Applications of Statistical Learning in Asset Allocation (2)

Hongshu Chen

4/28/2021

Mean-CVaR Optimization

Risk Measures Based On Loss (Return) Distribution

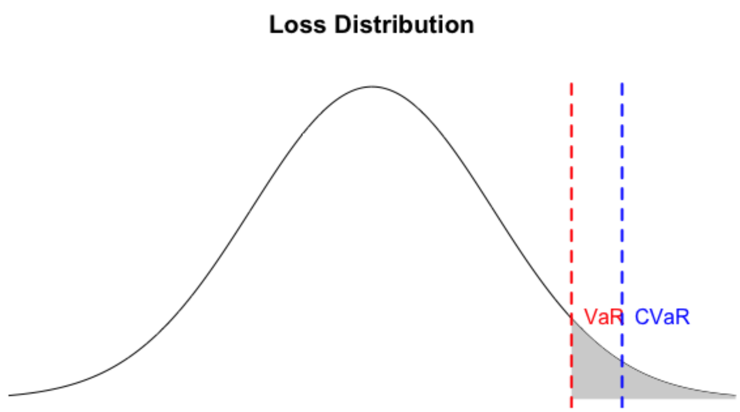
- Second moment:
  - Variance: $E((r-\mu)^2)$
  - Lower Semivariance: $E((r-\mu)^2 I(r<\mu))$, where $I(r<\mu)$ is an indicator function
- Tail:
  - Value at Risk (VaR)

$$\mathrm{VaR}(\alpha) = \inf\{x \in \mathbb{R} : P(l > x) \le 1-\alpha\}$$

  - **Conditional Value at Risk (CVaR)**

$$\mathrm{CVaR}(\alpha) = \frac{1}{1-\alpha}\int_{\mathrm{VaR}(\alpha)}^{+\infty} l p(l) dl$$

Why Not Mean-VaR?



- VaR is not a *coherent* risk measure, it does not satisfy *subadditivity* (i.e. diversification reduces risk) (McNeil, Frey, and Embrechts 2015)
- VaR gives a false sense of confidence
- Generally speaking, Mean-VaR is not convex

A Special Case

- If loss follows a normal distribution ($l \sim N(\mu, \sigma^2)$)
  - $\mathrm{VaR}(\alpha) = \mu + \sigma\Phi^{-1}(\alpha)$, where $\Phi^{-1}(\alpha)$ is the α-quantile of the standard normal distribution
  - $\mathrm{CVaR}(\alpha) = \mu + \sigma\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$, where ϕ is the density of the standard normal distribution
- When assets return follow a multivariate normal, portfolio return (loss) follows a normal distribution
- In that case, Mean-Variance, Mean-VaR and Mean-CVaR are equivalent, i.e., they all try to minimize portfolio volatility while satisfying the minimum return constraint

Mean-CVaR Optimization

Mean-CVaR is that it can be formulated as a linear programming problem (Rockafellar and Uryasev 2000) when the loss (return) distribution is discrete:

$$
\begin{gathered}
arg \min_{\mathbf{w},a,\mathbf{u}} a + \frac{\mathbf{1}^T\mathbf{u}}{k(1-\alpha)}, \\
s.t. \\
\mathbf{u} \geq \mathbf{0} \\
\mathbf{X}\mathbf{w} + a + \mathbf{u} \geq \mathbf{0} \\
\mu^T\mathbf{w} \geq r
\end{gathered}
$$

where $\mathbf{X}$ are scenario returns, k is the number of scenarios and α is the level of loss tail (e.g. 95%), μ is assets expected return and r is the portfolio required expected return. Value of the objective function is the $\mathrm{CVaR}(\alpha)$ of the optimal portfolio, and a is the corresponding portfolio $\mathrm{VaR}(\alpha)$.

A Mean-CVaR Example

- 7 assets: USLE, DME, EME, USIG, COM, REITS, CASH
- 20 years monthly returns
- Use *ImputeR* and Lasso to impute missing data
- Use historical average (with imputed data) as expeced return
- The required minimum monthly portfolio return is 30 bps
- Three optimal portfolios:
  1. Mean-Variance Optimization: Historical Variance-Covariance Matrix
  2. Mean-CVaR (95%) Optimization: Simulated Returns From Multivariate Normal Distribution (Historical Mean and Variance-Covariance Matrix)
  3. Mean-CVaR (95%) Optimization: Using Historical Returns

A Mean-CVaR Example (Cont.)

As expected, MVO and Mean-CVaR with simulated data from normal distribution generate the same optimal portfolio weights:

Portfolio Weights Comparison

	MVO	Mean-CVaR: Normal	Mean-CVar: Historical
USLE	0.07	0.07	0.00
DME	0.00	0.00	0.00
EME	0.00	0.00	0.00
USIG	0.68	0.68	0.79
COM	0.00	0.00	0.00
REITS	0.00	0.00	0.00
CASH	0.26	0.26	0.21

A Mean-CVaR Example (Cont.)

