Convex Optimization in Quantitative Finance

Kasper Johansson

University PhD Dissertation Defense

Department of Electrical Engineering

Advisor: Stephen Boyd

December 11, 2024

PhD overview

- **research focus:** convex optimization in quantitative finance
- **key areas of study:**
  - covariance prediction
  - portfolio construction
  - statistical arbitrage trading
  - retirement funding
  - hyperparameter learning
  - portfolio construction with crypto assets
- all problems formulated and solved through **convex optimization**
  - yields global solution (and optimality certificate)
  - fast and reliable (no need to tune parameters)
  - easily specified using domain-specific languages like CVXPY

Papers

- **K. Johansson**, M. Ogut, M. Pelger, T. Schmelzer, S. Boyd. A simple method for predicting covariance matrices of financial returns. *Foundations and Trends in Econometrics*, 2023.
- S. Boyd, **K. Johansson**, R. Kahn, P. Schiele, T. Schmelzer. Markowitz portfolio construction at seventy. *Journal of Portfolio Management, Harry Markowitz Special Issue*, 2024.
- **K. Johansson**, T. Schmelzer, S. Boyd. Finding moving-band statistical arbitrages via convex-concave optimization. *Optimization and Engineering*, 2024.
- **K. Johansson**, T. Schmelzer, S. Boyd. A Markowitz approach to managing a dynamic basket of moving-band statistical arbitrages. *Journal of Asset Management*, under review.
- S. Boyd, **K. Johansson**, P. Schiele. Convex optimization in quantitative finance. *Working paper*.
- **K. Johansson**, S. Boyd. Simple and effective portfolio construction with crypto assets. *Working paper*.
- **K. Johansson**, S. Boyd. A simple model predictive control policy for retirement funding. *Working paper*.
- **K. Johansson**, S. Boyd. Learning a hyperparameter policy for a trading algorithm. *Working paper*.

Papers: Today's focus

- **K. Johansson**, M. Ogut, M. Pelger, T. Schmelzer, S. Boyd. A simple method for predicting covariance matrices of financial returns. *Foundations and Trends in Econometrics*, 2023.
- S. Boyd, **K. Johansson**, R. Kahn, P. Schiele, T. Schmelzer. Markowitz portfolio construction at seventy. *Journal of Portfolio Management, Harry Markowitz Special Issue*, 2024.
- **K. Johansson**, T. Schmelzer, S. Boyd. Finding moving-band statistical arbitrages via convex-concave optimization. *Optimization and Engineering*, 2024.
- **K. Johansson**, T. Schmelzer, S. Boyd. A Markowitz approach to managing a dynamic basket of moving-band statistical arbitrages. *Journal of Asset Management*, under review.
- S. Boyd, **K. Johansson**, P. Schiele. Convex optimization in quantitative finance. *Working paper*.
- **K. Johansson**, S. Boyd. Simple and effective portfolio construction with crypto assets. *Working paper*.
- **K. Johansson**, S. Boyd. A simple model predictive control policy for retirement funding. *Working paper*.
- **K. Johansson**, S. Boyd. Learning a hyperparameter policy for a trading algorithm. *Working paper*.

Predicting covariance matrices: Challenges & contributions

challenges

- financial applications require covariance predictors that react to changing market conditions
- trade-off between stability and reactivity

contributions

- a simple and effective method for predicting reactive covariance matrices of financial returns
- a new method for evaluating a covariance predictor over changing market conditions
- extensive empirical study on several large data sets
- open-source implementation in Python:

  `https://github.com/cvxgrp/cov_pred_finance`

Financial returns

- $r_t \in \mathbf{R}^n$ is the vector of n financial asset returns over period t
- $t = 1, \ldots, T$ are the time periods (days, weeks, months, etc.)
- $(r_t)_i$ is the return of asset i over period t
- assets could be bonds, stocks, factors, etc.

Gaussian model

model: $r_t \sim \mathcal{N}(0, \Sigma_t)$

- can demean return data if needed
- for most daily, weekly, or monthly return data

$$\Sigma_t = \mathbf{E} r_t r_t^T - (\mathbf{E} r_t)(\mathbf{E} r_t)^T \approx \mathbf{E} r_t r_t^T$$

objective: find estimate $\hat{\Sigma}_t$ of Σ_t, based on $r_1, \ldots, r_{t-1}$

Evaluating covariance predictors

- predictions $\hat{\Sigma}_1, \ldots, \hat{\Sigma}_T$ evaluated on average log-likelihood

$$\frac{1}{2T} \sum_{t=1}^{T} \left(-n \log(2\pi) - \log\det \hat{\Sigma}_t - r_t^T \hat{\Sigma}_t^{-1} r_t \right)$$

