# Distributionally Robust Optimization

Melvyn Sim
Department of Analytics & Operations
NUS Business School

ZJU-NUS International Summer School 2020

# Optimization in Practice

**Model**

Input parameters: $\tilde{\boldsymbol{z}}$ RV

Decisions : $\boldsymbol{x}, \boldsymbol{y}(\tilde{\boldsymbol{z}})$

Output function: $f(\boldsymbol{x}, \boldsymbol{y}(\tilde{\boldsymbol{z}}), \tilde{\boldsymbol{z}})$

$\Downarrow$

**Decision Criterion:**

$$\rho(f(\boldsymbol{x}, \boldsymbol{y}(\tilde{\boldsymbol{z}}), \tilde{\boldsymbol{z}}))$$

WRITE THE RV INTO A NUMBER

$\Downarrow$

**Solver:**

$$(\boldsymbol{x}^*, \boldsymbol{y}^*(\cdot)) = \arg\max_{\boldsymbol{x}, \boldsymbol{y}()} \rho(f(\boldsymbol{x}, \boldsymbol{y}(\tilde{\boldsymbol{z}}), \tilde{\boldsymbol{z}}))$$

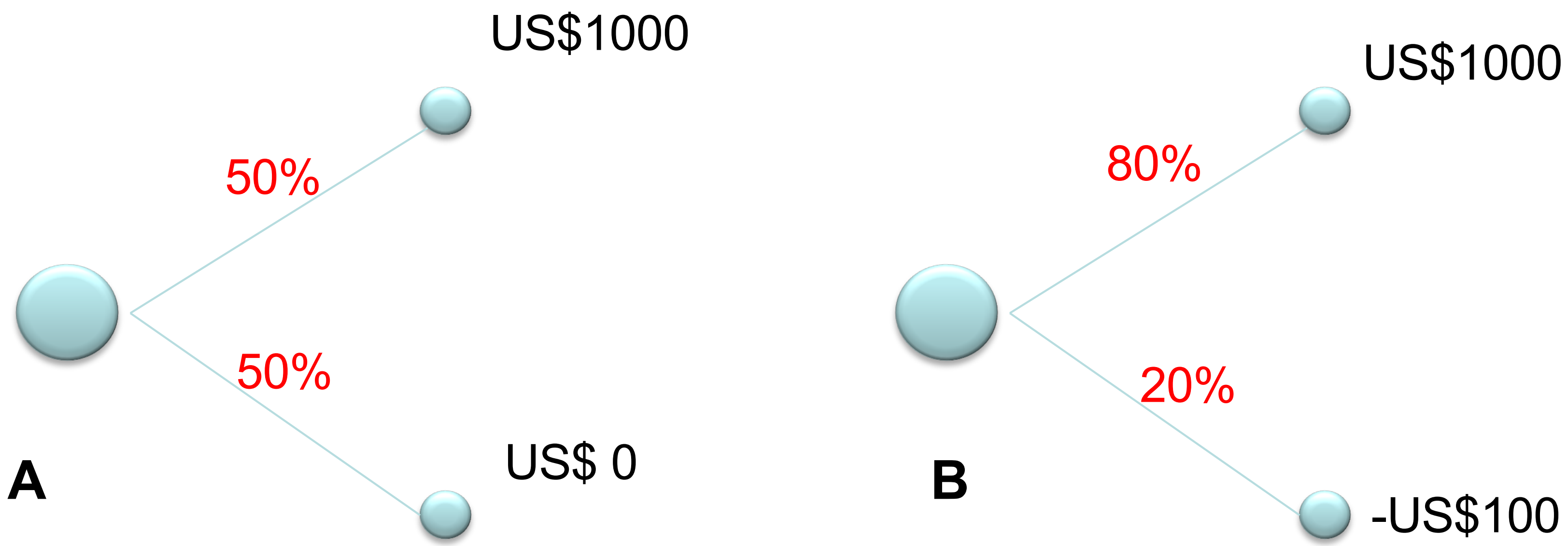
# Some Challenges

CPPLEX

- Fact: Most optimization problems used in practice are deterministic.
  - Optimization problems under uncertainty are generally computational intractable.
  - Lack of tractable framework that addresses uncertainty that can be broadly applied.

# Some Challenges

- Preference under uncertainty is more complex and subjective

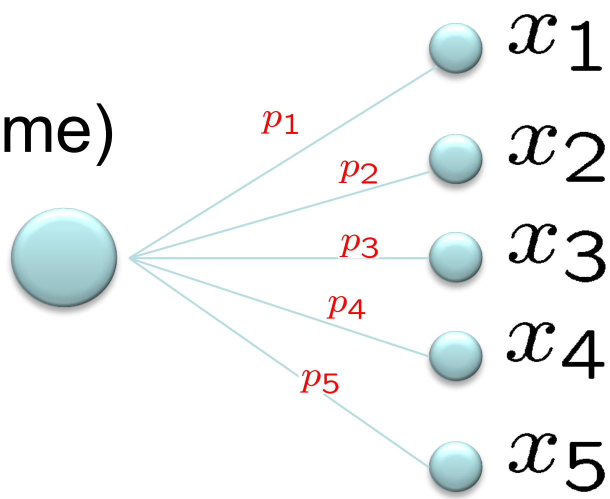


# Facets of Uncertainty

- Risk: Uncertainty with known frequency of occurrence distribution is known
- Ambiguity: Uncertainty without known frequency of occurrence
  - Also known as Knightian uncertainty, due to economist Frank Knight (1885-1972)

# Facets of Uncertainty: Risk

- Uncertainty described via Random Variables
  - Widely accepted method in math/statistics
  - Characterized by
    - Sample space (all possible outcomes, usually exponential or infinite)
    - Distributions (probability of outcome)



# Facets of Uncertainty: Risk

- Preference for Risk via Expected Utility
  - Based on the following Axioms (von Neumann/Morgenstern)

money

$$\rho(\tilde{v}) = \mathbb{E}[u(\tilde{v})]$$

utility

concave-risk averse

# Facets of Uncertainty: Risk

- Issues:
  - Practically prohibitive to obtained exact distributions
    - » Absence or limited historical data
    - » Reliability of historical data in predicting outcomes; non-stationary
    - » Difficulty of describing multivariate random variable

# A Portfolio Optimization Case Study

- 24 small cap stocks from different industry categories
- Historical returns from April 17 1998 to June 1, 2006
- Return and Covariance estimated from initial 80% of the data. Evaluate performance on last 20%.

# A Portfolio Optimization Case Study

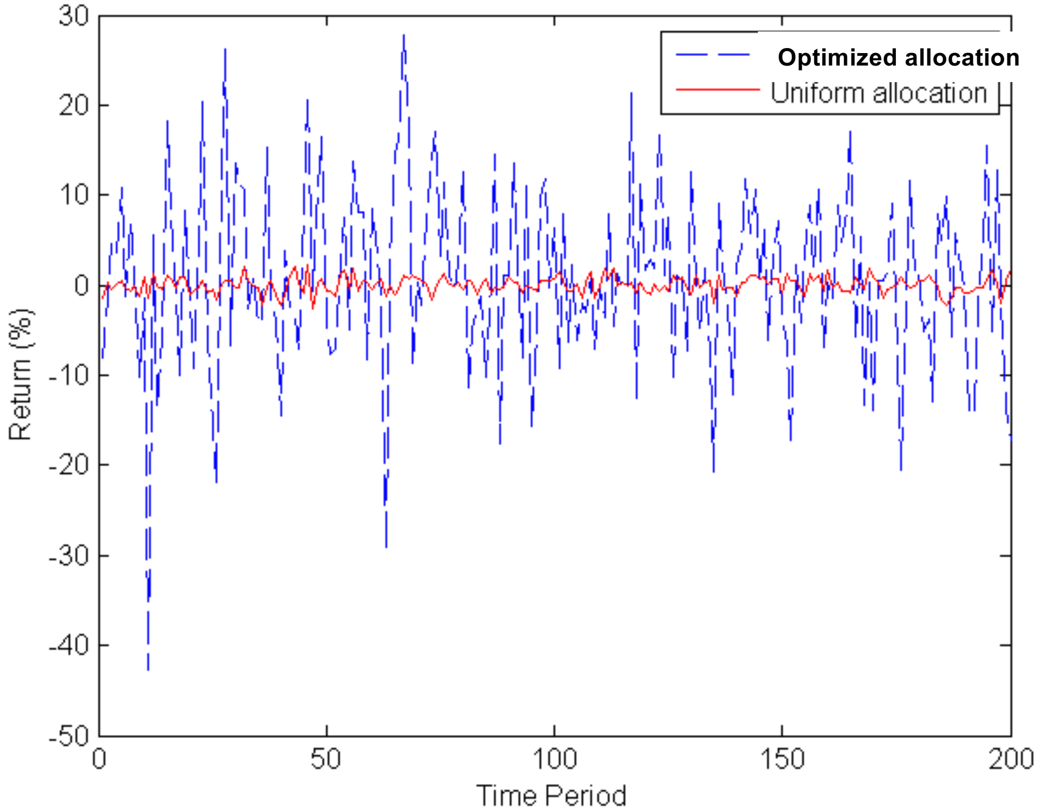
– Markowitz model

- $\tilde{\boldsymbol{r}}$: Vector of stock returns. Estimated Mean $\hat{\boldsymbol{\mu}}$ and Covariance $\hat{\boldsymbol{\Sigma}}$. (Estimated)

$$\begin{array}{rl} \text{min} & \boldsymbol{x}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{x} \\ \text{s.t.} & \widehat{\boldsymbol{\mu}}'\boldsymbol{x} = \frac{1}{n}\sum_{j=1}^{n} \widehat{\mu}_j \\ & \boldsymbol{x}'\mathbf{1} = 1, \end{array}$$

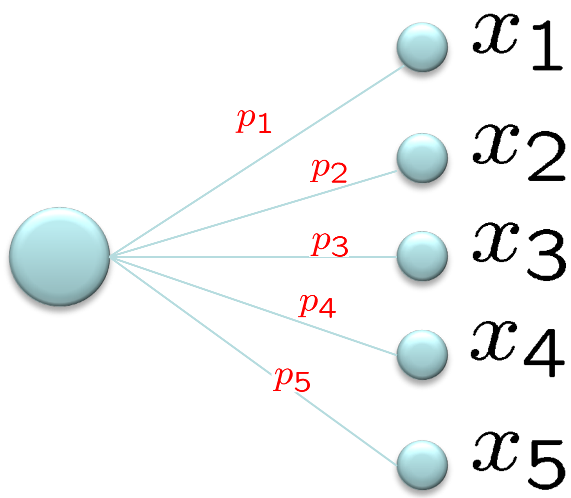
# A Portfolio Optimization Case Study

– Optimizing over unreliable historical data can be catastrophic!!!



