Chapter 6. The Mean-Variance Portfolio Theory
Asset Return

The rate of return on an asset:

\[ R = \frac{X_1}{X_0} \]  or  \[ r = \frac{X_1 - X_0}{X_0} = R - 1 \]

where \( X_0 \) is the amount invested, and \( X_1 \) is the amount received.

*Short-selling* of an asset is to own a negative unit of the asset.

*Example 6.1.* Suppose I decide to short 100 shares of stock in company CBA, which is currently selling for $10 per share. I borrowed 100 shares from my broker and sell these in the stock market, receiving $1,000. At the end of 1 year, the price of CBA has dropped to $9 per share. I buy back 100 shares for $900 and give these shares to my broker to repay the original loan. I made a profit of $100.

The rate of return of CBA is \((90 - 100)/100 = -10\%\), while shorting on CBA has the rate of return of \((-900 - (-1000))/1000 = 10\%\).
A portfolio is a collection of assets. For instance, if the proportion of Asset $i$ over the whole collection is $w_i$ (also known as the weight), then the portfolio is represented by a vector $(w_1, w_2, \ldots, w_n)$.

Return of portfolio

\[
R = \sum_{i=1}^{n} w_i R_i \quad \text{or} \quad r = \sum_{i=1}^{n} w_i r_i.
\]

<table>
<thead>
<tr>
<th>Security</th>
<th># of shares</th>
<th>price</th>
<th>total cost</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jazz, Inc.</td>
<td>100</td>
<td>$40</td>
<td>$4,000</td>
<td>25%</td>
</tr>
<tr>
<td>Classical, Inc.</td>
<td>400</td>
<td>$20</td>
<td>$8,000</td>
<td>50%</td>
</tr>
<tr>
<td>Rock, Inc.</td>
<td>200</td>
<td>$20</td>
<td>$4,000</td>
<td>25%</td>
</tr>
<tr>
<td>Portfolio total:</td>
<td></td>
<td></td>
<td>$16,000</td>
<td>100%</td>
</tr>
</tbody>
</table>
Mean and variance of a random variable $x$:

$$E[x] (= \bar{x}) \text{ and } \text{var}(x) = E[x - E[x]]^2 = E[x^2] - (E[x])^2.$$ 

Variance measures the volatility of a random variable.

Covariance of two random variables $x_1$ and $x_2$:

$$\text{cov}(x_1, x_2) = E[(x_1 - E[x_1])(x_2 - E[x_2])].$$

Covariance matrix of $n$ random variables $x_1, x_2, \ldots, x_n$:

$$\Sigma = (\sigma_{ij})_{n \times n}$$

where $\sigma_{ij} = \text{cov}(x_i, x_j)$.

Note that $\text{cov}(x_i, x_i) = \text{var}(x_i) = \sigma_i^2$. 

Shuzhong Zhang
We have

$$|\sigma_{ij}| \leq \sigma_i \sigma_j.$$ 

Correlation coefficient of $x_i$ and $x_j$:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}.$$ 

Two random variables are called uncorrelated if their correlation coefficient is zero; they are positively correlated if $\rho_{ij} > 0$, and negatively correlated if $\rho_{ij} < 0$. 

Shuzhong Zhang
Example. A roll of the die: The mean is
\[
\frac{1}{6} (1 + 2 + 3 + 4 + 5 + 6) = 3.5,
\]
and the variance is
\[
\sigma^2 = \frac{1}{6} (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - 3.5^2 = 2.92.
\]

Suppose that a die is rolled twice. The first roll is recorded as \(x\), the
second is \(y\) and the average of the two is recorded as \(z\). What is the
correlation between \(x\) and \(z\)?

