Estimation error reduction in portfolio optimization with CVaR

Gah-Yi Vahn*

Noureddine El Karoui† Andrew E.B. Lim*

*Department of IEOR
†Department of Statistics
University of California, Berkeley

Toronto 2011

Outline

The Portfolio Optimization Problem

The generic problem:

$$\begin{array}{rrcl} \min_{\mathbf{w}} & Risk(\mathbf{w}) & & \\ s.t. & Return(\mathbf{w}) & = & R \\ & \mathbf{w}'\mathbf{1} & = & 1 \end{array} \tag{1}$$

- Idea: minimize some notion of risk while preserving a guaranteed return level.
- Markowitz problem:
  - Assume some distribution P for asset returns: $\mathbf{X} \sim P$
  - Risk = variance of portfolio return: $Variance(\mathbf{w}'\mathbf{X})$
  - Reward = mean of portfolio return: $E(\mathbf{w}'\mathbf{X})$
- There are variations in characterizing portfolio risk
- We consider risk = conditional Value-at-Risk (CVaR) of the portfolio.

The Portfolio Optimization Problem

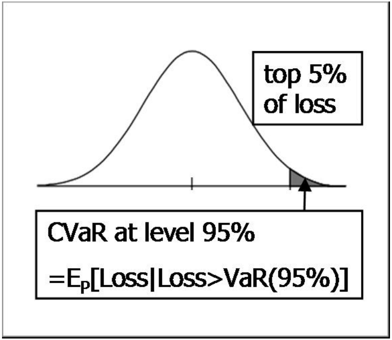
The generic problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & Risk(\mathbf{w}) & & \\ s.t. \quad & Return(\mathbf{w}) & = & \; R \\ & \mathbf{w}'\mathbf{1} & = & \; 1 \end{aligned} \tag{1}$$

- Idea: minimize some notion of risk while preserving a guaranteed return level.
- Markowitz problem:
  - Assume some distribution P for asset returns: $\mathbf{X} \sim P$
  - Risk = variance of portfolio return: $Variance(\mathbf{w}'\mathbf{X})$
  - Reward = mean of portfolio return: $E(\mathbf{w}'\mathbf{X})$
- There are variations in characterizing portfolio risk
- We consider risk = conditional Value-at-Risk (CVaR) of the portfolio.

Conditional Value-at-Risk

- CVaR at level 95%: the average loss in the top 5%
  - Tells you something about the loss tail
  - Also coherent [[Pflug(2000), Acerbi and Tasche (2001)]



CVaR Portfolio Optimization Problem

$$
\begin{aligned}
\min_{\mathbf{w}} \quad & CVaR(\mathbf{w}; \mathbf{X}, \beta) \\
s.t. \quad & \mathbf{w}^\top \mu = R \\
& \mathbf{w}^\top \mathbf{1} = 1
\end{aligned} \tag{2}
$$

- **X**: vector of asset returns $\stackrel{d}{=} P$
- solution: $\mathbf{w}_0$

The Empirical Problem

- But we don't know P, the distribution of $\mathbf{X}$
- Suppose we observe iid data $\mathcal{X}_n = X_1, \ldots, X_n \sim P$.
- Then solve the empirical problem:

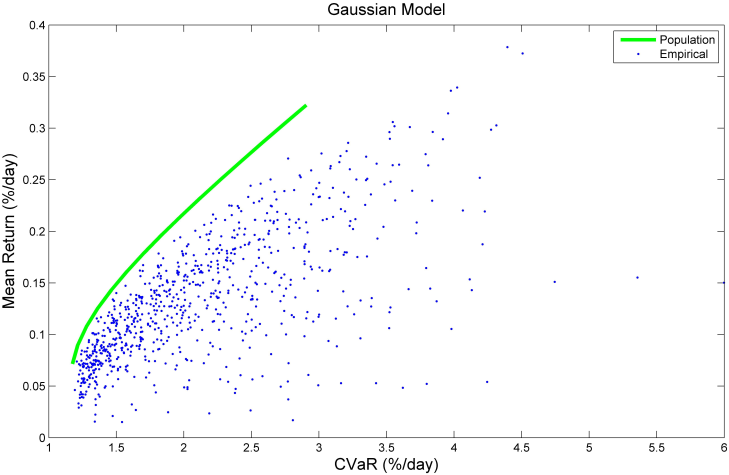
$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \widehat{CVaR}_n(\mathbf{w}; \mathcal{X}, \beta) \\ s.t. & \mathbf{w}'\hat{\mu}_n = R \\ & \mathbf{w}'\mathbf{1} = 1 \end{array} \tag{3}$$

