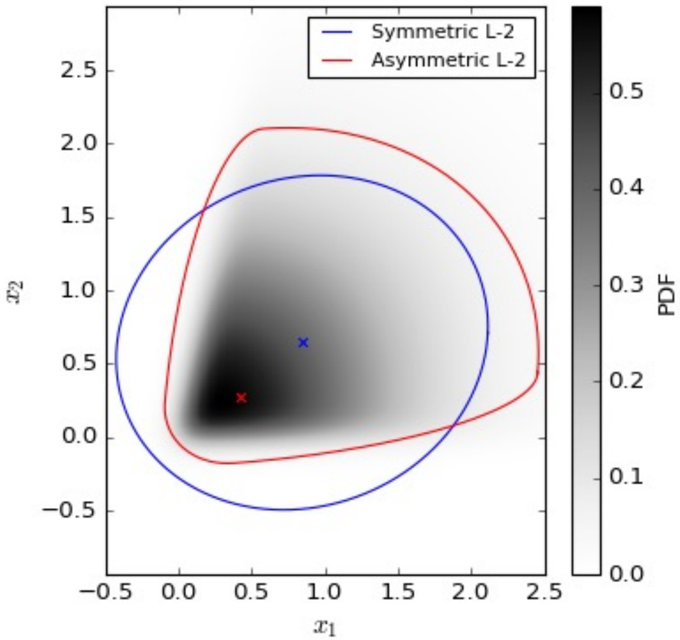
$$\mathbf{D}^\star = \mathbb{E}_{\widehat{\mathbb{P}}}\left[(\tilde{\boldsymbol{z}} - \widehat{\boldsymbol{z}}^\star)(\tilde{\boldsymbol{z}} - \widehat{\boldsymbol{z}}^\star)^\top\right]^{1/2}$$

Agenda

- Challenges of Optimizing Decisions Under Uncertainty
- A Tour of Robust Decision Models
- Estimating Parameters of Disparity Metric
- **Asymmetric Uncertainty Sets**
- Predictive Model with Side Information
- Numerical Studies

Asymmetric Confidence Set

$$\mathcal{B}(\hat{\boldsymbol{z}}, \mathrm{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}, \gamma) = \left\{ \boldsymbol{z} \in \mathbb{R}^L \mid \|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathrm{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathrm{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}))^+\| \leq \gamma \right\}.$$



How shall we estimate $\hat{\boldsymbol{z}}, \mathbf{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}$?

MCVS for Asymmetric Confidence Set

$$\mathrm{Vol}(\mathcal{B}(\hat{\boldsymbol{z}}, \mathbf{D}, \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}}, \gamma)) \propto \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D})$$

$$\begin{aligned} \inf \quad & \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \\ \text{s.t.} \quad & \mathbb{E}_{\widehat{\mathbb{P}}}\left[\psi(\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^- + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^+\|)\right] \leq 1 \\ & \mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r}) \\ & \hat{\boldsymbol{z}} \in \mathbb{R}^L, \ \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}} \in \mathbb{R}^L_{++}, \ \mathbf{D} \in \mathbb{S}^L_{++}. \end{aligned}$$

Not representable as a convex optimization problem!

MCVS for Asymmetric Confidence Set

$$
\begin{aligned}
\inf \quad & \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^{-} + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^{+}\|)\right] \leq 1 \\
& \mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r}) \\
& \hat{\boldsymbol{z}} \in \mathbb{R}^L, \ \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}} \in \mathbb{R}^L_{++}, \ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{aligned}
$$

