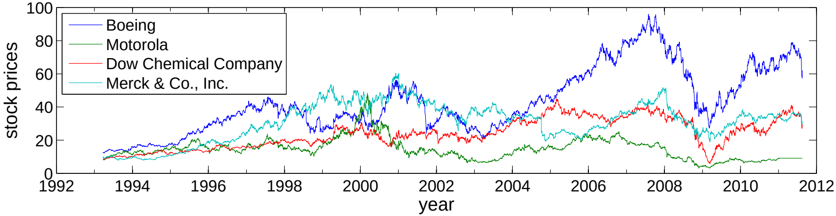
Addressing Model Ambiguity in the Expected Utility Framework

Erick Delage
CRC in Decision Making under Uncertainty
Associate professor, Dept. of Decision Sciences
HEC Montréal

Joint work with Yinyu Ye (Stanford University), Sharon Arroyo (Boeing Cie.), Benjamin Armbruster (Northwestern University)

Thursday, November 6th, 2014

The Portfolio Selection Problem



- An individual meets with his financial advisor to tell him he wishes to invest in a given industrial sector, country, etc.
- Since uncertain factors affect performance, a good portfolio is one where the risks of losses are best justified by the potential gains
- How can one trade-off optimally the risks and the returns taking into account his own perception of what is a serious risk?

Von Neumann-Morgenstern Expected Utility

If the investor agrees with the following axioms:

1. Completeness : He can order any two gambles
2. Transitivity : $\mathcal{H}_1 \succeq \mathcal{H}_2 \succeq \mathcal{H}_3 \ \Rightarrow \ \mathcal{H}_1 \succeq \mathcal{H}_3$
3. Continuity : If $\mathcal{H}_1 \succeq \mathcal{H}_2 \succeq \mathcal{H}_3$ then there is a p such that $\mathcal{H}_2 \sim p\mathcal{H}_1 + (1-p)\mathcal{H}_3$
4. Independence : If $\mathcal{H}_1 \succeq \mathcal{H}_2$ then $p\mathcal{H}_1 + (1-p)\mathcal{H}_3 \succeq p\mathcal{H}_2 + (1-p)\mathcal{H}_3$ for all p and $\mathcal{H}_3$

then the preference he expresses between any two gambles must be representable by an expected utility measure:

$$\mathcal{H}_1 \succeq \mathcal{H}_2 \ \ \Leftrightarrow \ \ E[u(\mathcal{H}_1)] \geq E[u(\mathcal{H}_2)]$$

Expected Utility Framework

When applying the expected utility framework to a decision problem:

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \quad \mathbb{E}\left[u(\mathsf{h}(\mathbf{x},\boldsymbol{\xi}))\right] ,$$

where $\mathbf{x}$ = decisions, $\boldsymbol{\xi}$ = uncertain parameters, $h(\mathbf{x},\boldsymbol{\xi})$ = profit,

Expected Utility Framework

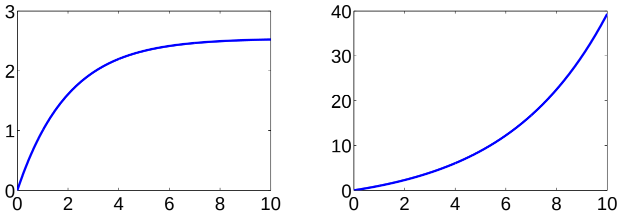
When applying the expected utility framework to a decision problem:

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \quad \mathbb{E}\left[u(\mathsf{h}(\mathbf{x},\boldsymbol{\xi}))\right] ,$$

where $\mathbf{x}$ = decisions, $\boldsymbol{\xi}$ = uncertain parameters, $h(\mathbf{x},\boldsymbol{\xi})$ = profit,

it is assumed that we know:

- The distribution of the random vector $\boldsymbol{\xi}$
- A utility function that matches investor's attitude to risk



Retirement Fund   Vegas Budget

Difficulties encountered in practice

Difficulties of developing an accurate probabilistic model:

- Need to collect enough observations
- Need to consult with experts of the field of practice
- Need to make simplifying assumptions
- Unforeseen events (e.g., economic crisis) might occur

Difficulties encountered in practice

Difficulties of developing an accurate probabilistic model:

- Need to collect enough observations
- Need to consult with experts of the field of practice
- Need to make simplifying assumptions
- Unforeseen events (e.g., economic crisis) might occur

Difficulties of developing an accurate utility function:

- Need to compare a large number of gambles
- Need to accept structural properties
- Perception might be biased

Robust Optimization Framework

Generally attributed to Ben-Tal & Nemirovski (1998), this framework implements a worst-case approach to dealing with model ambiguity.

$$\max_{\mathbf{x}\in\mathcal{X}}.\ \{h(\mathbf{x},\mathbf{y})\}_{\mathbf{y}\in\mathcal{Y}} \quad \rightarrow \quad \max_{\mathbf{x}\in\mathcal{X}}.\ \inf_{\mathbf{y}\in\mathcal{Y}} h(\mathbf{x},\mathbf{y})$$

Robust Optimization Framework

Generally attributed to Ben-Tal & Nemirovski (1998), this framework implements a worst-case approach to dealing with model ambiguity.

$$\max_{\mathbf{x}\in\mathcal{X}} \; \{h(\mathbf{x},\mathbf{y})\}_{\mathbf{y}\in\mathcal{Y}} \quad \rightarrow \quad \max_{\mathbf{x}\in\mathcal{X}} \; \inf_{\mathbf{y}\in\mathcal{Y}} h(\mathbf{x},\mathbf{y})$$

Success of the robust optimization relies in part on applying duality to combine inner and outer problems

$$\begin{aligned}
\max_{\mathbf{x}\in\mathcal{X}} \; \min_{\mathbf{y}:\mathbf{A}\mathbf{y}\leq\mathbf{b}} \; \mathbf{y}^T\mathbf{x} \;&\equiv\; \max_{\mathbf{x}\in\mathcal{X}} \; \min_{\mathbf{y}} \; \max_{\lambda\geq 0} \; \mathbf{y}^T\mathbf{x} + \lambda^T(\mathbf{A}\mathbf{y}-\mathbf{b}) \\
&\equiv\; \max_{\mathbf{x}\in\mathcal{X}} \; \max_{\lambda\geq 0} \; \min_{\mathbf{y}} \; \mathbf{y}^T\mathbf{x} + \lambda^T(\mathbf{A}\mathbf{y}-\mathbf{b}) \\
&\equiv\; \max_{\mathbf{x}\in\mathcal{X},\lambda\geq 0} \; -\mathbf{b}^T\lambda \;\; \text{s.t.} \;\; \mathbf{x} + \mathbf{A}^T\lambda = 0
\end{aligned}$$

Outline

Outline

Distributionally Robust Optimization

- Let's consider that the choice of F is ambiguous

Distributionally Robust Optimization

- Let's consider that the choice of F is ambiguous
- Use available information to define $\mathcal{D}$, such that $F \in \mathcal{D}$

