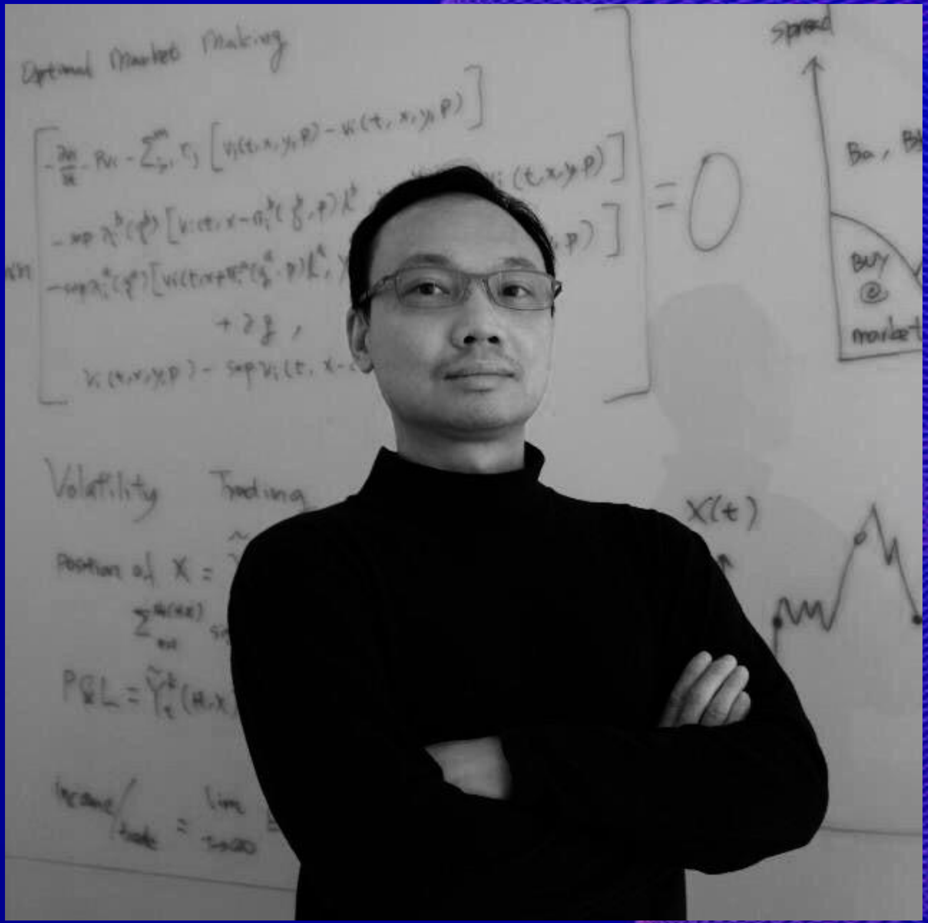
NMFinTech
WWW.NMFIN.TECH

New Technologies in Portfolio Optimization

Prof. Haksun Li

haksun.li@nmfin.tech

https://nmfin.tech



Speaker Profile

Dr. Haksun Li

- CEO, NM FinTech LTD.
- Vice Dean, Fudan University ZhongZhi Big Data Finance and Investment Institute
- (Ex-)Adjunct Professors, Industry Fellow, Advisor, Consultant with the National University of Singapore, Nanyang Technological University, Fudan University, the Hong Kong University of Science and Technology
- Quantitative Trader/Analyst, BNPP, UBS
- Ph.D., Artificial Intelligence, University of Michigan, Ann Arbor
- M.S., Financial Mathematics, University of Chicago
- B.S., Mathematics, University of Chicago

NM FinTech
WWW.NMFIN.TECH

Modern Portfolio Theory

Why Portfolio Optimization

FoF asset allocation

- How much capital to assign to each fund?

Portfolio asset allocation

- How much capital to assign to each strategy?

Alpha strategy asset allocation

- How much capital to assign to each stock?

Harry Markowitz

It all starts with
Markowitz in 1952…

Won the Nobel Memorial Prize in Economic Sciences in 1990.

- Standard textbook model
- Widely taught in universities
- MBA courses

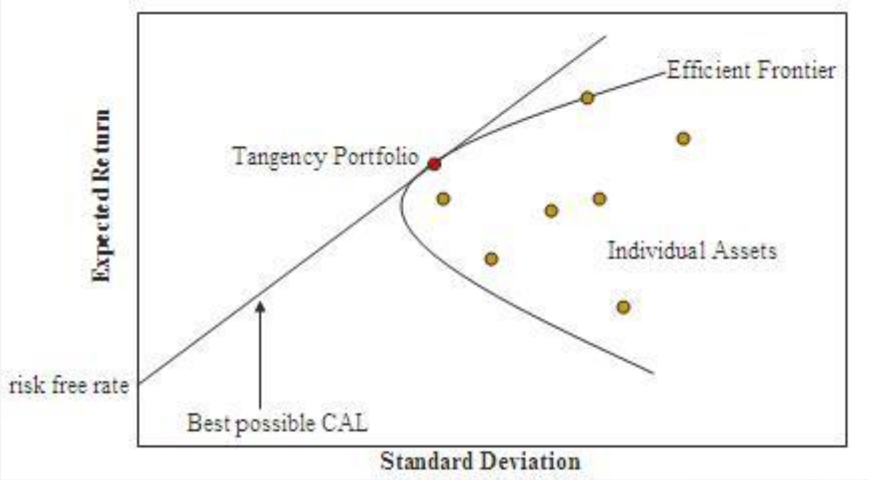
Modern Portfolio Theory – Insights

1. An asset's risk and return should be assessed by how it contributes to a portfolio's overall risk and return, but not just by itself.
2. Mean-Variance (MV) optimization
3. Investors are risk averse, meaning that given two portfolios that offer the same expected return, investors will prefer the less risky one.
4. An investor who wants higher expected returns must accept more risk.
5. An investor can have individual risk aversion characteristics in terms of the risk (tolerance) parameter.

Modern Portfolio Theory – Math

- $\max_{\boldsymbol{\omega}}\{\boldsymbol{\omega}\,\mathbf{E}(\boldsymbol{r}_{t+1}) - \boldsymbol{\lambda}\boldsymbol{\omega}'\boldsymbol{\Sigma}_t\boldsymbol{\omega}\}$
  - $\mathrm{E}(r_{t+1})$ is the expected return for the **next** period, a known quantity
  - Σ_t is the covariance matrix for the assets, a known matrix
  - where ω is the optimal portfolio weights, found by solving a deterministic problem
- **Constraints: $\boldsymbol{A\omega} \le \boldsymbol{b}$**
  - No short selling: $\omega \ge 0$
- **Alternatively, we have**
  - $\min_{\omega}\{\omega'\Sigma_t\omega - \lambda\omega\,\mathrm{E}(r_{t+1})\}$
- **Solution: Quadratic Programming**
- **NM:**
  - http://redmine.numericalmethod.com/projects/public/repository/svn-algoquant/show/core/src/main/java/com/numericalmethod/algoquant/model/portfoliooptimization/markowitz

Efficient Frontier



- Given
  - $\omega\,\mathrm{E}(r_{t+1}) = \mu$
- Find ω s.t.,
  - $\omega_{eff} = \underset{\omega}{\operatorname{argmin}}\{\lambda\omega'\Sigma_t\omega\}$

Problems with Markowitz Portfolio Optimization

NM FinTech
WWW.NMFIN.TECH

Estimation Problem with Markowitz's Theory

- Require the knowledge of means and covariances.
  - Too many parameters to estimate: $N + \frac{N^2+N}{2}$.
    - For $N = 300$, we have 45,450 parameters to estimate.
    - For $N = 3000$, we have 4,504,500 parameters to estimate.
    - Chopra & Ziemba (1993) shows that errors in means are about 10x as important as errors in variances, and errors in variances are about 2x important as errors in covariances.
  - Time varying. Tied to business cycles.

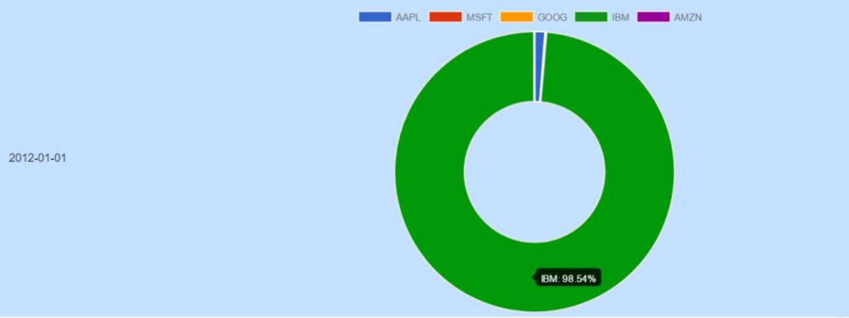
Problems with Sample Covariance Matrix

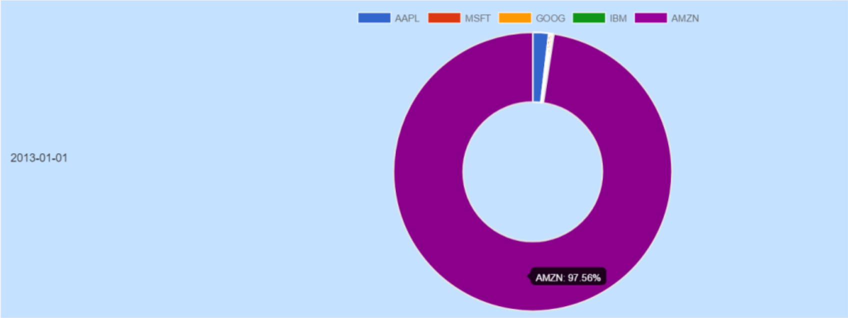
- A sample covariance matrix is often ill-conditioned, nearly singular, sometimes not even invertible and sometimes not even positive semi-definite.
  - dimension: p, number of samples: n
  - $\frac{p}{n} > 1$, matrix not invertible
  - $\frac{p}{n} < 1$ but not negligible, matrix ill-conditioned
- Linear dependency among stocks.
  - asynchronous data
  - incomplete data
  - artificial changes due to stress-tests
- Error Maximization:
  - Largest sample eigenvalues are systematically biased upwards.
  - Smallest sample eigenvalues are systematically biased downwards.
  - Inverting a sample covariance matrix increases significantly the estimation error.
  - Capital allocated to the extreme eigenvalues where they are most unreliable.

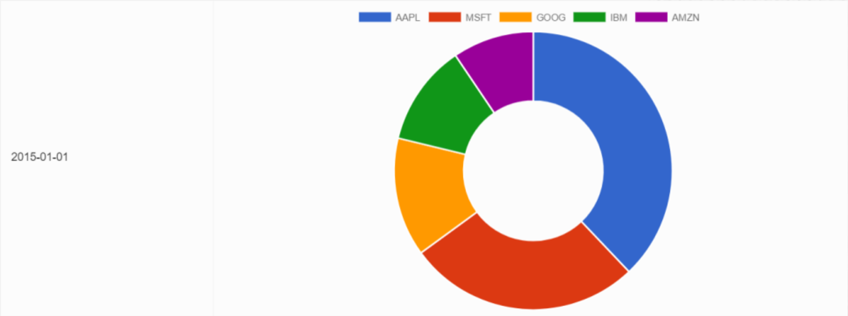
Problems with Sample Mean

- Sample mean is only an estimation using TWO data points, namely the TWO end points, regardless of how big the sample size is.
- Given a set of historical returns $\{r_1, \ldots, r_t\}$, the sample mean is
- $\bar{r} = \sum_{i=1}^{t} r_i$
- $\approx \sum_{i=1}^{t} \log(1 + r_i) = \sum_{i=1}^{t} \log(p_i) - \log(p_{i-1})$
- $= \log(p_t) - \log(p_0)$
- Assume returns follow Gaussian distribution.
  - Nassim Nicholas Taleb: After the stock market crash (in 1987), they rewarded two theoreticians, Harry Markowitz and William Sharpe, who built beautifully Platonic models on a Gaussian base, contributing to what is called Modern Portfolio Theory. Simply, if you remove their Gaussian assumptions and treat prices as scalable, you are left with hot air. The Nobel Committee could have tested the Sharpe and Markowitz models—they work like quack remedies sold on the Internet—but nobody in Stockholm seems to have thought about it.