Note that \(x\) and \(y\) are independent (hence uncorrelated). We have
\(\bar{x} = \bar{y} = \bar{z} = 3.5\), and
\[
\text{var } z = \mathbb{E} \left[ \left( \frac{(x - \bar{x}) + (y - \bar{y})}{2} \right)^2 \right] = \mathbb{E} \left[ \left( \frac{(x - \bar{x})^2 + 2(x - \bar{x})(y - \bar{y}) + (y - \bar{y})^2}{4} \right) \right] = \left( \sigma_x^2 + 2\rho_{xy}\sigma_x\sigma_y + \sigma_y^2 \right) / 4 = \sigma^2 / 2 = 1.46. \]
Moreover,
\[
\rho_{xz} = \frac{\mathbb{E}[(x - \bar{x})(z - \bar{z})]}{\sigma_x \sigma_z} = \frac{\sqrt{2}}{2} = 0.707.
\]
Example 6.7. Betting wheel:

\[
\begin{align*}
\bar{R}_1 &= \frac{1}{2} \times 3 + \frac{1}{2} \times 0 = \frac{3}{2}, \\
\sigma_1^2 &= \frac{1}{2} \times 3^2 - \bar{R}_1^2 = 2.25, \\
\sigma_{12} &= -\bar{R}_1 \bar{R}_2 = -1, \\
\sigma_{13} &= -\bar{R}_1 \bar{R}_3 = -1.5, \\
\bar{R}_2 &= \frac{1}{3} \times 2 + \frac{2}{3} \times 0 = \frac{2}{3}, \\
\sigma_2^2 &= \frac{1}{3} \times 2^2 - \bar{R}_2^2 = 0.889, \\
\sigma_{23} &= -\bar{R}_2 \bar{R}_3 = -0.67, \\
\bar{R}_3 &= \frac{1}{6} \times 6 + \frac{5}{6} \times 0 = 1 \\
\sigma_3^2 &= \frac{1}{6} \times 6^2 - 1 = 5 \\
\end{align*}
\]
Portfolio Mean and Variance

Suppose that there are \( n \) assets with rate of return \( r_1, r_2, \ldots, r_n \). Suppose that one holds a portfolio of these assets, where the weight on asset \( i \) is \( w_i, i = 1, 2, \ldots, n \). Then, the return on the portfolio is

\[
r = \sum_{i=1}^{n} w_i r_i.
\]

Its expected rate of return is

\[
E[r] = \sum_{i=1}^{n} w_i E[r_i].
\]
Its variance is

$$\text{var} \,(r) = E[(r - E[r])^2]$$

$$= E \left[ \left( \sum_{i=1}^{n} w_i (r_i - E[r_i]) \right)^2 \right]$$

$$= E \left[ \sum_{i,j=1}^{n} w_i w_j (r_i - E[r_i])(r_j - E[r_j]) \right]$$

$$= \sum_{i,j=1}^{n} \sigma_{ij} w_i w_j$$

$$= w^T \Sigma w.$$
Variance can be used to model volatility or risk of the portfolio.

Mean can be used to model the gain of the portfolio.

Diversification:

Suppose there are \( n \) uncorrelated assets, each with mean rate of return \( \mathbb{E}[R_i] = \mu \) and the variance \( \text{var}(R_i) = \sigma^2 \), \( i = 1, 2, \ldots, n \). Now consider a portfolio of \( w^T = (1/n, 1/n, \ldots, 1/n) \). Then, the mean of the portfolio remains to be

\[
\mathbb{E} \left[ \sum_{i=1}^{n} R_i w_i \right] = \mu
\]

but the variance is

\[
w^T \Sigma w = \sigma^2 / n.
\]
Efficient portfolio is a portfolio whose variance is minimum for the corresponding mean.

The Markowitz Model

Suppose that the desired level of expected return is \( \mu \):

\[
\min \quad \frac{1}{2} \sum_{i,j=1}^{n} \sigma_{ij} w_i w_j \\
\text{s.t.} \quad \sum_{i=1}^{n} w_i \bar{r}_i = \mu \\
\sum_{i=1}^{n} w_i = 1.
\]
A crash course on optimization

Consider the following model (unconstrained optimization):

\[(P_1) \quad \min \quad f(x) \]
\[\quad \text{s.t.} \quad x \in \mathbb{R}^n.\]

