Portfolio Optimization

Risk Parity Portfolios

Daniel P. Palomar (2025). *Portfolio Optimization: Theory and Application.* Cambridge University Press.

portfoliooptimizationbook.com

Latest update: April 14, 2025

Outline

Abstract

Markowitz's mean-variance portfolio optimizes the trade-off between expected return and risk, typically measured by variance or volatility. However, quantifying the portfolio risk with a single number is limiting. A more refined approach is to employ a risk profile that quantifies the risk contribution of each constituent asset, enabling better control over portfolio risk diversification, which will be explored in these slides (Palomar 2025, chap. 11).

Outline

Introduction

- **Markowitz's mean-variance portfolio optimization**:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{maximize}} & \boldsymbol{w}^\mathsf{T}\boldsymbol{\mu} - \frac{\lambda}{2}\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w} \\ \text{subject to} & \boldsymbol{w} \in \mathcal{W} \end{array}$$

  where λ is a risk-aversion hyper-parameter and $\mathcal{W}$ is the constraint set, e.g., $\mathcal{W} = \{\boldsymbol{w} \mid \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1, \boldsymbol{w} \geq \mathbf{0}\}$.

- **Limitations of variance as risk measure:**
  - Variance $\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}$ may not yield best out-of-sample performance.
  - Alternative risk measures are considered for improvement.
- **Risk profile characterization:**
  - Beyond a single risk number, assess risk contribution of each asset.
  - Enables control over portfolio risk diversification.
- **Risk parity portfolio:**
  - From simple forms with closed solutions to complex nonconvex formulations.
  - Wide range of numerical algorithms available for implementation.

Outline

From dollar to risk diversification

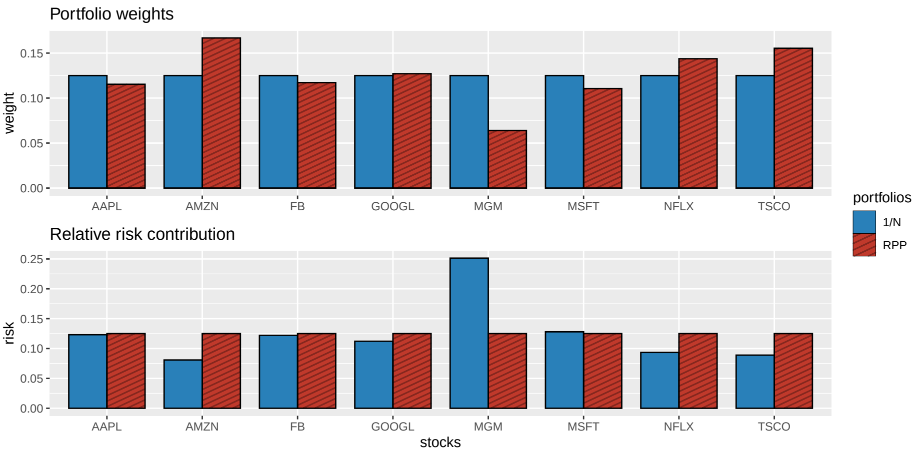
- **Risk parity investment approach:**
  - Focuses on equalizing risk contribution from each asset.
  - Shifts from dollar allocation to risk allocation.
- **Concept of risk diversification:**
  - Aims for assets to contribute equally to overall portfolio risk.
  - Enhances out-of-sample risk control and market downturn resistance.
- **Historical context:**
  - Traditional allocations like 60/40 stock/bond portfolios dominated by equity risk.
  - Risk parity emerged to address risk concentration issues.
- **Development and popularity:**
  - "All Weather" fund by Bridgewater Associates in 1996 initiated the practical application.
  - Term "risk parity" coined by Edward Qian in 2005 (Qian 2005).
  - Gained popularity post-2008 financial crisis.

From dollar to risk diversification

- **Skepticism and debate:**
  - Some managers question its effectiveness across all market conditions.
- **Academic and practitioner interest:**
  - Significant attention and numerous publications.
  - Textbooks for both practical (Qian 2016) and mathematical (Roncalli 2013) perspectives.
- **Illustration of diversification:**
  - $1/N$ portfolio obtain capital allocation diversification, not risk diversification.
  - Risk parity portfolio aims for balanced risk contribution across assets.

From dollar to risk diversification

Portfolio allocation and risk allocation for the $1/N$ portfolio and risk parity portfolio:



Outline

Risk contributions

- **Risk contribution in risk parity portfolio:**
  - Portfolio risk as sum of individual asset risk contributions:

$$\text{portfolio risk} = \sum_{i=1}^{N} \mathrm{RC}_i,$$

  - RC_i: risk contribution of the ith asset.

- **Alternative risk measures:**
  - Volatility, value-at-risk (VaR), conditional VaR (CVaR) are common risk measures.
  - For detailed discussion, see (Palomar 2025, chap. 10).

- **Euler's homogenous function theorem:**
  - For positively homogeneous functions of degree one:

$$f(\boldsymbol{w}) = \sum_{i=1}^{N} w_i \frac{\partial f}{\partial w_i}.$$

  - Applies to volatility, VaR, CVaR, but not variance.

Risk contributions

- **Risk contribution definitions:**
  - Risk Contribution (RC):

$$\mathrm{RC}_i = w_i \frac{\partial f(\boldsymbol{w})}{\partial w_i}.$$

  - Marginal Risk Contribution (MRC):

$$\mathrm{MRC}_i = \frac{\partial f(\boldsymbol{w})}{\partial w_i}.$$

  - Relative Risk Contribution (RRC):

$$\mathrm{RRC}_i = \frac{\mathrm{RC}_i}{f(\boldsymbol{w})},$$

    with $\sum_{i=1}^{N} \mathrm{RRC}_i = 1$.

Volatility risk contributions

- **Risk contribution for volatility:**
  - Risk Contribution (RC):

$$\mathrm{RC}_i = \frac{w_i(\boldsymbol{\Sigma}\boldsymbol{w})_i}{\sqrt{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w}}}$$

  - Marginal Risk Contribution (MRC):

$$\mathrm{MRC}_i = \frac{(\boldsymbol{\Sigma}\boldsymbol{w})_i}{\sqrt{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w}}}$$

  - Relative Risk Contribution (RRC):

$$\mathrm{RRC}_i = \frac{w_i(\boldsymbol{\Sigma}\boldsymbol{w})_i}{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w}}$$

- **Portfolio volatility decomposition:**
  - Portfolio volatility, $\sigma(\boldsymbol{w}) = \sqrt{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w}}$, decomposes as:

$$\sigma(\boldsymbol{w}) = \sum_{i=1}^{N} w_i \frac{\partial \sigma}{\partial w_i} = \sum_{i=1}^{N} \frac{w_i(\boldsymbol{\Sigma}\boldsymbol{w})_i}{\sqrt{\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w}}}$$

Outline

Problem formulation

- **Risk parity portfolio (RPP) or equal risk portfolio (ERP):**
  - Requires equal risk contributions from all assets:
$$\mathrm{RRC}_i = \frac{w_i(\boldsymbol{\Sigma}\boldsymbol{w})_i}{\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}} = \frac{1}{N}, \qquad i = 1,\dots,N.$$
  - Contrasts with the $1/N$ equally weighted portfolio (EWP) that equalizes dollar allocation.
- **Optimality under certain conditions:**
  - If assets have similar Sharpe ratios and correlations, RPP can align with Markowitz's mean-variance optimization.
  - RPP is unique and falls between minimum variance and equally weighted portfolios.
- **Risk budgeting portfolio (RBP):**
  - Allows for a specified risk profile allocation:
$$\mathrm{RRC}_i = \frac{w_i(\boldsymbol{\Sigma}\boldsymbol{w})_i}{\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}} = b_i, \qquad i = 1,\dots,N,$$
  - $\boldsymbol{b} = (b_1,\dots,b_N)$ represents the desired risk profile, normalized to sum to 1.

