

# AdvPortfolioPython

July 23, 2020

*Advanced Portfolio Construction and Analysis with Python*

<https://www.coursera.org/learn/advanced-portfolio-construction-python/>

Compiled by *ruud waij*

## 1 Week 1 Section 1

### 1.1 Video: Introduction to factor investing

#### 1.1.1 Smart beta

Another name for factor investing is **smart beta**. It consists of two parts:

- passive, beta: index investing (S&P 500)
- active: active trading, possibly outperforming the index.

An Active manager attempts to bring his or her skills to provide better risk adjusted returns than the benchmark index itself can deliver.

Factor investing is rule-based active investing without the subjective views of a manager.

*Indicization* refers to the trend of creating new indices that capture the portion of active management that is rules based and systematic, and in the long run should outperform the cap-weighted benchmark.

A *factor* is a variable that influences the returns of assets. It represents a commonality in the returns of assets, something outside of the individual asset.

Types of factors:

- macro factors: (industrial) growth, inflation
- statistical factors: information extracted from the data that may not be identifiable
- intrinsic factors or style factors: value-growth, momentum, low volatility ...

### 1.2 Video: Factor models and the CAPM

#### 1.2.1 Factor model

the factor model decomposes return  $R$  into the sum of the premia. The premium is the return that you (can) get in exchange for exposing yourself to that factor.

$$R_i = \beta_1 f_1 + \beta_2 f_2 + \dots + \beta_3 f_3 + \alpha + \varepsilon$$

- $R_i$ : the return

- $\beta$ : multiplier
- $f$ : return of a factor
- $\alpha$ : fixed component
- $\varepsilon$ : error term, the part of the return that factors cannot explain.

### 1.2.2 CAPM

CAPM stands for [capital asset pricing model](#) and it is a strange omission that this is not mentioned in the video!

$$E(r_i) - r_f = \frac{\text{cov}(r_i, r_m)}{\text{var}(r_m)}(E(r_m) - r_f)$$

$$E(r_i) - r_f = \beta_i(E(r_m) - r_f)$$

- $E(r_i) - r_f$ : excess return of an asset  $i$  over the risk-free rate  $r_f$
- $E(r_m) - r_f$ : excess return of the market  $m$  over the risk-free rate
- $\beta_i$ : the factor for asset  $i$ .  $\beta$  is a measure of the volatility of a security or portfolio compared to the market as a whole.
- $\text{cov}(r_i, r_m)$ : covariance of the return of asset  $i$  and the return of the market  $m$ .
- $\text{var}(r_m)$ : the variance of the return of the market  $m$ .

Since the excess return of an asset is based on the market return and the risk free rate, all assets (should, but don't) line up on a straight line: [the security market line](#) (see image below).

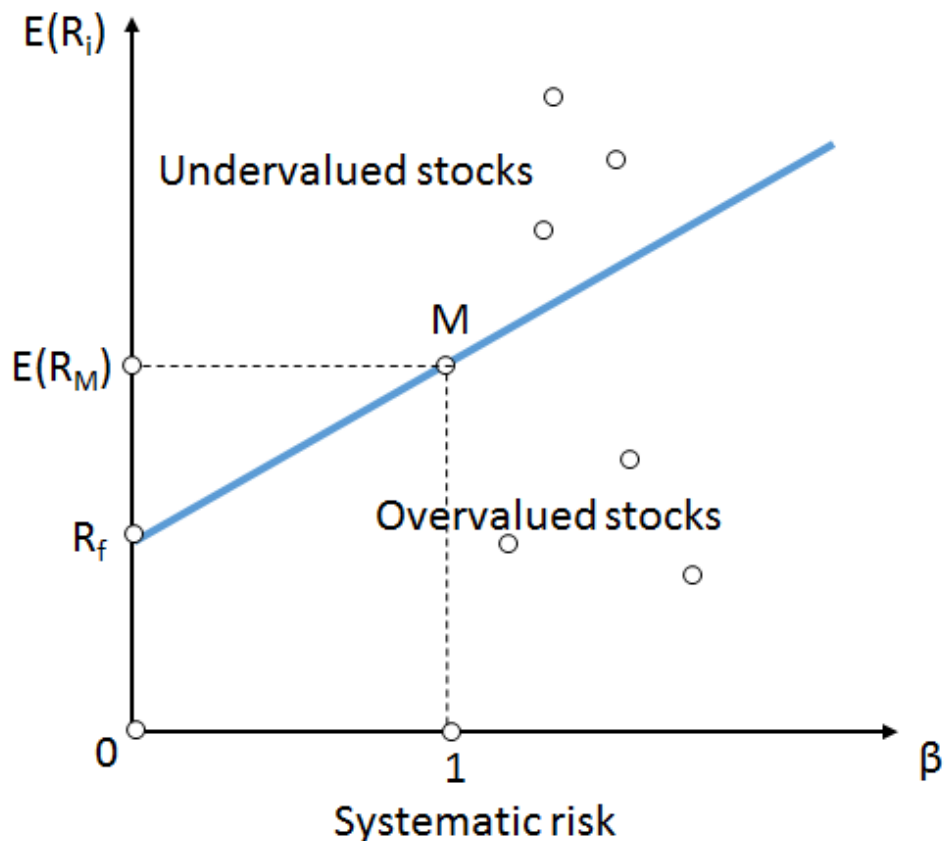


Image from Wikipedia

**Question:** According to the CAPM, the  $\alpha$  term in the CAPM Factor Model is

- Zero
- One
- Zero if Epsilon is Zero, One otherwise
- Depends on the Risk Free Rate

**Answer:**

- *Zero:* Correct. The CAPM predicts that the alpha term is zero
- *One:* This should not be selected. A Factor Model decomposes the return to Factor Returns, Alpha and an Error Term (epsilon) and the CAPM predicts that the excess return of a stock is a multiple of that stock's Beta relative to the market
- *Zero if Epsilon is Zero, One otherwise.:* This should not be selected. Epsilon is the Error Term
- *Depends on the Risk Free Rate.:* This should not be selected. The CAPM only uses the risk free rate to compute excess returns

The CAPM with just one factor is not accurate. That is why there are multi-factor models that change CAPM anomalies into regular factors.

## 1.3 Video: Multi-Factor models and Fama-French

### 1.3.1 Fama-French

The [Fama-French](#) model is an extension of the CAPM model.

**Size** In the Fama-French model the stocks are sorted according to market capitalization. Small-cap stocks on average outperform large-cap stocks. this is called the *size-effect* and is not explained by the CAPM-model (Section 1.2.2).

**Value vs. growth** The Fama-French model shows that [value stocks](#) outperform [growth stocks](#). This is probably why [book to price ratio](#) is briefly mentioned in the video.

**Question:** A Manager tells you that he concentrates his portfolio in Value stocks because Value outperforms Growth, and his portfolio has outperformed the S&P500 for the last 3 years. Assuming his statements are all True, which of the following statements can you conclude from this information.

- The portfolio will outperform the S&P500 next year -The portfolio will outperform the S&P500 next year if the [Value factor](#) has a positive risk premium next year
- If the manager does not show style drift AND the Value Factor generates a positive risk premium the next year, THEN the manager is likely to outperform, but it is not a certainty
- None of the Above

**Answer:** If the manager does not show style drift (Section 1.4.4) AND the Value Factor generates a positive risk premium the next year, THEN the manager is likely to outperform, but it is not a certainty. Correct

“The Factors Mimicking Portfolios are broad portfolio and it is possible to see a return that is different from the Factor Mimicking portfolio.” (This is not a very clear explanation. Factor mimicking is mentioned (not even explained) after this question. Sloppy work.)