- $\widehat{CVaR}_n(\mathbf{w}; \mathbf{X}, \beta)$: unbiased nonparametric estimator
- solution: $\hat{\mathbf{w}}_n$
- can be expressed as a LP [Rockafellar & Uryasev (2000)]
- but solution is very fragile [Lim, Shanthikumar & Vahn (2011)]

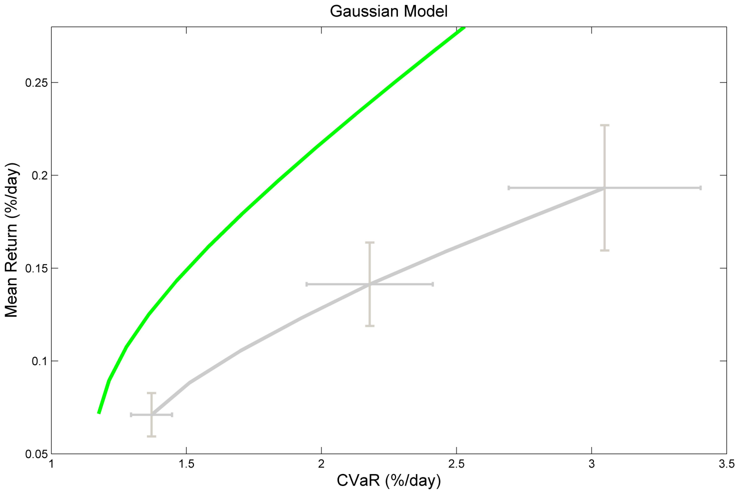
Example: Empirical Problem is Fragile

- Model: $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$
- Simulate 250 iid daily "observations" under the model
- Solve the empirical problem for $\hat{\mathbf{w}}_n$
- Plot realized return vs. realized CVaR of $\hat{\mathbf{w}}_n$
- Repeat (Monte Carlo) to get a distribution for the empirical solution

Example: why Empirical Problem is bad



Example: why Empirical Problem is bad



Research objective

- We want to shift closer to the population frontier with more reliability, i.e. less fluctuation
- Two methods:
  - Nonparametric: performance-based regularization
  - Parametric: alternative optimization for $\mathbf{X} \sim \mathit{Ellip}$

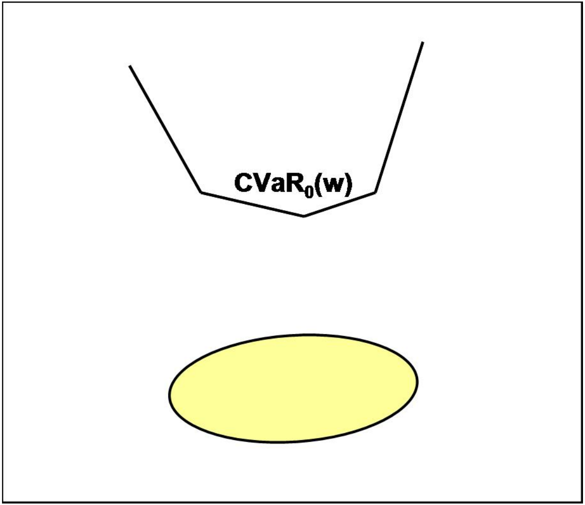
Research objective

- We want to shift closer to the population frontier with more reliability, i.e. less fluctuation
- Two methods:
  - Nonparametric: performance-based regularization
  - Parametric: alternative optimization for $\mathbf{X} \sim \mathit{Ellip}$

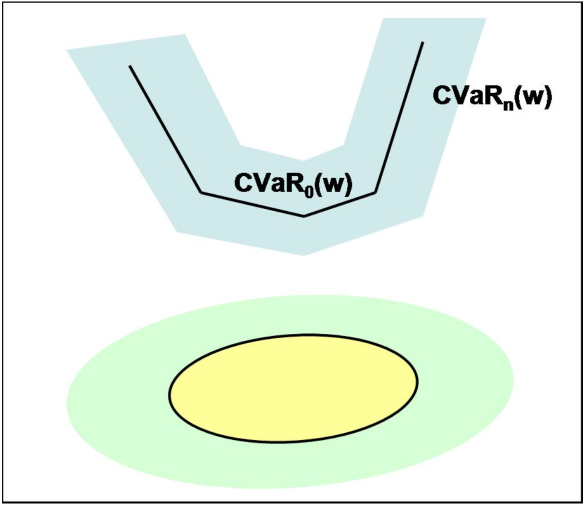
Nonparametric method

- Only assumption is that we observe iid returns $\mathbf{X}$
- Let's look at a picture for intuition

