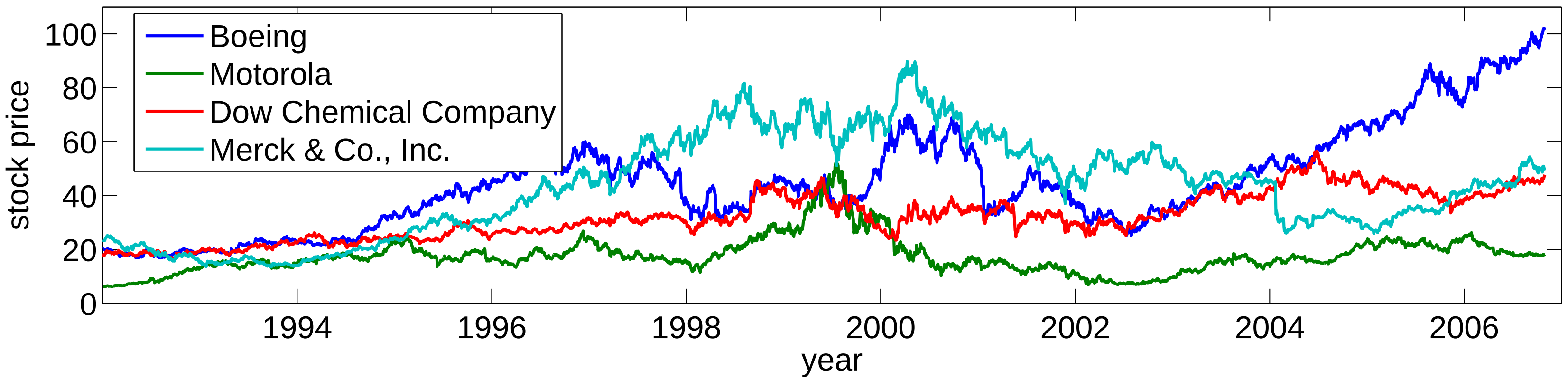
# Data-Driven Optimization for Portfolio Selection

Erick Delage, edelage@stanford.edu
Yinyu Ye, yinyu-ye@stanford.edu
Stanford University

October 2008

# Uncertainty in Portfolio Optimization

- One wants to design a portfolio of stocks
- Stock returns are highly uncertain
- Objective is to maximize daily gains in a "risk sensitive way"
- Difficulty : Little is known about the distribution of daily return for any stock $\xi \sim f_\xi$
- Hope : Benefit from having access to large amount of historical data to build a well-balanced portfolio



# Stochastic Programming Model

Assume that tomorrow's return is drawn randomly from a distribution $f_\xi$. Solve:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \mathbb{E}_{f_\xi}[u(\xi^\mathsf{T} x)] \quad ,$$

where $u(\cdot)$ is a concave utility function that reflects risk aversion.

Pros:

- Accounts explicitly for risk tolerance
- Somewhat tractable (sample average approx.)

Cons:

- It can be difficult to commit to a distribution $f_\xi$ simply based on historical data
- The optimal portfolio can be sensitive to the choice of $f_\xi$

# Dist. Robust Portfolio Optimization

Define a set of distributions $\mathcal{D}$ which is believed to contain $f_\xi$, then choose a portfolio that has highest expected utility with respect to the worst case distribution in $\mathcal{D}$. Hence, solving :

$$(DRPO) \quad \underset{x\in\mathcal{X}}{\text{maximize}} \left( \min_{f_\xi\in\mathcal{D}} \; \mathbb{E}_{f_\xi}[u(\xi^{\mathsf{T}}x)] \right) \quad .$$

In this talk:

- We propose a set $\mathcal{D}$ that constrains the mean, covariance matrix and support of $f_\xi$
- We suggests ways of constructing $\mathcal{D}$ based on historical data in order to be confident that it contains the true $f_\xi$
- We provide an efficient solution method for the resulting DRPO
- We present results using real stock market data

# Related Work

- DRPO with Perfect Moment Information [Popescu, 2007; Natarajan et al., 2008]:
  - For some known mean and covariance matrix, solve the DRPO accounting for all distributions that have such moments
  - Cons : Sensitive to estimation error in $\mu$ and $\Sigma$