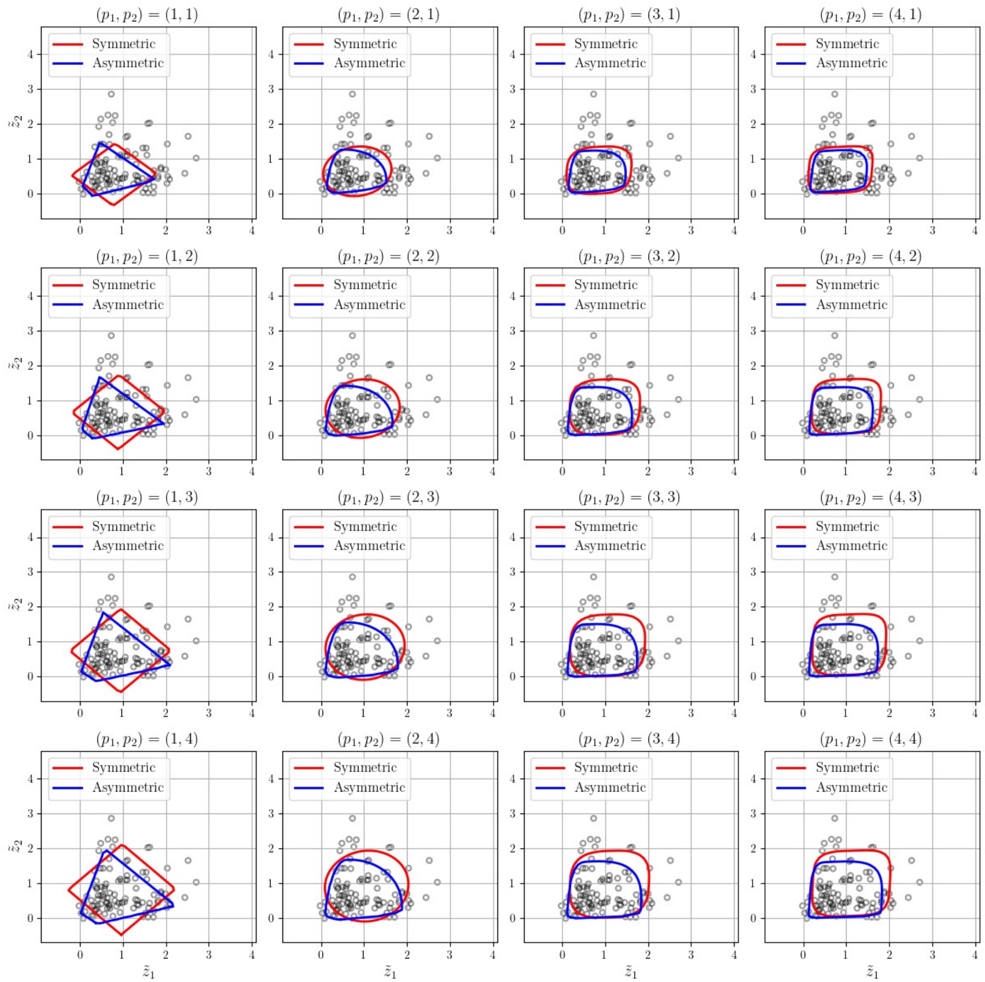
An alternating optimization strategy:

$$
\begin{aligned}
\inf \quad & \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^{-} + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^{+}\|)\right] \leq 1 \\
& \mathbf{D} \succeq \mathrm{Diag}(\boldsymbol{r}) \\
& \hat{\boldsymbol{z}} \in \mathbb{R}^L, \ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{aligned}
$$

Subproblems are conic representable!

$$
\begin{aligned}
\inf \quad & \prod_{\ell \in [L]} \frac{\underline{\sigma}_\ell + \bar{\sigma}_\ell}{2} \cdot \det(\mathbf{D}) \\
\text{s.t.} \quad & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|(\mathrm{Diag}^{-1}(\underline{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^{-} + (\mathrm{Diag}^{-1}(\bar{\boldsymbol{\sigma}})\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}))^{+}\|)\right] \leq 1 \\
& \underline{\boldsymbol{\sigma}}, \bar{\boldsymbol{\sigma}} \in \mathbb{R}^L_{++}
\end{aligned}
$$

MCVS for Asymmetric Confidence Set



Agenda

- Challenges of Optimizing Decisions Under Uncertainty
- A Tour of Robust Decision Models
- Estimating Parameters of Disparity Metric
- Asymmetric Uncertainty Sets
- **Predictive Model with Side Information**
- Numerical Studies

Predictive Model with Side Information

Consider the data-driven setting. We have a set of historical data with *side information*:

$$\mathcal{D} = \{(\omega, \boldsymbol{z}_\omega, \boldsymbol{s}_\omega)\}_{\omega \in [\Omega]}$$

where $\boldsymbol{z}_\omega \in \mathbb{R}^L$ is the observed outcome of the input vector in scenario ω, and $\boldsymbol{s}_\omega \in \mathbb{R}^L$ is the corresponding *side information*.

Predictive Model with Side Information

We focus on affine prediction map $\boldsymbol{\chi}: \mathbb{R}^S \to \mathbb{R}^L$,

$$\chi(\boldsymbol{s}) = \widehat{\boldsymbol{z}}_0 + \widehat{\mathbf{Z}}\boldsymbol{s},$$

where $\hat{\boldsymbol{z}}_0$ and $\hat{\mathbf{Z}}$ are parameters estimated from data.

