Portfolio selection with multiple risk measures

Garud Iyengar
Columbia University
Industrial Engineering and Operations Research

Joint work with Carlos Abad
Outline

• Portfolio selection and risk measures
  • Variance
  • Value at Risk
  • Coherent risk measures: Spectral risk measures

• Finite sample approximations
  • Linear programs: very large and ill-conditioned
  • Cannot handle multiple risk models in practice

• Our contribution: Fast first-order algorithm
  • Decomposition algorithm
  • Can handle very large instances
  • Can accommodate multiple risk models
Portfolio selection and risk measures

- Two objectives: Maximize expected return and minimize “risk”
  - Maximize return with a “risk” bound
  - Maximize a weighted combination of return and “risk”
Portfolio selection and risk measures

- Two objectives: Maximize expected return and minimize “risk”
  - Maximize return with a “risk” bound
  - Maximize a weighted combination of return and “risk”

- Markowitz mean-variance portfolio selection model (1952)
  - “risk” $\equiv$ variance of the return
  - Portfolio selection: convex quadratic program
  - Variance only appropriate for elliptical distributions
  - Does not model “tail” losses well
Portfolio selection and risk measures

- Two objectives: Maximize expected return and minimize “risk”
  - Maximize return with a “risk” bound
  - Maximize a weighted combination of return and “risk”

- Markowitz mean-variance portfolio selection model (1952)
  - “risk” $\equiv$ variance of the return
  - Portfolio selection: convex quadratic program
  - Variance only appropriate for elliptical distributions
  - Does not model “tail” losses well

- Value-at-Risk

\[ \text{VaR}_\beta(\tilde{L}) = \inf \{ v : \mathbb{P}(\tilde{L} \leq v) \geq \beta \} \]

- Mandated in the Basel-II accord
- Probability of tail losses but not magnitude
- Not a convex risk measure: portfolio selection hard
Coherent risk measures

- Coherent risk measures $\rho(\tilde{L})$ satisfy
  
  (i) **Monotonicity**: if $\tilde{X} \geq \tilde{Y}$, then $\rho(\tilde{X}) \geq \rho(\tilde{Y})$
  
  (ii) **Positive homogeneity**: for all $\alpha \geq 0$, $\rho(\alpha \tilde{X}) = \alpha \rho(\tilde{X})$
  
  (iii) **Convexity**: for all $0 \leq \alpha \leq 1$, 
  $$\rho(\alpha \tilde{X} + (1 - \alpha) \tilde{Y}) \leq \alpha \rho(\tilde{X}) + (1 - \alpha) \rho(\tilde{Y})$$
  
  (iv) **Cash-invariance**: for all $\alpha \in \mathbb{R}$, 
  $$\rho(\tilde{X} + \alpha) = \rho(\tilde{X}) + \alpha$$
Coherent risk measures

- Coherent risk measures $\rho(\tilde{L})$ satisfy
  \[ \rho(\tilde{L}) = \sup_{Q \in Q} \{ E_Q[\tilde{L}] \} \]
Coherent risk measures

- Coherent risk measures $\rho(\tilde{L})$ satisfy
  \[
  \rho(\tilde{L}) = \sup_{Q \in \mathcal{Q}} \left\{ E_Q[\tilde{L}] \right\}
  \]

- Expected Shortfall
  \[
  \text{ES}_\beta(\tilde{L}) = \frac{1}{1 - \beta} \int_\beta^1 \text{VaR}_p(\tilde{L}) dp
  \]
  Also called Average Value at Risk. Almost Conditional Value-at-Risk.
Coherent risk measures

- Coherent risk measures $\rho(\tilde{L})$ satisfy
  
  $$\rho(\tilde{L}) = \sup_{Q \in \mathcal{Q}} \{ E_Q[\tilde{L}] \}$$

- Expected Shortfall
  
  $$\text{ES}_\beta(\tilde{L}) = \frac{1}{1 - \beta} \int_\beta^1 \text{VaR}_p(\tilde{L}) dp$$

  Also called Average Value at Risk. Almost Conditional Value-at-Risk.