# Facets of Uncertainty: Ambiguity

- Probability distributions not completely known




$$(p_1, \ldots, p_5) \in \mathcal{F} \subseteq \{\boldsymbol{p} \mid \boldsymbol{p}'\mathbf{1} = 1, \boldsymbol{p} \geq 0\}$$

$$\mathbb{E}_{\mathbb{P}}[u(\tilde{v})] \in \left[\inf_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[u(\tilde{v})], \sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[u(\tilde{v})]\right]$$

# Ellsberg Paradox

- Box 1: 50 red balls and 50 blue balls
- Box 2: 100 red and blue balls with unknown proportions

Payoffs: US$1000 for choosing a red ball.
Which box will you choose?

# Ellsberg Paradox

– Box 1: 50 red balls and 50 blue balls
– Box 2: 100 red and blue balls with unknown proportions

Box 1 is generally preferred!

Ambiguity Aversion

# Ellsberg Paradox

– Box 1: 5 red balls and 95 blue balls
– Box 2: 100 red and blue balls with unknown proportions

Payoffs: US$1000 for choosing a red ball.
Which box will you choose?

# Preference under Ambiguity

Hurwitz Criterion:

worst case utility

best case utility

$$\rho_\alpha(\tilde{v}) = \alpha \inf_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}[u(\tilde{v})] + (1-\alpha) \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}[u(\tilde{v})]$$

- Risk preference determined by utility function, $u$
  - Concave utility: Risk aversion
  - Convex utility: Risk seeking
- Ambiguity preference determined by $\alpha \in [0,1]$
  - $\alpha = 1$ : Extreme ambiguity aversion
  - $\alpha = 0$ : Extreme ambiguity seeking
- Generally difficult to maximize. Lack of concavity, among others.

# Preference under Ambiguity

Decision Criterion in Robust Optimization:

$$\rho(\tilde{v}) = \inf_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[u(\tilde{v})]$$

- Concave utility: Risk aversion
- Extreme ambiguity aversion ( $\alpha = 1$ )
- Criterion is concave among uncertain choices.
- Convex preference and some tractability.

# Preference under Ambiguity

- Growing interests in ambiguity
  - Economic community
    - Maximin Expected Utility (Gilboa and Schmeidler)
    - Variational Preferences (Maccheroni, Marinacci, and Rustichini)
  - Finance & Insurance community
    - Coherent Risk Measure (Artner, Delbean, Eber and Health)
    - Convex Risk Measure (Foellmer and Schied)
    - Acceptability Index (Cherny and Madan)
  - Operation Research community
    - Robust Optimization (Ben-Tal, Bertsimas, El-Ghoui, Kuhn, Nemirovski, Sim, …)

# Newsvendor Problem

- Buy x items at cost c
- Items sold at price p
- Salvage price is zero
- Risk neutral

$$\begin{array}{ll} \max & -cx + \mathbb{E}[\min\{px, p\tilde{d}\}] \\ \text{s.t.} & x \geq 0. \end{array}$$

$$f(x, \tilde{d}) = \min\{px, p\tilde{d}\}$$

# Robust Newsvendor Problem

- Distributionally Robust Optimization approach when demand distribution is not fully known.
  - Risk Neutral and Ambiguity Averse

linear utility

$$
\begin{array}{ll}
\max & -cx + \inf_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}] \\
\text{s.t.} & x \geq 0.
\end{array}
$$

# Robust Newsvendor 1

- Only support of demand is known

family

$$\mathcal{F} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} \tilde{d} \sim \mathbb{P} \\ \mathbb{P}[\tilde{d} \in [20, 30]] = 1 \end{array} \right\}$$

$$\begin{array}{ll} \max & -cx + \inf_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}] \\ \text{s.t.} & x \geq 0. \end{array}$$

Since $\inf_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}] = \min\{px, p20\}$,
we have, $x^* = 20$.

# Robust Newsvendor 2

- Discrete demand with moments information

$$\mathcal{F} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} \tilde{d} \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\tilde{d}] = 15 \\ \mathbb{E}_{\mathbb{P}}[(\tilde{d} - 15)^2] = 25 \\ \mathbb{P}[\tilde{d} \in \{1, 2, .., 30\}] = 1 \end{array} \right\}$$

P(tilde-d=d)=qd is unknown

# Robust Newsvendor 2

$$
\begin{array}{ll}
\inf & \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}] \\
\text{s.t.} & \mathbb{E}_{\mathbb{P}}[\tilde{d}] = 15 \\
& \mathbb{E}_{\mathbb{P}}[(\tilde{d} - 15)^2] = 25 \\
& \mathbb{P}[\tilde{d} \in \{1, 2, .., 30\}] = 1
\end{array}
$$

$$
\begin{array}{ll}
\min & \sum_{d=1}^{30} \min\{px, pd\} q_d \\
\text{s.t.} & \sum_{d=1}^{30} d q_d = 15 \\
& \sum_{d=1}^{30} (d - 15)^2 q_d = 25 \\
& \sum_{d=1}^{30} q_d = 1 \\
& \boldsymbol{q} \geq 0.
\end{array}
$$

linear optimization problem

# Robust Newsvendor 2

- Primal

$$
\begin{aligned}
\min \quad & \sum_{d=1}^{30} \min\{px, pd\} q_d \\
\text{s.t.} \quad & \sum_{d=1}^{30} d q_d = 15 && \text{s2} \\
& \sum_{d=1}^{30} (d-15)^2 q_d = 25 && \text{s3} \\
& \sum_{d=1}^{30} q_d = 1 && \text{s1} \\
& \boldsymbol{q} \geq 0.
\end{aligned}
$$

- Dual

$$
\begin{aligned}
\max \quad & s_1 + 15 s_2 + 25 s_3 \\
\text{s.t.} \quad & s_1 + d s_2 + (d-15)^2 s_3 \leq \min\{px, pd\} \\
& \qquad \forall d \in \{1, \ldots, 20\} \\
& s_1, s_2, s_3 \in \mathbb{R}
\end{aligned}
$$

# Robust Newsvendor 2

$$
\begin{aligned}
\max \quad & -cx + \inf_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}] \\
\text{s.t.} \quad & x \geq 0.
\end{aligned}
$$

$$
\begin{aligned}
\max \quad & -cx + s_1 + 15s_2 + 25s_3 & \\
\text{s.t.} \quad & s_1 + ds_2 + (d-15)^2 s_3 \leq px & \forall d \in \{1, \ldots, 20\} \\
\text{s.t.} \quad & s_1 + ds_2 + (d-15)^2 s_3 \leq pd & \forall d \in \{1, \ldots, 20\} \\
& x \geq 0 & \\
& s_1, s_2, s_3 \in \mathbb{R} &
\end{aligned}
$$

- A Linear Optimization Model!

# Robust Newsvendor 3

- Continuous demand with convex piecewise moments information

$$
\mathcal{F} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} \tilde{d} \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\tilde{d}] = 15 \\ \mathbb{E}_{\mathbb{P}}[|\tilde{d} - 15|] \leq 5 \\ \mathbb{P}[\tilde{d} \in [0, 30]] = 1 \end{array} \right\}
$$

# Robust Newsvendor 3

$$\begin{aligned}
\inf \quad & \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}] \\
\text{s.t.} \quad & \mathbb{E}_{\mathbb{P}}[\tilde{d}] = 15 \\
& \mathbb{E}_{\mathbb{P}}[|(\tilde{d} - 15)|] \leq 5 \\
& \mathbb{P}[\tilde{d} \in [0, 30]] = 1
\end{aligned}$$

Some abuse of notation!

$$\begin{aligned}
\inf \quad & \int_0^{30} \min\{px, pz\} f(z) dz \\
\text{s.t.} \quad & \int_0^{30} z f(z) dz = 15 \\
& \int_0^{30} |(z - 15)| f(z) dz \leq 5 \\
& \int_0^{30} f(z) dz = 1 \\
& f(z) \geq 0 \ \ \forall z \in [0, 30]
\end{aligned}$$

# Robust Newsvendor 3

- Primal

$$\begin{array}{ll} \inf & \int_0^{30} \min\{px, pz\} f(z) dz \\ \text{s.t.} & \int_0^{30} z f(z) dz = 15 \\ & \int_0^{30} |(z-15)| f(z) dz \le 5 \\ & \int_0^{30} f(z) dz = 1 \\ & f(z) \ge 0 \quad \forall z \in [0, 30] \end{array}$$

s2

s3

s1

- Dual (Semi Infinite Programing)

aka Robust Counterpart

$$\begin{array}{ll} \sup & s_1 + 15 s_2 + 5 s_3 \\ \text{s.t.} & \boxed{\begin{array}{l} s_1 + z s_2 + |z-15| s_3 \le px \quad \forall z \in [0, 30] \\ s_1 + z s_2 + |z-15| s_3 \le pz \quad \forall z \in [0, 30] \end{array}} \\ & s_3 \le 0 \end{array}$$

# Robust Newsvendor 3

$$s_1 + zs_2 + |z - 15|s_3 \le px \quad \forall z \in [0, 30]$$

$$\Updownarrow$$

$$\max_{z \in [0,30]} \{zs_2 + |z - 15|s_3\} \le px - s_1$$

$$\Updownarrow$$

$$\max_{z \in [0,30], u \ge |z-15|} \{zs_2 + us_3\} \le px - s_1$$

Why “greater than” inequality?