### 1.3.2 Fama and French (1993)

The model includes a *small* and *value* factor to the *market* factor.

$$E[r_i] = r_f + \beta_{i,MKT}E[r_m - r_f] + \beta_{i,SMB}E[SMB] + \beta_{i,HML}E[HML]$$

- *MKT*: market factor
- *SMB*: small minus big stocks
- *HML*: high book/price (value) minus low book/price (growth)

Fama and French interpret the small stock effect and the value effect as being systematic factors.

SMB and HML are zero cost portfolios so the factors  $\beta_{i,SMB}$  and  $\beta_{i,HML}$  are centered around zero.

### 1.3.3 Other factors

In addition to *value* the following factors are recognized:

- low ‘vol’ (=volatility) beats high vol
- high quality beats low quality (??)
- [momentum](#)

Although *size* is relevant, it is not seen as a factor. The factors mentioned above are applied to small caps portfolios and large caps portfolios. (Probably, large caps portfolio are desirable, despite having a lower return.)

The factors can be used as diagnostic tools to decompose returns. This can be used to perform [style analysis](#) to determine investment behavior.

## 1.4 Video: Factor benchmarks and style analysis

### 1.4.1 Style analysis

Consider a portfolio with  $\beta = 1.3$

$$E(r_i - r_f) = \alpha + 1.3E(r_m - r_f)$$

$$E(r_i) = \alpha + [-0.3r_f + 1.3E(r_m)]$$

Factor benchmark  $[-0.3r_f + 1.3E(r_m)]$  is a short position of \$0.30 in cash (T-bills) and a leveraged position of \$1.30 in the market portfolio. This can be earned without intervention by an asset manager. The  $\alpha$  can be seen as the value that was added by the manager.

**Question:** Assume the risk free rate is 1% per year and the Stock Market returned 11% in a given year. An Active Manager “beat the market” and generated a 14% return and had a Beta of 1.3. Did the manager generate an  $\alpha = 0$ ;  $\alpha > 0$ ;  $\alpha < 0$ ?

**Answer:**  $r_f = 1\%$ ,  $r_m = 11\%$ ,  $r_i = 14\%$ ,  $\beta = 1.3$

$$E(r_i - r_f) = \alpha + \beta E(r_m - r_f)$$

$$\alpha = E(r_i - r_f) - \beta E(r_m - r_f) = 14 - 1 - 1.3 * (11 - 1) = 0$$

### 1.4.2 Sharpe style analysis

Model:

$$R_m = W_1R_{i1} + W_2R_{i2} + W_3R_{i3} + \alpha + \varepsilon$$

- $W_i$ : weight of a part  $i$  of the portfolio with  $\sum_i W_i = 1$  and all  $W_i > 0$ .
- $R_i$ : return of a particular type of investment (oil companies, european companies, it could be anything). It is an explanatory variable.

You run a [regression](#) to determine if  $\alpha > 0$ . If it is, that is due to the actions of the manager. The regression is solved through [quadratic programming](#), repeated for a sliding window of 1-3 years.

### 1.4.3 Quality of fit

Quality of fit:

$$PSEUDOR^2 = \frac{VAR(R_m) - VAR(\varepsilon)}{VAR(R_m)}$$

### 1.4.4 Style drift

As the time window moves, you can see the weights  $W_i$  change. This is called *style drift*.

## 2 Week 1 Section 2

### 2.1 video: Shortcomings of cap-weighted indices

#### 2.1.1 inefficiency of cap-weighted (cw) benchmarks

Portfolios that use *cw* indices are well within the [efficient frontier](#). they do not give the highest return for a given level of volatility. However, the portfolios on the efficient frontier suffer from [look-ahead bias](#).

[Platen and Rendek \(2010\)](#) found that the Sharpe-ratio of equally weighted (ew) portfolios was higher than that of the cap-weighted portfolio. The cap-weighted portfolio is not well diversified and holds unrewarded risk.

#### 2.1.2 Smart weighted benchmarks

- equally-weighted benchmarks
- minimum variance benchmarks
- risk parity benchmarks

#### 2.1.3 Monkeys!

[Clare, Motson and Thomas \(2013\)](#) found that randomly selecting and weighting stocks outperforms the cw index. [This story explains their approach](#).

### 2.2 Video: From cap-weighted benchmarks to smart-weighted benchmarks

#### 2.2.1 shortcomings of cap-weighted indices

- Cw indices have an inefficient diversification because high allocation to large cap stocks and [growth stocks](#). There are unrewarded [specific risks](#) that lead to a sub-optimal risk reward ratio.
- Cw indices provide an inefficient exposure to rewarded [systematic risks](#). As explained by Fama and French (Section [1.3.2](#)), small caps stocks and value stocks tend to have a higher reward.

#### 2.2.2 Smart factor benchmarks

A smart factor benchmark can be constructed using this two step process:

1. Select the type of factor exposure that you want to hold in your portfolio (value, size, momentum, volatility (see Section [1.3.3](#))).

Valuation	Size	Momentum	Volatility
Value	Large Cap	Past Winners	High Vol
Growth	Mid Cap	Past Losers	Low Vol
Book to market ratio	Freefloat adjusted market cap	Cumulative return over the past year	Vol of weekly returns

2. Select your preferred weighting scheme(s).

- Equally-Weighted
- Efficient Minimum Variance
- Risk Parity
- ...

### 3 Week 2 Section 1

#### 3.1 Video: The curse of dimensionality

##### 3.1.1 Problem

We have to estimate the parameters for all  $N$  portfolio constituents:

- $N$  return parameters
- $N$  volatility parameters
- $N * (N - 1)/2$  correlation parameters

This gives too much data. We might have to deal with 10 years of daily returns for 5000 stocks:  $10 * 250 * 5000$  parameters.

**Question:** What is the number of parameters required for mean-variance optimization based on the S&P 500 universe, which contains 500 stocks?

**Answer:**  $N = 500 : N + N + N * (N - 1)/2 \rightarrow 125750$

##### 3.1.2 Possible cures

- increase the sample size to estimate all parameters accurately
  - increase sample period
  - increase frequency
- decrease number of parameters
  - decrease number of assets  $N$ .
  - decrease the number of parameters for a fixed  $N$ .

##### 3.1.3 Extreme example 1: no model risk - high sample risk

Reduce the correlation parameters to a sample covariance estimate.

$$\hat{S}_{ij} = \frac{1}{t} \sum_{t=1}^T (R_{ijt} - \bar{R}_i)(R_{jt} - \bar{R}_j)$$

$$\bar{R}_i = \frac{1}{T} \sum_{t=1}^T R_{it}$$

$$\bar{R}_j = \frac{1}{T} \sum_{t=1}^T R_{jt}$$

- $\hat{S}_{ij}$ : sample covariance estimate
- $R_{it}$ : historical return of  $R_i$  at time  $t$ .

