Quantitative Risk Management

Asset Allocation and Risk Management

Martin B. Haugh

Department of Industrial Engineering and Operations Research
Columbia University

Outline

Review: Mean-Variance without a Riskfree Asset

Have n risky securities with corresponding return vector, $\mathbf{R}$, satisfying

$$\mathbf{R} \sim \mathsf{MVN}_n(\boldsymbol{\mu}, \Sigma).$$

Mean-variance portfolio optimization problem is formulated as:

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \Sigma w \tag{1}$$

$$\begin{aligned} \text{subject to} \qquad \mathbf{w}^\top \boldsymbol{\mu} &= p \\ \text{and} \qquad \mathbf{w}^\top \mathbf{1} &= 1. \end{aligned}$$

(1) is a quadratic program (QP)

- can be solved via standard Lagrange multiplier methods.

Note that specific value of p will depend on investor's level of **risk aversion**.

Review: Mean-Variance without a Riskfree Asset

- When we plot the mean portfolio return, p, against the corresponding minimized portfolio volatility / standard deviation we obtain the so-called portfolio frontier.

- Can also identify the portfolio having minimal variance among all risky portfolios: the minimum variance portfolio.

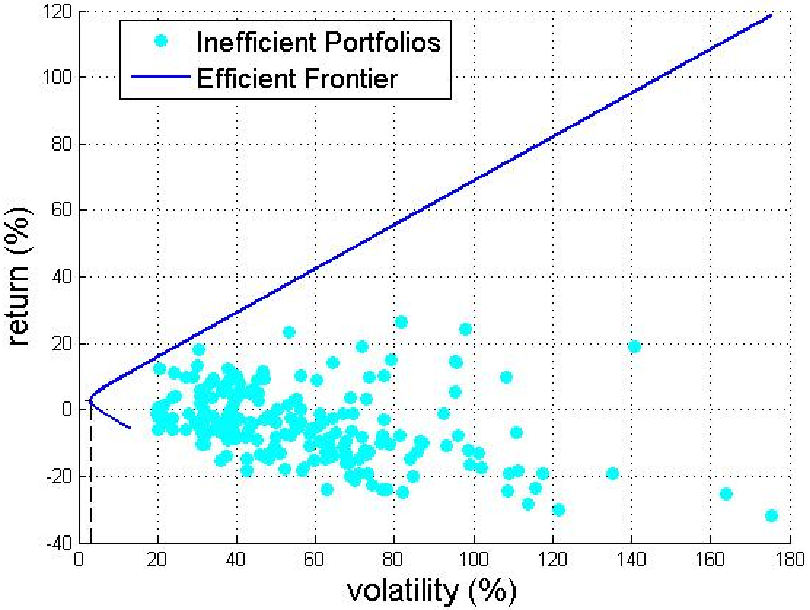
  Let $\bar{R}_{mv}$ denote expected return of minimum variance portfolio.

- Points on portfolio frontier with expected returns greater than $\bar{R}_{mv}$ are said to lie on the efficient frontier.

- Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be mean-variance efficient portfolios corresponding to expected returns p_1 and p_2, respectively, with $p_1 \neq p_2$.

- Can then be shown that **all** efficient portfolios can be obtained as linear combinations of $\mathbf{w}_1$ and $\mathbf{w}_2$

  - an example of a 2-fund theorem.


Inefficient Portfolios
Efficient Frontier
return (%)
volatility (%)
120
100
80
60
40
20
0
-20
-40
0
20
40
60
80
100
120
140
160
180

Review: Mean-Variance with a Riskfree Asset

Assume that there is a risk-free security available with risk-free rate equal to r_f.

Let $\mathbf{w} := (w_1, \ldots, w_n)^\top$ be the vector of portfolio weights on the n risky assets

- so $1 - \sum_{i=1}^n w_i$ is the weight on the risk-free security.

Investor's portfolio optimization problem may then be formulated as

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \Sigma \mathbf{w} \tag{2}$$

$$\text{subject to} \qquad \left(1 - \sum_{i=1}^n w_i\right) r_f \ + \ \mathbf{w}^\top \boldsymbol{\mu} \ = \ p.$$

Review: Mean-Variance with a Riskfree Asset

Optimal solution to (2) given by

$$\mathbf{w} \;=\; \xi\, \Sigma^{-1}(\boldsymbol{\mu} - r_f \mathbf{1}) \tag{3}$$

where $\xi := \sigma^2_{min}/(p - r_f)$ and

$$\sigma^2_{min} \;=\; \frac{(p - r_f)^2}{(\boldsymbol{\mu} - r_f \mathbf{1})^\top\, \Sigma^{-1}\, (\boldsymbol{\mu} - r_f \mathbf{1})} \tag{4}$$

