A Scalable Algorithm for Sparse Portfolios

The Sparse Markowitz Model [1]

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2\gamma}\boldsymbol{x}^\top\boldsymbol{x} + \frac{\sigma}{2}\boldsymbol{x}^\top\boldsymbol{\Sigma}\boldsymbol{x} - \boldsymbol{\mu}^\top\boldsymbol{x}, \quad (1)$$

$$\text{s.t. } \boldsymbol{l} \leq \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{u}, \boldsymbol{e}^\top\boldsymbol{x} = 1, \boldsymbol{x} \geq \boldsymbol{0}. \quad (2)$$

- Where $||\boldsymbol{x}||_0 \leq k$, to reduce transaction fees.
- γ enforces sparsity in big-M free manner.
- Big-M approach: introduce binary z where $x \leq z, e^\top z \leq k$. Yields weaker relaxations.

Main Contributions

- A **tractable** nonlinear transformation which decouples the discrete, continuous.
- A **scalable** cutting-plane method which solves real-world problem instances, including the S&P 500, Wilshire 5000.
- A **generalizable** approach which also solves facility location, network design, unit commitment, sparse learning problems [4].

Overview of the Approach

- Introduce new variables $\hat{x}_i = z_i x_i$, $\boldsymbol{z} \in \{0,1\}^n$, $\boldsymbol{Z} = \mathrm{Diag}(\boldsymbol{z})$.
- Perform a non-linear reformulation of (1) into:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n} \frac{1}{2\gamma}\boldsymbol{x}^\top\boldsymbol{x} + \frac{\sigma}{2}\boldsymbol{x}^\top\boldsymbol{Z}\boldsymbol{\Sigma}\boldsymbol{Z}\boldsymbol{x} - \boldsymbol{\mu}^\top\boldsymbol{Z}\boldsymbol{x}. \quad (5)$$

- Strengthened the formulation by omitting $\boldsymbol{Z}$ in $\frac{1}{2\gamma}\boldsymbol{x}^\top\boldsymbol{x}$ (if $z_i = 0$, best choice of $x_i = 0$).
- This is equivalent to $(\boldsymbol{A})$, via duality.

A Saddle-Point Reformulation

$$(\boldsymbol{A}): \min_{\boldsymbol{z}\in\{0,1\}^n : \sum_{i=1}^n z_i \leq k} f(\boldsymbol{z}) \text{ where } f(\boldsymbol{z}) := \max_{\boldsymbol{w}\in\mathbb{R}^n} h(\boldsymbol{w}) - \frac{\gamma}{2}\sum_i z_i w_i^2 \quad (3)$$

- $h(\boldsymbol{w})$ is concave in dual vars.
- Subgradients of $f(\boldsymbol{z})$ are given by $\frac{\partial f(\boldsymbol{z})}{\partial z_i} = \frac{-\gamma}{2} w_i^2$.
- Solvable via outer-approximation, using lazy callbacks.
- Invoking duality repeatedly gives a SOCP relaxation of (A); rediscovery of perspective relaxation of [2].

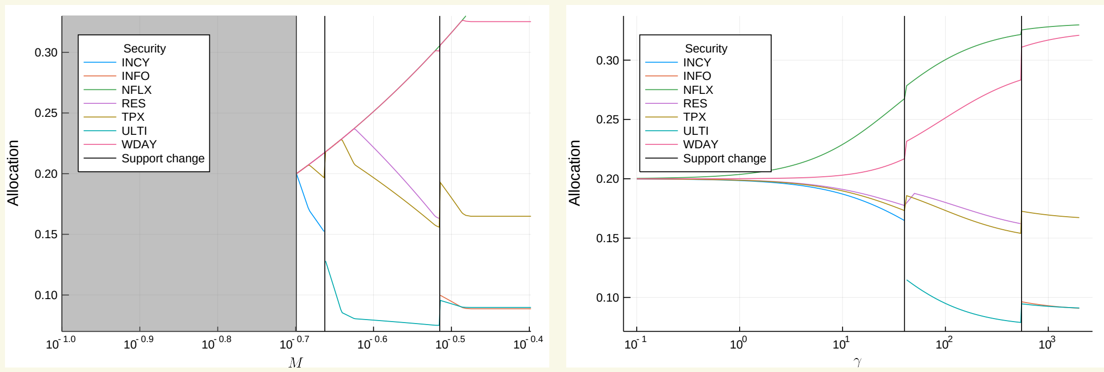
Analysis of SOCP Relaxation Quality [4]