  (larger values are better)
- best constant predictor is $\Sigma^{\text{emp}} = \frac{1}{T} \sum_{t=1}^{T} r_t r_t^T$
- **log-likelihood regret** is the difference between the log-likelihood of the best constant predictor and that of the predictors $\hat{\Sigma}_1, \ldots, \hat{\Sigma}_T$ (smaller values are better)
- the regret over multiple periods removes the effect of the log-likelihood of the empirical covariance varying due to changing market conditions

Exponentially weighted moving average (EWMA) predictor

$$\hat{\Sigma}_t = \alpha_t \sum_{\tau=1}^{t-1} \beta^{t-1-\tau} r_\tau r_\tau^T, \quad t = 2, 3, \ldots$$

- $\alpha_t = \left(\sum_{\tau=1}^{t-1} \beta^{t-1-\tau}\right)^{-1} = \frac{1-\beta}{1-\beta^{t-1}}$ is the normalizing constant
- $\beta \in (0, 1)$ is the forgetting factor, often expressed in terms of the half-life $H = -\log 2/\log \beta$

Iterated EWMA (IEWMA)

1. form initial diagonal estimate $\hat{\Sigma}_t^{(1)}$, with $\left(\hat{\Sigma}_t^{(1)}\right)_{ii}$ as EWMA of $(r_t)_i^2$ using half-life H^{vol}
2. compute volatility adjusted returns

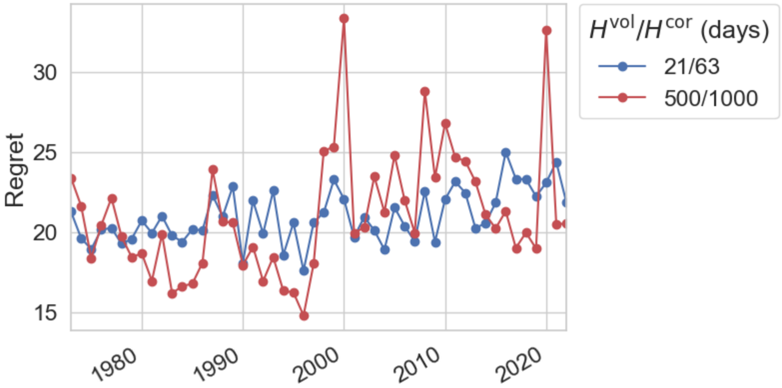
$$\tilde{r}_t = \left(\hat{\Sigma}_t^{(1)}\right)^{-1/2} r_t, \quad t = 1, \ldots, T$$

3. form $\hat{\Sigma}_t^{(2)}$ as EWMA covariance of $\tilde{r}_t$ using half-life H^{cor}
4. final estimate

$$\hat{\Sigma}_t = \left(\hat{\Sigma}_t^{(1)}\right)^{1/2} \hat{\Sigma}_t^{(2)} \left(\hat{\Sigma}_t^{(1)}\right)^{1/2}$$

- closely related to iterated covariance predictors [Barratt and Boyd, 2022]
- variation: let $\hat{\Sigma}_t^{(2)}$ be correlation matrix of $\tilde{r}_t$ [Engle, 2002]
- predictor parameters: H^{vol} and H^{cor}

IEWMA performance over time



- log-likelihood regret for a fast IEWMA and a slow IEWMA on 49 daily industry portfolios
- in volatile markets (2000, 2008, 2020) the fast IEWMA performs better (lower regret)
- in stable markets the slow IEWMA performs better

Our method

Dynamically weighted prediction combiner

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

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

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 $\hat{\Sigma}_t = \left(\hat{L}_t \hat{L}_t^T\right)^{-1}$

Choosing the weights via convex optimization

- choose weight vector 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 optimization problem [Boyd and Vandenberghe, 2004]

Combined multiple iterated EWMA (CM-IEWMA)

1. choose K half-life pairs $(H_k^{\text{vol}}, H_k^{\text{cor}})$, $k = 1, \ldots, K$
2. form the K IEWMA predictors $\hat{\Sigma}_t^{(k)}$ for these half-life pairs
3. combine the IEWMAs using the dynamically weighted prediction combiner to get the prediction $\hat{\Sigma}_t = \left(\hat{L}_t \hat{L}_t^T\right)^{-1}$

- predictor parameters: half-life pairs and lookback horizon N

Data set and experimental setup

- data: $n = 49$ daily industry portfolio returns 1970–2023, $T = 13{,}496$ trading days
- compare five covariance predictors
  - rolling window (RW) with a 500-day window
  - EWMA with 250-day half-life
  - IEWMA with half-lives $H^{\text{vol}}/H^{\text{cor}}$ of $125/250$ (in days)
  - MGARCH with parameters re-estimated annually
  - CM-IEWMA with $K = 5$ predictors with half-lives (in days):