- $T$ : the observed period
- $\bar{R}_i$ : the mean of  $R_i$  over time  $T$

### 3.1.4 Extreme example 2: low sample risk - high model risk

This approach uses a constant correlation model (CC) in which all  $N(N-2)/2$  individual correlation parameters  $\rho_{ij}$  between returns are replaced with a single correlation parameter  $\hat{\rho}$ .

$$\hat{\sigma}_{ij}^{CC} = \hat{\sigma}_i \hat{\sigma}_j \hat{\rho}$$

- $\hat{\sigma}_{ij}^{CC}$ : constant correlation model covariance estimate for  $i, j$
- $\hat{\sigma}_i$ : estimator for stock  $i$  volatility
- $\hat{\rho}$ : correlation parameter

This single correlation parameter is estimated by the average:

$$\hat{\rho} = \frac{1}{N(N-1)} \sum_{i,j=1; i \neq j}^N \hat{\rho}_{ij}$$

**Question:** What is the number of parameter estimates required for mean-variance optimization based on the S&P 500 universe, when using the constant correlation covariance matrix estimate?

**Answer:** We need 500 expected return estimates ( $\bar{R}_i$ ), 500 volatility parameter estimates ( $\hat{\sigma}_i$ ), and also one correlation parameter estimate ( $\hat{\rho}$ ).

## 3.2 Video: Estimating the Covariance Matrix with a Factor Model

### 3.2.1 Factor-based covariance estimate

Assume stock returns are driven by a limited set of factors:

$$R_{it} = \mu_i + \beta_{i1}F_{1t} + \dots + \beta_{ik}F_{kt} + \dots + \beta_{iK}F_{Kt} + \epsilon_{it}$$

- $R_{it}$ : return of asset  $i$  at time  $t$
- $\mu_i$ : ??
- $\beta_{ik}$ : sensitivity of asset  $i$  with respect to factor  $F_{kt}$
- $F_{kt}$ : factor
- $\epsilon_{it}$ : error term, the part that is not explained by the factor model.
- $K$ : the number of factors

Variance for the 2-factor case:

$$\sigma_i^2 = \beta_{i1}^2 \sigma_{F1}^2 + \beta_{i2}^2 \sigma_{F2}^2 + 2\beta_{i1}\beta_{i2} \text{Cov}(F_1, F_2) + \sigma_{\epsilon_i}^2$$

Covariance for the 2-factor case:

$$\sigma_{ij} = \beta_{i1}\beta_{j1}\sigma_{F1}^2 + \beta_{i2}\beta_{j2}\sigma_{F2}^2 + (\beta_{i1}\beta_{j2} + \beta_{i2}\beta_{j1})\text{Cov}(F_1, F_2) + \text{Cov}(\epsilon_{it}, \epsilon_{jt})$$



We assume that the error terms  $\epsilon_{it}$  and  $\epsilon_{jt}$  are uncorrelated. This means that the [specific risk](#) for these stocks are uncorrelated. We introduce *model risk* to reduce the number of parameters.

$$Cov(\epsilon_{it}, \epsilon_{jt}) = 0$$

### 3.2.2 Case with uncorrelated parameters

General composition of returns for  $K$  factors:

$$cov(R_i(t), R_j(t)) = \sum_{k=1}^K \beta_{ik} \beta_{jk} \sigma_{F_k}^2 + cov(\epsilon_i(t), \epsilon_j(t))$$

Assuming the error terms are uncorrelated  $cov(\epsilon_{it}, \epsilon_{jt}) = 0$ :

$$\sigma_{ij} = cov(R_i(t), R_j(t)) = \sum_{k=1}^K \beta_{ik} \beta_{jk} \sigma_{F_k}^2 \quad \text{for } i \neq j$$

$$\sigma_{ii} = \sigma_i^2 = cov(R_i(t), R_i(t)) = \sum_{k=1}^K \beta_{ik}^2 \sigma_{F_k}^2 \quad \text{for } i = j$$

**Question:** How many parameters do you need to estimate when using a 2-factor models for estimating the covariance matrix of a universe of 500 stocks?

**Answer:** We first need 500 volatility estimates  $\sigma_i$  for individual stock returns ( $i = 1 \dots 500$ ), plus 500 estimates of betas of stocks with respect to factor  $k = 1$  ( $\beta_{i1}$ ), 500 estimates of betas of stocks with respect to factor  $k = 2$  ( $\beta_{i2}$ ), and finally 2 volatility estimates for factor returns, which gives a total of  $500+500+500+2=1,502$ , which compares favorably to  $500 \times 499 / 2 = 124,750$  when using the sample covariance matrix estimate.

### 3.2.3 Choice of factor model

The simplest model is Sharpe's single-factor market model (1963):

$$R_{i,t} - r_{f,t} = \alpha_i + \beta_i(R_{M,t} - r_{f,t}) + \epsilon_{i,t}$$

There are three families of factor models: - explicit macro factor model with inflation, growth, interest rates, time spread (??) - explicit micro factor model with stock specific factors: country, industry, size, book-to-market. - implicit model with statistical factors: perform statistical analysis on data to determine orthogonal uncorrelated factors

## 3.3 Video: Honey I Shrank the Covariance Matrix!

### 3.3.1 Shrinkage

The curse of dimensionality can be handled by a new methodology: [shrinkage](#).

There is a trade-off between sample risk and model risk. Sample-based estimates for covariance parameters have lot of sample risk, too many parameters to estimate, but there is no model risk.

There are other methodologies like the constant correlation methodology (see Section 3.1.4) or factor based methodology (see Section 3.2.1) that suffer from a lower degree of sample risk, because they allow to reduce the number of parameters to estimate. But that came at the cost of introducing some kind of structure, and therefore there is some fair amount of model risk.

Statistical shrinkage mixes the two methodologies to deliver the optimal trade-off.

$$\hat{S}_{shrink} = \hat{\delta}^* \hat{F} + (1 - \hat{\delta}^*) \hat{S}$$

- $\hat{S}_{shrink}$ : covariance metrics parameter of the covariance matrix
- $\hat{\delta}^*$ : percentage to mix the two estimators described below
- $\hat{F}$ : the factor model-based estimator for the covariance matrix (with model risk)
- $\hat{S}$ : data based estimate for the covariance matrix (with sample risk)

**Question:** Consider two stocks with sample volatility estimates at 20% and 30%, respectively, and sample correlation at .75. Further assume that the average of the sample correlation estimates of all stocks in the universe is .5. What is for these two stocks the sample-based covariance estimate, the constant correlation covariance estimate and the covariance estimate based on statistical shrinkage with a shrinkage factor of 50%?

**Answer:** The sample-based estimate  $\hat{\sigma}_{1,2} = \sigma_1 \sigma_2 \rho_{1,2}$  is  $20\% * 30\% * 0.75 = 0.045$ . The constant correlation estimate  $\hat{\sigma}_{ij}^{CC} = \hat{\sigma}_i \hat{\sigma}_j \hat{\rho}$  is  $20\% * 30\% * 0.5 = 0.03$ . The shrinkage estimate  $\hat{\sigma}_{1,2}^{shrink} = \delta \hat{\sigma}_{1,2} + (1 - \delta) \hat{\sigma}_{ij}^{CC}$  with  $\delta = 0.5$  is  $(0.045 + 0.03)/2 = 0.0375$ .

Performing statistical shrinkage is formally equivalent to introducing min/max weight constraints.

Just LeDoit in Python: [sklearn.covariance.LedoitWolf](#)

## 4 Week 2 Section 2

### 4.1 Video: Portfolio Construction with Time-Varying Risk Parameters

#### 4.1.1 Estimating volatility

There is a curse of non-stationarity: the parameters vary over time.