- Mean Upper Semi-deviation ($\lambda \in [0, 1]$)
  
  $$\rho_{upper}(\tilde{L}) = E_P[\tilde{L}] + \lambda \left( E_P \left[ (\tilde{L} - E_P[\tilde{L}])^+ \right] \right)^{\frac{1}{2}}$$
Coherent risk measures

- Coherent risk measures $\rho(\tilde{L})$ satisfy
  \[ \rho(\tilde{L}) = \sup_{Q \in \mathcal{Q}} \{ \mathbb{E}_Q[\tilde{L}] \} \]

- Expected Shortfall
  \[ \text{ES}_\beta(\tilde{L}) = \frac{1}{1 - \beta} \int_\beta^1 \text{VaR}_p(\tilde{L}) \, dp \]
  Also called Average Value at Risk. Almost Conditional Value-at-Risk.

- Mean Upper Semi-deviation ($\lambda \in [0, 1]$)
  \[ \rho_{\text{upper}}(\tilde{L}) = \mathbb{E}_P[\tilde{L}] + \lambda \left( \mathbb{E}_P \left[ (\tilde{L} - \mathbb{E}_P[\tilde{L}])^+ \right] \right)^{\frac{1}{2}} \]

- Convex risk measures: portfolio optimization easy (in theory!)
Spectral risk measures

- Spectral Risk Measure (Acerbi (2002))

\[ M_\gamma(\tilde{L}) = \int_0^1 \text{ES}_\beta(\tilde{L}) \, d\gamma(\beta) \]

- \( \gamma : [0, 1] \rightarrow \mathbb{R}_+ \): probability measure

Bertsimas and Brown (2010): Distortion Risk Measures
Spectral risk measures

- Spectral Risk Measure (Acerbi (2002))

\[ M_\gamma(\tilde{L}) = \int_0^1 ES_\beta(\tilde{L}) d\gamma(\beta) \]

- \( \gamma : [0, 1] \rightarrow \mathbb{R}_+ : \) probability measure

Bertsimas and Brown (2010): Distortion Risk Measures

- Spectral Risk Measures \( \equiv \) Coherent, comonotone additive, and law-invariant
  - Law invariant risk measure: \( \tilde{X} = \tilde{Y} \) a.s. \( \Rightarrow \rho(\tilde{X}) = \rho(\tilde{Y}) \)
  - Sampling methods only work for law invariant risk measures
Spectral risk measures

- Spectral Risk Measure (Acerbi (2002))
  \[ M_\gamma(\tilde{L}) = \int_0^1 ES_\beta(\tilde{L}) \, d\gamma(\beta) \]
  
  - \( \gamma : [0, 1] \to \mathbb{R}_+ \): probability measure

Bertsimas and Brown (2010): Distortion Risk Measures

- Spectral Risk Measures \( \equiv \) Coherent, comonotone additive, and law-invariant
  
  - Law invariant risk measure: \( \tilde{X} = \tilde{Y} \) a.s. \( \Rightarrow \rho(\tilde{X}) = \rho(\tilde{Y}) \)
  
  - Sampling methods only work for law invariant risk measures

- **Law-invariant** risk measures \( \rho_{LI}(\tilde{L}) \) satisfy (Kusuoka 2001)
  \[ \rho_{LI}(\tilde{L}) = \max_{\gamma \in \Gamma} \{ M_\gamma(\tilde{L}) \} \]
Finite approximations

- $N$ samples of losses: $\ell = [\ell_1, \ldots, \ell_N]^\top$
Finite approximations

- $N$ samples of losses: $\ell = [\ell_1, \ldots, \ell_N]^\top$

- Finite sample approximation of ES:
  \[
  \text{ES}_\beta(\ell) = \max q^\top \ell, \\
  \text{s.t.} \quad 1^\top q = 1, \\
  0 \leq q \leq \frac{1}{(1-\beta)N} \cdot 1.
  \]

- A linear program
Finite approximations

- $N$ samples of losses: $\ell = [\ell_1, \ldots, \ell_N]^\top$

- Finite sample approximation of ES:
  
  $$ES_\beta(\ell) = \max q^\top \ell,$$
  
  s.t.  
  
  $$1^\top q = 1,$$
  
  $$0 \leq q \leq \frac{1}{(1 - \beta)N} \cdot 1.$$

- A linear program

- Spectral risk measures:
  
  $$M_\gamma(\ell) = \sum_{j=1}^{m} \gamma_j ES_{\beta_j}(\ell)$$
Finite approximations