Problems with Diversification







- Litterman & et al. (1992, 1999, 2003):
  - When unconstrained, portfolios will have large long and short positions.
  - When subject to long only constraint, capital is allocated only to a few assets.
- Best & Grauer (1991):
  - A small increase in expected return can consume half of the capital.

Problems with Constraints

- $\max_{\omega} \omega \, \mathrm{E}(r_{t+1})$, s.t.,
- $\omega' \Sigma_t \omega \leq \sigma_{MAX}$
- $1'\omega = 1$
- $\omega \geq 0$

Cannot model non-linear constraints such as market impact.

$\sum_{j \in S_i} \left| \omega_j^0 + \omega_j \right| \leq u_i$ for sector $i = 1, \ldots, S$

Minimizing variance

Market impact

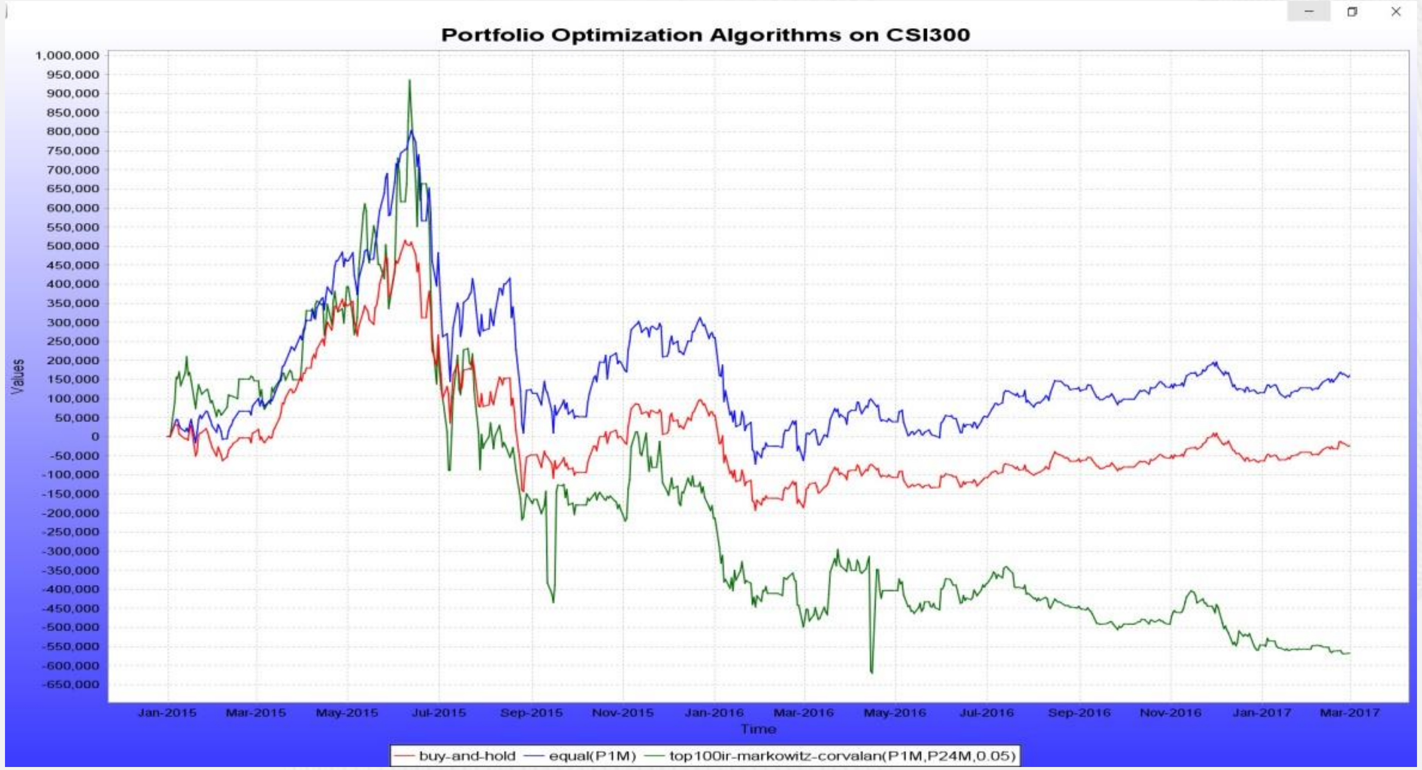
Diversification constraints (sector exposure)

Tax, transaction costs, etc.

Objective function:

$$\max_{\omega} \left\{ \omega \, \mathrm{E}(r_{t+1}) - \lambda_P \omega' \Sigma_t \omega - \lambda_M \sum_{j=1}^{n} \left(m_j |\omega_j|^{\frac{3}{2}} + c_j |\omega_j| \right) \right\}$$

Problem with Performance



- P&L is often worse than the 1/N strategy (equal weighting).

Comments on Markowitz

Wesley Gray: Although Markowitz did win a Nobel Prize, and this was partly based on his elegant mathematical solution to identifying mean-variance efficient portfolios, a funny thing happened when his ideas were applied in the real world: mean-variance performed poorly. The fact that a Nobel-Prize winning idea translated into a no-value-add-situation for investors is something to keep in mind when considering any optimization method for asset allocation ...complexity does not equal value!

Classical Solutions for Portfolio Optimization

Estimating Covariance Matrix

Solutions to Estimating Covariance – Dimension Reduction

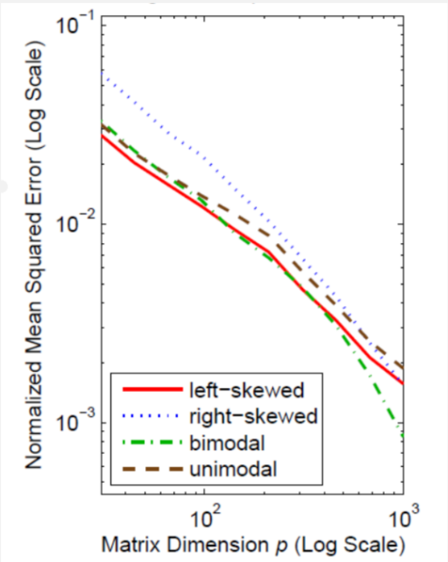
- Dimension reduction via multifactor models
  - Relate the *i*-th asset returns r_i to *k* factors $f_1, \ldots, f_k$ by
  - $r_i = \alpha_i + (f_1, \ldots, f_k)'\beta_i + \epsilon_i$
  - α_i, β_i are unknown regression parameters; ϵ_i unobserved random noise with mean 0 and are uncorrelated.
  - $\mathrm{Cov}\left(r_{it}, r_{jt}\right) = \beta_{it}' \, \mathrm{V}(f) \, \beta_{jt}' + \mathrm{Cov}\left(\epsilon_{it}, \epsilon_{jt}\right)$
  - E.g., alpha strategy, Fama-French model, CAPM, APT
  - NM:
    - http://redmine.numericalmethod.com/projects/public/repository/svn-algoquant/show/core/src/main/java/com/numericalmethod/algoquant/model/factormodel

Solutions to Estimating Covariance – Shrinkage Estimators

- Pull the extreme eigenvalues back to the mean.
- Ledoit and Wolf (2003, 2004):
  - $\hat{\Sigma} = \hat{\delta}\hat{F} + (1 - \hat{\delta})S$
    - $\hat{\delta}$ is an estimator of the optimal shrinkage constant
    - $\hat{F}$ is given by mean of the prior distribution or a structured covariance matrix, which has much fewer parameters than $N + \frac{N^2+N}{2}$.
    - S the sample covariance
  - NM:
    - https://nm.dev/html/javadoc/nmdev/dev/nm/stat/covariance/LedoitWolf2004.html
    - https://nm.dev/html/javadoc/nmdev/dev/nm/stat/covariance/nlshrink/LedoitWolf2016.html
- Ledoit and Wolf (2012): nonlinear shrinkage

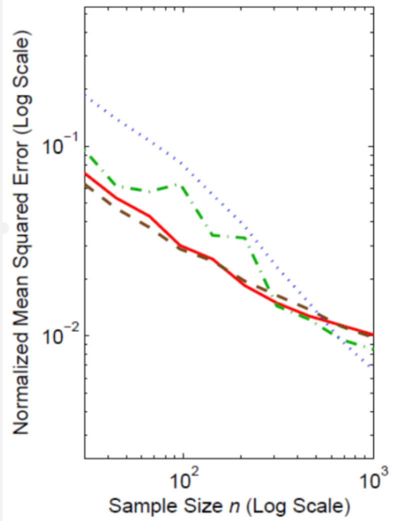
Non-Linear Shrinkage

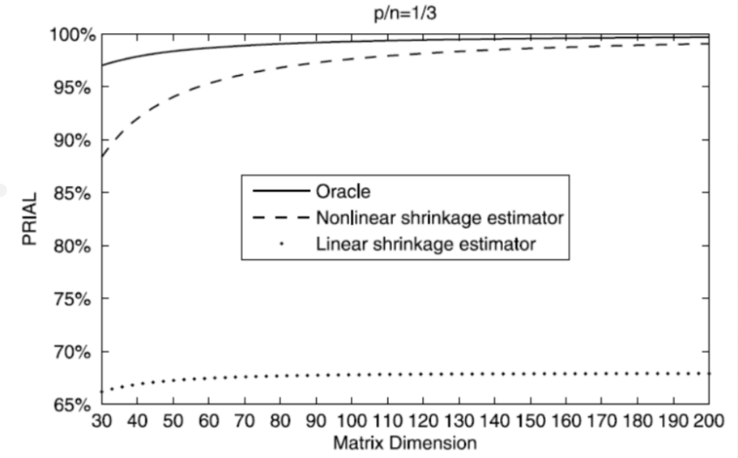
c=1/3



c=2

more dimension than data



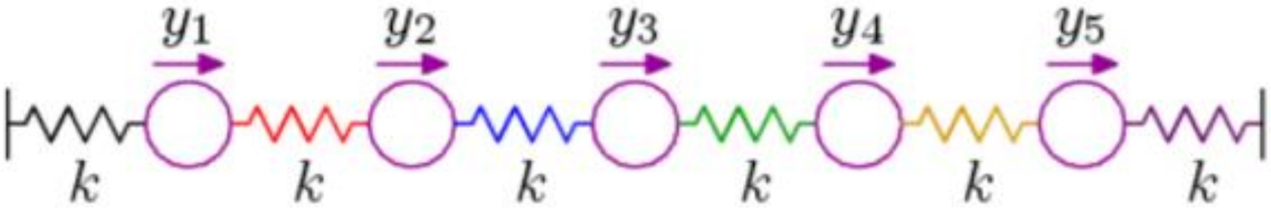


- Ledoit and Wolf (2016): non-linear shrinkage
- sample eigenvalues:
$$\lambda_{n,1} \le \lambda_{n,2} \le \cdots \le \lambda_{n,p}$$
- estimated true eigenvalues:
$$\tau_{n,1} \le \tau_{n,2} \le \cdots \le \tau_{n,p}$$
- estimation problem as minimization:
$$\widehat{\tau_n} = \underset{t\in[0,\infty]}{\operatorname{argmin}} \frac{1}{p}\sum_{1}^{p}\left[q_{n,p}^{i}(t) - \lambda_{n,i}\right]^2$$
- the best known method