- If $z^\star$ solves the SOCP relaxation of (**A**), a random rounding $z_j \sim \mathrm{Bernoulli}(z_i^\star)$ is ϵ-optimal with probability at least

$$1 - |\mathcal{R}| \exp\left(\frac{-2\epsilon^2}{\gamma^2 L^4 |\mathcal{R}|^2}\right) \quad (4)$$

- $|\mathcal{R}|$ is no. fractional entries in $z^\star$. L is bound on $|w_i^\star(z)|$.
- Randomized rounding scheme bounds the SOCP gap via the probabilistic method. The gap is $O(\frac{1}{\gamma}|\mathcal{R}| \ln \sqrt{|\mathcal{R}|})$, or smaller.

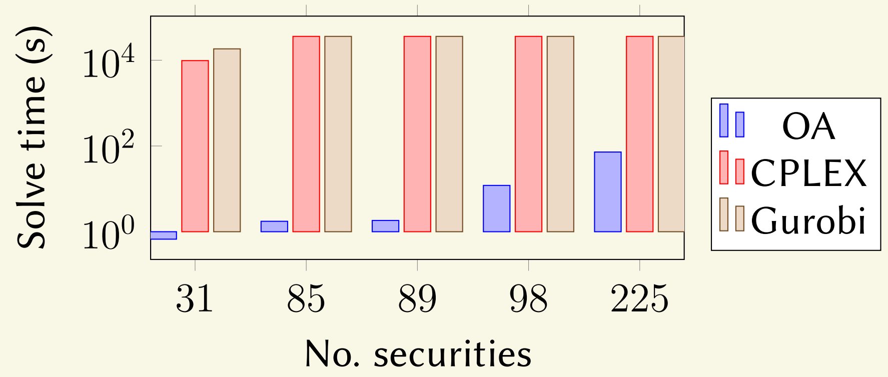
Big-M vs. Ridge Regularization [4]



- Sensitivity to M, γ for Russell 1000 with $k = 5$.
- M, γ play fundamentally same role. But γ-regularization is smoother, always feasible, while M can induce infeasibility.

Comparison With State-of-the-Art

- Comparison vs. big-M for OR-lib problems [3].



- 4 **orders of magnitude** speedup over big-M.

The Edge of Our Approach

Set of securities	Sparsity	Min Time (s)	Max Time (s)
S&P 500	$k = 50$	105 s	374 s
S&P 500	$k = 200$	169 s	370 s
Russell 1000	$k = 50$	$1,732$ s	$3,320$ s
Russell 1000	$k = 200$	$3,647$ s	4642 s
Wilshire 5000	$k = 50$	18 s	457 s
Wilshire 5000	$k = 200$	18 s	104 s

- We used a different solver for Wilshire 5000 SOCP bound.

References

[1] D. Bertsimas and R. Cory-Wright, "A scalable algorithm for sparse and robust portfolios," *arXiv preprint arXiv:1811.00138*, 2018.

[2] M. S. Aktürk, A. Atamtürk, and S. Gürel, "A strong conic quadratic reformulation for machine-job assignment with controllable processing times," *Oper. Res. Letters*, 2009.

[3] T.-J. Chang *et al.*, "Heuristics for cardinality constrained portfolio optimisation," *Comp. Oper. Res.*, 2000.

[4] D. Bertsimas, R. Cory-Wright, and J. Pauphilet, "A unified approach to mixed-integer optimization: Nonlinear formulations and scalable algorithms," *arXiv preprint arXiv:1907.02109*, 2019.

Dimitris Bertsimas, Ryan Cory-Wright

(Thanks to Jean Pauphilet) email: ryancw@mit.edu, website: ryancorywright.github.io

Operations Research Center, Massachusetts Institute of Technology

OPERATIONS RESEARCH CENTER