Markowitz variations

Diversification in style

Dr Thomas Schmelzer
Visiting Scholar Stanford University
and
Quant R&D Lead ADIA

Almost Academic Keynote
EQD Singapore
October 15

Between academia and industry

Integration is key between both worlds

With (from left):

- Trevor Hastie (RSA)
  - Statistical Learning and AI not only applied to finance
- Emmanuel Candes (FR)
  - Compressed sensing, reconstruction of data for machine learning algorithms.
- Stephen Boyd (USA)
  - Convex Programming used for portfolio construction

at the **Blackrock AI Lab**, Palo Alto

Harry Max Markowitz

August 24, 1927 – June 22, 2023

- Wrote the seminal ‘52 paper as a 24 year-old graduate student.
- Maximizing some expected profit subject to a risk constraint. Modern formulation but solvers did not exist in those days.
- For the risk term you could use a covariance matrix.
- **Read the original paper** before you throw mud.

The “first” paper: The 2-step process

- Markowitz distinguishes two steps:

  - **Estimation of expected returns and expected covariances**

  - **Formulation and solution of a quadratic program**

- Did not claim that a moving mean and sample covariance matrices should be used as expectations.
- He reiterated over the years that the estimation step is not his job.

His “second” paper (1955): Solving quadratic programs

- Is the algorithm he introduced today available?
- If not, can we reimplement it and make it widely available on GitHub?
- We need to understand the **efficient frontier**

RM-
2-
-1-

THE OPTIMIZATION OF QUADRATIC FUNCTIONS

SUBJECT TO LINEAR CONSTRAINTS

Harry Markowitz*

1 Suppose that variables $X_1, \ldots, X_N$ are to be chosen subject to constraints:

1) $\sum a_{ij} X_j = b_i$ $\quad i = 1, \ldots, m_1$

2) $\sum a_{ij} X_j \geq b_i$ $\quad i = m_1+1, \ldots, m$

3) $X_j \geq 0$ $\quad j = 1, \ldots, N_1$

he matrix (a_{ij}) $i = 1, \ldots, m_1$ has rank m_1, otherwise the system wou

nsistent or else there would be at least one redundant equation).

