Mechanical & Industrial Engineering
UNIVERSITY OF TORONTO

Generalized Risk Parity Portfolio Optimization: An ADMM Approach

CASCON | EVOKE

CASCON x EVOKE
Markham, ON
November 6, 2019

Giorgio Costa

University of Toronto

Department of Mechanical and Industrial Engineering

Joint work with Roy H. Kwon

Introduction

What is this presentation about?

- We wish to construct an optimal portfolio

  $\Rightarrow$ High return

  $\Rightarrow$ Low risk

  $\Rightarrow$ Well-diversified

  $\Rightarrow$ Flexibility for the investor

What is this presentation about?

- We wish to construct an optimal portfolio
  - $\Rightarrow$ High return
  - $\Rightarrow$ Low risk
  - $\Rightarrow$ Well-diversified
  - $\Rightarrow$ Flexibility for the investor
- Meeting these criteria is **difficult**
  - $\Rightarrow$ It may lead to non-convex problems

Could we design a model that:

- **Meets the investor's criteria**
- **Addresses non-convexity**

Assets and Portfolios

- An asset i has some expected return μ_i and variance (risk) σ_i^2
- A portfolio $\boldsymbol{x} \in \mathbb{R}^n$ is a collection of n financial assets

  $\Rightarrow x_i$ is our weight on asset i

Assets and Portfolios

- An asset i has some expected return μ_i and variance (risk) σ_i^2
- A portfolio $\boldsymbol{x} \in \mathbb{R}^n$ is a collection of n financial assets

  $\Rightarrow$ x_i is our weight on asset i

- The relevant measures of risk and return are

Assets	Portfolio
$\Rightarrow$ **Return**: $\boldsymbol{\mu} \in \mathbb{R}^n$	$\Rightarrow$ **Return**: $\mu_p = \boldsymbol{\mu}^T \boldsymbol{x}$
$\Rightarrow$ **Risk**: $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$	$\Rightarrow$ **Risk**: $\sigma_p^2 = \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x}$

Mean–Variance Optimization (MVO)

- Introduced by Markowitz (1952),

$$
\begin{aligned}
\min_{\boldsymbol{x}} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} - \lambda \boldsymbol{\mu}^T \boldsymbol{x} && \text{Min. risk and max. return} \\
\text{s.t.} \quad & \mathbf{1}^T \boldsymbol{x} = 1 && \text{Budget constraint} \\
& (\boldsymbol{x} \geq 0) && \text{Disallow short sales (optional)}
\end{aligned}
$$

Portfolio Optimization

Mean–Variance Optimization (MVO)

- Introduced by Markowitz (1952),

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^T \Sigma \boldsymbol{x} - \lambda \boldsymbol{\mu}^T \boldsymbol{x} && \text{Min. risk and max. return} \\ \text{s.t.} \quad & \mathbf{1}^T \boldsymbol{x} = 1 && \text{Budget constraint} \\ & (\boldsymbol{x} \geq 0) && \text{Disallow short sales (optional)} \end{aligned}$$

- Weaknesses:

  $\Rightarrow$ May lead to over-concentrated portfolios

  $\Rightarrow$ The estimated parameter $\boldsymbol{\mu}$ is very **noisy**

What is risk parity?

- Risk parity seeks to find portfolios based on a risk-weighted basis
- Does not require estimated returns as an input, improving stability
- Each asset contributes the same level of risk
- The resulting portfolio is well-diversified

Measuring the risk contribution per asset

- Decompose the portfolio variance

$$\sigma_p^2 = \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} = \sum_{i=1}^{n} x_i (\boldsymbol{\Sigma} \boldsymbol{x})_i$$

$\Rightarrow x_i(\boldsymbol{\Sigma}\boldsymbol{x})_i$ is the individual **risk contribution** of asset i

Measuring the risk contribution per asset

- Decompose the portfolio variance

$$\sigma_p^2 = \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} = \sum_{i=1}^{n} x_i (\boldsymbol{\Sigma} \boldsymbol{x})_i$$

$\Rightarrow x_i(\boldsymbol{\Sigma}\boldsymbol{x})_i$ is the individual risk contribution of asset i