Define  $\sigma_T$  as the volatility between day  $T$  and day  $T + 1$  as estimated at the end of day  $T$ :

$$\sigma_T^2 = \frac{1}{T} \sum_{t=1}^T (R_t - \bar{R})^2$$

$$\bar{R} = \frac{1}{T} \sum_{t=1}^T R_t$$

Assume that the mean  $\bar{R} = 0$  ( $R_t$  is centered around zero). The variance  $\sigma_T^2$  simplifies to:

$$\sigma_T^2 = \frac{1}{T} \sum_{t=1}^T R_t^2$$

**Question:** Consider the following stream of returns: +1%, -2%, -1%, +2%. What are the corresponding (arithmetic) average return and volatility estimates?

```
[1]: import pandas as pd
returns= pd.Series([0.01,-0.02,-0.01,0.02])
returns.mean(), returns.std(ddof=0)
```

```
[1]: (0.0, 0.015811388300841896)
```

**Answer:** Average return is  $(1\%-2\%-1\%+2\%)/4=0\%$ . Variance of returns is  $(1\%^2-2\%^2-1\%^2+2\%^2)/4=0.025\%$ . Volatility is square-root of variance:  $\sqrt{0.025\%}=1.58\%$ .

#### 4.1.2 Curse of non-stationarity

When trying to reduce sample risk (see Section 3.1.3), it is better to increase the frequency than increasing the time period in case of non-stationary return distributions.

#### 4.1.3 Expanding window analysis

As time goes by, you add data to calculate the estimate of the volatility.

#### 4.1.4 Rolling window analysis

As new data becomes available, you remove the oldest data to calculate the estimate of volatility. The size of the observation window remains the same.

**Question:** What type of data would give you the best estimation power for covariance matrix parameters, assuming constant parameters?

**Answer:** Weekly data for 5 year. This gives you  $52 \times 5 = 260$  data points. Of course, if risk parameters are time-varying, it may not be such a good idea to use data extending over such a long time period.

### 4.2 Video: Exponentially weighted average

#### 4.2.1 Historical volatility estimate

$$\sigma_T^2 = \frac{1}{T} \sum_{t=1}^T R_t^2 \quad \text{if } \bar{R} = 0$$

In the above formula each data point contributes with weight  $\alpha = \frac{1}{T}$

$$\sigma_T^2 = \sum_{t=1}^T \alpha_t R_t^2 \quad \text{where } \sum_{t=1}^T \alpha_t = 1$$

### 4.2.2 EWMA model

In an *exponentially weighted moving average* model (EWMA) the weights decline exponentially as we move back through time.

$$\alpha_t = \frac{\lambda^{T-t}}{\sum_{t=1}^T \lambda^{T-1}}$$

- $\lambda^T$ : decay factor ( $0 < \lambda < 1$ ), a low factor puts emphasis on recent data points.  $\lambda = 0.9$  has been found to be a good value.

Covariance parameter estimate:

$$\text{cov}(R_i, R_j) = \sum_{t=1}^T \alpha_t (R_{i,t} - \bar{R}_i)(R_{j,t} - \bar{R}_j)$$

Because EWMA puts emphasis on more recent data, we can use *expanding window analysis* (see Section 4.1.3) and do not have to rely on *rolling window analysis*. (On the other hand, why not limit the data that you use for calculations if part of the data contributes very little??).

The problem with rolling window analysis is that as long as the data point is within the rolling window, it matters, and whenever it is out of the rolling window, it doesn't matter at all. So the day before it gets out, it's as important as the most recent observation, the next day it's out of the rolling window and no longer matters. It is more intuitive to let the importance of each observation decrease gradually over time. We keep all observations and we use a weighting scheme that gives more importance to recent observations.

## 4.3 Video: ARCH and GARCH Models

### 4.3.1 ARCH model

ARCH stands for *autoregressive conditional heteroskedasticity*. 'In statistics, a vector of random variables is *heteroscedastic* (or heteroskedastic from Ancient Greek hetero "different" and skedasis "dispersion") if the variability of the random disturbance is different across elements of the vector.'

In an *ARCH(T) model* we assign some weight to the long-run variance  $V_L$ .

$$\sigma_T^2 = \gamma V_L + \sum_{t=1}^T \alpha_t R_t^2$$

where

$$\gamma + \sum_{t=1}^T \alpha_t = 1$$

*ARCH(1) model*

$$\sigma_T^2 = \gamma V_L + \alpha R_T^2$$

where

$$\gamma + \alpha = 1$$

### 4.3.2 GARCH model

**GARCH** stands for *generalized autoregressive conditional heteroskedasticity model*.

In GARCH(1,1) we additionally assign some weight to the previous variance estimate to capture *volatility clustering*. Levels of volatility values are clustered in time.

$$\sigma_T^2 = \gamma V_L + \alpha R_T^2 + \beta \sigma_{T-1}^2$$

with

$$\gamma + \alpha + \beta = 1$$

- $\sigma_T^2$ : new estimate of volatility
- $\gamma V_L$ : contribution of long term volatility
- $\alpha R_T^2$ : contribution of last return
- $\beta \sigma_{T-1}^2$ : contribution of previous estimate of volatility

**Question:** Suppose the estimation of a GARCH(1,1) model on daily data gives:

$$\sigma_T^2 = 0.000002 + 0.13R_T^2 + 0.86\sigma_{T-1}^2$$

and also suppose the last daily estimate of the volatility is 1.6% per day and the most recent percentage change in the market variable is 1%. (The change in the market variable is the return! I wonder if *variable* is a typo and *value* was meant.) What is the new daily volatility estimate?

**Answer:**  $\gamma V_L = 0.000002$     $\alpha = 0.13$     $\beta = 0.86$

Mind that the GARCH formula uses variance, not volatility (=standard deviation)

```
[9]: import math
alpha=0.13
beta=0.86
gammaV=0.000002
return_val= 0.01
volatility=0.016
math.sqrt(gammaV + alpha * return_val*return_val + beta * volatility*volatility)
```

```
[9]: 0.015334927453366058
```

### 4.3.3 Variations on GARCH

In model **GARCH(P,Q)**  $p$  is the number of past return data points and  $q$  is the number of contributing previous volatility estimates to compute the volatility estimate.  $\omega$  is the contribution of the long term volatility.

$$\sigma_T^2 = \omega + \sum_{i=1}^p \alpha_i R_{T-i}^2 + \sum_{j=1}^q \beta_j \sigma_{T-j}^2$$

$$\omega = \gamma V_L$$

To take into account that volatility changes over time, we introduce additional parameters:  $\alpha_1 \dots \alpha_p \beta_1 \dots \beta_q \gamma$ . This increases the curse of dimensionality (see Section ??).

#### 4.3.4 Factor Garch

The *orthogonal (O)GARCH model* is a factor model for explaining co-variance terms between two different assets, and it is a factor model with uncorrelated orthogonal factors. It only allows for time variation in the variance of the factors.

$$\hat{\sigma}_{ij}^{OGARCH} = \hat{\sigma}_{ij}(t) = \sum_{k=1}^K \hat{\beta}_{ik} \hat{\beta}_{jk} \hat{\sigma}_{F_k}^2(t)$$

**Question:** How many parameters do you need to estimate when using a 2-factor models with GARCH(1,1) model for the volatility of each one of the two factors? **Not mentioned:** you have a 500 stock portfolio.

**Answer:** We first need 500 volatility estimates for individual stock returns, plus 500 estimates of betas of stocks with respect to factor 1, 500 estimates of betas of stocks with respect to factor 2, and finally 3 GARCH parameter estimates for each factor, which gives a total of  $500+500+500+2 \times 3=1,506$ , which is not much more than if we had assumed constant volatility parameters.