- $N$ samples of losses: $\ell = [\ell_1, \ldots, \ell_N]^\top$

- Finite sample approximation of ES:

$$ES_\beta(\ell) = \max \mathbf{q}^\top \ell,$$

s.t. $\begin{align*}
\mathbf{1}^\top \mathbf{q} &= 1, \\
0 &\leq \mathbf{q} \leq \frac{1}{(1-\beta)N} \cdot \mathbf{1}.
\end{align*}$

- A linear program

- Spectral risk measures:

$$M_\gamma(\ell) = \sum_{j=1}^{m} \gamma_j ES_{\beta_j}(\ell)$$

- Law invariant risk measures:

$$\rho(\ell) = \max_{\gamma \in \Gamma} \left\{ \sum_{j=1}^{m} \gamma_j ES_{\beta_j}(\ell) \right\}$$
Mean-spectral risk portfolio selection problem

- $n$ assets: portfolio $\mathbf{x} \in \mathbb{R}^n$ with $\mathbf{1}^\top \mathbf{x} = 1$.

- $m$ different risk models
  - $\mathbf{L}_k \in \mathbb{R}^{N_k \times n}$: Loss matrix for the $k$-th model
  - $\rho_k(\mathbf{L}_k \mathbf{x}) = \sum_{j=1}^{d_k} \gamma_j^{(k)} \text{ES}_{\beta_j^{(k)}}(\mathbf{L}_k \mathbf{x})$

Why bother with multiple risk models?

- Good, bad, ugly historical return periods
- Risk models with different periods
- Robustness to parameters

Portfolio selection problem

$$\begin{align*}
\text{max} & \quad \mathbf{1}^\top \mathbf{x} - \lambda \| \mathbf{x} \|_1 \\
\text{s.t.} & \quad \rho_k(\mathbf{L}_k \mathbf{x}) \leq \alpha_k, \quad k = 1, \ldots, m \\
& \quad \mathbf{1}^\top \mathbf{x} = 1, \quad \| \mathbf{x} \|_\infty \leq B
\end{align*}$$
Mean-spectral risk portfolio selection problem

- **n** assets: portfolio \( \mathbf{x} \in \mathbb{R}^n \) with \( \mathbf{1}^\top \mathbf{x} = 1 \).

- **m** different risk models
  - \( \mathbf{L}_k \in \mathbb{R}^{N_k \times n} \): Loss matrix for the \( k \)-th model
  - \( \rho_k(\mathbf{L}_k \mathbf{x}) = \sum_{j=1}^{d_k} \gamma_{j}^{(k)} \text{ES}_{\beta_{j}^{(k)}}(\mathbf{L}_k \mathbf{x}) \)

- Why bother with multiple risk models?
  - Good, bad, ugly historical return periods
  - Risk models with different periods
  - Robustness to parameters
Mean-spectral risk portfolio selection problem

- \( n \) assets: portfolio \( \mathbf{x} \in \mathbb{R}^n \) with \( \mathbf{1}^\top \mathbf{x} = 1 \).

- \( m \) different risk models
  - \( \mathbf{L}_k \in \mathbb{R}^{N_k \times n} \): Loss matrix for the \( k \)-th model
  - \( \rho_k(\mathbf{L}_k \mathbf{x}) = \sum_{j=1}^{d_k} \gamma_j^{(k)} \text{ES}_{\beta_j^{(k)}}(\mathbf{L}_k \mathbf{x}) \)

- Why bother with multiple risk models?
  - Good, bad, ugly historical return periods
  - Risk models with different periods
  - Robustness to parameters

- Portfolio selection problem

\[
\begin{align*}
\max & \quad \mathbf{\mu}^\top \mathbf{x} - \lambda \|\mathbf{x}\|_1 \\
\text{s.t.} & \quad \rho_k(\mathbf{L}_k \mathbf{x}) \leq \alpha_k, \quad k = 1, \ldots, m, \\
& \quad \mathbf{1}^\top \mathbf{x} = 1, \quad \|\mathbf{x}\|_\infty \leq B.
\end{align*}
\]
LP formulation

- LP duality (Rockafellar and Uryasev)

\[
\text{ES}_\beta(L) = \min_z \left\{ z + \frac{1}{(1 - \beta)N} \cdot \sum_{j=1}^{N} (L_j - z)^+ \right\}
\]
LP formulation

- LP duality (Rockafellar and Uryasev)

\[
\text{ES}_\beta(L) = \min \left\{ z + \frac{1}{(1 - \beta)N} \sum_{j=1}^{N} (L_j - z)^+ \right\}
\]

