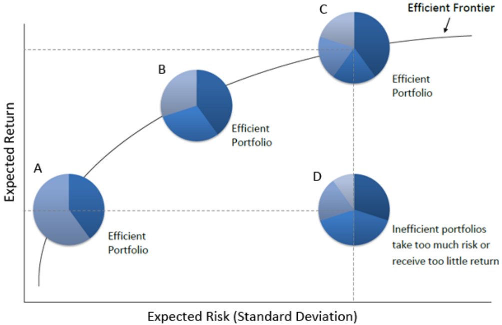
DATA SCIENCE
The Markowitz's model

Antonio VIOLI

antonio.violi@unirc.it

The efficient frontier

- How can we determine the efficient frontier?
- By using the mathematical programming tools!



The Markotwiz's model

- Reward and risk (measured by variance) are the two criteria that should be considered in defining efficient portfolios
- For a given level of expected return, determine the portfolio with minimum variance
- For a given level of risk, determine the portfolio with maximum expected return

H. Markowitz 1990 Nobel Memorial Prize in Economic Sciences

The Markotwiz's model

- Let N denote the universe of assets
- For each asset $i = \ldots, N$ let
  - $\bar{r}_i$ the expected rate of return of asset i
  - σ_{ij} the covariance of the return of asset i and j
- We denote by x_i the decision variable associated with the percentage of money invested in asset i
- We assume that the decision variables are non negative (no short selling are allowed)

The Markotwiz's model

- Let γ be a threshold on the expected portfolio return
- The Markowitz's model is as follows

$$
\begin{aligned}
\min z \quad &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij} x_i x_j \\
&\sum_{i=1}^{N} \bar{r}_i x_i \geq \gamma \\
&\sum_{i=1}^{N} x_i = 1 \\
&x_i \geq 0 \qquad i = 1, \ldots, N
\end{aligned}
$$

- The Markowitz's model belongs to the class of nonlinear (quadratic) programming problem
- By solving several instances of the Markowitz's model associated with different γ values we obtain different points in the $(\bar{r} - \sigma)$ plane
- By taking the convex combinations of these points, we get the efficient frontier

The Markotwiz's model: Compact formulation

We denote by

- $\bar{\mathbf{r}}$ the vector of expected returns
- $\mathbf{e}$ the vector of 1
- $\mathbf{x}$ the vector of decision variables
- V the variance-covariance matrix

$$
\begin{aligned}
\min z \quad & = \mathbf{x}^T V \mathbf{x} \\
& \bar{\mathbf{r}}^T \mathbf{x} \geq \gamma \\
& \mathbf{e}^T \mathbf{x} = 1 \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned}
$$

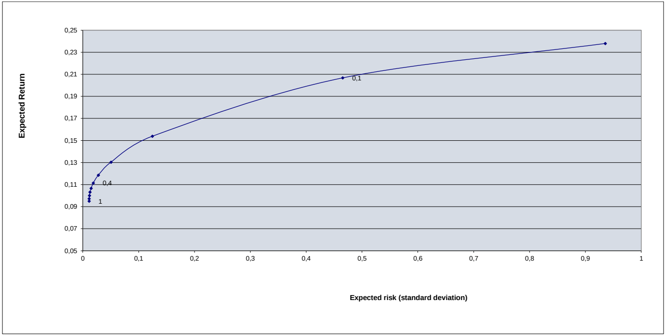
The mean-risk formulation

- Rather than considering different γ values, the efficient frontier can be determined by adopting a bi-objective approach
- Let λ denote a given scalar taking values in the interval [0,1]
- The mean-variance formulation is as follows

$$
\begin{aligned}
\min z \quad & = \lambda * (\mathbf{x}^T V \mathbf{x}) - (1-\lambda) * (\bar{\mathbf{r}}^T \mathbf{x}) \\
& \mathbf{e}^T \mathbf{x} = 1 \\
& \mathbf{x} \geq \mathbf{0}
\end{aligned}
$$

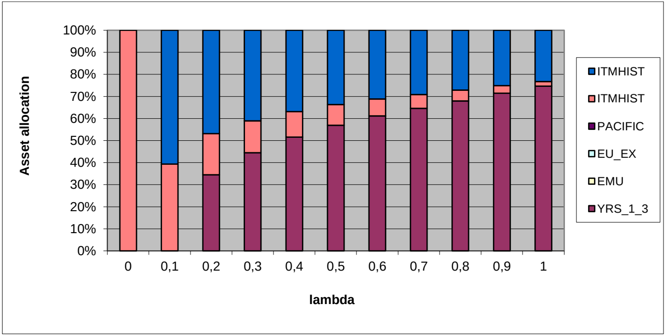
- The parameter λ measures the risk aversion attitude
  - λ values close to 1 are used to model a high risk aversion level
  - λ values close to 0 denote a risk neutrality position

The efficient frontier



The optimal allocation

- The higher λ, the more diversified the portfolio



Limitations of Markowitz Portfolio Theory

The basic Markowitz model presents several limitations(drawback)

- Data estimation
- Lack of real features
- Inadequacy of variance as risk measure
- Single–period perspective

Data estimation

- The information required by the Markowitz model grows substantially as the number N of assets incresaes
- For example, for a portfolio of small size (N = 50), it is necessary to calculate:
  1. N = 50 estimates of expected returns
  2. N = 50 estimates of the variance
  3. $$\frac{N * (N-1)}{2} = 1255$$
     estimates of covariance
- 1325 total estimates