- **Risk parity**: Take a least-squares approach

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} \big(x_i (\boldsymbol{\Sigma}\boldsymbol{x})_i - x_j (\boldsymbol{\Sigma}\boldsymbol{x})_j \big)^2 \\ \text{s.t.} \quad & \mathbf{1}^T \boldsymbol{x} = 1 \\ & \boldsymbol{x} \geq \mathbf{0} \end{aligned}$$

Non-convexity of risk parity

- **Problem**: The objective is non-convex

  $\Rightarrow$ In standard quadratic notation: $x_i(\boldsymbol{\Sigma}\boldsymbol{x})_i = \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x}$

  $\Rightarrow$ $\boldsymbol{A}_i \in \mathbb{R}^{n\times n}$ captures the individual risk contribution of asset i

  $\Rightarrow$ The matrices $\boldsymbol{A}_i$ are indefinite

Non-convexity of risk parity

- **Problem**: The objective is non-convex

  $\Rightarrow$ In standard quadratic notation: $x_i(\boldsymbol{\Sigma x})_i = \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x}$

  $\Rightarrow$ $\boldsymbol{A}_i \in \mathbb{R}^{n \times n}$ captures the individual risk contribution of asset i

  $\Rightarrow$ The matrices $\boldsymbol{A}_i$ are indefinite

- **Solution**: Disallow short selling

  $\Rightarrow$ This limits the investor's possibilities

Formulate a new optimization problem

Formulate a new optimization problem

- We seek a desirable portfolio

  $\Rightarrow$ Minimize risk and maximize return

  $\Rightarrow$ Risk-based diversification

  $\Rightarrow$ Short selling allowed

Generalized Risk Parity

Formulate a new optimization problem

- We seek a desirable portfolio

  ⇒ Minimize risk and maximize return

  ⇒ Risk-based diversification

  ⇒ Short selling allowed

- Generalized Risk Parity (GRP)

$$
\begin{aligned}
\min_{\boldsymbol{x},\ \theta} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} - \lambda \boldsymbol{\mu}^T \boldsymbol{x} \\
\text{s.t.} \quad & (1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad i = 1, \ldots, n \\
& \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad i = 1, \ldots, n \\
& \boldsymbol{1}^T \boldsymbol{x} = 1
\end{aligned}
$$

Generalized Risk Parity

A closer look at the risk diversification constraints

$$
\begin{aligned}
(1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad & i = 1, ..., n \\
\boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad & i = 1, ..., n
\end{aligned}
$$

- $\theta \in \mathbb{R}$ is an auxiliary variable
- $c \in \mathbb{R}_+$ is a user-defined risk diversification parameter

Generalized Risk Parity

A closer look at the risk diversification constraints

$$(1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad i = 1, ..., n$$
$$\boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad i = 1, ..., n$$

- $\theta \in \mathbb{R}$ is an auxiliary variable.
- $c \in \mathbb{R}_+$ is a user-defined risk diversification parameter.

  $\Rightarrow c = 0$ enforces perfect risk parity

  $\Rightarrow c > 1$ reverts to MVO

Adding robustness to the portfolio return

- μ is a **noisy** estimate

Generalized Risk Parity

Adding robustness to the portfolio return

- μ is a **noisy** estimate

- Robust GRP

$$
\begin{aligned}
\min_{\boldsymbol{x},\ \theta} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} - \lambda \left(\boldsymbol{\mu}^T \boldsymbol{x} - \omega \| \boldsymbol{\Omega}^{1/2} \boldsymbol{x} \|_2 \right) \\
\text{s.t.} \quad & (1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad i = 1, ..., n \\
& \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad i = 1, ..., n \\
& \mathbf{1}^T \boldsymbol{x} = 1
\end{aligned}
$$

- $\boldsymbol{\Omega} \in \mathbb{R}^{n \times n}$ and $\omega \in \mathbb{R}_+$ quantify the estimation error around $\boldsymbol{\mu}$