is a linear function $R = \sum r_j X_j$. The payoff coefficients r_j are n

t the time the X_j are chosen. The r_j are, rather, random variables

pected values μ_j and covariances σ_{jk} (including variances σ_{jj} =

ected value of R is

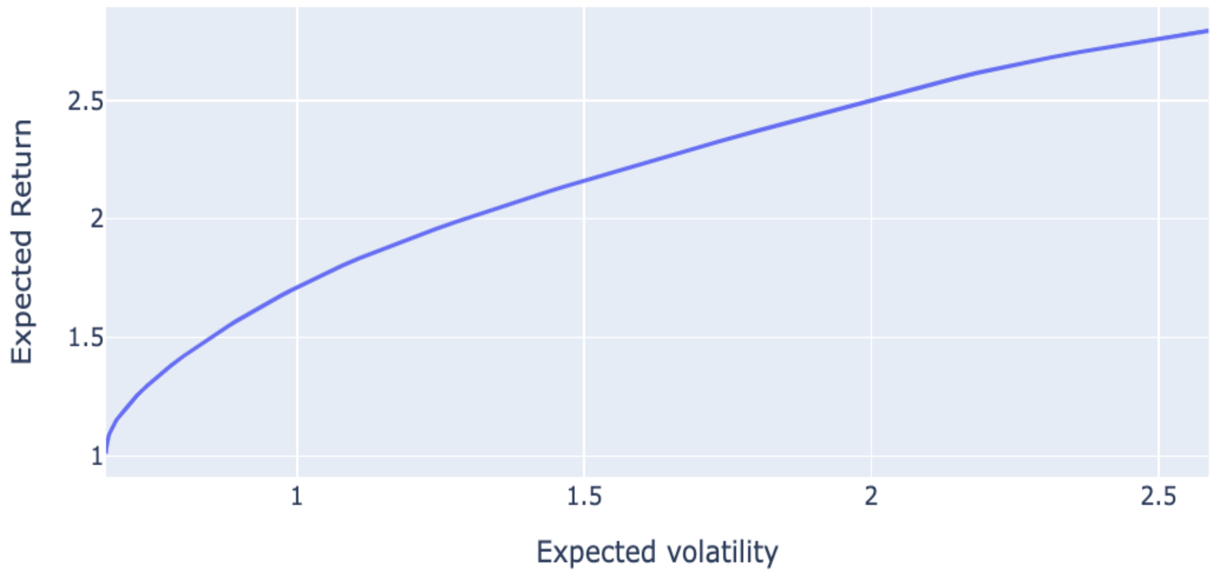
4) $E = \sum \mu_j X_j$

iance of R is

5) $V = \sum\sum \sigma_{jk} X_j X_k$

Computing an efficient frontier

The “brute force” approach of today



Loop over different values of target variance

Compute max. expected return by solving

1952

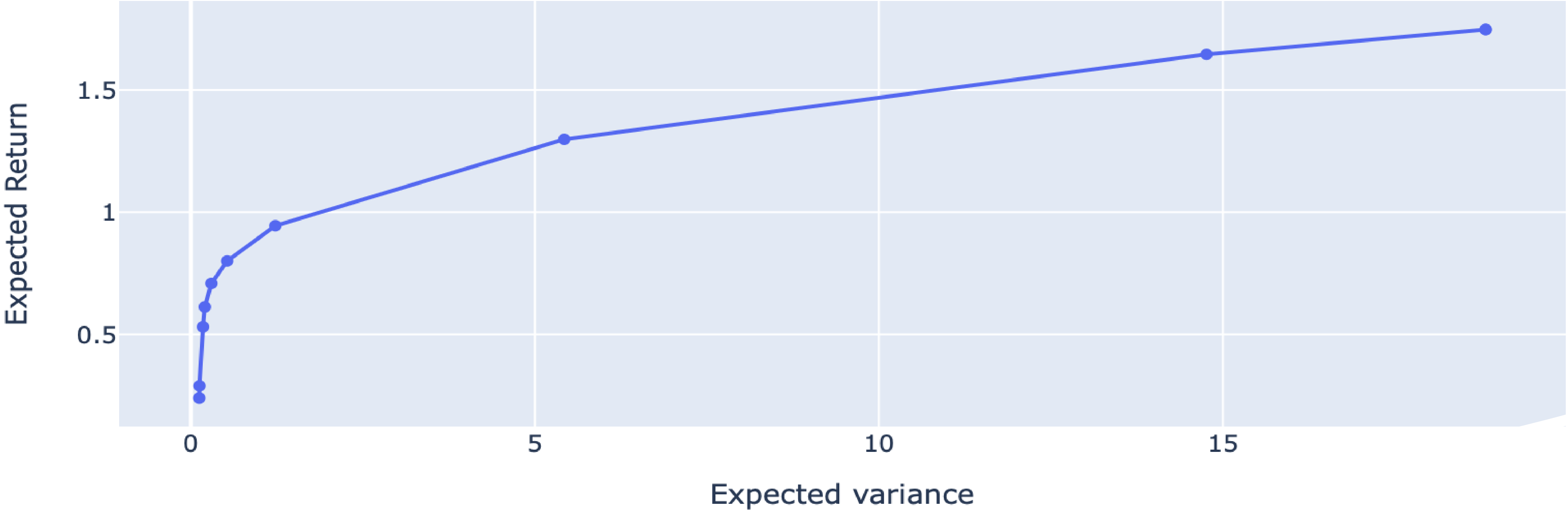
$$\begin{array}{ll} \text{maximize} & \mu^T w \\ \text{subject to} & w^T \Sigma w \leq (\sigma^{\text{tar}})^2, \\ & \mathbf{1}^T w = 1, \end{array}$$

Solution may not exist if target variance is too small (see Sin 7)

Collect the point (target variance, expected return)

Possible, but too expensive for 1955

Can we compute a minimal set of points reflecting the entire frontier?

