MATH4512 – Fundamentals of Mathematical Finance

Topic Two — Mean variance portfolio theory

2.1 Mean and variance of portfolio return

2.2 Markowitz mean-variance formulation

2.3 Two-fund Theorem

2.4 Inclusion of the risk free asset: One-fund Theorem

2.1 Mean and variance of portfolio return

Single-period investment model – Asset return

Suppose that you purchase an asset at time zero, and 1 year later you sell the asset. The **total return** on your investment is defined to be

$$\text{total return} = \frac{\text{amount received}}{\text{amount invested}}.$$

If X_0 and X_1 are, respectively, the amounts of money invested and received and R is the total return, then

$$R = \frac{X_1}{X_0}.$$

The **rate of return** is defined by

$$r = \frac{\text{amount received} - \text{amount invested}}{\text{amount invested}} = \frac{X_1 - X_0}{X_0}.$$

It is clear that

$$R = 1 + r \quad \text{and} \quad X_1 = (1 + r)X_0.$$

- Amount received X_1 = dividend received during the investment period + terminal asset value.

  Note that both dividends and terminal asset value are uncertain.

- We treat r as a random variable, characterized by its probability distribution. For example, in the discrete case, we have

rate of return	r_1	r_2	$\cdots$	r_n
probability of occurrence	p_1	p_2	$\cdots$	p_n

- Two important statistics (discrete random variable)

$$\text{mean} = \overline{r} = \sum_{i=1}^{n} r_i p_i$$

$$\text{variance} = \sigma^2(r) = \sum_{i=1}^{n} (r_i - \overline{r})^2 p_i.$$

Statement of the problem

- A *portfolio* is defined by allocating fractions of initial wealth to individual assets. The fractions (or weights) must sum to one (some of these weights may be negative, corresponding to short selling).

- Return is quantified by portfolio's expected rate of return;
  Risk is quantified by variance of portfolio's rate of return.

Goal: Maximize return for a given level of risk;
or minimize risk for a given level of return.

(i) How do we determine the optimal portfolio allocation?

(ii) The characterization of the set of optimal portfolios (minimum variance funds and efficient funds).

Limitations in the mean variance portfolio theory

- Only the *mean* and *variance* of rates of returns are taken into consideration in the mean-variance portfolio analysis. The higher order moments (like the skewness) of the probability distribution of the rates of return are irrelevant in the formulation.

- Indeed, only the Gaussian (normal) distribution is fully specified by its mean and variance. Unfortunately, the rates of return of risky assets are not Gaussian in general.

- *Calibration* of parameters in the model is always challenging.
  – Sample mean: $\widehat{r} = \frac{1}{n}\sum_{t=1}^{n} r_t$.
  – Sample variance $\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{t=1}^{n}(r_t - \widehat{r})^2$,
  where r_t is the historical rate of return observed at time $t, t = 1, 2, \cdots, n$.

Short sales

- It is possible to sell an asset that you do not own through the process of **short selling**, or **shorting**, the asset. You then sell the borrowed asset to someone else, receiving an amount X_0. At a later date, you repay your loan by purchasing the asset for, say, X_1 and return the asset to your lender. Short selling is profitable if the asset price declines.

- When short selling a stock, you are essentially duplicating the role of the issuing corporation. You sell the stock to raise immediate capital. If the stock pays dividends during the period that you have borrowed it, you too must pay that same dividend to the person from whom you borrowed the stock.

Return associated with short selling

We *receive* X_0 initially and *pay* X_1 later, so the outlay (支出) is $-X_0$ and the final receipt (進款，收入) is $-X_1$, and hence the total return is

$$R = \frac{-X_1}{-X_0} = \frac{X_1}{X_0}.$$

The minus signs cancel out, so we obtain the same expression as that for purchasing the asset. The return value R applies algebraically to both purchases and short sales.

We can write

$$-X_1 = -X_0 R = -X_0(1 + r)$$

to show how the final receipt is related to the initial outlay.

Example of short selling transaction

Suppose I short 100 shares of stock in company CBA. This stock is currently selling for \$10 per share. I borrow 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. Because the stock price fell, this has been a favorable transaction for me. I made a profit of \$100.

The rate of return is clearly negative as $r = -10\%$.

Shorting converts a negative rate of return into a profit because the original investment is also negative.

Portfolio weights

Suppose now that n different assets are available. We form a **portfolio** of these n assets. Suppose that this is done by apportioning an amount X_0 among the n assets. We then select amounts $X_{0i}, i = 1, 2, \cdots, n$, such that $\sum_{i=1}^{n} X_{0i} = X_0$, where X_{0i} represents the amount invested in the i^{th} asset. If we are allowed to sell an asset short, then some of the X_{0i}'s can be negative.

We write

$$X_{0i} = w_i X_0, \quad i = 1, 2, \cdots, n$$

where w_i is the **weight** of asset i in the portfolio. Clearly,

$$\sum_{i=1}^{n} w_i = 1$$

and some w_i's may be negative if short selling is allowed.

Portfolio return

Let R_i denote the total return of asset i. Then the amount of money generated at the end of the period by the i^{th} asset is $R_i X_{0i} = R_i w_i X_0$.

The total amount received by this portfolio at the end of the period is therefore $\sum_{i=1}^{n} R_i w_i X_0$. The overall total return of the portfolio is

$$R_P = \frac{\sum_{i=1}^{n} R_i w_i X_0}{X_0} = \sum_{i=1}^{n} w_i R_i.$$

Since $\sum_{i=1}^{n} w_i = 1$, we have

$$r_P = R_P - 1 = \sum_{i=1}^{n} w_i (R_i - 1) = \sum_{i=1}^{n} w_i r_i.$$

Covariance of a pair of random variables

When considering two or more random variables, their mutual dependence can be summarized by their **covariance**.

Let x_1 and x_2 be a pair random variables with expected values $\overline{x}_1$ and $\overline{x}_2$, respectively. The covariance of this pair of random variables is defined to be the expectation of the product of deviations from the respective mean of x_1 and x_2:

$$\text{cov}(x_1, x_2) = E[(x_1 - \overline{x}_1)(x_2 - \overline{x}_2)].$$

The covariance of two random variables x and y is denoted by σ_{xy}. We write $\text{cov}(x_1, x_2) = \sigma_{12}$. By symmetry, $\sigma_{12} = \sigma_{21}$, where

$$\sigma_{12} = E[x_1 x_2 - \overline{x}_1 x_2 - x_1 \overline{x}_2 + \overline{x}_1 \overline{x}_2] = E[x_1 x_2] - \overline{x}_1 \overline{x}_2.$$

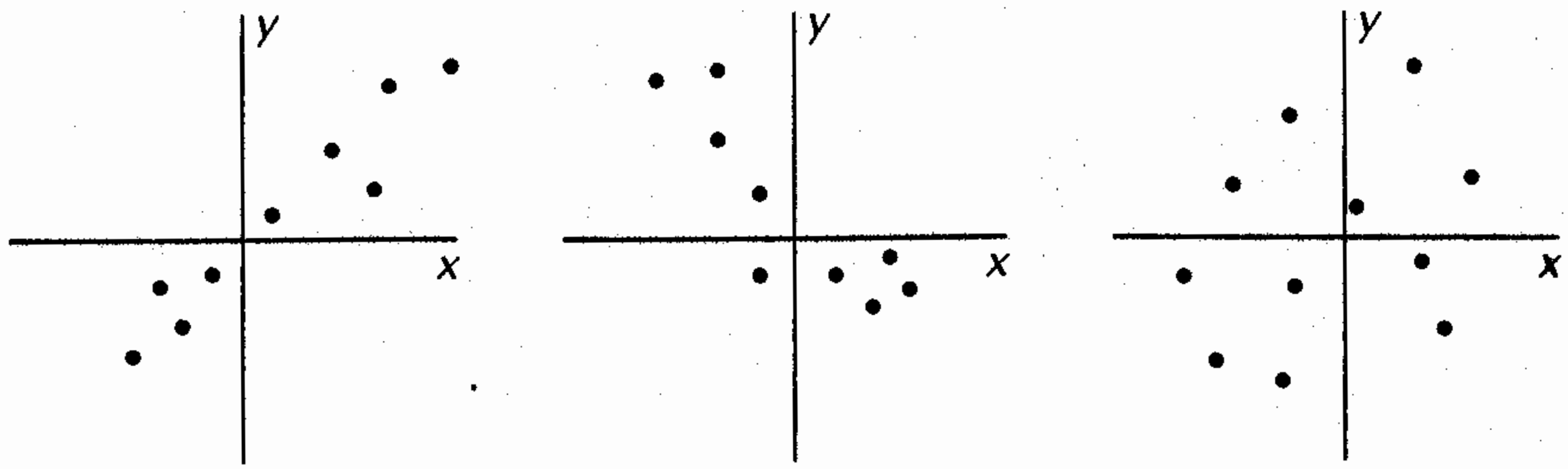
Correlation

- If the two random variables x_1 and x_2 have the property that $\sigma_{12} = 0$, then they are said to be **uncorrelated**.

- If the two random variables are independent, then they are uncorrelated. When x_1 and x_2 are independent, $E[x_1 x_2] = \overline{x}_1 \overline{x}_2$ so that $\text{cov}(x_1, x_2) = 0$.

- If $\sigma_{12} > 0$, then the two variables are said to be **positively correlated**. In this case, if one variable is above its mean, the other is likely to be above its mean as well.

- On the other hand, if $\sigma_{12} < 0$, the two variables are said to be **negatively correlated**.



(a) Positively correlated (b) Negatively correlated (c) Uncorrelated

When x_1 and x_2 are positively correlated, a positive deviation from mean of one random variable has a higher tendency to have a positive deviation from mean of the other random variable.

The **correlation coefficient** of a pair of random variables is defined as

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2}.$$

It can be shown that $|\rho_{12}| \leq 1$.

This would imply that the covariance of two random variables satisfies

$$|\sigma_{12}| \leq \sigma_1 \sigma_2.$$

If $\sigma_{12} = \sigma_1 \sigma_2$, the variables are **perfectly correlated**. In this situation, the covariance is as large as possible for the given variance. If one random variable were a fixed positive multiple of the other, the two would be perfectly correlated.

