Estimation of the Mean–Variance Portfolio Model

- In the mean–variance framework, the optimal portfolio weight vector, $x^*$, is a function of the investor’s preference parameters, $c$, and the first two moments of the return distribution, i.e., the mean vector $\mu$ and the covariance matrix $\Sigma$, that is,

$$x^* = x(c, \mu, \Sigma).$$

- As $\mu$ and $\Sigma$ are not directly observable, they have to be estimated, using either subjective judgement (e.g., based on fundamental information) or historical data (i.e., statistical methods).
• If we plug-in the estimates of means and covariances, $\hat{\mu}$ and $\hat{\Sigma}$, into our solution for the optimal weight (1), we get estimates of the optimal portfolio weights, $\hat{x}^*$,

$$\hat{x}^* = x(c, \hat{\Sigma}, \hat{\mu}).$$

• The estimation error of the parameter estimates is transferred to the portfolio weights.

• Therefore, estimated optimal weights are almost certainly different from the true optimal weights.
Sample Moments

- The most straightforward method to estimate $\mu$ and $\Sigma$ is simply to use the sample mean and the sample covariance matrix,

$$
\hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} R_t, \quad \hat{\Sigma} = \frac{1}{T - 1} \sum_{t=1}^{T} (R_t - \hat{\mu})(R_t - \hat{\mu})',
$$

where $T$ is the number of (usually monthly) return observations for each asset.

- These estimators are rarely used in practice.

- Portfolio weights calculated on the basis of the sample moments have been documented to be quite imprecise for realistic sample sizes.

- This is the case in particular if the number of assets under consideration is large, relative to the number of historical return observations (a common situation in practice).
Concerning the volatility part, as the number of assets increases, the number of distinct elements of the covariance matrix increases at a quadratic rate, as it has $N(N + 1)/2$ independent elements (apart from the requirement of positive definiteness).

For example, with 500 assets, the covariance matrix involves 125250 unique elements.

As optimal portfolio weights depend on all these parameters, the portfolio weights are estimated with a lot of error.
Moreover, the portfolio weights tend to be rather sensitive to the inputs (expected returns and variances/covariances).

That is, small changes in the inputs may produce relatively large changes in the estimated optimal portfolio weights.

Unfortunately, the most extreme (and thus “influential”) estimated coefficients are likely to be extreme simply because they have large estimation errors.

Extreme (“unrealistic”) and unstable portfolio weights may be the result.
• A potential remedy is to impose some structure on the covariance matrix to reduce the large fraction of random noise in the sample moments.

• This gives rise to a more parsimonious parameterization of the covariance matrix which, hopefully, helps to efficiently filter out the systematic information from historical correlations, and thus to reduce estimation error and improve forecasting performance.
• Candidate methods for improved (statistical) estimation of risk premiums include, for example,
  – shrinkage estimation,
  – incorporation of economic models (as in the approach of Black and Littermann).

• To structure covariance matrices, factor models (or index models) are frequently used.

• The factors may be observable
  – observable (linear regression)
  – extracted directly from the returns using a statistical technique such as principal component analysis (PCA)
Simulation Experiment (similar to Jobson and Korkie (1980))

• Consider monthly returns of 24 stocks in the DAX over the period 1996–2001.

• Compute the sample mean and covariance matrix from these returns and take these estimates as the “true” values (to have a “realistic” example).

• Then simulate from the “true” return distribution, assuming iid normality.

• For each simulated sample, compute the sample mean and the sample covariance matrix, and compute the tangency portfolio using these quantities (assuming \( r_f = 0 \)).

• Then, using these estimated tangency portfolios, we compute the associated Sharpe Ratio implied by this weight vector and the true moments.
Realized Sharpe Ratios are on average considerably lower than the corresponding measure for the true tangency portfolio, indicating substantial economic loss due to estimation error.
$T = 36$, i.e. 3 years of monthly data

$T = 60$, i.e. 5 years of monthly data

$T = 120$, i.e. 10 years of monthly data
The first and simplest index model is the single-index model (SIM) proposed by Sharpe (1963).

This is based on the observation that, when the markets go up (bull markets), most stocks tend to increase in price, and when markets go down (bear markets), most stocks tend to decrease in price.

Thus, there seems to be a strong common (market) factor driving the joint movement of asset returns.

That is, one important reason why asset returns are correlated is because of a common response to market changes.

In the SIM, it is assumed that the only reason why stocks vary together, systematically, is because of their common comovement with the market.
The return of asset \( i, i = 1, \ldots, N \), is described by
\[
R_i = \alpha_i + \beta_i R_M + \epsilon_i, \quad i = 1, \ldots, N \tag{2}
\]
\[
E(\epsilon_i) = 0, \quad \text{Var}(\epsilon_i) = \sigma_{\epsilon_i}^2, \quad i = 1, \ldots \tag{3}
\]
\[
\text{Cov}(\epsilon_i, \epsilon_j) = 0, \quad i \neq j. \tag{4}
\]

Expected return and variance of the market return will be denoted by \( E(R_M) = \mu_M \) and \( \text{Var}(R_M) = \sigma_M^2 \), and we assume
\[
\text{Cov}(R_M, \epsilon_i) = 0, \quad i = 1, \ldots, N. \tag{5}
\]

This structure implies that
\[
E(R_i) = \alpha_i + \beta_i \mu_M, \quad i = 1, \ldots, N, \tag{6}
\]
\[
\text{Var}(R_i) = \beta_i^2 \sigma_M^2 + \sigma_{\epsilon_i}^2, \quad i = 1, \ldots, N, \tag{7}
\]
\[
\text{Cov}(R_i, r_j) = \beta_i \beta_j \sigma_M^2, \quad i, j = 1, \ldots, N, \quad i \neq j. \tag{8}
\]
• For the covariance structure of the returns, given by (8), Assumption (4) is crucial, as it implies that the only reason for asset $i$ and asset $j$ moving together is their joint dependence on the market return $R_{Mt}$.