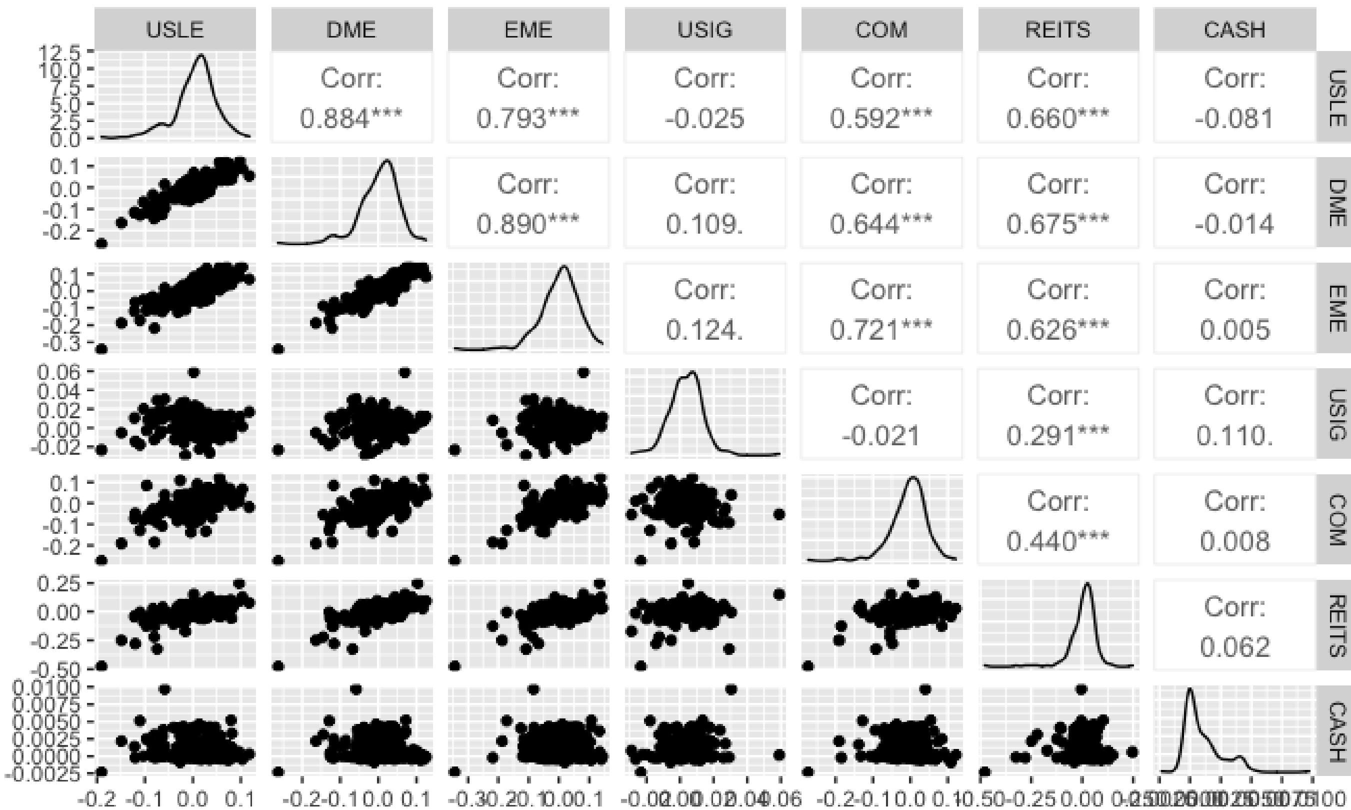
And the estimated CVaR of optimal portfolio is slightly larger with Mean-CVaR based on historical returns:

Optimal Portfolio Monthly VaR/CVaR Esitmates

	Mean-CVaR: Normal	Mean-CVar: Historical
VaR (95%)	0.0093	0.0092
CVaR (95%)	0.0124	0.0146

A Mean-CVaR Example (Cont.)

Results in previous slides make sense because historical asset returns do not follow a normal distribution and most of them are negatively skewed with a heavy left tail.



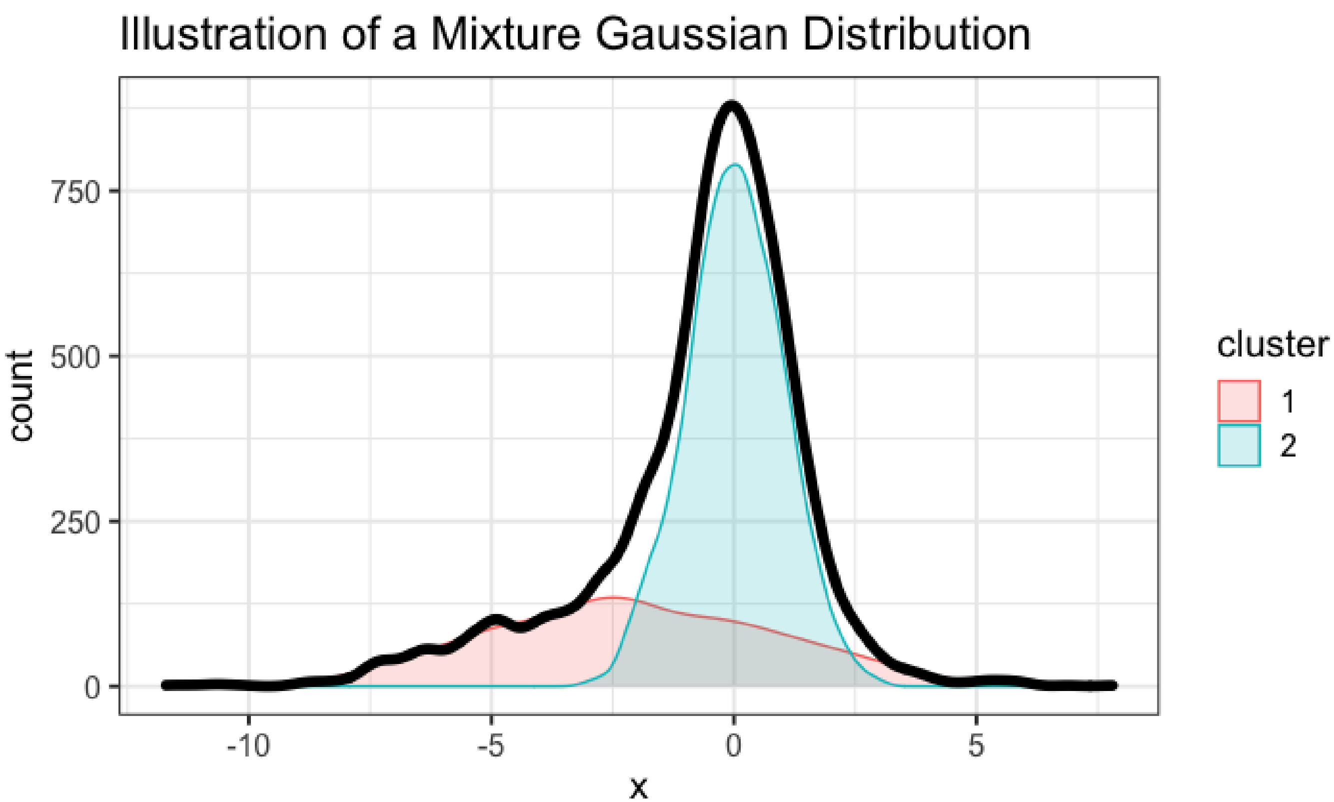
Return Simulation with Gaussian Mixture Model (GMM)