H^{vol}	21	63	125	250	500
H^{cor}	63	125	250	500	1000

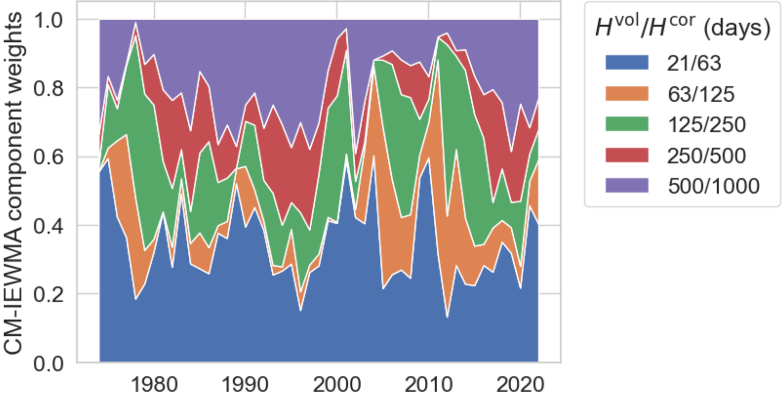
- results on other data sets like stocks and factors are qualitatively similar

Log-likelihood regret

Predictor	Average	Std. dev.	Max.
RW	20.4	6.9	72.8
EWMA	19.4	6.2	70.1
IEWMA	18.2	3.6	41.4
MGARCH	17.9	3.0	32.8
CM-IEWMA	**16.9**	**2.4**	**28.4**

- metrics on quarterly regret (over 212 quarters)
- CM-IEWMA performs best

CM-IEWMA component weights π



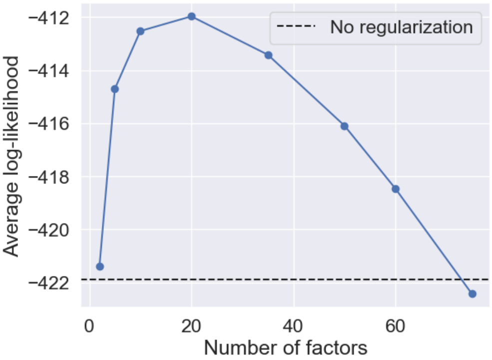
- average weight π_i, $i = 1, \ldots, 5$ on the five IEWMAs each year
- substantial weight is put on slow IEWMAs most years
- during and following volatile periods we see significant weight increase on fast IEWMAs

Extension: Factor covariance model

$$\Sigma_t = F_t \Sigma_t^{\mathrm{f}} F_t^T + D_t$$

- $F_t \in \mathbf{R}^{n \times k}$ is matrix of factor loadings
- k is number of factors, typically with $k \ll n$
- Σ_t^{f} is $k \times k$ factor covariance matrix
- D_t is diagonal matrix of unexplained (idiosyncratic) variances
- a strong regularizer which can give better return covariance estimates
- factors constructed by many methods, like principal component analysis (PCA) or by hand

Benefits of factor covariance model



- average regret on a 238-asset universe from 2000–2023
- factors estimated by PCA every year using previous two years of data

Papers: Today's focus

- **K. Johansson**, M. Ogut, M. Pelger, T. Schmelzer, S. Boyd. A simple method for predicting covariance matrices of financial returns. *Foundations and Trends in Econometrics*, 2023.
- S. Boyd, **K. Johansson**, R. Kahn, P. Schiele, T. Schmelzer. Markowitz portfolio construction at seventy. *Journal of Portfolio Management, Harry Markowitz Special Issue*, 2024.
- **K. Johansson**, T. Schmelzer, S. Boyd. Finding moving-band statistical arbitrages via convex-concave optimization. *Optimization and Engineering*, 2024.
- **K. Johansson**, T. Schmelzer, S. Boyd. A Markowitz approach to managing a dynamic basket of moving-band statistical arbitrages. *Journal of Asset Management*, under review.
- S. Boyd, **K. Johansson**, P. Schiele. Convex optimization in quantitative finance. *Working paper*.
- **K. Johansson**, S. Boyd. Simple and effective portfolio construction with crypto assets. *Working paper*.
- **K. Johansson**, S. Boyd. A simple model predictive control policy for retirement funding. *Working paper*.
- **K. Johansson**, S. Boyd. Learning a hyperparameter policy for a trading algorithm. *Working paper*.

Markowitz portfolio construction: Challenges & contributions

challenges

- Markowitz portfolio construction balances risk and return through convex optimization
- the basic version can be sensitive to estimation errors, often producing impractical portfolios

contributions

- collect minimal set of constraints and extensions from prior work to address practical issues
  - constraints on leverage, turnover, etc. [Grinold & Kahn, 2000]
  - address uncertainty with robust optimization [Ben-Tal, El Ghaoui, & Nemirovski, 2009]
  - incorporate soft constraints in optimization problems [Bertsimas & Brown, 2011]
- novel method for how to prioritize constraints
- extension preserves convexity
- extensive empirical evaluation on historical data