- Robust Markowitz Model [Goldfarb et al., 2003]:
  - Use historical data to define an uncertainty set $\mathcal{U}$ for $\mu$ and $\Sigma$ and solve a robust Markowitz model:

$$\underset{x \in \mathcal{X}}{\text{maximize}} \left( \min_{(\mu,\Sigma) \in \mathcal{U}} \mu^{\mathsf{T}} x - \alpha x^{\mathsf{T}} \Sigma x \right)$$

  - Although this problem can be solved efficiently, it is ambiguous how it relates to a true measure of risk

# Describing Distribution Uncertainty

Even when $f_\xi$ is not known exactly, we believe that one can often assume that the distribution lies in a set of the form:

$$\mathcal{D}(\gamma) = \left\{ f_\xi \left| \begin{array}{l} \mathbb{P}(\xi \in \mathcal{S}) = 1 \\ (\mathbb{E}[\xi] - \hat{\mu})^{\mathsf{T}} \hat{\Sigma}^{-1} (\mathbb{E}[\xi] - \hat{\mu}) \leq \gamma_1 \\ z^{\mathsf{T}} \mathbb{E}[(\xi - \hat{\mu})(\xi - \hat{\mu})^{\mathsf{T}}] z \leq (1 + \gamma_2) z^{\mathsf{T}} \hat{\Sigma} z \ , \ \forall z \end{array} \right. \right\} .$$

# Confidence region for $f_\xi$

Given that:

- $\hat{\mu}$ and $\hat{\Sigma}$ are empirical estimates based on $M$ independent samples drawn from $f_\xi$
- $\mathcal{S}$ is contained in a ball of radius $R$

Then, for some $\bar{\gamma}_1 = O(\frac{R^2}{M}\log(1/\delta))$ and some $\bar{\gamma}_2 = O(\frac{R^2}{\sqrt{m}}\sqrt{\log(1/\delta)})$, we can show that

$$\mathbb{P}(f_\xi \in \mathcal{D}(\gamma)) \geq 1 - \delta \quad .$$

Hence, if one solves the DRPO with $\mathcal{D}(\bar{\gamma})$ then he is confident that the resulting portfolio will perform well on the actual distribution $f_\xi$.

# Practical Parametrization

- In practice, historical samples are not identically distributed over the whole history
- Instead, assume data is identically distributed over sub-periods of size $M$
- Build $\mathcal{D}(\gamma_1, \gamma_2)$ as follows:
  - Use $M$ most recent samples to estimate $(\hat{\mu}, \hat{\Sigma})$
  - Choose $\gamma_1$ and $\gamma_2$ such that over $1-\delta$ percent of the pairs of contiguous periods of $M$ samples:

$$\begin{aligned} (\hat{\mu}_2 - \hat{\mu}_1)^{\mathsf{T}} \hat{\Sigma}_1^{-1} (\hat{\mu}_2 - \hat{\mu}_1) &\leq \gamma_1 \\ \hat{\Sigma}_2 + (\hat{\mu}_2 - \hat{\mu}_1)(\hat{\mu}_2 - \hat{\mu}_1)^{\mathsf{T}} &\preceq (1+\gamma_2)\hat{\Sigma}_1 \end{aligned}$$

- Our experiments suggest such a procedure is robust without being too conservative

# Solving the DRPO Problem

**Theorem 1. :** *Given that the utility function has the piecewise linear concave form :*

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

*then the distributionally robust portfolio optimization problem:*

$$\underset{x \in \mathcal{X}}{\text{maximize}} \;\; \left( \min_{f_\xi \in \mathcal{D}(\gamma)} \; \mathbb{E}_{f_\xi}[u(\xi^{\mathsf{T}} x)] \right)$$

1. *can be solved in polynomial time as long as $\mathcal{S}$ is convex*
2. *can be solved in $O(K^{3.5} n^{6.5})$ given that $\mathcal{S}$ is ellipsoidal*

# Solving the DRPO Problem

If $\mathcal{S}$ takes the form:

$$\mathcal{S} = \{\xi | (\xi - \xi_0)^\mathsf{T} A(\xi - \xi_0) \leq \rho\} \ \ , A \succeq 0$$