Distributionally Robust Optimization

- Let's consider that the choice of F is ambiguous
- Use available information to define $\mathcal{D}$, such that $F \in \mathcal{D}$
- Distributionally Robust Optimization values the lowest performing possible stochastic model

$$(DRO) \qquad \underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \; \inf_{F\in\mathcal{D}} \mathbb{E}_F[u(\mathrm{h}(\mathbf{x},\boldsymbol{\xi}))]$$

Distributionally Robust Optimization

- Let's consider that the choice of F is ambiguous
- Use available information to define $\mathcal{D}$, such that $F \in \mathcal{D}$
- Distributionally Robust Optimization values the lowest performing possible stochastic model

$$(DRO) \qquad \underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \ \inf_{F\in\mathcal{D}} \mathbb{E}_F[u(\mathrm{h}(\mathbf{x},\boldsymbol{\xi}))]$$

- 1958: H. Scarf introduces DRO
- 2007: I. Popescu solves μ-Σ portfolio prob.
- 2010: Bertsimas et al. solves linear μ-Σ prob.
- 2010: Goh et al. approx. linear $\mathcal{S}$-μ-Σ prob.
- 2010: Delage et al. solves concave-convexe $\mathcal{S}$-$\tilde{\mu}$-$\tilde{\Sigma}_{\max}$ problems
- 2014: Wiesemann et al. solves or approximates conic prob.

Three Assumptions

Let's make three assumptions about $\mathbb{E}[u(\mathrm{h}(\mathbf{x},\boldsymbol{\xi}))]$.

1. The profit function is concave in $\mathbf{x}$ and convex in $\boldsymbol{\xi}$
   - In portfolio optimization, $h(\mathbf{x},\boldsymbol{\xi}) = \boldsymbol{\xi}^T\mathbf{x}$
2. The utility function is piecewise linear concave :

$$u(y) = \min_{1\leq k\leq K} \quad a_k y + b_k \quad ,$$

3. The information about F is captured by

$$\mathcal{D}(\gamma) = \left\{ F \left| \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \mathcal{S}) = 1 \\ \|\mathbb{E}[\boldsymbol{\xi}] - \hat{\mu}\|^2_{\hat{\Sigma}^{-1/2}} \leq \gamma_1 \\ \mathbb{E}[(\boldsymbol{\xi} - \hat{\boldsymbol{\mu}})(\boldsymbol{\xi} - \hat{\boldsymbol{\mu}})^{\mathsf{T}}] \preceq (1+\gamma_2)\hat{\Sigma} \end{array} \right. \right\}$$

The DRO is a tractable

- The DRO problem with $\mathcal{D}(\gamma)$ is equivalent to

$$\begin{aligned} \max_{\mathbf{x},\,\mathbf{Q},\mathbf{q},r}. \quad & r - \left(\gamma_2 \hat{\Sigma} + \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\mathsf{T}\right) \bullet \mathbf{Q} - \hat{\boldsymbol{\mu}}^\mathsf{T}\mathbf{q} - \sqrt{\gamma_1}\,\|\hat{\Sigma}^{1/2}(\mathbf{q} + 2\mathbf{Q}\hat{\boldsymbol{\mu}})\| \\ \text{s.t.} \quad & r \leq u(\mathrm{h}(\mathbf{x},\boldsymbol{\xi})) + \boldsymbol{\xi}^\mathsf{T}\mathbf{q} + \boldsymbol{\xi}^\mathsf{T}\mathbf{Q}\boldsymbol{\xi}\ \forall\,\boldsymbol{\xi} \in \mathcal{S} \\ & \mathbf{Q} \succeq 0 \end{aligned}$$

-

The DRO is a tractable

- The DRO problem with $\mathcal{D}(\gamma)$ is equivalent to

$$\begin{aligned}
\max_{\mathbf{x},\,\mathbf{Q},\mathbf{q},r} \quad & r - \left(\gamma_2\hat{\Sigma} + \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\mathsf{T}\right) \bullet \mathbf{Q} - \hat{\boldsymbol{\mu}}^\mathsf{T}\mathbf{q} - \sqrt{\gamma_1}\, \|\hat{\Sigma}^{1/2}(\mathbf{q} + 2\mathbf{Q}\hat{\boldsymbol{\mu}})\| \\
\text{s.t.} \quad & r \leq a_k(\mathrm{h}(\mathbf{x},\boldsymbol{\xi})) + b_k + \boldsymbol{\xi}^\mathsf{T}\mathbf{q} + \boldsymbol{\xi}^\mathsf{T}\mathbf{Q}\boldsymbol{\xi} \;\forall\, \boldsymbol{\xi} \in \mathcal{S}, \forall\, k \quad (\star) \\
& \mathbf{Q} \succeq 0
\end{aligned}$$

The DRO is a tractable

- The DRO problem with $\mathcal{D}(\gamma)$ is equivalent to

$$\begin{aligned} \max_{\mathbf{x},\,\mathbf{Q},\mathbf{q},r}. \quad & r - \left(\gamma_2 \hat{\Sigma} + \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^{\mathsf{T}}\right) \bullet \mathbf{Q} - \hat{\boldsymbol{\mu}}^{\mathsf{T}}\mathbf{q} - \sqrt{\gamma_1}\,\|\hat{\Sigma}^{1/2}(\mathbf{q} + 2\mathbf{Q}\hat{\boldsymbol{\mu}})\| \\ \text{s.t.} \quad & r \leq a_k(\mathrm{h}(\mathbf{x}, \boldsymbol{\xi})) + b_k + \boldsymbol{\xi}^{\mathsf{T}}\mathbf{q} + \boldsymbol{\xi}^{\mathsf{T}}\mathbf{Q}\boldsymbol{\xi} \;\forall\, \boldsymbol{\xi} \in \mathcal{S}, \forall\, k \quad (\star) \\ & \mathbf{Q} \succeq 0 \end{aligned}$$

- For portfolio selection, if $\mathcal{S}$ = polygon or ellipsoid, then DRO equivalent to semi-definite program.
  E.g., when $\mathcal{S} = \mathbb{R}^m$, constraint $(\star)$ can be replaced by

$$\begin{bmatrix} \mathbf{Q} & (\mathbf{q} + a_k\mathbf{x})/2 \\ (\mathbf{q} + a_k\mathbf{x})^{\mathsf{T}}/2 & b_k - r \end{bmatrix} \succeq 0 \,,\ \forall\, k$$

Distributionally Robust Portfolio Optimization

Let's consider the case of portfolio optimization:

$$\max_{\mathbf{x}\in\mathcal{X}} . \ \min_{F\in\mathcal{D}} \mathbb{E}_F[u(\boldsymbol{\xi}^\mathsf{T}\mathbf{x})] \ ,$$

where x_i is how much is invested in stock i with future return ξ_i.