Conversely, if $\sigma_{12} = -\sigma_1 \sigma_2$, the two variables exhibit **perfect negative correlation**.

Mean rate of return of a portfolio

Suppose that there are N assets with (random) rates of return $r_1, r_2, \cdots, r_N$, and their expected values $E[r_1] = \overline{r}_1, E[r_2] = \overline{r}_2, \cdots$, $E[r_N] = \overline{r}_N$. The rate of return of the portfolio in terms of the rate of return of the individual assets is

$$r_P = w_1 r_1 + w_2 r_2 + \cdots + w_n r_N,$$

so that

$$\begin{aligned} E[r_P] &= \overline{r}_P = w_1 E[r_1] + w_2 E[r_2] + \cdots + w_n E[r_N] \\ &= w_1 \overline{r}_1 + w_2 \overline{r}_2 + \cdots + w_n \overline{r}_N. \end{aligned}$$

The portfolio's mean rate of return is simply the weighted average of the mean rates of return of the assets. Note that a negative rate of return r_i of asset i with negative weight w_i (short selling) contributes positively to $\overline{r}_P$.

Variance of portfolio's rate of return

We denote the variance of the return of asset i by σ_i^2, the variance of the return of the portfolio by σ_P^2, and the covariance of the return of asset i with that of asset j by σ_{ij}. Portfolio variance is given by

$$
\begin{aligned}
\sigma_P^2 &= E[(r_P - \overline{r}_P)^2] \\
&= E\left[\left(\sum_{i=1}^N w_i r_i - \sum_{i=1}^N w_i \overline{r}_i\right)^2\right] \\
&= E\left[\left(\sum_{i=1}^N w_i (r_i - \overline{r}_i)\right)\left(\sum_{j=1}^N w_j (r_j - \overline{r}_j)\right)\right] \\
&= E\left[\sum_{i=1}^N \sum_{j=1}^N w_i w_j (r_i - \overline{r}_i)(r_j - \overline{r}_j)\right] \\
&= \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}.
\end{aligned}
$$

Zero correlation

Suppose that a portfolio is constructed by taking equal portions of N of these assets; that is, $w_i = \frac{1}{N}$ for each i. The overall rate of return of this portfolio is

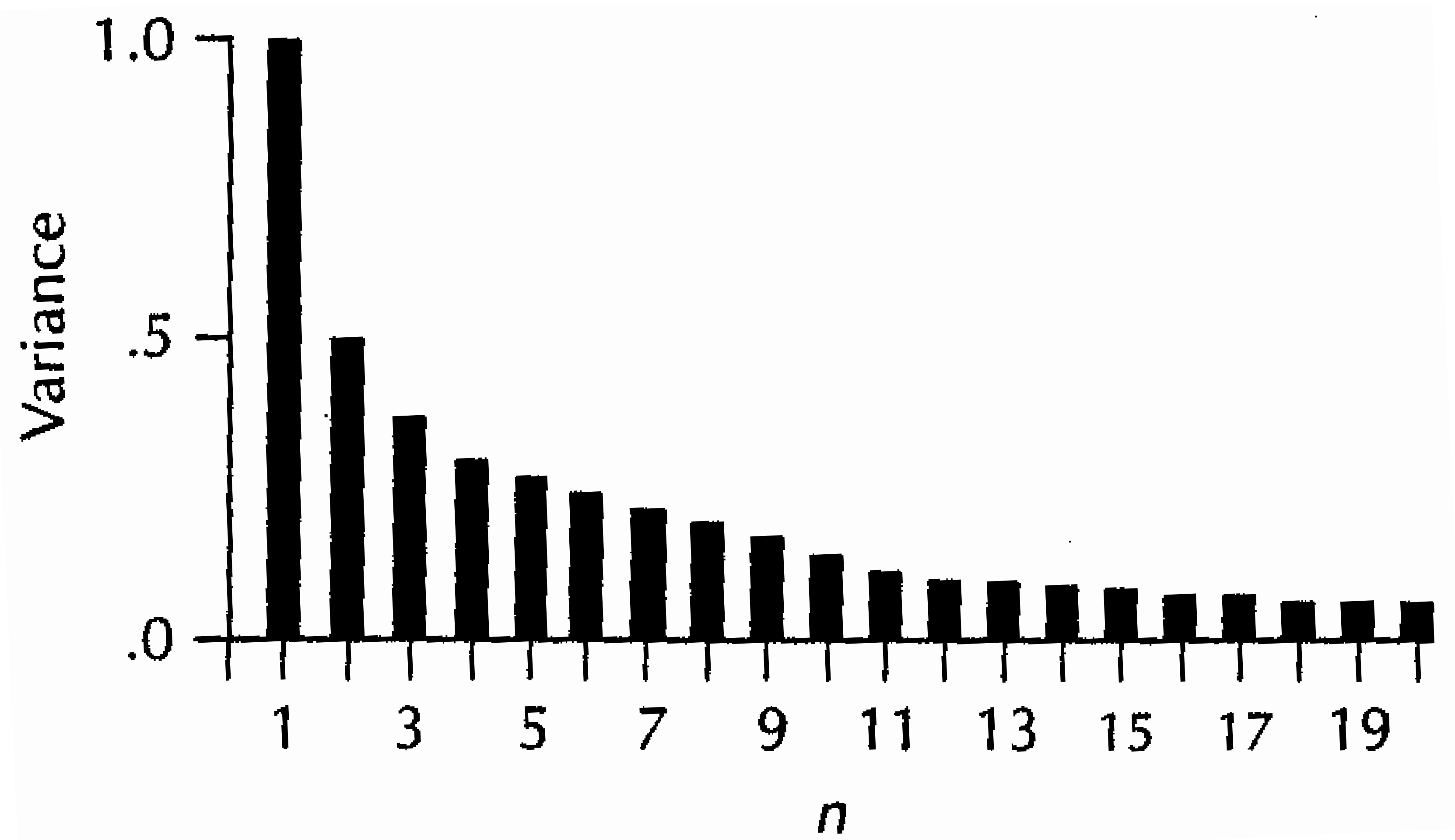
$$r_P = \frac{1}{N} \sum_{i=1}^{N} r_i.$$

Let σ_i^2 be the variance of the rate of return of asset i. When the rates of return are uncorrelated, the corresponding variance is

$$\text{var}(r_P) = \frac{1}{N^2} \sum_{i=1}^{N} \sigma_i^2 = \frac{\sigma_{aver}^2}{N}.$$

The variance decreases rapidly as N increases.

Uncorrelated assets



When the rates of return of assets are uncorrelated, the variance of a portfolio can be made very small.

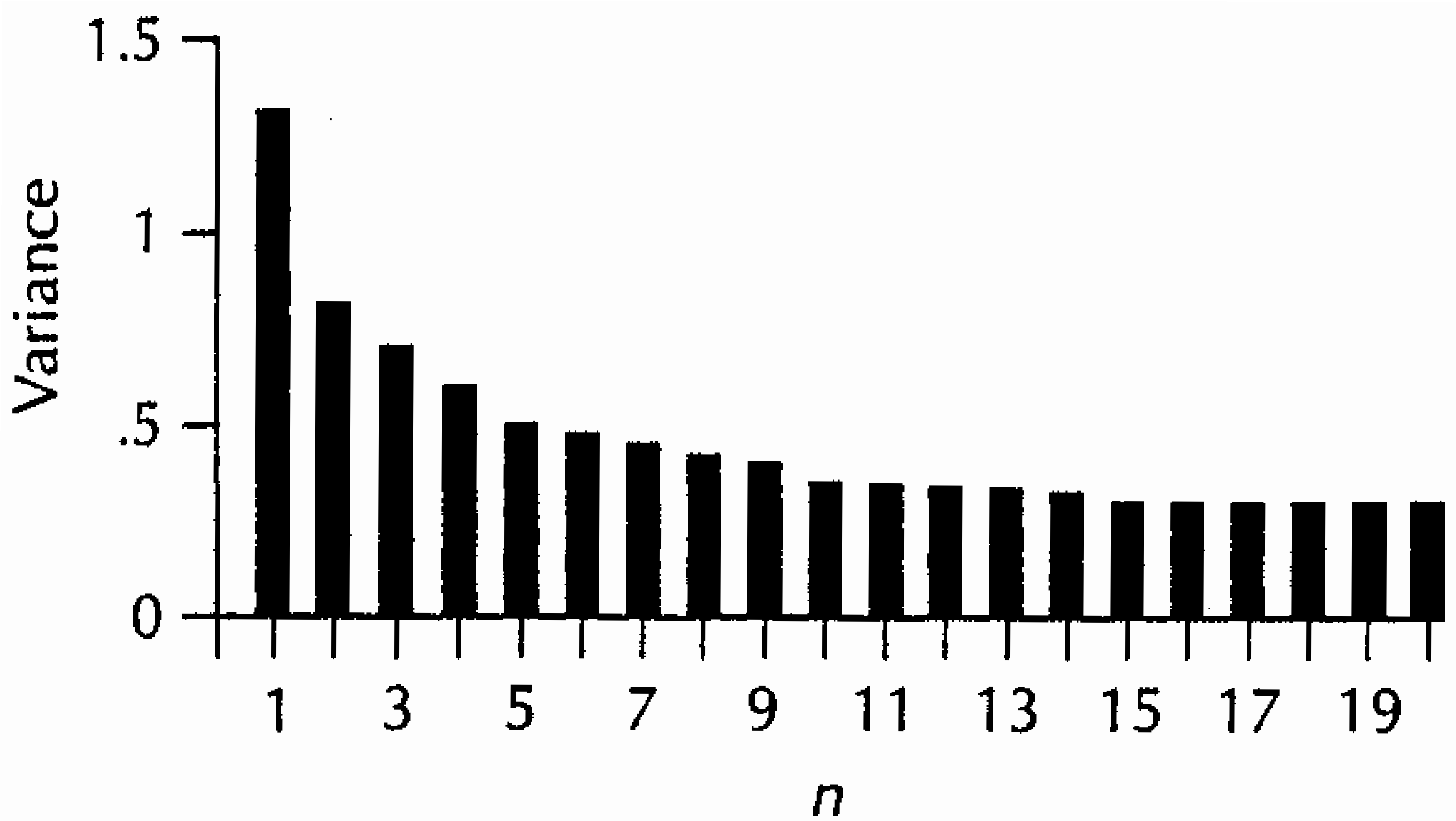
Non-zero correlation

We form a portfolio by taking equal portions of $w_i = \frac{1}{N}$ of these assets. In this case,