## 5 Week 3 Section 1

### 5.1 Video: Lack of Robustness of Expected Return Estimates

*Sample based information is, unfortunately, close to useless when it comes to expected return estimation. Sample based expected return can be very sample dependent. Small changes in the sample will lead to large changes in the sample based estimate for expected returns. Sample based expected return estimators are extremely noisy, especially for high volatility portfolios. So, the confidence intervals are very large, we have very little confidence that the estimator that we come up with is of any meaningfulness.*

#### 5.1.1 Frequentist versus Bayesian statics.

In [Frequentist statistics](#) information is only gathered from taking samples. [Bayesian statistics](#) uses prior knowledge to draw conclusions.

#### 5.1.2 Bayesian statistics and statistical shrinkage again!

The sample mean estimate might be improved by shrinking the individual means to the *grand sample mean*.

$$\bar{\mu}_i = \frac{1}{T} \sum_{t=0}^{T-1} R_{t,t+1}^i$$

$$\bar{\mu} = \frac{1}{N} \sum_{i=1}^N \bar{\mu}_i$$

$$\hat{\mu}_i = \delta \bar{\mu} + (1 - \delta) \bar{\mu}_i$$

- $N$ : number of stocks
- $T$ : the number of (equally spaced in time) samples for each stock  $i$ .
- $\bar{\mu}_i$ : sample based average return for each stock  $i$  between  $t$  and  $t + 1$  (this is noisy data).
- $\bar{\mu}$ : the average of all  $\bar{\mu}_i$  (grand mean)
- $\hat{\mu}_i$ : expected return of each stock  $i$  of the  $N$  stocks, using shrinkage
- $\delta$ : shrinkage factor (0..1)

**Question:** Consider 3 assets with sample means equal to 10%, 15% and 20%, and assume a shrinkage factor  $\delta=50\%$ . What is the shrinkage estimator for the expected return on these 3 assets?

```
[12]: import pandas as pd
sample_means= pd.DataFrame([0.1,0.15,0.2])
grand_mean= sample_means.mean()
delta= 0.5
shrinkage_means= delta*grand_mean + (1-delta)* sample_means
shrinkage_means
```

```
[12]:      0
0  0.125
1  0.150
2  0.175
```

**Answer:** Grand mean =  $(10\%+15\%+20\%)/3 = 15\%$ . Shrinkage estimator for asset 1:  $50\% \times 10\% + 50\% \times 15\% = 12.5\%$ . Shrinkage estimator for asset 2:  $50\% \times 15\% + 50\% \times 15\% = 15\%$ . Shrinkage estimator for asset 3:  $50\% \times 20\% + 50\% \times 15\% = 17.5\%$ .

## 5.2 Video: Agnostic Priors on Expected Return Estimates

To maximize the Sharpe Ratio of a portfolio  $SR_p$ , we need we need the expected return of the portfolio  $\mu_p$ , for which we need the expected returns of the individual components.

$$SR_p \equiv \frac{\mu_p - r}{\sigma_p}$$

### 5.2.1 First agnostic prior: expected returns are all equal

We can use as prior knowledge that the expected return of each component is equal to the grand mean (see Section 5.1.2). The estimate for all returns is now equal, independent of their volatility.

### 5.2.2 Second agnostic prior: Sharpe ratios are all equal

Sharpe ratios are constant across assets. Excess expected return (return minus risk free return) is proportional to volatility.

$$\mu_i - r_f = \lambda \sigma_i$$

- $\lambda$ : Sharpe ratio

$$SR_p = \frac{\sum_{i=1}^N w_i(\mu_i - r_f)}{\sqrt{\sum_{i,j=1}^N w_i w_j \sigma_{ij}}} = \frac{\sum_{i=1}^N w_i(\mu - r_f)}{\sqrt{\sum_{i,j=1}^N w_i w_j \sigma_{ij}}} = \lambda \frac{\sum_{i=1}^N w_i \sigma_i}{\sqrt{\sum_{i,j=1}^N w_i w_j \sigma_{ij}}}$$

- $\sum_{i=1}^N w_i \sigma_i$ : weighted average of the component volatilities
- $\sqrt{\sum_{i,j=1}^N w_i w_j \sigma_{ij}}$ : portfolio volatility

We can maximize the portfolio Sharpe ratio by maximizing the numerator given by the weighted average of the volatilities divided by portfolio volatility. This ratio is known as *diversification ratio*. We do not have to know what the value of  $\lambda$  is.

**Question:** What is the Sharpe ratio of a portfolio of an equally-weighted portfolio of two stocks with volatility respectively equal to 20% and 30%, and a pairwise correlation .6, assuming that they both have a 70% Sharpe ratio?

- $w_1 = 0.5$  (weight)
- $w_2 = 0.5$
- $\sigma_1 = 0.2$  (volatility)
- $\sigma_2 = 0.3$
- $\rho = 0.6$  (correlation)
- $\lambda = 0.7$  (Sharpe ratio)

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} \Rightarrow \text{cov}(X, Y) = \rho(X, Y)\sqrt{\text{var}(X)\text{var}(Y)}$$

$$\text{var}(X) = \sigma_X^2$$

$$SR_p = \lambda \frac{\sum_{i=1}^N w_i \sigma_i}{\sqrt{\sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}}}$$

```
[13]: import math
w1= 0.5
w2= 0.5
sigma1= 0.2
sigma2= 0.3
rho= 0.6
lambdaa=0.7

var1= sigma1*sigma1
var2= sigma2*sigma2
cov12= rho*sigma1*sigma2

# see SRp formula above
SRp= lambdaa* (w1*sigma1 + w2*sigma2)/ math.sqrt( w1*w1*var1 +2*w1*w2*cov12+
↪w2*w2*var2 )
```



[13]: 0.7787397791074733

**Answer:** 77.87%

### 5.2.3 Rewarded versus unrewarded risk

Asset pricing theory suggests that only **systematic risk** is rewarded. **Specific risk** can be diversified away.

*Well, in this context, we may want to assume that all stocks have the same Sharpe ratio. We may want to assume that there's a relationship between excess expected return and not total risk, but the systematic part of volatility. So in other words, we may want to decompose volatility in terms of specific risk and systematic risk, and relate and come up with a better estimate for expected returns by relating it to systematic risk as opposed to relating it to total risk.*

**Question:** Assume that a stock index and a bond index have the same Sharpe ratio, and that this common Sharpe ratio is equal to 50%. Further assume that interest rate is 2% and that stock index volatility is 20% and bond index volatility is 10%. Calculate the stock index expected return and the bond index expected return.

$$\mu = \lambda\sigma + r_f$$

- $\mu$ : expected return
- $\lambda = 0.5$
- $r_f = 0.02$
- $\sigma_{bi} = 0.1$
- $\sigma_{si} = 0.2$

**Answer:** bond:  $0.5 \cdot 0.1 + 0.02 = 7\%$ ; stock:  $0.5 \cdot 0.2 + 0.02 = 12\%$

## 5.3 Video: Using Factor Models to Estimate Expected Returns

### 5.3.1 CAPM-based expected return estimates

*If CAPM (see Section 1.2.2) is the true asset pricing model, then the excess expected return is proportional to  $\beta$ .*

$$\mu_i - r_f = \beta_i(\mu_M - r_f)$$

- $\mu_i$ : expected return of stock  $i$
- $r_f$ : riskfree rate
- $\beta_i$ : factor for stock  $i$
- $\mu_M$ : market return
- $\mu_M - r_f$ : market risk premium for systematic risk

The **Treynor ratio**:

$$TR_i = \frac{\mu_i - r_f}{\beta_i} = \mu_M - r_f$$

The Treynor ratio has an identical value for all stocks:  $\mu_M - r_f$ .