QuEST Transformation

- $Q_{n,p}(\boldsymbol{t}) = \left(q_{n,p}^{1}(\boldsymbol{t}), \dots, q_{n,p}^{p}(\boldsymbol{t})\right)$
  - mapping from true eigenvalues to sample eigenvalues
- $q_{n,p}^{i}(\boldsymbol{t}) = p \int_{(i-1)/p}^{i/p} \left(F_{n,p}^{t}\right)^{-1}(v)dv$
- $\left(F_{n,p}^{t}\right)^{-1}(v)$: the inverse of $F_{n,p}^{t}(x)$
- $F_{n,p}^{t}(x) = \begin{cases} \max\left(1 - \frac{n}{p}, \frac{1}{p}\sum_{i=1}^{p} 1_{\{t_i=0\}}\right), x = 0 \\ \lim\limits_{\eta \to 0+} \frac{1}{\pi} \int_{-\infty}^{x} \mathrm{Im}\left[m_{n,p}^{t}(\xi + i\eta)\right] d\xi \end{cases}$
  - the limiting empirical cdf of sample eigenvalues
- $m_{n,p}^{t}(z)$ is the unique solution to this equation
  - $m = \frac{1}{p}\sum_{i=1}^{p} \frac{1}{t_i\left(1 - \frac{p}{n} - \frac{p}{n} zm\right) - z}$

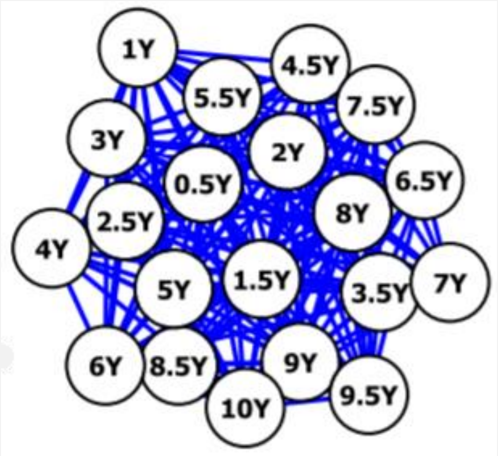
Inverse Covariance Matrix vs Covariance Matrix

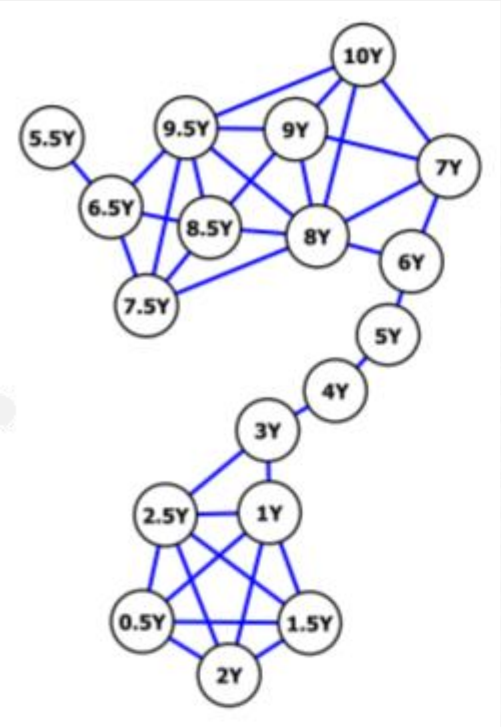


inverse-covariance matrix or covariance matrix?

$$\mathbf{K}^{-1} = \frac{k}{T}\begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} \qquad \mathbf{K} = \frac{T}{k}\begin{bmatrix} 0.83 & 0.67 & 0.50 & 0.33 & 0.17 \\ 0.67 & 1.33 & 1.00 & 0.67 & 0.33 \\ 0.50 & 1.00 & 1.50 & 1.00 & 0.50 \\ 0.33 & 0.67 & 1.00 & 1.33 & 0.67 \\ 0.17 & 0.33 & 0.50 & 0.67 & 0.83 \end{bmatrix}$$

Solutions to Estimating Covariance – Covariance Selection





- $\max_X \log\det X - \mathrm{Tr}(\Sigma X) - \rho \,\mathrm{Card}\, X$
- Awoye, OA; (2016): Graphical LASSO
- Dempster (1972): the covariance structure of a multivariate normal population can be simplified by setting elements of the inverse of the covariance matrix to zero
- NM:
  - https://nm.dev/html/javadoc/nmdev/dev/nm/stat/covariance/covarianceselection/lasso/CovarianceSelectionGLASSOFAST.html
  - https://nm.dev/html/javadoc/nmdev/dev/nm/stat/covariance/covarianceselection/lasso/CovarianceSelectionLASSO.html

Estimating Mean

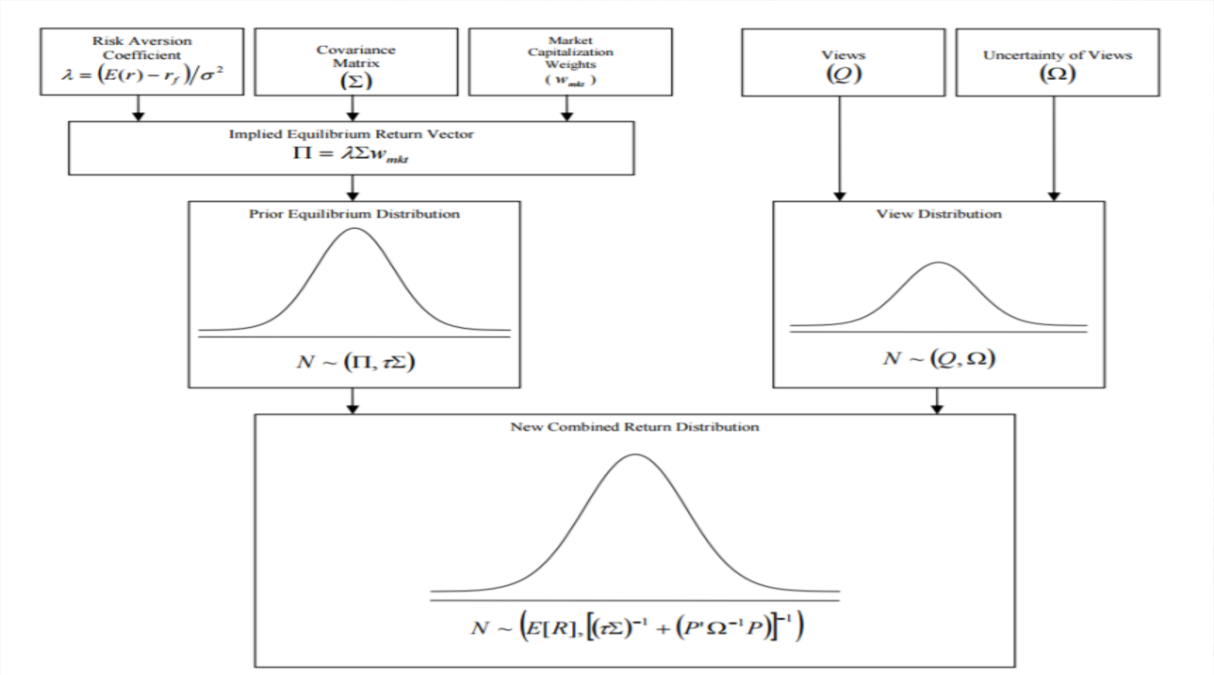
Solutions to Estimating Mean – Statistical Methods

1. Trading signals
2. Multifactor models
3. Shrinkage

$r_i = \alpha_i + (f_1, \dots, f_k)'\beta_i + \epsilon_i$

NM:
http://redmine.numericalmethod.com/projects/public/repository/svn-algoquant/show/core/src/main/java/tech/nmfin/algoquant/model

Solutions to Estimating Mean – Black-Litterman

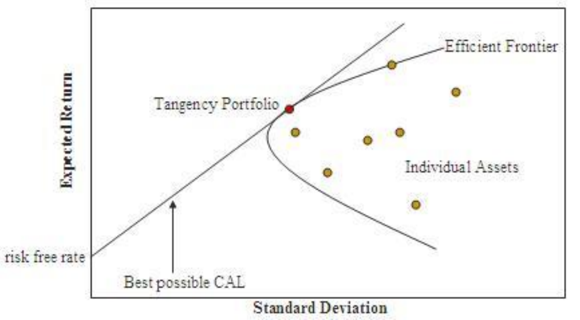


- Combined Return Vector
  - $\mathrm{E}(R) = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q]$
    - P: a matrix that identifies the assets involved in the views $(K \times N)$
    - Ω: a diagonal covariance matrix of error terms from the expressed views representing the uncertainties in each view $(K \times K)$
    - Π: the implied equilibrium return vector $(N \times 1)$
    - Q: the view vector $(K \times 1)$

Diversification

Solutions to Diversification – Almost Efficient Portfolios

- MVO intends to give an optimized portfolio in terms of risk-reward
- MVO does not intend to give a diversified portfolio
- Many portfolios on the efficient frontier are indeed concentrated
- However, there are many well diversified portfolios within a small neighborhood of the efficient frontier
- Almost Efficient Portfolios:
  - $\max_{\omega} D(\omega)$ s.t., (*D* is the diversification criterion.)
    - $\sqrt{\omega'\Sigma\omega} \leq \sigma^{\text{eff}} + \Delta\sigma$, relaxation of portfolio variance
    - $R^{\text{eff}} - \Delta R \leq \omega' r$, relaxation of portfolio expected return
    - $1'\omega = 1$



- NM:
  - https://nmfin.tech/2013/06/19/solving-the-corner-solution-problem-of-portfolio-optimization/
  - https://nm.dev/html/javadoc/nmdev/tech/nmfin/portfoliooptimization/corvalan2005/Corvalan2005.html

Solutions to Diversification – Using Constraints

- **Black-Litterman**
- **Diversification constraints, e.g.,**
  - lower and upper bounds
  - sector exposure

Non-Linear Constraints

NM FinTech
WWW.NMFIN.TECH

Second Order Conic Programming

- $\min_x f'x$, s.t.,
  - $\|A_i x + b_i\|_2 \leq c_i' x + d_i,\ i = 1, \ldots, m$
  - $Fx = g$
- LP, QP
- Solution: interior point method

Second Order Conic Programming with Constraints

- Market impact
  - $\sum_{j=1}^{n}\left(m_j|\omega_j|^{\frac{3}{2}}\right) \le t_2$

$$
\begin{aligned}
||0||_2 &\le t_2 - \sum_{j=1}^{n} \beta_j, \\
||0||_2 &\le \bar{x}_j - (y_j - w_j^0),\ j = 1, \cdots, n, \\
||0||_2 &\le \bar{x}_j - (-y_j + w_j^0),\ j = 1, \cdots, n, \\
|| \begin{pmatrix} \bar{x}_j \\ \frac{\beta_j}{m_j} - \frac{s_j}{2} \end{pmatrix} ||_2 &\le \frac{\beta_j}{m_j} + \frac{s_j}{2},\ j = 1, \cdots, n, \\
|| \begin{pmatrix} s_j \\ \frac{1 - \bar{x}_j}{2} \end{pmatrix} ||_2 &\le \frac{1 + \bar{x}_j}{2},\ j = 1, \cdots, n.
\end{aligned}
$$

- Diversification constraints (sector exposure)
  - $\sum_{j \in S_i}\left|\omega_j^0 + \omega_j\right| \le u_i$ for sector $i = 1, \dots, S$

$$||0||_2 \le -\sum_{j \in S_i} \bar{y}_j + u_i \iff ||A_i^\top z + C_i||_2 \le b_i^\top z + d_i, \quad i = 1, \cdots, k$$