Problem formulation

- **Formulation of RBP:**
  - Find $\boldsymbol{w} \geq \mathbf{0}$, with $\mathbf{1}^\mathsf{T}\boldsymbol{w} = 1$, that satisfies:

$$w_i(\boldsymbol{\Sigma}\boldsymbol{w})_i = b_i\,\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}, \qquad i = 1, \dots, N.$$

  - This is a feasibility problem with constraints but no explicit objective.

- **Approaches to solving RBP:**
  - Naive diagonal formulation.
  - Vanilla convex formulation.
  - General nonconvex formulation.

- **Practical implementation:**
  - R package `riskParityPortfolio`
  - Python package `riskparityportfolio`

Formulation with shorting

- **Typical RPP constraints:**
  - No shorting allowed: $\boldsymbol{w} \geq \mathbf{0}$.
  - Shorting introduces complexity in resolution methods.
- **Shorting pattern known a priori:**
  - If shorting pattern is predefined, problem simplification is possible.
  - $\boldsymbol{s} = (s_1, \dots, s_N)$ indicates long $(s_i = 1)$ or short $(s_i = -1)$ positions.
- **Portfolio relation with shorting pattern:**
  - Actual portfolio $\boldsymbol{w}$ related to a virtual no-shorting portfolio $\tilde{\boldsymbol{w}} \geq \mathbf{0}$:
  $$\boldsymbol{w} = \boldsymbol{s} \odot \tilde{\boldsymbol{w}}$$
  - Risk remains equivalent:
  $$\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w} = \tilde{\boldsymbol{w}}^\mathsf{T}\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{w}},$$
  where $\tilde{\boldsymbol{\Sigma}} = \mathrm{Diag}(\boldsymbol{s})\boldsymbol{\Sigma}\mathrm{Diag}(\boldsymbol{s})$.
- **Risk budgeting with shorting:**
  - Risk budgeting equations for virtual portfolio $\tilde{\boldsymbol{w}}$:
  $$\tilde{w}_i(\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{w}})_i = b_i\,\tilde{\boldsymbol{w}}^\mathsf{T}\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{w}}, \qquad i = 1, \dots, N.$$

Formulation with group risk parity

- **Concept of group risk parity:**
  - Risk contributions of assets within the same group (e.g., industry or sector) are considered collectively.
- **Group definition:**
  - K groups, $\mathcal{G}_1, \ldots, \mathcal{G}_K$, partition the N assets.
  - Each group $\mathcal{G}_k$ contains assets that are treated as a single entity in terms of risk.
- **Group risk contribution:**
  - Risk contribution from the kth group:

$$\mathrm{RC}_{\mathcal{G}_k} = \sum_{i \in \mathcal{G}_k} w_i \frac{\partial \sigma}{\partial w_i} = \sum_{i \in \mathcal{G}_k} \frac{w_i (\boldsymbol{\Sigma} \boldsymbol{w})_i}{\sqrt{\boldsymbol{w}^\mathsf{T} \boldsymbol{\Sigma} \boldsymbol{w}}}$$

- **Risk budgeting for groups:**
  - Risk budgeting equations for groups:

$$\sum_{i \in \mathcal{G}_k} w_i (\boldsymbol{\Sigma} \boldsymbol{w})_i = b_k \, \boldsymbol{w}^\mathsf{T} \boldsymbol{\Sigma} \boldsymbol{w}, \qquad k = 1, \ldots, K.$$

  - b_k represents the risk budget for group k.

Formulation with risk factors

- **Factor model for returns:**
$$\boldsymbol{r}_t = \boldsymbol{\alpha} + \boldsymbol{B}\boldsymbol{f}_t + \boldsymbol{\epsilon}_t,$$
  - $\boldsymbol{f}_t$: K factors (with $K \ll N$).
  - $\boldsymbol{\alpha}$: "alpha".
  - $\boldsymbol{B}$: matrix of "betas" for different factors.
  - $\boldsymbol{\epsilon}_t$: residual.
- **Risk contribution from factors:**
  - Defined for the kth factor as:
$$\mathrm{RC}_k = \frac{(\boldsymbol{B}^\mathsf{T}\boldsymbol{w})_k(\boldsymbol{B}^\dagger\boldsymbol{\Sigma}\boldsymbol{w})_k}{\sqrt{\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}}},$$
    where $\boldsymbol{B}^\dagger$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{B}$.
- **Risk budgeting in factor model:**
  - Risk budgeting equations for factors:
$$(\boldsymbol{B}^\mathsf{T}\boldsymbol{w})_k(\boldsymbol{B}^\dagger\boldsymbol{\Sigma}\boldsymbol{w})_k = b_k\, \boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}, \qquad k = 1, \dots, K.$$
  - b_k: risk budget for the kth factor.

Outline

Naive diagonal formulation

- **Risk budgeting equations with diagonal covariance:**
  - For diagonal covariance matrix $\boldsymbol{\Sigma} = \mathsf{Diag}(\boldsymbol{\sigma}^2)$:

$$w_i^2\sigma_i^2 = b_i \sum_{j=1}^{N} w_j^2\sigma_j^2, \qquad i = 1, \dots, N$$

  - Simplifies to:

$$w_i = \frac{\sqrt{b_i}}{\sigma_i}\sqrt{\sum_{j=1}^{N} w_j^2\sigma_j^2}, \qquad i = 1, \dots, N.$$

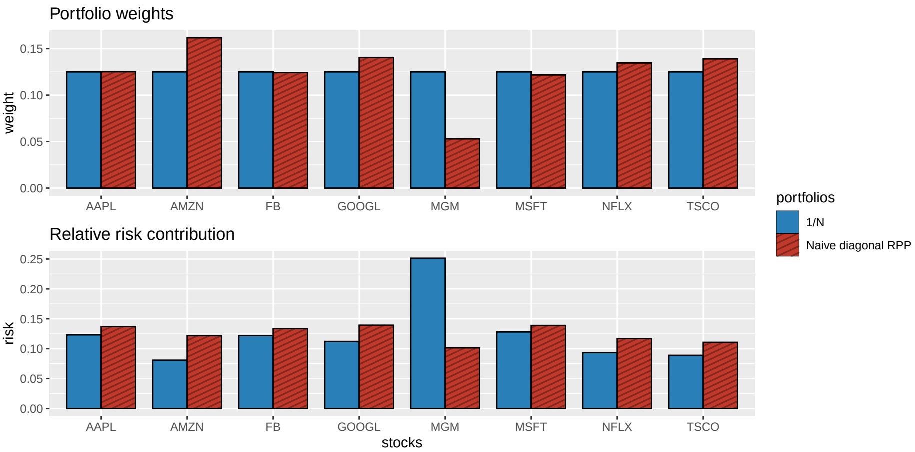
- **Inverse volatility portfolio (IVolP):**
  - Portfolio weights inversely proportional to asset volatilities.
  - Lower weights to high-volatility assets, higher weights to low-volatility assets.
  - Results in equal volatility contribution from each asset for $b_i = 1/N$.

Naive diagonal formulation

- **General nondiagonal covariance matrix:**
  - No closed-form solution available; optimization required.
  - Diagonal solution serves as a "naive" approach.
- **Portfolio allocation and risk contribution:**
  - The $1/N$ portfolio allocates capital equally across assets.
  - However, it results in unequal risk contributions.
- **Naive risk parity portfolio:**
  - Achieves a more balanced risk contribution among assets.
  - Not perfectly equalized due to ignoring off-diagonal covariance matrix elements.

Example: Naive RPP vs. $1/N$ Portfolio

Portfolio allocation and risk contribution of the $1/N$ portfolio and naive RPP:



Outline

Vanilla convex formulations

- **Risk budgeting equations:**
  - Given by:
$$w_i\left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i = b_i\,\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}, \qquad i = 1, \ldots, N$$
  - With constraints $\mathbf{1}^\mathsf{T}\boldsymbol{w} = 1$ and $\boldsymbol{w} \geq \mathbf{0}$.