• The first part of (7) is also often referred to as the *systematic* risk (which is related to the general tendency of the market), whereas the second part is the *unsystematic* (idiosyncratic, specific) risk, which is not related to market factors.
• In contrast to the market–related, systematic risk, the specific risk can be diversified away.

• Consider an equally, weighted portfolio, i.e., a portfolio with weights $x_i = 1/N$, $i = 1, \ldots, N$.

• Then the portfolio variance is, assuming the SIM correctly describes the covariance structure,

$$
\sigma_p^2 = \frac{1}{N^2} \sum_{i=1}^{N} (\beta_i^2 \sigma_M^2 + \sigma_{\epsilon_i}^2) + \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j \neq i} \beta_i \beta_j \sigma_M^2
$$

$$
= \left( \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_i \beta_j \right) \sigma_M^2 + \frac{1}{N^2} \sum_{i=1}^{N} \sigma_{\epsilon_i}^2
$$

$$
= \left( \frac{1}{N} \sum_{i=1}^{N} \beta_i \right)^2 \sigma_M^2 + \frac{1}{N^2} \sum_{i=1}^{N} \sigma_{\epsilon_i}^2.
$$
• Now

\[ \frac{1}{N^2} \sum_{i=1}^{N} \sigma_{\epsilon_i}^2 \leq \frac{\max\{\sigma_{\epsilon_1}^2, \ldots, \sigma_{\epsilon_N}^2\}}{N} \xrightarrow{N \to \infty} 0, \]

provided the variances of the unsystematic risks are bounded.

• Hence, for large \( N \),

\[ \sigma_{pt}^2 \approx \left( \frac{1}{N} \sum_{i=1}^{N} \beta_i \right)^2 \sigma_M^2 = \overline{\beta}_p^2 \sigma_M^2, \]

where

\[ \overline{\beta}_p = \frac{1}{N} \sum_{i=1}^{N} \beta_i \]

is the portfolio’s \( \beta \).

• That is, the market risk cannot be diversified away.
Multi-Index Models

- The single–index model is straightforwardly extended to $k$ factors.

- In this case, with $k$ factors $f_i$, $i = 1, \ldots, k$,

$$R_i = \alpha_i + \sum_{j=1}^{k} \beta_{ij} f_i + \epsilon_i, \quad i = 1, \ldots, N,$$

or, in obvious notation,

$$r = \alpha + B f + \epsilon.$$

The covariance matrix of returns is

$$\Sigma = B \Sigma_f B' + \Sigma_\epsilon,$$

where $\Sigma_f$ is the covariance matrix of the factors, which may or may not be diagonal.
For example, in industry index models, we start with the basic SIM and add industry indices to capture industry effects.

E.g., two steel stocks will most likely have positive correlation between their returns, even after the effects of the market have been removed.

To avoid noisy estimates, it may be reasonable that returns of each firm are affected only by the market and one industry factor (in contrast to all industry factors).

For firm $i$ in industry $j$, we then have

$$R_i = \alpha_i + \beta_{im}I_m + \beta_{ij}I_j + \epsilon_i,$$

where $I_m$ and $I_j$ are the market and industry factors, respectively.
• Assuming for the moment that the factors are uncorrelated ($\Sigma_f$ is diagonal), the covariance between securities $i$ and $\ell$ is then

$$\sigma_{i\ell} = \beta_{im}\beta_{\ell m}\sigma_m^2 + \beta_{ij}\beta_{\ell j}\sigma_I^2_i$$

for firms in the same industry $j$, and

$$\sigma_{i\ell} = \beta_{im}\beta_{\ell m}\sigma_m^2$$

for firms in different industries.

• macroeconomic variables, country–specific variables
• In contrast to the sample covariance matrix, the matrix provided by an index model will usually be nonsingular.

• For example, when $N = 1000$, we would need more than 1000 observations on past return vectors to get a nonsingular sample covariance matrix.

• So data requirements are also reduced when index models are used.
Common Correlation Models (CCMs)

- Such models have been proposed by Elton and Gruber (1973).

- They are extremely simple but tend to outperform the historical correlation matrix.

- The idea is that “historical data only contain information concerning the mean correlation coefficients and that observed pairwise differences from the average are random or sufficiently unstable“ (EG, 1973).

- Thus, the best way to forecast future correlations is to use the average of the observed historical sample correlation coefficients.

- In the simplest case, this reduces the number of parameters to be estimated for the correlation matrix from $N(N-1)/2$ to one.
More Sophisticated CCMs

- The idea of a common correlation between all pairs of assets may be too restrictive, and grouping techniques will be more appropriate (e.g., common correlation among stocks from the same industry).

- For example, if there were two industries, steels and chemicals, we may assume that the correlation between all steels is constant ($\rho_{SS}$) and the correlation between all chemicals is constant ($\rho_{CC}$), but potentially different from $\rho_{SS}$.

- Also, the correlation between members of the steel group and those of the chemical group is constant ($\rho_{CS}$).
In international portfolio diversification, the *National Mean Model* has been shown to be potentially useful, where every intra–country pairwise correlation is constant, and every correlation between assets of two countries, say $A$ and $B$, is equal to $\rho_{AB}$.

This was first observed by Eun and Resnick (1984).