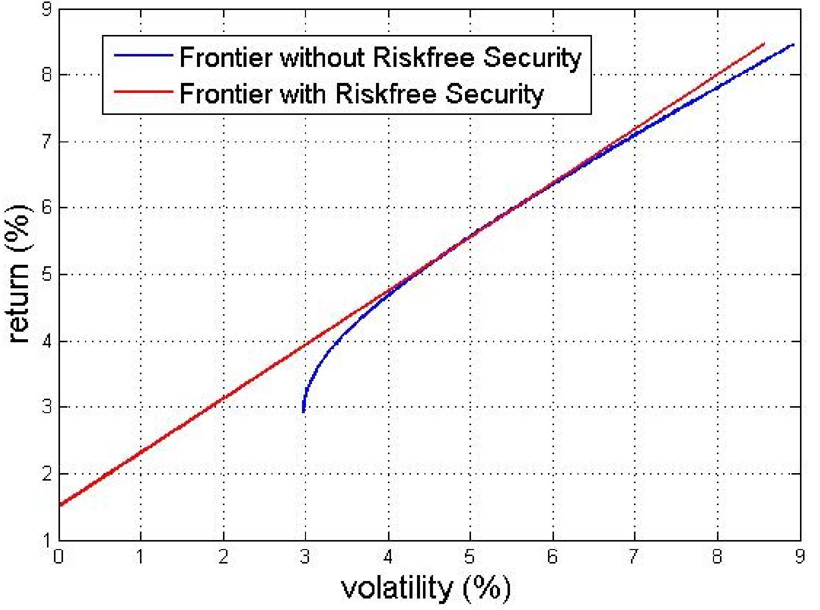
is the minimized variance.

While ξ (or p) depends on investor's level of risk aversion it is often inferred from the market portfolio.

Taking square roots across (4) we obtain

$$\sigma_{min}(p) \;=\; \frac{(p - r_f)}{\sqrt{(\boldsymbol{\mu} - r_f \mathbf{1})^\top\, \Sigma^{-1}\, (\boldsymbol{\mu} - r_f \mathbf{1})}} \tag{5}$$

– so the efficient frontier $(\sigma_{min}(p),\, p)$ is linear when we have a risk-free security:



Frontier without Riskfree Security
Frontier with Riskfree Security
return (%)
volatility (%)
1
2
3
4
5
6
7
8
9
0
1
2
3
4
5
6
7
8
9

Review: Mean-Variance with a Riskfree Asset

In fact suppose $r_f < \bar{R}_{mv}$.

Efficient frontier then becomes a straight line that is **tangent** to the risky efficient frontier and with a y-intercept equal to the risk-free rate.

We also then have a 1-fund theorem:

> Every investor will optimally choose to invest in a combination of the risk-free security and the tangency portfolio.

Review: Mean-Variance with a Riskfree Asset

Recall the optimal solution to mean-variance problem given by:

$$\mathbf{w} \;=\; \xi \, \Sigma^{-1}(\boldsymbol{\mu} - r_f \mathbf{1}) \tag{6}$$

where $\xi := \sigma^2_{min}/(p - r_f)$ and

$$\sigma^2_{min} \;=\; \frac{(p - r_f)^2}{(\boldsymbol{\mu} - r_f \mathbf{1})^\top \, \Sigma^{-1} \, (\boldsymbol{\mu} - r_f \mathbf{1})} \tag{7}$$

is the minimized variance.

The tangency portfolio, $\mathbf{w}^*$, is given by (6) except that it must be **scaled** so that its components sum to 1

- this scaling removes the dependency on p.

Question: Describe the efficient frontier if no-borrowing is allowed.

Weaknesses of Traditional Mean-Variance Analysis

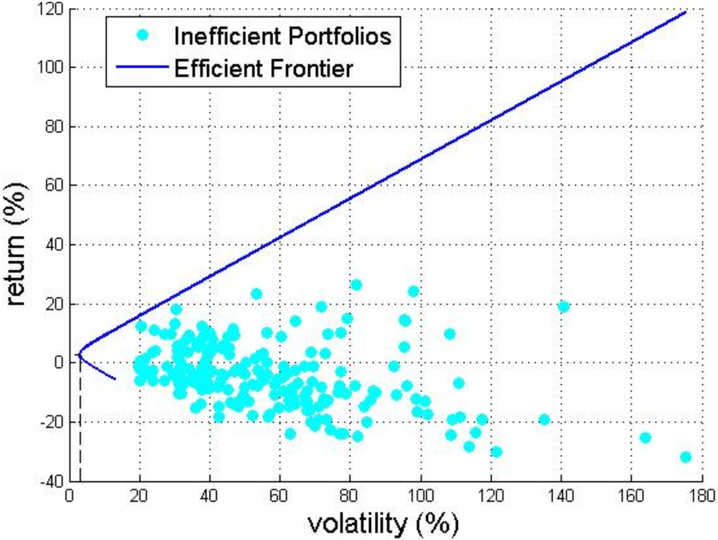
Traditional mean-variance analysis has many weaknesses when applied naively in practice.

For example, it often produces **extreme** portfolios combining extreme shorts with extreme longs

- portfolio managers generally do not trust these extreme weights as a result.

This problem is typically caused by **estimation errors** in the mean return vector and covariance matrix.