Factor models

- The randomness displayed by the returns of N assets often can be traced back to a smaller number of underlying basic sources of randomness (termed factor)
- For common stocks, the main factor is a proper market index
- If we denote by $\widetilde{r}_m$ the random variable associated with the index return, then we assume that

$$\widetilde{r}_i = f(\widetilde{r}_m)$$

Single-Factor model

- We assume that the rates of return and the factor are related by the following equation:

$$\widetilde{r_i} = \alpha_i + \beta_i * \widetilde{r_m} + e_i$$

  - α_i and β_i are fixed constants
  - e_i are random quantities which represent errors

Single-Factor model: assumptions

We shall assume that

- The errors have zero mean

$$\mathbb{E}[e_i] = 0$$

- The errors are uncorrelated with the market index

$$\mathbb{E}[e_i(\widetilde{r}_m - \overline{r}_m)] = 0$$

- The errors are uncorrelated

$$\mathbb{E}[e_i e_j] = 0 \quad j \neq i$$

Single-Factor model

The parameters for the mean-variance analysis are

$$\bar{r}_i = \alpha_i + \beta_i \bar{r}_m$$

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{e_i}^2$$

$$\sigma_{ij} = \beta_i \beta_j \sigma_m^2$$

Single-Factor model

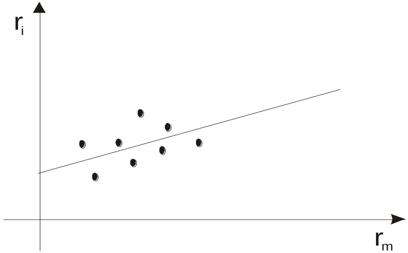
These equations reveal the primary advantage of a factor model

- For a portfolio of small size (N = 50) we have to calculate:
  - 1 estimate of the expected rate of the market index return
  - 1 estimate of the variance of the return of the market index
  - 50 estimates of the α, β and the variance of the errors
- 152 total estimates
- In general for N assets

$$3N + 2$$

Parameter estimation

- The values of β_i are estimated using the method of least squares
- Given a set of observations, represented as points in the plane (r_i, r_m), we look for the "best" straight line that interpolates them
- "Better" in the sense of minimizing the mean square error between the observed data and the estimated



Parameter estimation

- Let us denote by M the number of historical observations (r_{ik}, r_{mk}) with $k = 1, \dots, M$

$$e_{ik} = r_{ik} - \alpha_i - \beta_i r_{mk}$$

- We look for the α_i and β_i values that minimize

$$\frac{1}{M} \sum_{k=1}^{M} e_{ik}^2$$

- By solving

$$\beta_i = \frac{\sigma_{im}}{\sigma_m^2}$$

$$\alpha_i = \bar{r}_i - \beta_i \bar{r}_m$$

Single-factor model

- Portfolio variance

$$\sigma_p^2 = \sum_{i=1}^{N} x_i^2 \beta_i^2 \sigma_m^2 + \sum_{i=1}^{N} x_i^2 \sigma_{e_i}^2 + \sum_{i=1}^{N} \sum_{j=1 \; j \neq i}^{N} x_i x_j \beta_i \beta_j \sigma_m^2$$

- Expected return of portfolio

$$\sum_{i=1}^{N} x_i \alpha_i + \sum_{i=1}^{N} x_i \beta_i \bar{r}_m \geq \gamma$$

Diversification limits and transactions costs

- A general problem with the mean-variance model for risk-averse investors is that the optimal portfolios often contain a large number of individual stocks, making them very expensive to purchase and to manage
- A natural question is how to limit the number of stocks in the portfolio
- There are at least two obvious ways to do so
  - Simply limit the number of stocks present in the portfolio.
  - Explicitly modeling transactions costs subtracting them out of the expected returns

Diversification limits

- Let y_i denote a binary variables taking the value 1 if asset i is included on the portfolio and 0 otherwise
- We assume that here are lower l_i and upper u_i limitations on the fraction of capital invested capital in each asset

$$x_i \leq u_i y_i$$

$$x_i \geq l_i y_i$$

- We may want to impose that the number of assets in our portfolio is at most K

$$\sum_{i=1}^{N} y_i \leq K$$

Modeling transaction costs

- Typically financial institutions may charge a fixed fee each time they execute a transaction involving a particular type of security.
- Such a fee may have a limiting effect on the number of different securities in the optimal portfolio, and is significant enough in most real-world applications to be considered as a necessary part of the model.
- To model transaction cost we introduce a binary variable y_i taking the value 1 if asset i is included on the portfolio and 0 otherwise

Modeling transaction costs

- Transaction cost determine a reduction of the portfolio expected return

$$\sum_{i=1}^{N} \bar{r}_i x_i - \sum_{i=1}^{N} f * y_i$$

- Note that variables x_i should be related to the y_i

$$x_i \leq u_i y_i$$

$$x_i \geq l_i y_i$$

l_i and u_i may be eventually set to 0 and 1, respectively

Remarks

- A further problem with the traditional Markowitz model is the return distribution assumption on which it is implicitly based
- In the model, only the first two moments of the distribution, i.e. the expected value and the variance, are considered while higher moments such as skewness and kurtosis are ignored.
- Traditional risk/return optimization therefore assumes that returns follow a normal distribution.
- Empirical studies demonstrate, however, that normal distribution, especially during periods of market turbulence, significantly underestimates the likelihood of strongly negative returns.
- Empirical return distributions are not symmetric and exhibit fat tails.
- One way of resolving the problem is to incorporate forward-looking performance scenarios

Scenario-based asset allocation

The scenario-based approach poses a number of important issues

- **How realistic scenarios could be generated ?**
  As we shall see in the course, we can adopt simulation techniques
- **Once generated, how scenarios can be included in the mathematical models?**
  We have to introduce some sophisticated modeling paradigms

The scenario–based approach will allow us to elaborate on different risk measures, overcoming the intrinsic limitations of the variance