Correlation Structure of Stock Returns

Joel R. Barber
Portfolio Expected Return and Risk

1. $X_1, X_2, ..., X_N$ portfolio weights on assets 1, 2, ..., $N$
2. $R_i$ rate of return on asset $i = 1, 2, ..., N$
3. $\sigma_{ij} = \text{cov}(R_i, R_j)$
4. $\sigma_{ii} = \sigma_i^2 = \text{var}(R_i)$
5. $E[R_i]$ is the expected return on asset $i$
6. expected portfolio return:

$$E[R_p] = \sum_{i=1}^{N} X_i E[R_i]$$

7. portfolio variance:

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} X_i X_j \sigma_{ij}$$
covariance Matrix

\[ V = \begin{bmatrix}
\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\
\sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN}
\end{bmatrix} \]

problem

\[
\begin{align*}
\min_X & \quad \sigma^2_p(X_1, \ldots, X_N) \\
\text{st} & \quad \sum X_i E[R_i] = K \\
& \quad \sum X_i = 1 \\
& \quad X_1, \ldots, X_N \geq 0
\end{align*}
\]
Number of Inputs

1. $N$ expected returns
2. $N$ variances
3. $(N^2 - N)/2$ covariances
4. Suppose $N = 100$
5. 100 expected returns
6. 100 variances
7. 4950 covariances
8. Must estimate 5150 parameters
9. Number of parameters increases with $N^2$
Correlation Structure of Stock Returns

1. empirical observation: stocks tend to move up and down together
2. some portion of the common movement can be explained by factors common to all stocks
3. the remainder is unexplained
4. Stock returns can be divided into a systematic and unsystematic components
Assume systematic component is explained by a linear combination of $K \ll N$ factors:

$$b_1 f_1 + b_2 f_2 + ... b_K f_K$$

where $b_1, ..., b_K$ are constants and $f_1, ..., f_K$ are random variables.

Note factor loadings $b_1, ..., b_K$ are different for each stock and indicate the sensitivity of return to a given systematic risk factor.

Define unsystematic component as what is left:

$$\varepsilon = R - (b_1 f_1 + b_2 f_2 + ... b_K f_K)$$

Assume unsystematic component of any two stocks $i$ and $j$ is uncorrelated: $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$.

Then covariance matrix can be determined by $N \times K$ factor loadings.

Suppose $N = 100$ and $K = 2$. Then you need to estimate 200 factor loadings: two coefficients for each of a 100 stocks.
Market Model

1. used to analyze market risk
2. technique allows us to measure the market and unique risk of a stock
3. not a theoretical model
4. specification

\[ R = \alpha + \beta R_m + \varepsilon, \]

where

\[ \beta = \frac{\text{cov}(R_m, R)}{\sigma_m^2} \]

and

\[ \alpha = E[R] - \beta E[R_m] \]
Properties of residual

1. **Assumptions**
   1. $E[\varepsilon] = 0$
   2. $\text{cov}(R_m, \varepsilon) = 0$

2. **Results**
   1. $E[R] = \alpha + \beta E[R_m]$
   2. $\sigma^2 = \beta^2 \sigma^2_m + \text{var}(\varepsilon)$
   3. R-squared $= \frac{\beta^2 \sigma^2_m}{\sigma^2}$

3. **The total variance equals the systematic plus unsystematic variance:**
   
   $$TV = \beta^2 \sigma^2_m + \text{var}(\varepsilon)$$
Example

1. \( \text{cov}(R, R_m) = 24\%^2, \sigma_m = 4\%, \sigma = 10\%, E[R] = 12\%, \text{ and } E[R_m] = 10\% \)

2. \( \beta = 24/4^2 = 1.5 \)

3. \( \alpha = 12 - 1.5 \times 10 = -3\% \)

4. \( \text{systematic variance} = 1.5^2 \times 4^2 = 36\%^2 \)

5. \( \text{systematic standard deviation} = 6\% \)

6. \( \text{unsystematic variance} = 10^2 - 36 = 64\%^2 \)

7. \( \text{unsystematic standard deviation} = 8\% \)

8. \( \text{R-squared} = 36/100 = .36 \)

9. \( \text{corr}(R, R_m) = .6 \)
Does security fall on security market line (SML)?

1. The answer depends upon the risk-free rate $R_f$
2. Falls on SML if $\alpha = (1 - \beta)R_f$
3. So if beta is greater than one alpha is negative
4. Above if $\alpha > (1 - \beta)R_f$
5. In above example, with $\alpha = -3$ and $\beta = 1.5$, if $R_f = 6\%$
   1. $(1 - \beta)R_f = (1 - 1.5)6 = -3$
   2. So the expected return falls on security market line:

$$E[R] = 6 + 1.5(10 - 6) = 12\%$$
Assume the market model holds for each stock and for each pair of stocks the residuals are uncorrelated: \( \text{cov}(\varepsilon_i, \varepsilon_j) = 0 \).

Then for \( i \neq j \),

\[
\text{cov}(R_i, R_j) = \beta_i \beta_j \sigma_m^2
\]

What about \( i = j \)?

\[
\sigma_i^2 = \beta_i^2 \sigma_m^2 + \text{var}(\varepsilon_i)
\]

Now the covariance matrix can be determined by \( N \) betas and \( N \) total variances.
1. Return on a portfolio

\[ R_p = \alpha_p + \beta_p R_m + \varepsilon_p \]

where portfolio alpha, beta, and residual are the weighted average of the market model values.

2. Expected return on a portfolio

\[ E[R_p] = \alpha_p + \beta_p E[R_m] \]

where portfolio alpha and beta are the weighted average of the component alphas and betas.

3. Variance of a portfolio

\[ \sigma^2_p = \beta_p^2 \sigma^2_m + \text{var}(\varepsilon_p) \]

4. Unsystematic portfolio variance

\[ \text{var}(\varepsilon_p) = \sum_{i=1}^{N} X_i^2 \text{var}(\varepsilon_i) \]
$100$ invested in GM with a $\beta = 1.5$, $\sigma = 10\%$ and $300$ invested in FPL with a $\beta = .75$ and $\sigma = 6\%$.

GM: $\text{var}(\varepsilon) = 64$ (from last example)

FPL: $\text{var}(\varepsilon) = 36 - (3/4)^2 \times 16 = 27$

$\text{var}(\varepsilon_p) = (1/4)^2(64) + (3/4)^2(27) = \frac{307}{16} = 19.19$

portfolio beta = $(1/4)(1.5) + (3/4)(.75) = .9375$

systematic variance = $.9375^2 \times 16 = 14.06$

total variance = $14.06 + 19.19 = 33.25$

total standard deviation = $5.67\%$

lower than GM or FPL

R-squared = $14.06/33.25 = 0.42286$