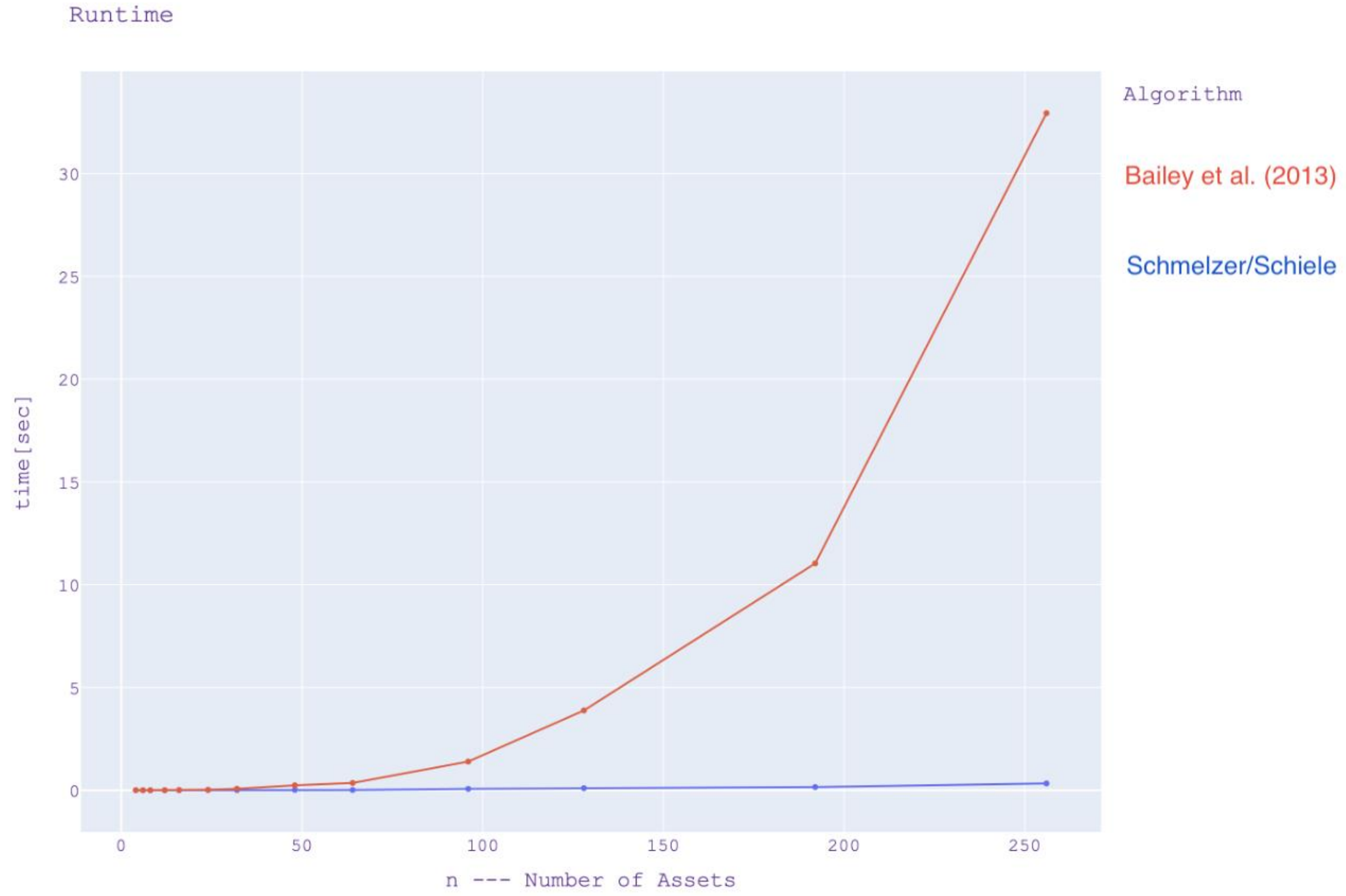


Efficient frontier over variance

- An efficient frontier is piecewise linear, e.g. each point is a linear interpolation between its two adjacent turning points.
- The critical line algorithm is computing all turning points going from the point of highest return to the minimum variance portfolio.