$$
\begin{aligned}
\text{var}(r_P) &= E\left[\sum_{i=1}^{N} \frac{1}{N}(r_i - \overline{r})\right]^2 \\
&= \frac{1}{N^2} E\left\{\left[\sum_{i=1}^{N}(r_i - \overline{r})\right]\left[\sum_{j=1}^{N}(r_j - \overline{r})\right]\right\} \\
&= \frac{1}{N^2}\sum_{i,j}\sigma_{ij} = \frac{1}{N^2}\left\{\sum_{i=j}\sigma_{ij} + \sum_{i\neq j}\sigma_{ij}\right\} \\
&= \frac{1}{N^2}\{N(\sigma_i^2)_{aver} + (N^2 - N)(\sigma_{ij})_{aver}\} \\
&= \frac{1}{N}\left[(\sigma_i^2)_{aver} - (\sigma_{ij})_{aver}\right] + (\sigma_{ij})_{aver}.
\end{aligned}
$$

The covariance terms remain when we take $N \to \infty$. Also, var(r_P) may be decreased by choosing assets that are negatively correlated by noting the presence of the term $\left(1 - \dfrac{1}{N}\right)(\sigma_{ij})_{aver}$.

Correlated assets



If returns of assets are correlated, there is likely to be a lower limit to the portfolio variance that can be achieved. This is because the term $(\sigma_{ij})_{aver}$ remains in var(r_P) even when $N \to \infty$.

2.2 Markowitz mean-variance formulation

We consider a single-period investment model. Suppose there are N risky assets, whose rates of return are given by the random variables $r_1, \cdots, r_N$, where

$$r_n = \frac{S_n(1) - S_n(0)}{S_n(0)}, \quad n = 1, 2, \cdots, N.$$

Here, time-0 stock price $S_n(0)$ is known while time-1 stock price $S_n(1)$ is random, $n = 1, 2, \cdots, N$. Let $\boldsymbol{w} = (w_1 \cdots w_N)^T, w_n$ denotes the proportion of wealth invested in asset n, with $\sum_{n=1}^{N} w_n = 1$. The rate of return of the portfolio r_P is

$$r_P = \sum_{n=1}^{N} w_n r_n.$$

Assumption

The two vectors $\boldsymbol{\mu} = (\overline{r}_1 \quad \overline{r}_2 \cdots \overline{r}_N)^T$ and $\mathbf{1} = (1 \quad 1 \cdots 1)^T$ are linearly independent. If otherwise, the mean rates of return are equal and so the portfolio return can only be the common mean rate of return. Under this degenerate case, the portfolio choice problem becomes a simpler minimization problem.

The first two moments of r_P are

$$\mu_P = E[r_P] = \sum_{n=1}^{N} E[w_n r_n] = \sum_{n=1}^{N} w_n \mu_n, \text{ where } \mu_n = \overline{r}_n,$$

and

$$\sigma_P^2 = \text{var}(r_P) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \text{cov}(r_i, r_j) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{ij}.$$

Covariance matrix

Let Ω denote the covariance matrix so that

$$\sigma_P^2 = \boldsymbol{w}^T \Omega \boldsymbol{w},$$

where Ω is symmetric and $(\Omega)_{ij} = \sigma_{ij} = \text{cov}(r_i, r_j)$. For example, when $n = 2$, we have

$$(w_1 \quad w_2) \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = w_1^2 \sigma_1^2 + w_1 w_2 (\sigma_{12} + \sigma_{21}) + w_2^2 \sigma_2^2.$$

Since portfolio variance σ_P^2 must be non-negative, so the covariance matrix must be symmetric and semi-positive definite. The eigenvalues are all real non-negative.

- Recall that $\det \Omega$ = product of eigenvalues and Ω^{-1} exists if and only if $\det \Omega \neq 0$. In our later discussion, we always assume Ω to be symmetric and positive definite (avoiding the unlikely event where one of the eigenvalues is zero) so that Ω^{-1} always exists. Note that Ω^{-1} is also symmetric and positive definite.

Sensitivity of σ_P^2 with respect to w_k

By the product rule in differentiation

$$\begin{aligned}\frac{\partial \sigma_P^2}{\partial w_k} &= \sum_{j=1}^{N}\sum_{i=1}^{N}\frac{\partial w_i}{\partial w_k}w_j\sigma_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\frac{\partial w_j}{\partial w_k}\sigma_{ij} \\ &= \sum_{j=1}^{N} w_j\sigma_{kj} + \sum_{i=1}^{N} w_i\sigma_{ik}.\end{aligned}$$

Since $\sigma_{kj} = \sigma_{jk}$, we obtain

$$\frac{\partial \sigma_P^2}{\partial w_k} = 2\sum_{j=1}^{N} w_j\sigma_{kj} = 2(\Omega \boldsymbol{w})_k,$$

where $(\Omega \boldsymbol{w})_k$ is the k^{th} component of the vector $\Omega \boldsymbol{w}$. Alternatively, we may write

$$\bigtriangledown \sigma_P^2 = 2\Omega \boldsymbol{w},$$

where $\bigtriangledown$ is the gradient operator. This partial derivative gives the sensitivity of the portfolio variance with respect to the weight of a particular asset.

Remarks

1. The portfolio risk of return is quantified by σ_P^2. In the mean-variance analysis, only the first two moments are considered in the portfolio investment model. Earlier investment theory prior to Markowitz only considered the maximization of μ_P without σ_P.

2. The measure of risk by variance would place equal weight on the upside and downside deviations. In reality, positive deviations should be more welcomed.

3. The assets are characterized by their random rates of return, $r_i, i = 1, \cdots, N$. In the mean-variance model, it is assumed that their first and second order moments: μ_i, σ_i and σ_{ij} are all known. In the Markowitz mean-variance formulation, we would like to determine the choice variables: $w_1, \cdots, w_N$ such that σ_P^2 is minimized for a given preset value of μ_P.

Two-asset portfolio

Consider a portfolio of two assets with known means $\overline{r}_1$ and $\overline{r}_2$, variances σ_1^2 and σ_2^2, of the rates of return r_1 and r_2, together with the correlation coefficient ρ, where $\text{cov}(r_1, r_2) = \rho\sigma_1\sigma_2$.

Let $1-\alpha$ and α be the weights of assets 1 and 2 in this two-asset portfolio, so $\boldsymbol{w} = (1-\alpha \quad \alpha)^T$.

Portfolio mean: $\overline{r}_P = (1-\alpha)\overline{r}_1 + \alpha\overline{r}_2$,

Portfolio variance: $\sigma_P^2 = (1-\alpha)^2\sigma_1^2 + 2\rho\alpha(1-\alpha)\sigma_1\sigma_2 + \alpha^2\sigma_2^2$.

Note that $\overline{r}_P$ is not affected by ρ while σ_P^2 is dependent on ρ.

assets' mean and variance

	Asset A	Asset B
Mean return (%)	10	20
Variance (%)	10	15

Portfolio mean[a] and variance[b] for weights and asset correlations

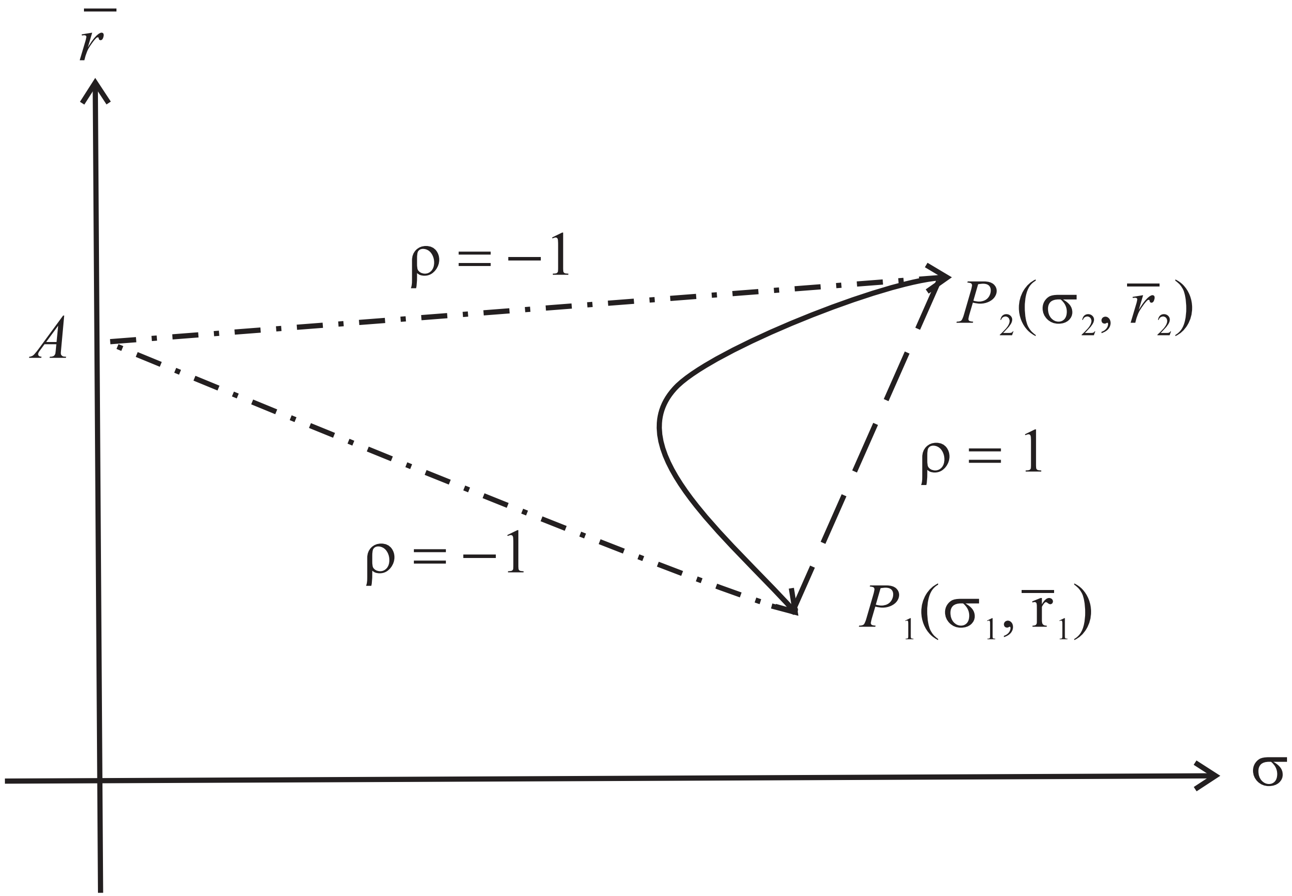
weight		$\rho = -1$		$\rho = -0.5$		$\rho = 0.5$		$\rho = 1$	
w_A	$w_B = 1 - w_A$	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
1.0	0.0	10.0	10.00	10.0	10.00	10.0	10.00	10.0	10.00
0.8	0.2	12.0	3.08	12.0	5.04	12.0	8.96	12.0	10.92
0.5	0.5	15.0	0.13	15.0	3.19	15.0	9.31	15.0	12.37
0.2	0.8	18.0	6.08	18.0	8.04	18.0	11.96	18.00	13.92
0.0	1.0	20.0	15.00	20.0	15.00	20.0	15.00	20.0	15.00