# Robust Newsvendor 3

$$\max_{z \in [0,30], u \geq |z-15|} \{zs_2 + us_3\}$$

$$\Updownarrow$$

$$\begin{array}{lll} \max & zs_2 + us_3 & \\ \text{s.t.} & z \leq 30 & : r_1 \\ & -z \leq 0 & : r_2 \\ & z - u \leq 15 & : t_1 \\ & -z - u \leq -15 & : t_2 \end{array}$$

min 30r1+15t1-15t2

s.t. r1-r2+t1+t2=s2

-t1-t2=s3

r1,r2,s1,s2 greater than 0

# Robust Newsvendor 3

$$\begin{array}{ll} \max & zs_2 + us_3 \\ \text{s.t.} & z \leq 30 \\ & -z \leq 0 \\ & z - u \leq 15 \\ & -z - u \leq -15 \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \min & 30r_1 + 15t_1 - 15t_2 \\ \text{s.t.} & r_1 - r_2 + t_1 - t_2 = s_2 \\ & -t_1 - t_2 = s_3 \\ & r_1, r_2, t_1, t_2 \geq 0 \end{array}$$

# Robust Newsvendor 3

$$\max_{z\in[0,30],u\geq|z-15|} \{zs_2 + us_3\} \leq px - s_1$$

$$\Updownarrow$$

$$\begin{array}{ll} \min & \{30r_1 + 15t_1 - 15t_2\} \leq px - s_1 \\ r_1, r_2, t_1, t_2 \ : & \\ r_1 - r_2 + t_1 - t_2 = s_2 & \\ -t_1 - t_2 = s_3 & \\ r_1, r_2, t_1, t_2 \geq 0 & \end{array}$$

$$\Updownarrow$$

$$\begin{array}{c} 30r_1 + 15t_1 - 15t_2 \leq px - s_1 \\ \exists r_1, r_2, t_1, t_2 \ : \\ r_1 - r_2 + t_1 - t_2 = s_2 \\ -t_1 - t_2 = s_3 \\ r_1, r_2, t_1, t_2 \geq 0 \end{array}$$

# Robust Newsvendor 3

- Similarly

$$s_1 + zs_2 + |z - 15|s_3 \leq pz \quad \forall z \in [0, 30]$$

$$\Updownarrow$$

$$\max_{z \in [0,30], u \geq |z-15|} \{z(s_2 - p) + us_3\} \leq -s_1$$

$$\Updownarrow$$

$$\begin{gathered} 30\bar{r}_1 + 15\bar{t}_1 - 15\bar{t}_2 \leq -s_1 \\ \exists \bar{r}_1, \bar{r}_2, \bar{t}_1, \bar{t}_2 \quad : \\ \bar{r}_1 - \bar{r}_2 + \bar{t}_1 - \bar{t}_2 = s_2 - p \\ -\bar{t}_1 - \bar{t}_2 = s_3 \\ \bar{r}_1, \bar{r}_2, \bar{t}_1, \bar{t}_2 \geq 0 \end{gathered}$$

# Robust Newsvendor 3

$$\begin{array}{ll} \max & -cx + \inf\limits_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}] \\ \text{s.t.} & x \geq 0. \end{array}$$

$$\begin{array}{ll} \max & -cx + s_1 + 15s_2 + 5s_3 \\ \text{s.t.} & 30r_1 + 15t_1 - 15t_2 \leq px - s_1 \\ & r_1 - r_2 + t_1 - t_2 = s_2 \\ & -t_1 - t_2 = s_3 \\ & 30\bar{r}_1 + 15\bar{t}_1 - 15\bar{t}_2 \leq -s_1 \\ & \bar{r}_1 - \bar{r}_2 + \bar{t}_1 - \bar{t}_2 = s_2 - p \\ & -\bar{t}_1 - \bar{t}_2 = s_3 \\ & r_1, r_2, t_1, t_2, \bar{r}_1, \bar{r}_2, \bar{t}_1, \bar{t}_2 \geq 0 \\ & s_1, s_2 \in \mathbb{R} \\ & s_3 \leq 0, x \geq 0 \end{array}$$

- A Linear Optimization Problem

多面鲁棒副本

# Polyhedral Robust Counterpart

$$\boldsymbol{y}'\boldsymbol{z} \leq t \quad \forall \boldsymbol{z} \in \mathcal{U}$$

$$\Updownarrow$$

$$\max_{\boldsymbol{z}\in\mathcal{U}} \boldsymbol{y}'\boldsymbol{z} \leq t$$

# Uncertainty Set

- Polyhedral uncertainty set

$$\mathcal{U} = \{\boldsymbol{z} \mid \exists \boldsymbol{u} \ : \ \boldsymbol{C}\boldsymbol{z} + \boldsymbol{D}\boldsymbol{u} \leq \boldsymbol{d}\}$$

- Why include u?

# The Price of Robustness

- Budget of Uncertainty Set

$$\mathcal{U}(r) = \{\boldsymbol{z} \mid \|\boldsymbol{z}\|_\infty \leq \cancel{r}\,1, \|\boldsymbol{z}\|_1 \leq r\}$$

$$= \left\{ \boldsymbol{z} \;\middle|\; \begin{array}{l} \exists \boldsymbol{u} \quad : \\ \boldsymbol{z} \leq \boldsymbol{u} \\ -\boldsymbol{z} \leq \boldsymbol{u} \\ \boldsymbol{0} \leq \boldsymbol{u} \leq \boldsymbol{1} \\ \boldsymbol{1}'\boldsymbol{u} \leq \cancel{r}\ r \end{array} \right\}$$

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1. | **SHOCK-WAVES ON THE HIGHWAY**<br>By: RICHARDS, PI<br>OPERATIONS RESEARCH Volume: 4 Issue: 1 Pages: 42-51 Published: 1956 | 140 | 182 | 168 | 155 | 78 | 1973 | 30.35 |
| 2. | **EFFECTIVE HEURISTIC ALGORITHM FOR TRAVELING-SALESMAN PROBLEM**<br>By: LIN, S; KERNIGHAN, BW<br>OPERATIONS RESEARCH Volume: 21 Issue: 2 Pages: 498-516 Published: 1973 | 90 | 82 | 81 | 76 | 33 | 1781 | 37.10 |
| 3. | **The price of robustness**<br>By: Bertsimas, D; Sim, M<br>OPERATIONS RESEARCH Volume: 52 Issue: 1 Pages: 35-53 Published: JAN-FEB 2004 | 215 | 196 | 216 | 222 | 152 | 1738 | 102.24 |
| 4. | **ALGORITHMS FOR THE VEHICLE-ROUTING AND SCHEDULING PROBLEMS WITH TIME WINDOW CONSTRAINTS**<br>By: SOLOMON, MM<br>OPERATIONS RESEARCH Volume: 35 Issue: 2 Pages: 254-265 Published: MAR-APR 1987 | 114 | 141 | 151 | 141 | 83 | 1702 | 50.06 |
| 5. | **SCHEDULING OF VEHICLES FROM CENTRAL DEPOT TO NUMBER OF DELIVERY POINTS**<br>By: CLARKE, G; WRIGHT, JW<br>OPERATIONS RESEARCH Volume: 12 Issue: 4 Pages: 568 -& Published: 1964 | 116 | 104 | 119 | 125 | 60 | 1686 | 29.58 |
| 6. | **A PROOF FOR THE QUEUING FORMULA - L=LAMBDA-W**<br>By: LITTLE, JDC<br>OPERATIONS RESEARCH Volume: 9 Issue: 3 Pages: 383-387 Published: 1961 | 48 | 72 | 61 | 61 | 28 | 1391 | 23.18 |
| 7. | **REGRET IN DECISION-MAKING UNDER UNCERTAINTY**<br>By: BELL, DE<br>OPERATIONS RESEARCH Volume: 30 Issue: 5 Pages: 961-981 Published: 1982 | 37 | 64 | 76 | 70 | 45 | 1168 | 29.95 |
| 8. | **OPTIMUM LOCATIONS OF SWITCHING CENTERS + ABSOLUTE CENTERS + MEDIANS OF GRAPH**<br>By: HAKIMI, SL<br>OPERATIONS RESEARCH Volume: 12 Issue: 3 Pages: 450 -& Published: 1964 | 64 | 59 | 65 | 77 | 34 | 1134 | 19.89 |
| 9. | **DECOMPOSITION PRINCIPLE FOR LINEAR-PROGRAMS**<br>By: DANTZIG, GB; WOLFE, P<br>OPERATIONS RESEARCH Volume: 8 Issue: 1 Pages: 101-111 Published: 1960 | 47 | 49 | 47 | 45 | 23 | 1012 | 16.59 |
| 10. | **ROBUST OPTIMIZATION OF LARGE-SCALE SYSTEMS**<br>By: MULVEY, JM; VANDERBEI, RJ; ZENIOS, SA<br>OPERATIONS RESEARCH Volume: 43 Issue: 2 Pages: 264-281 Published: MAR-APR 1995 | 83 | 80 | 72 | 85 | 47 | 951 | 36.58 |

# Polyhedral Robust Counterpart

$$\mathcal{U} = \{\boldsymbol{z} \mid \exists \boldsymbol{u} \ : \ \boldsymbol{C}\boldsymbol{z} + \boldsymbol{D}\boldsymbol{u} \leq \boldsymbol{d}\}$$

$$\max_{\boldsymbol{z} \in \mathcal{U}} \boldsymbol{y}'\boldsymbol{z} \leq t$$

# Polyhedral Robust Counterpart

$$\left\{(\boldsymbol{y},t)\,\middle|\,\boldsymbol{y}'\boldsymbol{z} \le t \quad \forall \boldsymbol{z} \in \mathcal{U}\right\}$$

$$\Updownarrow$$

$$\left\{(\boldsymbol{y},t)\,\middle|\,\begin{array}{l}\exists p \ge \mathbf{0} \quad : \\ \boldsymbol{p}'\boldsymbol{d} \le t \\ \boldsymbol{C}'\boldsymbol{p} = \boldsymbol{y} \\ \boldsymbol{D}'\boldsymbol{p} = \mathbf{0}\end{array}\right\}$$

# Conic Optimization

- LP standard from and its dual

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \end{aligned} \qquad \begin{aligned} \max \quad & \boldsymbol{b}'\boldsymbol{p} \\ \text{s.t.} \quad & \boldsymbol{A}'\boldsymbol{p} \leq \boldsymbol{c} \end{aligned}$$

- Is there a standard form framework for convex optimization problem?