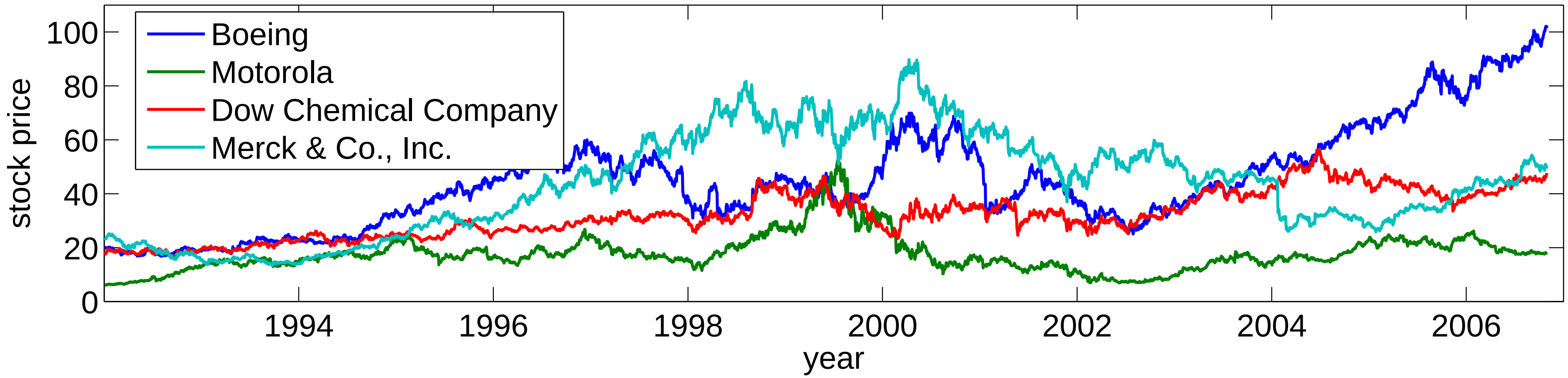
then the DRPO reduces to the Semi-Definite Program:

$$\begin{array}{ll} \underset{x,Q,q,t,P,p,s,\tau}{\text{minimize}} & \gamma_2 \, \mathbf{trace}(\hat{\Sigma}Q) - \hat{\mu}^\mathsf{T} Q \hat{\mu} + t + \mathbf{trace}(\hat{\Sigma}P) - 2\hat{\mu}^\mathsf{T} p + \gamma_1 s \\ \text{subject to} & \begin{bmatrix} P & p \\ p^\mathsf{T} & s \end{bmatrix} \succeq 0 \ , \ \ p = -q/2 - Q\hat{\mu} \\ & \begin{bmatrix} Q & q/2 + a_k x/2 \\ q^\mathsf{T}/2 + a_k x^\mathsf{T}/2 & t + b_k \end{bmatrix} \succeq -\tau_k \begin{bmatrix} A & -A\xi_0 \\ -\xi_0^\mathsf{T} A & \xi_0^\mathsf{T} A \xi_0 - \rho \end{bmatrix} \ , \ \forall k \\ & \tau_k \geq 0 \ , \ \ \forall \, k \ \ , \ \ Q \succeq 0 \ \ , \ \ x \in \mathcal{X} \ \ , \end{array}$$

which can be solved efficiently using an interior point algorithm.

# Experiments with Historical Data

30 stocks were tracked over horizon (1992-2007)