Consider again the same efficient frontier of risky securities that we saw earlier:



In practice, investors can never compute this frontier since they do not know the true mean vector and covariance matrix of returns.

The best we can hope to do is to **approximate** it. But how might we do this?

Weaknesses of Traditional Mean-Variance Analysis

One approach would be to simply estimate the mean vector and covariance matrix using historical data.

Each of dashed black curves in next figure is an **estimated** frontier that we computed by:

(i) simulating $m = 24$ sample returns from the true distribution
 - which, in this case, was assumed to be multivariate normal.

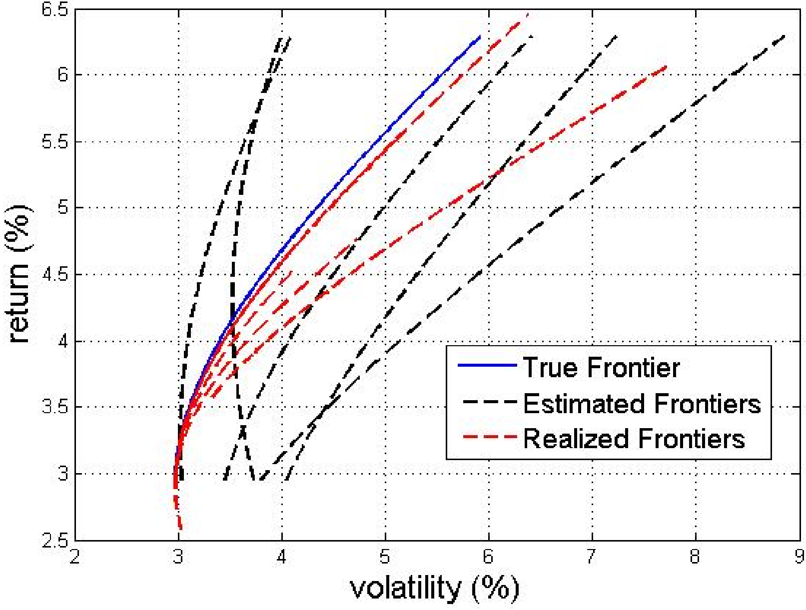
(ii) estimating the mean vector and covariance matrix from this simulated data

(iii) using these estimates to generate the (estimated) frontier.

The blue curve in the figure is the true frontier computed using the true mean vector and covariance matrix.

First observation is that the estimated frontiers are **random** and can differ greatly from the true frontier;

- an estimated frontier may lie below, above or may intersect the true frontier.



6.5
6
5.5
5
4.5
4
3.5
3
2.5
2
3
4
5
6
7
8
9
return (%)
volatility (%)
True Frontier
Estimated Frontiers
Realized Frontiers

Weaknesses of Traditional Mean-Variance Analysis

An investor who uses such an estimated frontier to make investment decisions may end up choosing a poor portfolio.

But just how poor?

The dashed red curves in the figure are the **realized** frontiers

- the true mean - volatility tradeoff that results from making decisions based on the estimated frontiers.

In contrast to the estimated frontiers, the realized frontiers must always (why?) lie **below** the true frontier.

In the figure some of the realized frontiers lie very close to the true frontier and so in these cases an investor would do very well.

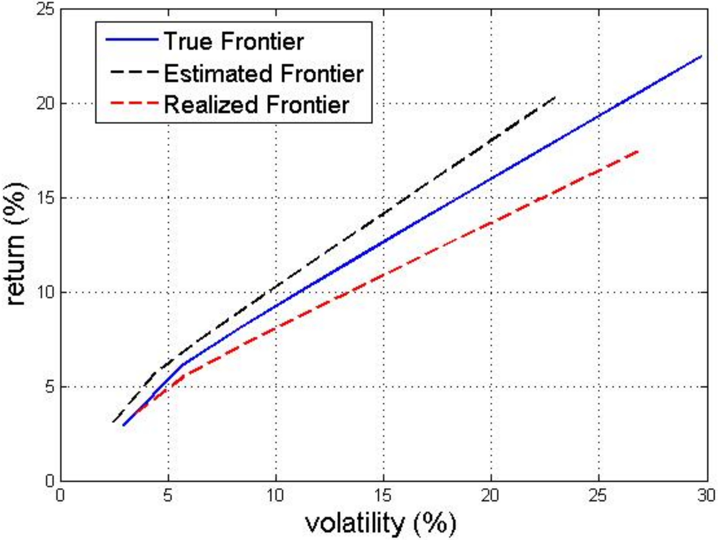
But in other cases the realized frontier is far from the (unobtainable) true efficient frontier.

Weaknesses of Traditional Mean-Variance Analysis

As a result of these weaknesses, portfolio managers traditionally had little confidence in mean-variance analysis and therefore applied it rarely in practice.

Efforts to overcome these problems include the use of better estimation techniques such as the use of shrinkage estimators, robust estimators and Bayesian techniques such as the Black-Litterman framework.

In addition to mitigating the problem of extreme portfolios, the Black-Litterman framework allows users to specify their own **subjective views** on the market in a consistent and tractable manner.