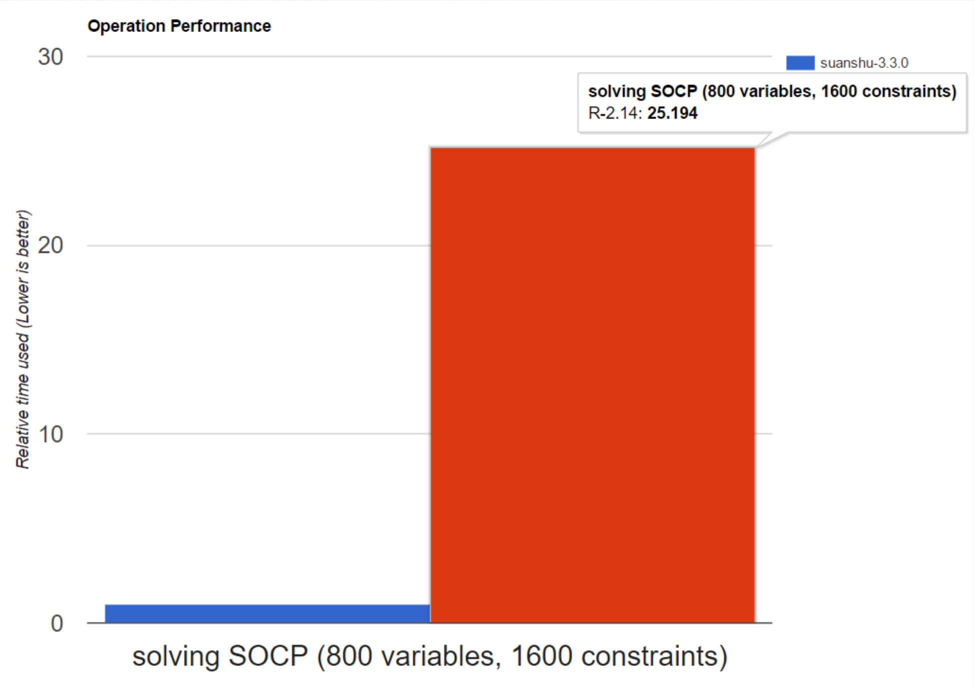
$$A_i^\top = 0_{1 \times n},\ C_i = 0,\ b_i = -\sum_{j \in S_i} e_j,\ d_i = u_i,\ z = \bar{y},$$

- Many other constraints can be modeled as SOCP constraints.
- NM has a collection of them.

- com.numericalmethod.suanshu.model.portfoliooptimization.socp.constraints
  - SOCPMaximumLoan.java
  - SOCPNoTradingList1.java
  - SOCPSectorNeutrality.java
  - SOCPSelfFinancing.java
  - SOCPZeroValue.java
- com.numericalmethod.suanshu.model.portfoliooptimization.socp.constraints.ybar
  - SOCPNoTradingList2.java
  - SOCPSectorExposure.java

Optimizer Comparison

- Numerical Method Optimizers
  - 25 times faster than free optimizers, e.g., R
- MOSEK
- Gurobi
- CPLEX
- XPRESS

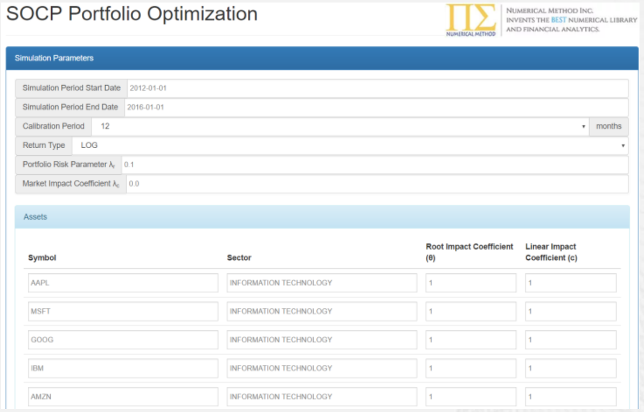


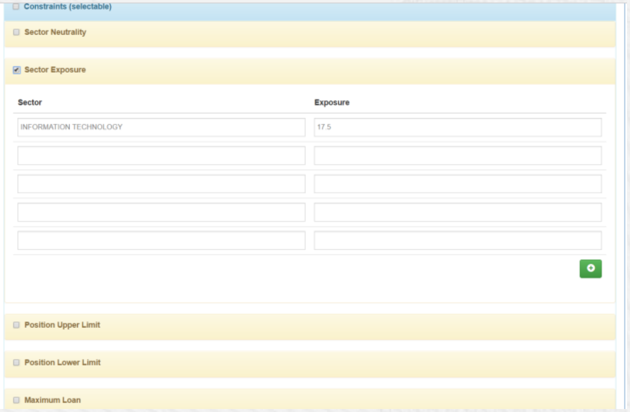
NM Portfolio Optimization System

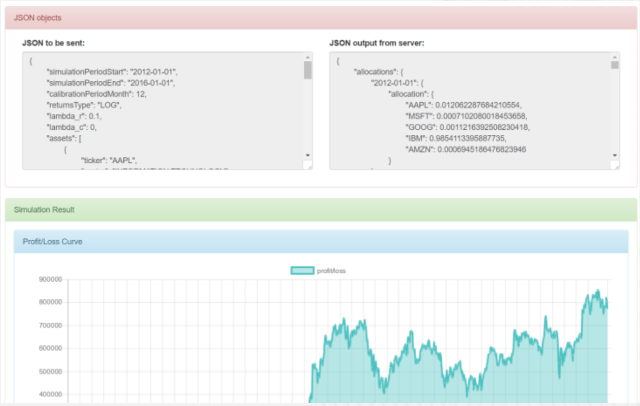
Solution to Performance – Better Estimations

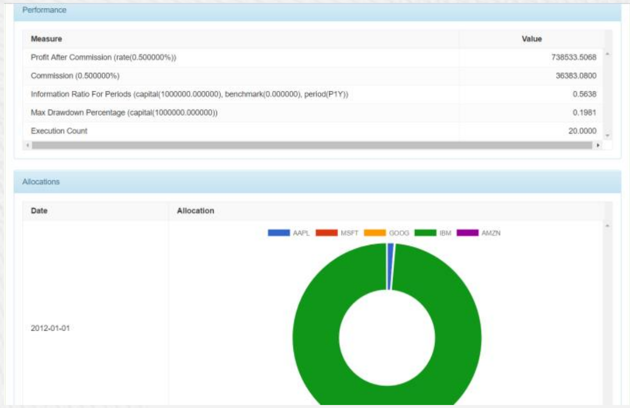
- We combine all the NM modules and algorithms to create better MVO models.
  - Better mean estimation
  - Better covariance estimation
  - Better constraint modeling
  - Better diversification criterion
- NM MVO comparison framework:
- http://redmine.numericalmethod.com/projects/public/repository/svn-algoquant/show/core/src/main/java/tech/nmfin/algoquant/model/portfoliooptimization/simulation

NM Portfolio Optimizer









- https://portoptim.nmfin.tech/

SuanShu AlgoQuant QuantitativeModels CQI NM FinTech WWW.NMFIN.TECH

Stochastic Portfolio Optimization Models

NM FinTech
WWW.NMFIN.TECH

Root Cause for Model Failure

- The historical returns that we observe are random. They are just one realization of the true/unknown/unobservable joint probability distribution of the assets traded.
- From the same joint probability distribution, what we observe is just one possibility of the unknown random process. We could have observed a different reality.
- Any estimators, e.g., mean, covariance, that we estimate from the past returns are therefore also random.
  - Classical theories never account for that randomness.

The Missing Portfolio Variance

- We write portfolio variance as such
  - $\mathrm{Var}(W) = E[\mathrm{Var}(W|R_n)] + \mathrm{Var}(\mathrm{E}(W|R_n))$
    $= \mathrm{E}(w^T \Sigma_n w) + \mathrm{Var}(w^T \mu_n)$
  - Using Σ_n to replace Σ and assuming that we know μ and Σ essentially ignores the second term.
  - That is the root cause of all Markowitz based model.

Portfolio Optimization Fundamentally Stochastic

- A “good” portfolio optimization model should therefore be fundamentally stochastic, dealing explicitly with the fact that the observed past returns are random.
- We offer two implementations in AlgoQuant:
  - Solving a stochastic optimization problem assuming mean and covariance are random (Lai and etc. 2011)
  - A functional approach to iteratively improve over a baseline model (Tsang and He 2020)

Stochastic Optimization

- We should optimize a portfolio with respect to the true/underlying/unknown/stochastic joint probably distribution.
- All classical theories optimize a portfolio with respect to the (random) observed past data and assume that they are true and deterministic.

Stochastic Optimization

NM FinTech
WWW.NMFIN.TECH

Mean and Covariance Cannot Be Known

Markowitz Objective:

- $\max_{w}\{\mathrm{E}(w' r_{n+1}) - \lambda \,\mathrm{Var}(w' r_{n+1})\}$
- can be transformed into a "mean versus second moment" optimization problem
  w(η)
  $= \arg\min_{w}\{\lambda \mathrm{E}[(w^T r_n + 1)^2] - \eta \mathrm{E}(w^T r_n + 1)\}$

Cause for Failure: require the knowledge of expected returns and covariances

- Estimation of covariance may be possible
- Estimation of expected returns is probably the single most challenging problem in finance

Let $W = w' r_{n+1}$, W the portfolio return, we have

- $$\begin{aligned}\mathrm{Var}(W) &= \mathrm{E}(\mathrm{Var}(W|R_n)) + \mathrm{Var}(\mathrm{E}(W|R_n)) \\ &= \mathrm{E}(w'\Sigma_n w) + \mathrm{Var}(w'\mu_n)\end{aligned}$$
- When expected returns and covariances are known, the 2nd term is mistakenly 0

SuanShu AlgoQuant QuantitativeModels CQI NM FinTech WWW.NMFIN.TECH

Stochastic Optimization

- $\max_{w}\{\mathrm{E}(w'r_{n+1}) - \lambda\,\mathrm{Var}(w'r_{n+1})\}$
  - Note w is now random because its value depends on the (past) observations which are themselves random
  - Thus, w is now inside the expectation and variance operators; fundamentally different from classical theories, where w is outside the operators
  - $w'r_{n+1}$ is random
  - We can add in other constraints, e.g., long-only, cardinality
- The fundamental innovation is that
  - we assume that expected returns and covariance are not available and therefore are random
  - w is random and depends on the (random) past returns
  - This is a radically different approach than most theories developed in the last 60 years

Solution to Standard Stochastic Optimization Problem

- A standard stochastic control problem in the Bayesian setting
  - $\max_{a \in A} \mathrm{E}\, \mathrm{g}(X, \theta, a)$
  - $g(X, \theta, a)$ is the reward when action a is taken; the maximization is over the action space A
- The key to its solution is the law of conditional expectations
  - $Eg(X, \theta, a) = E\{E[g(X, \theta, a)|X]\}$
  - The stochastic optimization problem can be solved by choosing a to maximize the posterior reward
  - However, this cannot be applied to maximizing nonlinear functions such as $[Eg(X, \theta, a)]^2$ in our problem