We can simulate return scenarios from Mixture Gaussian distribution instead. Let’s use a univariate mixture Gaussian distribution as an example, its probability density function is as below:

$$p(x) = \sum_{i=1}^{k} \pi_i N(\mu_i, \sigma_i^2),$$

where $\sum_{i=1}^{k} \pi_i = 1$.

Illustration of GMM



GMM Parameter Estimation

Parameters of a GMM θ can be estimated as θ^* by the EM algorithm (Benaglia et al. 2009):

- E-step: $Q(\theta|\theta^*) = E(logP_\theta(Z|X, \theta^*))$
- M-step: $\theta^* = arg\max_\theta Q(\theta|\theta^*)$
- X is the observed data and Z is the latent cluster membership variable.

GMM With Two Clusters

GMM Cluster Marginal Probabilities

Cluster 1	Cluster 2
0.92	0.08

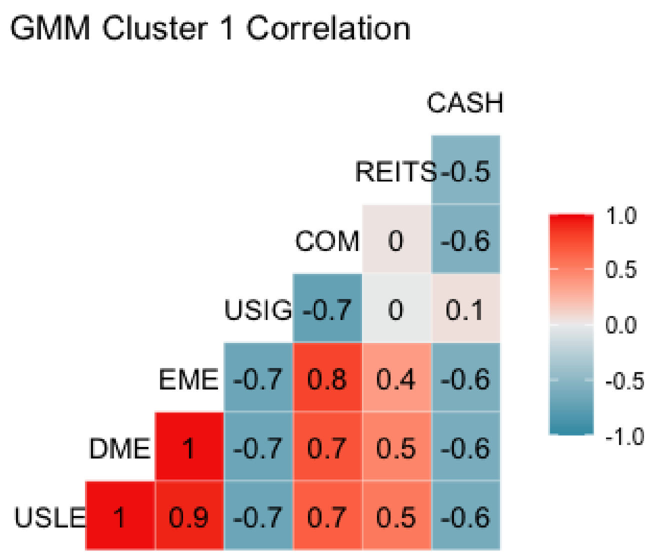
GMM Cluster Annualized Average Return

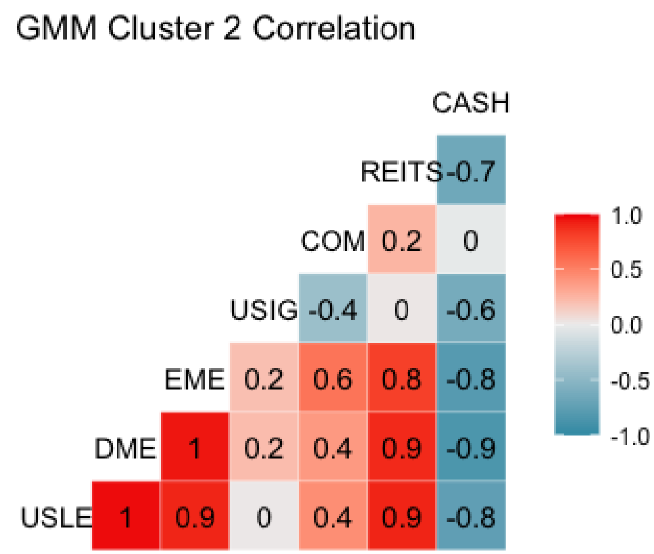
	USLE	DME	EME	USIG	COM	REITS	CASH
Cluster 1	0.1211	0.0962	0.1443	0.0402	0.0218	0.1615	0.0121
Cluster 2	-0.6431	-0.7014	-1.0142	0.0666	-0.9611	-1.1031	0.0214

GMM Cluster Annualized Volatility

	USLE	DME	EME	USIG	COM	REITS	CASH
Cluster 1	0.13	0.15	0.18	0.03	0.14	0.15	0.00
Cluster 2	0.23	0.30	0.37	0.07	0.27	0.60	0.01

GMM With Two Clusters (Cont.)





GMM With Two Clusters (Cont.)

We see that equity returns are higher and volatility are lower in Cluster 1, we can consider it as the “risk on” or “low volatility” time periods, while Cluster 2 is “risk off” or “high volatility” time periods.

A Mean-CVaR Example (Cont.)

Now we use the fitted GMM to generate scenario returns. The resulted Mean-CVaR optimal portfolio weights and CVaR estimates are closer to the ones based on historical returns.

Portfolio Weights Comparison

	MVO	Mean-CVaR: Normal	Mean-CVar: Historical	Mean-CVaR: GMM
USLE	0.07	0.07	0.00	0.03
DME	0.00	0.00	0.00	0.00
EME	0.00	0.00	0.00	0.00
USIG	0.68	0.68	0.79	0.73
COM	0.00	0.00	0.00	0.00
REITS	0.00	0.00	0.00	0.00
CASH	0.26	0.26	0.21	0.23

Optimal Portfolio Monthly VaR/CVaR Esitmates

	Mean-CVaR: Normal	Mean-CVar: Historical	Mean-CVaR: GMM
VaR (95%)	0.0093	0.0092	0.0084
CVaR (95%)	0.0124	0.0146	0.0136

Return Simulation with Hidden Markov Model (HMM)

In the example above with GMM, we assume all the observations are independent from each other, but we know it is usually not the case with financial time series. It is reasonable to assume that if at time t, we are in Cluster 1 (or the “risk on” regime), there is a higher chance that at time $t+1$, we are still in the same regime. And we can assume that the latent cluster (regime) membership z follows a Markov process,

$$P(z_{t+1}|z_1, z_2, \dots, z_t) = P(z_{t+1}|z_t)$$

i.e. the probability distribution of z_{t+1} depends on and only depends on the value of z_t.