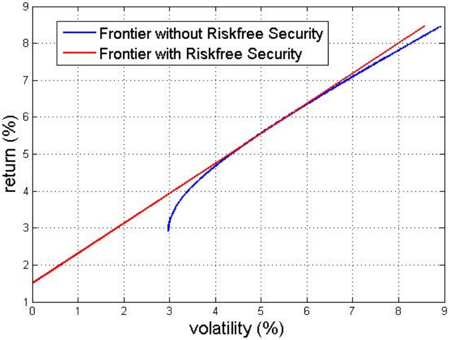


– Figure displays a robustly estimated frontier.

It lies much closer to the true frontier

- this is also the case with the corresponding realized frontier.

Review: The Capital Asset Pricing Model (CAPM)



If every investor is a mean-variance optimizer then each of them will hold the same tangency portfolio of risky securities in conjunction with a position in the risk-free asset.

Because the tangency portfolio is held by all investors and because markets must clear, we can identify this portfolio as the market portfolio.

The efficient frontier is then termed the capital market line (CML).

Review: The Capital Asset Pricing Model (CAPM)

Now let R_m and $\bar{R}_m$ denote the return and expected return, respectively, of the market, i.e. tangency, portfolio.

Central insight of the Capital Asset-Pricing Model is that in equilibrium the riskiness of an asset is not measured by the standard deviation (or variance) of its return but by its beta:

$$\beta := \frac{\mathsf{Cov}(R, R_m)}{\mathsf{Var}(R_m)}.$$

In particular, there is a linear relationship between the expected return, $\bar{R} = \mathsf{E}[R]$, of any security (or portfolio) and the expected return of the market portfolio:

$$\bar{R} \ = \ r_f \ + \ \beta \, (\bar{R}_m - r_f). \tag{8}$$

Review: The Capital Asset Pricing Model (CAPM)

The CAPM is one of the most famous models in all of finance.

Even though it arises from a simple one-period model, it provides considerable insight to the problem of asset-pricing.

For example, it is well-known that riskier securities should have higher expected returns in order to compensate investors for holding them. But how do we measure risk?

According to the CAPM, security risk is measured by its beta which is proportional to its covariance with the market portfolio

- a very important insight.

This does not contradict the mean-variance formulation of Markowitz where investors do care about return variance

- indeed, we derived the CAPM from mean-variance analysis!

The Black-Litterman Framework

We now assume that the $n \times 1$ vector of excess returns, $\mathbf{X}$, is multivariate normal conditional on knowing the mean excess return vector, $\boldsymbol{\mu}$.

In particular, we assume

$$\mathbf{X} \mid \boldsymbol{\mu} \;\sim\; \mathsf{MVN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{9}$$

where the $n \times n$ covariance matrix, $\boldsymbol{\Sigma}$, is again assumed to be known.

In contrast to the mean-variance approach, we now assume that $\boldsymbol{\mu}$ is also random and satisfies

$$\mu \;\sim\; \mathsf{MVN}_n(\boldsymbol{\pi}, \mathbf{C}) \tag{10}$$

where:

- $\boldsymbol{\pi}$ is an $n \times 1$ vector
- $\mathbf{C}$ is an $n \times n$ matrix
- Both $\boldsymbol{\pi}$ and $\mathbf{C}$ are assumed to be known.

The Black-Litterman Framework

In practice, it is common to take

$$\mathbf{C} = \tau \boldsymbol{\Sigma}$$

where τ is a given subjective constant that quantifies our level of certainty regarding the true value of $\boldsymbol{\mu}$.

In order to specify π, Black-Litterman invoked the CAPM and set

$$\boldsymbol{\pi} = \lambda \boldsymbol{\Sigma} \, \mathbf{w}_m \tag{11}$$

where:

- $\mathbf{w}_m$ is the observable market portfolio
- And λ measures the average level of risk aversion in the market.

Note that (11) can be justified by identifying λ with $1/\xi$ in (3).

Worth emphasizing that we are free to choose $\boldsymbol{\pi}$ in any manner we choose
- but of course our choice of $\boldsymbol{\pi}$ should be reasonable!

Subjective Views

Black-Litterman allows the user to express subjective views on the vector of mean excess returns, $\boldsymbol{\mu}$.

This is achieved by specifying a $k \times n$ matrix $\mathbf{P}$, a $k \times k$ covariance matrix $\boldsymbol{\Omega}$ and then defining

$$\mathbf{V} := \mathbf{P}\boldsymbol{\mu} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon} \sim \mathsf{MVN}_k(\mathbf{0}, \boldsymbol{\Omega})$ independently of $\boldsymbol{\mu}$.

Common in practice to set $\boldsymbol{\Omega} = \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top / c$ for some scalar, $c > 0$

- c represents the level of confidence we have in our views.

Specific views are then expressed by conditioning on $\mathbf{V} = \mathbf{v}$ where $\mathbf{v}$ is a given $k \times 1$ vector.

Note that we can only express linear views on the expected excess returns

- these views can be absolute or relative, however.