The issue of non-convexity still remains

$$
\begin{aligned}
(1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} &\geq 0, \quad i = 1, ..., n \\
\boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta &\geq 0, \quad i = 1, ..., n
\end{aligned}
$$

The issue of non-convexity still remains

$$(1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad i = 1, ..., n$$
$$\boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad i = 1, ..., n$$

- **Relax the problem into a Semidefinite Program (SDP)**

Semidefinite relaxation

- Introduce a new variable $\boldsymbol{X} \in \mathbb{R}^{n \times n}$

  $\Rightarrow$ **Non-convex**: $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$

Semidefinite relaxation

- Introduce a new variable $\boldsymbol{X} \in \mathbb{R}^{n\times n}$

$\Rightarrow$ Non-convex: $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$

$\Rightarrow$ **Convex relaxation**: $\boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^T$

$$\begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{bmatrix} \succeq \begin{bmatrix} x_1^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_nx_1 & x_nx_2 & \cdots & x_n^2 \end{bmatrix}$$

Semidefinite relaxation of the problem

- Introduce a new variable $\boldsymbol{X} \in \mathbb{R}^{n \times n}$

  $\Rightarrow$ Non-convex: $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$

  $\Rightarrow$ Convex relaxation: $\boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^T$

- Taking the Schur complement, $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^{\boldsymbol{T}} & 1 \end{bmatrix} \succeq 0$.

SDP Relaxation

Semidefinite relaxation of the problem

- Introduce a new variable $\boldsymbol{X} \in \mathbb{R}^{n \times n}$

  $\Rightarrow$ Non-convex: $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$

  $\Rightarrow$ Convex relaxation: $\boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^T$

- Taking the Schur complement, $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^T & 1 \end{bmatrix} \succeq 0$.

- Align the input parameters with the dimensions of $\boldsymbol{Y}$

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{\Sigma} & -\frac{\lambda}{2}\boldsymbol{\mu} \\ -\frac{\lambda}{2}\boldsymbol{\mu}^T & \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{B_i} = \begin{bmatrix} \boldsymbol{A}_i & \boldsymbol{0} \\ \boldsymbol{0}^T & 0 \end{bmatrix} \text{ for } i = 1, ..., n.$$

Relax the GRP model into a SDP

Original $\Rightarrow$

$$
\begin{aligned}
\min_{\boldsymbol{x},\ \theta} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} - \lambda \left(\boldsymbol{\mu}^T \boldsymbol{x} - \omega \|\boldsymbol{\Omega}^{1/2} \boldsymbol{x}\|_2 \right) \\
\text{s.t.} \quad & (1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad i = 1, ..., n \\
& \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad i = 1, ..., n \\
& \mathbf{1}^T \boldsymbol{x} = 1
\end{aligned}
$$

SDP Relaxation

Relax the GRP model into a SDP

Original $\Rightarrow$

$$\begin{aligned}\min_{\boldsymbol{x},\ \theta} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} - \lambda \left(\boldsymbol{\mu}^T \boldsymbol{x} - \omega \| \boldsymbol{\Omega}^{1/2} \boldsymbol{x} \|_2 \right) \\ \text{s.t.} \quad & (1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad i = 1, \ldots, n \\ & \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad i = 1, \ldots, n \\ & \mathbf{1}^T \boldsymbol{x} = 1\end{aligned}$$

SDP $\Rightarrow$

$$\begin{aligned}\min_{\boldsymbol{Y},\ \theta} \quad & \mathrm{Tr}(\boldsymbol{Q}\boldsymbol{Y}) + \lambda\omega \| \boldsymbol{\Omega}^{1/2} \mathbf{Y}_{1:n,n+1} \|_2 \\ \text{s.t.} \quad & (1+c)\theta - \mathrm{Tr}(\boldsymbol{B}_i \boldsymbol{Y}) \geq 0, \quad i = 1, \ldots, n \\ & \mathrm{Tr}(\boldsymbol{B}_i \boldsymbol{Y}) - (1-c)\theta \geq 0, \quad i = 1, \ldots, n \\ & \sum_{i=1}^{n} Y_{i,n+1} = 1 \\ & Y_{n+1,n+1} = 1 \\ & \boldsymbol{Y} \succeq 0\end{aligned}$$