By adding this assumption to the GMM above, we get a Hidden Markov Model (HMM) (Rabiner and Juang 1986) with Gaussian emission distribution. We can still use EM algorithm to estimate HMM’s model parameters.

HMM With Two Regimes

HMM Regime Marginal Probabilities

Regime 1	Regime 2
0.09	0.91

HMM Transition Probabilities

	Regime 1	Regime 2
Regime 1	0.64	0.04
Regime 2	0.36	0.96

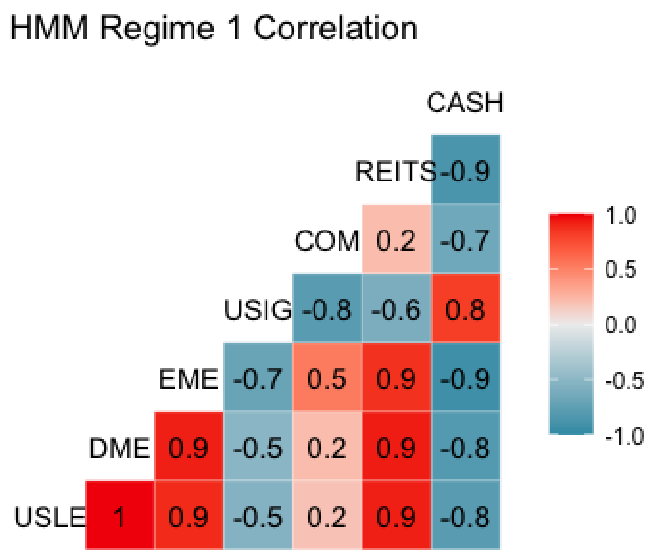
HMM Regime Annualized Average Return

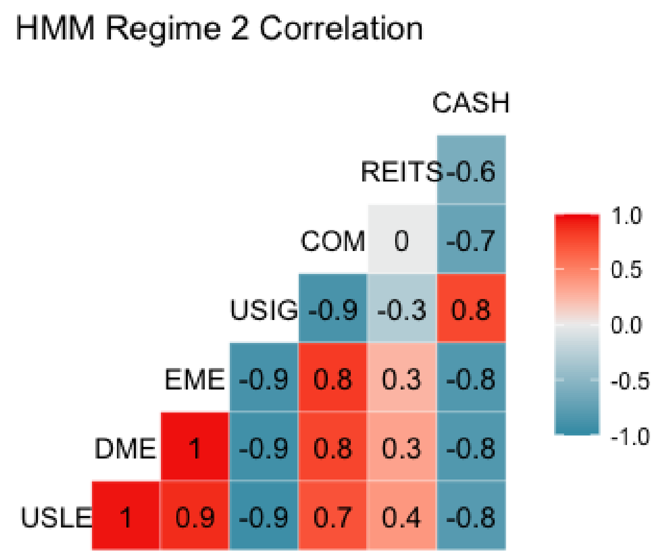
	USLE	DME	EME	USIG	COM	REITS	CASH
Regime 1	-0.5530	-0.7051	-0.8784	0.0515	-0.7005	-0.9441	0.0157
Regime 2	0.1263	0.1116	0.1522	0.0412	0.0137	0.1690	0.0125

HMM Regime Annualized Volatility

	USLE	DME	EME	USIG	COM	REITS	CASH
Regime 1	0.24	0.29	0.36	0.10	0.32	0.56	0.07
Regime 2	0.13	0.15	0.18	0.04	0.13	0.15	0.02

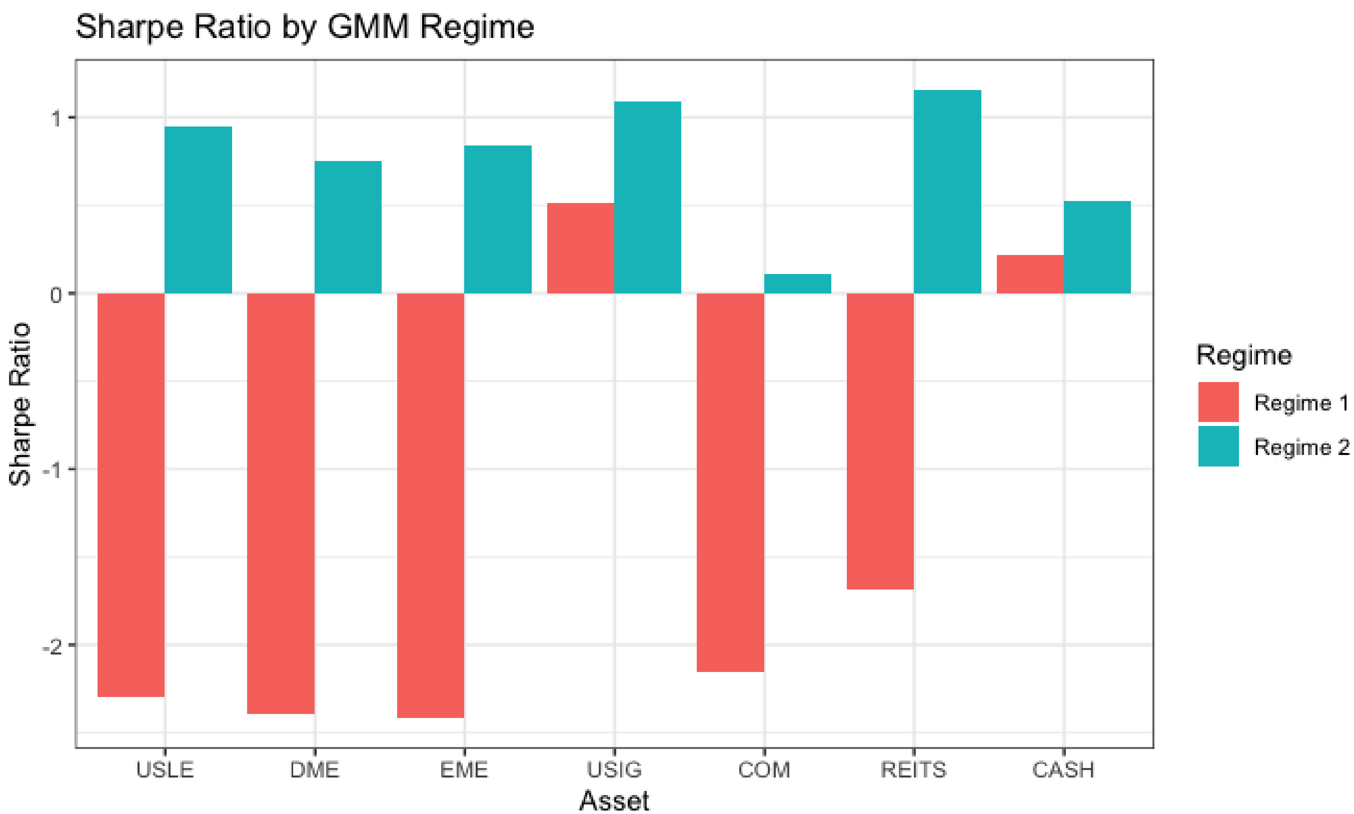
HMM With Two Regimes (Cont.)