Basic Markowitz optimization

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

- variable $w \in \mathbf{R}^n$ of portfolio weights
- $\mu \in \mathbf{R}^n$ and $\Sigma \in \mathbf{S}^n_{++}$ are asset return mean and covariance
- σ^{tar} is target (per period) volatility
- basic form goes back to [Markowitz, 1952]

```
w = cp.Variable(n)
objective = mu.T @ w
constraints = [cp.quad_form(w, Sigma) <= sigma**2, cp.sum(w) == 1]
prob = cp.Problem(cp.Maximize(objective), constraints)
prob.solve()
```

Critiques of Markowitz optimization

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

- sensitivity to data errors and estimation uncertainty
- risk symmetry
- maximizing expected utility versus mean-variance
- statistical assumptions: assumes Gaussian returns, and ignores higher moments
- greedy method, only looks one step ahead

Critiques of Markowitz optimization

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

- sensitivity to data errors and estimation uncertainty
- risk symmetry
- maximizing expected utility versus mean-variance
- statistical assumptions: assumes Gaussian returns, and ignores higher moments
- greedy method, only looks one step ahead

- we address the first issue of sensitivity to data errors and estimation uncertainty
- the other critiques seem less relevant in practice [Luxenberg and Boyd, 2023]

Adding practical constraints and objective terms

- include cash holdings c, previous holdings w^{pre}, trades $z = w - w^{\text{pre}}$
- account for (convex) holding costs ϕ^{hold} and trading costs ϕ^{trade}
- limit weights, cash, trades, turnover $T = \|z\|_1$, and leverage $L = \|w\|_1$

$$\begin{array}{ll} \text{maximize} & \mu^T w - \gamma^{\text{hold}} \phi^{\text{hold}}(w, c) - \gamma^{\text{trade}} \phi^{\text{trade}}(z) \\ \text{subject to} & \mathbf{1}^T w + c = 1, \quad z = w - w^{\text{pre}}, \\ & w^{\text{min}} \leq w \leq w^{\text{max}}, \quad c^{\text{min}} \leq c \leq c^{\text{max}}, \quad L \leq L^{\text{tar}}, \\ & z^{\text{min}} \leq z \leq z^{\text{max}}, \quad T \leq T^{\text{tar}}, \\ & \|\Sigma^{1/2} w\|_2 \leq \sigma^{\text{tar}} \end{array}$$

Adding practical constraints and objective terms

- include cash holdings c, previous holdings w^{pre}, trades $z = w - w^{\text{pre}}$
- account for (convex) holding costs ϕ^{hold} and trading costs ϕ^{trade}
- limit weights, cash, trades, turnover $T = \|z\|_1$, and leverage $L = \|w\|_1$

$$
\begin{array}{ll}
\text{maximize} & \mu^T w - \gamma^{\text{hold}} \phi^{\text{hold}}(w, c) - \gamma^{\text{trade}} \phi^{\text{trade}}(z) \\
\text{subject to} & \mathbf{1}^T w + c = 1, \quad z = w - w^{\text{pre}}, \\
& w^{\text{min}} \leq w \leq w^{\text{max}}, \quad c^{\text{min}} \leq c \leq c^{\text{max}}, \quad L \leq L^{\text{tar}}, \\
& z^{\text{min}} \leq z \leq z^{\text{max}}, \quad T \leq T^{\text{tar}}, \\
& \|\Sigma^{1/2} w\|_2 \leq \sigma^{\text{tar}}
\end{array}
$$

remaining challenges (and solutions)

- Σ is estimated in factor covariance form; estimating μ is difficult and typically proprietary
- optimization is sensitive to errors in μ and Σ (use robustification)
- constraints may lead to infeasibility or unnecessary trading (use soft constraints)

Computational benefits of factor model

- with factor model, cost of portfolio optimization reduced from $O(n^3)$ to $O(nk^2)$ flops [Boyd and Vandenberghe, 2004]
- easily exploited in modeling languages like CVXPY
- timings for Clarabel open source solver:

		solve time (s)	
assets n	factors k	factor model	full covariance
100	10	0.002	0.040
300	20	0.010	0.700
1000	30	0.080	25.600
3000	50	0.600	460.000

Robustifying Markowitz

- basic Markowitz optimization can be sensitive to estimation errors in μ, Σ
- replace mean return $\mu^T w$ with worst-case return

$$R^{\text{wc}} = \min\{(\mu+\delta)^T w \mid |\delta| \leq \rho\} = \mu^T w - \rho^T |w|$$

  where $\rho \geq 0$ is vector of mean return uncertainties
- replace risk $w^T \Sigma w$ with worst-case risk

$$\begin{aligned}(\sigma^{\text{wc}})^2 &= \max\{w^T(\Sigma+\Delta)w \mid |\Delta_{ij}| \leq \kappa(\Sigma_{ii}\Sigma_{jj})^{1/2}\} \\ &= \sigma^2 + \kappa\left(\sum_{i=1}^n \Sigma_{ii}^{1/2}|w_i|\right)^2\end{aligned}$$

  where $\kappa \geq 0$ represents covariance uncertainty
- easily handled by CVXPY