- **Change of variable**
  - Define $\boldsymbol{x} = \boldsymbol{w}/\sqrt{\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}}$.
  - Rewrite equations as:
$$x_i\left(\boldsymbol{\Sigma}\boldsymbol{x}\right)_i = b_i$$
  - Vector form:
$$\boldsymbol{\Sigma}\boldsymbol{x} = \boldsymbol{b}/\boldsymbol{x}$$
  - Portfolio recovery by normalizing $\boldsymbol{x}$:
$$\boldsymbol{w} = \boldsymbol{x}/(\mathbf{1}^\mathsf{T}\boldsymbol{x}).$$

Vanilla convex formulations

- **Correlation matrix reformulation:**
  - Rewrite in terms of correlation matrix $\boldsymbol{C}$:

$$\boldsymbol{C}\tilde{\boldsymbol{x}} = \boldsymbol{b}/\tilde{\boldsymbol{x}},$$

  - $\boldsymbol{C} = \boldsymbol{D}^{-1/2}\Sigma\boldsymbol{D}^{-1/2}$, with $\boldsymbol{D} = \mathrm{Diag}(\boldsymbol{\sigma}^2)$.
  - $\boldsymbol{x} = \tilde{\boldsymbol{x}}/\boldsymbol{\sigma}$.

- **Numerical benefits:**
  - Normalizing returns with respect to asset volatilities can improve numerical stability.

Vanilla convex formulations: Direct resolution via root finding

- **Nonlinear equations system:**
  - System defined by $\boldsymbol{\Sigma x} = \boldsymbol{b}/\boldsymbol{x}$.
  - Interpreted as finding roots of $F(\boldsymbol{x}) = \boldsymbol{\Sigma x} - \boldsymbol{b}/\boldsymbol{x}$.
  - Goal: Solve $F(\boldsymbol{x}) = \boldsymbol{0}$.
- **Root finding in practice:**
  - Utilize general-purpose nonlinear multivariate root finders.
  - Available in most programming languages.
- **Root-finding with budget constraint:**
  - Include budget constraint $\boldsymbol{1}^\mathsf{T}\boldsymbol{w} = 1$ in function:

$$F(\boldsymbol{w}, \lambda) = \begin{bmatrix} \boldsymbol{\Sigma w} - \lambda \boldsymbol{b}/\boldsymbol{w} \\ \boldsymbol{1}^\mathsf{T}\boldsymbol{w} - 1 \end{bmatrix}.$$

- **Programming tools:**
  - **R:** Use `multiroot()` from package `rootSolve` for multivariate root finding.
  - **Matlab:** Use `fsolve()` for solving systems of nonlinear equations.

Convex reformulations

- **Convex optimization for risk budgeting:**
  - Risk budgeting equations can be solved through convex optimization, revealing hidden convexity.

- **Spinu's convex formulation:** (Spinu 2013)

$$\underset{\boldsymbol{x}\geq\mathbf{0}}{\text{minimize}} \quad \frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x} - \boldsymbol{b}^\mathsf{T}\log(\boldsymbol{x}).$$

- **Equivalence to risk budgeting:**
  - Gradient set to zero matches risk budgeting equation:

$$\boldsymbol{\Sigma}\boldsymbol{x} = \boldsymbol{b}/\boldsymbol{x}.$$

Convex reformulations

- **Roncalli's convex formulation:** (Roncalli 2013)

$$\underset{\boldsymbol{x}\geq\boldsymbol{0}}{\text{minimize}} \quad \sqrt{\boldsymbol{x}^T\boldsymbol{\Sigma}\boldsymbol{x}} - \boldsymbol{b}^T\log(\boldsymbol{x}).$$

  - Gradient zero leads to a form similar to risk budgeting equation after renormalization.

- **Maillard, Roncalli, and Teiletche's convex formulation:** (Maillard, Roncalli, and Teiletche 2010)

$$\underset{\boldsymbol{x}\geq\boldsymbol{0}}{\text{minimize}} \quad \sqrt{\boldsymbol{x}^T\boldsymbol{\Sigma}\boldsymbol{x}}, \quad \text{subject to} \quad \boldsymbol{b}^T\log(\boldsymbol{x}) \geq c.$$

  - Minimizes volatility with a diversification constraint.

Convex reformulations

- **Kaya and Lee's convex formulation:** (Kaya and Lee 2012)

$$\underset{\boldsymbol{x}\geq\boldsymbol{0}}{\text{maximize}} \quad \boldsymbol{b}^\mathsf{T}\log(\boldsymbol{x}), \quad \text{subject to} \quad \sqrt{\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}} \leq \sigma_0.$$

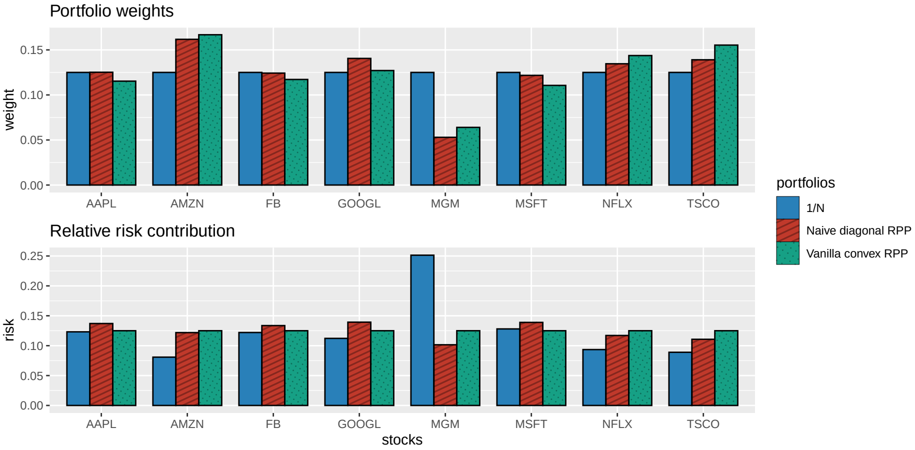
  - Gradient of Lagrangian matches risk budgeting equation after renormalization.

- **Solving convex formulations:**
  - General-purpose solvers can be used, available in programming languages like R (`optim()`) and Matlab (`fmincon()`).
  - Tailored algorithms can offer simple and efficient solutions.

- **Key Insight:**
  - These convex formulations provide different perspectives on achieving risk parity through optimization, each with its unique advantages and interpretations.

Example

Portfolio allocation and risk contribution of the vanilla convex RPP compared to benchmarks:



Vanilla convex formulations: Algorithms

- **Iterative algorithms:**
  - Develop practical algorithms for Spinu's and Roncalli's formulations.
  - Generate a sequence of iterates $\boldsymbol{x}^0, \boldsymbol{x}^1, \boldsymbol{x}^2, \ldots$
  - Important to have a good initial point $\boldsymbol{x}^0$ that attempts to approximate the solution to the nonlinear equations $\boldsymbol{\Sigma x} = \boldsymbol{b}/\boldsymbol{x}$.

- **Initial point options:** Crucial for the convergence and efficiency of the algorithms.
  - **Naive diagonal solution:**
$$\boldsymbol{x}^0 = \sqrt{\boldsymbol{b}}/\boldsymbol{\sigma}.$$
  - **Diagonal row-sum heuristic:**
$$\boldsymbol{x}^0 = \sqrt{\boldsymbol{b}}/\sqrt{\boldsymbol{\Sigma 1}}.$$

Vanilla convex formulations: Newton's method

- **Newton's method iteration:**

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathsf{H}f(\mathbf{x}^k)^{-1}\nabla f(\mathbf{x}^k).$$

- **Gradient and Hessian for Spinu's formulation:** $\left(f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{x} - \mathbf{b}^\mathsf{T}\log(\mathbf{x})\right)$

$$\begin{aligned}\nabla f(\mathbf{x}) &= \boldsymbol{\Sigma}\mathbf{x} - \mathbf{b}/\mathbf{x}\\ \mathsf{H}f(\mathbf{x}) &= \boldsymbol{\Sigma} + \mathrm{Diag}(\mathbf{b}/\mathbf{x}^2).\end{aligned}$$

- **Application to RPP:**
  - Newton's method can be applied to solve the risk parity portfolio optimization problem.
  - The method uses the gradient and Hessian of the objective function to iteratively improve the solution.
- **Reference for Newton's method:**
  - Detailed study of Newton's method for risk parity portfolio in (Spinu 2013).
  - For a general overview of gradient methods, see (Palomar 2025, Appendix B).