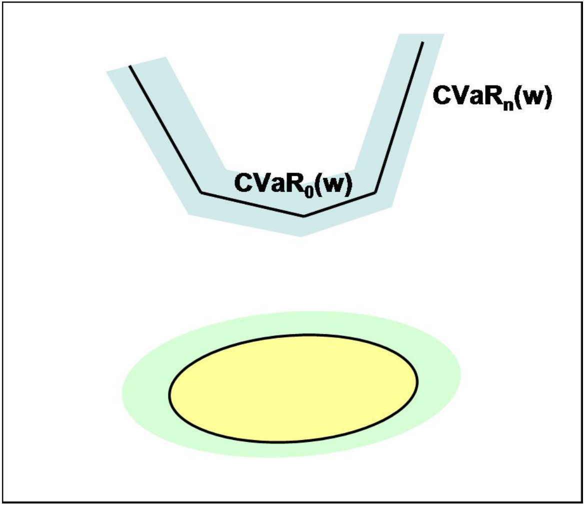
Schematic of Population Problem



Schematic of Empirical Problem



Schematic of Empirical Problem



Nonparametric: Performance-based Regularization

- What's an appropriate penalty function?
- Let us consider penalizing the variance of $\widehat{CVaR}_n(\mathbf{w}; \mathcal{X}_n, \beta)$ and $\mathbf{w}'\hat{\mu}_n$

$$
\begin{array}{lrl}
\min_{\mathbf{w}} & \widehat{CVaR}_n(\mathbf{w}; \mathcal{X}, \beta) & \\
s.t. & \mathbf{w}'\hat{\mu}_n = & R \\
 & \mathbf{w}'\mathbf{1} = & 1 \\
 & \widehat{VAR1}(\mathbf{w}) \leq & U_1 \\
 & \widehat{VAR2}(\mathbf{w}) \leq & U_2
\end{array}
$$

where $\widehat{VAR1}$, $\widehat{VAR2}$ are sample variances of $\widehat{CVaR}_n(\mathbf{w})$ and $\mathbf{w}'\hat{\mu}_n$

- Theorem: The regularized problem with $\widehat{VAR1}$, $\widehat{VAR2}$ penalty functions is QCQP
- Hence, can be solved efficiently

Nonparametric: Performance-based Regularization

- What's an appropriate penalty function?
- Let us consider penalizing the variance of $\widehat{CVaR}_n(\mathbf{w}; \mathcal{X}_n, \beta)$ and $\mathbf{w}'\hat{\mu}_n$

$$
\begin{array}{lrl}
\min\limits_{\mathbf{w}} & \widehat{CVaR}_n(\mathbf{w}; \mathcal{X}, \beta) & \\
s.t. & \mathbf{w}'\hat{\mu}_n = & R \\
 & \mathbf{w}'\mathbf{1} = & 1 \\
 & \widehat{VAR1}(\mathbf{w}) \leq & U_1 \\
 & \widehat{VAR2}(\mathbf{w}) \leq & U_2
\end{array}
$$

where $\widehat{VAR1}$, $\widehat{VAR2}$ are sample variances of $\widehat{CVaR}_n(\mathbf{w})$ and $\mathbf{w}'\hat{\mu}_n$

- Theorem: The regularized problem with $\widehat{VAR1}$, $\widehat{VAR2}$ penalty functions is QCQP
- Hence, can be solved efficiently

Nonparametric: Performance-based Regularization

- What's an appropriate penalty function?
- Let us consider penalizing the variance of $\widehat{CVaR}_n(\mathbf{w}; \mathcal{X}_n, \beta)$ and $\mathbf{w}'\hat{\boldsymbol{\mu}}_n$

$$
\begin{aligned}
\min_{\mathbf{w}} \quad & \widehat{CVaR}_n(\mathbf{w}; \mathcal{X}, \beta) \\
s.t. \quad & \mathbf{w}'\hat{\boldsymbol{\mu}}_n = R \\
& \mathbf{w}'\mathbf{1} = 1 \\
& \widehat{VAR1}(\mathbf{w}) \leq U_1 \\
& \widehat{VAR2}(\mathbf{w}) \leq U_2
\end{aligned}
$$

where $\widehat{VAR1}$, $\widehat{VAR2}$ are sample variances of $\widehat{CVaR}_n(\mathbf{w})$ and $\mathbf{w}'\hat{\boldsymbol{\mu}}_n$