Distributionally Robust Portfolio Optimization

Let's consider the case of portfolio optimization:

$$\max_{\mathbf{x}\in\mathcal{X}}. \ \min_{F\in\mathcal{D}} \mathbb{E}_F[u(\boldsymbol{\xi}^\top \mathbf{x})] \ ,$$

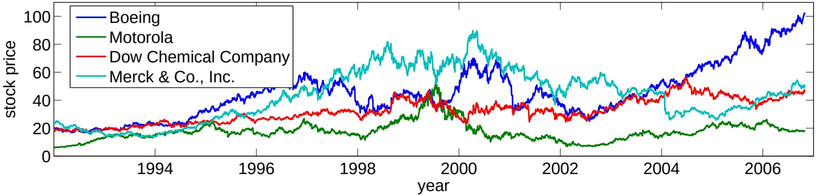
where x_i is how much is invested in stock i with future return ξ_i.

Does the robust solution perform better than solution of expected utility problem with fixed $\hat{F}$?

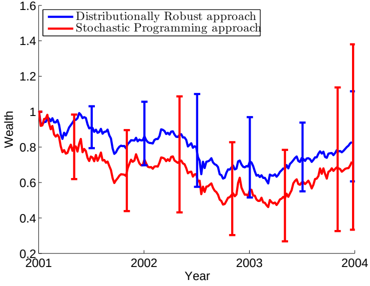
$$\mathcal{D} = \mathcal{D}(\gamma) \qquad \text{vs.} \qquad \mathcal{D} = \{\hat{F}\}$$

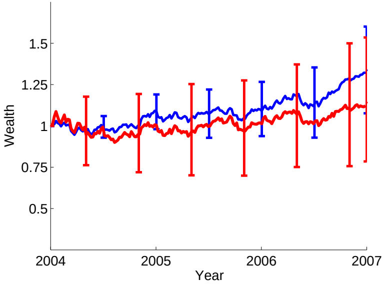
Experiments in Portfolio Optimization

30 stocks tracked over years 1992-2007 using Yahoo! Finance



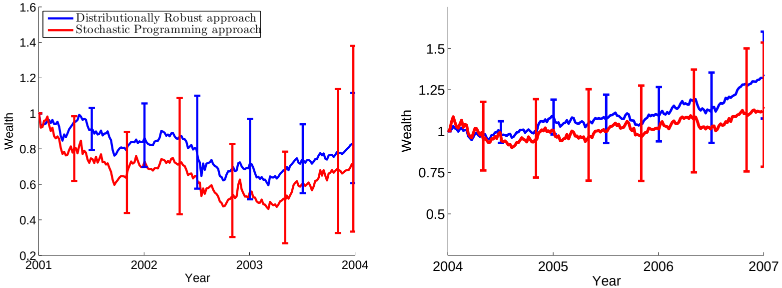
Wealth Evolution for 300 Experiments





- 10% and 90% percentiles are indicated periodically.

Wealth Evolution for 300 Experiments



- 10% and 90% percentiles are indicated periodically.
- 79% of time, the DRO outperformed the stoch. prog.
- 67% improvement on average using DRO with $\mathcal{D}(\gamma)$

Other applications

- Territory partitioning for multi-vehicle routing problem with J. G. Carlsson ▸ Details
- Fleet mix optimization problem with S. Arroyo and Y. Ye ▸ Details
- Quadratic knapsack problem with J. Cheng and A. Lisser
- Multi-item newsvendor problem with A. Ardestani-Jaafari

Outline

How does One Assesses Risk Tolerance ?

Here are two questions from a survey proposed by [Grable & Lytton, Financial Services Review (1999)].

How does One Assesses Risk Tolerance ?

Here are two questions from a survey proposed by [Grable & Lytton, Financial Services Review (1999)].

- You have just finished saving for a "once-in-a-lifetime" vacation. Three weeks before you plan to leave, you lose your job. You would:
  1. Cancel the vacation
  2. Take a much more modest vacation
  3. Go as scheduled, reasoning that you need the time to prepare for a job search
  4. Extend your vacation, because this might be your last chance to go first-class

How does One Assesses Risk Tolerance ?

Here are two questions from a survey proposed by [Grable & Lytton, Financial Services Review (1999)].

- You have just finished saving for a "once-in-a-lifetime" vacation. Three weeks before you plan to leave, you lose your job. You would:
  1. Cancel the vacation
  2. Take a much more modest vacation
  3. Go as scheduled, reasoning that you need the time to prepare for a job search
  4. Extend your vacation, because this might be your last chance to go first-class
- In addition to whatever you own, you have been given $2000. You are now asked to choose between:
  1. A sure loss of $500
  2. A 50% chance to lose $1,000 and a 50% chance to lose nothing

Common Utility Function Estimation Techniques I

Parametric approach:

- One assumes that the function has a specific parametric form
  - Negative exponential utility (CARA)
  - Power utility (CRRA)
  - HARA (incr/decreasing absolute/relative risk aversion)

$$u(y) = \frac{1-\eta}{\eta}\left(\frac{ay}{1-\eta} + b\right)^{\eta}$$

- Ask enough questions to identify the parameters

Common Utility Function Estimation Techniques I

Parametric approach:

- One assumes that the function has a specific parametric form
  - Negative exponential utility (CARA)
  - Power utility (CRRA)
  - HARA (incr/decreasing absolute/relative risk aversion)

$$u(y) = \frac{1-\eta}{\eta}\left(\frac{ay}{1-\eta} + b\right)^{\eta}$$

- Ask enough questions to identify the parameters

- Why it might fail ?
  - Investor must commit to global structure
  - Reduces investor's personal ambitions & perception to a few parameters

Common Utility Function Estimation Techniques II

Non-parametric approach:

- Identify the certainty equivalents of a list of lotteries. By answering:
  - What is the smallest amount c_i of money you would take instead of playing a lottery $\mathcal{L}_i$?
- Find a piecewise linear utility function that satisfies these certainty equivalents

$$E[u(\mathcal{L}_i)] = c_i \;\; \forall i$$

Common Utility Function Estimation Techniques II

Non-parametric approach:

- Identify the certainty equivalents of a list of lotteries. By answering:
  - What is the smallest amount c_i of money you would take instead of playing a lottery $\mathcal{L}_i$?
- Find a piecewise linear utility function that satisfies these certainty equivalents

$$E[u(\mathcal{L}_i)] = c_i \;\; \forall i$$

- Why it might fail ?
  - Which/How many questions should be asked ?
  - Expresses risk neutrality between breakpoints
  - Ignores how confident we are in the final choice of $u(\cdot)$

Applying the Robust Optimization Framework

Information can be used to characterize a set $\mathcal{U}$ of plausible utility functions. I.e., any $u(\cdot)$ such that:

- Global info:
  - Risk aversion : $u(\cdot)$ is concave
  - Prudence : $u(\cdot)$ is convex
  - S-shaped : $u(\cdot)$ convex-concave
- Local info :

$$E[u(\mathcal{W}_k)] \geq E[u(\mathcal{Y}_k)] \;\; \forall k$$

Applying the Robust Optimization Framework

Information can be used to characterize a set $\mathcal{U}$ of plausible utility functions. I.e., any $u(\cdot)$ such that:

- Global info:
  - Risk aversion : $u(\cdot)$ is concave
  - Prudence : $u(\cdot)$ is convex
  - S-shaped : $u(\cdot)$ convex-concave
- Local info :

$$E[u(\mathcal{W}_k)] \geq E[u(\mathcal{Y}_k)] \;\; \forall k$$

Unfortunately, a direct application of RO is meaningless

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \;\; \inf_{u\in\mathcal{U}} \; \mathbb{E}\left[u(\mathsf{h}(\mathbf{x},\xi))\right] \;\; = \;\; -\infty$$

Applying the Robust Optimization Framework

Information can be used to characterize a set $\mathcal{U}$ of plausible utility functions. I.e., any $u(\cdot)$ such that:

- Global info:
  - Risk aversion : $u(\cdot)$ is concave
  - Prudence : $u(\cdot)$ is convex
  - S-shaped : $u(\cdot)$ convex-concave
- Local info :

$$E[u(\mathcal{W}_k)] \geq E[u(\mathcal{Y}_k)] \;\; \forall k$$

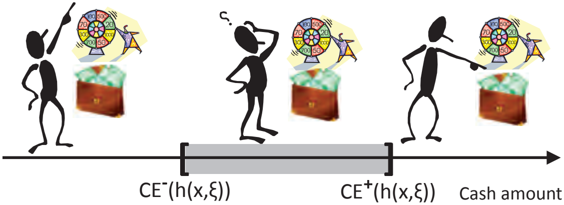
Unfortunately, a direct application of RO is meaningless

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \;\; \inf_{u\in\mathcal{U}} \; \mathbb{E}\left[u(\mathsf{h}(\mathbf{x},\boldsymbol{\xi}))\right] \;\; = \;\; -\infty$$

Even if we force $u(0) = 0$ $u(1) = 1$, model promotes risk neutrality.

Certainty Equivalent Ambiguity

- Given a decision $\mathbf{x}$, we can start by defining an interval $[CE^{-}(h(x,\xi)), CE^{+}(h(\mathbf{x},\xi))]$ of plausible minimum certain return that decision maker would prefer to random profit $h(\mathbf{x},\xi)$



Robust Certainty Equivalent Approach

Theorem

Identifying the decision that maximizes the lowest perceived CE

$$\max_{\mathbf{x}\in\mathcal{X}} . \quad \inf_{u\in\mathcal{U}} CE_u(h(x,\xi))$$

can be done efficiently.

Robust Certainty Equivalent Approach

Theorem

Identifying the decision that maximizes the lowest perceived CE

$$\max_{\mathbf{x}\in\mathcal{X}}. \quad \inf_{u\in\mathcal{U}} CE_u(h(x,\xi))$$

can be done efficiently.

Proof:

- Objective is quasiconcave and reduces to

$$\max_{\mathbf{x},t}. \quad t \quad \text{s.t.} \quad CE(h(\mathbf{x},\boldsymbol{\xi})) \geq t \;\; \forall\, u \in \mathcal{U}$$

or equiv. $\max_{t} \quad t \quad \text{s.t.} \quad \max_{\mathbf{x}\in\mathcal{X}} \inf_{u\in\mathcal{U}} E[u(h(\mathbf{x},\boldsymbol{\xi}))] - u(t) \geq 0$

- Infimum over $u \in \mathcal{U}$ can be reduced to finite dimensional program so duality can be applied ▸ details

Accounting for elicitation errors

- Since “to err is human”, we should account for mislabeling of the compared lotteries
- Hence, that for some rand perception noise ϵ, we have that

$$E[u(W_k)] + \epsilon_k \geq E[u(Y_k)]$$

Accounting for elicitation errors

- Since "to err is human", we should account for mislabeling of the compared lotteries
- Hence, that for some rand perception noise ϵ, we have that

$$E[u(W_k)] + \epsilon_k \geq E[u(Y_k)]$$

- In that case, one could consider $u(\cdot)$ plausible as long as $\exists \delta \geq 0$ such that $\sum_k \delta_k \leq \Gamma$ and that

$$E[u(W_k)] + \delta_k \geq E[u(Y_k)] \ \forall \ k$$

- This can easily be incorporated to the model

Numerical experiments

- Obtained from Yahoo! Finance historical stock returns for 350 companies from 1993 to 2011
- Ran extensive amount of trials using last 50 weekly returns to decide investment among 10 assets for next week
- In each experiment, the investor has an unknown risk averse utility function and compares up to 80 pairs of gambles

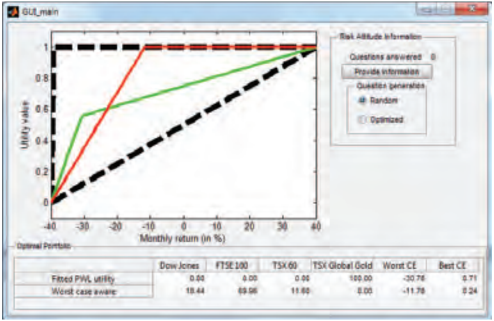
Experimental Results

- Using the wrong utility function can mislead the choice of portfolio
- The investor's unknown utility function satisfies $\frac{-u''(y)y}{u'(y)} = \frac{20}{y}$

Function used	Portfolio's true certainty equivalent (in perc. point)		
	5 questions	20 questions	80 questions
Exponential	-0.05	-0.12	**-0.13**
Fitted PWL	**-0.60**	-0.11	0.05
Worst-case	-0.14	-0.08	0.06
Worst-case prudent	-0.13	-0.05	0.08
True	0.12	0.12	0.12

A Tool for Interacting with Investors

- Framework can be used for active learning of investors risk attitude



Outline

Conclusion & Future Work

- There is some wisdom in accounting for ambiguity about the expected utility model
  - Disregarding this knowledge might be misleading

Conclusion & Future Work

- There is some wisdom in accounting for ambiguity about the expected utility model
  - Disregarding this knowledge might be misleading
- Accounting for distribution ambiguity or ambiguity in risk preferences isn't computationally demanding
  - It remains to verify whether both ambiguities can be accounted for jointly

Conclusion & Future Work

- There is some wisdom in accounting for ambiguity about the expected utility model
  - Disregarding this knowledge might be misleading
- Accounting for distribution ambiguity or ambiguity in risk preferences isn't computationally demanding
  - It remains to verify whether both ambiguities can be accounted for jointly
- Studying the sensitivity of optimal solution with respect to modeled ambiguity can be helpful
  - Value of stochastic modeling
  - Guidance for risk tolerance assessment

Bibliography I

Ardestani-Jaafari, A., E. Delage. 2014. Robust optimization of a class of bi-convex functions with application to inventory problems. Draft.

Armbruster, B., E. Delage. 2014. Decision making under uncertainty when preference information is incomplete. *Management Science* .

Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research* **23**(4) 769–805.

Bertsimas, D., X. V. Doan, K. Natarajan, C. P. Teo. 2010. Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research* **35**(3) 580–602.

Carlsson, J. G., E. Delage. 2013. Robust partitioning for stochastic multi-vehicle routing. *Operations Research* **61**(3) 727–744.