[a] The mean is calculated as $E(R) = w_A 10 + (1 - w_A)20$.

[b] The variance is calculated as $\sigma_P^2 = w_A^2 10 + (1 - w_A)^2 15 + 2 w_A (1 - w_A) \rho \sqrt{10}\sqrt{15}$ where ρ is the assumed correlation coefficient and $\sqrt{10}$ and $\sqrt{15}$ are standard deviations of the returns of the two assets, respectively.

Observation: A lower variance is achieved for a given mean when the correlation of the pair of assets' returns becomes more negative.

We represent the two assets in a mean-standard deviation diagram



As α varies, $(\sigma_P, \overline{r}_P)$ traces out a conic curve in the σ-$\overline{r}$ plane. With $\rho = -1$, it is possible to have $\sigma_P = 0$ for some suitable choice of weight α. Note that $P_1(\sigma_1, \overline{r}_1)$ corresponds to $\alpha = 0$ while $P_2(\sigma_2, \overline{r}_2)$ corresponds to $\alpha = 1$.

Consider the special case where $\rho = 1$,

$$\begin{aligned}\sigma_P(\alpha;\rho=1) &= \sqrt{(1-\alpha)^2\sigma_1^2 + 2\alpha(1-\alpha)\sigma_1\sigma_2 + \alpha^2\sigma_2^2} \\ &= (1-\alpha)\sigma_1 + \alpha\sigma_2.\end{aligned}$$

Since $\overline{r}_P$ and σ_P are linear in α, and if we choose $0 \leq \alpha \leq 1$, then the portfolios are represented by the straight line joining $P_1(\sigma_1, \overline{r}_1)$ and $P_2(\sigma_2, \overline{r}_2)$.

When $\rho = -1$, we have

$$\sigma_P(\alpha;\rho=-1) = \sqrt{[(1-\alpha)\sigma_1 - \alpha\sigma_2]^2} = |(1-\alpha)\sigma_1 - \alpha\sigma_2|.$$

Since both $\overline{r}_P$ and σ_P are also linear in α, $(\sigma_P, \overline{r}_P)$ traces out linear line segments. When α is small (close to zero), the corresponding point is close to $P_1(\sigma_1, \overline{r}_1)$. The line AP_1 corresponds to

$$\sigma_P(\alpha;\rho=-1) = (1-\alpha)\sigma_1 - \alpha\sigma_2.$$

The point A corresponds to $\alpha = \dfrac{\sigma_1}{\sigma_1+\sigma_2}$. It is a point on the vertical axis which has zero value of σ_P. Also, see point 5 in the Appendix.

The quantity $(1-\alpha)\sigma_1 - \alpha\sigma_2$ remains positive until $\alpha = \dfrac{\sigma_1}{\sigma_1+\sigma_2}$. When $\alpha > \dfrac{\sigma_1}{\sigma_1+\sigma_2}$, the locus traces out the upper line AP_2 corresponding to $\sigma_P(\alpha;\rho=-1) = \alpha\sigma_2 - (1-\alpha)\sigma_1$. In summary, we have

$$\sigma_P(\alpha;\rho=-1) = \begin{cases} (1-\alpha)\sigma_1 - \alpha\sigma_2 & \alpha \le \frac{\sigma_1}{\sigma_1+\sigma_2} \\ \alpha\sigma_2 - (1-\alpha)\sigma_1 & \alpha > \frac{\sigma_1}{\sigma_1+\sigma_2} \end{cases}.$$

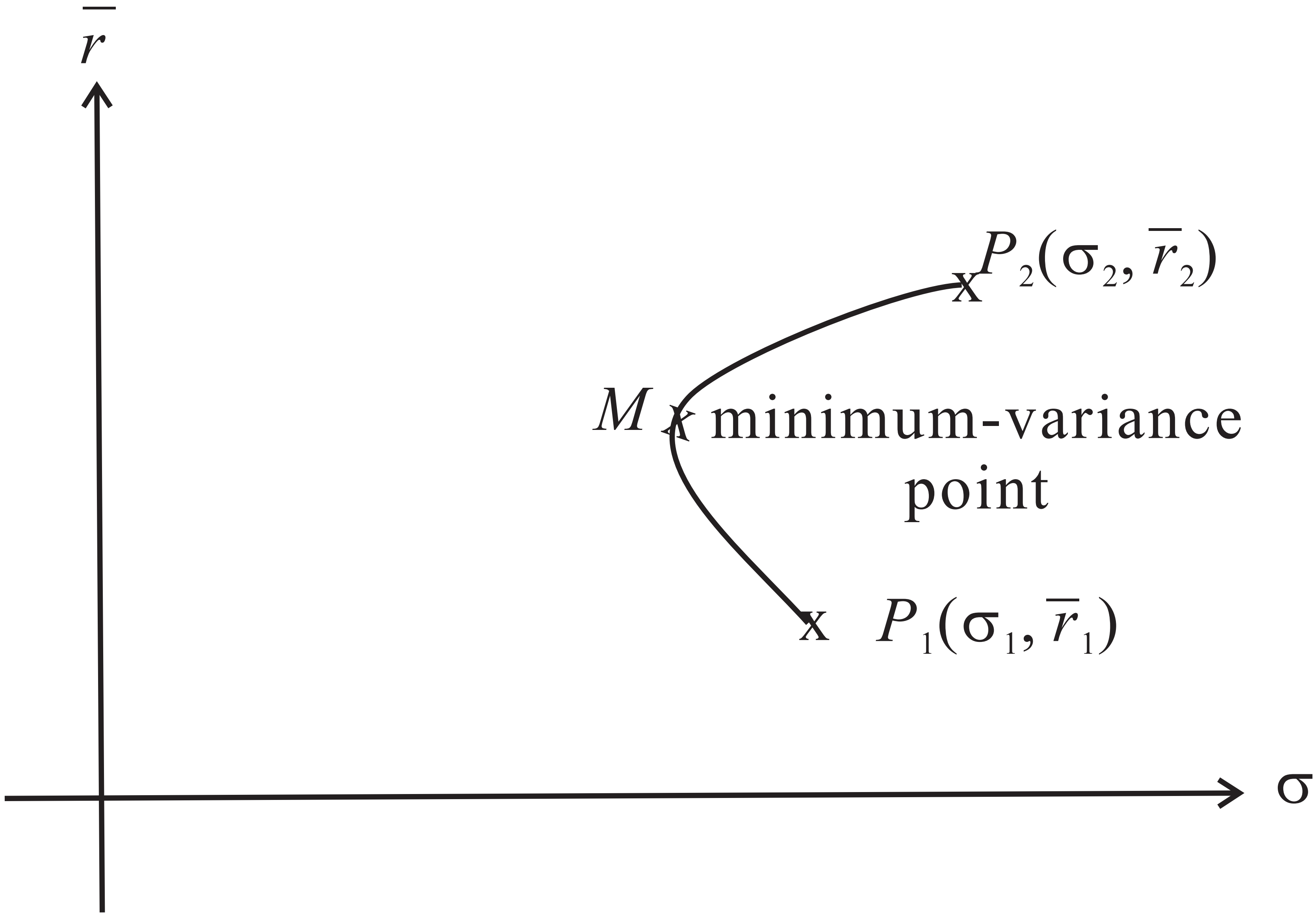
Suppose $-1 < \rho < 1$, the minimum variance point on the curve that represents various portfolio combinations is determined by

$$\frac{\partial\sigma_P^2}{\partial\alpha} = -2(1-\alpha)\sigma_1^2 + 2\alpha\sigma_2^2 + 2(1-2\alpha)\rho\sigma_1\sigma_2 = 0$$

giving

$$\alpha = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}.$$

Mean-standard deviation diagram



Formulation of Markowitz's mean-variance analysis

$$\text{minimize } \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij}$$

subject to $\sum_{i=1}^{N} w_i \overline{r}_i = \mu_P$ and $\sum_{i=1}^{N} w_i = 1$. Given the target expected rate of return of portfolio μ_P, we find the optimal portfolio strategy that minimizes σ_P^2. The constraint: $\sum_{i=1}^{N} w_i = 1$ refers to the strategy of putting *all* wealth into investment of risky assets.