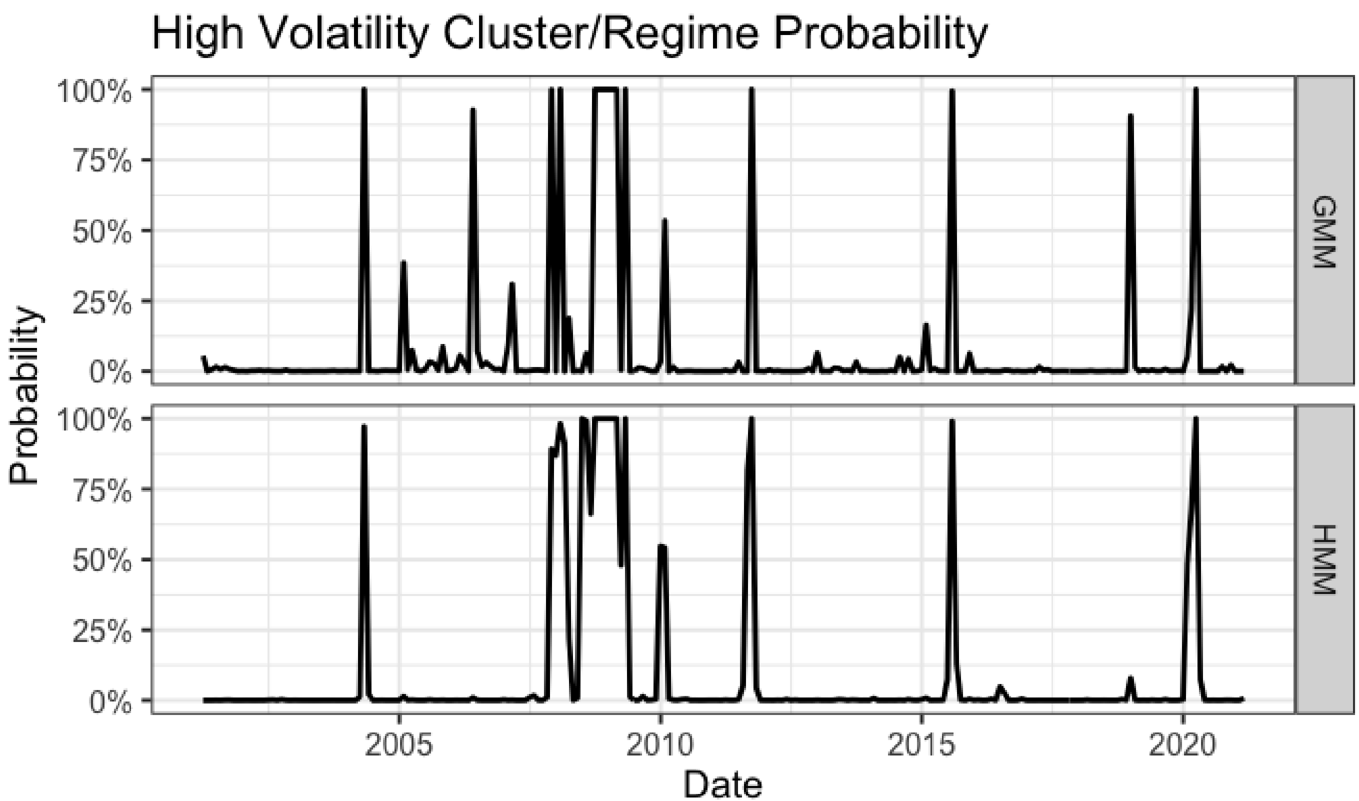
HMM With Two Regimes (Cont.)

Regime 1 is the “risk off” regime and Regime 2 is the “risk on” regime.



GMM vs. HMM

The classifications from HMM is less volatile than GMM.



A Mean-CVaR Example (Cont.)