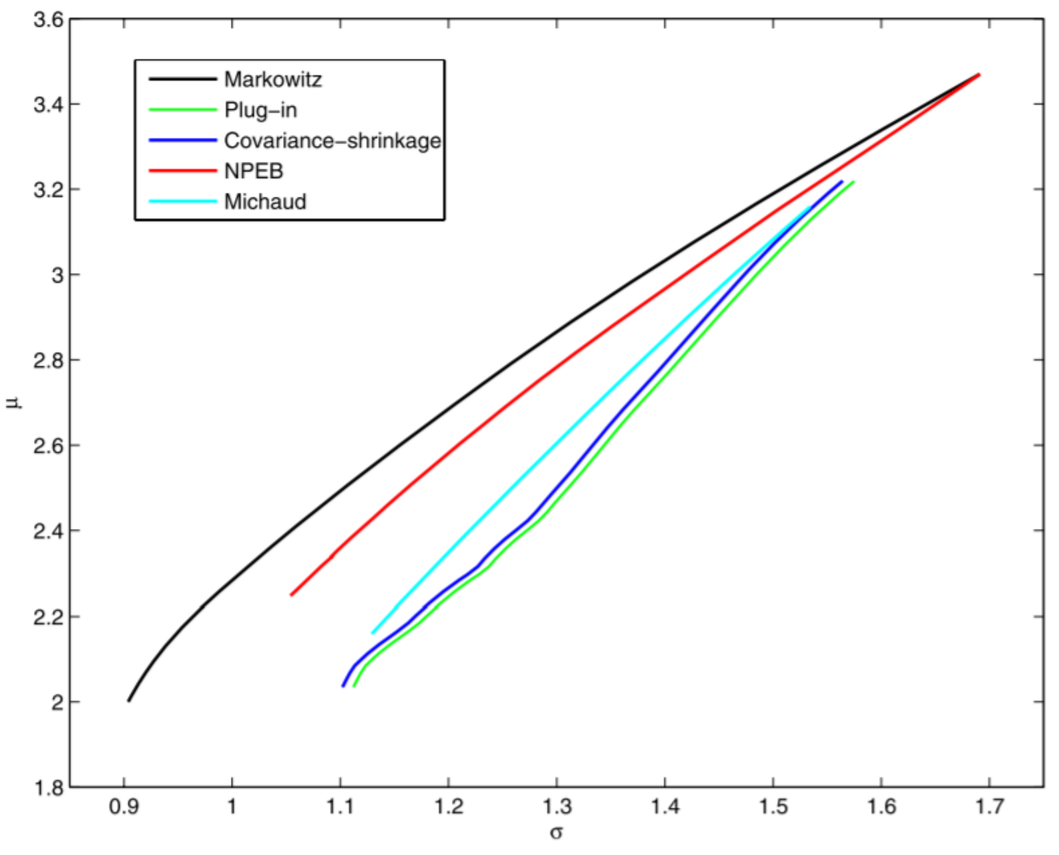
Conversion to Standard Problem

- Let $W = w^T r_{n+1}$
- Note that $E(W) - \lambda Var(W) = h(EW, EW2)$
  - where $h(x, y) = x + \lambda x^2 - \lambda y$
- Let $W_B = w_B^T r_{n+1}$ and $\eta = 1 + 2\lambda E(W_B)$
  - where w_B is the Bayes weight vector
- $h(EW, EW^2) - h(EW_B, EW_B^2)$
  $= \{E(W) - E(W_B)\} - \lambda\{E(W^2) - E(W_B^2)\} + \lambda\{(EW)^2 - (EW_B)^2\}$
  $= \eta\{E(W) - E(W_B)W\} + \lambda\{E(W_B^2) - E(W^2)\} + \lambda\{E(W) - E(WB)\}^2$
  $\geq \{\lambda(W_B^2) - \eta E(W_B)\} - \{\lambda E(W^2) - \eta E(W)\}$
- Hence $\lambda E(W_B^2) - \eta E(W_B) \leq \lambda E(W^2) - \eta E(W)$
- Therefore, we want to find the η that minimizes the equation.

Solution to Stochastic Portfolio Optimization

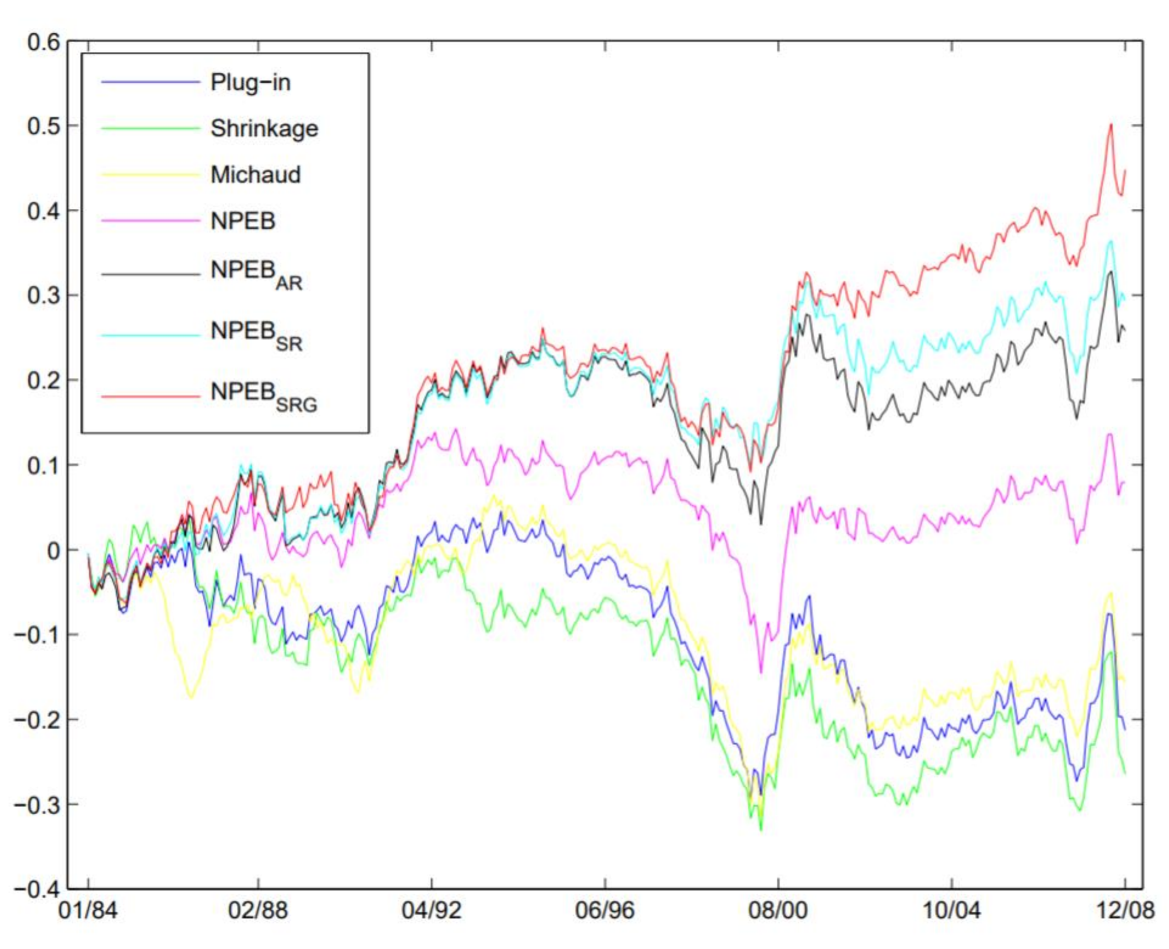
- We solve an equivalent problem
  - $\max_{\eta}\{\mathrm{E}(w(\eta)' r_{n+1}) - \lambda \,\mathrm{Var}(w(\eta)' r_{n+1})\}$
- where w(η) is the solution of the stochastic optimization problem
  - $w(\eta) = \arg\min_{w}\{\lambda \mathrm{E}[(w^T r_{n+1})^2] - \eta \mathrm{E}(w^T r_{n+1})\}$
  - $\eta = 1 + 2\lambda \mathrm{E}(WB)$

Solution to Performance – Unknown Mean and Unknown Covariance



(σ, μ) *curves of different portfolios.*

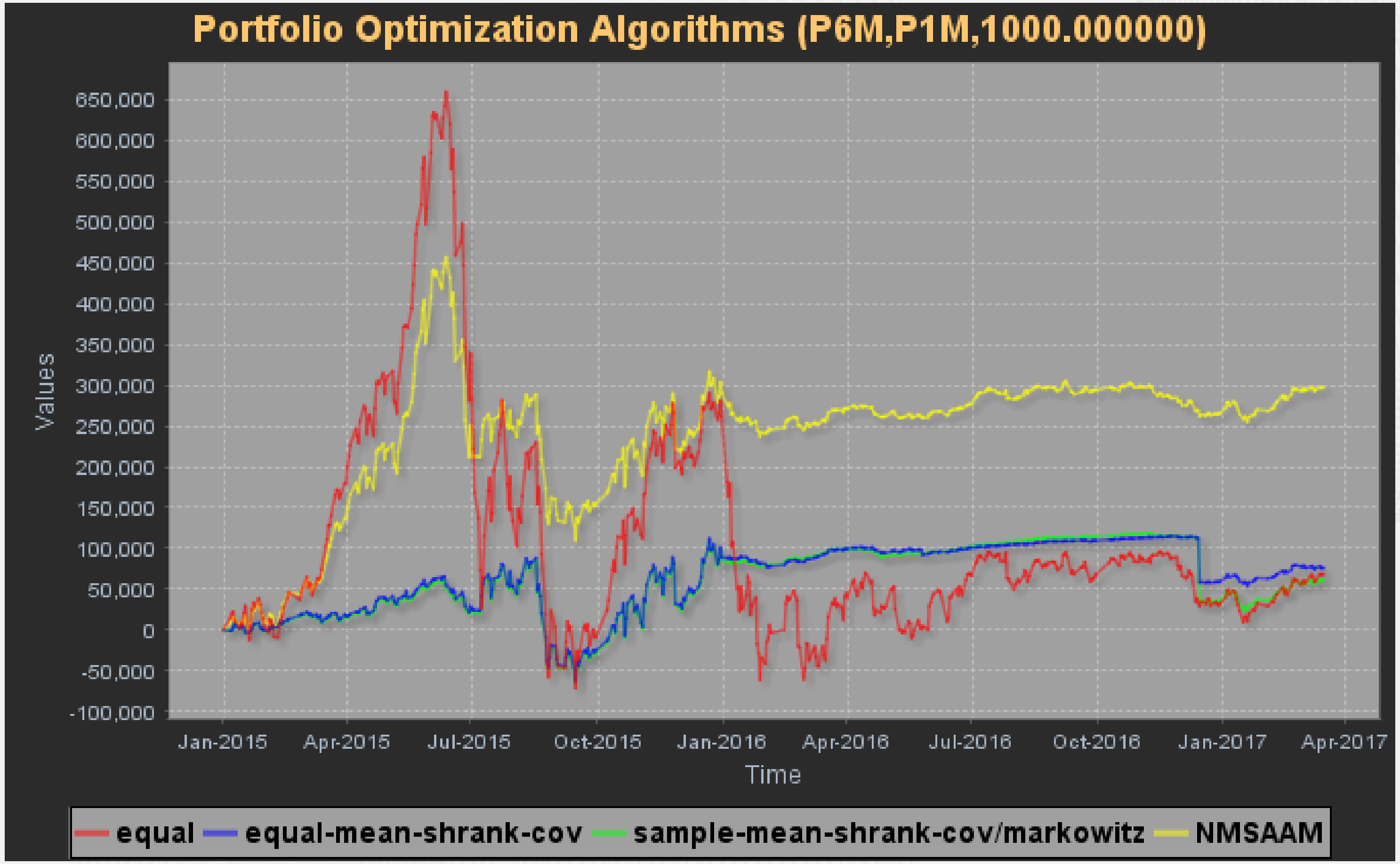
Realized Cumulative Returns Over Time – Unknown Mean and Unknown Covariance



Empirical Results

- Using a pool of ETFs provided by a bank in China, we compare the asset allocation results of SAAM to 1/N (and other models)
- Our results show that SAAM out performs 1/N in a wide range of parameters, hence stability
- SAAM properties:
  - When the market rallies big, it rides the trend but not to the extreme extend, giving good and risk-controlled profits
  - **When the market crashes, it limits the drawdown**
  - When the market walks sideway (since 2016/1), it has a much smaller volatility and smoother pnl curve