Softening constraints

- **soft constraints** allow limited violations of constraints, based on priority
- to soften a constraint $f \leq f^{\text{max}}$, replace it with a penalty term $\gamma(f - f^{\text{max}})_+$ in the objective
- in Markowitz risk, leverage, and turnover can be softened, giving three priority parameters

$$\gamma^{\text{risk}}, \quad \gamma^{\text{lev}}, \quad \gamma^{\text{turn}}$$

- the softened problem reduces unnecessary trading and is always feasible

Softening constraints

- **soft constraints** allow limited violations of constraints, based on priority
- to soften a constraint $f \leq f^{\text{max}}$, replace it with a penalty term $\gamma(f - f^{\text{max}})_+$ in the objective
- in Markowitz risk, leverage, and turnover can be softened, giving three priority parameters

$$\gamma^{\text{risk}}, \quad \gamma^{\text{lev}}, \quad \gamma^{\text{turn}}$$

- the softened problem reduces unnecessary trading and is always feasible

choosing priority parameters

- can be chosen or initialized based on Lagrange multipliers of hard constrained problem
- *e.g.*, as 80th percentile of recorded multipliers over a historical period
- fast solve time enables backtesting to fine-tune parameters

Data and experimental setup

- S&P 100 stocks, data gathered daily from 2000-01-04 to 2023-09-22
- exclude stocks without data for the full period (gives $n = 74$ assets)
- simulated but realistic mean predictions, and EWMA covariance
- priority parameters retuned each year based on the previous two years of data

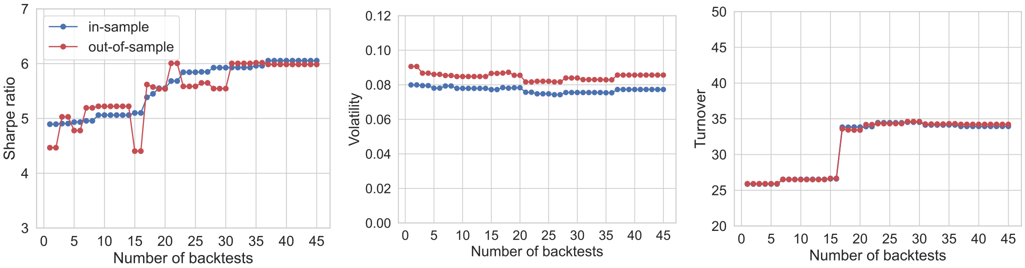
- **focus**: relative performance comparison of methods, not real portfolio construction

Parameter tuning (in-sample)

0. initialize $\gamma^{\text{hold}} = \gamma^{\text{trade}} = 1$, and $\gamma^{\text{risk}}, \gamma^{\text{lev}}, \gamma^{\text{turn}}$ based on hard constraint Lagrange multipliers
1. cycle through parameters, increasing the parameter ($+25\%$), one at a time
2. keep changes if all of the following hold:
   - the in-sample Sharpe ratio increases
   - the in-sample annualized turnover is no more than 100
   - the in-sample maximum leverage is no more than 2
   - the in-sample annualized volatility is no more than 15%

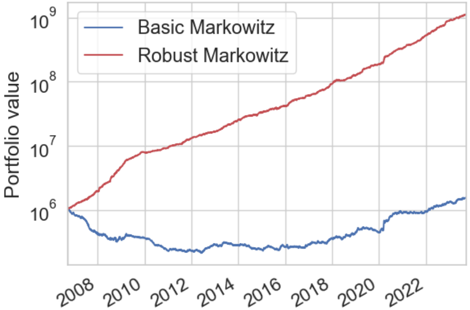
   if not, decrease the parameter (-20%) and check if the metrics improve
3. repeat 1–2 until convergence

Tuning



- figures show typical effect of tuning on Sharpe ratio, volatility, and turnover
- in-sample: April 19, 2016 to March 19, 2018
- out-of-sample: March 20, 2018 to March 4, 2019
- Sharpe ratio increases from around 4.5 to 6.0, while other metrics stay within bounds

Portfolio performance



Metric	Basic	Robust
Return	3.5%	38.1%
Risk	14.4%	8.6%
Sharpe	0.2	4.6
Drawdown	80%	6%

- out-of-sample portfolio performance for basic Markowitz and robust Markowitz

Conclusions

- convex optimization shown to be effective in quantitative finance
- covariance prediction
  - introduced a simple and effective convex optimization-based predictor
  - requires minimal tuning, is interpretable, lightweight, and effective
  - outperforms popular benchmark models
- Markowitz portfolio construction
  - extended to include practical constraints (*e.g.*, leverage, turnover, ...)
  - addressed estimation errors with robust optimization
  - leveraged soft constraints to reduce trading and ensure feasibility
- thesis also considers: statistical arbitrage trading, retirement funding, hyperparameter tuning, and crypto assets