Examples of Subjective Views

e.g. *"the Italian index will rise by* 10%*"*

- an absolute view.

e.g. *"the U.S. will outperform Germany by* 12%*"*

- a relative view.

Both of these views could be expressed by setting

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

and $\mathbf{v} = (10\%, 12\%)^\top$.

If the i^{th} view, v_i, is more qualitative in nature, then could express it as

$$v_i = \mathbf{P}_{i.}\boldsymbol{\pi} + \eta\sqrt{(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top)_{i,i}}$$

for some $\eta \in \{-\beta, -\alpha, \alpha, \beta\}$ representing "very bearish", "bearish", "bullish" and "very bullish", respectively.

The Posterior Distribution

Goal now is to compute the conditional distribution of $\boldsymbol{\mu}$ given the views, $\mathbf{V} = \mathbf{v}$.

Conceptually, we assume that $\mathbf{V}$ has been observed equal to $\mathbf{v}$ and we now want to determine the distribution of $\boldsymbol{\mu}$ conditional on these observations.

We have the following result which follows from Baye's Theorem:

Proposition: The conditional distribution of $\boldsymbol{\mu}$ given $\mathbf{v}$ is multivariate normal. In particular, we have

$$\boldsymbol{\mu} \mid \mathbf{V} = \mathbf{v} \ \sim \ \mathsf{MVN}_n(\boldsymbol{\mu}_{bl}, \boldsymbol{\Sigma}_{bl})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{bl} &:= \boldsymbol{\pi} + \mathbf{C}\,\mathbf{P}^\top (\mathbf{P}\mathbf{C}\,\mathbf{P}^\top + \boldsymbol{\Omega})^{-1} (\mathbf{v} - \mathbf{P}\boldsymbol{\pi}) \\ \boldsymbol{\Sigma}_{bl} &:= \mathbf{C} - \mathbf{C}\,\mathbf{P}^\top (\mathbf{P}\mathbf{C}\,\mathbf{P}^\top + \boldsymbol{\Omega})^{-1} \mathbf{P}\mathbf{C}. \end{aligned}$$

The Posterior Distribution

Of course what we really care about is the conditional distribution of $\mathbf{X}$ given $\mathbf{V} = \mathbf{v}$.

Using results of previous Proposition it is easy to show that

$$\mathbf{X} \mid \mathbf{v} \sim \mathsf{MVN}(\boldsymbol{\mu}_{bl}, \boldsymbol{\Sigma}^x_{bl})$$

where

$$\boldsymbol{\Sigma}^x_{bl} := \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{bl}.$$

Once we have expressed our views, the posterior distribution can be computed and then used in any asset allocation setting.

The Posterior Distribution

Question: Suppose we choose $\boldsymbol{\pi}$ in accordance with (11) and then solve a mean-variance portfolio optimization problem using the posterior distribution of $\mathbf{X}$.

Why are we much less likely in general to obtain the extreme corner solutions that are often obtained in the traditional implementation of the mean-variance framework (where historical returns are used to estimate expected returns and covariances)?

Question: What happens to the posterior distribution of $\mathbf{X}$ as $\boldsymbol{\Omega} \rightarrow \infty$?

- Does this make sense?

An Extension of Black-Litterman: Views on Risk Factors

In the Black-Litterman framework, views are expressed on $\boldsymbol{\mu}$, the vector of mean excess returns.

May be more meaningful, however, to express views directly on the market outcome, $\mathbf{X}$.

e.g. More natural for a portfolio manager to express views regarding $\mathbf{X}$ rather than $\boldsymbol{\mu}$.

Straightforward to model this situation

- in fact can view $\mathbf{X}$ more generally as a vector of (changes in) risk factors.

Can allow these risk factors to influence the portfolio return in either a linear or **non-linear** manner.

An Extension of Black-Litterman: Views on Risk Factors

No longer model $\boldsymbol{\mu}$ as a random vector but instead assume $\boldsymbol{\mu} \equiv \boldsymbol{\pi}$ so that

$$\mathbf{X} \sim \mathsf{MVN}_n(\boldsymbol{\pi}, \boldsymbol{\Sigma})$$

Express our linear views on the market via a matrix, $\mathbf{P}$, and assume that

$$\mathbf{V}|\mathbf{X} \sim \mathsf{MVN}_k(\mathbf{PX}, \boldsymbol{\Omega})$$

where again $\boldsymbol{\Omega}$ represents our uncertainty in the views.

Letting $\mathbf{v}$ denote our realization of our view $\mathbf{V}$, can compute the posterior distribution of $\mathbf{X}$.