- Portfolio selection problem

\[
\begin{align*}
\max & \quad \mu^\top x - \lambda \|x\|_1 \\
\text{s.t.} & \quad \sum_{\ell=1}^{d_k} \gamma_{k\ell} \left( z_{k\ell} + \frac{1}{(1 - \beta_{k\ell})N_k} \sum_{j=1}^{N_k} \left( (L_{k\ell}x)_j - z_{k\ell} \right)^+ \right) \leq \alpha_k, \quad \forall k \\
1^\top x & = 1, \quad \|x\|_\infty \leq B.
\end{align*}
\]
LP formulation

• LP duality (Rockafellar and Uryasev)

\[
\text{ES}_\beta(L) = \min_z \left\{ z + \frac{1}{(1 - \beta)N} \cdot \sum_{j=1}^{N} (L_j - z)^+ \right\}
\]

• Portfolio selection problem

\[
\text{max } \mu^\top x - \lambda \|x\|_1
\]
\[
\text{s.t. } \sum_{\ell=1}^{d_k} \gamma_{k\ell} \left( z_{k\ell} + \frac{1}{(1 - \beta_{k\ell})N_k} \sum_{j=1}^{N_k} ((L_{k\ell} x)_j - z_{k\ell})^+ \right) \leq \alpha_k, \quad \forall k
\]
\[
1^\top x = 1, \quad \|x\|_\infty \leq B.
\]

• Complexity of LP = \(\mathcal{O}((mdN + n)^3)\)
  
  • \(n = 100, \ m = 5\) models, \(d = 3, \ N = 10,000\): \(mdN + n = 150,000\).
  
  • LP is very badly ill-conditioned
Penalty formulation

- “Decouple” the risk measures: reduces complexity.
“Decouple” the risk measures: reduces complexity.

Penalty formulation:

\[
\min \eta \left( \lambda \|x\|_1 - \mu^\top x \right) + \left( \max_{1 \leq k \leq m} \{ \rho_k (L_k x) - \alpha_k \} \right)^+ \\
\text{s.t. } 1^\top x = 1, \quad \|x\|_\infty \leq B
\]

Solve for a decreasing sequence of values of \( \eta \)
Penalty formulation

- “Decouple” the risk measures: reduces complexity.

- Penalty formulation:

\[
\min \eta \left( \lambda \|x\|_1 - \mu^\top x \right) + \left( \max_{1 \leq k \leq m} \left\{ \rho_k(L_kx) - \alpha_k \right\} \right)^+
\]

s.t. \( 1^\top x = 1, \quad \|x\|_\infty \leq B \)

- Solve for a decreasing sequence of values of \( \eta \)

- The objective is non-smooth:
  - \( \max\{\cdot\} \) is non-smooth
  - \( \rho_k \) contains ES terms that are non-smooth
Penalty formulation

- “Decouple” the risk measures: reduces complexity.

- Penalty formulation:

\[
\min \eta \left( \lambda \|x\|_1 - \mu^\top x \right) + \left( \max_{1 \leq k \leq m} \{ \rho_k(L_k x) - \alpha_k \} \right)^+ \\
\text{s.t. } 1^\top x = 1, \quad \|x\|_\infty \leq B
\]

- Solve for a decreasing sequence of values of \( \eta \)

- The objective is non-smooth:
  - \( \max \{ \cdot \} \) is non-smooth
  - \( \rho_k \) contains ES terms that are non-smooth

- Sub-gradient algorithms very slow!
Smooth the non-smooth function $g(x)$

- Smooth the max term: $\max_{1 \leq k \leq m}\{x_k\}$

  \[
  \max_{1 \leq k \leq m}\{x_k\} = \max \sum_{k=1}^{m} u_k x_k \\
  \text{s.t. } 1^\top u = 1, u \geq 0.
  \]
Smooth the non-smooth function $g(x)$

- Smooth the max term: $\max_{1 \leq k \leq m} \{x_k\}$

$$
\text{max}_\delta(x) = \max \sum_{k=1}^{m} u_k x_k - \frac{\delta}{2} \|u\|_2^2 \\
\text{s.t. } 1^\top u = 1, u \geq 0.
$$
Smooth the non-smooth function $g(x)$

- Smooth the max term: $\max_{1 \leq k \leq m}\{x_k\}$
  
  $$\max_{\delta}(x) = \max \sum_{k=1}^{m} u_k x_k - \frac{\delta}{2} \|u\|_2^2$$

  s.t. $1^\top u = 1, u \geq 0$.