Empirical Results {rebalance=6M, data=1M}



SAAM Stability

rebalance	window	riskAverse	algorithm	Profit After Com	Commission (0.2	Information Ratio	Max Drawdown I	BEAT benchmark	rebalance	window	riskAverse	algorithm	Profit After (	Commission	Information Ratio F	Max Drawdc
P6M	P9M	0.1	NMSAAM	**$ 232,292.77**	$ 20,281.10	1.128315787	41.89%	1	P6M	P9M	0.1	equal	**$31,549.45**	$16,181.39	0.122890693	44.34%
P6M	P3M	1	NMSAAM	**$ 173,566.64**	$ 19,505.77	0.319557842	36.74%	1	P6M	P3M	1	equal	**$59,338.54**	$16,172.34	0.344467733	43.99%
P6M	P1M	0.1	NMSAAM	**$ 94,333.03**	$ 20,268.17	0.216543966	27.77%	1	P6M	P1M	0.1	equal	**$52,142.09**	$16,115.82	0.134276425	44.06%
P6M	P3M	0.1	NMSAAM	**$ 86,539.34**	$ 18,613.82	0.240494521	37.00%	1	P6M	P3M	0.1	equal	**$59,338.54**	$16,172.34	0.344467733	43.99%
P6M	P6M	0.1	NMSAAM	**$ 69,081.43**	$ 18,224.59	0.165673319	33.91%	1	P6M	P6M	0.1	equal	**$37,666.03**	$16,249.19	0.039309408	44.21%
P6M	P6M	1	NMSAAM	**$ 68,408.35**	$ 18,258.34	0.182644292	33.24%	1	P6M	P6M	1	equal	**$37,666.03**	$16,249.19	0.039309408	44.21%
P6M	P12M	1	NMSAAM	**$ 67,882.62**	$ 18,425.08	0.268401727	22.86%	1	P6M	P12M	1	equal	**$29,682.01**	$16,114.07	0.06498543	43.83%
P6M	P9M	1	NMSAAM	**$ 61,818.51**	$ 18,209.09	1.631970017	31.70%	1	P6M	P9M	1	equal	**$31,549.45**	$16,181.39	0.122890693	44.34%
P6M	P12M	0.1	NMSAAM	**$ 50,359.00**	$ 18,163.79	0.234910757	21.58%	1	P6M	P12M	0.1	equal	**$29,682.01**	$16,114.07	0.06498543	43.83%
P6M	P1M	1	NMSAAM	**$ 47,450.11**	$ 20,160.90	0.119946006	31.74%	0	P6M	P1M	1	equal	**$52,142.09**	$16,115.82	0.134276425	44.06%
			avg	**$ 95,173.18**	$ 19,011.06	**0.450845823**	**31.84%**					avg	**$42,075.62**	$16,166.56	**0.141185938**	**44.09%**
			stdev	$ 60,150.56	$ 931.40							stdev	$12,321.51	$ 52.51		
			max	**$ 232,292.77**	$ 20,281.10							max	**$59,338.54**	$16,249.19		
			min	$ 47,450.11	$ 18,163.79							min	$29,682.01	$16,114.07		

- From the table, we see that for {rebalance = 6M}, almost all (except 1) parameter sets out perform 1/N
  - The average PnL is 2.26x of that of equal-weighting
  - The max profit is 3.91x, the min profit 1.6x
  - The Sharpe ratio is 0.45 vs. 0.14
- We repeat the analysis for the whole (reasonable) parameter space
  - SAAM out performs 1/N 55.56% in terms of final PnL
  - Average PnL $64,265 vs. $53,052, max drawdown 27.72% vs 42.79%， Sharpe ratio 0.26 vs. 0.21
- **Conclusion: SAAM consistently out performs equal-weighting and is the probably the best asset allocation theory to date**

Dynamic Index

- Using SZ50 China as the benchmark, we compare the NM index to it from 2005/1 to 2017/12
- Our results show that
  - SAAM out performs the market index in an absolute majority of wide range of parameters hence stability
  - and in terms of a number of measures hence lower risk

Empirical Results (SZ50, China)

	SZ50	all parameters	selected plateau
average return (daily)	0.06%	0.06%	0.06%
stdev return (daily)	1.79%	1.62%	
SR (annualized)	0.4855	0.5983	0.6178
better (SR)	0	93.75%	100.00%
dSR/SR	0.00%	23.41%	27.11%
better (cumpnl)	0	73.44%	100.00%
dPnl/PnL	0	8.63%	18.87%

- Even without filtering and comparing using all (reasonable) parameters, our dynamic portfolios already outperform SZ50 consistently
  - 93.75% of our portfolios better than the index in terms of Sharpe
  - 73.44% of our portfolios better than the index in terms of final PnL
- By selecting a stable parameter plateau, our result shows that NM index is unambiguously superior than the original index
  - 100% of our portfolios better than the index in terms of Sharpe and PnL
  - Sharpe ratio 27.11% better than the index
  - Final PnL 18.87% better than the index
- In summary, not only does NM index provides a better return in terms of PnL, it also provides a lower risk alternative than the index in terms of lower volatility and lower max drawdown

Functional Optimization

Functional weighting maximizer

- Remind that $\mathrm{Var}(W) = \mathrm{E}(\mathrm{Var}(W|R_n)) + \mathrm{Var}(\mathrm{E}(W|R_n)) = \mathrm{E}(w'\Sigma_n w) + \mathrm{Var}(w'\mu_n)$
- If r_t are i.i.d., it is proved that best weight is a constant vector. However, r_t are not independent. Engle and Bollerslev (1986): volatility clustering and strong autocorrelations of squared returns.
- The point is that the optimal weight w, as a function of past returns, is also random

Iterative Improvements of a Portfolio

- The main idea is to
  - simulate past returns by bootstrap to estimate the underlying joint probability distribution
  - and use gradient descent to search for a better weighting iteratively

Model

- Conditional mean: $u_n(S_n) = E(r_{n+1}|S_n)$
- Conditional second moment: $V_n(S_n) = E(r_{n+1}r_{n+1}^T|S_n)$
- Then $E(w^T r_{n+1}) = E(w^T u_n)$ and $E((w^T r_{n+1})^2) = E(w^T V_n w)$
  - We assume that w is a function of the random S_n
- Let $G(w) = F\left(E(w^T u_n), E(w^T V_n w)\right)$, the objective function
  - $\nabla G(w) = \nabla F(E(w^T u_n), E(w^T V_n w))$, the total derivative
- By Taylor Expansion, we have:
  - $\mathrm{G}(w + t\delta) = G(\mathrm{w}) + \nabla G(w)^{\mathrm{T}}\left(tE(\delta^T u_n), 2tE(w^T V_n \delta) + t^2 E(\delta^T \mathrm{V}_n \delta)\right) + O(t^2)$, for small $t > 0$
- If we choose the search direction $\delta(S_n) = (u_n(S_n), 2V_n(S_n)\, w(S_n))$,
  - then $\nabla G(w)$ and δ does not equal to zero almost for sure
  - the first order term of δ becomes $\mathrm{E}(\|\delta\|^2) > 0$
  - hence $G(w + t\delta) > G(w)$ for positive small t

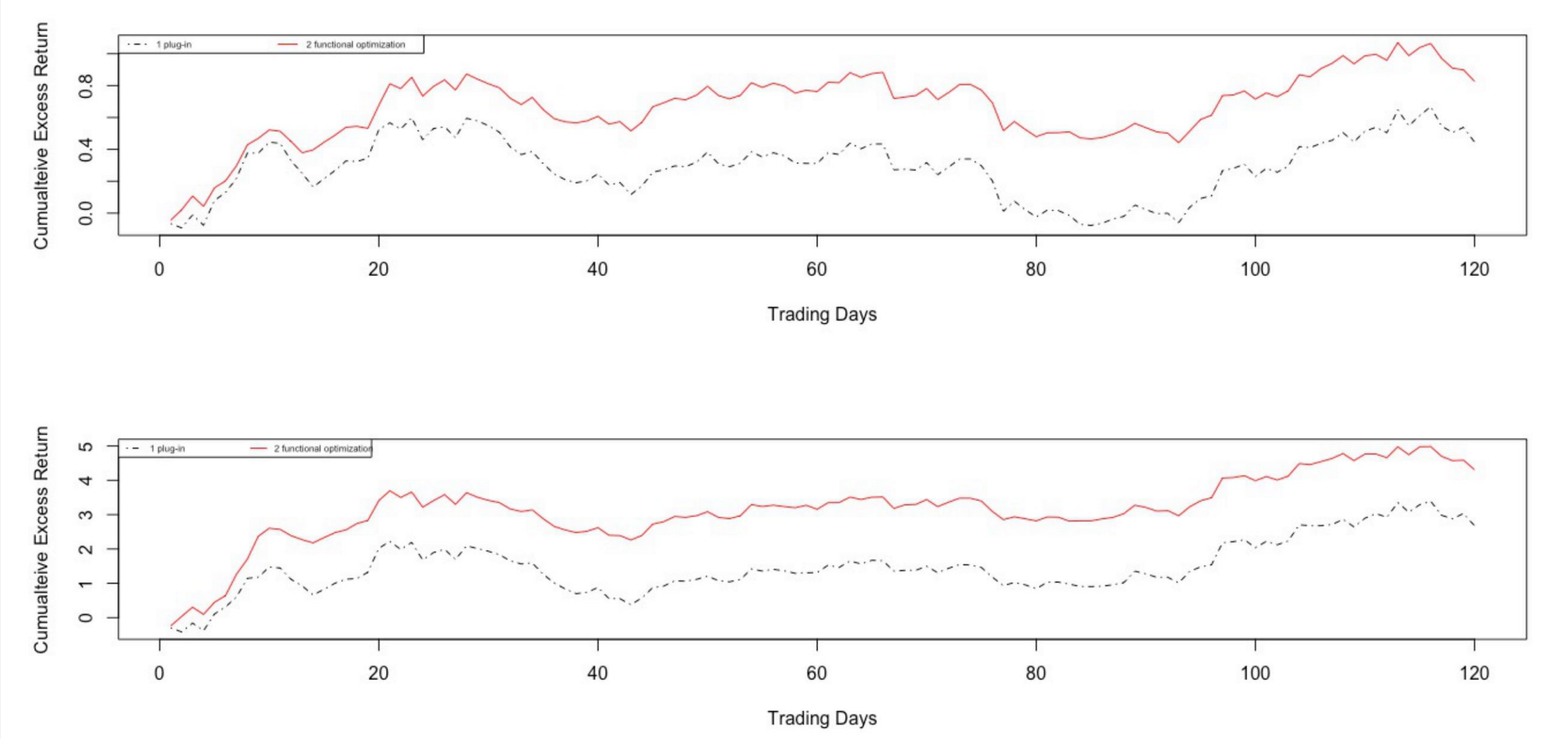
Gradient Search

- Use simulation to compute conditional expectations.
  - bootstrap samples $\{r_{b1}^*, r_{b2}^*, \ldots, r_{bn}^*\}$ $(\mathrm{b} = 1,2, \ldots \mathrm{B})$ drawn with replacement from the observed sample $\{r_1, r_2, \ldots, r_n\}$
  - u_n and V_n are estimated for each sample path
  - $U_0 = \frac{1}{B}\sum_{b=1}^{B} w_0\left(S_n^b\right)^T u_n\left(S_n^b\right) \approx E(w_0^T u_n)$
  - $V_0 = \frac{1}{B}\sum_{b=1}^{B} w_0(S_n^b)^T V_n(S_n^b) w_0(S_n^b) \approx E(w_0^T V_n w_0)$
- One step:
  - $\delta^T = P * \left(u_n(S_n), 2V_n(S_n)w_0(S_n)\right)\nabla G(w_0)$
  - $P = (I - \mathbf{1}\mathbf{1}^{\mathrm{T}})/p$, the projection matrix
    - Note that $P^T\mathbf{1} = 0$ which means $\delta^T\mathbf{1} = 0$, so that we can add it to the original weight
  - $\nabla G(w) = \nabla F\left(E(w^T u_n), E(w^T V_n w)\right) = \nabla F(U_0, V_0)$
  - Choose t such that $\mathrm{G}(w_0 + t\delta) > \mathrm{G}(w_0)$
- Multi-steps: replace w_0 by $w_0 + t\delta$; repeat the above