Predictive Model with Side Information

The robust constraint with *side information*:

$$f(\boldsymbol{x}, \boldsymbol{z}) \leq \tau + \kappa(\Delta(\boldsymbol{z}, \hat{z}_0 + \hat{\mathbf{Z}}\boldsymbol{s}) - \gamma)^+ \quad \forall \boldsymbol{z} \in \mathcal{Z}$$

Given side information $\boldsymbol{s}$, determine decision on $\boldsymbol{x}$.
Example in unit commitment:

- $\boldsymbol{s}$: Weather information, AI or expert predictions
- $\boldsymbol{z}$: Energy demands
- $\boldsymbol{x}$: Production levels of energy

Predictive Model with Side Information

We utilize the empirical distribution:

$$\hat{\mathbb{P}}\left\{(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{s}}) = (\boldsymbol{z}_\omega, \boldsymbol{s}_\omega)\right\} = 1/\Omega \ \forall \omega \in [\Omega].$$

to estimate parameters $\mathbf{D}$, $\hat{\boldsymbol{z}}_0$ and $\hat{\mathbf{Z}}$.

MVCS with side information:

$$\begin{array}{ll} \inf & \det(\mathbf{D}) \\ \text{s.t.} & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}_0 - \hat{\mathbf{Z}}\tilde{\boldsymbol{s}})\|)\right] \leq 1 \\ & \mathbf{D} \succeq \text{Diag}(\boldsymbol{r}) \\ & \hat{\boldsymbol{z}}_0 \in \mathbb{R}^L, \ \hat{\mathbf{Z}} \in \mathbb{R}^{L \times S}, \ \mathbf{D} \in \mathbb{S}^L_{++}. \end{array} \quad \Longleftrightarrow \quad \begin{array}{ll} \sup & (\det(\mathbf{Q}))^{1/L} \\ \text{s.t.} & \mathbb{E}_{\hat{\mathbb{P}}}\left[\psi(\|\mathbf{Q}\tilde{\boldsymbol{z}} - \boldsymbol{q} - \mathbf{P}\tilde{\boldsymbol{s}}\|)\right] \leq 1 \\ & \mathbf{Q} \preceq \text{Diag}^{-1}(\boldsymbol{r}) \\ & \boldsymbol{q} \in \mathbb{R}^L, \ \mathbf{P} \in \mathbb{R}^{L \times S}, \ \mathbf{Q} \in \mathbb{S}^L_{++}. \end{array}$$

- Given the optimal solution of RHS, we can get the optimal solution of LHS:

$$\mathbf{D}^\star = \mathbf{Q}^{\star-1}, \ \hat{\boldsymbol{z}}_0^\star = \mathbf{D}^\star \boldsymbol{q}^\star, \ \text{and} \ \hat{\mathbf{Z}}^\star = \mathbf{D}^\star \mathbf{P}^\star.$$

Generalizes Ordinary Least Square (OLS)

- Optimal solution can coincide with OLS:

$$
\begin{array}{ll}
\inf & \det(\mathbf{D}) \\
\text{s.t.} & \mathbb{E}_{\widehat{\mathbb{P}}}\left[(\|\mathbf{D}^{-1}(\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}_0 - \hat{\mathbf{Z}}\tilde{\boldsymbol{s}})\|_2)^2\right] \leq 1 \\
& \hat{\boldsymbol{z}}_0 \in \mathbb{R}^L,\ \hat{\mathbf{Z}} \in \mathbb{R}^{L\times S},\ \mathbf{D} \in \mathbb{S}^L_{++}.
\end{array}
$$

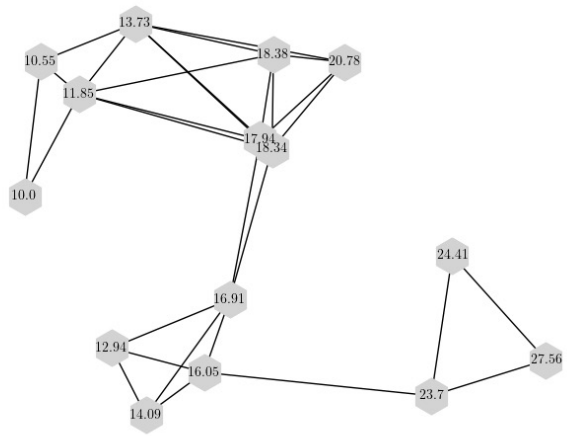
- Statistical properties yet to be discovered.