Applying of Baye's Theorem leads to

$$\mathbf{X}|\mathbf{V} = \mathbf{v} \sim \mathsf{MVN}_n(\boldsymbol{\mu}_{mar}, \boldsymbol{\Sigma}_{mar})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{mar} &:= \boldsymbol{\pi} + \boldsymbol{\Sigma}\,\mathbf{P}^\top \left(\mathbf{P}\boldsymbol{\Sigma}\,\mathbf{P}^\top + \boldsymbol{\Omega}\right)^{-1} (\mathbf{v} - \mathbf{P}\boldsymbol{\pi}) \qquad (12) \\ \boldsymbol{\Sigma}_{mar} &:= \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\,\mathbf{P}^\top (\mathbf{P}\boldsymbol{\Sigma}\,\mathbf{P}^\top + \boldsymbol{\Omega})^{-1}\,\mathbf{P}\boldsymbol{\Sigma}. \qquad (13) \end{aligned}$$

Note that $\mathbf{X}|\mathbf{v} \rightarrow \mathsf{MVN}_n(\boldsymbol{\pi}, \boldsymbol{\Sigma})$, the reference model, as $\boldsymbol{\Omega} \rightarrow \infty$.

Scenario Analysis

Scenario analysis corresponds to the situation where $\mathbf{\Omega} \rightarrow \mathbf{0}$

- the case in which the user has complete confidence in his view.

A better interpretation, however, is that the user simply wants to understand the posterior distribution when she conditions on certain outcomes or scenarios.

In this case it is immediate from (12) and (13) that

$$\mathbf{X} | \mathbf{V} = \mathbf{v} \ \sim \ \mathsf{MVN}_n(\boldsymbol{\mu}_{scen}, \boldsymbol{\Sigma}_{scen})$$

$$\begin{aligned}
\boldsymbol{\mu}_{scen} &:= \boldsymbol{\pi} \ + \ \boldsymbol{\Sigma}\, \mathbf{P}^\top \left(\mathbf{P}\boldsymbol{\Sigma}\, \mathbf{P}^\top\right)^{-1} (\mathbf{v} - \mathbf{P}\boldsymbol{\pi}) \\
\boldsymbol{\Sigma}_{scen} &:= \boldsymbol{\Sigma} \ - \ \boldsymbol{\Sigma}\, \mathbf{P}^\top (\mathbf{P}\boldsymbol{\Sigma}\, \mathbf{P}^\top)^{-1}\, \mathbf{P}\boldsymbol{\Sigma}.
\end{aligned}$$

Question: What happens to $\boldsymbol{\mu}_{scen}$ and $\boldsymbol{\Sigma}_{scen}$ when $\mathbf{P}$ is the identity matrix?

- Is this what you would expect?

Mean-VaR Portfolio Optimization

Consider the situation where an investor maximizes the expected return on her portfolio subject to a constraint on the VaR of the portfolio.

Given the failure of VaR to be sub-additive, it is not surprising that such an exercise could result in a very unbalanced and undesirable portfolio.

We demonstrate this using the setting of Example 2 of the *Risk Measures, Risk Aggregation and Capital Allocation* lecture notes.

- Consider an investor who has a budget of V that she can invest in $n = 100$ defaultable corporate bonds.
- Probability of a default over the next year is identical for all bonds and is equal to 2%.
- We assume that defaults of different bonds are independent from one another.
- Current price of each bond is $\$100$ and if there is no default, a bond will pay $\$105$ one year from now.

Mean-VaR Portfolio Optimization

- If the bond defaults then there is no repayment.
- Let

$$\Lambda_V \ := \ \{\boldsymbol{\lambda} \in \mathbb{R}^n \ : \ \boldsymbol{\lambda} \geq \mathbf{0}, \ \sum_{i=1}^{100} 100\lambda_i \ = \ V\} \tag{14}$$

  denote the set of all possible portfolios with current value V.

- Note that (14) implicitly rules out short-selling or borrowing additional cash.
- Let $L(\boldsymbol{\lambda})$ denote the loss on the portfolio one year from now and suppose the objective is to solve

$$\max_{\boldsymbol{\lambda} \in \Lambda_V} \ \mathsf{E}\left[-L(\boldsymbol{\lambda})\right] \ - \ \beta \, \mathsf{VaR}_\alpha\left(L(\boldsymbol{\lambda})\right) \tag{15}$$

  where $\beta > 0$ is a measure of **risk aversion**.

Mean-VaR Portfolio Optimization

This problem is easily solved: $\mathsf{E}[L(\boldsymbol{\lambda})]$ is identical for every $\boldsymbol{\lambda} \in \Lambda_V$ since the individual loss distributions are identical across all bonds.

Selecting the optimal portfolio therefore amounts to choosing the vector $\boldsymbol{\lambda}$ that minimizes $\mathsf{VaR}_\alpha(L(\boldsymbol{\lambda}))$.

When $\alpha = .95$ we have already seen that the optimal solution is to invest all funds, i.e. V, into just one bond.

Clearly this is not a well diversified portfolio!

As MFE point out, problems of the form (15) occur frequently in practice

- particularly in the context of risk-adjusted performance measurement.

Mean-VaR Portfolio Optimization

The following Proposition should not be too surprising given the subadditivity of VaR with elliptically distributed risk factors.

Proposition: Suppose $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$ with $\text{Var}(X_i) < \infty$ for all i and let $\mathcal{W} := \{\mathbf{w} \in \mathbb{R}^n \ : \ \sum_{i=1}^n w_i \ = \ 1\}$ denote the set of possible portfolio weights.