- Easy QP: can be solved by a 1 dimensional search
- $\nabla \max_{\delta}(x) = u^*$

Smoothed function $g_{\nu}(x) = \max_{\delta}(\{d_k \sum_{\ell=1}^L \gamma_k \ell \cdot \max_{\beta, \nu}(\ell)\} - \alpha)$

$m_k = 1$
Smooth the non-smooth function $g(x)$

- Smooth the max term: $\max_{1 \leq k \leq m} \{x_k\}$

$$
\max_{\delta}(x) = \max \sum_{k=1}^{m} u_k x_k - \frac{\delta}{2} \|u\|_2^2 \\
\text{s.t. } 1^T u = 1, u \geq 0.
$$

- Easy QP: can be solved by a 1 dimensional search
  - $\nabla \max_{\delta}(x) = u^*$

- Smooth the Expected Shortfall term

$$
\text{ES}_{\beta, \nu}(\ell) = \max \ q^T \ell - \frac{\nu}{2} \|q\|_2^2 \\
\text{s.t. } 1^T u = 1, 0 \leq q \leq \frac{1}{(1-\beta)N} \cdot 1.
$$

- Harder QP: can still be solved by a 1 dimensional search
  - $\nabla \text{ES}_{\beta, \nu}(\ell) = q^*$
Smooth the non-smooth function \( g(x) \)

- Smooth the max term: \( \max_{1 \leq k \leq m} \{ x_k \} \)
  \[
  \max_{\delta}(x) = \max \sum_{k=1}^{m} u_k x_k - \frac{\delta}{2} \| u \|_2^2
  \]
  s.t. \( 1^\top u = 1, u \geq 0 \).

- Easy QP: can be solved by a 1 dimensional search
  \( \nabla \max_{\delta}(x) = u^* \)

- Smooth the Expected Shortfall term
  \[
  \text{ES}_{\beta,\nu}(\ell) = \max q^\top \ell - \frac{\nu}{2} \| q \|_2^2
  \]
  s.t. \( 1^\top u = 1, 0 \leq q \leq \frac{1}{(1-\beta)N} \cdot 1 \).

- Harder QP: can still be solved by a 1 dimensional search
  \( \nabla \text{ES}_{\beta,\nu}(\ell) = q^* \)

- Smoothed function
  \[
  g_{\nu\delta}(x) = \max_{\delta} \left( \sum_{\ell=1}^{d_k} \gamma_{k\ell} \text{ES}_{\beta_{k\ell},\nu}(L_k x) - \alpha \right)_{k=1}^{m}
  \]
FISTA for fixed $\eta$

- **Smoothed** Penalty formulation:

$$\min \eta \left( \lambda \|x\|_1 - \mu^T x \right) + g_{\nu\delta}(x)$$

subject to

$$1^T x = 1, \|x\|_\infty \leq B.$$
FISTA for fixed $\eta$

- **Smoothed** Penalty formulation:

$$\min \eta \left( \lambda \|x\|_1 - \mu^\top x \right) + g_{\nu \delta}(x)$$

s.t. $1^\top x = 1, \|x\|_\infty \leq B$,

- **Proximal** gradient algorithm: In every iteration we need to solve

$$\min \eta \lambda \|x\|_1 + h(x; y^{(k)})$$

s.t. $1^\top x = 1, \|x\|_\infty \leq B$

- $h(x; y^{(k)}) = (\eta \mu + \nabla g_{\nu \delta}(y^{(k)}))^\top (x - y^{(k)}) + \frac{L}{2} \|x - y^{(k)}\|_2^2$

- $\nabla g_{\nu \delta}: O(dmN)$ complexity
FISTA for fixed $\eta$

- **Smoothed** Penalty formulation:
  \[
  \min \eta \left( \lambda \|x\|_1 - \mu ^\top x \right) + g_{\nu \delta}(x)
  \]
  \[
  \text{s.t. } 1^\top x = 1, \|x\|_{\infty} \leq B,
  \]