- Theorem: The regularized problem with $\widehat{VAR1}$, $\widehat{VAR2}$ penalty functions is QCQP
- Hence, can be solved efficiently

Parametric: alternative optimization

- If $\mathbf{X} \sim Ellip(\boldsymbol{\mu}, Y, \Sigma)$, i.e. $\mathbf{X} \sim \boldsymbol{\mu} + Y\Sigma^{1/2}U$
  - $\boldsymbol{\mu}$ is the mean vector,
  - U is unif. distributed on the p-dim sphere of radius 1
  - Y is a non-negative random variable independent of U.
  - Special case: $Y = ||Z_p||$, $Z_p \sim \mathcal{N}(0, I_p)$ then $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$.
- FACT: $\boxed{CVaR(\mathbf{w}; \mathbf{X}, \beta) = C\sqrt{\mathbf{w}'\Sigma\mathbf{w}} - \mathbf{w}'\boldsymbol{\mu}}$
- So Markowitz and mean-CVaR are equivalent in population (truth)
- Why not solve empirical Markowitz instead of empirical mean-CVaR?

Parametric: alternative optimization

- If $\mathbf{X} \sim Ellip(\boldsymbol{\mu}, Y, \Sigma)$, i.e. $\mathbf{X} \sim \boldsymbol{\mu} + Y\Sigma^{1/2}U$
  - $\boldsymbol{\mu}$ is the mean vector,
  - U is unif. distributed on the p-dim sphere of radius 1
  - Y is a non-negative random variable independent of U.
  - Special case: $Y = ||Z_p||$, $Z_p \sim \mathcal{N}(0, I_p)$ then $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$.
- FACT: $\boxed{CVaR(\mathbf{w}; \mathbf{X}, \beta) = C\sqrt{\mathbf{w}'\Sigma\mathbf{w}} - \mathbf{w}'\boldsymbol{\mu}}$
- So Markowitz and mean-CVaR are equivalent in population (truth)
- Why not solve empirical Markowitz instead of empirical mean-CVaR?

Evaluation of methods

- Want to compare location and variability of efficient frontiers:
- Empirical ($\hat{\mathbf{w}}_n$) vs Regularized ($\hat{\mathbf{w}}_{reg}$) vs Markowitz ($\hat{\mathbf{w}}_{Mark}$)

- Newsflash: we can do the comparison theoretically by comparing the asymptotic distributions of the solutions. Stay tuned!

- But, let's just stay with comparison via Monte Carlo, as before

Evaluation of methods

- Want to compare location and variability of efficient frontiers:
- Empirical ($\hat{\mathbf{w}}_n$) vs Regularized ($\hat{\mathbf{w}}_{reg}$) vs Markowitz ($\hat{\mathbf{w}}_{Mark}$)

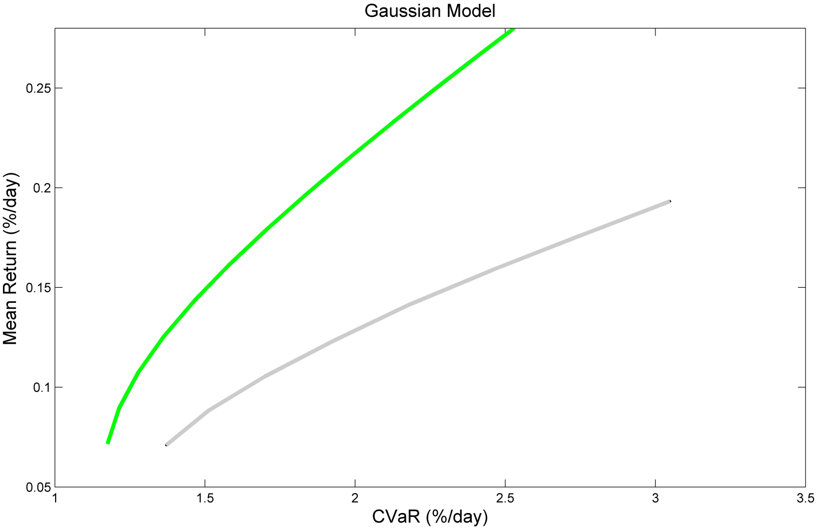
- Newsflash: we can do the comparison theoretically by comparing the asymptotic distributions of the solutions. Stay tuned!