Relax the GRP model into a SDP

Original $\Rightarrow$

$$
\begin{aligned}
\min_{\boldsymbol{x},\,\theta} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} - \lambda \left(\boldsymbol{\mu}^T \boldsymbol{x} - \omega \|\boldsymbol{\Omega}^{1/2} \boldsymbol{x}\|_2 \right) \\
\text{s.t.} \quad & (1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad i = 1, \ldots, n \\
& \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad i = 1, \ldots, n \\
& \mathbf{1}^T \boldsymbol{x} = 1
\end{aligned}
$$

SDP $\Rightarrow$

$$
\begin{aligned}
\min_{\boldsymbol{Y},\,\theta} \quad & \left.\operatorname{Tr}(\boldsymbol{Q}\boldsymbol{Y}) + \lambda\omega \|\boldsymbol{\Omega}^{1/2} \mathbf{Y}_{1:n,n+1}\|_2 \quad \right\} f(\boldsymbol{Y}) \\
\text{s.t.} \quad & \left.\begin{aligned}
& (1+c)\theta - \operatorname{Tr}(\boldsymbol{B}_i \boldsymbol{Y}) \geq 0, \quad i = 1, \ldots, n \\
& \operatorname{Tr}(\boldsymbol{B}_i \boldsymbol{Y}) - (1-c)\theta \geq 0, \quad i = 1, \ldots, n \\
& \sum_{i=1}^{n} Y_{i,n+1} = 1 \\
& Y_{n+1,n+1} = 1 \\
& \boldsymbol{Y} \succeq 0
\end{aligned} \right\} \mathcal{S}
\end{aligned}
$$

How do we solve the original problem?

- Recall $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^{\boldsymbol{T}} & 1 \end{bmatrix} \succeq \mathbf{0}$
- If $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$, we recover the original problem

How do we solve the original problem?

- Recall $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^T & 1 \end{bmatrix} \succeq \mathbf{0}$
- If $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$, we recover the original problem
- This is equivalent to a rank-1 constraint

$$\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T \iff \operatorname{rank}(\boldsymbol{Y}) = 1$$

SDP Relaxation

How do we solve the original problem?

- Recall $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^T & 1 \end{bmatrix} \succeq \boldsymbol{0}$

- If $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$, we recover the original problem

- This is equivalent to a rank-1 constraint

$$\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T \iff \operatorname{rank}(\boldsymbol{Y}) = 1$$

- **Idea: Approximate the rank-1 condition**

Sequential Tightening

Approximate the rank-1 condition

- Use the Alternating Direction Method of Multipliers (ADMM)
- Transfer the rank-1 requirement to a new variable $\boldsymbol{Z}$

  $\Rightarrow \boldsymbol{Z} = \boldsymbol{Y}$

  $\Rightarrow \mathrm{rank}(\boldsymbol{Z}) = 1$

Sequential Tightening

Approximate the rank-1 condition

- Use the Alternating Direction Method of Multipliers (ADMM)
- Transfer the rank-1 requirement to a new variable Z

  $\Rightarrow \boldsymbol{Z} = \boldsymbol{Y}$

  $\Rightarrow \mathrm{rank}(\boldsymbol{Z}) = 1$

- Reformulate as an Augmented Lagrangian

$$L(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\Lambda}) = f(\boldsymbol{Y}) + \frac{\rho}{2}\left\|\boldsymbol{Y} - \left(\boldsymbol{Z} - \frac{1}{\rho}\boldsymbol{\Lambda}\right)\right\|_F^2$$

  $\Rightarrow \rho \in \mathbb{R}_+$ is a tuning parameter

  $\Rightarrow \boldsymbol{\Lambda} \in \mathbb{R}^{(n+1)\times(n+1)}$ is the dual variable of the constraint $\boldsymbol{Z} = \boldsymbol{Y}$

ADMM algorithm

- Iterate through the steps:

  1) Convex $\boldsymbol{Y}$-minimization: $$\boldsymbol{Y}^{k+1} = \operatorname*{argmin}_{\boldsymbol{Y},\theta \, \in \, \mathcal{S}} L(\boldsymbol{Y}, \boldsymbol{Z}^k, \boldsymbol{\Lambda}^k)$$

ADMM algorithm

- Iterate through the steps:

  $\mathbf{1})$ Convex $\boldsymbol{Y}$-minimization: $\boldsymbol{Y}^{k+1} = \underset{\boldsymbol{Y},\theta \,\in\, \mathcal{S}}{\operatorname{argmin}} \, L(\boldsymbol{Y}, \boldsymbol{Z}^k, \boldsymbol{\Lambda}^k)$

  $\mathbf{2})$ Non-convex $\boldsymbol{Z}$-minimization: $\boldsymbol{Z}^{k+1} = \underset{\operatorname{rank}(\boldsymbol{Z})=1}{\operatorname{argmin}} \, L(\boldsymbol{Y}^{k+1}, \boldsymbol{Z}, \boldsymbol{\Lambda}^k)$

ADMM algorithm

- Iterate through the steps:

  $\mathbf{1})$ Convex $\boldsymbol{Y}$-minimization: $\boldsymbol{Y}^{k+1} = \underset{\boldsymbol{Y},\theta \in \mathcal{S}}{\operatorname{argmin}} L(\boldsymbol{Y}, \boldsymbol{Z}^k, \boldsymbol{\Lambda}^k)$

  $\mathbf{2})$ Non-convex $\boldsymbol{Z}$-minimization: $\boldsymbol{Z}^{k+1} = \underset{\operatorname{rank}(\boldsymbol{Z})=1}{\operatorname{argmin}} L(\boldsymbol{Y}^{k+1}, \boldsymbol{Z}, \boldsymbol{\Lambda}^k)$

  $\mathbf{3})$ Dual variable $\boldsymbol{\Lambda}$-update: $\boldsymbol{\Lambda}^{k+1} = \boldsymbol{\Lambda}^k + \rho(\boldsymbol{Y}^{k+1} - \boldsymbol{Z}^{k+1})$

ADMM algorithm

- Iterate through the steps:

$$1)\ \text{Convex } \boldsymbol{Y}\text{-minimization:} \qquad \boldsymbol{Y}^{k+1} = \underset{\boldsymbol{Y},\theta\,\in\,\mathcal{S}}{\operatorname{argmin}}\, L(\boldsymbol{Y}, \boldsymbol{Z}^k, \boldsymbol{\Lambda}^k)$$

$$2)\ \text{Non-convex } \boldsymbol{Z}\text{-minimization:} \qquad \boldsymbol{Z}^{k+1} = \underset{\operatorname{rank}(\boldsymbol{Z})=1}{\operatorname{argmin}}\, L(\boldsymbol{Y}^{k+1}, \boldsymbol{Z}, \boldsymbol{\Lambda}^k)$$

$$3)\ \text{Dual variable } \Lambda\text{-update:} \qquad \boldsymbol{\Lambda}^{k+1} = \boldsymbol{\Lambda}^k + \rho(\boldsymbol{Y}^{k+1} - \boldsymbol{Z}^{k+1})$$

- We can efficiently solve the **non-convex** step

Sequential Tightening

ADMM algorithm

- Iterate through the steps:

$$
\begin{aligned}
&1)\ \text{Convex } \boldsymbol{Y}\text{-minimization:} && \boldsymbol{Y}^{k+1} = \underset{\boldsymbol{Y},\theta\,\in\,\mathcal{S}}{\operatorname{argmin}}\, L(\boldsymbol{Y}, \boldsymbol{Z}^k, \boldsymbol{\Lambda}^k)\\
&2)\ \text{Non-convex } \boldsymbol{Z}\text{-minimization:} && \boldsymbol{Z}^{k+1} = \underset{\operatorname{rank}(\boldsymbol{Z})=1}{\operatorname{argmin}}\, L(\boldsymbol{Y}^{k+1}, \boldsymbol{Z}, \boldsymbol{\Lambda}^k)\\
&3)\ \text{Dual variable } \Lambda\text{-update:} && \boldsymbol{\Lambda}^{k+1} = \boldsymbol{\Lambda}^k + \rho(\boldsymbol{Y}^{k+1} - \boldsymbol{Z}^{k+1})
\end{aligned}
$$