Acknowledgments

Acknowledgments

- Stephen Boyd

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn
- collaborators: Giray Ogut, Philippe Schiele, Thomas Schmelzer

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn
- collaborators: Giray Ogut, Philippe Schiele, Thomas Schmelzer
- cvxgrp

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn
- collaborators: Giray Ogut, Philippe Schiele, Thomas Schmelzer
- cvxgrp
- many friends: Daniel, Max, Giray, Pelle, Otto, Gustaf, Wille, Jacob, David, Gustav, Axel, Johan, Filip, Vincent, Jesper, Michael, Jonatan, and more

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn
- collaborators: Giray Ogut, Philippe Schiele, Thomas Schmelzer
- cvxgrp
- many friends: Daniel, Max, Giray, Pelle, Otto, Gustaf, Wille, Jacob, David, Gustav, Axel, Johan, Filip, Vincent, Jesper, Michael, Jonatan, and more
- Maya P ♥

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn
- collaborators: Giray Ogut, Philippe Schiele, Thomas Schmelzer
- cvxgrp
- many friends: Daniel, Max, Giray, Pelle, Otto, Gustaf, Wille, Jacob, David, Gustav, Axel, Johan, Filip, Vincent, Jesper, Michael, Jonatan, and more
- Maya P ♥
- mom and dad

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn
- collaborators: Giray Ogut, Philippe Schiele, Thomas Schmelzer
- cvxgrp
- many friends: Daniel, Max, Giray, Pelle, Otto, Gustaf, Wille, Jacob, David, Gustav, Axel, Johan, Filip, Vincent, Jesper, Michael, Jonatan, and more
- Maya P ♥
- mom and dad
- grandma and grandpa

Acknowledgments

- Stephen Boyd
- my PhD committee: Mert Pilanci, Markus Pelger, Sanjay Lall, Ron Kahn
- collaborators: Giray Ogut, Philippe Schiele, Thomas Schmelzer
- cvxgrp
- many friends: Daniel, Max, Giray, Pelle, Otto, Gustaf, Wille, Jacob, David, Gustav, Axel, Johan, Filip, Vincent, Jesper, Michael, Jonatan, and more
- Maya P ♥
- mom and dad
- grandma and grandpa
- finally, my best friend and little brother

Thank you!

Appendix

Covariance prediction: Some practical extensions and variations

- realized covariance
  - uses intraperiod returns
- large universes
  - when n is larger than 100 or so
- smoothing
  - penalize variation in covariance estimate

Realized covariance

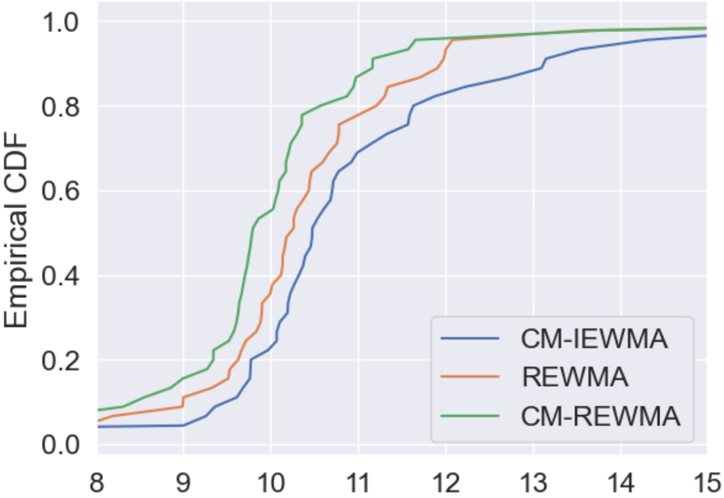
- $r_t \in \mathbf{R}^{n \times m}$ return matrix at time t, with columns that are m intraperiod return vectors
- $C_t = r_t r_t^T$ realized covariance at time t
- realized EWMA (REWMA):

$$\hat{\Sigma}_t = \alpha_t \sum_{\tau=1}^{t-1} \beta^{t-1-\tau} C_\tau, \quad t = 2, 3, \ldots,$$

- CM-REWMA combines REWMAs with different half-lives

Realized covariance empirical results

- $n = 39$ stocks and $m = 77$ intraperiod returns, January 2 2004 to December 30 2016
- CM-IEWMA gives improvement here too



Large universes

- in practice, the number of assets n can be very large
- we describe two closely related methods for large universes
  - traditional factor model
  - fitting a factor model to a (given) covariance matrix
- computational cost of portfolio optimization reduced from $\mathcal{O}(n^3)$ to $\mathcal{O}(nk^2)$ when using a k-factor model [Boyd and Vandenberghe, 2004]