- But, let's just stay with comparison via Monte Carlo, as before

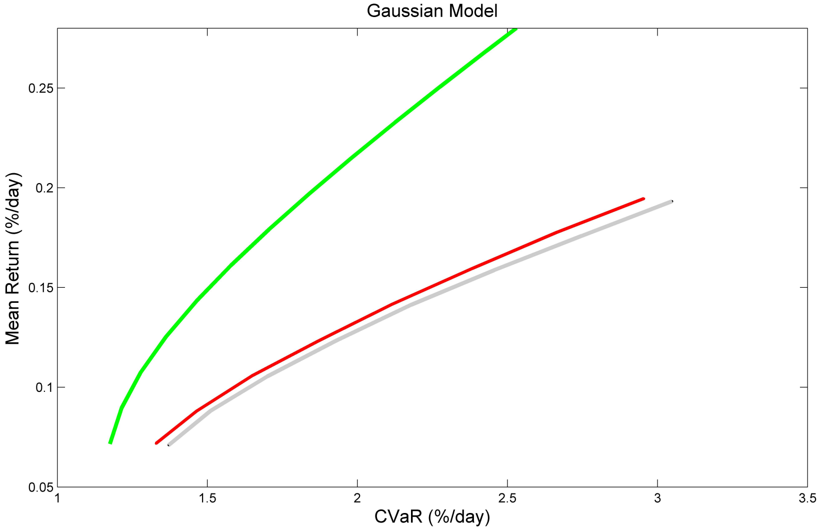
Evaluation of methods

- Want to compare location and variability of efficient frontiers:
- Empirical ($\hat{\mathbf{w}}_n$) vs Regularized ($\hat{\mathbf{w}}_{reg}$) vs Markowitz ($\hat{\mathbf{w}}_{Mark}$)
- Newsflash: we can do the comparison theoretically by comparing the asymptotic distributions of the solutions. Stay tuned!
- But, let's just stay with comparison via Monte Carlo, as before