Vanilla convex formulations: Cyclical coordinate descent algorithm

- **Algorithm overview:**
  - Minimize function $f(\boldsymbol{x})$ cyclically for each element x_i (not parallel update).
  - Other elements of $\boldsymbol{x} = (x_1, \ldots, x_N)$ are held fixed during minimization.
  - Known as block coordinate descent (BCD) (Palomar 2025, Appendix B).
- **Elementwise minimization for Spinu's formulation:** $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x} - \boldsymbol{b}^\mathsf{T}\log(\boldsymbol{x})$

$$\underset{x_i \geq 0}{\text{minimize}} \quad \frac{1}{2}x_i^2\boldsymbol{\Sigma}_{ii} + x_i(\boldsymbol{x}_{-i}^\mathsf{T}\boldsymbol{\Sigma}_{-i,i}) - b_i \log x_i$$

  - $\boldsymbol{x}_{-i}$: variable $\boldsymbol{x}$ without ith element.
  - $\boldsymbol{\Sigma}_{-i,i}$: ith column of $\boldsymbol{\Sigma}$ without ith element.
- **Closed-form solution:**
  - Solve second order equation for x_i:

$$\boldsymbol{\Sigma}_{ii}x_i^2 + (\boldsymbol{x}_{-i}^\mathsf{T}\boldsymbol{\Sigma}_{-i,i})x_i - b_i = 0,$$

  - Positive solution:

$$x_i = \frac{-\boldsymbol{x}_{-i}^\mathsf{T}\boldsymbol{\Sigma}_{-i,i} + \sqrt{(\boldsymbol{x}_{-i}^\mathsf{T}\boldsymbol{\Sigma}_{-i,i})^2 + 4\boldsymbol{\Sigma}_{ii}b_i}}{2\boldsymbol{\Sigma}_{ii}}.$$

Vanilla convex formulations: Parallel update via MM

- **Majorization-minimization (MM) framework overview:** (Sun, Babu, and Palomar 2017) (Palomar 2025, Appendix B)
  - Solves optimization problems by iteratively solving simpler surrogate problems.
  - Surrogate problems are designed to majorize (upper-bound) the objective function.

- **Decoupling elements with MM:**
  - The term $\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}$ couples all elements of $\boldsymbol{x}$, complicating parallel updates.
  - MM framework allows for decoupling by using a particular majorizer for $\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}$.

- **Majorizer for $\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}$:**

$$\frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x} \leq \frac{1}{2}(\boldsymbol{x}^k)^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}^k + (\boldsymbol{\Sigma}\boldsymbol{x}^k)^\mathsf{T}(\boldsymbol{x}-\boldsymbol{x}^k) + \frac{\lambda_{\max}}{2}(\boldsymbol{x}-\boldsymbol{x}^k)^\mathsf{T}(\boldsymbol{x}-\boldsymbol{x}^k),$$

  where $\lambda_{\max}$ is the largest eigenvalue of $\boldsymbol{\Sigma}$.

Vanilla convex formulations: Parallel update via MM

- **Majorized problem for Spinu's formulation:**

$$\underset{\boldsymbol{x}\geq\boldsymbol{0}}{\text{minimize}} \quad \frac{\lambda_{\max}}{2}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{x} + \boldsymbol{x}^{\mathsf{T}}(\boldsymbol{\Sigma} - \lambda_{\max}\boldsymbol{I})\boldsymbol{x}^{k} - \boldsymbol{b}^{\mathsf{T}}\log(\boldsymbol{x}),$$

  - Solving this majorized problem simplifies the optimization.
- **Solution to majorized problem:**
  - Second order equation for x_i:

$$\lambda_{\max}x_i^2 + ((\boldsymbol{\Sigma} - \lambda_{\max}\boldsymbol{I})\boldsymbol{x}^{k})_i x_i - b_i = 0$$

  - Positive solution:

$$x_i = \frac{-((\boldsymbol{\Sigma} - \lambda_{\max}\boldsymbol{I})\boldsymbol{x}^{k})_i + \sqrt{((\boldsymbol{\Sigma} - \lambda_{\max}\boldsymbol{I})\boldsymbol{x}^{k})_i^2 + 4\lambda_{\max}b_i}}{2\lambda_{\max}}.$$

- **Advantages of MM:**
  - Allows for parallel updates by decoupling the elements of $\boldsymbol{x}$.
  - Simplifies the optimization problem, making it more tractable.

Vanilla convex formulations: Parallel update via SCA

- **SCA framework overview:** (Scutari et al. 2014) (Palomar 2025, Appendix B)
  - Solves optimization problems by iteratively solving simpler surrogate problems.
  - Surrogate problems approximate the original objective function, making optimization more tractable.

- **Decoupling elements with SCA:**
  - The term $\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}$ couples all elements of $\boldsymbol{x}$, complicating parallel updates.
  - SCA allows for decoupling by using a surrogate for $\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}$.

- **Surrogate for $\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}$:**

$$\frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x} \approx \frac{1}{2}(\boldsymbol{x}^k)^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{x}^k + (\boldsymbol{\Sigma}\boldsymbol{x}^k)^\mathsf{T}(\boldsymbol{x} - \boldsymbol{x}^k) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^k)^\mathsf{T}\mathrm{Diag}(\boldsymbol{\Sigma})(\boldsymbol{x} - \boldsymbol{x}^k)$$

where $\mathrm{Diag}(\boldsymbol{\Sigma})$ is a diagonal matrix with the diagonal of $\boldsymbol{\Sigma}$.

Vanilla convex formulations: Parallel update via SCA

- **Surrogate problem for Spinu's formulation:**

$$\underset{\mathbf{x}\geq\mathbf{0}}{\text{minimize}} \quad \frac{1}{2}\mathbf{x}^{\mathsf{T}}\text{Diag}(\boldsymbol{\Sigma})\mathbf{x} + \mathbf{x}^{\mathsf{T}}(\boldsymbol{\Sigma} - \text{Diag}(\boldsymbol{\Sigma}))\mathbf{x}^{k} - \mathbf{b}^{\mathsf{T}}\log(\mathbf{x}),$$

  - Solving this surrogate problem simplifies the optimization.
- **Solution to surrogate problem:**
  - Second order equation for x_i:

$$\boldsymbol{\Sigma}_{ii}x_i^2 + ((\boldsymbol{\Sigma} - \text{Diag}(\boldsymbol{\Sigma}))\mathbf{x}^{k})_i x_i - b_i = 0$$