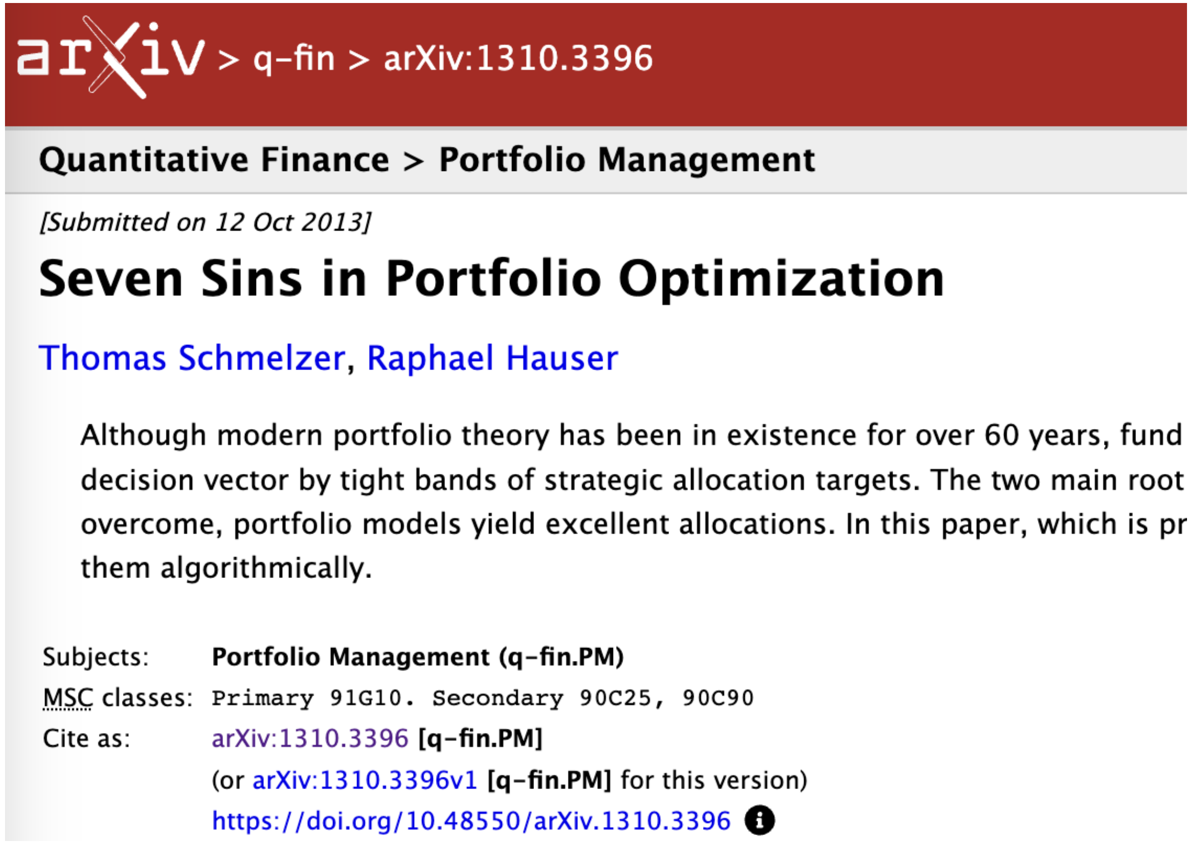
Reimplementing Markowitz' ideas fast(er)
With Philipp Schiele (Citadel)

We have found a didactic implementation in a paper from 2013. The authors kindly shared their source code with us:



Problems? It’s not Markowitz’s fault.

- It’s the **user.**
- Examples:
  - The eigenvalues are negative or too small.
  - The problem is not convex.
  - …
- **Further examples: See the paper ➔**
- **This paper describes what not to do**
- We are working on an updated version of it. **Please share your stories.**



arXiv > q-fin > arXiv:1310.3396

Quantitative Finance > Portfolio Management

[Submitted on 12 Oct 2013]

Seven Sins in Portfolio Optimization

Thomas Schmelzer, Raphael Hauser

Although modern portfolio theory has been in existence for over 60 years, fund decision vector by tight bands of strategic allocation targets. The two main root overcome, portfolio models yield excellent allocations. In this paper, which is pr them algorithmically.