- We can efficiently solve the **non-convex** step

$\Rightarrow$ Singular Value Decomposition of $\boldsymbol{Y}^{k+1} + \frac{1}{\rho}\boldsymbol{\Lambda}^k$

$$\boldsymbol{Z}^{k+1} = s_1 \boldsymbol{v}_1 \boldsymbol{v}_1^T,$$

$\Rightarrow$ $s_1 \in \mathbb{R}$ and $\boldsymbol{v}_1 \in \mathbb{R}^{n+1}$ are the top singular value and vector

ADMM algorithm

- As we iterate, we close the distance between Y and Z
- Z is the closest rank-1 approximation of Y

ADMM algorithm

- As we iterate, we close the distance between Y and Z
- Z is the closest rank-1 approximation of Y
- At convergence, $\mathrm{rank}(\boldsymbol{Y}) = \mathrm{rank}(\boldsymbol{Z}) = 1$

  $\Rightarrow \boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$

  $\Rightarrow$ We solve the original problem

Sequential Tightening

ADMM algorithm

- As we iterate, we close the distance between $\boldsymbol{Y}$ and $\boldsymbol{Z}$
- $\boldsymbol{Z}$ is the closest rank-1 approximation of $\boldsymbol{Y}$
- At convergence, $\mathrm{rank}(\boldsymbol{Y}) = \mathrm{rank}(\boldsymbol{Z}) = 1$

  $\Rightarrow \boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$

  $\Rightarrow$ We solve the original problem

- By tightening the lower bound, we can attain feasibility

  $\Rightarrow$ We converge to a highly quality local optimum

Experimental setup

- Two experiments with $n = 33$ and $n = 50$

- Data

  $\Rightarrow$ U.S. stocks belonging to the S&P 500 index

  $\Rightarrow$ Weekly rates of return from 01-Jan-2007 to 31-Dec-2009

Numerical Experiments

Experimental setup

- Two experiments with $n = 33$ and $n = 50$
- Data
  - $\Rightarrow$ U.S. stocks belonging to the S&P 500 index
  - $\Rightarrow$ Weekly rates of return from 01-Jan-2007 to 31-Dec-2009
- Competing models
  - $\Rightarrow$ Robust MVO
  - $\Rightarrow$ **ADMM**
  - $\Rightarrow$ Non-convex GRP
  - $\Rightarrow$ Non-convex GRP (warm)

Numerical Experiments

Points to keep in mind

- Measures of performance

  $\Rightarrow$ Objective value

  $\Rightarrow$ Coefficient of variation of the asset risk contributions

  $\Rightarrow$ Runtime

Numerical Experiments

Points to keep in mind

- Measures of performance
  - ⇒ **Objective value** $\}$ Lower is better
  - ⇒ Coefficient of variation of the asset risk contributions
  - ⇒ Runtime

Numerical Experiments

Points to keep in mind

- Measures of performance

  $\Rightarrow$ Objective value

  $\Rightarrow$ **Coefficient of Variation (CV)** $\}$ Lower = more diversification