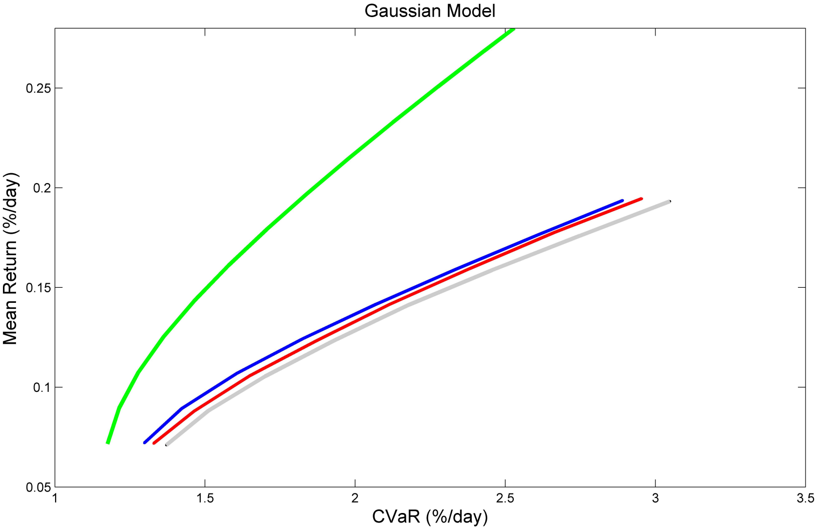
$\mathbf{X} \sim$ *Gaussian*



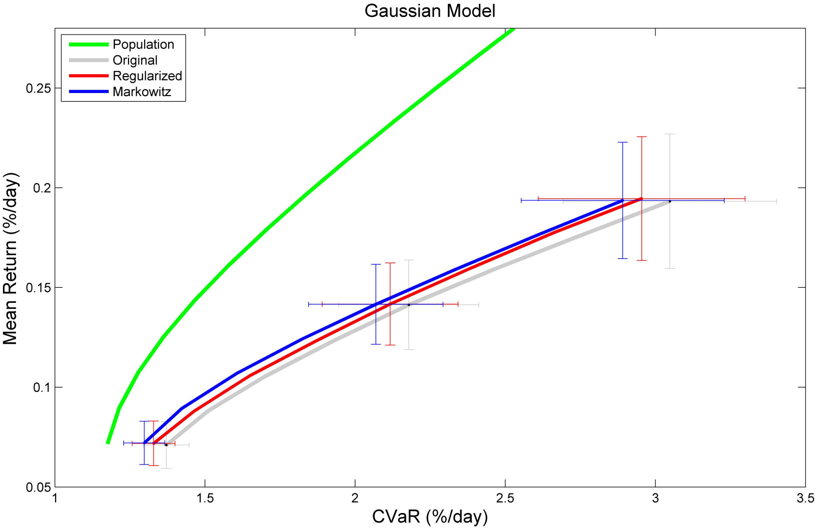
$\mathbf{X} \sim$ *Gaussian*



$\mathbf{X} \sim$ *Gaussian*



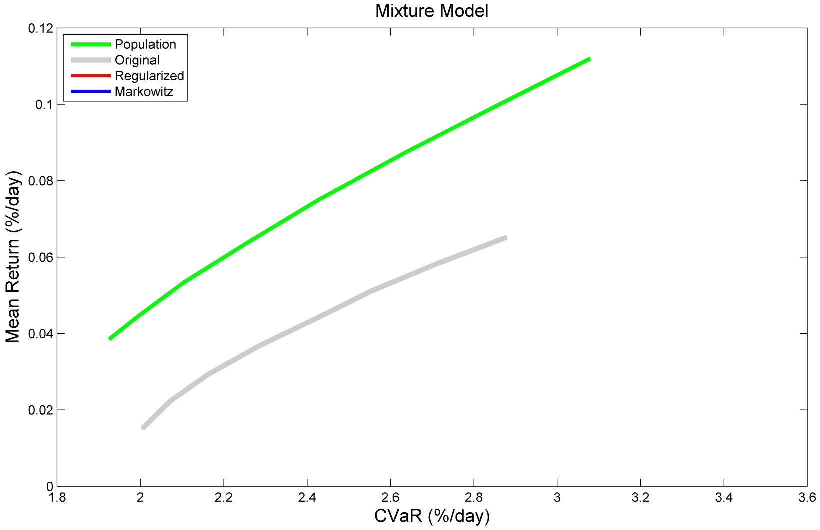
$\mathbf{X} \sim$ *Gaussian*



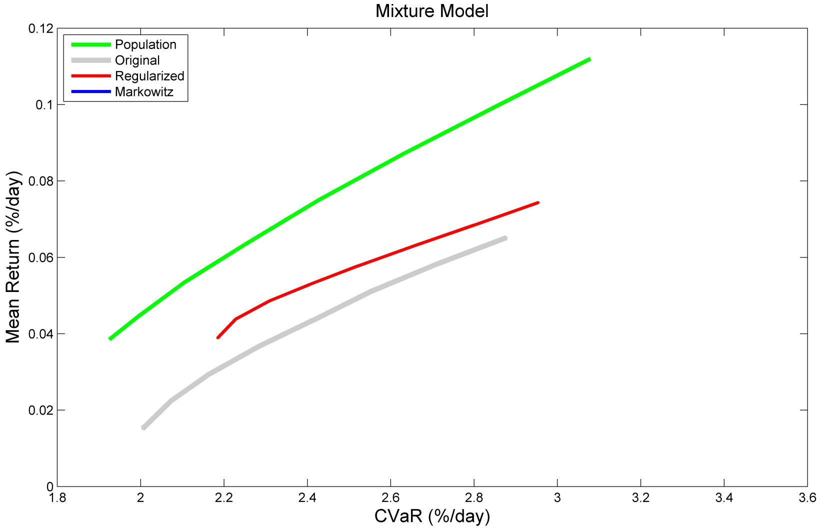
Summary of results

- Empirical $<$ Regularization $\leq$ Markowitz
- Both methods: we're closer to the population frontier with less fluctuation!
- For $\mathbf{X} \sim$ *Ellip*, similar result
- But can Regularization $>$ Markowitz?