Cheng, J., E. Delage, A. Lisser. 2014. Distributionally robust stochastic knapsack problem. *Journal on Optimization* **24**(3) 1485–1506.

Delage, E., S. Arroyo, Y. Ye. 2014. The value of stochastic modeling in two-stage stochastic programs with cost uncertainty. *Operations Research (forthcoming)* .

Bibliography II

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

Goh, J., M. Sim. 2010. Distributionally robust optimization and its tractable approximations. *Operations Research* **58** 902–917.

Grable, J., R. H. Lytton. 1999. Financial risk tolerance revisited: the development of a risk assessment instrument. *Financial Services Review* **8** 163–181.

Popescu, I. 2007. Robust mean-covariance solutions for stochastic optimization. *Operations Research* **55**(1) 98–112.

von Neumann, J., O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press.

Wiesemann, W., D. Kuhn, M. Sim. 2014. Distributionally robust convex optimization. *Operations Research* .

Questions & Comments ...

... Thank you!

Dealing with model ambiguity: Ellsberg Paradox

Consider an urn with 30 blue balls and 60 other balls that are either red or yellow (you don't know how many are red or yellow).

Dealing with model ambiguity: Ellsberg Paradox

Consider an urn with 30 blue balls and 60 other balls that are either red or yellow (you don't know how many are red or yellow).

Experiment 1: Choose among the following two gambles

- Gamble A: If you draw a blue ball, then you win 100$
- Gamble B: If you draw a red ball, then you win 100$

Dealing with model ambiguity: Ellsberg Paradox

Consider an urn with 30 blue balls and 60 other balls that are either red or yellow (you don't know how many are red or yellow).

Experiment 1: Choose among the following two gambles

- Gamble A: If you draw a blue ball, then you win 100$
- Gamble B: If you draw a red ball, then you win 100$

Experiment 2: Choose among the following two gambles

- Gamble C: If you draw blue or yellow ball, then you win 100$
- Gamble D: If you draw red or yellow ball, then you win 100$

Dealing with model ambiguity: Ellsberg Paradox

Consider an urn with 30 blue balls and 60 other balls that are either red or yellow (you don't know how many are red or yellow).

Experiment 1: Choose among the following two gambles

- Gamble A: If you draw a blue ball, then you win 100$
- Gamble B: If you draw a red ball, then you win 100$

Experiment 2: Choose among the following two gambles

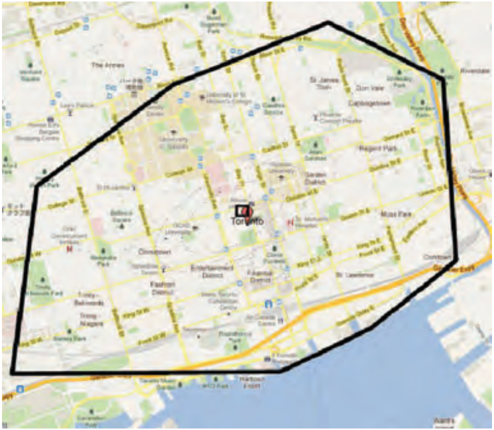
- Gamble C: If you draw blue or yellow ball, then you win 100$
- Gamble D: If you draw red or yellow ball, then you win 100$

If you clearly prefer Gamble A & D, then you are averse to model ambiguity

Outline

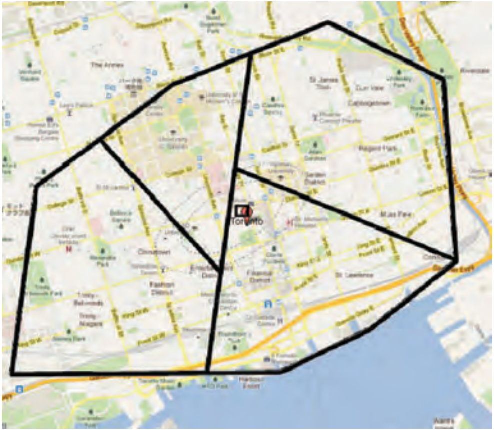
Multi-Vehicle Routing on a Planar Region

- Divide a planar region into K subregions, each serviced by a different vehicle, so that the total workload be most evenly distributed among the fleet



Multi-Vehicle Routing on a Planar Region

- Divide a planar region into K subregions, each serviced by a different vehicle, so that the total workload be most evenly distributed among the fleet



Multi-Vehicle Routing on a Planar Region

- Divide a planar region into K subregions, each serviced by a different vehicle, so that the total workload be most evenly distributed among the fleet

Multi-Vehicle Routing on a Planar Region

- Divide a planar region into K subregions, each serviced by a different vehicle, so that the total workload be most evenly distributed among the fleet

Distributionally Robust Partitioning

- Given $\mathcal{D}$, we partition so that the largest workload over the worst distribution of demand points is as small as possible

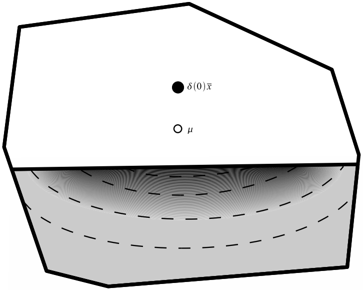
$$\min_{\{\mathcal{R}_1,\mathcal{R}_2,\ldots,\mathcal{R}_K\}} \sup_{F\in\mathcal{D}} \left\{ \max_i \mathbb{E}[TSP(\{\xi_1,\xi_2,\ldots,\xi_N\}\cap\mathcal{R}_i)] \right\} ,$$

Distributionally Robust Partitioning

- Given $\mathcal{D}$, we partition so that the largest workload over the worst distribution of demand points is as small as possible

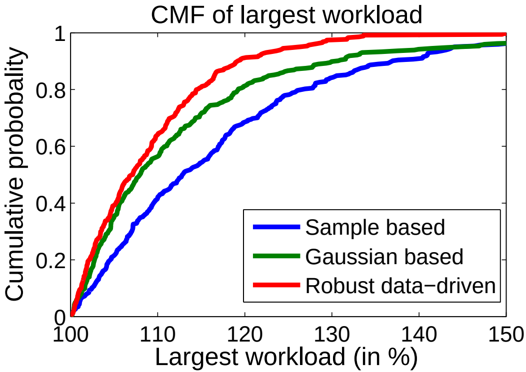
$$\min_{\{\mathcal{R}_1,\mathcal{R}_2,\ldots,\mathcal{R}_K\}} \sup_{F\in\mathcal{D}} \left\{ \max_i \mathbb{E}[TSP(\{\xi_1,\xi_2,\ldots,\xi_N\}\cap\mathcal{R}_i)] \right\},$$

- A side product is to characterize for any partition what is a worst-case distribution of demand locations



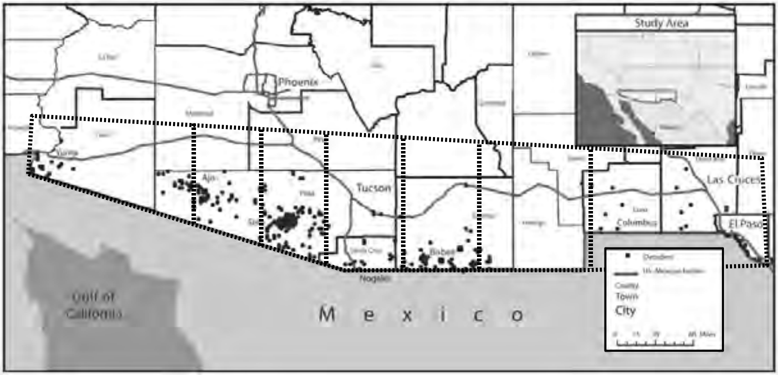
Distributionally Robust Partitioning

We simulated three partition schemes on a set of randomly generated parcel delivery problems where the territory needed to be divided into two regions and the demand is drawn from a mixture of truncated Gaussian distribution



Border Patrol Workload Partitioning

Robust partitions of the USA-Mexico border obtained using our branch & bound algorithm.