# Generalized Inequality

- Polyhedron

$$\boldsymbol{Ax} \geq \boldsymbol{b}$$

- Nice Characteristics of "$\geq$"
  - Reflexibility: $\boldsymbol{a} \geq \boldsymbol{a}$
  - Antisymmetry: $\boldsymbol{a} \geq \boldsymbol{b}$ and $\boldsymbol{b} \geq \boldsymbol{a}$, then $\boldsymbol{a} = \boldsymbol{b}$.
  - Transitivity: $\boldsymbol{a} \geq \boldsymbol{b}$ and $\boldsymbol{b} \geq \boldsymbol{c}$, then $\boldsymbol{a} \geq \boldsymbol{c}$.
  - Linear operations compatibility:
    - Homogeneity: If $\boldsymbol{a} \geq \boldsymbol{b}$, $\lambda\boldsymbol{a} \geq \lambda\boldsymbol{b}$, for all $\lambda \geq 0$
    - Additivity: If $\boldsymbol{a} \geq \boldsymbol{b}$ and $\boldsymbol{c} \geq \boldsymbol{d}$ then $\boldsymbol{a} + \boldsymbol{c} \geq \boldsymbol{b} + \boldsymbol{d}$.

# Generalized Inequality

- Fact: A lot of nice characteristic of LP comes from the characteristics of "$\geq$"
- Generalization of such characteristic: Conic formulation

$$\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \geq \boldsymbol{0} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \in \mathcal{K}$$

$$\mathcal{K} = \mathbb{R}^m_+$$

# Generalized Inequality

- $\mathcal{K}$ is a cone iff for all $\boldsymbol{x} \in \mathcal{K}$,

$$\lambda \boldsymbol{x} \in \mathcal{K}, \forall \lambda \geq 0.$$

- Conic constraint:

$$\boldsymbol{A}\boldsymbol{x} \succeq_{\mathcal{K}} \boldsymbol{b} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \succeq_{\mathcal{K}} \boldsymbol{0} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \in \mathcal{K}$$

# Generalized Inequality

- Strict inequality

$$\boldsymbol{A}\boldsymbol{x} \succeq_{\mathcal{K}} \boldsymbol{b} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \succeq_{\mathcal{K}} \boldsymbol{0} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \in \mathrm{int}\mathcal{K}$$

- Example

$$\boldsymbol{A}\boldsymbol{x} > \boldsymbol{b} \Leftrightarrow \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} > \boldsymbol{0}$$

# Generalized Inequality

- Not all cone will have nice characteristics.
- Regular cone
  - Closed
    - A sequence of vectors in the cone has a limit
  - Convex
  - Pointed
    - No lines can be drawn.
  - Non-empty interior
    - Full dimensional

# Common Cones

- Three important regular cones:

  1. $\mathbb{R}^n_+$
  2. Second order cone

$$\mathcal{L}^{n+1} = \left\{ (x_0, \underbrace{x_1, \ldots, x_n}_{\boldsymbol{x}}) \,|\, \|\boldsymbol{x}\|_2 \leq x_0 \right\}$$

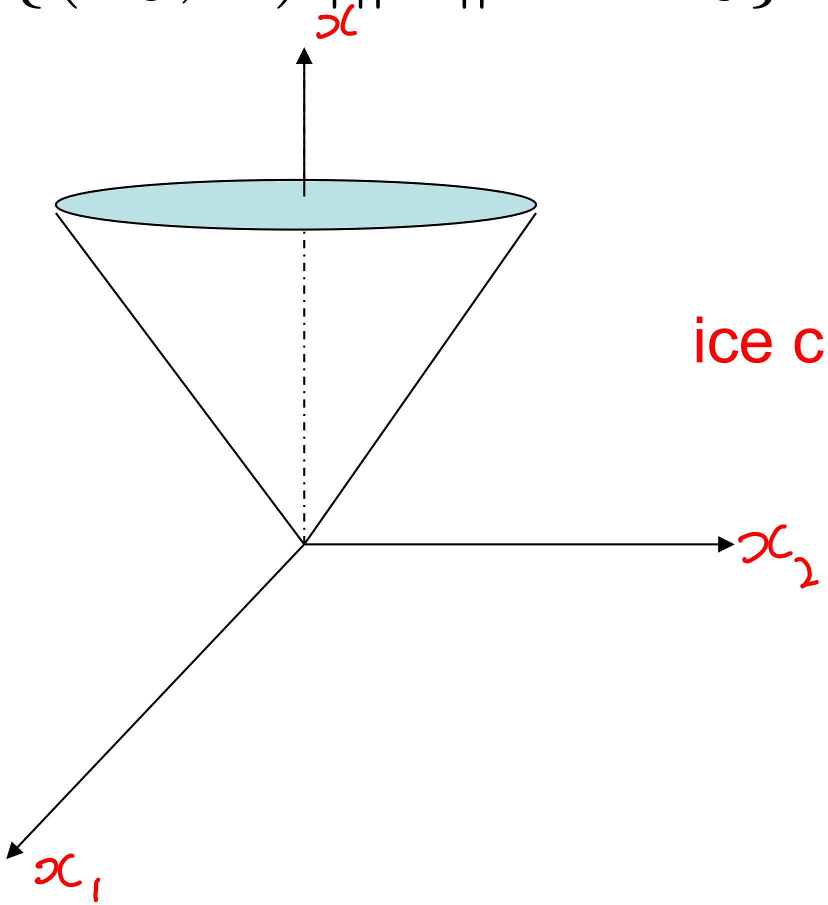
  3. 

$$\mathcal{S}^n_+ = \{\boldsymbol{X} | \boldsymbol{X} \text{ symmetric positive semidefinite matrix } \}$$

# Second Order Cone

$$\mathcal{L}^{n+1} = \{(x_0, \boldsymbol{x}) \mid \|\boldsymbol{x}\|_2 \leq x_0\}$$

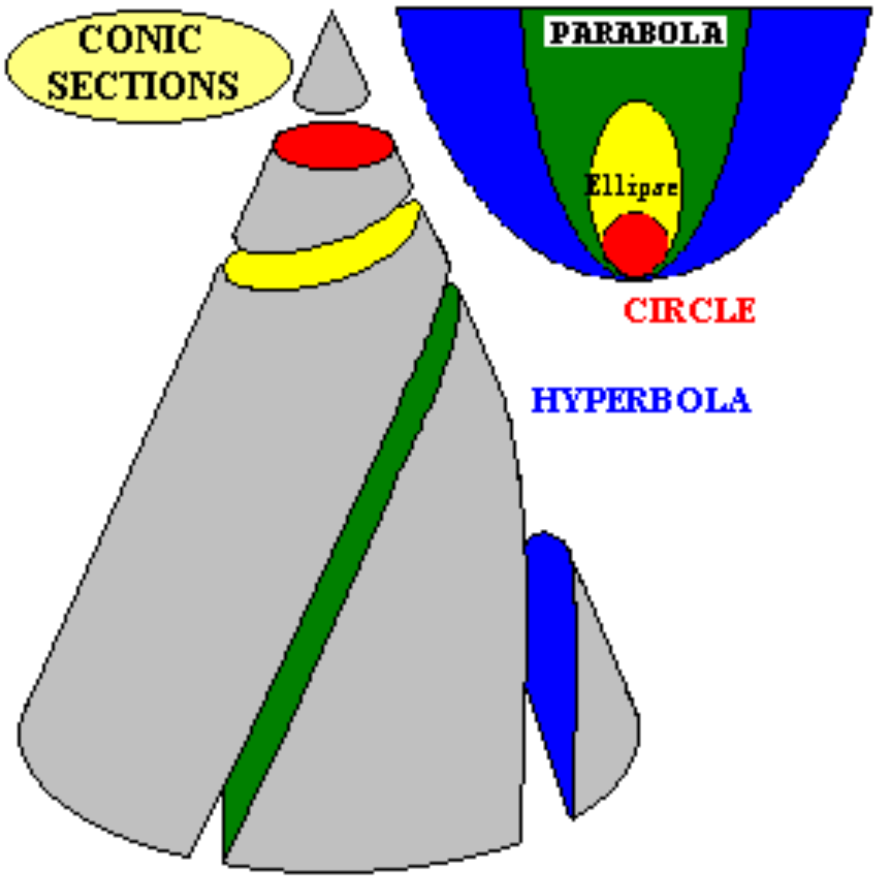
Lawrence cone

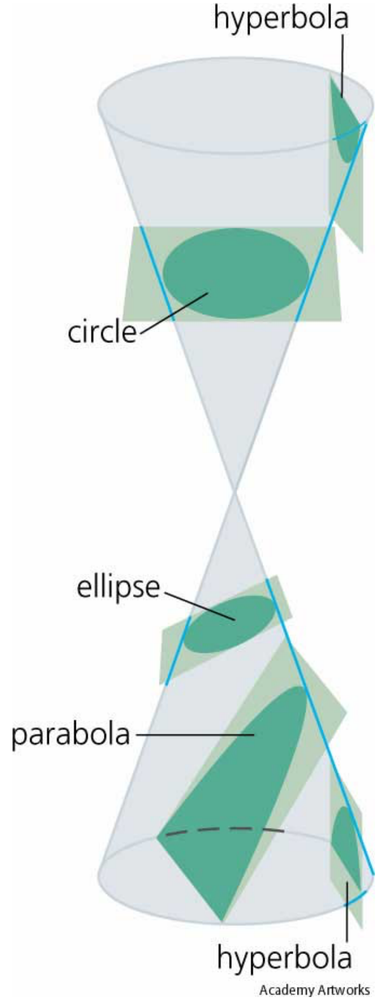



ice cream cone

# Second Order Cone

- Modeling powers of second order cone.