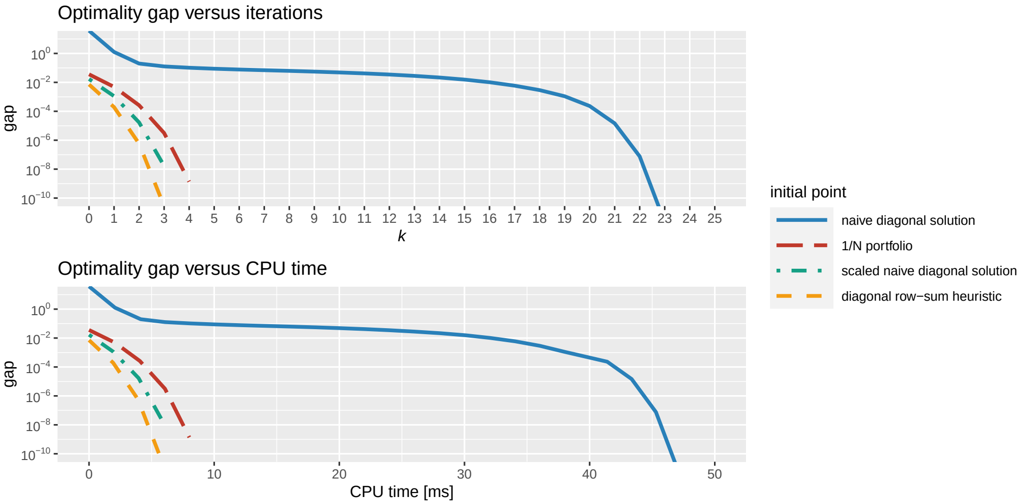
  - Positive solution:

$$x_i = \frac{-((\boldsymbol{\Sigma} - \text{Diag}(\boldsymbol{\Sigma}))\mathbf{x}^{k})_i + \sqrt{((\boldsymbol{\Sigma} - \text{Diag}(\boldsymbol{\Sigma}))\mathbf{x}^{k})_i^2 + 4\boldsymbol{\Sigma}_{ii}b_i}}{2\boldsymbol{\Sigma}_{ii}}.$$

- **Advantages of SCA:**
  - Allows for parallel updates by decoupling the elements of $\mathbf{x}$.
  - Simplifies the optimization problem, making it more tractable.

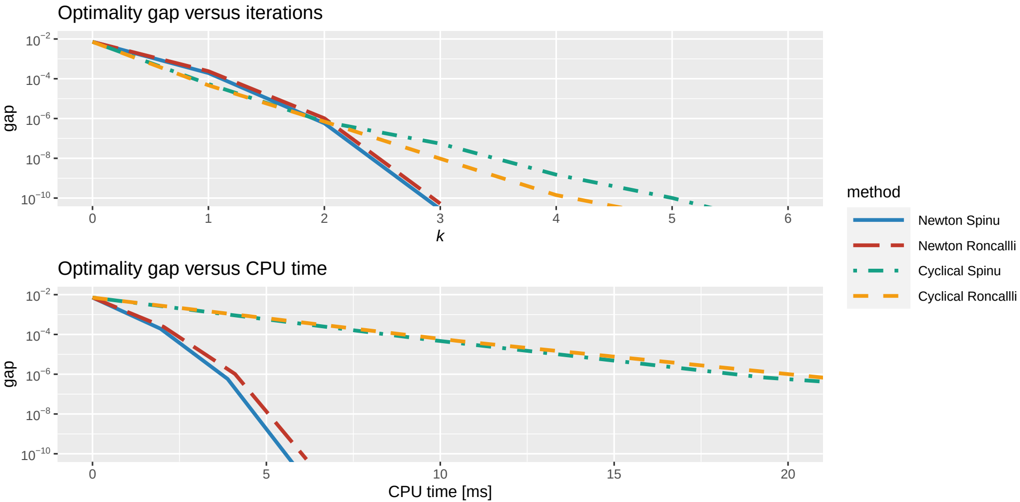
Numerical experiments: Effect of initial point

Effect of the initial point in Newton's method for Spinu's RPP formulation:



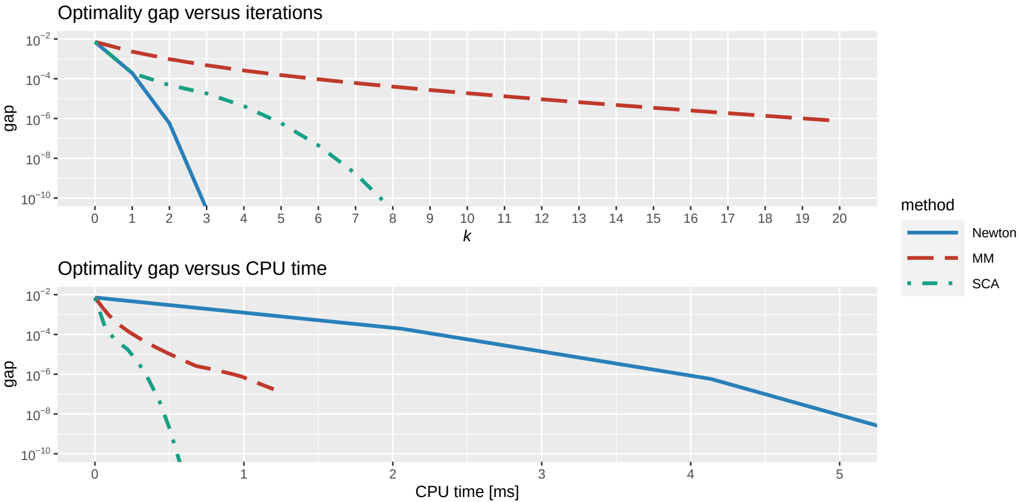
Numerical experiments: Newton vs cyclical optimization

Difference between Newton and cyclical optimization for Spinu's and Roncalli's:



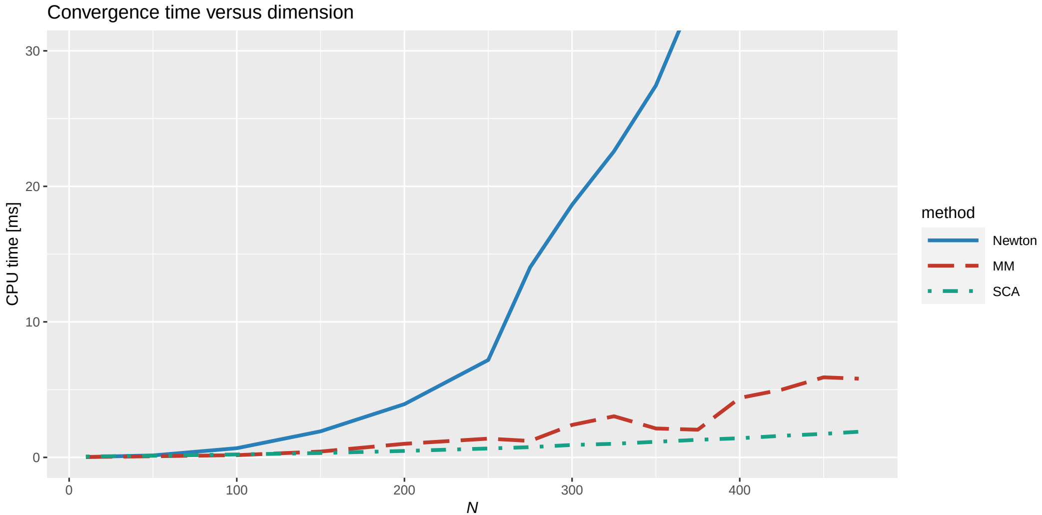
Numerical experiments: Final comparison

Convergence of different algorithms for the vanilla convex RPP:



Numerical experiments: Final comparison

Computational cost versus dimension N of different algorithms for the vanilla convex RPP:



Outline

General nonconvex formulations

- **Risk parity with expected return:**
  - Enhanced risk parity considers expected return within the risk parity framework.
  - Addresses criticism of risk parity's focus on risk over performance.
- **Vanilla formulation:**
  - Vanilla convex formulations focused on basic portfolio constraints.
  - Convex reformulations optimal for risk budgeting equations:

$$w_i \left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i = b_i \boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}, \quad i = 1, \dots, N.$$

- **Realistic scenarios with additional constraints:**
  - Portfolio managers often have extra constraints (turnover, market-neutral, maximum-position, etc.).
  - Additional objectives like maximizing expected return or minimizing variance/volatility.
  - Convex formulations no longer applicable; nonconvex formulations required.