Agenda

- Challenges of Optimizing Decisions Under Uncertainty
- A Tour of Robust Decision Models
- Estimating Parameters of Disparity Metric
- Asymmetric Uncertainty Sets
- Predictive Model with Side Information
- **Numerical Studies**

Resource Allocation for Emergency Service

- L nodes with uncertain demands z_ℓ, $\ell \in [L]$
- Here-and-now decision: x_ℓ inventory at location $\ell \in [L]$
- Location ℓ is connected to nodes in $\mathcal{I}_\ell \subseteq [L]$.



Attribute function: Largest demand shortfall

$$f(\boldsymbol{x}, \boldsymbol{z}) = \begin{cases} \min & v \\ \text{s.t.} & \sum_{j \in \mathcal{I}_\ell} y_{\ell j} \leq x_i & \forall \ell \in [L], \\ & \sum_{i \in \mathcal{I}_\ell} y_{i\ell} \geq z_\ell - v & \forall \ell \in [L] \\ & y_{\ell j} \geq 0 & \forall j \in \mathcal{I}_\ell,\ \ell \in [L] \\ & v \in \mathbb{R}. \end{cases}$$

Demands are fulfilled if and only if $f(\boldsymbol{x}, \boldsymbol{z}) \leq 0$.

Empirical Optimization

Resource allocation that covers all historical demands

$$V_{\mathsf{EO}} = \left\{ \begin{array}{ll} \min & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}_\omega) \le 0 \quad \forall \omega \in [\Omega] \\ & \boldsymbol{x} \in \mathcal{X}. \end{array} \right.$$

$$= \left\{ \begin{array}{lll} \min & \boldsymbol{c}^\top \boldsymbol{x} & \\ \text{s.t.} & \sum_{j \in \mathcal{I}_\ell} y_{\ell j \omega} \le x_i & \forall \ell \in [L],\ \omega \in [\Omega] \\ & \sum_{i \in \mathcal{I}_\ell} y_{i \ell \omega} \ge z_{\ell \omega} & \forall \ell \in [L],\ \omega \in [\Omega] \\ & y_{\ell j \omega} \ge 0 & \forall j \in \mathcal{I}_\ell,\ \ell \in [L],\ \omega \in [\Omega] \\ & \boldsymbol{x} \in \mathcal{X}. & \end{array} \right.$$

Budgeted Robust Optimization Model

$$\left\{\begin{array}{lll} \min & \boldsymbol{c}^\top \boldsymbol{x} & \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}) \le 0 & \forall \boldsymbol{z} \in \mathcal{Z}(\gamma) \\ & \boldsymbol{x} \in \mathcal{X}. & \end{array}\right.$$

Lifted uncertainty set:

$$\bar{\mathcal{Z}}(\gamma) = \left\{ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \mathcal{Z} \times \mathbb{R}^L_+ \times \mathbb{R}^L_+ \times \mathbb{R}_+ \;\middle|\; \begin{array}{l} \bar{\boldsymbol{w}} - \underline{\boldsymbol{w}} = \mathrm{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}) \\ \|\underline{\boldsymbol{w}} + \bar{\boldsymbol{w}}\| \le u \le \gamma \end{array} \right\}.$$

- $\mathcal{A}(2L+1)$ is the set of all affine functions:
  - Affine recourse adaption on the on lifted variables

Tolerated Robust Satisficing Model

$$\begin{cases} \min & \kappa \\ \text{s.t.} & f(\boldsymbol{x}, \boldsymbol{z}) \leq \kappa \Delta(\boldsymbol{z}, \hat{\boldsymbol{z}}) \quad \forall \boldsymbol{z} \in \mathcal{Z} \\ & \boldsymbol{c}^\top \boldsymbol{x} \leq V_1 \\ & \boldsymbol{x} \in \mathcal{X}, \kappa \geq 0 \end{cases}$$