$S_n^{(b)}$ generation

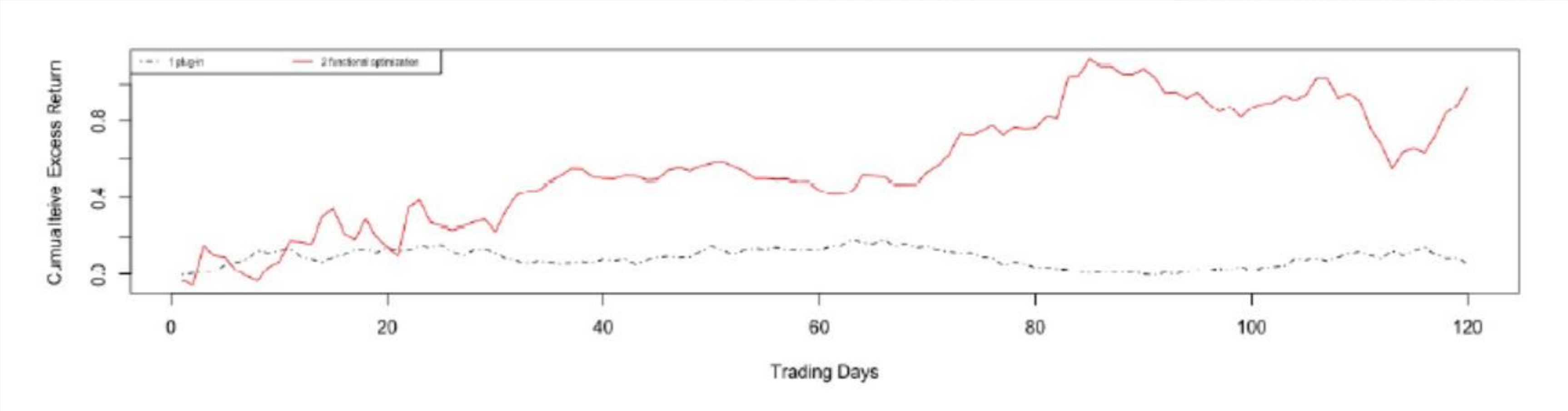
- Block bootstrap
  - Group $\{r_1, \ldots, r_n\}$ into blocks that contain consecutive r_t and sample on those blocks so that certain level of correlation among the resampled returns is remained
- Parametric bootstrap
  - For example, if we know that r_t follow AR(1) model with $\boldsymbol{\xi_t} = \{\xi_{1,t}, \ldots, \xi_{n,t}\}$ are independent, then we can apply bootstrap to sample $\{\boldsymbol{\xi_1}^{(b)}, \ldots, \boldsymbol{\xi_n}^{(b)}\}$ by drawing with replacement from $\{\boldsymbol{\xi_1}, \ldots, \boldsymbol{\xi_n}\}$ and then generate $S_n^{(b)} = \{r_t^{(b)}: r_{i,t}^{(b)} = a_i(S_n) + b_i(S_n) r_{i,t}^{(b)} \xi_{i,t}^{(b)}, \mathrm{t} = 1, \ldots, \mathrm{n}\}$

Empirical results on SP500



- The upper panel is the time series plot of realized cumulative excess returns over S&P500 with the constraints $\sum wi = 1$ and $wi \geq -0.2$.
  The lower panel is the time series plot of realized cumulative excess returns over S&P500 with the constraints $\sum wi = 1$ and $wi \geq -1$.

A Potential New Dynamic SP500 Index



- The upper panel is the time series plot of realized cumulative excess returns over S&P500 with the constraints $\sum wi = 1$ and $wi \geq 0$, hence Long Only.

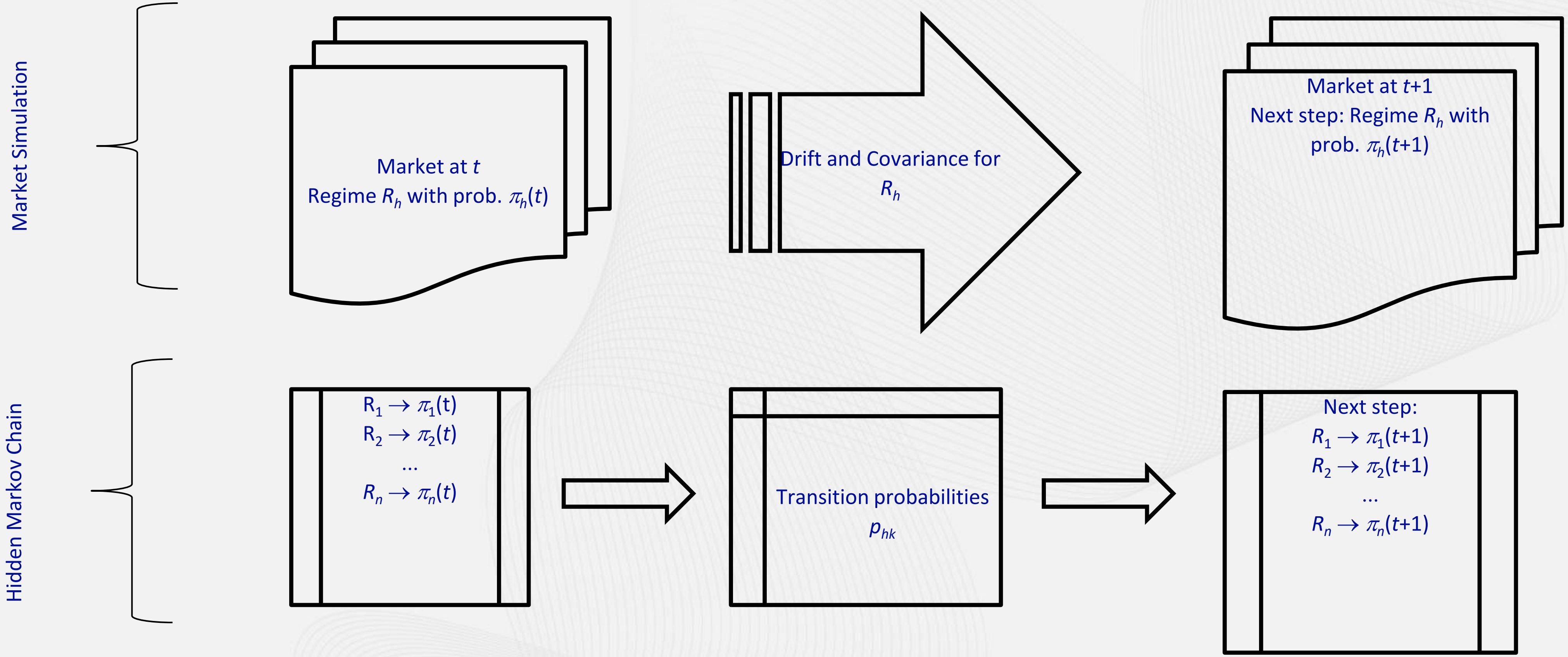
Model Validation and Risk Management

NM FinTech
WWW.NMFIN.TECH

Model Validation

- We should optimize a portfolio with respect to the underlying joint probability distribution of the assets, rather than historical returns.
- We do simulation to check the risk of a portfolio under many different scenarios including extreme stressed situations.

Markov Chain Monte-Carlo



SuanShu AlgoQuant QuantitativeModels CQI
NM FinTech WWW.NMFIN.TECH

Hidden Markov Model

Dynamics in the Market Variable Space

- Finite Number of Regimes $R_1,\ldots,R_m$
  - Covariance Matrix Γ_k
  - Drift Vector μ_k
  - $R_k \sim N(\mu_k,\Gamma_k)$
- Transition Probability Matrix
  - $P(t,t+1) = (p_{hk})$
  - p_{hk} = Probability of transitioning from R_h to R_k
  - $\forall h, \quad \sum_{k=1}^{m} p_{hk} = 1$
  - $P(t,t+n) = P^n$

Simulation: Markov Chain Monte-Carlo

- Euler Scheme:
  - Discretized Time $t_0,\ldots,t_n$
  - Pick new regime $R(t_{i+1})$ according to P applied to current regime $R(t_i)$
  - Simulate Market Evolution following $R(t_{i+1})$
- Gaussian Mixture $\Rightarrow$ "Fat Tails"
- Asymptotically Gaussian
  - $\mu_\infty = \sum_{k=1}^{m} \pi_k \mu_k$
  - $\Gamma_\infty = \sum_{k=1}^{m} \pi_k \Gamma_k$
  - $P'\pi = \pi \quad \pi = (\pi_1,\ldots,\pi_m)$

MCMC Calibration: Regime Identification

1. Determine Breakpoints $\{t_1,\dots,t_n\}$ and "homogenous" periods $J_k = [t_{k-1}, t_k]$

2. Calibrate a Gaussian distribution $N(\mu_k,\Gamma_k)$ over consecutive periods J_k

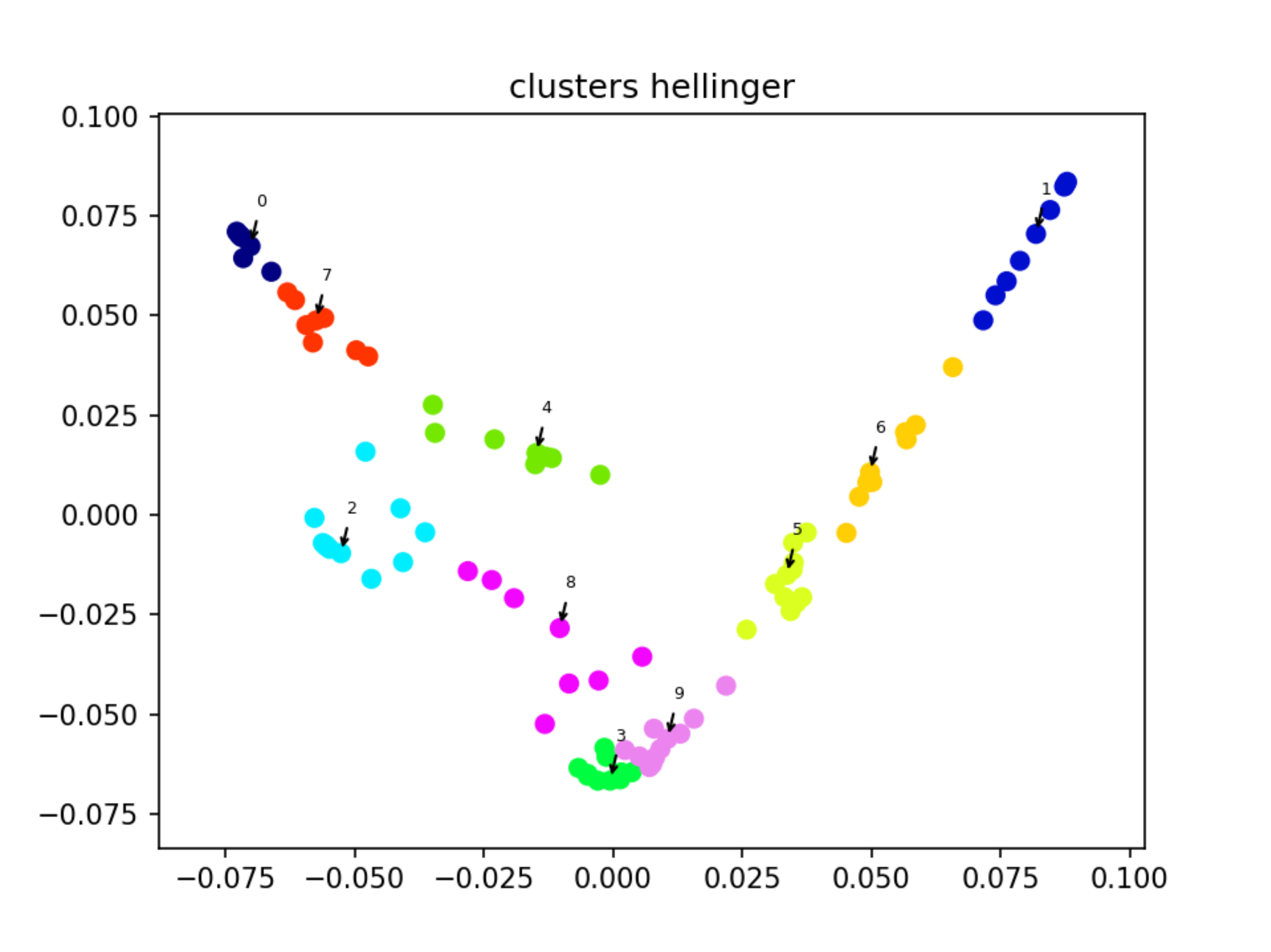
3. Clusterize the set of $\Phi_k = (\mu_k, \Gamma_k)$, using some information distance (K-L, Tsallis, Hellinger…)

4. Issue: these distributions lie in a high dimensional space $n(n + 3)/2 \Rightarrow$ No recurrence
   i. Reduce dimension by describing the market with fewer indices
   ii. Project Φ_k onto a lower dimensional space, approximately preserving distances, using *Spectral Embedding*
      - Common practice: focus only on volatility…
   iii. The projection now has some *recurrence*.