Solution

We form the Lagrangian

$$L = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} - \lambda_1 \left(\sum_{i=1}^{N} w_i - 1\right) - \lambda_2 \left(\sum_{i=1}^{N} w_i \overline{r}_i - \mu_P\right),$$

where λ_1 and λ_2 are the Lagrangian multipliers.

We differentiate L with respect to w_i and the Lagrangian multipliers, then set all the derivatives be zero.

$$\frac{\partial L}{\partial w_i} = \sum_{j=1}^{N} \sigma_{ij} w_j - \lambda_1 - \lambda_2 \overline{r}_i = 0, \quad i = 1, 2, \cdots, N; \tag{1}$$

$$\frac{\partial L}{\partial \lambda_1} = \sum_{i=1}^{N} w_i - 1 = 0; \tag{2}$$

$$\frac{\partial L}{\partial \lambda_2} = \sum_{i=1}^{N} w_i \overline{r}_i - \mu_P = 0. \tag{3}$$

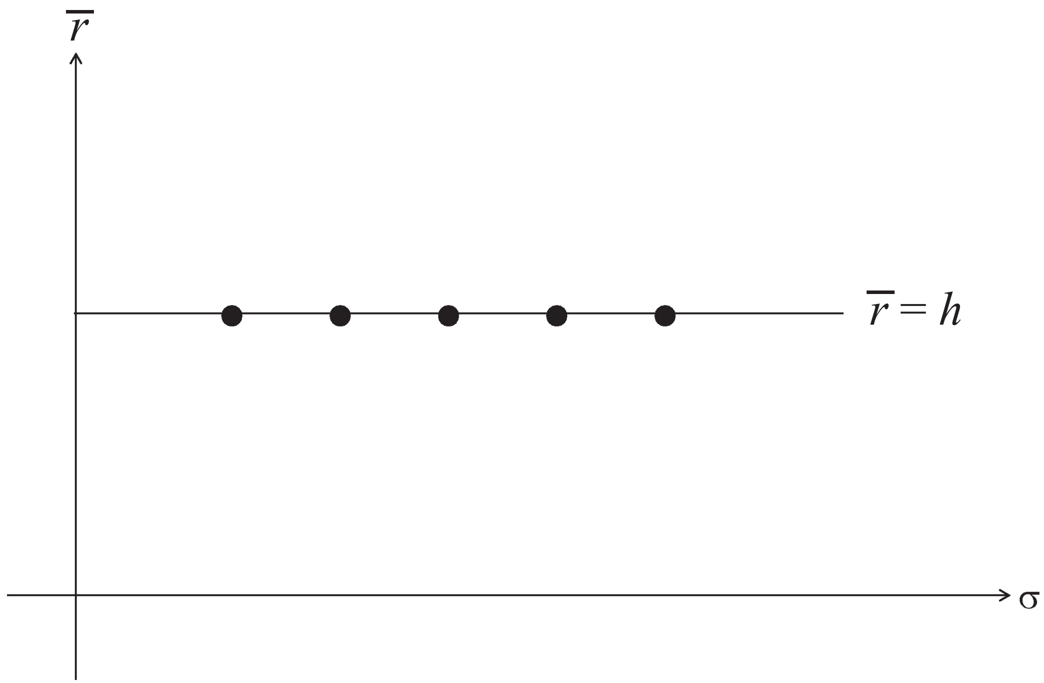
From Eq. (1), we deduce that the optimal portfolio vector weight $\boldsymbol{w}^*$ admits solution of the form

$$\Omega \boldsymbol{w}^* = \lambda_1 \mathbf{1} + \lambda_2 \boldsymbol{\mu} \text{ or } \boldsymbol{w}^* = \Omega^{-1}(\lambda_1 \mathbf{1} + \lambda_2 \boldsymbol{\mu})$$

where $\mathbf{1} = (1 \quad 1 \cdots 1)^T$ and $\boldsymbol{\mu} = (\overline{r}_1 \quad \overline{r}_2 \cdots \overline{r}_N)^T$.

Degenerate case

Consider the case where all assets have the same expected rate of return, that is, $\boldsymbol{\mu} = h\mathbf{1}$ for some constant h. In this case, the solution to Eqs. (2) and (3) gives $\mu_P = h$. The assets are represented by points that all lie on the horizontal line: $\overline{r} = h$.



In this case, the expected portfolio return cannot be arbitrarily prescribed. Actually, we have to take $\mu_P = h$, so the constraint on the expected portfolio return becomes irrelevant.

Solution procedure

To determine λ_1 and λ_2, we apply the two constraints:

$$\begin{aligned} 1 &= \mathbf{1}^T\Omega^{-1}\Omega \boldsymbol{w}^* = \lambda_1 \mathbf{1}^T\Omega^{-1}\mathbf{1} + \lambda_2 \mathbf{1}^T\Omega^{-1}\boldsymbol{\mu} \\ \mu_P &= \boldsymbol{\mu}^T\Omega^{-1}\Omega \boldsymbol{w}^* = \lambda_1 \boldsymbol{\mu}^T\Omega^{-1}\mathbf{1} + \lambda_2 \boldsymbol{\mu}^T\Omega^{-1}\boldsymbol{\mu}. \end{aligned}$$

Writing $a = \mathbf{1}^T\Omega^{-1}\mathbf{1}, b = \mathbf{1}^T\Omega^{-1}\boldsymbol{\mu}$ and $c = \boldsymbol{\mu}^T\Omega^{-1}\boldsymbol{\mu}$, we have two equations for λ_1 and λ_2:

$$1 = \lambda_1 a + \lambda_2 b \text{ and } \mu_P = \lambda_1 b + \lambda_2 c.$$

Solving for λ_1 and λ_2:

$$\lambda_1 = \frac{c - b\mu_P}{\Delta} \quad \text{and} \quad \lambda_2 = \frac{a\mu_P - b}{\Delta},$$

where $\Delta = ac - b^2$. Provided that $\boldsymbol{\mu} \neq h\mathbf{1}$ for some scalar h, we then have $\Delta \neq 0$.

Solution to the minimum portfolio variance

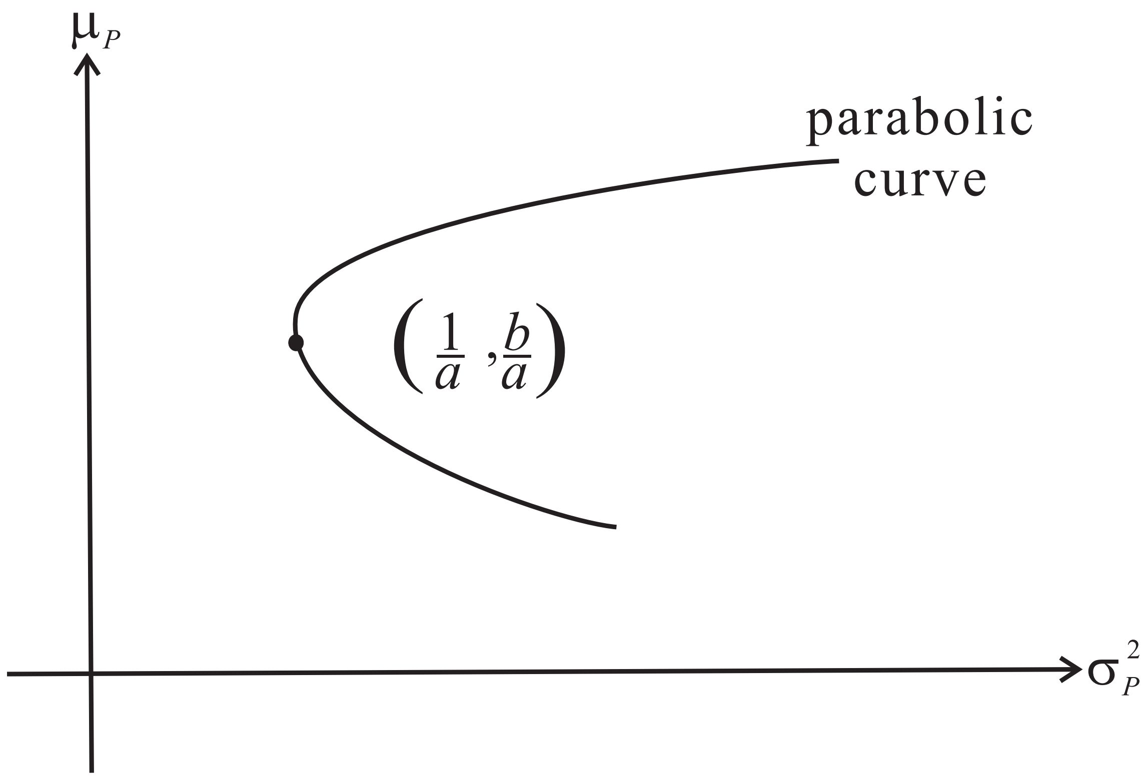
- Both λ_1 and λ_2 have dependence on μ_P, where μ_P is the target mean prescribed in the variance minimization problem.

- The minimum portfolio variance for a given value of μ_P is given by

$$\begin{aligned}\sigma_P^2 &= \boldsymbol{w}^{*T}\Omega\boldsymbol{w}^* = \boldsymbol{w}^{*T}(\lambda_1\mathbf{1} + \lambda_2\boldsymbol{\mu}) \\ &= \lambda_1 + \lambda_2\mu_P = \frac{a\mu_P^2 - 2b\mu_P + c}{\Delta}.\end{aligned}$$

- $\sigma_P^2 = \boldsymbol{w}^T\Omega\boldsymbol{w} \geq 0$, for all $\boldsymbol{w}$, so Ω is guaranteed to be semi-positive definite. In our subsequent analysis, we assume Ω to be positive definite. Given that Ω is positive definite, so does Ω^{-1}, we have $a > 0, c > 0$ and Ω^{-1} exists. By virtue of the Cauchy-Schwarz inequality, $\Delta > 0$. Since a and Δ are both positive, the quantity $a\mu_P^2 - 2b\mu_P + c$ is guaranteed to be positive (since the quadratic equation has no real root, a result from highschool mathematics).