Portfolio Weights Comparison

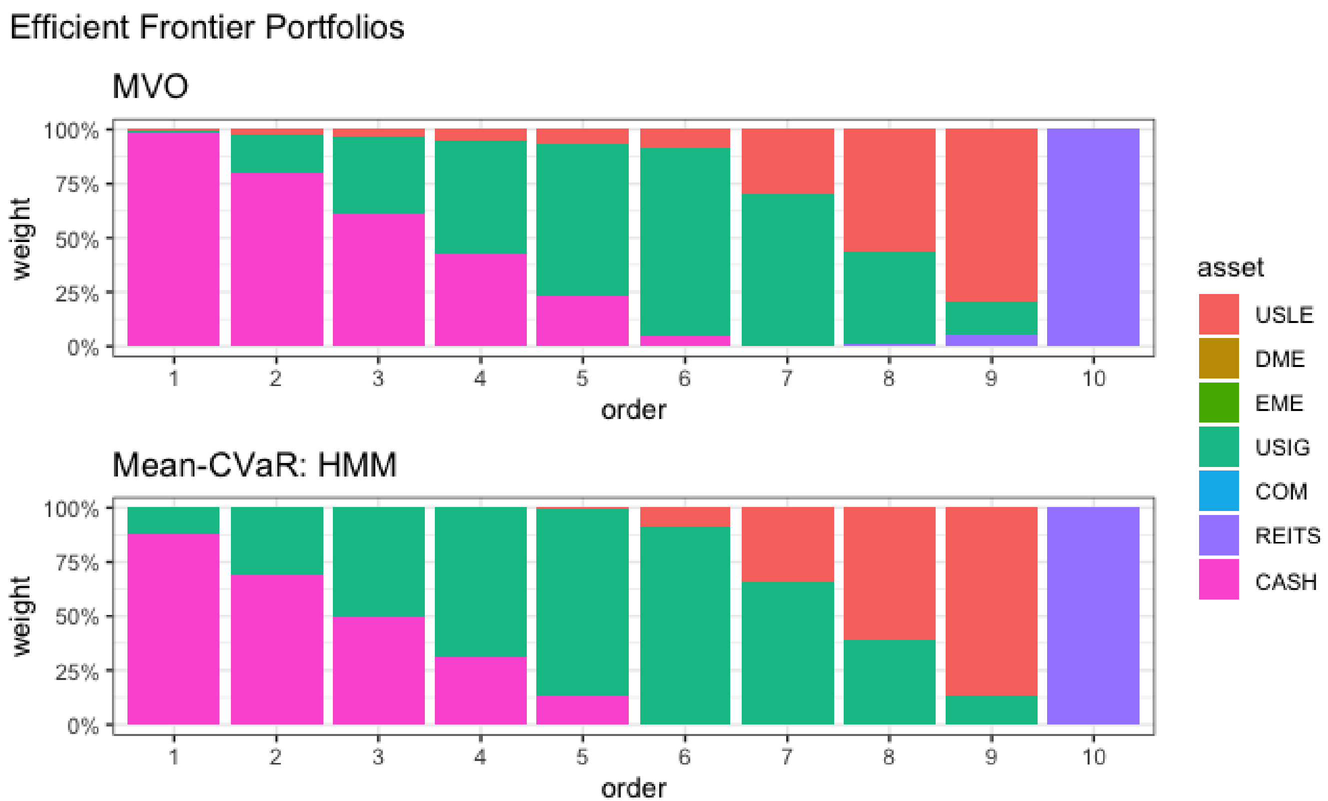
	MVO	Mean-CVaR: Normal	Mean-CVar: Historical	Mean-CVaR: GMM	Mean-CVaR: HMM
USLE	0.07	0.07	0.00	0.03	0.00
DME	0.00	0.00	0.00	0.00	0.00
EME	0.00	0.00	0.00	0.00	0.00
USIG	0.68	0.68	0.79	0.73	0.79
COM	0.00	0.00	0.00	0.00	0.00
REITS	0.00	0.00	0.00	0.00	0.00
CASH	0.26	0.26	0.21	0.23	0.21

Optimal Portfolio Monthly VaR/CVaR Esitmates

	Mean-CVaR: Normal	Mean-CVar: Historical	Mean-CVaR: GMM	Mean-CVaR: HMM
VaR (95%)	0.0093	0.0092	0.0084	0.0152
CVaR (95%)	0.0124	0.0146	0.0136	0.0243

The optimal portfolio weights are the same for “Mean-CVaR: Historical” and “Mean-CVaR: HMM”, but the VaR and CVaR estimates from the latter are much larger.

A Mean-CVaR Example (Cont.)



Risk Parity

Contribution to Risk (Volatility)

Volatility is a positive-homogeneous function of portfolio weight, according to Euler's rule (McNeil, Frey, and Embrechts 2015), contribution to risk (volatility) is calculated as:

$$\frac{\partial\sqrt{\mathbf{w}^T\Sigma\mathbf{w}}}{\partial\mathbf{w}}\circ\mathbf{w} = \frac{\Sigma\mathbf{w}}{\sigma}\circ\mathbf{w} \propto \Sigma\mathbf{w}\circ\mathbf{w},$$

where $\sigma = \sqrt{\mathbf{w}^T\Sigma\mathbf{w}}$.

Is 60/40 Well Disersified?

```
return data history: 2003-11-30 ~ 2021-02-28
```

```
volatility:
```

```
  USLE   USIG
0.1475 0.0371
```

```
correlation:
```

```
     USLE USIG
USLE 1.00 0.05
USIG 0.05 1.00
```

Risk Contribution to 60/40 Portfolio

	Risk Contribution	Relative Risk Contribution (%)
USLE	0.0873	96.55
USIG	0.0031	3.45

Risk Parity Basics

- Coined by Edward Qian (Qian 2016)
- Seeks to achieve equal risk contribution from different assets/assets groups/factors
- Does not rely on expected return forecasts
- Often uses leverage
- Popular among institutional investors

Problem Statement

$$arg \min_{\mathbf{w}} (\frac{\Sigma\mathbf{w} \circ \mathbf{w}}{\mathbf{w}^T\Sigma\mathbf{w}} - \mathbf{b})^T (\frac{\Sigma\mathbf{w} \circ \mathbf{w}}{\mathbf{w}^T\Sigma\mathbf{w}} - \mathbf{b}),$$
$$s.t.$$
$$\sqrt{\mathbf{w}^T\Sigma\mathbf{w}} = \sigma$$

where $\mathbf{b}$ is the assets risk budget vector (e.g. $[1/n, 1/n, \ldots, 1/n]^T$) and $\mathbf{1}^T\mathbf{b} = 1$, σ is the volatility target.