**Theorem of Fermat:** Suppose \(x^*\) is an optimal solution of \((P_1)\). Then \(\nabla f(x^*) = 0\).

**Basic facts:**

- For a quadratic function \(f(x) = x^TQx + b^Tx\), its gradient is \(\nabla f(x) = 2Qx + b\).
- If \(Q\) is positive semidefinite, then \(f(x)\) is a convex function. In the convex case, the optimality condition is **necessary and sufficient**.
Consider the following model (equality constrained optimization):

\[
(P_2) \quad \min \quad f(x) \\
\text{s.t.} \quad h_i(x) = 0, \ i = 1, 2, \ldots, p.
\]

Assume that \( \nabla g_i(x) \)'s are always linearly independent.

*Theorem of Lagrange*: Suppose \( x^* \) is an optimal solution of \((P_2)\). Then there exist \( \lambda_i, \ i = 1, \ldots, p \), such that

\[
\nabla f(x^*) - \sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) = 0.
\]

If \( f \) is convex and \( h_i \)'s are affine linear, then the above optimality condition is *necessary and sufficient*. 
Finally, consider the following (equality and inequality constrained) model:

\[
(P_3) \quad \min \ f(x) \\
\text{s.t.} \quad h_i(x) = 0, \ i = 1, 2, \ldots, p; \\
g_j(x) \leq 0, \ j = 1, 2, \ldots, r.
\]

Suppose the above has an optimal solution \( x^* \). Moreover, let \( I(x) = \{ j \mid g_j(x) = 0 \} \). Under the condition that the vectors \( \{ \nabla h_i(x), \ i = 1, \ldots, p; \ \nabla g_j(x), \ j \in I(x) \} \) are always linearly independent. Then the so-called KKT theorem holds.
Theorem of Karusk, Kuhn, and Tucker: For any optimal solution $x^*$ (for $(P_3)$), there exist $\lambda_i, i = 1, ..., p,$ and $\mu_j \geq 0, j = 1, ..., r,$ such that

$$\nabla f(x^*) - \sum_{i=1}^{p} \lambda_i \nabla h_i(x^*) + \sum_{i=1}^{r} \mu_j \nabla g_j(x^*) = 0,$$

and $\mu_j g_j(x^*) = 0,$ for all $j = 1, 2, ..., r.$

If $f$ and $g_j$’s are convex, and $h_i$’s are affine linear, then the KKT optimality condition is necessary and sufficient.
Now let us return to the Markowitz model.

The solution for the Markowitz model:

\[
\begin{align*}
\sum_{i=1}^{n} \sigma_{ij} w_j - \lambda_1 \bar{r}_i - \lambda_2 &= 0, \quad \text{for } i = 1, 2, \ldots, n \\
\sum_{i=1}^{n} w_i \bar{r}_i &= \mu \\
\sum_{i=1}^{n} w_i &= 1.
\end{align*}
\]

In the vector form, let \( e \) be the all-one vector, we have

\[
w = \lambda_1 \Sigma^{-1} \bar{r} + \lambda_2 \Sigma^{-1} e
\]

where

\[
\lambda_1 = \frac{c \mu - b}{ac - b^2} \quad \text{and} \quad \lambda_2 = \frac{-b \mu + a}{ac - b^2}
\]

with \( a = \bar{r}^T \Sigma^{-1} \bar{r}, \) \( b = \bar{r}^T \Sigma^{-1} e, \) and \( c = e^T \Sigma^{-1} e. \)
The above result leads to the so-called *two-fund* theorem:

No matter what the desired level of expected return $\mu$, the efficient portfolio can also be seen as a combination of two portfolios

$$\frac{1}{b} \Sigma^{-1} \bar{r} \text{ and } \frac{1}{c} \Sigma^{-1} e.$$  

If the expected rate of return is $\mu$, then the holding for the first portfolio is $\frac{bc\mu - b^2}{ac - b^2}$ and the holding for the second portfolio is $\frac{-bc\mu + ac}{ac - b^2}$.  