The set of minimum variance portfolios is represented by a parabolic curve in the $\sigma_P^2 - \mu_P$ plane. The parabolic curve is generated by varying the value of the parameter μ_P. Note that $\frac{1}{a} > 0$ while $\frac{b}{a}$ may become negative under some extreme adverse cases of negative mean rates of return.



Non-optimal portfolios are represented by points which must fall on the right side of the parabolic curve.

Global minimum variance portfolio

Given μ_P, we obtain $\lambda_1 = \dfrac{c - b\mu_P}{\Delta}$ and $\lambda_2 = \dfrac{a\mu_P - b}{\Delta}$, and the optimal weight $\boldsymbol{w}^* = \Omega^{-1}(\lambda_1\mathbf{1} + \lambda_2\boldsymbol{\mu}) = \dfrac{c - b\mu_P}{\Delta}\Omega^{-1}\mathbf{1} + \dfrac{a\mu_P - b}{\Delta}\Omega^{-1}\boldsymbol{\mu}$.

To find the global minimum variance portfolio, we set

$$\frac{d\sigma_P^2}{d\mu_P} = \frac{2a\mu_P - 2b}{\Delta} = 0$$

so that $\mu_P = b/a$ and $\sigma_P^2 = 1/a$. Correspondingly, $\lambda_1 = 1/a$ and $\lambda_2 = 0$. The weight vector that gives the global minimum variance portfolio is found to be

$$\boldsymbol{w}_g = \lambda_1\Omega^{-1}\mathbf{1} = \frac{\Omega^{-1}\mathbf{1}}{a} = \frac{\Omega^{-1}\mathbf{1}}{\mathbf{1}^T\Omega^{-1}\mathbf{1}}.$$

Note that $\boldsymbol{w}_g$ is independent of $\boldsymbol{\mu}$. Obviously, $\boldsymbol{w}_g^T\mathbf{1} = 1$ due to the normalization factor $\mathbf{1}^T\Omega^{-1}\mathbf{1}$ in the denominator. As a check, we have $\mu_g = \boldsymbol{\mu}^T\boldsymbol{w}_g = \dfrac{b}{a}$ and $\sigma_g^2 = \boldsymbol{w}_g^T\Omega\boldsymbol{w}_g = \dfrac{1}{a}$.

Example

Given the variance matrix

$$\Omega = \begin{pmatrix} 2 & 0.5 & 0 \\ 0.5 & 3 & 0.5 \\ 0 & 0.5 & 2 \end{pmatrix},$$

find $\boldsymbol{w}_g$. This can be obtained effectively by solving

$$\begin{aligned} 2v_1 + 0.5v_2 &= 1 \\ 0.5v_1 + 3v_2 + 0.5v_3 &= 1 \\ 0.5v_2 + 2v_3 &= 1. \end{aligned}$$

Here, $\boldsymbol{v} = (v_1 \ v_2 \ v_3)^T$ gives $\Omega^{-1}\mathbf{1}$. Due to symmetry between asset 1 and asset 3 since $\sigma_1^2 = \sigma_3^2$ and $\sigma_{12} = \sigma_{32}$, etc., we expect $v_1 = v_3$.

The above system reduces to

$$\begin{aligned} 2v_1 + 0.5v_2 &= 1 \\ v_1 + 3v_2 &= 1 \end{aligned}$$

giving $v_1 = v_3 = \frac{5}{11}$ and $v_2 = \frac{2}{11}$. Lastly, by normalization to sum of weights equals 1, we obtain

$$\boldsymbol{w}_g = \begin{pmatrix} \frac{5}{12} & \frac{2}{12} & \frac{5}{12} \end{pmatrix}^T.$$

Two-parameter ($\lambda_1 - \lambda_2$) family of minimum variance portfolios

Recall $\boldsymbol{w}^* = \lambda_1 \Omega^{-1}\mathbf{1} + \lambda_2 \Omega^{-1}\boldsymbol{\mu}$, so the minimum variance portfolios (frontier funds) are seen to be generated by a linear combination of $\Omega^{-1}\mathbf{1}$ and $\Omega^{-1}\boldsymbol{\mu}$, where $\boldsymbol{\mu} \neq h\mathbf{1}$ so that $\Omega^{-1}\mathbf{1}$ and $\lambda^{-1}\boldsymbol{\mu}$ are independent.

It is not surprising to see that $\lambda_2 = 0$ corresponds to $\boldsymbol{w}_g^*$ since the constraint on the target mean vanishes when λ_2 is taken to be zero. In this case, we minimize risk while paying no regard to the target mean, thus the global minimum variance portfolio is resulted.

Suppose we normalize $\Omega^{-1}\boldsymbol{\mu}$ by b and define

$$\boldsymbol{w}_d = \frac{\Omega^{-1}\boldsymbol{\mu}}{b} = \frac{\Omega^{-1}\boldsymbol{\mu}}{\mathbf{1}^T\Omega^{-1}\boldsymbol{\mu}}.$$

Obviously, $\boldsymbol{w}_d$ also lies on the frontier since it is a member of the family of minimum variance portfolio with $\lambda_1 = 0$ and $\lambda_2 = \dfrac{1}{b}$.

The corresponding expected rate of return μ_d and σ_d^2 are given by

$$\mu_d = \boldsymbol{\mu}^T \boldsymbol{w}_d = \frac{c}{b}$$

$$\sigma_d^2 = \frac{\left(\Omega^{-1}\boldsymbol{\mu}\right)^T \Omega \left(\Omega^{-1}\boldsymbol{\mu}\right)}{b^2} = \frac{\boldsymbol{\mu}^T \Omega^{-1} \boldsymbol{\mu}}{b^2} = \frac{c}{b^2}.$$

Since $\Omega^{-1}\mathbf{1} = a\boldsymbol{w}_g$ and $\Omega^{-1}\boldsymbol{\mu} = b\boldsymbol{w}_d$, the weight of any frontier fund (minimum variance fund) can be represented by

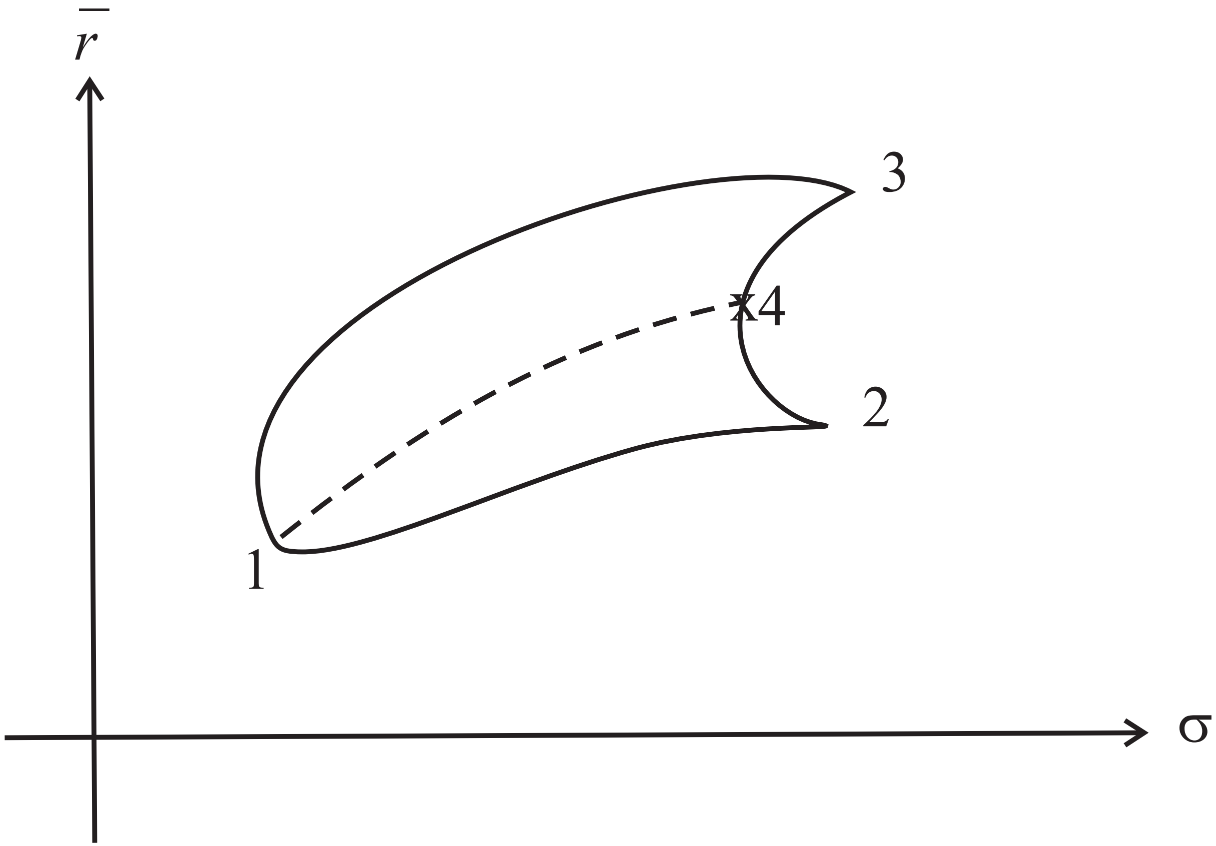
$$\boldsymbol{w}^* = (\lambda_1 a)\boldsymbol{w}_g + (\lambda_2 b)\boldsymbol{w}_d = \frac{c - b\mu_P}{\Delta} a\boldsymbol{w}_g + \frac{a\mu_P - b}{\Delta} b\boldsymbol{w}_d.$$

This provides the motivation of the *Two-Fund Theorem*. The above representation indicates that the optimal portfolio weight $\boldsymbol{w}^*$ depends on μ_P set by the investor.

- Any minimum variance fund can be generated by an appropriate combination of the two funds corresponding to $\boldsymbol{w}_g$ and $\boldsymbol{w}_d$ (see Sec. 2.3: Two-fund Theorem).