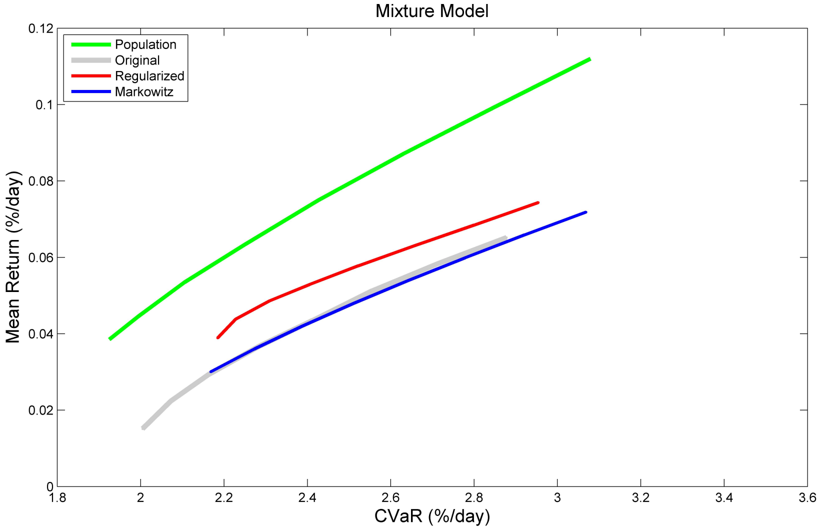
$\mathbf{X} \sim I(0.95)Gaussian + I(0.05)Neg.Exp$



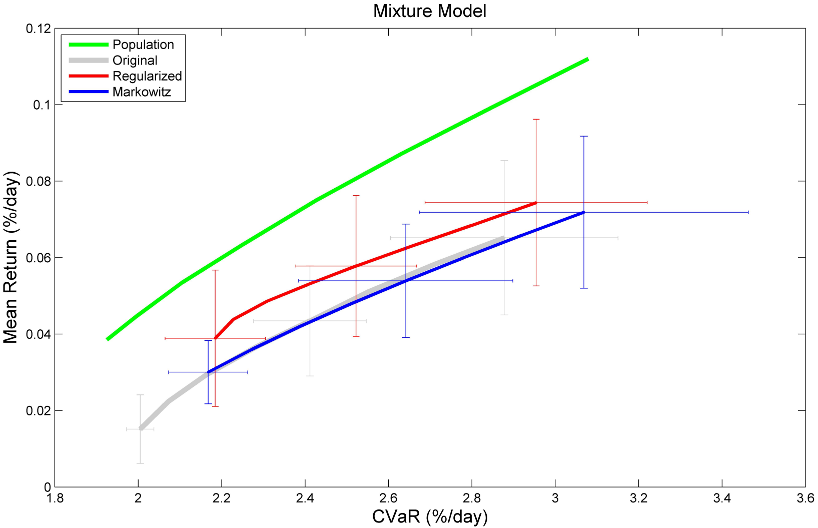
$\mathbf{X} \sim I(0.95)Gaussian + I(0.05)Neg.Exp$



$\mathbf{X} \sim I(0.95)Gaussian + I(0.05)Neg.Exp$



$\mathbf{X} \sim I(0.95)Gaussian + I(0.05)Neg.Exp$



Conclusion

- Empirical portfolio optimization is BAD
- Can improve upon the empirical solution
  - Nonparametric: regularization helps
  - Parametric (more info): If $\mathbf{X} \sim \mathit{Ellip}$, Markowitz is better
- Directions for future work:
  - Performance-based regularization of the empirical Markowitz problem?
  - Consider different penalty functions
  - Combine with known statistical methods of variance reduction (e.g. bootstrapping, sub-sampling)

Conclusion

- Empirical portfolio optimization is BAD
- Can improve upon the empirical solution
  - Nonparametric: regularization helps
  - Parametric (more info): If $\mathbf{X} \sim \mathit{Ellip}$, Markowitz is better
- Directions for future work:
  - Performance-based regularization of the empirical Markowitz problem?
  - Consider different penalty functions
  - Combine with known statistical methods of variance reduction (e.g. bootstrapping, sub-sampling)

Conclusion

- Empirical portfolio optimization is BAD
- Can improve upon the empirical solution
  - Nonparametric: regularization helps
  - Parametric (more info): If $\mathbf{X} \sim \mathit{Ellip}$, Markowitz is better
- Directions for future work:
  - Performance-based regularization of the empirical Markowitz problem?
  - Consider different penalty functions
  - Combine with known statistical methods of variance reduction (e.g. bootstrapping, sub-sampling)