Let V be the current portfolio value so that $L(\mathbf{w}) = V \sum_{i=1}^n w_i X_i$ is the (linearized) portfolio loss.

The subset of portfolios giving expected return m is given by

$$\Theta := \{\mathbf{w} \in \mathcal{W} \ : \ -\mathbf{w}^\top \boldsymbol{\mu} \ = \ m\}.$$

Then if ϱ is any positive homogeneous and translation invariant risk measure that depends only on the distribution of risk we have

$$\underset{\mathbf{w} \in \Theta}{\operatorname{argmin}} \ \varrho(L(\mathbf{w})) \ = \ \underset{\mathbf{w} \in \Theta}{\operatorname{argmin}} \ \text{Var}(L(\mathbf{w}))$$

where $\text{Var}(\cdot)$ denotes variance. $\square$

Mean-CVaR Optimization

Now assume the decision maker needs to strike a balance between CVaR and expected return.

Could be formulated as either minimizing CVaR subject to a constraint on expected return.

Or as maximizing expected return subject to a constraint (or constraints) on CVaR.

Mean-CVaR Optimization: Notation

- Suppose there are a total of N securities with initial price vector, $\mathbf{p}_0$, at time $t = 0$.
- Let $\mathbf{p}_1$ denote the random security price vector at date $t = 1$.
- Let f denote the PDF of $\mathbf{p}_1$
  - don't need $\mathbf{p}_1$ to have a PDF as long as we can generate samples of $\mathbf{p}_1$.
- Let $\mathbf{w}$ denote vector of portfolio holdings that are chosen at date $t = 0$.

The **loss function** then given by $l(\mathbf{w}, \mathbf{p}_1) \; = \; \mathbf{w}^\top (\mathbf{p}_0 - \mathbf{p}_1)$.

Let $\Psi(\mathbf{w}, \alpha) =$ probability that loss function does not exceed the threshold, α.

Recalling that $\text{VaR}_\beta(\mathbf{w})$ is the β-quantile of the loss distribution, we see that

$$\Psi\left(\mathbf{w}, \text{VaR}_\beta(\mathbf{w})\right) \; = \; \beta.$$

Mean-CVaR Optimization

Also have familiar definition of CVaR_β as the expected portfolio loss conditional on the portfolio loss exceeding VaR_β:

$$\text{CVaR}_\beta(\mathbf{w}) = \frac{1}{1-\beta} \int_{l(\mathbf{w},\mathbf{p}_1) > \text{VaR}_\beta(\mathbf{w})} l(\mathbf{w},\mathbf{p}_1)\, f(\mathbf{p}_1)\, d\mathbf{p}_1. \tag{16}$$

Difficult to optimize CVaR using (16) due to the presence of VaR_β on the right-hand-side.

Key contribution of Rockafeller and Uryasev (2000) was to define the simpler function

$$F_\beta(\mathbf{w},\alpha) := \alpha + \frac{1}{1-\beta} \int_{l(\mathbf{w},\mathbf{p}_1) > \alpha} \left(l(\mathbf{w},\mathbf{p}_1) - \alpha\right) f(\mathbf{p}_1)\, d\mathbf{p}_1 \tag{17}$$

- can be used instead of (16) to model CVaR due to following proposition:

Mean-CVaR Optimization

Proposition: The function $F_\beta(\mathbf{w}, \alpha)$ is convex with respect to α. Moreover,

1. Minimizing $F_\beta(\mathbf{w}, \alpha)$ with respect to α gives $\text{CVaR}_\beta(\mathbf{w})$ and
2. $\text{VaR}_\beta(\mathbf{w})$ is a minimum point.

That is

$$\text{CVaR}_\beta(\mathbf{w}) \;=\; F_\beta\left(\mathbf{w}, \text{VaR}_\beta(\mathbf{w})\right) \;=\; \min_\alpha F_\beta(\mathbf{w}, \alpha).$$

Proof: See Rockafeller and Uryasev (2000). □

Can use this proposition to optimize CVaR_β over $\mathbf{w}$ and to simultaneously calculate VaR_β.

This follows since

$$\min_{\mathbf{w} \in \mathcal{W}} \text{CVaR}_\beta(\mathbf{w}) \;=\; \min_{\mathbf{w} \in \mathcal{W}, \alpha} F_\beta(\mathbf{w}, \alpha) \tag{18}$$

where $\mathcal{W}$ is the subset of $\mathbb{R}^N$ denoting the set of feasible portfolios.

Mean-CVaR Optimization

Can therefore optimize CVaR_β and compute the corresponding VaR_β by minimizing $F_\beta(\mathbf{w}, \alpha)$ with respect to both $\mathbf{w} \in \mathcal{W}$ and α.

And because $l(\mathbf{w}, \mathbf{p}_1)$ is linear and therefore convex in $\mathbf{w}$, it can be shown that $F_\beta(\mathbf{w}, \alpha)$ is also convex with respect to $\mathbf{w}$.