General nonconvex formulations

- **Approximate satisfaction of risk budgeting equations:**

$$w_i\left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i \approx b_i\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}, \quad i=1,\ldots,N.$$

- **Measures of approximation error:**
  - Sum of squared relative risk-contribution errors:

$$\sum_{i=1}^{N}\left(\frac{w_i\left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i}{\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}} - b_i\right)^2$$

  - Sum of squared risk-contribution errors:

$$\sum_{i=1}^{N}\left(\frac{w_i\left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i}{\sqrt{\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}}} - b_i\sqrt{\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}}\right)^2$$

  - Sum of squared volatility-scaled risk-contribution errors:

$$\sum_{i=1}^{N}\left(w_i\left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i - b_i\boldsymbol{w}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{w}\right)^2$$

General nonconvex formulations

- **Herfindahl index for risk concentration:**

$$h(\boldsymbol{w}) = \sum_{i=1}^{N} \left(\frac{w_i \frac{\partial f}{\partial w_i}}{f(\boldsymbol{w})} \right)^2$$

  - Indicates risk diversification, with $1/N \leq h(\boldsymbol{w}) \leq 1$.
  - Smaller index implies more diversified risk.

- **Alternative norms for error measurement:**
  - ℓ_1-norm, ℓ_∞-norm, Huber's robust penalty function, etc.
  - Leads to various portfolio formulations with different convergence behaviors.

- **Application:**
  - These measures and formulations are used to create portfolios that balance risk diversification with performance objectives, accommodating a range of constraints and preferences.

General nonconvex formulations

- **Maillard, Roncalli, and Teiletche's formulation:** (Maillard, Roncalli, and Teiletche 2010)

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \displaystyle\sum_{i,j=1}^{N} \left(w_i \left(\boldsymbol{\Sigma w} \right)_i - w_j \left(\boldsymbol{\Sigma w} \right)_j \right)^2 \\ \text{subject to} & \boldsymbol{w} \in \mathcal{W} \end{array}$$

- **Alternative reformulation with dummy variable:**

$$\begin{array}{ll} \underset{\boldsymbol{w},\theta}{\text{minimize}} & \displaystyle\sum_{i=1}^{N} \left(w_i \left(\boldsymbol{\Sigma w} \right)_i - \theta \right)^2 \\ \text{subject to} & \boldsymbol{w} \in \mathcal{W} \end{array}$$

where the optimal θ is $\theta = \frac{1}{N} \boldsymbol{w}^{\mathsf{T}} \boldsymbol{\Sigma w}$.

General nonconvex formulations

- **Bruder and Roncalli's formulation:** (Bruder and Roncalli 2012)

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \sum_{i=1}^{N}\left(\dfrac{w_i\left(\boldsymbol{\Sigma}\mathbf{w}\right)_i}{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}-b_i\right)^2\\ \text{subject to} & \mathbf{w}\in\mathcal{W}.\end{array}$$

- **Minimization of the Herfindahl index:**

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \sum_{i=1}^{N}\left(\dfrac{w_i(\boldsymbol{\Sigma}\mathbf{w})_i}{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}\right)^2\\ \text{subject to} & \mathbf{w}\in\mathcal{W},\end{array}$$

which can be seen as particular case of Bruder and Roncalli's formulation with $b_i = 0$.

Numerical Issues and Recommendations

- **Maillard et al.'s double-summation formulation:**
  - Can suffer from numerical issues due to very small squared terms.
  - Covariance matrix $\boldsymbol{\Sigma}$ may need artificial scaling.

- **Preferred formulations for numerical stability:**
  - Bruder and Roncalli's formulation.
  - Minimization of the Herfindahl index.
  - Based on normalized terms, offering better numerical stability.

- **Application:**
  - These formulations are used to create risk parity portfolios that also consider additional constraints and objectives, such as expected return, while maintaining numerical stability.

Unified formulation

- **General Formulation:** (Feng and Palomar 2015)

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \sum_{i=1}^{N} g_i(\mathbf{w})^2 + \lambda F(\mathbf{w}) \\ \text{subject to} & \mathbf{w} \in \mathcal{W} \end{array}$$

- **Concentration Error Measure ($g_i(\mathbf{w})$):**
  - Represents the deviation of the ith asset's risk contribution from its target budget b_i.
  - Example:

$$g_i(\mathbf{w}) = \frac{w_i\left(\mathbf{\Sigma}\mathbf{w}\right)_i}{\mathbf{w}^\mathsf{T}\mathbf{\Sigma}\mathbf{w}} - b_i,$$

- **Preference Function ($F(\mathbf{w})$):**
  - Encapsulates additional objectives, such as maximizing expected return or minimizing variance.
  - Example:

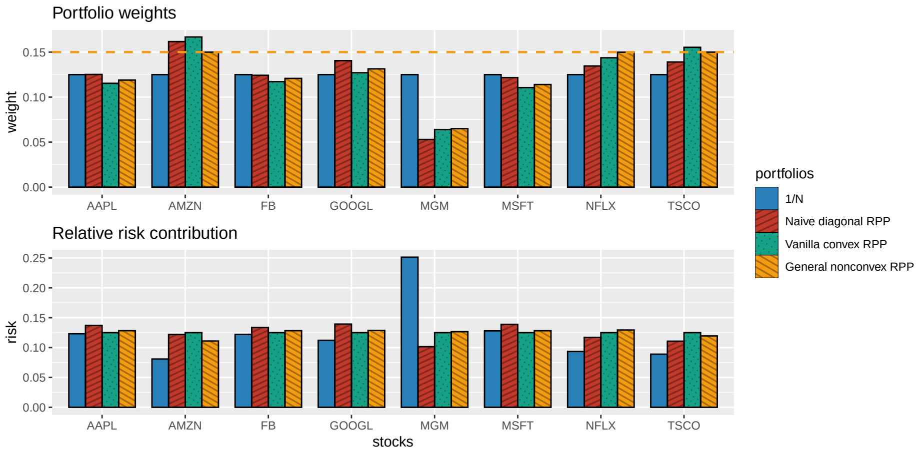
$$F(\mathbf{w}) = -\mathbf{w}^\mathsf{T}\boldsymbol{\mu} + \frac{1}{2}\mathbf{w}^\mathsf{T}\mathbf{\Sigma}\mathbf{w}$$

Unified formulation

- **Trade-off hyper-parameter (λ):**
  - Balances between minimizing concentration errors and optimizing the preference function.
- **Versatility of the formulation:**
  - Capable of incorporating various risk parity formulations and additional objectives.
  - Adaptable to different error measures and preference functions.
- **Challenges in algorithm design:**
  - Nonconvexity of the term $\sum_{i=1}^{N} g_i(\boldsymbol{w})^2$ complicates the development of algorithms.
  - Requires sophisticated optimization techniques to navigate the nonconvex landscape.
- **Significance:**
  - This unified formulation offers a comprehensive framework for risk parity portfolio construction.
  - It allows for the integration of risk management with performance optimization, accommodating a wide range of portfolio management preferences and constraints.

Numerical experiments

Portfolio allocation and risk contribution of general nonconvex RPP (with $w_i \leq 0.15$) compared to benchmarks ($1/N$ portfolio, naive diagonal RPP, and vanilla convex RPP):



General nonconvex formulations: Algorithms

- **Iterative algorithms:**
  - General-purpose solvers can address previous nonconvex formulations.
  - Iterative algorithms can be developed for efficiency producing a sequence of iterates: $\boldsymbol{w}^0, \boldsymbol{w}^1, \boldsymbol{w}^2, \dots$

- **Initial point:**
  - Initial point for algorithms can be the solution from vanilla convex formulation.
  - Must ensure feasibility with all constraints in $\mathcal{W}$.
  - Alternatively, use the $1/N$ portfolio as a simpler initial point.

Algorithms: SCA

- **Original difficult problem:**

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \mathcal{X} \end{array}$$

  where f is the (possibly nonconvex) objective function and $\mathcal{X}$ is the convex feasible set.