# Second Order Cone

- Second order cone

$$\boldsymbol{x}'\boldsymbol{x} \leq st \quad (s, t \geq 0)$$

$$\Updownarrow$$

$$\left\| \begin{array}{c} \boldsymbol{x} \\ (s-t)/2 \end{array} \right\|_2 \leq (s+t)/2, \Leftrightarrow \begin{bmatrix} (s+t)/2 \\ \boldsymbol{x} \\ (s-t)/2 \end{bmatrix} \succeq_{\boldsymbol{L}^{n+2}} \mathbf{0}$$

# Second Order Cone

$$\boldsymbol{x}'M\boldsymbol{x} + \boldsymbol{b}'\boldsymbol{x} + c \le 0$$

$$\boldsymbol{x}'\boldsymbol{x} \le st$$

$$\Updownarrow$$

$$\left\| \begin{array}{c} \boldsymbol{x} \\ (s-t)/2 \end{array} \right\|_2 \le (s+t)/2$$

# Second Order Cone

- Examples

$$x^4 \leq y$$ more powerful

$$\Updownarrow$$

$$\exists s : x^2 \leq s, s^2 \leq y$$

$$x^8 \leq y$$

$$\Updownarrow$$

$$\exists s_1, s_2 : x^2 \leq s_1, s_1^2 \leq s_2, s_2^2 \leq y$$

# Second Order Cone

- How about?

$$|x|^3 \leq y$$

- Idea: $\exists \alpha : |x| \leq \alpha, \alpha^4 \leq y\alpha$

# Second Order Cone

- Yet another example:

$$x^{3/2} \leq y \qquad\qquad x, y \geq 0$$

- How about?

$$x^{2/3} \leq y \qquad\qquad x, y \geq 0$$

# Second Order Cone

- And another example:

$$\|\boldsymbol{x}\|_4 \leq y$$

- For y>0, we have equivalently

$$\sum_i (x_i/y)^4 \leq 1$$

# Second Order Cone

- …cont
  - Check y=0

$$\sum_i (x_i/y)^4 \leq 1$$

$$\frac{x_i^2}{y^2} \leq \frac{t_i}{y} \qquad \frac{t_i^2}{y^2} \leq \frac{s_i}{y} \quad \forall i$$

$$\sum_i \frac{s_i}{y} \leq 1$$

# Exponential & Power Cones

- New cones supported by MOSEK.

  1. Exponential cone

  $$\mathcal{K}_{\exp} = \mathrm{cl}\{(x_1, x_2, x_3) \mid x_1 \geq x_2 \exp(x_3/x_2), x_1, x_2 > 0\}$$

  2. Power cone

  $$\mathcal{K}_{\boldsymbol{\alpha}} = \mathrm{cl}\{(\boldsymbol{x}, y) \mid |y| \leq x_1^{\alpha_1} \cdots x_M^{\alpha_M}, \boldsymbol{x} \geq \mathbf{0}\}$$

  for $\boldsymbol{\alpha} > 0$, $\sum_{i \in [M]} \alpha_i = 1$.

# Phi-Divergence

- $\boldsymbol{q}$ is the reference distribution.

| **Divergence** | $\phi(t), t \geq 0$ | $F_\phi(\boldsymbol{p}, \boldsymbol{q})$ |
|---|---|---|
| Kullback-Leibler | $t \log t - t + 1$ | $\sum p_m \log(p_m/q_m)$ |
| Burg entropy | $-\log t + t - 1$ | $\sum q_m \log(q_m/p_m)$ |
| J-divergence | $(t-1)\log t$ | $\sum (p_m - q_m)\log(p_m/q_m)$ |
| $\chi^2$-distance | $\frac{1}{t}(t-1)^2$ | $\sum \frac{(p_m-q_m)^2}{p_m}$ |
| Modified $\chi^2$-distance | $(t-1)^2$ | $\sum \frac{(p_m-q_m)^2}{q_m}$ |
| Hellinger distance | $(\sqrt{t}-1)^2$ | $\sum (\sqrt{p_m} - \sqrt{q_m})^2$ |
| $\chi$-divergence of order $\theta > 1$ | $\lvert t-1\rvert^\theta$ | $\sum q_m \lvert 1 - p_m/q_m\rvert^\theta$ |
| Variation distance | $\lvert t-1\rvert$ | $\sum q_m \lvert 1 - p_m/q_m\rvert$ |
| Cressie and Read | $\frac{1-\theta+\theta t - t^\theta}{\theta(1-\theta)}, \theta \neq 0, 1$ | $\frac{1}{\theta(1-\theta)}\left(1 - \sum p_m^\theta q_m^{1-\theta}\right)$ |

# Conic Optimization

- Generalizing convex optimization

Given a convex set nonempty, full dimensional and bounded set $\mathcal{X}$

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \end{array}$$

is the same as

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & y = 1 \\ & (\boldsymbol{x}, y) \in \mathcal{K} \end{array}$$

where

$$\mathcal{K} = \text{cl}\{(\boldsymbol{x}, y) \mid \boldsymbol{x}/y \in \mathcal{X}, y > 0\}$$

# Conic Optimization

- Conic Optimization framework

$$\begin{array}{ll} \min & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}_i\boldsymbol{x} \succeq_{\mathcal{K}_i} \boldsymbol{b} \quad i \in [M] \end{array}$$

- "Standard form"

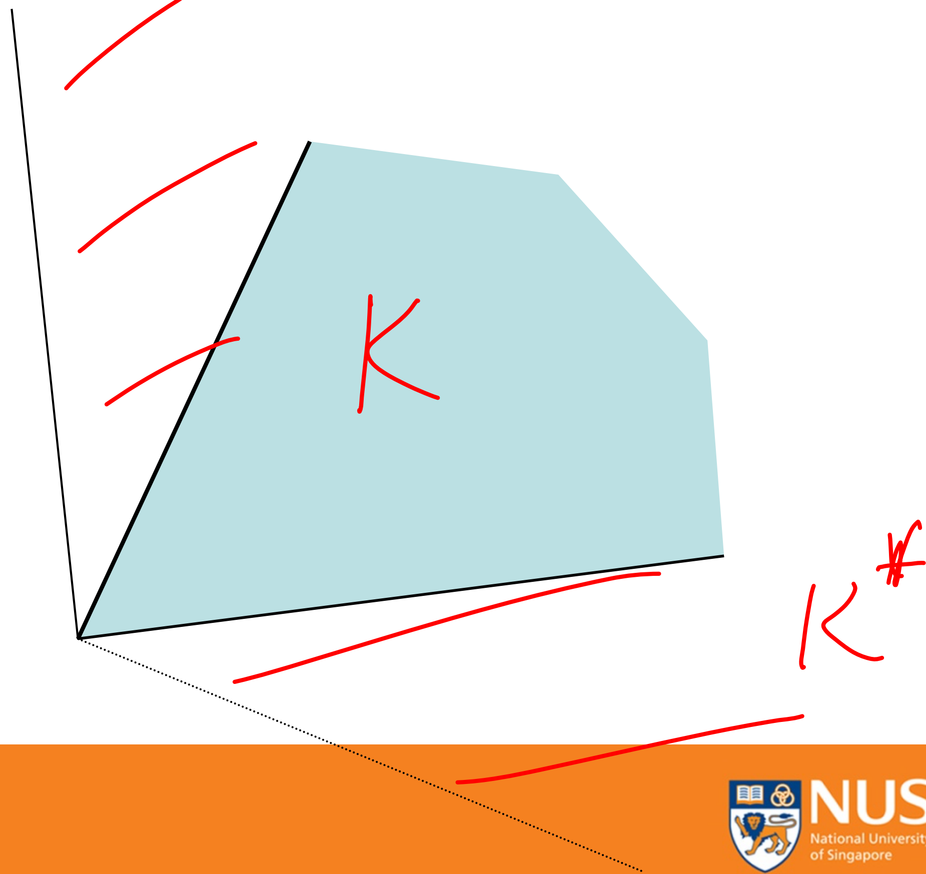
$$\begin{array}{lll} \min & \displaystyle\sum_{j\in[M]} \boldsymbol{c}_j'\boldsymbol{x}_j & \\ \text{s.t.} & \displaystyle\sum_{j\in[M]} \boldsymbol{A}_j\boldsymbol{x}_j = \boldsymbol{b} & \\ & \boldsymbol{x}_i \succeq_{\mathcal{K}_i} \boldsymbol{0} & i \in [M] \end{array}$$

# Conic Duality

- Dual Cone

$$\mathcal{K}^* = \{\boldsymbol{y} : \boldsymbol{y}'\boldsymbol{x} \geq 0 \ \ \forall \boldsymbol{x} \in \mathcal{K}\}$$

$$\boldsymbol{x}'\boldsymbol{y} \geq 0 \quad \forall \boldsymbol{x} \succeq_{\mathcal{K}} \mathbf{0}, \boldsymbol{y} \succeq_{\mathcal{K}^*} \mathbf{0}$$



# Conic Duality

- Dual Cone Properties

$$\mathcal{K}^* = \{\boldsymbol{y} \mid \boldsymbol{y}'\boldsymbol{x} \geq 0 \forall \boldsymbol{x} \in \mathcal{K}\}$$

Consider $\mathcal{K} \subset \mathbb{R}^n$ a non-empty set and the set $\mathcal{K}^* = \{\boldsymbol{z} \in \mathbb{R}^n | \boldsymbol{z}'\boldsymbol{x} \geq 0, \forall \boldsymbol{x} \in \mathcal{K}\}$. Then we have that

1. $\mathcal{K}^*$ is a closed convex cone
2. If $\text{int}\mathcal{K} \neq \emptyset$ then $\mathcal{K}^*$ is pointed
3. If $\mathcal{K}$ is a closed convex pointed cone, then $\text{int}\mathcal{K}^* \neq \emptyset$
4. If $\mathcal{K}$ is a closed convex cone, then so is $\mathcal{K}^*$ and $(\mathcal{K}^*)^* = \mathcal{K}$

# Conic Duality

- The three regular cones are self-dual, *i.e.*, $\mathcal{K} = \mathcal{K}^*$.