We can use a nonlinear solver solve this problem, but when the portfolio is long only (i.e. $\mathbf{w} \geq \mathbf{0}$), we can formulate it as a convex optimization problem (Roncalli 2013).

Risk Parity Optimization

$$
\begin{gathered}
arg \min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^T \Sigma \mathbf{w}, \\
s.t. \\
\mathbf{b}^T \log \mathbf{w} \geq c
\end{gathered}
$$

where c is an arbitrary positive number. The optimal $\mathbf{w}$ satisfies the Karush-Kuhn-Tucker conditions (Boyd and Vandenberghe 2004), that is there exists a multiplier $\lambda \geq 0$ such that,

$$
L = \frac{1}{2}\mathbf{w}^T \Sigma \mathbf{w} + \lambda(c - \mathbf{b}^T \log \mathbf{w})
$$

$$
\frac{\partial L}{\partial \mathbf{w}} = \Sigma \mathbf{w} - \lambda \frac{\mathbf{b}}{\mathbf{w}} = 0
$$

Risk Parity Optimization (Cont.)

We can rearrange the equation to show that $\mathbf{w}$ satisfies the risk budgeting constraint:

$$\Sigma\mathbf{w} = \lambda\frac{\mathbf{b}}{\mathbf{w}}$$

$$\Sigma\mathbf{w} \circ \mathbf{w} = \lambda\mathbf{b}$$

$$\frac{\Sigma\mathbf{w} \circ \mathbf{w}}{\mathbf{w}^T\Sigma\mathbf{w}} = \frac{\lambda\mathbf{b}}{\lambda\mathbf{1}^T\mathbf{b}} = \mathbf{b}$$

It is trivial to show that we can rescale $\mathbf{w}$ such that $\sqrt{\mathbf{w}^T\Sigma\mathbf{w}} = \sigma$ while still satisfying the risk budgeting constraint.

Connection to MVO

Solution to MVO:

$$\mathbf{w} = \frac{1}{\lambda}\Sigma^{-1}\mu$$

substitute it into the risk contribution equation:

$$\frac{\frac{1}{\lambda^2}\mu \circ (\Sigma^{-1}\mu)}{\sigma} \propto \mu \circ (\Sigma^{-1}\mu) = \Gamma^{-1} \begin{bmatrix} \mu_1^2/\sigma_1^2 \\ \mu_2^2/\sigma_2^2 \\ \vdots \\ \mu_n^2/\sigma_n^2 \end{bmatrix}$$

where Γ is the correlation matrix.

This suggests that MVO and Risk Parity are equivalent when assets have the **same Sharpe Ratio** and off diagonal entries of Γ are identical, i.e., **identical correlation**, (or whenever the vector of ones is one of Γ's eigenvectors). The implied assumption of equal Sharpe Ratio is an important insight, even though it does not hold empirically.

A Two Assets Risk Parity Example

Risk Parity Portfolio Weights

	Unlevered	Levered (Matching 60/40 Volatility)
USLE	0.2	0.42
USIG	0.8	1.68

When there are **only two assets** or **assets correlations are all zero**, risk parity portfolio weights are **inversely proportional to assets volatility**.

Risk Parity Example 2

6 asset classes (USLE/USSE/DME/EME/USIG/COM):

Risk Parity (Equal Contribution)

	Weight	Relative Risk Contribution
USLE	0.0776	0.1667
USSE	0.0614	0.1667
DME	0.0629	0.1667
EME	0.0508	0.1667
USIG	0.6655	0.1667
COM	0.0818	0.1667

Even though each of the assets contributes 1/6 of the portfolio risk, but if we add the risk contribution from the 4 equity asset classes together, they contribute to 2/3 of the portfolio, while fixed income and commodities only contributes 1/6 each.

Risk Parity Example 2 (Cont.)

A K-means clustering analysis confirms the 4 equity asset classes are in one cluster. We can instead do a risk parity based on the clustering results:

```
Kmeans Clustering:


USLE USSE  DME  EME USIG  COM
   3    3    3    3    2    1
```

Group Risk Parity

	Weight	Relative Risk Contribution
USLE	0.0361	0.0833
USSE	0.0292	0.0833
DME	0.0279	0.0833
EME	0.0218	0.0833
USIG	0.7607	0.3333
COM	0.1243	0.3333

Risk Parity Example 3

12 assets: USLE/USSE/DME/EME/USIG/CORP/USHY/DMB/EMB/TIPS/REITS/COM

Risk Parity

	Weight	Relative Risk Contribution
USLE	0.0454	0.0833
USSE	0.0357	0.0833
DME	0.0361	0.0833
EME	0.0291	0.0833
USIG	0.2482	0.0833
CORP	0.0953	0.0833
USHY	0.0462	0.0833
DMB	0.2129	0.0833
EMB	0.0510	0.0833
TIPS	0.1230	0.0833
REITS	0.0283	0.0833
COM	0.0489	0.0833

Risk Parity Example 3 (Cont.)