Feasible set

Given N risky assets, we can form various portfolios from these N assets. We plot the point $(\sigma_P, \overline{r}_P)$ that represents a particular portfolio in the $\sigma - \overline{r}$ diagram. The collection of these points constitutes the feasible set or feasible region.



Argument to show that the collection of the points representing $(\sigma_P, \overline{r}_P)$ *of a 3-asset portfolio generates a solid region in the* σ-$\overline{r}$ *plane*

- Consider a 3-asset portfolio, the various combinations of assets 2 and 3 sweep out a curve between them (the particular curve taken depends on the correlation coefficient ρ_{23}).

- A combination of assets 2 and 3 (labelled 4) can be combined with asset 1 to form a curve joining 1 and 4. As 4 moves between 2 and 3, the family of curves joining 1 and 4 sweep out a solid region.

Properties of the feasible regions

1. For a portfolio with at least 3 risky assets (not perfectly correlated and with different means), the feasible set is a solid two-dimensional region.

2. The feasible region is *convex to the left*. Any combination of two portfolios also lies in the feasible region. Indeed, the left boundary of a feasible region is a hyperbola (as solved by the Markowitz constrained minimization model).

Locate the efficient and inefficient investment strategies

- Since investors prefer the lowest variance for the same expected return, they will focus on the set of portfolios with the smallest variance for a given mean, or the mean-variance frontier (collection of minimum variance portfolios).

- The mean-variance frontier can be divided into two parts: an efficient frontier and an inefficient frontier.

- The efficient part includes the portfolios with the highest mean for a given variance.

Minimum variance set and efficient funds

The left boundary of a feasible region is called the *minimum variance set*. The most left point on the minimum variance set is called the *global minimum variance point*. The portfolios in the minimum variance set are called the *frontier funds*.

For a given level of risk, only those portfolios on the *upper half* of the efficient frontier with a higher return are desired by investors. They are called the *efficient funds*.

A portfolio $\boldsymbol{w}^*$ is said to be mean-variance efficient if there exists no portfolio $\boldsymbol{w}$ with $\mu_P \geq \mu_P^*$ and $\sigma_P^2 \leq {\sigma_P^*}^2$, except itself. That is, you cannot find a portfolio that has a higher return and lower risk than those of an efficient portfolio. The funds on the inefficient frontier do not exhibit the above properties.

Example – Uncorrelated assets with short sales constraint

Suppose there are three uncorrelated assets. Each has variance 1, and the mean rates of return are $1, 2$ and 3 (in percentage points), respectively. We have $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1$ and $\sigma_{12} = \sigma_{23} = \sigma_{13} = 0$; that is $\Omega = I$.

The first order conditions give $\Omega \boldsymbol{w} = \lambda_1 \mathbf{1} + \lambda_2 \boldsymbol{\mu}$, $\boldsymbol{\mu}^T \boldsymbol{w} = \mu_P$ and $\mathbf{1}^T \boldsymbol{w} = 1$, so we obtain

$$
\begin{aligned}
w_1 - \lambda_2 - \lambda_1 &= 0 \\
w_2 - 2\lambda_2 - \lambda_1 &= 0 \\
w_3 - 3\lambda_2 - \lambda_1 &= 0 \\
w_1 + 2w_2 + 3w_3 &= \mu_P \\
w_1 + w_2 + w_3 &= 1.
\end{aligned}
$$

By eliminating w_1, w_2, w_3, we obtain two equations for λ_1 and λ_2

$$
\begin{aligned}
14\lambda_2 + 6\lambda_1 &= \mu_P \\
6\lambda_2 + 3\lambda_1 &= 1.
\end{aligned}
$$

These two equations can be solved to yield $\lambda_2 = \frac{\mu_P}{2} - 1$ and $\lambda_1 = 2\frac{1}{3} - \mu_P$. The portfolio weights are expressed in terms of μ_P:

$$\begin{aligned} w_1 &= \frac{4}{3} - \frac{\mu_P}{2} \\ w_2 &= \frac{1}{3} \\ w_3 &= \frac{\mu_P}{2} - \frac{2}{3}. \end{aligned}$$

The standard deviation of r_P at the solution is $\sqrt{w_1^2 + w_2^2 + w_3^2}$, which by direct substitution gives

$$\sigma_P = \sqrt{\frac{7}{3} - 2\mu_P + \frac{\mu_P^2}{2}}.$$

The minimum-variance point is, by symmetry, at $\mu_P = 2$, with $\sigma_P = \sqrt{3}/3 = 0.58$. When $\mu_P = 2$, we obtain

$$w_1 = w_2 = w_3 = \frac{1}{3}.$$

Short sales not allowed (adding the constraints: $w_i \geq 0, i = 1, 2, 3$)

Unlike the unrestricted case of allowing short sales, we now impose $w_i \geq 0, i = 1, 2, 3$. As a result, μ_P can only lie between $1 \leq \mu_P \leq 3$ [recall $\mu_P = w_1 + 2w_2 + 3w_3$]. The lower bound is easily seen since $\mu_P = (w_1 + w_2 + w_3) + (w_2 + 2w_3) = 1 + w_2 + 2w_3 \geq 1$ since $w_2 \geq 0$ and $w_3 \geq 0$. Also, μ_P cannot go above 3 as the maximum value of μ_P can only be achieved by choosing $w_3 = 1$, $w_1 = w_2 = 0$. For certain range of μ_P, some of the optimal portfolio weights may become negative when there is no short sales constraint.

It is instructive to consider seperately, the following 3 intervals for μ_P: $\left[1, \dfrac{4}{3}\right]$, $\left[\dfrac{4}{3}, \dfrac{8}{3}\right]$ and $\left[\dfrac{8}{3}, 3\right]$.

$1 \le \mu_P \le \frac{4}{3}$	$\frac{4}{3} \le \mu_P \le \frac{8}{3}$	$\frac{8}{3} \le \mu_P \le 3$
$w_1 = 2 - \mu_P$	$\frac{4}{3} - \frac{\mu_P}{2}$	0
$w_2 = \mu_P - 1$	$\frac{1}{3}$	$3 - \mu_P$
$w_3 = 0$	$\frac{\mu_P}{2} - \frac{2}{3}$	$\mu_P - 2$

$$\sigma_P = \sqrt{2\mu_P^2 - 6\mu_P + 5} \quad \sqrt{\frac{7}{3} - 2\mu_P + \frac{\mu_P^2}{2}} \quad \sqrt{2\mu_P^2 - 10\mu_P + 13}.$$

- From $w_3 = \frac{\mu_P}{2} - \frac{2}{3}$, we deduce that when $1 \leq \mu_P < \frac{4}{3}, w_3$ becomes negative in the minimum variance portfolio when short sales are allowed. This is truly an inferior investment choice as the investor sets μ_P to be too low while asset 3 has the highest mean rate of return. When short sales are not allowed, we expect to have "$w_3 = 0$" in the minimum variance portfolio. The problem reduces to two-asset portfolio model and the corresponding optimal weights w_1 and w_2 can be easily obtained by solving

$$\begin{aligned} w_1 + 2w_2 &= \mu_P \\ w_1 + w_2 &= 1. \end{aligned}$$

- Similarly, when $\frac{8}{3} \leq \mu_P \leq 3$, it is optimal to choose $w_1 = 0$. In this case, the investor sets μ_P to be too high while asset 1 has the lowest mean rate of return.

- When $\frac{4}{3} \leq \mu_P \leq \frac{8}{3}$, we have the same solution as the case without the short sales constraint. This is because the solutions to the weights happen to be non-negative under the unconstrained case. The short sales constraint becomes redundant.

2.3 Two-fund Theorem

Take any two frontier funds (portfolios), then any combination of these two frontier funds remains to be a frontier fund. Indeed, any frontier portfolio can be duplicated, in terms of mean and variance, as a combination of these two frontier funds. In other words, all investors seeking frontier portfolios need only invest in various combinations of these two funds. This property can be extended to a combination of efficient funds (frontier funds that lie on the upper portion of the efficient frontier)?

Remark

This is analogous to the concept of a basis of $\mathbb{R}^2$ with two independent basis vectors. Any vector in $\mathbb{R}^2$ can be expressed as a unique linear combination of the basis vectors. Choices of bases of $\mathbb{R}^2$ can be $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$, $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$, etc.

Proof of the Two-fund Theorem

Let $\boldsymbol{w}^1 = (w_1^1 \cdots w_n^1), \lambda_1^1, \lambda_2^1$ and $\boldsymbol{w}^2 = (w_1^2 \cdots w_n^2)^T, \lambda_1^2, \lambda_2^2$ be two known solutions to the minimum variance formulation with expected rates of return μ_P^1 and μ_P^2, respectively. By setting μ_P equal μ_P^1 and μ_P^2 successively, both solutions satisfy

$$\sum_{j=1}^{n} \sigma_{ij} w_j - \lambda_1 - \lambda_2 \overline{r}_i = 0, \quad i = 1, 2, \cdots, n \tag{1}$$

$$\sum_{i=1}^{n} w_i \overline{r}_i = \mu_P \tag{2}$$

$$\sum_{i=1}^{n} w_i = 1. \tag{3}$$

We would like to show that $\alpha \boldsymbol{w}_1 + (1-\alpha) \boldsymbol{w}_2$ is a solution corresponds to the expected rate of return $\alpha \mu_P^1 + (1-\alpha)\mu_P^2$.

For example, $\mu_P^1 = 2\%$, $\mu_P^2 = 4\%$, and we set μ_P to be 2.5%, then $\alpha = 0.75$.

1. The new weight vector $\alpha \boldsymbol{w}^1 + (1-\alpha)\boldsymbol{w}^2$ is a legitimate portfolio with weights that sum to one.