Therefore if the constraint set $\mathcal{W}$ is convex, we can minimize CVaR_β by solving a smooth convex optimization problem

- very efficient numerical techniques are available for solving these problems.

Note that the problem formulation in (18) includes the problem where we minimize CVaR subject to constraining the mean portfolio return.

The Linear Programming Formulation

Suppose now the PDF f is not available to us or that it is simply difficult to work with.

Suppose, however, that we can easily generate samples, $\mathbf{p}_1^{(1)}, \ldots, \mathbf{p}_1^{(J)}$ from f

- typically values are $J \approx 10,000$ or $J \approx 50,000$
- but appropriate value depends on dimension $n = \#$ of securities

Can then approximate $F_\beta(\mathbf{w}, \alpha)$ with $\tilde{F}_\beta(\mathbf{w}, \alpha)$ defined as

$$\tilde{F}_\beta(\mathbf{w}, \alpha) \ := \ \alpha \ + \ v \sum_{j=1}^{J} \left(l(\mathbf{w}, \mathbf{p}_1^{(j)}) - \alpha \right)^+ .$$

where $v := 1/((1-\beta)J)$.

If constraint set $\mathcal{W}$ is convex, then minimizing CVaR_β amounts to solving a non-smooth convex optimization problem.

The Linear Programming Formulation

In our setting $l(\mathbf{w}, \mathbf{p}_1^{(j)})$ is in fact linear so we can solve it using LP methods.

In fact we have the following formulation for minimizing CVaR_β subject to $\mathbf{w} \in \mathcal{W}$:

$$\min_{\alpha, \mathbf{w}, \mathbf{z}} \alpha \; + \; v \sum_{j=1}^{J} z_j$$

subject to

$$\begin{array}{rcll} \mathbf{w} & \in & \mathcal{W} & \\ l(\mathbf{w}, \mathbf{p}_1^{(j)}) \; - \; \alpha & \leq & z_j & \text{for } \; j = 1, \ldots, J \\ z_j & \geq & 0 & \text{for } \; j = 1, \ldots, J. \end{array}$$

If the constraint set $\mathcal{W}$ is linear then we have an LP formulation and standard LP methods can be used applied

- but there are much more efficient methods.

Maximizing Expected Return Subject to CVaR Constraints

Suppose we want to maximize the expected return on the portfolio subject to m different CVaR constraints.

This problem may be formulated as

$$\max_{\mathbf{w}\in\mathcal{W}} \mathsf{E}\left[-l(\mathbf{w},\mathbf{p}_1)\right]$$

subject to

$$\mathsf{CVaR}_{\beta_i}(\mathbf{w}) \ \leq \ C_i \quad \text{for} \quad i=1,\ldots,m.$$

However, we can take advantage of the proposition to instead formulate the problem as

$$\max_{\alpha_1,\ldots,\alpha_m,\mathbf{w}\in\mathcal{W}} \mathsf{E}\left[-l(\mathbf{w},\mathbf{p}_1)\right]$$

subject to

$$F_{\beta_i}(\mathbf{w},\alpha_i) \ \leq \ C_i \quad \text{for} \quad i=1,\ldots,m. \tag{19}$$

Note also that if the i^{th} CVaR constraint in (19) is binding then the optimal α_i^* is equal to VaR_{β_i}.

Beware of the Bias!

Suppose we minimize CVaR subject to some portfolio constraints by generating J scenarios and then solving the resulting discrete version of the problem.

Let $\mathbf{w}^*$ and CVaR^* denote the optimal portfolio holdings / weights and the optimal objective function.

Suppose now we estimate the out-of-sample portfolio CVaR by running a Monte-Carlo simulation to generate portfolio losses using $\mathbf{w}^*$.

Let $\widehat{\text{CVaR}}(\mathbf{w}^*)$ be the estimated CVaR.

Question: How do you think $\widehat{\text{CVaR}}(\mathbf{w}^*)$ will compare with CVaR^*?

Answer: After some consideration(!) it should be clear that we would expect ...

This bias can be severe – therefore always prudent to estimate $\widehat{\text{CVaR}}(\mathbf{w}^*)$!

Parameter Estimation – Very Important!!!

The various methods we've described all require good estimators of means, covariances etc. – without good estimators, the **methodology is of no use!**

Should be clear from mean-variance review that naive estimators typically perform very poorly – generally due to insufficient data.

Much easier to estimate covariance matrices than mean return vectors.

Covariance matrices can be estimated using robust estimators, e.g. via Kendall's τ, or shrinkage estimators, e.g. arising from factor model or Bayesian estimators.

Depending on time horizon, it may be very necessary to use time series methods, e.g. a multivariate GARCH factor model, to construct these estimators.

Very difficult to estimate mean returns and **naive** estimators based on historical data should **not** be used!

Robust, shrinkage, and time series methods may again be useful together with subjective views (imposed via a Bayesian framework) based on other analysis for estimating mean vectors

- no easy way to get rich but common sense should help avoid catastrophes!