▸ Back to DRO applications

Outline

The Robustness of the Deterministic Solution

If we are risk neutral we might not even need distribution information

Theorem

The solution of

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \quad \mathbb{E}[h(\mathbf{x},\mu)]$$

is optimal with respect to

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \quad \inf_{F\in\mathcal{D}(\mu,\Psi)} \quad \mathbb{E}_F[h(\mathbf{x},\xi)] \ ,$$

for any set of convex functions Ψ *with*

$$\mathcal{D}(\mu,\Psi) = \left\{ F \middle| \begin{array}{l} \mathbb{E}[\xi] = \mu \\ \mathbb{E}[\psi(\xi)] \leq 0 \ , \ \forall\, \psi \in \Psi \end{array} \right\} .$$

The Value of Stochastic Modeling

Consider the situation:

1. We know of a set $\mathcal{D}$ such that $F \in \mathcal{D}$
2. We have a candidate solution $\mathbf{x}_1$ in mind
3. Is it worth developing a stochastic model: $\mathcal{D} \rightarrow F$?
   (a) If yes, then develop a model & solve it
   (b) Otherwise, implement $\mathbf{x}_1$

The Value of Stochastic Modeling

Consider the situation:

1. We know of a set $\mathcal{D}$ such that $F \in \mathcal{D}$
2. We have a candidate solution $\mathbf{x}_1$ in mind
3. Is it worth developing a stochastic model: $\mathcal{D} \rightarrow F$?
   (a) If yes, then develop a model & solve it
   (b) Otherwise, implement $\mathbf{x}_1$

The Value of Stochastic Modeling ($\mathcal{VSM}$) gives an optimistic estimate of the value of obtaining perfect information about F.

$$\mathcal{VSM}(\mathbf{x}_1) \;:=\; \sup_{F \in \mathcal{D}} \left\{ \max_{\mathbf{x}_2} \mathbb{E}_F[\mathrm{h}(\mathbf{x}_2, \boldsymbol{\xi})] - \mathbb{E}_F[\mathrm{h}(\mathbf{x}_1, \boldsymbol{\xi})] \right\}$$

The Value of Stochastic Modeling

Consider the situation:

1. We know of a set $\mathcal{D}$ such that $F \in \mathcal{D}$
2. We have a candidate solution $\mathbf{x}_1$ in mind
3. Is it worth developing a stochastic model: $\mathcal{D} \to F$?
   (a) If yes, then develop a model & solve it
   (b) Otherwise, implement $\mathbf{x}_1$

The Value of Stochastic Modeling ($\mathcal{VSM}$) gives an optimistic estimate of the value of obtaining perfect information about F.

$$\mathcal{VSM}(\mathbf{x}_1) \;:=\; \sup_{F \in \mathcal{D}} \left\{ \max_{\mathbf{x}_2} \mathbb{E}_F[\mathrm{h}(\mathbf{x}_2, \boldsymbol{\xi})] - \mathbb{E}_F[\mathrm{h}(\mathbf{x}_1, \boldsymbol{\xi})] \right\}$$

Theorem

Unfortunately, evaluating $\mathcal{VSM}(\mathbf{x}_1)$ exactly is NP-hard in general.

Bounding the Value of Stochastic Modeling

Theorem

If $\mathcal{S} \subseteq \{\boldsymbol{\xi} \,|\, \|\boldsymbol{\xi}\|_1 \leq \rho\}$, an upper bound can be evaluated in $O(d^{3.5} + d\, T_{DCP})$ using:

$$\begin{aligned} \mathcal{UB}(\mathbf{x}_1, \bar{\mathbf{y}}_1) \;:=\; \min_{s,\mathbf{q}} \quad & s + \boldsymbol{\mu}^{\mathsf{T}}\mathbf{q} \\ \text{s.t.} \quad & s \geq \alpha(\rho\mathbf{e}_i) - \rho\mathbf{e}_i^{\mathsf{T}}\mathbf{q} \,,\ \forall\, i \in \{1, ..., d\} \\ & s \geq \alpha(-\rho\mathbf{e}_i) + \rho\mathbf{e}_i^{\mathsf{T}}\mathbf{q} \,,\ \forall\, i \in \{1, ..., d\}\,, \end{aligned}$$

where $\alpha(\boldsymbol{\xi}) = \max_{\mathbf{x}_2} h(\mathbf{x}_2, \boldsymbol{\xi}) - h(\mathbf{x}_1, \boldsymbol{\xi}; \bar{\mathbf{y}}_1)$.

Bounding the Value of Stochastic Modeling

Theorem

If $\mathcal{S} \subseteq \{\boldsymbol{\xi} \,|\, \|\boldsymbol{\xi}\|_1 \leq \rho\}$, an upper bound can be evaluated in $O(d^{3.5} + d\, T_{DCP})$ using:

$$\begin{aligned} \mathcal{UB}(\mathbf{x}_1, \bar{\mathbf{y}}_1) \;:=\; \min_{s,\mathbf{q}} \quad & s + \boldsymbol{\mu}^{\mathsf{T}}\mathbf{q} \\ \text{s.t.} \quad & s \geq \alpha(\rho \mathbf{e}_i) - \rho \mathbf{e}_i^{\mathsf{T}}\mathbf{q} \,,\, \forall\, i \in \{1, ..., d\} \\ & s \geq \alpha(-\rho \mathbf{e}_i) + \rho \mathbf{e}_i^{\mathsf{T}}\mathbf{q} \,,\, \forall\, i \in \{1, ..., d\}\,, \end{aligned}$$

where $\alpha(\boldsymbol{\xi}) = \max_{\mathbf{x}_2} h(\mathbf{x}_2, \boldsymbol{\xi}) - h(\mathbf{x}_1, \boldsymbol{\xi}; \bar{\mathbf{y}}_1)$.

- $\mathcal{UB}$ only uses information about $\boldsymbol{\mu}$ and $\mathcal{S}$
- $\mathcal{UB}$ simplifies the structure of $\mathcal{S}$
- $\mathcal{UB}$ assumes the candidate decision $\mathbf{y}_1$ cannot adapt to $\boldsymbol{\xi}$

Are Airlines Adventurous in their Fleet Acquisition?