**Question:** Assume that a stock has an expected return of 13% and a volatility of 20%. Further assume that the risk-free rate is 3%, and the stock beta is 1. What is the Sharpe ratio (SR) and Treynor ratio (TR) for this stock?

- $\mu_i = 0.13$
- $\sigma_i = 0.2$
- $r_f = 0.03$
- $\beta_i = 1$

$$SR_i = \frac{\mu_i - r_f}{\sigma_i}$$

$$TR_i = \frac{\mu_i - r_f}{\beta_i}$$

**Answer:** SR= (0.13-0.03)/0.2= 50% TR= (0.13-0.03)/1= 10%

CAPM is not the true asset pricing model.

**Question:** Assume that last year volatility is 20% and last year performance is 15% for stock 1, while it is 30% and 18% respectively for stock 2. Further assume that stock 1 beta is 1.1 and stock 2 beta is 0.8. Which stock should have the highest expected return if the CAPM is correct?

$$\mu_i - r_f = \beta_i(\mu_M - r_f) \Rightarrow \mu_i = \beta_i(\mu_M - r_f) + r_f$$

**Answer:** according to CAPM (see Section 1.2.2) excess return is fully predicted by the value of  $\beta$ . Stock 1 has the higher value of  $\beta$ . Historic return values have no predictive power. CAPM does not consider volatility.

### 5.3.2 Expected return estimates with multi-factor models

With multi-factor models excess expected returns are given by a combination of risk exposures times factor premia.

Stephen Ross Arbitrage pricing theory:

$$\mu_i - r_f = \sum_{k=1}^K \beta_{ik}(\mu_k - r_f) = \sum_{k=1}^K \beta_{ik} \lambda_k \sigma_k \quad \text{with } \lambda_k = \frac{\mu_k - r_f}{\sigma_k}$$

- $\mu_i - r_f$ : excess expected return for stock  $i$
- $\beta_{ik}(\mu_k - r_f)$ : the contribution of the excess expected return of stock  $k$  to that of  $i$
- $\beta_{ik}$ : the factor of stock  $k$  for stock  $i$

We need to be able to estimate the excess returns  $u_k - r_f$  of the components, which is equally difficult as estimating the expected return of  $\mu_i$ .

Different approaches: 1. *agnostic approach*: assume that all factors have the same Sharpe ratio  
2. *frequentist approach*: determine the Sharpe ratios by looking at the longest possible sample. 3.

*Active approach:* the use of qualitative or quantitative analysis by a portfolio manager (back to voodoo)

**Question:** Assume that a stock has a beta of 1 with respect to factor 1 and a beta of .5 with respect to factor 2. Further assume that the excess expected return is 6% on factor 1 and 8% on factor 2, and that the risk-free rate is 2%. What is the expected return on the stock, assuming that these two factors are the only rewarded factors?

- $\beta_1 = 1$
- $\beta_2 = 0.5$
- $\mu_1 - r_f = 0.06$
- $\mu_2 - r_f = 0.08$
- $r_f = 0.02$

**Answer:** the expected return  $\beta_1(\mu_1 - r_f) + \beta_2(\mu_2 - r_f) + r_f = 1 * 0.06 + 0.5 * 0.08 + 0.02 = 12\%$

## 6 Week 3 Section 2

### 6.1 Video: Extracting Implied Expected Returns

In the [Black-Litterman model](#) (BL) you have active views of expected returns and those are used in portfolio construction.

#### 6.1.1 Finding an anchor point

BL uses a preferred benchmark as a starting point (anchor) in the portfolio construction process. If the portfolio manager does not have a lot of confidence in his views, the portfolio should mainly consist of assets from the benchmark.

BL uses Bayesian analysis.

#### 6.1.2 Neutral/implied expected returns

The neutral prior distribution is obtained by reverse engineering, assuming the benchmark is the optimal portfolio.

Engineering of portfolio construction: - start with expected returns  $\mu_i$  and covariances  $\sigma_{i,j}$  - determine the weights  $w_i^*$  of the maximum Sharpe ratio portfolio (denoted by \*).

$$(\mu_i, \sigma_{i,j})_{i,j=1\dots N} \longrightarrow (w_i^*)_{i,j=1\dots N}$$

Unfortunately, there are no meaningful estimates of expected returns. We cannot determine stable Sharpe ratio maximizing weights  $w_i^*$ .

Instead we do reverse engineering (reverse portfolio optimization) and start with weights for a selected benchmark (cap-weighted, equally weighted, any benchmark).

- take as input:
  - the covariance parameters  $\sigma_{i,j}$
  - the weights of the preferred benchmark  $w_i^{benchmark}$ .
- Calculate (*extract*) the (implied) expected returns (vector  $\Pi$ ) that you would get from using these weights.

$$(\sigma_{i,j}, w_i^{benchmark})_{i,j=1\dots N} \longrightarrow \Pi = (\mu_i^{implied})_{i=1\dots N}$$

**Question:** Assume two uncorrelated stocks with volatility 10% and 15% respectively. Further assume that the risk-free rate is 0%. What can we say about the neutral expected returns consistent with an equally-weighted benchmark portfolio?

**Answer:** The official answer claims that we know that the Sharpe ratio maximizing weights are proportional to  $(\mu_i - r)/\sigma_i^2$ . I have no **recollection** of that.

$$\mu_1/\sigma_1^2 = \mu_2/\sigma_2^2 = 50\% \quad \mu_1/\mu_2 = \sigma_2^2/\sigma_1^2 = 2.25$$

*Bayesian prior* is that true expected returns are centered at market implied values denoted by  $\Pi$ .

$$\mu = \Pi + \varepsilon^e$$

- $\varepsilon^e \sim N(0, \tau\Sigma)$  (0 ??)
- $\tau$  is a scalar indicating the uncertainty of the prior
- $\Sigma = (\sigma_{ij})_{i,j=1\dots N}$  is the covariance matrix

## 6.2 Video: Introducing Active Views

The Black-Litterman approach mixes market (benchmark) implied expected returns with the manager's active views about those expected returns.

### 6.2.1 Active view

An *active view* is expressed as a statement that the expected return on a given asset within a portfolio  $P$  has a normal distribution with mean  $Q$  and standard deviation  $\Omega$ .  $Q$  is a vector of mean values of  $K$  views in the portfolio. The view may concern an asset or a set of assets.

$$P\mu = Q + \varepsilon^v \quad \text{where } \varepsilon^v \sim N(0, \Omega)$$

- $N$  is the total number of assets
- $N(0, \Omega)$  is the normal distribution??
- $K$  is the total number of views that will be expressed.
- $P$  is a  $K \times N$  matrix that identifies the assets involved in the views.
- $Q$  is a  $K$ -vector of expected returns on these portfolios or assets (??).
- $\Omega$  is a  $K \times N$  matrix of error terms in the views (confidence levels)
- $\varepsilon^v$ : uncertainty in the views

### 6.2.2 Black-Litterman model

The Black-Litterman model combines benchmark implied expected returns  $\Pi$  (prior return vector):

$$\mu = \Pi + \varepsilon^e \quad \text{where } \varepsilon^e \sim N(0, \tau\Sigma)$$

with active views:

$$P\mu = Q + \varepsilon^v \quad \text{where } \varepsilon^v \sim N(0, \Omega)$$

In a Bayesian framework the new expected returns  $\bar{\mu}$  (posterior return vector) can be written as:

$$\bar{\mu} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1} [(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q]$$

- $\bar{\mu}$ : expected returns vector
- $\Pi$ : the bench-marked implied expected return
- $Q$ : the active views

## 6.3 Video: Black-Litterman Analysis

### 6.3.1 Expected returns - no views

The historical mean is a noisy estimate for expected return. The expected return based on the CAPM-model is less noisy.