  $\Rightarrow$ Runtime

Points to keep in mind

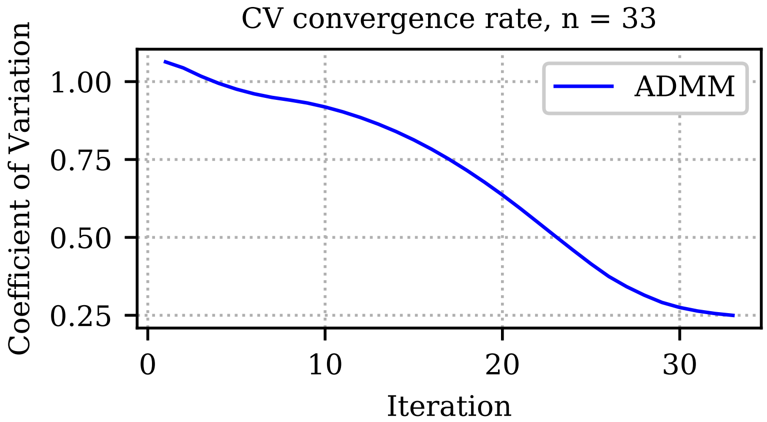
- Measures of performance

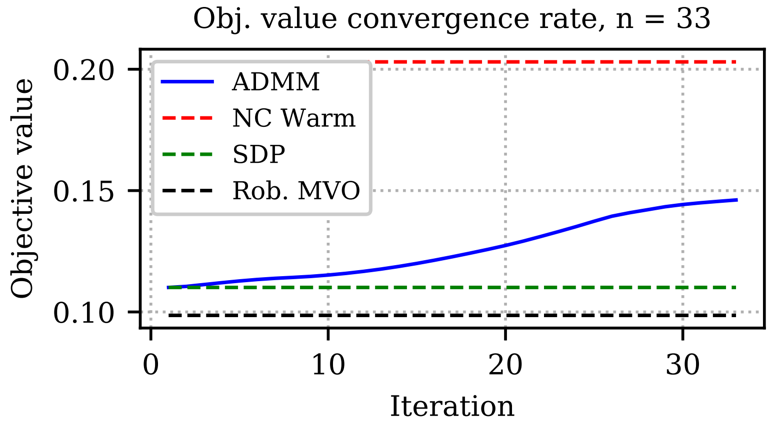
  ⇒ Objective value

  ⇒ Coefficient of Variation (CV)