Traditional factor model

- model: $r_t = F_t f_t + z_t, \quad t = 1, 2, \ldots,$
  - $F_t \in \mathbf{R}^{n \times k}$ factor loadings
  - $f_t \in \mathbf{R}^k$ factor returns
  - $z_t \in \mathbf{R}^n$ idiosyncratic return
- we end up with covariance of low-rank plus diagonal form

$$\Sigma_t = F_t \Sigma_t^{\mathrm{f}} F_t^T + E_t$$

  - Σ_t^{f} factor return covariance
  - E_t diagonal matrix of idiosyncratic variances
- never have to store $n \times n$ covariance

Fitting a factor model to a covariance matrix

- given covariance Σ
- find one in factor form, $\hat{\Sigma} = FF^T + E$, such that the Kullback-Leibler divergence between $\mathcal{N}(0, \Sigma)$ and $\mathcal{N}(0, \hat{\Sigma})$,

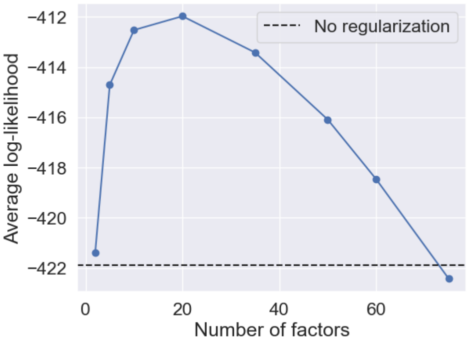
$$\mathcal{K}(\Sigma, \hat{\Sigma}) = \frac{1}{2}\left(\log \frac{\det \hat{\Sigma}}{\det \Sigma} - n + \mathbf{Tr}\, \hat{\Sigma}^{-1}\Sigma\right)$$

  is minimized
- equivalent to maximizing the expected log-likelihood of $r \sim N(0, \Sigma)$ under the model $\mathcal{N}(0, \hat{\Sigma})$
- can be solved via the expectation maximization algorithm (suggested and derived by Emmanuel Candès)

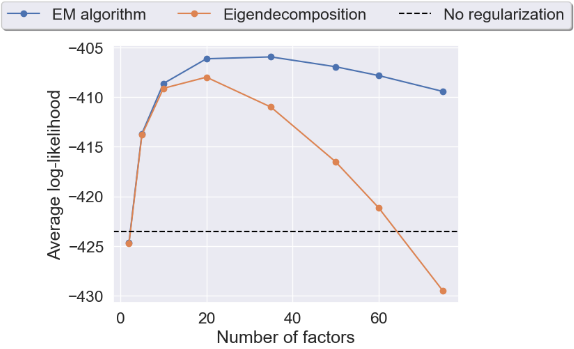
Large universes: empirical setup

- 238 US stocks over 5787 trading days
- traditional factor model
  - create factor model using PCA on two years of data, refitted annually
  - we use k factors and use the CM-IEWMA with half-lives (in days) $H^{\text{vol}}/H^{\text{cor}}$ of $\lceil k/2 \rceil / k$, $k/3k$, and $3k/6k$, to compute the factor covariance
- fitting factor model to covariance
  - use CM-IEWMA directly with half-lives (in days) $H^{\text{vol}}/H^{\text{cor}}$ of 63/125, 125/250, 250/500, and 500/1000
  - approximate CM-IEWMA predictor using factor model

Large universes: empirical results



traditional factor model



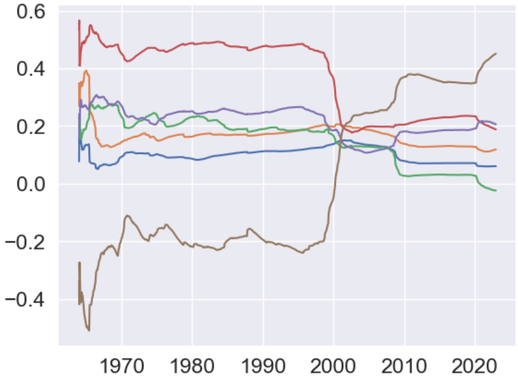
fitting factor model to covariance

Smooth covariance predictions

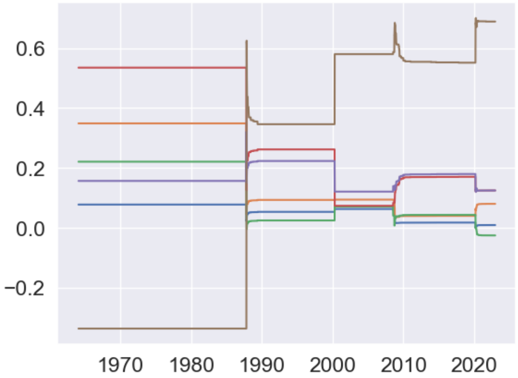
- given predictions $\hat{\Sigma}_t$, $t = 1, 2, \ldots,$
- let $\hat{\Sigma}_t^{\text{sm}}$ be the EWMA of $\hat{\Sigma}_t$
  - equivalent to minimizing
  $$\left\|\hat{\Sigma}_t^{\text{sm}} - \hat{\Sigma}_t\right\|_F^2 + \lambda \left\|\hat{\Sigma}_t^{\text{sm}} - \hat{\Sigma}_{t-1}^{\text{sm}}\right\|_F^2,$$
  where λ is a smoothing parameter
  - yields smooth covariance predictions
- with regularizer $\lambda\|\hat{\Sigma}_t^{\text{sm}} - \hat{\Sigma}_{t-1}^{\text{sm}}\|_F$, we obtain piecewise constant predictions
- smoothing can lead to reduced trading and improved portfolio performance