- **Proximal** gradient algorithm: In every iteration we need to solve
  \[
  \min \eta \lambda \|x\|_1 + h(x; y^{(k)})
  \]
  \[
  \text{s.t. } 1^\top x = 1, \|x\|_{\infty} \leq B
  \]
  \[
  h(x; y^{(k)}) = (\eta \mu + \nabla g_{\nu \delta}(y^{(k)}))^\top (x - y^{(k)}) + \frac{L}{2} \|x - y^{(k)}\|_2^2
  \]
  \[
  \nabla g_{\nu \delta}: O(dmN) \text{ complexity}
  \]
  \[
  \ell_1\text{-penalized separable QP}
  \]
  \[
  \text{Number of variables equal to number of assets}
  \]
  \[
  \text{Efficiently solvable even with side constraints}
  \]
FISTA for fixed $\eta$

- **Smoothed** Penalty formulation:
  \[
  \min \eta \left( \lambda \|x\|_1 - \mu^\top x \right) + g_{\nu\delta}(x)
  \]
  s.t. $1^\top x = 1, \|x\|_\infty \leq B$,

- **Proximal** gradient algorithm: In every iteration we need to solve
  \[
  \min \eta \lambda \|x\|_1 + h(x; y^{(k)})
  \]
  s.t. $1^\top x = 1, \|x\|_\infty \leq B$

- $h(x; y^{(k)}) = (\eta \mu + \nabla g_{\nu\delta}(y^{(k)}))^\top (x - y^{(k)}) + \frac{L}{2} \|x - y^{(k)}\|_2^2$

- $\nabla g_{\nu\delta}$: $O(dmN)$ complexity

- $\ell_1$-penalized separable QP
  - Number of variables equal to number of assets
  - Efficiently solvable even with side constraints

- Complexity: $O(mdN + n^3)$ compared to $O((mdN + n)^3)$
Other portfolio selection problems

- Weighted sparse mean-spectral risk

\[
\max \mu^\top x - \lambda \|x\|_1 - \sum_{k=1}^m \eta_k \rho_k(L_k x)
\]
\[
\text{s.t. } 1^\top x = 1, \|x\|_\infty \leq B.
\]

- Sparse mean-max spectral risk portfolio selection problem

\[
\max \mu^\top x - \lambda \|x\|_1 - \eta \left( \max_{k=1,\ldots,m} \rho_k(L_k x) \right)
\]
\[
\text{s.t. } 1^\top x = 1, \|x\|_\infty \leq B.
\]

- Suppose Kusuoka representation set \( \Gamma = \text{conv} (\gamma_1, \ldots, \gamma_m) \). Then

\[
\rho(\tilde{X}) = \max_{1 \leq k \leq m} \{ M_{\gamma_k}(\tilde{X}) \}
\]

Method extends to law-invariant coherent risk measures.
Problem scaling results

- Compared the performance of our FISTA based code and **Gurobi**

<table>
<thead>
<tr>
<th>assets</th>
<th>risk models</th>
<th># ES</th>
<th>samples</th>
<th>max err(%)</th>
<th>Average CPU Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gurobi</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>3</td>
<td>100</td>
<td>0.22</td>
<td>0.118</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>3</td>
<td>500</td>
<td>0.41</td>
<td>1.222</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1000</td>
<td>0.01</td>
<td>1.994</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1500</td>
<td>0.00</td>
<td>2.924</td>
</tr>
<tr>
<td>100</td>
<td>5</td>
<td>3</td>
<td>1000</td>
<td>0.00</td>
<td>29.494</td>
</tr>
<tr>
<td>100</td>
<td>5</td>
<td>3</td>
<td>5000</td>
<td>0.00</td>
<td>242.974</td>
</tr>
<tr>
<td>100</td>
<td>5</td>
<td>3</td>
<td>10000</td>
<td>0.00</td>
<td>373.729</td>
</tr>
<tr>
<td>100</td>
<td>5</td>
<td>3</td>
<td>15000</td>
<td><strong>x.xxx</strong></td>
<td><strong>x.xxx</strong></td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>3</td>
<td>5000</td>
<td>1.65</td>
<td>38378.000</td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>3</td>
<td>10000</td>
<td><strong>x.xxx</strong></td>
<td><strong>x.xxx</strong></td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>3</td>
<td>15000</td>
<td><strong>x.xxx</strong></td>
<td><strong>x.xxx</strong></td>
</tr>
</tbody>
</table>

- **x.xxx** = Gurobi exited without computing a solution
Derivative portfolio selection


- 4 correlated assets and
  - 12 vanilla European calls and puts
  - 12 binary European calls and puts
  - The option prices are computed using the Black-Scholes formula.