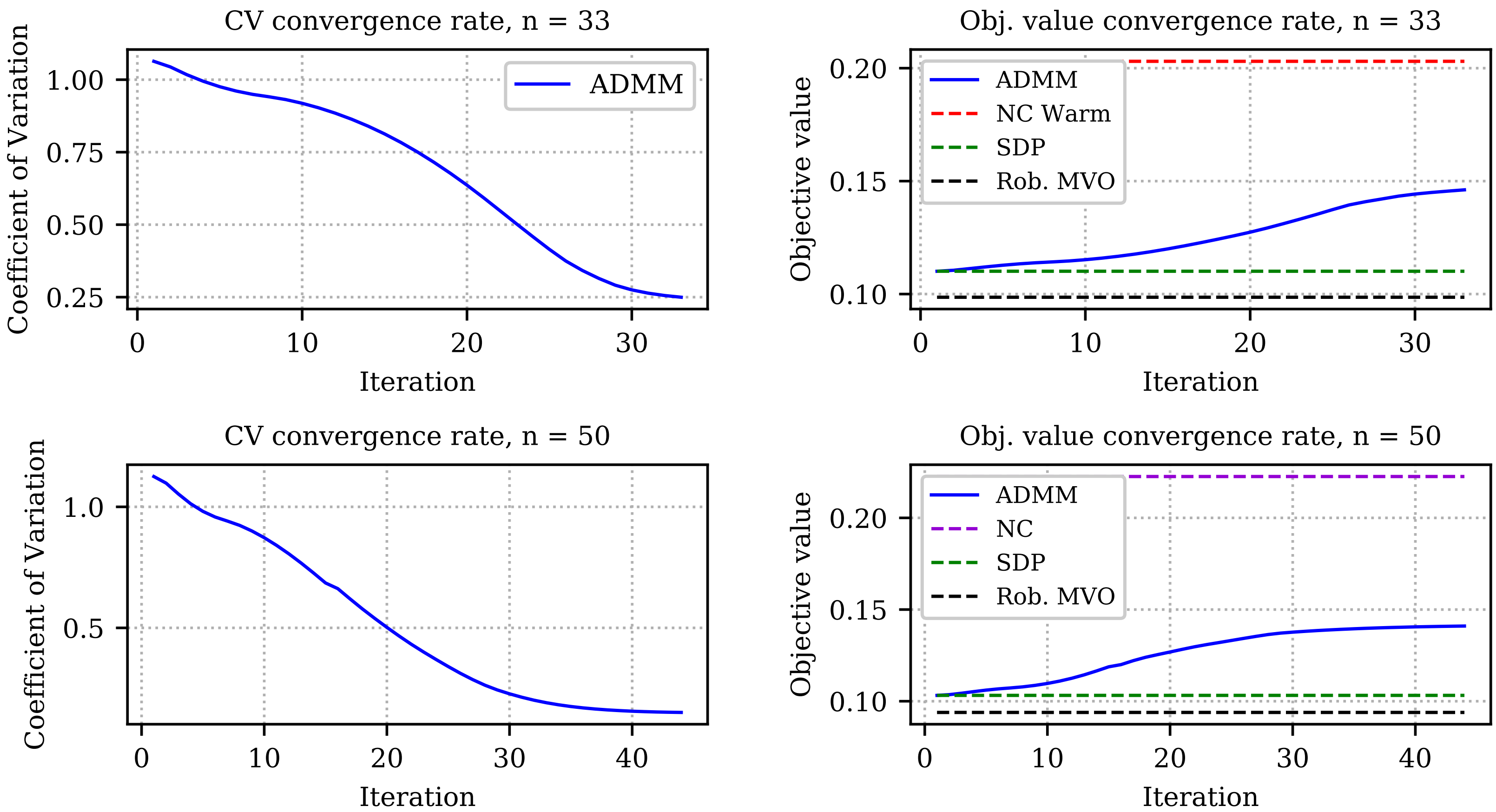
  ⇒ **Runtime** $\}$ Lower is better

CV and Objective Value





CV and Objective Value



Summary of results

	$n = 33$, $c = 0.25$, $\lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.099	0.536	0.203	0.146
CV	1.72	0.258	0.253	0.250
Runtime (sec)	0.026	0.049	0.061	7.16

	$n = 50$, $c = 0.15$, $\lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.094	0.223	0.259	0.141
CV	1.75	0.150	0.150	0.151
Runtime (sec)	0.040	0.101	0.103	42.88

Summary of results

	$n = 33,\ c = 0.25,\ \lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.099	0.536	**0.203**	**0.146**
CV	1.72	0.258	0.253	0.250
Runtime (sec)	0.026	0.049	0.061	7.16

	$n = 50,\ c = 0.15,\ \lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.094	**0.223**	0.259	**0.141**
CV	1.75	0.150	0.150	0.151
Runtime (sec)	0.040	0.101	0.103	42.88

Summary of results

	$n = 33$, $c = 0.25$, $\lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.099	0.536	0.203	0.146
CV	1.72	0.258	**0.253**	**0.250**
Runtime (sec)	0.026	0.049	0.061	7.16

	$n = 50$, $c = 0.15$, $\lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.094	0.223	0.259	0.141
CV	1.75	0.150	**0.150**	**0.151**
Runtime (sec)	0.040	0.101	0.103	42.88

Summary of results

	$n = 33,\ c = 0.25,\ \lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.099	0.536	0.203	0.146
CV	1.72	0.258	0.253	0.250
Runtime (sec)	0.026	**0.049**	0.061	**7.16**
	$n = 50,\ c = 0.15,\ \lambda = 0.1$			
	MVO	Non-Convex	NC (Warm)	**ADMM**
Obj. Value	0.094	0.223	0.259	0.141
CV	1.75	0.150	0.150	0.151
Runtime (sec)	0.040	**0.101**	0.103	**42.88**

Conclusion

What was our goal?

- We wanted to address

  $\Rightarrow$ Risk–return profile

  $\Rightarrow$ Risk-based diversification

  $\Rightarrow$ Short selling (flexibility)

What was our goal?

- We wanted to address
  - $\Rightarrow$ Risk–return profile
  - $\Rightarrow$ Risk-based diversification
  - $\Rightarrow$ Short selling (flexibility)
- Meeting these criteria is difficult
  - $\Rightarrow$ We have a **non-convex** problem

Conclusion

Our contribution

- Proposed a generalized risk parity model
- Addressed the non-convexity of risk parity

  $\Rightarrow$ Imposed a rank-1 constraint via ADMM