Subjects:	**Portfolio Management (q-fin.PM)**
MSC classes:	Primary 91G10. Secondary 90C25, 90C90
Cite as:	arXiv:1310.3396 **[q-fin.PM]**
	(or arXiv:1310.3396v1 **[q-fin.PM]** for this version)
	https://doi.org/10.48550/arXiv.1310.3396

Solutions: Modern Markowitz in practice

a roadmap

Markowitz Portfolio Construction at Seventy

S. Boyd, K. Johansson, R. Kahn, P. Schiele, and T. Schmelzer

(Alphabetical author order.)

Journal of Portfolio Management, 50(8), 117–160, July 2024.

- Current version
- Code to reproduce results in paper

Stephen Boyd* Kasper Johansson Ronald Kahn Philipp Schiele
Thomas Schmelzer

January 5, 2024

Abstract

More than seventy years ago Harry Markowitz formulated portfolio construction as an optimization problem that trades off expected return and risk, defined as the standard deviation of the portfolio returns. Since then the method has been extended to include many practical constraints and objective terms, such as transaction cost or leverage limits. Despite several criticisms of Markowitz's method, for example its sensitivity to poor forecasts of the return statistics, it has become the dominant quantitative method for portfolio construction in practice. In this article we describe an extension of Markowitz's method that addresses many practical effects and gracefully handles the uncertainty inherent in return statistics forecasting. Like Markowitz's original formulation, the extension is also a convex optimization problem, which can be solved with high reliability and speed.

This paper describes what could be done

From the paper

The paper "Markowitz Portfolio Construction at Seventy" revisits Harry Markowitz's 1952 portfolio theory, which trades off expected return against risk. The authors propose extending the original method to handle practical constraints like transaction costs and estimation uncertainty. They **focus on making the optimization more robust by incorporating regularization and convex optimization techniques**. This generalization, called "Markowitz++," aims to address the sensitivity of Markowitz's model to forecast errors and improve its practical application while **retaining the core principles of balancing risk and return**

Markowitz vs Markowitz++

1952

$$\begin{array}{ll} \text{maximize} & \mu^T w \\ \text{subject to} & w^T \Sigma w \leq (\sigma^{\text{tar}})^2, \\ & \mathbf{1}^T w = 1, \end{array}$$

Inputs include expected returns, μ, and covariance matrix, Σ. Also the target risk level, σ^{tar}

This is only one way to write this down.

2024

$$\begin{array}{ll} \text{maximize} & R^{\text{wc}} - \gamma^{\text{hold}} \phi^{\text{hold}}(w, c) - \gamma^{\text{trade}} \phi^{\text{trade}}(z) \\ \text{subject to} & \mathbf{1}^T w + c = 1, \quad z = w - w^{\text{pre}}, \\ & w^{\min} \leq w \leq w^{\max}, \quad L \leq L^{\text{tar}}, \quad c^{\min} \leq c \leq c^{\max} \\ & z^{\min} \leq z \leq z^{\max}, \quad T \leq T^{\text{tar}}, \\ & \sigma^{\text{wc}} \leq \sigma^{\text{tar}}, \end{array}$$

Many more inputs: R^{wc} are the worst-case mean return forecasts = $\text{Min}\{(\mu+\delta)^T w \mid |\delta|<\rho\}$.

That turns out to be $R - \rho^T|w|$. We have similar worst-case estimates of risk, σ^{wc}.

Other input parameters include γ^{hold} and γ^{trade}, which scale holding and trading costs;

and constraints on holdings, w, leverage, cash, turnover, and trading amounts.

Source: Talk by Ronald Kahn for UBS US Quant conference

BlackRock.