1. $\mathbb{R}^n_+$
2. Second order cone

$$\boldsymbol{L}^{n+1} = \left\{ (x_0, \underbrace{x_1, \ldots, x_n}_{\boldsymbol{x}}) \middle| \|\boldsymbol{x}\|_2 \leq x_0 \right\}$$

3. 

$$\boldsymbol{S}^n_+ = \{\boldsymbol{X} | \boldsymbol{X} \text{ symmetric positive semi definite matrix } \}$$

# Conic Duality

- Dual cones

1. Exponential cone

$$\mathcal{K}_{\exp} = \mathrm{cl}\{(x_1, x_2, x_3) \mid x_1 \geq x_2 \exp(x_3/x_2), x_1, x_2 > 0\}$$

$$\mathcal{K}^*_{\exp} = \mathrm{cl}\{(y_1, y_2, y_3) \mid -y_1 \log(-y_1/y_3) + y_1 \leq y_2, y_1 < 0, y_3 > 0\}$$

2. Power cone

$$\mathcal{K}_{\boldsymbol{\alpha}} = \{(\boldsymbol{x}, y) \mid |y| \leq x_1^{\alpha_1} \cdots x_M^{\alpha_M}, \boldsymbol{x} \geq 0\}$$

for $\boldsymbol{\alpha} > 0$, $\sum_{i \in [M]} \alpha_i = 1$.

$$\mathcal{K}^*_{\boldsymbol{\alpha}} = \{(\boldsymbol{u}, v) \mid |v| \leq (u_1/\alpha_1)^{\alpha_1} \cdots (u_1/\alpha_M)^{\alpha_M}, \boldsymbol{u} \geq \mathbf{0}\}$$

# Conic Duality

- Conic Duality

$$\begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \succeq_{\mathcal{K}} \boldsymbol{b} \end{aligned}$$

$$\begin{aligned} \max \quad & \boldsymbol{b}'\boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{y} = \boldsymbol{c} \\ & \boldsymbol{y} \succeq_{\mathcal{K}^*} \boldsymbol{0} \end{aligned}$$

# Conic Duality

- Conic Duality

$$
\begin{aligned}
\min \quad & \sum_{j \in [M]} \boldsymbol{c}_j' \boldsymbol{x}_j \\
\text{s.t.} \quad & \sum_{j \in [M]} \boldsymbol{A}_j \boldsymbol{x}_j = \boldsymbol{b} \\
& \boldsymbol{x}_i \succeq_{\mathcal{K}_i} \boldsymbol{0}
\end{aligned}
$$

# Conic Duality

- Weak Duality

$$
\begin{array}{ll}
\min & \boldsymbol{c}'\boldsymbol{x} \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} \succeq_{\mathcal{K}} \boldsymbol{b}
\end{array}
$$

$$
\begin{array}{ll}
\max & \boldsymbol{b}'\boldsymbol{y} \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{y} = \boldsymbol{c} \\
 & \boldsymbol{y} \succeq_{\mathcal{K}^*} \boldsymbol{0}
\end{array}
$$

- For any $\boldsymbol{x}$ feasible in the primal and $\boldsymbol{y}$ feasible in the dual, then $\boldsymbol{c}'\boldsymbol{x} \geq \boldsymbol{y}'\boldsymbol{b}$

# Conic Duality

- Strong Duality?

If the primal is finite and optimal, does it necessarily imply that $\boldsymbol{c}'\boldsymbol{x} = \boldsymbol{y}'\boldsymbol{b}$?

$$\begin{array}{ll} \min & x_2 \\ \text{s.t.} & \sqrt{x_1^2 + x_2^2} \leq x_1 \end{array}$$

$$\begin{array}{ll} \min & x_2 \\ \text{s.t.} & \begin{pmatrix} x_1 \\ x_1 \\ x_2 \end{pmatrix} \succeq_{\boldsymbol{L}^3} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad : \quad \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix} \end{array}$$

$$\begin{array}{ll} \max & 0 \\ \text{s.t.} & \begin{bmatrix} y_0 + y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ & \sqrt{y_1^2 + y_2^2} \leq y_0 \end{array}$$

# Conic Duality

- $$Z_1 = \inf \quad \boldsymbol{c}'\boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} \succeq_{\mathcal{K}} \boldsymbol{b}$$ $$Z_2 = \sup \quad \boldsymbol{b}'\boldsymbol{y} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{y} = \boldsymbol{c},\ \boldsymbol{y} \succeq_{\mathcal{K}^*} \boldsymbol{0}$$

- , Consider the conic problem $(CP)$ and its dual problem $(CD)$

  1. The dual to $(CD)$ is equivalent to $(CP)$
  2. For any $\boldsymbol{x}$ feasible for $(CP)$ and $\boldsymbol{y}$ feasible for $(CD)$ we have that $\boldsymbol{c}'\boldsymbol{x} \geq \boldsymbol{b}'\boldsymbol{y}$
  3. If $(CP)$ is bounded below and $\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \in \text{int}\mathcal{K}$ for some $\boldsymbol{x}$, then $(CD)$ is solvable and $Z_1 = Z_2$. (analogous result if $(CD)$ is bounded above and strictly feasible).
  4. If either $(CP)$ or $(CD)$ is bounded and strictly feasible; then any primal dual pair $(\boldsymbol{x}, \boldsymbol{y})$ is an optimal solution
     (a) if and only if $\boldsymbol{c}'\boldsymbol{x} = \boldsymbol{b}'\boldsymbol{y}$
     (b) if and only if $\boldsymbol{y}'(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = 0$

# Conic Robust Counterpart

$$\underbrace{\boldsymbol{y}'\boldsymbol{z}} \le t \qquad \forall \boldsymbol{z} \in \mathcal{U}$$

$$\Updownarrow$$

$$\max_{\boldsymbol{z} \in \mathcal{U}} \boldsymbol{y}'\boldsymbol{z} \le t$$

# Conic Robust Counterpart

$$\max_{\boldsymbol{z}\in\mathcal{U}} \boldsymbol{x}'\boldsymbol{z} \le t$$

- Conic uncertainty set

$$\mathcal{U} = \{\boldsymbol{z} \ : \ \exists \boldsymbol{u}, \boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{u} \preceq_K \boldsymbol{b}\}$$

$$\begin{aligned} \max_{\boldsymbol{z}\in\mathcal{U}} \boldsymbol{x}'\boldsymbol{z} = \quad & \min \quad \boldsymbol{b}'\boldsymbol{p} \\ & \text{s.t} \quad \boldsymbol{A}'\boldsymbol{p} = \boldsymbol{x} \\ & \qquad\ \ \boldsymbol{B}'\boldsymbol{p} = \boldsymbol{0} \\ & \qquad\ \ \boldsymbol{p} \succeq_{K^*} \boldsymbol{0} \end{aligned}$$

# Conic Robust Counterpart

$$\boldsymbol{x}'\boldsymbol{z} \le t \qquad \forall \boldsymbol{z} \in \mathcal{U}$$

$$\Updownarrow$$

$$\begin{aligned} \exists \boldsymbol{p} \quad : \boldsymbol{b}'\boldsymbol{p} &\le t \\ \boldsymbol{A}'\boldsymbol{p} &= \boldsymbol{x} \\ \boldsymbol{B}'\boldsymbol{p} &= \boldsymbol{0} \\ \boldsymbol{p} &\succeq_{K^*} \boldsymbol{0} \end{aligned}$$

- Conic feasible set with finite decision variables!

# Robust Newsvendor 4

- Continuous demand with convex moments information

$$\mathcal{F} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} \tilde{d} \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\tilde{d}] = 15 \\ \mathbb{E}_{\mathbb{P}}[(\tilde{d} - 15)^2] \leq 25 \\ \mathbb{P}[\tilde{d} \in [0, 30]] = 1 \end{array} \right\}$$

# Robust Newsvendor 4

- Primal

$$
\begin{array}{ll}
\inf & \int_0^{30} \min\{px, pz\} f(z) dz \\
\text{s.t.} & \int_0^{30} z f(z) dz = 15 \\
& \int_0^{30} (z-15)^2 f(z) dz \leq 25 \\
& \int_0^{30} f(z) dz = 1 \\
& f(z) \geq 0 \quad \forall z \in [0, 30]
\end{array}
$$

- Dual (Semi Infinite Programing)

$$
\begin{array}{ll}
\sup & s_1 + 15 s_2 + 25 s_3 \\
\text{s.t.} & s_1 + z s_2 + (z-15)^2 s_3 \leq px \quad \forall z \in [0, 30] \\
& s_1 + z s_2 + (z-15)^2 s_3 \leq pz \quad \forall z \in [0, 30] \\
& s_3 \leq 0
\end{array}
$$

# Robust Newsvendor 4

$$s_1 + zs_2 + (z-15)^2 s_3 \le px \quad \forall z \in [0,30]$$

$$\Updownarrow$$

$$\max_{z\in[0,30],\,u\ge(z-15)^2} \{zs_2 + us_3\} \le px - s_1$$

Convex but nonlinear optimization.
Can we generalize the duality?