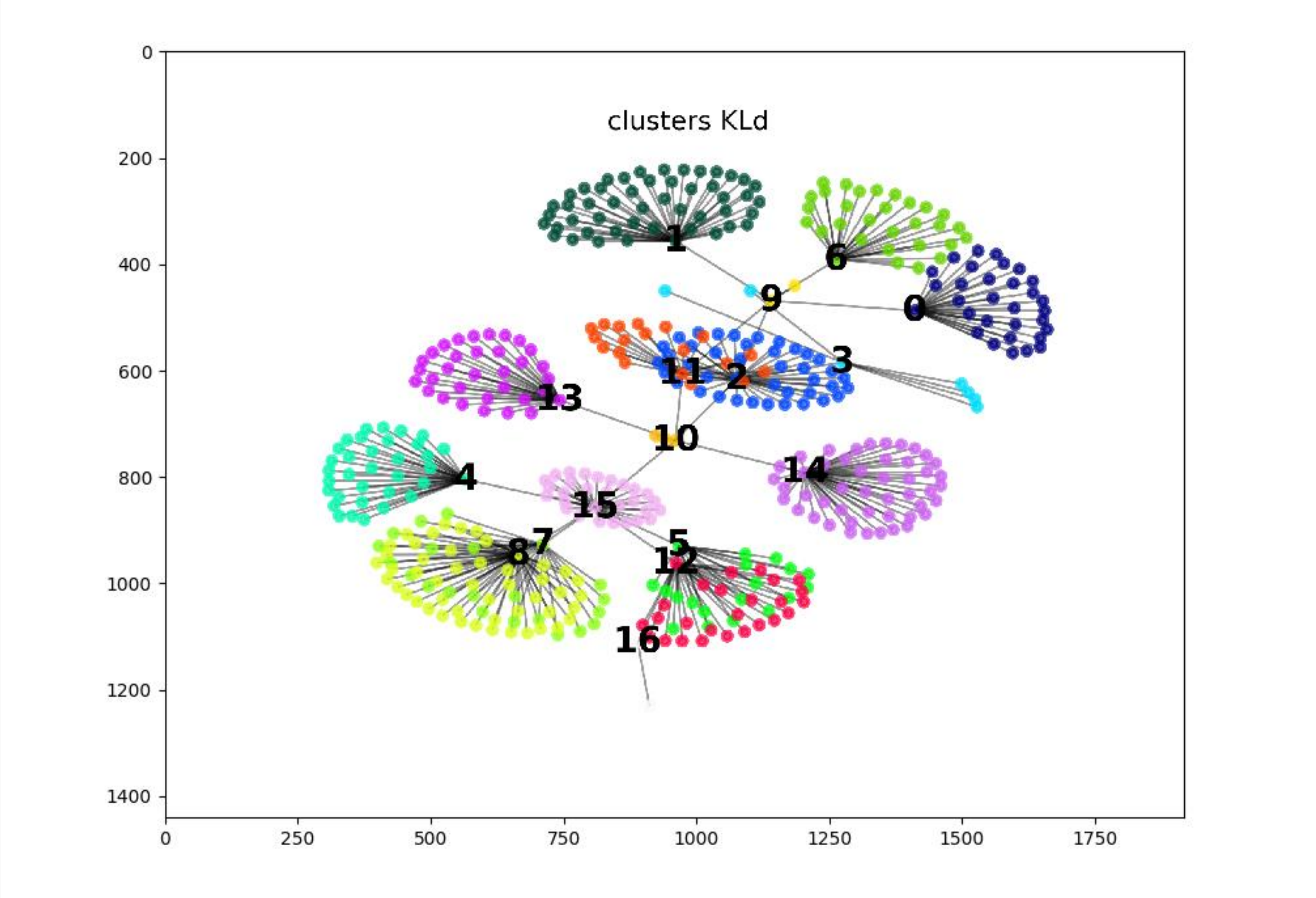
5. Estimate Transition Probabilities
   i. Baum-Welch algorithm: too imprecise
   ii. SVM or EM provide more accurate results
   iii. Depends on time spent within a regime

6. Crisis Prediction: Mild regime that is likely to transit to a wild one

Spectral Embedding and Clustering



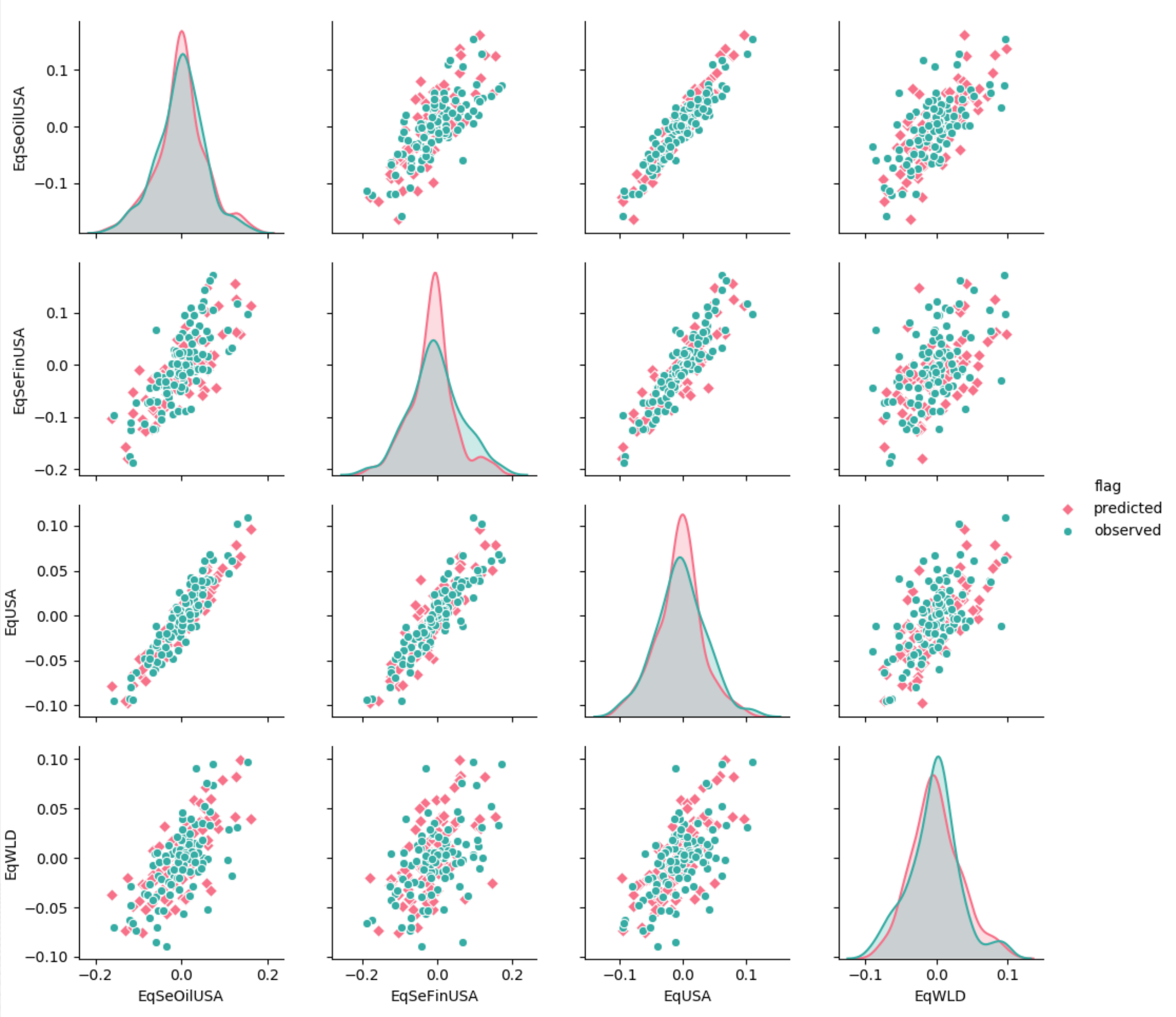
Clustering by Tree Algorithm



Regime Sequence

Multi-Regime Simulation vs. Observed

- Joint distribution of 4 risk factors:
  - SP500
  - SP Sector Financials
  - SP Sector Oil Companies
  - MSCI World
- We can see that simulated returns match very well the observed returns in terms of extreme situations



Pink: Multi-Gaussian Simulation (MCMC)

Green: Actual Returns

Portfolio Multi-Regime Optimization

- Identification of Current Regime
- MCMC simulated over a given period of time (e.g. one month)
  - Expected Return of each asset
  - Monte-Carlo simulation $\Rightarrow$ VaR of any virtual portfolio (Gaussian mixture)
- Maximization of Expected Return / VaR
- Rebalancing based on dynamic criterion (profit taking) and/or signal strength
- Benchmark: Same with 1 regime = Markowitz, 1/N, etc.

Applications

NM FinTech
WWW.NMFIN.TECH

Applications

1 Indices

2 Partial Indices

3 More effective portfolio hedging

Alternatives to SPX, CSI, HSI, FTSE, Euronext, DAX, NIKKEI, Nifty with higher Sharpe ratio and lower volatility

Adding a cardinality constraint will enable good enough tracking of an index using only a subset of the stocks

NM Index

Applying SAAM to the same set of constituent stocks of an index, we are essentially creating a new index

NM index, compared to existing index, is significantly more stable, lower volatility/risk and higher Sharpe ratio

Provide (risk adverse) investors with an alternative market investment opportunity

- Investors buy the same set of stocks but with different weights or capital allocation
- Because of significantly lower drawdown, can safely do leverage to gear up performance, if wanted

NM

Commercialization

01

Underwrite NM index funds/ETFs

- They are more attractive to risk adverse investors, such as banks, insurance companies, pension funds, who have little risk appetite
- Funds/EFTs tailored to the risk appetite of different investors
- More effective tools for portfolio hedging

02

Using NM index as the underlying asset, underwrite call options to investors who want principal protection

03

Using NM index as the underlying asset, underwrite futures, calls, puts

04

Using NM index as the underlying asset, underwrite derivatives and structured products

NMFinTech
WWW.NMFIN.TECH

For more information about our technologies or collaboration, please contact

Prof. Haksun Li

haksun.li@nmfin.tech