Parameters:
Problems with any covariance matrix

- You got lost? Good moment to rejoin the talk.
- How aggressivly shall we update a covariance matrix? How far do we look back in time?
- It should be chilled to avoid trading costs yet reactive?
- Fundamental question: **Are we allocating risk or are we allocating capital?** Try to work with correlation rather than covariance matrices.
- An approach based on covariance matrices may not be the best choice at all, e.g. when combining risk premia strategies with an asymmetric return profile. Explore the cVaR.

Solution

Let’s take multiple estimators

A Simple Method for Predicting Covariance Matrices of Financial Returns

K. Johansson, M. Ogut, M. Pelger, T. Schmelzer, and S. Boyd

Foundations and Trends in Econometrics, 12(4):324–407, 2023.

- Final version
- Talk slides
- Code

- Use a convex combination of multiple matrices at all times
- Dynamically adjust the coefficients to be more reactive when markets move faster
- Formulate this as a convex program as we have access to robust solvers

Step back: Estimate one covariance matrix

- Estimating a covariance matrix is challenging when
  - A large number of assets have to be addressed
  - The distrubtion of returns has fat tails
  - Data is not synchronous or a lot of it is missing
- Common approaches in industry
  - Rolling window
  - External sourcing (e.g. buy a factor risk model)
  - Ignore all cross-correlations and estimate diagonal only
  - MGARCH
  - **DCC light**

DCC (Engle) light

- Dynamic conditional correlations.
- Asset returns have fat tails.

Step 1: Estimate the volatility and adjust + winsorize returns. Use an exponentially weighted mean. Vol adj returns are no longer fat tailed

Step 2: Estimate a correlation matrix and multiply with volatilities (from both sides) to get back to a covariance matrix (if needed).

- **To estimate multiple covariance matrices we use DCC light with different parameter choices for the windows.**

From the paper: Dynamically weighted prediction combiner

1. start with K covariance predictors $\hat{\Sigma}_t^{(k)}, \quad k = 1, \dots, K$
2. Cholesky factorizations of associated precision matrices

$$\left(\hat{\Sigma}_t^{(k)}\right)^{-1} = \hat{L}_t^{(k)} (\hat{L}_t^{(k)})^T, \quad k = 1, \dots, K$$

3. create convex combination

$$\hat{L}_t = \sum_{k=1}^{K} \pi_k \hat{L}_t^{(k)},$$

   where $\pi_k \geq 0$ and $\sum_{k=1}^{K} \pi_k = 1$

4. recover covariance predictor as $\hat{\Sigma}_t = \left(\hat{L}_t \hat{L}_t^T\right)^{-1}$

From the paper: Choose weights via convex optimization

- choose weights π at time t to maximize log-likelihood over past N time-steps

$$\begin{array}{ll}
\text{maximize} & \sum_{j=1}^{N} \left(\sum_{i=1}^{n} \log \hat{L}_{t-j,ii} - (1/2)\|\hat{L}_{t-j}^{T} r_{t-j}\|_2^2 \right) \\
\text{subject to} & \hat{L}_\tau = \sum_{j=1}^{K} \pi_j \hat{L}_\tau^{(j)}, \quad \tau = t-1, \ldots, t-N \\
& \pi \geq 0, \quad \mathbf{1}^T \pi = 1,
\end{array}$$