(I probably somehow missed it, but I do not remember seeing the CAPM expected return clearly explained. I found the explanation on [Investopedia](#).)

We start by extracting the expected return from the benchmark portfolio (cw, ew or any other benchmark). If we compare the CAPM expected returns with the benchmark expected returns, the correlation is high.

The BL approach can be extended to any benchmark. For this just derive in the first step implied expected returns consistent with the EW portfolio being the Max Sharpe Ratio portfolio.

### 6.3.2 Introducing active views

The BL model allows active views to be expressed in absolute or relative terms.

We are going to assume that the confidence levels for the views are going to be proportional to the variance of the prior. Just as in one of the classical implementation of the Black-Litterman model that is displayed in a paper by [He and Litterman](#) (version with good layout!) in 1999.

## 7 Week 4 Section 1

### 7.1 Naive Diversification

#### 7.1.1 Proverbial definition for diversification

Diversification is having a well-balanced allocation of your dollars to different assets or securities.

#### 7.1.2 Mean goal versus end goal

- *Mean goal*: well-balanced portfolio
- *End goal*: well rewarded portfolio

Diversification diversifies away some of the unrewarded risk within the portfolio and that allows us to achieve the highest possible reward per unit of risk.

### 7.1.3 Effective number of constituents (ENC)

What is the number of *meaningful* allocations to assets, small fractions do not help with diversification.

$$ENC \equiv \left( \sum_{i=1}^N w_i^2 \right)^{-1}$$

- $ENC$ : Effective number of constituents
- $w_i$ : the weight of asset  $i$  in the portfolio
- $N$ : the number of assets in the portfolio

**Example**, extreme case 1: fully concentrated portfolio with  $w_1 = 1$  and no allocation to the other  $N - 1$  assets. This gives  $ENC = \frac{1}{1^2 + 0^2 + \dots + 0^2} = 1$ .

**Example**, extreme case 2: equally weighted portfolio with  $w_i = \frac{1}{N}$ . This gives  $ENC = \frac{1}{(\frac{1}{N})^2 + \dots + (\frac{1}{N})^2} = N$ .

The equally weighted portfolio gives the maximum value for  $ENC$ .

The  $ENC$  for the *S&P500* (with 500 stocks) is typically about 4 times smaller than the nominal number of constituents. In this course,  $ENC$  is expressed as a percentage:  $ENC_{S\&P500} = 26.9\%$ .

**Question:** What is the effective number of constituents ( $ENC$ ) for a portfolio invested 20% in a 20% volatility asset and 80% in a 10% volatility asset, assuming that these assets are uncorrelated?

```
[1]: w1= 0.2
      w2= 0.8
      ENC= 1/(w1*w1+w2*w2)
      ENC
```

```
[1]: 1.4705882352941173
```

**Answer:**  $ENC = 1.47$

## 7.2 Video: Scientific diversification

### 7.2.1 GMV

In order to maximize the return, give a certain risk budget, we would like to hold the *Maximum Sharpe Ratio* (MSR) portfolio. It is difficult to reliably get the expected return values needed to calculate the portfolio weights. Instead we can get the *Global Minimum Variance* (GMV) portfolio, this portfolio does not require estimated return estimates. (See image below.)

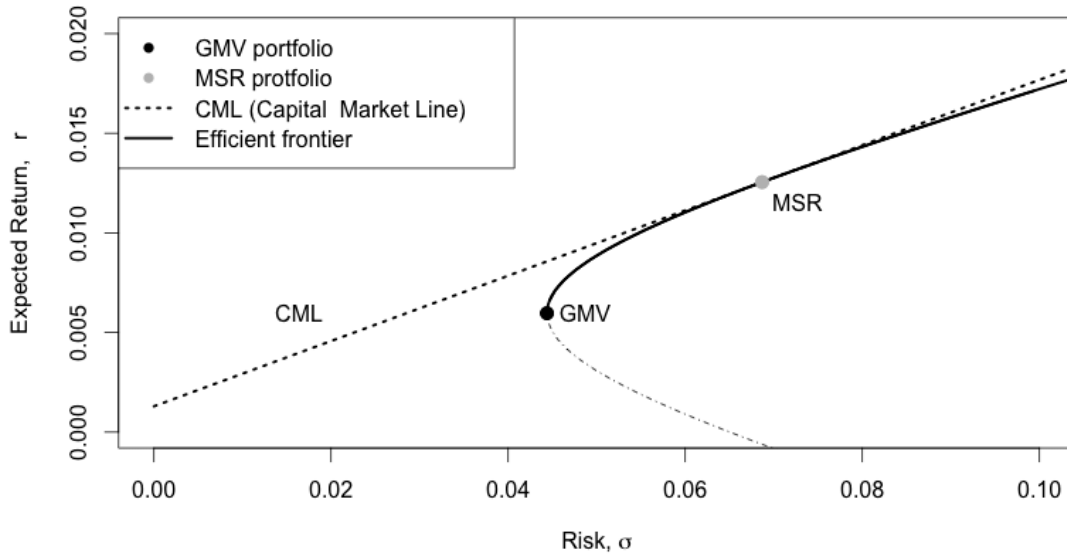


Image from ResearchGate

Derivation of the Global Minimum Variance portfolio

$$\text{Min}\sigma_p^2 = \text{Min} \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} = \text{Min} \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_i \sigma_j \rho_{ij}$$

### 7.2.2 Problems with GMV portfolios

GMV assumes equal expected returns for all assets in the portfolio, which is unlikely. When expected returns are equal, the MSR and GMV portfolio coincide. This means we can use the quadratic optimizer to find the MSR portfolio (as we did in [the first course](#)).

The optimizer will not optimize on high returns (since they are all equal), but will over-weight low volatility components. the result of this is that GMV is not a well balanced portfolio. It has been found that GMV is not consistently better than the equally weighted portfolio.

### 7.2.3 Improving GMV portfolios

There are several ways to improve the performance of GMV:

**Minimum value for ENC** GMV with an imposed minimum value for ENC (see Section [7.1.3](#)) mixes scientific diversification with naive portfolio diversification (see Section [7.1](#)).

**Same volatility** Impose that all assets have the same volatility.

$$\text{Min} \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_i \sigma_j \rho_{ij} \longrightarrow \text{Min} \sigma^2 \sum_{i=1}^N \sum_{j=1}^N w_i w_j \rho_{ij} \quad \text{if } \sigma_i = \sigma \text{ for } i = 1, \dots, N$$

The optimizer now searches for maximally decorrelated assets in the portfolio.

### 7.3 Video: Measuring risk contributions

#### 7.3.1 Shortcomings of ENC

Well balanced portfolios in terms of dollar contributions can be highly concentrated in terms of risk contributions. It is useful to know the contribution of each asset to the risk of the portfolio.