- **Successive convex approximation (SCA) method:** (Scutari et al. 2014) (Palomar 2025, Appendix B)
  - Approximates a difficult optimization problem by a sequence of simpler problems:

$$\boldsymbol{x}^{k+1} = \arg\min_{\boldsymbol{x} \in \mathcal{X}} \tilde{f}\left(\boldsymbol{x}; \boldsymbol{x}^{k}\right), \qquad k = 0, 1, 2, \ldots$$

    where $\tilde{f}\left(\boldsymbol{x}; \boldsymbol{x}^{k}\right)$ approximates $f(\boldsymbol{x})$ around the current point $\boldsymbol{x}^{k}$.
  - Produces a sequence of iterates $\boldsymbol{x}^{0}, \boldsymbol{x}^{1}, \boldsymbol{x}^{2}, \ldots$ that converge to $\boldsymbol{x}^{\star}$.
  - For convergence, we need a smoothing step to avoid oscillations ($\gamma^{k} \in (0, 1]$):

$$\left.\begin{aligned} \hat{\boldsymbol{x}}^{k+1} &= \arg\min_{\boldsymbol{x} \in \mathcal{X}} \tilde{f}\left(\boldsymbol{x}; \boldsymbol{x}^{k}\right) \\ \boldsymbol{x}^{k+1} &= \boldsymbol{x}^{k} + \gamma^{k}\left(\hat{\boldsymbol{x}}^{k+1} - \boldsymbol{x}^{k}\right) \end{aligned}\right. \qquad k = 0, 1, 2, \ldots$$

Algorithms: SCA

- **Application of SCA method:**
  - Unified formulation objective function:

$$U(\mathbf{w}) = \sum_{i=1}^{N} g_i(\mathbf{w})^2 + \lambda F(\mathbf{w}).$$

  - Convexification by linearizing $g_i(\mathbf{w})$ around $\mathbf{w}^k$:

$$g_i(\mathbf{w}) \approx g_i(\mathbf{w}^k) + \nabla g_i(\mathbf{w}^k)^\mathsf{T} \left(\mathbf{w} - \mathbf{w}^k\right).$$

  - Surrogate function:

$$\tilde{U}(\mathbf{w}, \mathbf{w}^k) = \sum_{i=1}^{N} \left(g_i(\mathbf{w}^k) + \nabla g_i(\mathbf{w}^k)^\mathsf{T} \left(\mathbf{w} - \mathbf{w}^k\right)\right)^2 + \lambda F(\mathbf{w}) + \frac{\tau}{2} \left\|\mathbf{w} - \mathbf{w}^k\right\|_2^2.$$

- **Approximated QP Formulation:**

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \frac{1}{2}\mathbf{w}^{\mathsf{T}}\mathbf{Q}^{k}\mathbf{w} + \mathbf{w}^{\mathsf{T}}\mathbf{q}^{k} + \lambda F(\mathbf{w}) \\ \text{subject to} & \mathbf{w} \in \mathcal{W}, \end{array}$$

where

$$\mathbf{Q}^{k} \triangleq 2\left(\mathbf{J}^{k}\right)^{\mathsf{T}}\mathbf{J}^{k} + \tau\mathbf{I},$$

$$\mathbf{q}^{k} \triangleq 2\left(\mathbf{J}^{k}\right)^{\mathsf{T}}\mathbf{g}^{k} - \mathbf{Q}^{k}\mathbf{w}^{k},$$

and

$$\mathbf{g}^{k} \triangleq \left[g_{1}(\mathbf{w}^{k}), \ldots, g_{N}(\mathbf{w}^{k})\right]^{\mathsf{T}}$$

$$\mathbf{J}^{k} \triangleq \begin{bmatrix} \nabla g_{1}(\mathbf{w}^{k})^{\mathsf{T}} \\ \vdots \\ \nabla g_{N}(\mathbf{w}^{k})^{\mathsf{T}} \end{bmatrix}.$$

Algorithms: SCA

Successive Convex optimization for RIsk Parity portfolio (SCRIP) (Feng and Palomar 2015)

Initialization:

- Start with an initial portfolio $\boldsymbol{w}^0$ within the feasible set $\mathcal{W}$.
- Define sequence $\{\gamma^k\}$.

Repeat (kth iteration):

1. Calculate risk concentration terms $\boldsymbol{g}^k$ and Jacobian matrix $\boldsymbol{J}^k$ for current point $\boldsymbol{w}^k$.
2. Solve approximated QP problem and keep solution as $\hat{\boldsymbol{w}}^{k+1}$.
3. Update the portfolio as $\boldsymbol{w}^{k+1} \leftarrow \boldsymbol{w}^k + \gamma^k(\hat{\boldsymbol{w}}^{k+1} - \boldsymbol{w}^k)$.
4. $k \leftarrow k+1$.

Until: The solution converges to the optimal portfolio.

Algorithms: ALM

- **Alternate linearization method (ALM) overview:**
  - Proposed in (Bai, Scheinberg, and Tütüncü 2016) for solving Maillard's formulation with a single summation.
  - Objective function:

$$F(\mathbf{w}, \theta) = \sum_{i=1}^{N} \left(w_i \left(\boldsymbol{\Sigma}\mathbf{w}\right)_i - \theta\right)^2 = \sum_{i=1}^{N} \left(\mathbf{w}^\mathsf{T}\mathbf{M}_i\mathbf{w} - \theta\right)^2,$$

    where $\mathbf{M}_i$ contains the ith-row of $\boldsymbol{\Sigma}$ and zeros elsewhere.

- **ALM strategy:**
  - Introduce variable $\mathbf{y}$, redefine objective as

$$F(\mathbf{w}, \mathbf{y}, \theta) = \sum_{i=1}^{N} \left(\mathbf{w}^\mathsf{T}\mathbf{M}_i\mathbf{y} - \theta\right)^2,$$

    subject to $\mathbf{y} = \mathbf{w}$.
  - Sequentially optimize $\mathbf{w}$, $\mathbf{y}$, and θ using two QP approximations.

Algorithms: ALM

- **QP approximations in ALM:**
  - First QP approximation:

$$Q^1(\boldsymbol{w}, \boldsymbol{y}^k, \theta) = F(\boldsymbol{w}, \boldsymbol{y}^k, \theta) + \nabla_2 F(\boldsymbol{y}^k, \boldsymbol{y}^k, \theta)^\mathsf{T}(\boldsymbol{w} - \boldsymbol{y}^k) + \frac{1}{2\mu}\|\boldsymbol{w} - \boldsymbol{y}^k\|_2^2$$

  - Second QP approximation:

$$Q^2(\boldsymbol{w}^{k+1}, \boldsymbol{y}, \theta) = F(\boldsymbol{w}^{k+1}, \boldsymbol{y}, \theta) + \nabla_1 F(\boldsymbol{w}^{k+1}, \boldsymbol{w}^{k+1}, \theta)^\mathsf{T}(\boldsymbol{y} - \boldsymbol{w}^{k+1}) + \frac{1}{2\mu}\|\boldsymbol{y} - \boldsymbol{w}^{k+1}\|_2^2$$

- **Gradient calculations for ALM:**
  - Gradient with respect to $\boldsymbol{w}$:

$$\nabla_1 F(\boldsymbol{w}, \boldsymbol{y}, \theta) = 2\sum_{i=1}^{N}\left(\boldsymbol{w}^\mathsf{T}\boldsymbol{M}_i\boldsymbol{y} - \theta\right)\boldsymbol{M}_i\boldsymbol{y}$$

  - Gradient with respect to $\boldsymbol{y}$:

$$\nabla_2 F(\boldsymbol{w}, \boldsymbol{y}, \theta) = 2\sum_{i=1}^{N}\left(\boldsymbol{w}^\mathsf{T}\boldsymbol{M}_i\boldsymbol{y} - \theta\right)\boldsymbol{M}_i^\mathsf{T}\boldsymbol{w}.$$