# Robust Newsvendor 4

$$\max_{z\in[0,30],u\geq(z-15)^2} \{zs_2 + us_3\}$$

$$\Updownarrow$$

$$\begin{array}{ll} \max & zs_2 + us_3 \\ \text{s.t.} & 0 \leq z \leq 30 \\ & \sqrt{(15-z)^2 + \left(\frac{u-1}{2}\right)^2} \leq \frac{u+1}{2} \end{array}$$

$$\Updownarrow$$

$$\begin{array}{ll} \max & zs_2 + us_3 \\ \text{s.t.} & 0 \leq z \leq 30 \\ & \begin{pmatrix} u/2 \\ -z \\ u/2 \end{pmatrix} \succeq_{\boldsymbol{L}^3} \begin{pmatrix} -1/2 \\ -15 \\ 1/2 \end{pmatrix} \end{array}$$

# Robust Newsvendor 4

$$
\begin{array}{ll}
\max & z s_2 + u s_3 \\
\text{s.t.} & z \leq 30 \\
& -z \leq 0 \\
& \begin{pmatrix} -u/2 \\ z \\ -u/2 \end{pmatrix} \preceq_{\boldsymbol{L}^3} \begin{pmatrix} 1/2 \\ 15 \\ -1/2 \end{pmatrix}
\end{array}
\quad \Leftrightarrow \quad
\begin{array}{ll}
\min & 30 r_1 + \frac{1}{2} t_0 + 15 t_1 - \frac{1}{2} t_2 \\
\text{s.t.} & r_1 - r_2 + t_1 = s_2 \\
& -\frac{1}{2} t_0 - \frac{1}{2} t_2 = s_3 \\
& (t_0, t_1, t_2) \succeq_{\boldsymbol{L}^3} \mathbf{0}
\end{array}
$$

# Robust Newsvendor 4

$$\begin{aligned}\max \quad & -cx + \inf_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}[\min\{px, p\tilde{d}\}]\\ \text{s.t.} \quad & x \geq 0.\end{aligned}$$

$$\begin{aligned}\sup \quad & -cx + s_1 + 15s_2 + 25s_3\\ \text{s.t.} \quad & 30r_1 + \tfrac{1}{2}t_0 + 15t_1 - \tfrac{1}{2}t_2 \leq px - s_1\\ & r_1 - r_2 + t_1 = s_2\\ & -\tfrac{1}{2}t_0 - \tfrac{1}{2}t_2 = s_3\\ & 30r_1 + \tfrac{1}{2}\bar{t}_0 + 15\bar{t}_1 - \tfrac{1}{2}\bar{t}_2 \leq -s_1\\ & \bar{r}_1 - \bar{r}_2 + \bar{t}_1 = s_2 - p\\ & -\tfrac{1}{2}\bar{t}_0 - \tfrac{1}{2}\bar{t}_2 = s_3\\ & (t_0, t_1, t_2) \succeq_{\boldsymbol{L}^3} \mathbf{0}\\ & (\bar{t}_0, \bar{t}_1, \bar{t}_2) \succeq_{\boldsymbol{L}^3} \mathbf{0}\\ & r_1, r_1\bar{r}_1, \bar{r}_2 \in \mathbb{R}\\ & s_3 \leq 0, x \geq 0\\ & s_1, s_2 \in \mathbb{R}\end{aligned}$$

- A Second Order Conic Optimization Problem (SOCP)

# Beyond Newsvendor

- Multivariate random variables
  - Scenarios are likely exponential in size
- Potentially large number of decision variables

# Ambiguity Set

$$\mathcal{F} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} \tilde{\boldsymbol{z}} \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\boldsymbol{g}(\tilde{\boldsymbol{z}})] \leq \boldsymbol{\mu} \\ \mathbb{P}[\tilde{\boldsymbol{z}} \in \mathcal{W}] = 1 \end{array} \right\}$$

$$\mathcal{U} = \{(\boldsymbol{z}, \boldsymbol{u}) \mid \boldsymbol{g}(\boldsymbol{z}) \leq \boldsymbol{u}, \boldsymbol{z} \in \mathcal{W}\}$$

Assume lifted uncertainty set $\mathcal{U}$ is conic representable and “Slater Condition” is satisfied.

# Lifted Ambiguity Set

$$\mathcal{G} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} (\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{u}}) \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{u}}] \leq \boldsymbol{\mu} \\ \mathbb{P}[(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{u}}) \in \mathcal{U}] = 1 \end{array} \right\}$$

$$\mathcal{U} = \{(\boldsymbol{z}, \boldsymbol{u}) \mid \boldsymbol{g}(\boldsymbol{z}) \leq \boldsymbol{u}, \boldsymbol{z} \in \mathcal{W}\}$$

- $\tilde{\boldsymbol{u}}$ : Lifted/auxiliary random variables
- $\mathcal{F}$ is equivalent to the set of marginal distributions of $\tilde{\boldsymbol{z}}$ under $\mathcal{G}$

$$\sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x}, \tilde{\boldsymbol{z}})] = \sup_{\mathbb{P} \in \mathcal{G}} \mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x}, \tilde{\boldsymbol{z}})]$$

# Lifted Ambiguity Set

- Standardize moment-based ambiguity set
  - Ensure tractability
  - Idea from conic optimization models
- Expectation constraints are restricted to linear functions
- Support set is a conic representable set

# Adaptive Robust Optimization

$\boldsymbol{z}$

First stage

$\boldsymbol{x}$

$\boldsymbol{c}'\boldsymbol{x}$

$\boldsymbol{x} \in X_1$

Second stage

$\boldsymbol{y}$

$$
\begin{array}{rll}
Q(\boldsymbol{x},\boldsymbol{z}) = & \min & \boldsymbol{d}'\boldsymbol{y} \\
& \text{s.t.} & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \boldsymbol{b}(\boldsymbol{z}) \\
& & \boldsymbol{y} \in \mathbb{R}^{N_2}
\end{array}
$$

# Adaptive Robust Optimization

- Second stage optimization problem is an LOP

$$
\begin{aligned}
Q(\boldsymbol{x}, \boldsymbol{z}) = \quad \min \quad & \boldsymbol{d}'\boldsymbol{y} \\
\text{s.t.} \quad & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \boldsymbol{b}(\boldsymbol{z}) \\
& \boldsymbol{y} \in \mathbb{R}^{N_2}
\end{aligned}
$$

- Issue: May not necessarily be feasible.
- Complete recourse:

For any $\boldsymbol{t}$ there exists $\boldsymbol{y}$ such that $\boldsymbol{B}\boldsymbol{y} \geq \boldsymbol{t}$.

  – Strong condition that ensures recourse feasibility

# Adaptive Robust Optimization

- Ambiguity Averse Risk Neutral Model

$$
\begin{aligned}
Z^*_{DRO} = \quad \min \quad & \boldsymbol{c}'\boldsymbol{x} + \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x}, \tilde{\boldsymbol{z}})] \\
\text{s.t.} \quad & \boldsymbol{x} \in X_1
\end{aligned}
$$

  - Generalization of Ambiguity Averse Newsvendor
    - Eg: Multiproduct Newsvendor

# Adaptive Robust Optimization

- Assuming second stage objective is well defined and finite

$$
\begin{aligned}
Q(\boldsymbol{x},\boldsymbol{z}) = \quad & \min \quad \boldsymbol{d}'\boldsymbol{y} \\
& \text{s.t.} \quad \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \boldsymbol{b}(\boldsymbol{z}) \\
& \qquad\;\; \boldsymbol{y} \in \mathbb{R}^{N_2} \\
= \quad & \max \quad (\boldsymbol{b}(\boldsymbol{z}) - \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x})'\boldsymbol{p} \\
& \text{s.t.} \quad \boldsymbol{p} \in \text{Extreme Points of } P
\end{aligned}
$$

where $P = \{\boldsymbol{p} \;:\; \boldsymbol{B}'\boldsymbol{p} = \boldsymbol{d}, \boldsymbol{p} \geq \boldsymbol{0}\}$

- Assume data is affinity dependent on $\boldsymbol{z} \in \mathbb{R}^{I_1}$:

$$
\boldsymbol{A}(\boldsymbol{z}) = \boldsymbol{A}^0 + \sum_{k\in[I_1]} \boldsymbol{A}^k z_k, \boldsymbol{b}(\boldsymbol{z}) = \boldsymbol{b}^0 + \sum_{k\in[I_1]} \boldsymbol{b}^k z_k
$$

# Adaptive Robust Optimization

- $Q(\boldsymbol{x}, \boldsymbol{z})$ is a piecewise biconvex function of $\boldsymbol{x}$ and $\boldsymbol{z}$

$$Q(\boldsymbol{x}, \boldsymbol{z}) = \max_{i \in V} (\boldsymbol{b}(\boldsymbol{z}) - \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x})' \boldsymbol{p}^i$$

where $p^i$, $i \in V$ are the extreme points of $\{\boldsymbol{p} \ : \ \boldsymbol{B}'\boldsymbol{p} = \boldsymbol{d}, \boldsymbol{p} \geq \boldsymbol{0}\}$

# Distributionally Robust Counterpart

$$\sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x},\tilde{\boldsymbol{z}})] \leq y$$

$$\Updownarrow$$

$$\mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x},\tilde{\boldsymbol{z}})] \leq y \quad \forall \mathbb{P} \in \mathcal{F}$$

- Like classical robust optimization, the tractability of distributionally robust counterparts depend on the choice of ambiguity set, $\mathcal{F}$

# Distributionally Robust Counterpart

- Assume conditions of strong duality hold:

$$
\begin{array}{rll}
& \sup_{\mathbb{P}\in\mathcal{G}} & \mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x},\tilde{\boldsymbol{z}})] \\
= & \min & r + \boldsymbol{s}'\boldsymbol{\mu} \\
& \text{s.t.} & r + \boldsymbol{s}'\boldsymbol{u} \geq Q(\boldsymbol{x},\boldsymbol{z}) \quad \forall(\boldsymbol{z},\boldsymbol{u}) \in \mathcal{U} \\
& & \boldsymbol{s} \geq \boldsymbol{0}, r \in \mathbb{R}
\end{array}
$$

$$
\mathcal{G} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} (\tilde{\boldsymbol{z}},\tilde{\boldsymbol{u}}) \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{u}}] \leq \boldsymbol{\mu} \\ \mathbb{P}[(\tilde{\boldsymbol{z}},\tilde{\boldsymbol{u}}) \in \mathcal{U}] = 1 \end{array} \right\}
$$

$$
\mathcal{U} = \{(\boldsymbol{z},\boldsymbol{u}) \mid \boldsymbol{g}(\boldsymbol{z}) \leq \boldsymbol{u}, \boldsymbol{z} \in \mathcal{W}\}
$$

# Adaptive Robust Optimization

- Exact solution

$$
\begin{aligned}
Z^*_{DRO} = \quad & \min \quad \boldsymbol{c}'\boldsymbol{x} + \sup_{\mathbb{P}\in\mathcal{G}} \mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x},\tilde{\boldsymbol{z}})] \\
& \text{s.t.} \quad \boldsymbol{x} \in X_1 \\
= \quad & \min \quad \boldsymbol{c}'\boldsymbol{x} + r + \boldsymbol{s}'\boldsymbol{\mu} \\
& \text{s.t.} \quad r + \boldsymbol{s}'\boldsymbol{u} \geq (\boldsymbol{b}(\boldsymbol{z}) - \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x})'\boldsymbol{p}^i \quad \forall(\boldsymbol{z},\boldsymbol{u}) \in \mathcal{U}, \forall i \in V \\
& \qquad \boldsymbol{s} \geq \boldsymbol{0}, r \in \mathbb{R} \\
& \qquad \boldsymbol{x} \in X_1
\end{aligned}
$$

  - Transform DRO to a standard Robust Linear Optimization problem
  - Tractability depends on the number of extreme points of P, which can be exponential
  - Exact solutions can be derived for portfolio optimization and newsvendor problems.