- Fleet composition is a difficult decision problem:
  - Fleet contracts are signed 10 to 20 years ahead of schedule.
  - Many factors are still unknown at that time:
    e.g., passenger demand, fuel prices, etc.

Are Airlines Adventurous in their Fleet Acquisition?

- Fleet composition is a difficult decision problem:
  - Fleet contracts are signed 10 to 20 years ahead of schedule.
  - Many factors are still unknown at that time:
    e.g., passenger demand, fuel prices, etc.
- Yet, most airline companies sign these contracts based on a single scenario of what the future may be.

Are Airlines Adventurous in their Fleet Acquisition?

- Fleet composition is a difficult decision problem:
  - Fleet contracts are signed 10 to 20 years ahead of schedule.
  - Many factors are still unknown at that time:
    e.g., passenger demand, fuel prices, etc.
- Yet, most airline companies sign these contracts based on a single scenario of what the future may be.
- Are airlines companies being neglecful?

Mathematical formulation for Fleet Mix Problem

The fleet composition problem is a stochastic mixed integer LP

$$\max_{\mathbf{x}}. \ \mathbb{E}\,[-\underbrace{\mathbf{o}^{\mathsf{T}}\mathbf{x}}_{\text{ownership cost}} + \underbrace{h(\mathbf{x}, \tilde{\mathbf{p}}, \tilde{\mathbf{c}}, \tilde{\mathbf{L}})}_{\text{future profits}}\,] \ ,$$

Fleet mix → $\mathbf{x}$

Mathematical formulation for Fleet Mix Problem

The fleet composition problem is a stochastic mixed integer LP

$$\max_{\mathbf{x}}.\ \mathbb{E}[-\underbrace{\mathbf{o}^T\mathbf{x}}_{\text{ownership cost}} + \underbrace{h(\mathbf{x},\tilde{\mathbf{p}},\tilde{\mathbf{c}},\tilde{\mathbf{L}})}_{\text{future profits}}]\ ,$$

Fleet mix → $\mathbf{x}$

with $h(\mathbf{x},\tilde{\mathbf{p}},\tilde{\mathbf{c}},\tilde{\mathbf{L}}) :=$

$$\max_{z\geq 0, y\geq 0, w} \sum_k \Big(\sum_i \overbrace{\tilde{p}_i^k w_i^k}^{\text{flight profit}} - \overbrace{\tilde{c}_k(z_k - x_k)^+}^{\text{rental cost}} + \overbrace{\tilde{L}_k(x_k - z_k)^+}^{\text{lease revenue}} \Big)$$

$$\text{s.t.}\quad w_i^k \in \{0,1\}\ ,\ \forall\, k, \forall\, i \quad \&\quad \sum_k w_i^k = 1\ ,\ \forall\, i \quad \Big\}\ \text{Cover}$$

$$y^k_{g\in\text{in}(v)} + \sum_{i\in\text{arr}(v)} w_i^k = y^k_{g\in\text{out}(v)} + \sum_{i\in\text{dep}(v)} w_i^k\ ,\ \forall\, k, \forall\, v \quad \Big\}\ \text{Balance}$$

$$z_k = \sum_{v\in\{v|\text{time}(v)=0\}} \big(y^k_{g\in\text{in}(v)} + \sum_{i\in\text{arr}(v)} w_i^k\big)\ ,\quad \forall k \quad \Big\}\ \text{Count}$$

Experiments in Fleet Mix Optimization

We experimented with three test cases :

1. 3 types of aircrafts, 84 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [4\%, 53\%]$
2. 4 types of aircrafts, 240 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [2\%, 20\%]$
3. 13 types of aircrafts, 535 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [2\%, 58\%]$

Experiments in Fleet Mix Optimization

We experimented with three test cases :

1. 3 types of aircrafts, 84 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [4\%, 53\%]$
2. 4 types of aircrafts, 240 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [2\%, 20\%]$
3. 13 types of aircrafts, 535 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [2\%, 58\%]$

Results:

Test cases	CPU Time		DRO sub-optimality	
	DRO	SP with $\hat{F}$	Under $\hat{F}$	$\forall\, F \in \mathcal{D}$
#1	0.6 s	3 min	0.001%	$< 6\%$
#2	1 s	14 min	0.001%	$< 1\%$
#3	5 s	21 h	0.003%	$< 7\%$

Experiments in Fleet Mix Optimization

We experimented with three test cases :

1. 3 types of aircrafts, 84 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [4\%, 53\%]$
2. 4 types of aircrafts, 240 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [2\%, 20\%]$
3. 13 types of aircrafts, 535 flights, $\sigma_{\tilde{p}_i}/\mu_{\tilde{p}_i} \in [2\%, 58\%]$

Results:

Test cases	CPU Time		DRO sub-optimality	
	DRO	SP with $\hat{F}$	Under $\hat{F}$	$\forall\, F \in \mathcal{D}$
#1	0.6 s	3 min	0.001%	$< 6\%$
#2	1 s	14 min	0.001%	$< 1\%$
#3	5 s	21 h	0.003%	$< 7\%$

Conclusions:

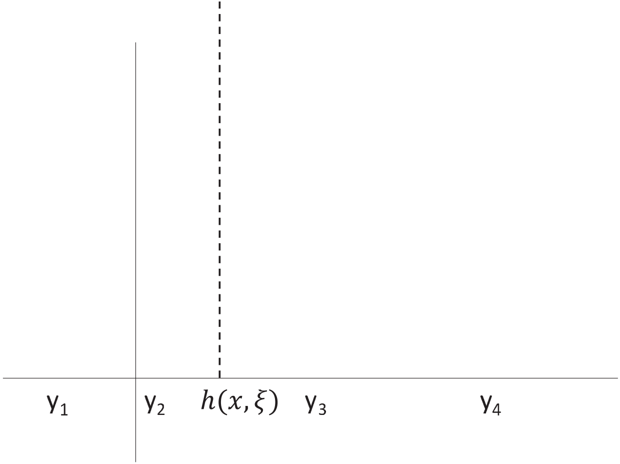
- It's wasteful to invest more than 7% of profits in extra info

Back to DRO applications

Outline

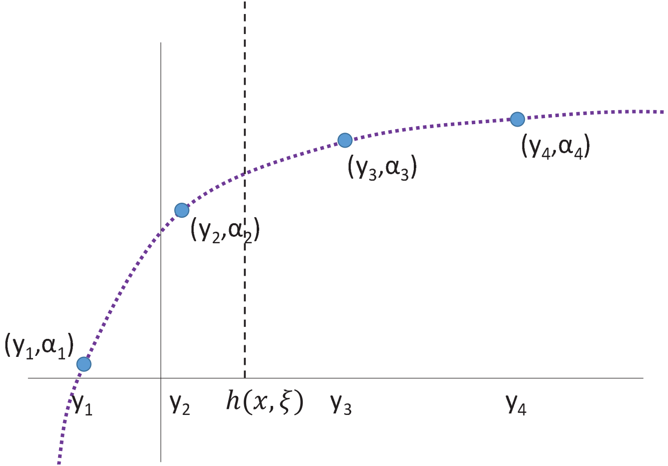
Constructing the Worst-case Utility I

- Define $\mathcal{S} = \{y_1, y_2, ..., y_N\}$ contains support of $\mathcal{W}_k$ and $\mathcal{Y}_k$, and t.
- Define the values $\alpha_i := u(y_i)$