*Shuzhong Zhang*
If there is a restriction that no short selling is allowed, then the model becomes

$$\min \quad \frac{1}{2} \sum_{i,j=1}^{n} \sigma_{ij} w_i w_j$$

s.t.  
$$\sum_{i=1}^{n} w_i \bar{r}_i = \mu$$
$$\sum_{i=1}^{n} w_i = 1$$
$$w_i \geq 0, \ i = 1, 2, \ldots, n.$$  

Such model is known as *quadratic optimization*, and can be readily solved numerically.

However, the two-fund theorem does not hold in this case.
\textbf{Inclusion of a risk-free asset}

Suppose that there is a risk-free asset with rate of return $r_f$.

Then, the model becomes

$$\min \quad \frac{1}{2} \sum_{i,j=1}^{n} \sigma_{ij} w_i w_j$$

s.t.\hspace{0.5cm} \sum_{i=1}^{n} w_i \bar{r}_i + w_0 r_f = \mu$$

$$\sum_{i=1}^{n} w_i + w_0 = 1.$$

Or,

$$\min \quad \frac{1}{2} \sum_{i,j=1}^{n} \sigma_{ij} w_i w_j$$

s.t.\hspace{0.5cm} \sum_{i=1}^{n} w_i \bar{r}_i + (1 - \sum_{i=1}^{n} w_i) r_f = \mu.$$
Solving the optimality condition gives us the portion on the risky assets to be

$$w = \frac{\mu - r_f}{(\bar{\mu} - r_f e)^T \Sigma^{-1} (\bar{\mu} - r_f e)} \Sigma^{-1} (\bar{\mu} - r_f e).$$

The one-fund theorem:

In the presence of a risk-free asset, the optimal portfolio is a combination of a fixed portfolio of risky assets and the risk-free asset.
The (minimum) variance of the optimal portfolio is

\[
\frac{(\mu - r_f)^2}{(\bar{r} - r_{fe})^T \Sigma^{-1} (\bar{r} - r_{fe})}.
\]

In that case, the efficient frontier in the mean expect return $\mu$ and the minimum standard deviation $\sigma$ will form a linear relationship:

\[
\mu - r_f = \kappa \sigma,
\]

where

\[
\kappa = \sqrt{(\bar{r} - r_{fe})^T \Sigma^{-1} (\bar{r} - r_{fe})}.
\]
Example 6.10. Consider the following data

<table>
<thead>
<tr>
<th>Securities</th>
<th>Covariance</th>
<th>$\bar{r}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.30 0.93 0.62 0.74 -0.23</td>
<td>15.10</td>
</tr>
<tr>
<td>2</td>
<td>0.93 1.40 0.22 0.56 0.26</td>
<td>12.50</td>
</tr>
<tr>
<td>3</td>
<td>0.62 0.22 1.80 0.78 -0.27</td>
<td>14.7</td>
</tr>
<tr>
<td>4</td>
<td>0.74 0.56 0.78 3.40 -0.56</td>
<td>9.02</td>
</tr>
<tr>
<td>5</td>
<td>-0.23 0.26 -0.27 -0.56 2.60</td>
<td>17.68</td>
</tr>
</tbody>
</table>

One easily computes that in this case we have

$$\frac{1}{b} \Sigma^{-1} \bar{r} = [0.1583, 0.1554, 0.3143, 0.0379, 0.3341]^T$$

and

$$\frac{1}{c} \Sigma^{-1} e = [0.0883, 0.2509, 0.2824, 0.1038, 0.2748]^T$$

The two fund theorem asserts that any efficient portfolio will be a combination of these two portfolios.
If the risk-free rate of return is $r_f = 10\%$, then the portfolio on the risky assets part corresponding to the one-fund theorem is

$$\frac{1}{e^T \Sigma^{-1}(\bar{r} - r_f e)} \Sigma^{-1}(\bar{r} - r_f e)$$

$$= [0.3172, -0.0611, 0.3865, -0.1113, 0.4687]^T.$$  

Any other efficient portfolio is a combination of that portfolio in combination with the risk-free asset.