# Another Perspective

$$\begin{aligned} Q(\boldsymbol{x},\boldsymbol{z}) = \quad & \min \quad \boldsymbol{d}'\boldsymbol{y} \\ & \text{s.t.} \quad \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \geq \boldsymbol{b}(\boldsymbol{z}) \\ & \qquad\ \ \boldsymbol{y} \in \mathbb{R}^{N_2} \end{aligned}$$

$$\begin{aligned} Z^*_{DRO} = \quad & \min \quad \boldsymbol{c}'\boldsymbol{x} + \sup_{\mathbb{P}\in\mathcal{G}} \mathbb{E}_{\mathbb{P}}[Q(\boldsymbol{x},\tilde{\boldsymbol{z}})] \\ & \text{s.t.} \quad \boldsymbol{x} \in X_1 \\ = \quad & \min \quad \boldsymbol{c}'\boldsymbol{x} + \sup_{\mathbb{P}\in\mathcal{G}} \mathbb{E}_{\mathbb{P}}[\boldsymbol{d}'\boldsymbol{y}(\tilde{\boldsymbol{z}})] \\ & \text{s.t.} \quad \boldsymbol{A}(\boldsymbol{z}) + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{z}) \geq \boldsymbol{b}(\boldsymbol{z}) \quad \forall(\boldsymbol{z},\boldsymbol{u}) \in \mathcal{U} \\ & \qquad\ \ \boldsymbol{x} \in X_1 \\ & \qquad\ \ \boldsymbol{y} \text{ measurable function, } \boldsymbol{y} : \mathbb{R}^{I_1} \mapsto \mathbb{R}^{N_2} \end{aligned}$$

- Optimization over **recourse function**, $\boldsymbol{y}$ is generally not a tractable format due to potentially infinite number of decision variables.

# Tractable Recourse Adaptation

- Tractable recourse adaption by restricting the recourse function y.
- No recourse adaptation
  - Simplest function
    - Not responsive to uncertainties
    - Can be very conservative

$$\boldsymbol{y}(\boldsymbol{z}) = \boldsymbol{y}^0$$

# Affine Recourse Adaptation

- Easily extendable to multiperiod models to capture non-anticipativity
- Practical: Can be incorporated in optimization software
  - ROME, AIMMS, ROC,RSOME

$$\boldsymbol{y}(\boldsymbol{z}) = \boldsymbol{y}^0 + \sum_{k \in [I_1]} \boldsymbol{y}^k z_k$$

# Affine Recourse Adaptation

$$
\begin{aligned}
Z^*_{ARA} = \quad \min \quad & \boldsymbol{c}'\boldsymbol{x} + \sup_{\mathbb{P}\in\mathcal{G}} \mathbb{E}_{\mathbb{Q}}[\boldsymbol{d}'\boldsymbol{y}(\tilde{\boldsymbol{z}})] \\
\text{s.t} \quad & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{z}) \geq \boldsymbol{b}(\boldsymbol{z}) \quad \forall \boldsymbol{z} \in \mathcal{W} \\
& \boldsymbol{x} \in X_1 \\
& \boldsymbol{y}(\boldsymbol{z}) \text{ Affine fn of } \boldsymbol{z}
\end{aligned}
$$

$$Z^*_{DRO} \leq Z^*_{ARA}$$

# Affine Recourse Adaptation

$$
\begin{array}{rcll}
Z^*_{ARA} = & \min & \boldsymbol{c}'\boldsymbol{x} + r + \boldsymbol{s}'\boldsymbol{\mu} & \\
& \text{s.t} & r + \boldsymbol{s}'\boldsymbol{u} \geq \boldsymbol{d}'\boldsymbol{y}(\boldsymbol{z}) & \forall (\boldsymbol{z}, \boldsymbol{u}) \in \mathcal{U} \\
& & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{z}) \geq \boldsymbol{b}(\boldsymbol{z}) & \forall \boldsymbol{z} \in \mathcal{W} \\
& & \boldsymbol{x} \in X_1 & \\
& & \boldsymbol{s} \geq \boldsymbol{0}, r \in \mathbb{R} & \\
& & \boldsymbol{y}(\boldsymbol{z}) \text{ Affine fn of } \boldsymbol{z} &
\end{array}
$$

# On Affine Recourse Adaptation

- May lead to infeasible solution even when the problem has complete recourse

$$\mathcal{F} = \left\{ \mathbb{P} \mid \tilde{z} \sim \mathbb{P}, \mathbb{E}_{\mathbb{P}}[\tilde{z}] = 0, \mathbb{E}_{\mathbb{P}}[\tilde{z}^2] \leq 1 \right\}$$

$$\sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[|\tilde{z}|] = 1$$

$$\begin{array}{llll} Z^*_{ARA} = & \min & \sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[y(\tilde{z})] & \\ & \text{s.t} & y(z) \geq z & \forall z \in \mathbb{R} \\ & & y(z) \geq -z & \forall z \in \mathbb{R} \\ & & y(z) := y^0 + y^1 z & \end{array}$$

# On Affine Recourse Adaptation

- Under ARA
  - Problem becomes infeasible

$$\underbrace{y^0 + y^1 z \geq z \quad \forall z \in \Re}_{y^1 = 1}$$

$$\underbrace{y^0 + y^1 z \geq -z \quad \forall z \in \Re}_{y^1 = -1}$$

  - ARA can be infeasible even if the problem has complete recourse!!

# Piecewise Affine Recourse Adaptation

- Piecewise ARA that improves upon ARA and while keeping the model tractable
  - Chen, Sim, Sun and Zhang (2006) and Goh and Sim (2010)
  - Implemented in software ROME, Robust Optimzation Made Easy by Goh and Sim.

# ARA with Lifted Random Variables

- Incorporate lifted/auxiliary random variables of Lifted Ambiguity Set in ARA

$$\mathcal{G} = \left\{ \mathbb{P} \;\middle|\; \begin{array}{l} (\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{u}}) \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{u}}] \leq \boldsymbol{\mu} \\ \mathbb{P}[(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{u}}) \in \mathcal{U}] = 1 \end{array} \right\}$$

$$\boldsymbol{y}(\boldsymbol{z}, \boldsymbol{u}) = \boldsymbol{y}^0 + \sum_{k \in [I_1]} \boldsymbol{y}_1^k z_k + \sum_{k \in [I_2]} \boldsymbol{y}_2^k u_k$$

# Lifted ARA

$$
\begin{aligned}
Z^*_{LARA} = \quad & \min \quad \boldsymbol{c}'\boldsymbol{x} + \sup_{\mathbb{P}\in\mathcal{G}} \mathbb{E}_{\mathbb{P}}[\boldsymbol{d}'\boldsymbol{y}(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{u}})] \\
& \text{s.t} \quad \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{z}, \boldsymbol{u}) \geq \boldsymbol{b}(\boldsymbol{z}) \quad \forall (\boldsymbol{z}, \boldsymbol{u}) \in \mathcal{U} \\
& \qquad\ \ \boldsymbol{x} \in X_1 \\
& \qquad\ \ \boldsymbol{y}(\boldsymbol{z}, \boldsymbol{u}) \text{ Affine fn of } (\boldsymbol{z}, \boldsymbol{u})
\end{aligned}
$$

# Lifted ARA

$$
\begin{array}{rll}
Z^*_{LARA} = & \min & \boldsymbol{c}'\boldsymbol{x} + r + \boldsymbol{s}'\boldsymbol{\mu} \\
& \text{s.t} & r + \boldsymbol{s}'\boldsymbol{u} \geq \boldsymbol{d}'\boldsymbol{y}(\boldsymbol{z},\boldsymbol{u}) \qquad \forall(\boldsymbol{z},\boldsymbol{u}) \in \mathcal{U} \\
& & \boldsymbol{A}(\boldsymbol{z})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{z},\boldsymbol{u}) \geq \boldsymbol{b}(\boldsymbol{z}) \quad \forall(\boldsymbol{z},\boldsymbol{u}) \in \mathcal{U} \\
& & \boldsymbol{x} \in X_1 \\
& & \boldsymbol{s} \geq \boldsymbol{0}, r \in \mathbb{R} \\
& & \boldsymbol{y}(\boldsymbol{z},\boldsymbol{u}) \text{ Affine fn of } (\boldsymbol{z},\boldsymbol{u})
\end{array}
$$

# ARA with Lifted Random Variables

- Lifted ARA achieves tighter bound than ARA

$$Z^*_{DRO} \leq Z^*_{LARA} \leq Z^*_{ARA}$$

- ARA bound is finite for complete recourse problems
- Improve upon deflected ARA.
- Recover exact solution if complete recourse and one-dimensional recourse function
  - Exact result for Newsvendor problem