Numerical experiments

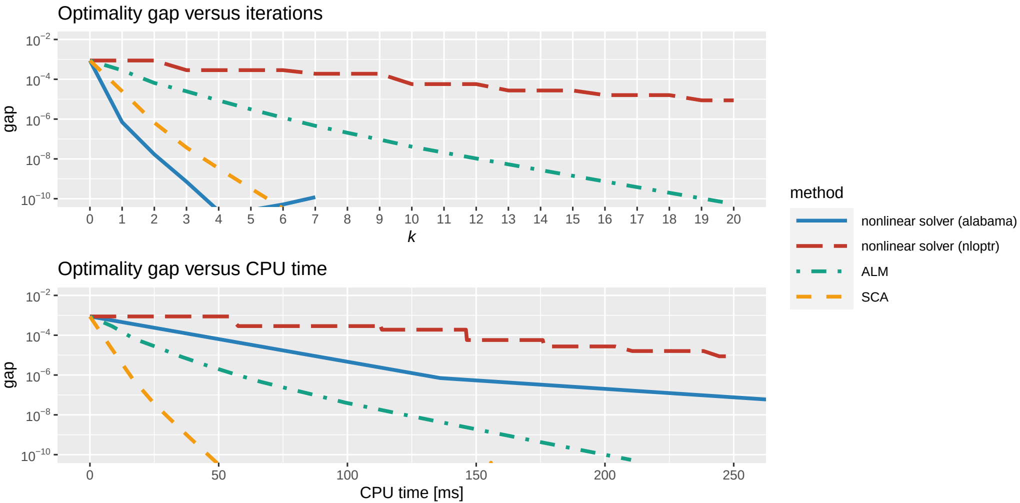
- **Nonconvex formulation challenges:**
  - Initial nonconvex problem:

$$\begin{array}{ll}\underset{\mathbf{w},\theta}{\text{minimize}} & \sum_{i=1}^{N}\left(w_i\left(\boldsymbol{\Sigma}\mathbf{w}\right)_i-\theta\right)^2\\ \text{subject to} & \mathbf{w}\in\mathcal{W}\end{array}$$

  - Squared terms $w_i\left(\boldsymbol{\Sigma}\mathbf{w}\right)_i$ can be numerically unstable when small.
- **Numerical stability heuristic:**
  - Scale up covariance matrix $\boldsymbol{\Sigma}$ by a large factor (e.g., 10^4) to mitigate numerical issues.
  - Suggested in (Mausser and Romanko 2014).
- **Preferred formulation for stability:**
  - Use normalized terms for better numerical stability: $w_i\left(\boldsymbol{\Sigma}\mathbf{w}\right)_i/(\mathbf{w}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{w})$

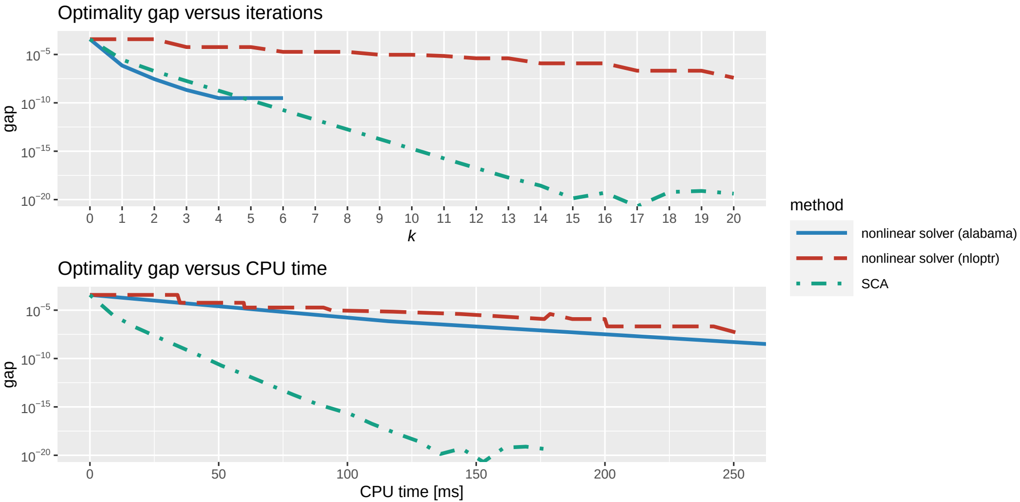
$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \sum_{i=1}^{N}\left(\dfrac{w_i\left(\boldsymbol{\Sigma}\mathbf{w}\right)_i}{\mathbf{w}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{w}}-b_i\right)^2\\ \text{subject to} & \mathbf{w}\in\mathcal{W}.\end{array}$$

Convergence of algorithms for nonconvex RPP formulation in terms of $w_i\left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i$:



Numerical experiments

Convergence of algorithms for nonconvex RPP formulation in terms of $w_i\left(\boldsymbol{\Sigma}\boldsymbol{w}\right)_i / (\boldsymbol{w}^{\mathsf{T}}\boldsymbol{\Sigma}\boldsymbol{w})$:



Outline

Summary

Diversification is key in portfolio design, as the saying goes, "don't put all your eggs in one basket." Some key points:

- The $1/N$ portfolio effectively diversifies capital allocation, but risk parity portfolios offer a more advanced strategy by diversifying risk allocation.
- Risk parity portfolios express the risk measure (e.g., volatility) as the sum of individual risk contributions from each asset, providing refined risk control compared to using a single overall portfolio risk value.
- Risk parity formulations have three levels of complexity:
  - **Naive diagonal formulation**: assumes a diagonal covariance matrix, simplifying to the inverse-volatility portfolio (ignoring asset correlations);
  - **Vanilla convex formulations**: consider simple long-only portfolios, rewritten in convex form with efficient algorithms; and
  - **General nonconvex formulations**: admit realistic constraints and extended objective functions, becoming nonconvex problems requiring careful resolution (still with efficient iterative algorithms).

References I

Bai, X., K. Scheinberg, and R. Tütüncü. 2016. "Least-Squares Approach to Risk Parity in Portfolio Selection." *Quantitative Finance* 16 (3): 357–76.

Bruder, B., and T. Roncalli. 2012. "Managing Risk Exposures Using the Risk Budgeting Approach." University Library of Munich, Germany.

Feng, Y., and D. P. Palomar. 2015. "SCRIP: Successive Convex Optimization Methods for Risk Parity Portfolios Design." *IEEE Transactions on Signal Processing* 63 (19): 5285–5300.

Kaya, H., and W. Lee. 2012. "Demystifying Risk Parity." *SSRN Electronic Journal*. `https://dx.doi.org/10.2139/ssrn.1987770`.

Maillard, S., T. Roncalli, and J. Teiletche. 2010. "The Properties of Equally Weighted Risk Contribution Portfolios." *Journal of Portfolio Management* 36 (4): 60–70.

Mausser, H., and O. Romanko. 2014. "Computing Equal Risk Contribution Portfolios." *Journal of Research and Development* 58 (4): 5:1–12.

Palomar, D. P. 2025. *Portfolio Optimization: Theory and Application*. Cambridge University Press.

References II

Qian, E. 2005. “Risk Parity Portfolios: Efficient Portfolios Through True Diversification.” *PanAgora Asset Management*.

———. 2016. *Risk Parity Fundamentals*. CRC Press.

Roncalli, T. 2013. *Introduction to Risk Parity and Budgeting*. Chapman & Hall/CRC Financial Mathematics Series.

Scutari, G., F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang. 2014. “Decomposition by Partial Linearization: Parallel Optimization of Multi-Agent Systems.” *IEEE Transactions on Signal Processing* 62 (3): 641–56.

Spinu, F. 2013. “An Algorithm for Computing Risk Parity Weights.” *SSRN Electronic Journal*. `https://dx.doi.org/10.2139/ssrn.2297383`.

Sun, Y., P. Babu, and D. P. Palomar. 2017. “Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning.” *IEEE Transactions on Signal Processing* 65 (3): 794–816.