RSOME Code

$$\bar{\mathcal{Z}} = \left\{ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \mathcal{Z} \times \mathbb{R}^L_+ \times \mathbb{R}^L_+ \times \mathbb{R}_+ \;\middle|\; \begin{array}{l} \bar{\boldsymbol{w}} - \underline{\boldsymbol{w}} = \mathrm{D}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}) \\ \|_{p1} \underline{\boldsymbol{w}} + \bar{\boldsymbol{w}}\| \le u \le \bar{u} \end{array} \right\}.$$

$$\begin{array}{lll}
\min & \kappa & \\
\text{s.t.} & \displaystyle\sum_{j \in \mathcal{I}_\ell} y_{\ell j}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \le x_\ell & \forall \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}} \\
& \kappa u \ge z_\ell - \displaystyle\sum_{i \in \mathcal{I}_\ell} y_{i\ell}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) & \forall \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}} \\
& y_{\ell j}(\underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \ge 0 & \forall j \in \mathcal{I}_\ell,\ \ell \in [L],\ (\boldsymbol{z}, \underline{\boldsymbol{w}}, \bar{\boldsymbol{w}}, u) \in \bar{\mathcal{Z}} \\
& y_{\ell j} \in \mathcal{A}(2L+1) & \forall j \in \mathcal{I}_\ell,\ \ell \in [L] \\
& \boldsymbol{c}^\top \boldsymbol{x} \le V_1 & \\
& \boldsymbol{x} \in \mathcal{X}, \kappa \ge 0. &
\end{array}$$

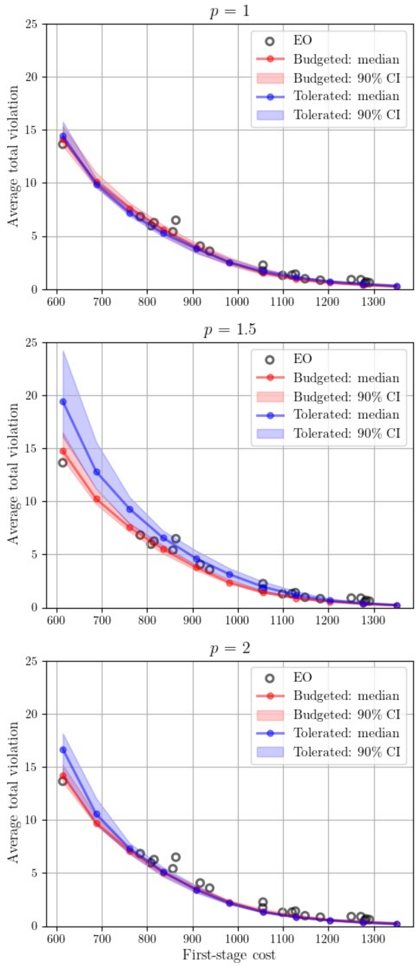
```python
import numpy as np
import rsome as rso
from rsome import ro
from rsome import cpt_solver as cpt

def rs_decision(c, arcs, xbar, tau, zin, p1=2, p2=2, display=True):

    I = len(c)
    ij= np.array([(i, i) for i in range(I)] +
                 arcs + [a[::-1] for a in arcs])
    K = ij.shape[0]
    q, Q = mvcs(zin, p1, p2, display)
    gbar = np.linalg.norm(zin@Q - q, p1, axis=1).max()

    model = ro.Model()
    z = model.rvar(I)
    w1 = model.rvar(I)
    w2 = model.rvar(I)
    u = model.rvar()
    zset = (w1 - w2 == Q@z - q,
            rso.norm(w1 + w2, p1) <= u, u <= gbar,
            w1 >= 0, w2 >= 0,
            z >= 0, z <= 100)

    x = model.dvar(I)
    kappa = model.dvar()
    y = model.ldr(K)
    y.adapt(w1)
    y.adapt(w2)
    y.adapt(u)

    model.minmax(kappa, zset)
    model.st(c @ x <= tau)
    for i in range(I):
        model.st(y[ij[:, 0] == i].sum() <= x[i])
        model.st(z[i] - y[ij[:, 1] == i].sum() <= kappa * u)
    model.st(x >= 0, x <= xbar, y >= 0)
    model.solve(cpt, display=display)

    objval, xs = model.get(), x.get()
    return objval, xs
```

Out of Sample Evaluations

Comparison results with different norms.

- Average total violation:
  the expected number of unfulfilled demands.

.



References

- Lou, Z., Chen, Z., Sim, M., Xie, J., & Xiong, P. (2024). Estimation and Prediction Procedures for Unified Robust Decision Models. *Available at SSRN 4890089*.

Thank You！

Melvyn SIM
melvynsim@nus.edu.sg