Smooth covariance predictions empirical results

- minimum variance portfolios on five Fama-French factor returns
- portfolio weights for smooth and piecewise constant covariances



smooth



piecewise constant

Try it out!

```
from cvx.covariance.combination import from_ewmas
halflife_pairs = [(10, 21), (21, 63), (63, 125)]
combinator = from_ewmas(returns, halflife_pairs)
covariances = {}
for predictor in combinator.solve(window=10):
    covariances[predictor.time] = predictor.covariance
```

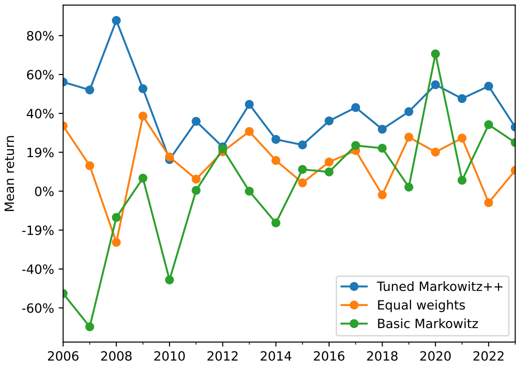
https://github.com/cvxgrp/cov_pred_finance

Taming Markowitz

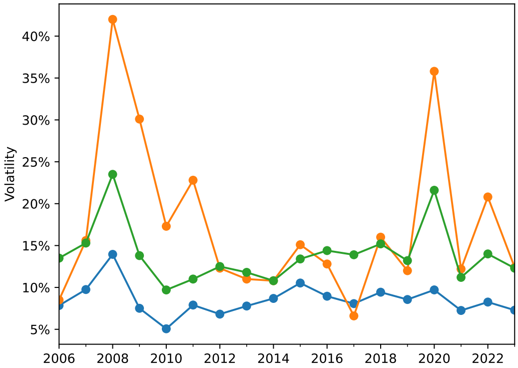
	Return	Volatility	Sharpe	Turnover	Leverage	Drawdown
Equal weight	14.1%	20.1%	0.66	1.2	1.0	50.5%
Basic Markowitz	3.7%	14.5%	0.19	1145.2	9.3	78.9%
Weight-limited	20.2%	11.5%	1.69	638.4	5.1	30.0%
Leverage-limited	22.9%	11.9%	1.86	383.6	1.6	14.9%
Turnover-limited	19.0%	11.8%	1.54	26.1	6.5	25.0%
Robust	15.7%	9.0%	1.64	458.8	3.2	24.7%
Markowitz++	38.6%	8.7%	4.32	28.0	1.8	7.0%
Tuned Markowitz++	41.8%	8.8%	4.65	38.6	1.6	6.4%

- portfolio performance for various modifications to basic Markowitz optimization
- Markowitz++ refers to our proposed extension

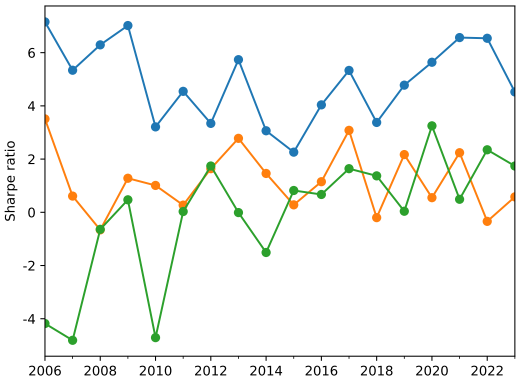
Annual returns



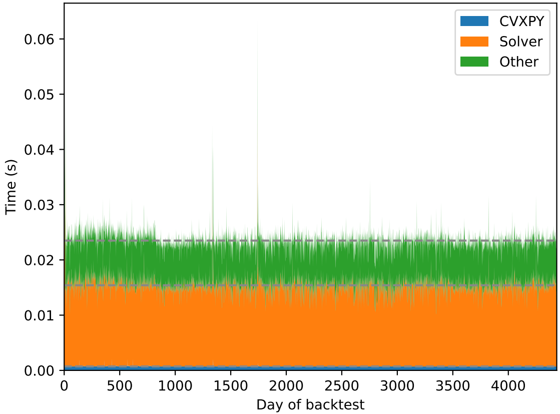
Annual volatilities



Annual Sharpes



Solve times


CVXPY
Solver
Other
Time (s)
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0
500
1000
1500
2000
2500
3000
3500
4000
Day of backtest