- **Nominal** portfolio:
  - $N = 25,000$ samples using **one** risk model: $\sigma^2 = \sigma_0^2$
Derivative portfolio selection


- 4 correlated assets and
  - 12 vanilla European calls and puts
  - 12 binary European calls and puts
  - The option prices are computed using the Black-Scholes formula.

- **Nominal** portfolio:
  - $N = 25,000$ samples using **one** risk model: $\sigma^2 = \sigma_0^2$

- **Robust** portfolio
  - $N = 25,000$ with **three** risk models: $\sigma^2 = [1.05, 1, 0.95] \sigma_0^2$
  - This problem is intractable for Gurobi.
Derivative portfolio selection


- 4 correlated assets and
  - 12 vanilla European calls and puts
  - 12 binary European calls and puts
  - The option prices are computed using the Black-Scholes formula.

- **Nominal** portfolio:
  - \( N = 25,000 \) samples using **one** risk model: \( \sigma^2 = \sigma_0^2 \)

- **Robust** portfolio
  - \( N = 25,000 \) with **three** risk models: \( \sigma^2 = [1.05, 1, 0.95]\sigma_0^2 \)
  - This problem is intractable for Gurobi.

- Test: Risk budget violation on 10 sets of \( N = 25,000 \) samples.
Derivative portfolio: Numerical results

- Low risk $\alpha = 1$ and high sparsity $\lambda = \lambda_0$

<table>
<thead>
<tr>
<th>prob</th>
<th>$\mu^\top x$</th>
<th>$\rho_{test} &gt; \alpha$</th>
<th>$\max \rho_{test}$</th>
<th>$\mu_5^\top x$</th>
<th>$x_i \neq 0$</th>
<th>cpu time</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom</td>
<td>0.0181</td>
<td>5</td>
<td>1.0133</td>
<td>0.0180</td>
<td>7</td>
<td>352.69</td>
</tr>
<tr>
<td>rob</td>
<td>0.0176</td>
<td>0</td>
<td>0.9831</td>
<td>0.0175</td>
<td>4</td>
<td>887.13</td>
</tr>
</tbody>
</table>

- Low risk $\alpha = 1$ and low sparsity $\lambda = \lambda_0/16$

<table>
<thead>
<tr>
<th>prob</th>
<th>$\mu^\top x$</th>
<th>$\rho_{test} &gt; \alpha$</th>
<th>$\max \rho_{test}$</th>
<th>$\mu_5^\top x$</th>
<th>$x_i \neq 0$</th>
<th>cpu time</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom</td>
<td>0.0228</td>
<td>6</td>
<td>1.0197</td>
<td>0.0193</td>
<td>42</td>
<td>1078.80</td>
</tr>
<tr>
<td>rob</td>
<td>0.0221</td>
<td>0</td>
<td>0.9857</td>
<td>0.0190</td>
<td>48</td>
<td>3130.00</td>
</tr>
</tbody>
</table>

- High risk $\alpha = 3$ and low sparsity $\lambda = \lambda_0/16$

<table>
<thead>
<tr>
<th>prob</th>
<th>$\mu^\top x$</th>
<th>$\rho_{test} &gt; \alpha$</th>
<th>$\max \rho_{test}$</th>
<th>$\mu_5^\top x$</th>
<th>$x_i \neq 0$</th>
<th>cpu time</th>
</tr>
</thead>
<tbody>
<tr>
<td>nom</td>
<td>0.0645</td>
<td>6</td>
<td>3.0697</td>
<td>0.0550</td>
<td>68</td>
<td>1066.90</td>
</tr>
<tr>
<td>rob</td>
<td>0.0626</td>
<td>0</td>
<td>2.9709</td>
<td>0.0535</td>
<td>64</td>
<td>2898.00</td>
</tr>
</tbody>
</table>
Conclusions

- Fast first-order algorithm for portfolio selection with multiple spectral risk constraints

- Each step of algorithm ≡ separable convex QP in number of assets.

- The algorithm very efficient both in theory and in practice.
  - Can prove a theoretical complexity bound
  - Tested algorithm with $n = 200$, $N = 25,000$ and $m = 5$ risk models

- Even MATLAB implementation is superior to state-of-art LP solver!