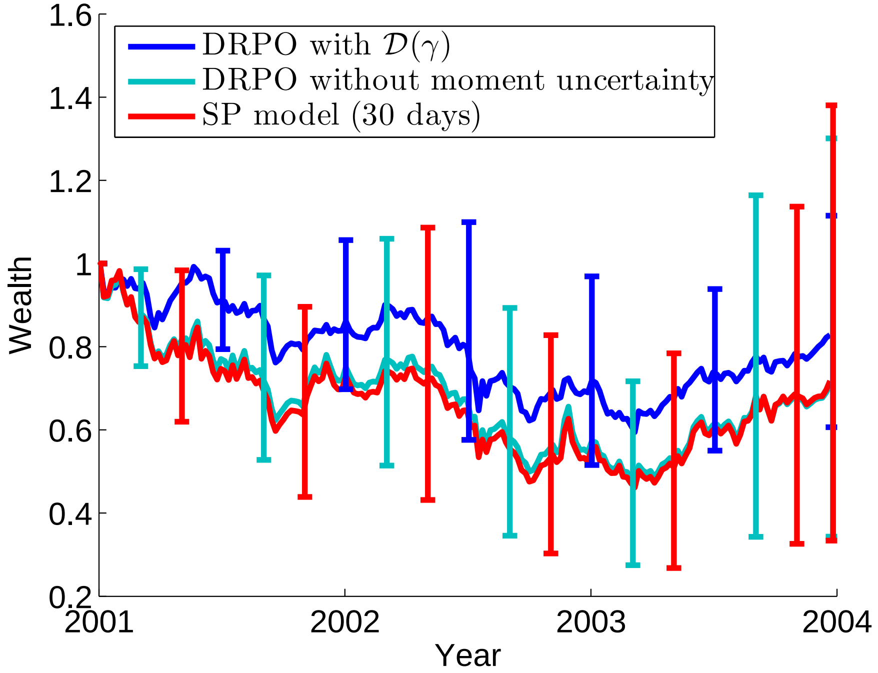
# Experiments with Historical Data

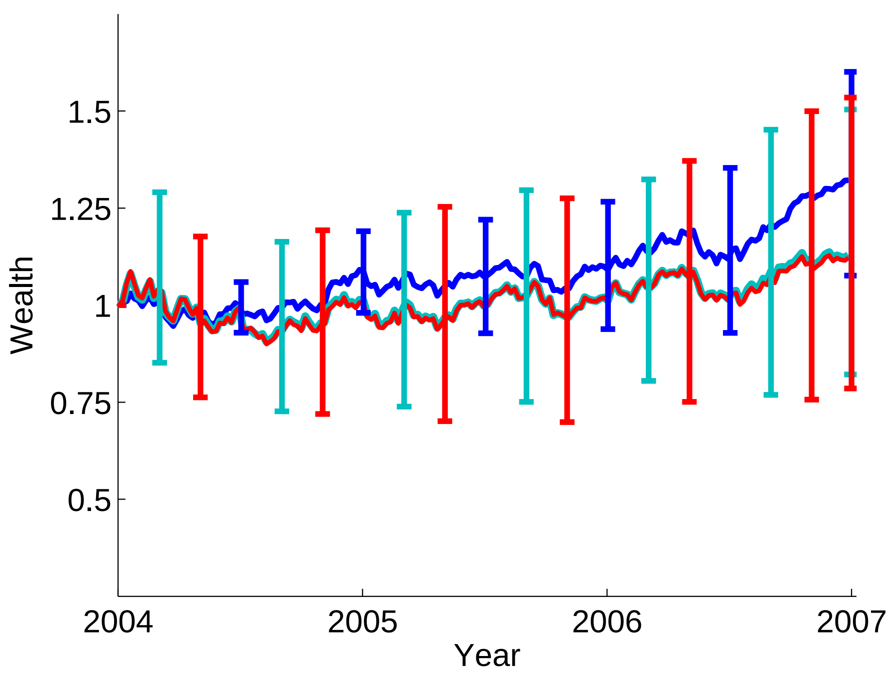
An experiment consists of trading 4 stocks over (2001-07).

- Use (1992-2001) to choose $\gamma_1$ and $\gamma_2$
- Update portfolio on daily basis
- Estimate $\hat{\mu}$ and $\hat{\Sigma}$ based on a 30 days period
- DRPO with $\mathcal{D}(\gamma)$ is compared to :
  - DRPO without moment uncertainty
  - Stochastic Program using empirical distribution over last 30 days

# Experimental Results I

Comparison of wealth evolution in 300 experiments conducted over the years 2001-2007. For each model, the periodical 10% and 90% percentiles of wealth are indicated.



# Experimental Results II

In finer details:

| Method | 2001-2004 | | 2004-2007 | |
|---|---|---|---|---|
| | Avg. yearly return | 10-perc. | Avg. yearly return | 10-perc. |
| DRPO with $\mathcal{D}(\gamma)$ | 0.944 | 0.846 | 1.102 | 1.025 |
| DRPO w/o moment uncertainty | 0.700 | 0.334 | 1.047 | 0.936 |
| SP model (30 days) | 0.908 | 0.694 | 1.045 | 0.923 |

- 79% of the time, our DRPO outperformed both models
- On average accounting for moment uncertainty led to a relative gain of 1.67

# Summary

- Derived a DRPO which accounts for limited distribution information present in historical data
- Proposed a set $\mathcal{D}(\bar{\gamma})$ with probabilistic guarantees in data-driven problems
- Empirically justified the need to account for distribution & moment uncertainty in portfolio optimization

- We encourage using a distributionally robust criterion as an objective or constraint; hence, hedge against the risks of making investment decisions that rely on an inaccurate probabilistic model

# Questions & Comments ...

... Thank you!