**Example:** 50% allocation in stock 1 with 30% volatility; 50% allocation in 10% volatility bond 2, correlation=0. What is the variance of the portfolio? (Watch out, we are mixing variance and volatility!)

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_i \sigma_j \rho_{ij}$$

$w_1 = 0.5$ ,  $w_2 = 0.5$ ,  $\sigma_1 = 0.3$ ,  $\sigma_2 = 0.1$  and  $\rho_{ij} = 0$  for  $i \neq j$ .

$$\sigma_p^2 = 0.5^2 \times 0.3^2 + 0.5^2 \times 0.1^2 = 0.025$$

The variance of stock 1 *in the portfolio* is 50%<sup>2</sup> of 30%<sup>2</sup>. The risk contribution for stock 1 is  $p_1 = \frac{0.5^2 \times 0.3^2}{\sigma_p^2} = 90\%$ .

Now assume, for this example, that the correlation  $\rho_{1,2} = 0.25$ .

$$\sigma_p^2 = 0.5^2 \times 0.3^2 + 0.5^2 \times 0.1^2 + 2 \times 0.5 \times 0.5 \times 0.3 \times 0.1 \times 0.25 = 0.0288$$

#### 7.3.2 Allocating the correlated component

The correlated component of the variance is split by the fraction of contribution to the portfolio.

**Question:** Consider a portfolio invested at 50% in a 30% volatility stock and 50% in a 10% volatility bond. What are the risk contributions in case the correlation between the stock and the bond returns is 0.5.

```
[21]: w1= 0.5; w1_2= w1*w1
      w2= 0.5; w2_2= w2*w2
      vol1= 0.3; vol1_2= vol1*vol1 #var1= vol1_2
      vol2=0.1; vol2_2= vol2*vol2
      rho= 0.5
      var_p= w1_2*vol1_2 + 2*w1*w2*vol1*vol2*rho + w2_2*vol2_2
      p1= (w1_2* vol1_2 + w1_2 * vol1 * vol2 * rho) / var_p
      p1
```

[21]: 0.8076923076923076

**Answer:**  $p_1 \simeq 81\%$



## 8 Week 4 Section 2

### 8.1 Video: Simplified risk parity portfolios

#### 8.1.1 Risk parity portfolio

*Risk parity portfolio* is a portfolio with equal risk contribution from both assets. It maximizes the ENC (see Section 7.3.1) applied to risk contributions, as opposed to dollar contributions. This version of ENC is called *ENCB*, known as the *effective number of correlated bets*.

**Question:** Assume an annualized volatility of 15.1% on a broad equity index and annualized volatility of 4.6% on a broad bond index, with a 0.2 correlation. What is the ENCB for the 60/40 portfolio?

```
[4]: w1= 0.6; w1_2= w1*w1 # 1 = equity
      w2= 0.4; w2_2= w2*w2 # 2 = bond
      vol1= 0.151; vol1_2= vol1*vol1 #var1= vol1_2
      vol2=0.046; vol2_2= vol2*vol2
      rho= 0.2
      var_p= w1_2*vol1_2 + 2*w1*w2*vol1*vol2*rho + w2_2*vol2_2
      p1= (w1_2* vol1_2 + w1_2 * vol1 * vol2 * rho) / var_p # risk contribution 1
      p2= (w2_2* vol2_2 + w2_2 * vol1 * vol2 * rho) / var_p # risk contribution 2
      # ENCB
      1/(p1*p1+p2*p2)
```

```
[4]: 1.1147824352230271
```

**Answer:**  $ENCB = 1.15$

In the two asset case, risk parity weights are proportional to the inverse of the volatilities. The risk contributions of asset 1 and asset 2 are equal if their weights  $w_{1,2}$  satisfy the following equation:

$$\frac{w_1}{w_2} = \frac{\sigma_2}{\sigma_1}$$

#### 8.1.2 Interesting properties

The risk parity portfolio is also known as the *equal risk contribution portfolio* (ERC). It is an inverse volatility weighted portfolio if all pairwise correlations are equal (important for portfolios with more than two assets). In general, pairwise correlations will not be equal!

**Question:** Assume that a portfolio is invested in a 20% volatility asset and in a 10% volatility asset, and further assume that these assets are uncorrelated. What is the risk parity allocation for this portfolio?

**Answer:**  $w_1 = 0.2 \quad w_2 = 0.1 \quad \frac{w_1}{w_2} = \frac{\sigma_2}{\sigma_1} = \frac{0.1}{0.2} \longrightarrow w_2 = 2 * w_1$

## 8.2 Video: Risk Parity Portfolios

### 8.2.1 General expressions

The portfolio variance  $\sigma_p^2$  is the sum of the asset  $i$  variances and covariances of  $i, j$ :

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} = \sum_{i=1}^N w_i \sigma_i^2 + \sum_{i=1}^N \sum_{i \neq j}^N w_i w_j \sigma_{ij}$$

Contribution  $p_i$  of asset  $i$  to the risk of the portfolio:

$$p_i = \frac{w_i^2 \sigma_i^2 + \sum_{j \neq i}^N w_i w_j \sigma_{ij}}{\sigma_p^2}$$

The effective number of correlated bets  $ENCB$  (similar to ENC):

$$ENCB = \left( \sum_{i=1}^N p_i^2 \right)^{-1}$$

### 8.2.2 Risk parity portfolio

*Risk parity portfolio*: choose portfolio (of  $N$  assets) with weights  $w_i$  so as to equalize risk contributions  $p_i = \frac{1}{N}$ , or equivalently, maximize  $ENCB$ . This is also known as *equal risk contribution* (ERC) portfolio.

There is no analytical way to achieve risk parity, you have to use a numerical approach. Assets having the same volatility levels is not sufficient to create a risk parity portfolio because you have to take the pairwise correlation factors into account.

Risk parity is a naïve diversification strategy. No attempt is made to maximize the Sharpe ratio.

## 8.3 Video: Comparing Diversification Options

### 8.3.1 Competing portfolio construction schemes

For US large cap stocks (1987-2018) we compare 4 schemes:

- cap-weighted (cw) portfolio
- equally-weighted (ew) portfolio
- equal risk contribution (erc) portfolio (see Section 8.2.2)
- global minimum variance portfolio (gmv) (see Section 7.2.1)

For the gmv portfolio, the minimum weight that has been chosen is  $1/(3 \times 500)$  (one third of the ew weight) and the maximum weight is  $3/500$  (three times the ew weight).

Compared to cw, ew has a higher volatility but a better ENCB value (see Section 8.2.1). ENC is much better by design.

Compared to cw, erc has a better performance, lower [max drawdown](#), better [Sharpe ratio](#) and better ENC (see Section 7.1.3). ENCB is much better by design.

Compared to cw, gmv has better performance, lower volatility, half of the max drawdown, more than double the Sharpe ratio, better ENC and ENCB.

For US large cap stocks (1987-2018), gmv has performed the best. This is because of the [bear markets](#) after the tech bubble (2003) and the subprime crisis (2008).

#### 8.4 Video: Lab-session Risk Contributions and Risk Parity

The word 'risk' derives from the early Italian [risicare](#), which means 'to dare'. In this sense, risk is a choice rather than a fate. The actions we dare to take, which depend on how free we are to make choices, are what the story of risk is all about.

*Peter L. Bernstein, Against the Gods: The Remarkable Story of Risk*