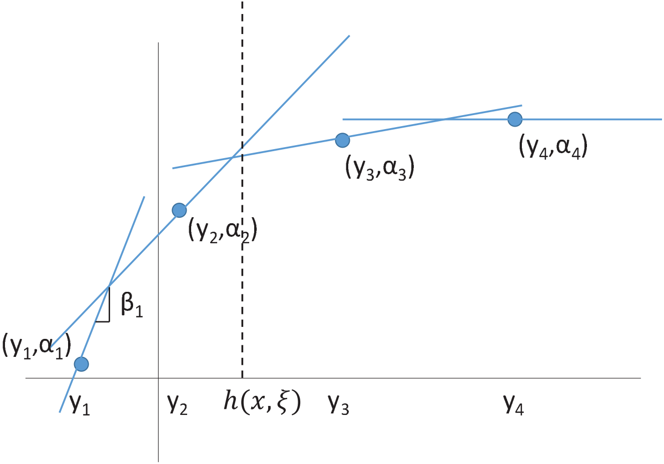
Constructing the Worst-case Utility I

- Define $\mathcal{S} = \{y_1, y_2, ..., y_N\}$ contains support of $\mathcal{W}_k$ and $\mathcal{Y}_k$, and t.
- Define the values $\alpha_i := u(y_i)$



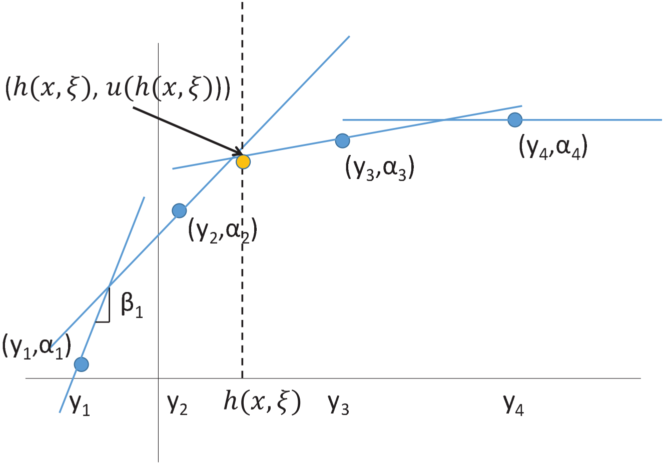
Constructing the Worst-case Utility I

- Define $\mathcal{S} = \{y_1, y_2, ..., y_N\}$ contains support of $\mathcal{W}_k$ and $\mathcal{Y}_k$, and t.
- Define the values $\alpha_i := u(y_i)$



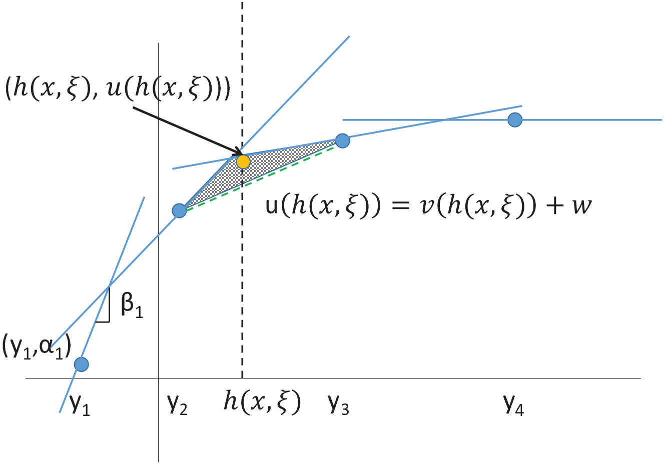
Constructing the Worst-case Utility II

- Once all $(y_i, u(y_i))$ are fixed, identify the worst-case utility value for $u(h(\mathbf{x}, \xi))$.



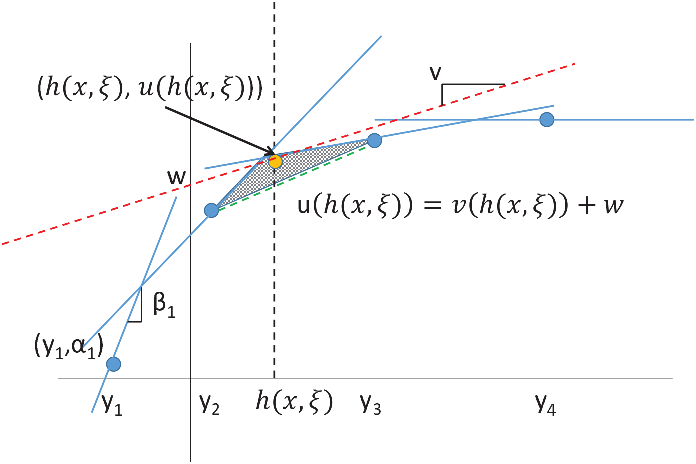
Constructing the Worst-case Utility II

- Once all $(y_i, u(y_i))$ are fixed, identify the worst-case utility value for $u(h(\mathbf{x}, \xi))$.



Constructing the Worst-case Utility II

- Once all $(y_i, u(y_i))$ are fixed, identify the worst-case utility value for $u(h(\mathbf{x}, \xi))$.



LP reformulation of $\inf_{u\in\mathcal{U}} \; E[u(h(\mathbf{x},\xi))] - u(t)$

We wish to find an $\mathbf{x}$ s.t. the following finite dimensional LP has a positive optimal value:

$$\begin{aligned}
\min_{\alpha,\beta,v,w} \quad & \sum_i p_i(v_i h(\mathbf{x},\xi^i) + w_i) - \alpha_t \\
\text{s.t.} \quad & v_i y_i + w_i \geq \alpha_j \;\; \forall\, i,j \quad \text{(Risk aversion at } h(\mathbf{x},\xi^i)) \\
& \sum_j P(\mathcal{W}_k = y_j)\alpha_j \geq \sum_j P(\mathcal{Y}_k = y_j)\alpha_j \;\; \forall\, k \quad \text{(Local pref's)} \\
& \alpha_{j+1} \leq \alpha_j + \beta_j(y_{j+1} - y_j) \;\; \forall\, j \quad \text{(Risk aversion at } y_j\text{'s)} \\
& \alpha_{j-1} \leq \alpha_j + \beta_j(y_{j-1} - y_j) \;\; \forall\, j \\
& v \geq 0, \beta \geq 0 \quad \text{(Monotonicity)}
\end{aligned}$$

After taking the dual of this LP, we can join the maximization with $\mathbf{x} \in \mathcal{X}$

Back to talk