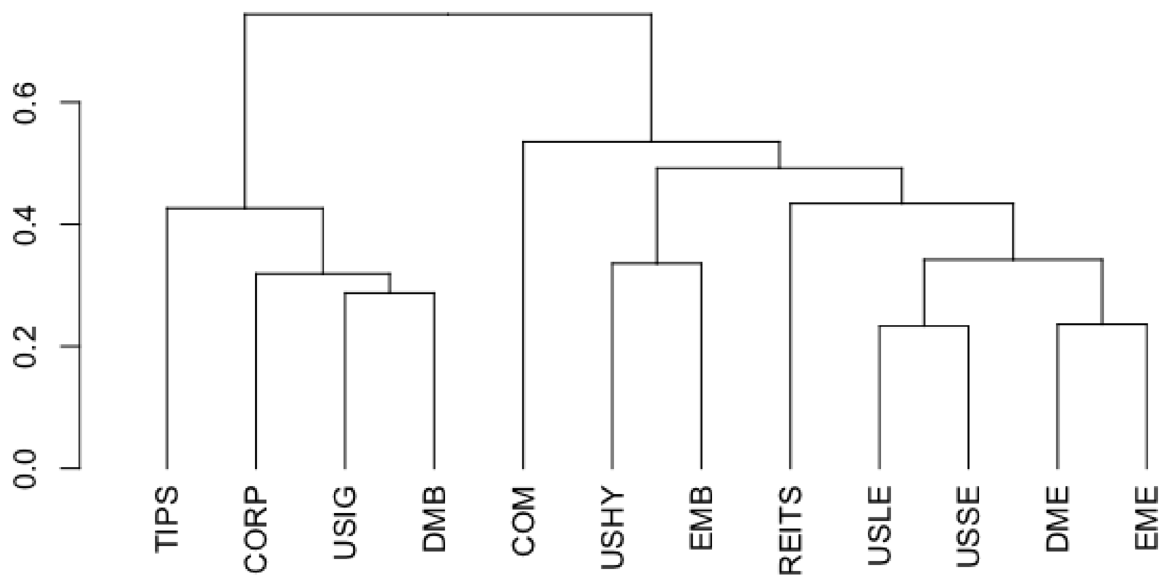
```
K-means clustering:

 USLE  USSE   DME   EME  USIG  CORP  USHY   DMB   EMB  TIPS REITS   COM
    2     2     2     2     1     1     4     1     4     1     5     3
```

Group Risk Parity

	Weight	Relative Risk Contribution
USLE	0.0317	0.05
USSE	0.0242	0.05
DME	0.0253	0.05
EME	0.0201	0.05
USIG	0.2125	0.05
CORP	0.0764	0.05
USHY	0.0651	0.10
DMB	0.1752	0.05
EMB	0.0742	0.10
TIPS	0.0984	0.05
REITS	0.0760	0.20
COM	0.1208	0.20

Risk Parity Example 3 (Cont.)

We can also make use results of hierarchical clustering:



Hierarchical risk parity (Lohre, Rother, and Schafer, n.d.) iteratively builds risk parity portfolios from leafs to the root of the tree.

Risk Parity Example 3 (Cont.)

Hierarchical Risk Parity

	Weight	Relative Risk Contribution
USLE	0.0101	0.0147
USSE	0.0075	0.0137
DME	0.0087	0.0170
EME	0.0070	0.0179
USIG	0.1428	0.0560
CORP	0.1417	0.1304
USHY	0.0439	0.0736
DMB	0.1447	0.0614
EMB	0.0512	0.0823
TIPS	0.2991	0.2521
REITS	0.0245	0.0622
COM	0.1187	0.2186

Number of Independent Bets

- Everything else being equal, we would like the portfolio to be as diversified as possible
- Meucci (2009) quantifies diversification by measuring the number of independent bets
  - PCA on assets variance-covariance matrix (Σ)
  - Calculate portfolio relative risk contribution (p_i) from each PC
  - Calculate entropy: $Entropy = -\sum_i p_i \log p_i$
  - $N = e^{Entropy}$
- Similar to the inverse of Herfindahl-Hirschman Index ($1/\sum_i p_i^2$)

Final Remarks

Other Frameworks

- Minimum variance:

$$\arg\min_{\mathbf{w}} \mathbf{w}^T \Sigma \mathbf{w}$$

- Maximum diversification ratio:

$$\arg\max_{\mathbf{w}} \frac{\sum_i w_i \sigma_i}{\sqrt{\mathbf{w}^T \Sigma \mathbf{w}}}$$

- Target volatility:

$$\arg\max_{\mathbf{w}} \mu^T \mathbf{w}$$

$$s.t. \sqrt{\mathbf{w}^T \Sigma \mathbf{w}} = c$$

Takeaways

- Asset allocation is
  - a growing field in the asset management industry
  - important, interesting and challenging
  - quantitative in nature
- Build a solid foundation
  - Finance
  - Optimization
  - Statistics
  - Coding
- Be a lifetime learner

Thank You!

- Thank you for your time and patience
- Your feedback is valuable and please take the survey

References

Benaglia, Tatiana, Didier Chauveau, David R. Hunter, and Derek S. Young. 2009. "Mixtools: An R Package for Analyzing Mixture Models." *Journal of Statistical Software, Articles* 32 (6): 1–29. https://doi.org/10.18637/jss.v032.i06.

Boyd, Stephen, and Lieven Vandenberghe. 2004. *Convex Optimization*. USA: Cambridge University Press.

Lohre, Harald, Carsten Rother, and Kilian Axel Schafer. n.d. "Hierarchical Risk Parity: Accounting for Tail Dependencies in Multi-Asset Multi-Factor Allocation." https://ssrn.com/abstract=3513399.

McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2015. *Quantitative Risk Management: Concepts, Techniques and Tools Revised edition*. Princeton University Press.

Meucci, Attilio. 2009. "Managing Diversification." *Risk* 22 (5): 74–79.

Qian, Edward E. 2016. *Risk Parity Fundamentals*. Campman; Hall/CRC.

Rabiner, L., and B. Juang. 1986. "An Introduction to Hidden Markov Models." *IEEE ASSP Magazine* 3 (1): 4–16. https://doi.org/10.1109/MASSP.1986.1165342.

Rockafellar, R. Tyrrell, and Stanislav Uryasev. 2000. "Optimization of Conditional Value-at-Risk." *Journal of Risk* 2: 21–41.

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