- convex problem that can be solved quickly and reliably by many methods

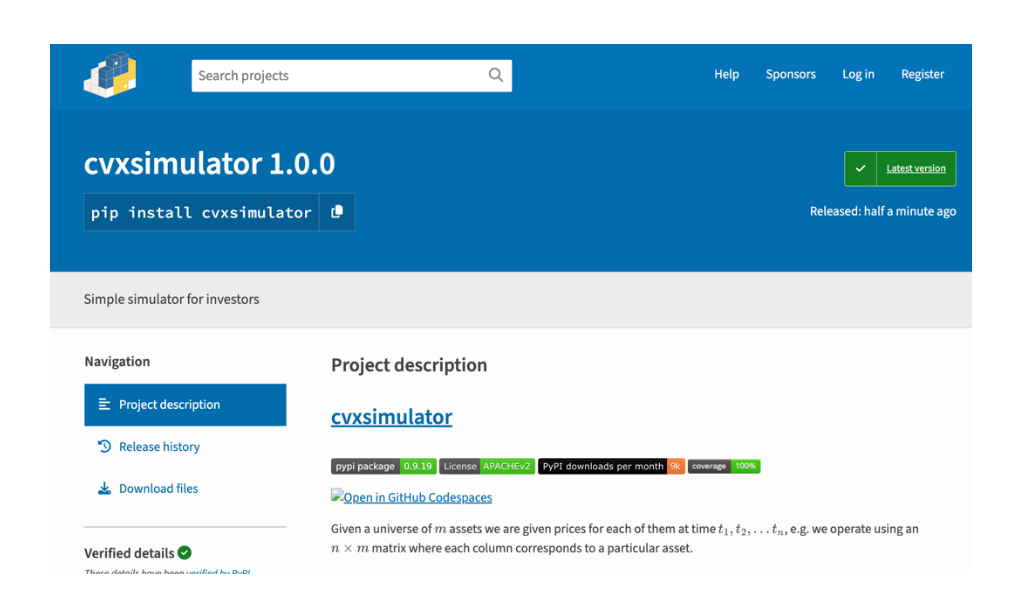
Summary

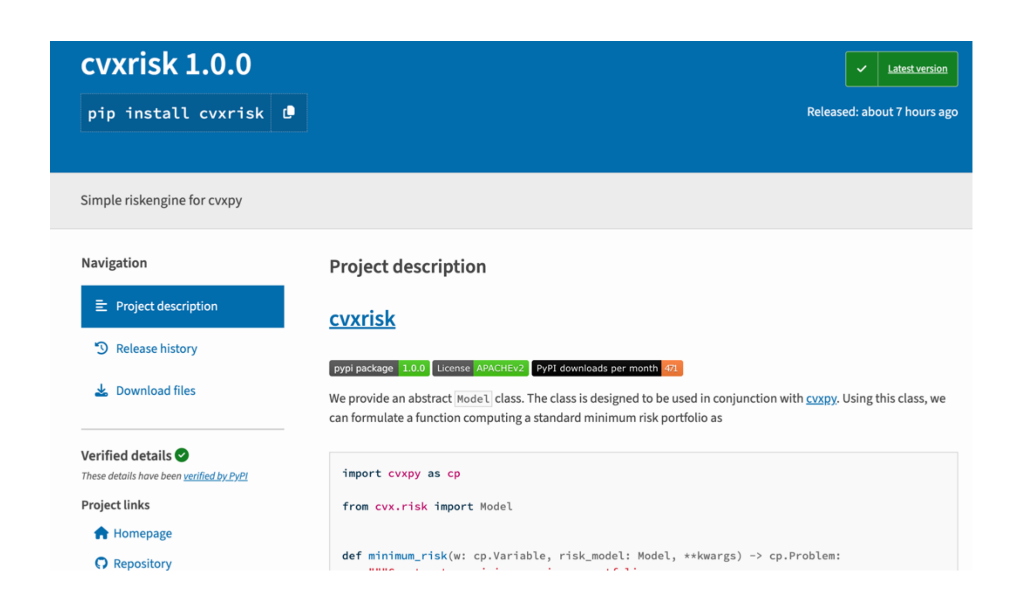
- introduced a covariance predictor for financial returns
- relies on solving a small convex optimization problem
- requires little or no tuning or fitting
- interpretable, lightweight, and practically effective
- outperforms popular EWMA and is comparable to MGARCH

Library development

in the cvx family

- We create open-source libraries to simplify your experiments with convex programming
- Examples:
  - cvxrisk
  - cvxsimulator
  - cvxcovariance
  - cvxcla
  - pmog
  - …





Conclusions

- Markowitz is sometimes misunderstood. He introduced the reference framework still in use today.
- More than 70 years after his seminal paper we understand its limitations and use cases.
- Convex programming in the spirit of Harry Max Markowitz is still very much alive and is doing well.
- We suggest some mild extensions and provide recipes to bypass the classic problems and issues.
- We then revisit the estimation of input parameters and apply convex programming also on that level.
- We finish with pointing to some of our more recent additions of software into the cvxpy family.