2. Check the condition on the expected rate of return

$$\begin{aligned}
& \sum_{i=1}^{n} \left[\alpha w_i^1 + (1-\alpha) w_i^2\right] \overline{r}_i \\
= \; & \alpha \sum_{i=1}^{n} w_i^1 \overline{r}_i + (1-\alpha) \sum_{i=1}^{n} w_i^2 \overline{r}_i \\
= \; & \alpha \mu_P^1 + (1-\alpha)\mu_P^2.
\end{aligned}$$

3. Eq. (1) is satisfied by $\alpha \boldsymbol{w}^1 + (1-\alpha)\boldsymbol{w}^2$ since the system of equations is linear. The corresponding λ_1 and λ_2 are given by

$$\lambda_1 = \alpha\lambda_1^1 + (1-\alpha)\lambda_1^2 \quad \text{and} \quad \lambda_2 = \alpha\lambda_2^1 + (1-\alpha)\lambda_2^2.$$

4. Given μ_P, the appropriate portion α is determined by

$$\mu_P = \alpha\mu_P^1 + (1-\alpha)\mu_P^2.$$

Global minimum variance portfolio $\boldsymbol{w}_g$ and the counterpart $\boldsymbol{w}_d$

For convenience, we choose the two frontier funds to be $\boldsymbol{w}_g$ and $\boldsymbol{w}_d$. To obtain the optimal weight $\boldsymbol{w}^*$ for a given μ_P, we solve for α using $\alpha\mu_g+(1-\alpha)\mu_d=\mu_P$ and $\boldsymbol{w}^*$ is then given by $\alpha\boldsymbol{w}_g+(1-\alpha)\boldsymbol{w}_d$. Recall $\mu_g=b/a$ and $\mu_d=c/b$, so $\alpha=\dfrac{(c-b\mu_P)a}{\Delta}$.

Proposition

Any minimum variance portfolio with the target mean μ_P can be uniquely decomposed into the sum of two portfolios

$$\boldsymbol{w}_P^*=\alpha\boldsymbol{w}_g+(1-\alpha)\boldsymbol{w}_d$$

where $\alpha=\dfrac{c-b\mu_P}{\Delta}a$.

Indeed, any two minimum variance portfolios $\boldsymbol{w}_u$ and $\boldsymbol{w}_v$ on the frontier can be used to substitute for $\boldsymbol{w}_g$ and $\boldsymbol{w}_d$. Suppose

$$\begin{aligned}\boldsymbol{w}_u &= (1-u)\boldsymbol{w}_g + u\boldsymbol{w}_d \\ \boldsymbol{w}_v &= (1-v)\boldsymbol{w}_g + v\boldsymbol{w}_d\end{aligned}$$

we then solve for $\boldsymbol{w}_g$ and $\boldsymbol{w}_d$ in terms of $\boldsymbol{w}_u$ and $\boldsymbol{w}_v$. Recall

$$\boldsymbol{w}_P^* = \lambda_1\Omega^{-1}\mathbf{1} + \lambda_2\Omega^{-1}\boldsymbol{\mu}$$

so that

$$\begin{aligned}\boldsymbol{w}_P^* &= \lambda_1 a\boldsymbol{w}_g + (1-\lambda_1 a)\boldsymbol{w}_d \\ &= \frac{\lambda_1 a + v - 1}{v-u}\boldsymbol{w}_u + \frac{1-u-\lambda_1 a}{v-u}\boldsymbol{w}_v,\end{aligned}$$

whose sum of coefficients remains to be 1 and $\lambda_1 = \dfrac{c - b\mu_P}{\Delta}$.

Convex combination of efficient portfolios

Any convex combination (that is, weights are non-negative) of efficient portfolios is also an efficient portfolio.

Proof

Let $w_i \geq 0$ be the weight of the efficient fund i whose random rate of return is r_e^i. Recall that $\dfrac{b}{a}$ is the expected rate of return of the global minimum variance portfolio.

It suffices to show that such convex combination has an expected rate of return greater than $\frac{b}{a}$ in order that the combination of funds remains to be efficient.

Since $E\left[r_e^i\right] \geq \dfrac{b}{a}$ for all i as all these funds are efficient and $w_i \geq 0$, $i = 1, 2, \ldots, n$, we have

$$\sum_{i=1}^{n} w_i E\left[r_e^i\right] \geq \sum_{i=1}^{n} w_i \frac{b}{a} = \frac{b}{a}.$$

Example

Means, variances, and covariances of the rates of return of 5 risky assets are listed:

Security	covariance, σ_{ij}					mean, $\overline{r}_i$
1	2.30	0.93	0.62	0.74	−0.23	15.1
2	0.93	1.40	0.22	0.56	0.26	12.5
3	0.62	0.22	1.80	0.78	−0.27	14.7
4	0.74	0.56	0.78	3.40	−0.56	9.02
5	−0.23	0.26	−0.27	−0.56	2.60	17.68

Recall that $\boldsymbol{w}^*$ has the following closed form solution

$$\begin{aligned}\boldsymbol{w}^* &= \frac{c - b\mu_P}{\Delta}\Omega^{-1}\mathbf{1} + \frac{a\mu_P - b}{\Delta}\Omega^{-1}\boldsymbol{\mu} \\ &= \alpha\boldsymbol{w}_g + (1-\alpha)\boldsymbol{w}_d,\end{aligned}$$

where $\alpha = (c - b\mu_P)\dfrac{a}{\Delta}$. Here, α satisfies

$$\mu_P = \alpha\mu_g + (1-\alpha)\mu_d = \alpha\left(\frac{b}{a}\right) + (1-\alpha)\frac{c}{b}.$$

We compute $\boldsymbol{w}_g^*$ and $\boldsymbol{w}_d^*$ through finding $\Omega^{-1}\mathbf{1}$ and $\Omega^{-1}\boldsymbol{\mu}$, then normalize by enforcing the condition that their weights are summed to one.

1. To find $\boldsymbol{v}^1 = \Omega^{-1}\mathbf{1}$, we solve the system of equations
$$\sum_{j=1}^{5} \sigma_{ij} v_j^1 = 1, \quad i = 1, 2, \cdots, 5.$$
Normalize the component v_i^1's so that they sum to one
$$w_i^1 = \frac{v_i^1}{\sum_{j=1}^5 v_j^1}.$$
After normalization, this gives the solution to $\boldsymbol{w}_g$. Why?

We first solve for $\boldsymbol{v}^1 = \Omega^{-1}\mathbf{1}$ and later divide $\boldsymbol{v}^1$ by the sum of components, $\mathbf{1}^T\boldsymbol{v}^1$. This sum of components is simply equal to a, where

$$a = \mathbf{1}^T\Omega^{-1}\mathbf{1} = \sum_{j=1}^{N} v_j^1.$$

2. To find $\boldsymbol{v}^2 = \Omega^{-1}\boldsymbol{\mu}$, we solve the system of equations:

$$\sum_{j=1}^{5} \sigma_{ij} v_j^2 = \overline{r}_i, \quad i = 1, 2, \cdots, 5.$$

Normalize v_i^2's to obtain w_i^2. After normalization, this gives the solution to $\boldsymbol{w}_d$. Also, $b = \mathbf{1}^T\Omega^{-1}\boldsymbol{\mu} = \sum_{j=1}^{N} v_j^2$ and $c = \boldsymbol{\mu}^T\Omega^{-1}\boldsymbol{\mu} = \sum_{j=1}^{N} \overline{r}_j v_j^2$.

security	$\boldsymbol{v}^1$	$\boldsymbol{v}^2$	$\boldsymbol{w}_g$	$\boldsymbol{w}_d$
1	0.141	3.652	0.088	0.158
2	0.401	3.583	0.251	0.155
3	0.452	7.284	0.282	0.314
4	0.166	0.874	0.104	0.038
5	0.440	7.706	0.275	0.334
mean			14.413	15.202
variance			0.625	0.659
standard deviation			0.791	0.812

Recall $\boldsymbol{v}^1 = \Omega^{-1}\mathbf{1}$ and $\boldsymbol{v}^2 = \Omega^{-1}\boldsymbol{\mu}$ so that

- sum of components in $\boldsymbol{v}^1 = \mathbf{1}^T\Omega^{-1}\mathbf{1} = a$
- sum of components in $\boldsymbol{v}^2 = \mathbf{1}^T\Omega^{-1}\boldsymbol{\mu} = b$.

Note that $\boldsymbol{w}_g = \boldsymbol{v}^1/a$ and $\boldsymbol{w}_d = \boldsymbol{v}^2/b$.

Relation between $\boldsymbol{w}_g$ and $\boldsymbol{w}_d$

Both $\boldsymbol{w}_g$ and $\boldsymbol{w}_d$ are frontier funds with

$$\mu_g = \frac{\boldsymbol{\mu}^T \Omega^{-1} \mathbf{1}}{a} = \frac{b}{a} \quad \text{and} \quad \mu_d = \frac{\boldsymbol{\mu}^T \Omega^{-1} \boldsymbol{\mu}}{b} = \frac{c}{b}.$$

Their variances are

$$\begin{aligned} \sigma_g^2 &= \boldsymbol{w}_g^T \Omega \boldsymbol{w}_g = \frac{(\Omega^{-1}\mathbf{1})^T \Omega (\Omega^{-1}\mathbf{1})}{a^2} = \frac{1}{a}, \\ \sigma_d^2 &= \boldsymbol{w}_d^T \Omega \boldsymbol{w}_d = \frac{(\Omega^{-1}\boldsymbol{\mu})^T \Omega (\Omega^{-1}\boldsymbol{\mu})}{b^2} = \frac{c}{b^2}. \end{aligned}$$

Difference in expected returns $= \mu_d - \mu_g = \dfrac{c}{b} - \dfrac{b}{a} = \dfrac{\Delta}{ab}$. Note that $\mu_d > \mu_g$ if and only if $b > 0$.

Also, difference in variances $= \sigma_d^2 - \sigma_g^2 = \dfrac{c}{b^2} - \dfrac{1}{a} = \dfrac{\Delta}{ab^2} > 0$.

Covariance of the portfolio returns for any two minimum variance portfolios

The random rates of return of u-portfolio and v-portfolio are given by

$$r_P^u = \boldsymbol{w}_u^T \boldsymbol{r} \text{ and } r_P^v = \boldsymbol{w}_v^T \boldsymbol{r},$$

where $\boldsymbol{r} = (r_1 \cdots r_N)^T$ is the random rate of return vector. First, for the two special frontier funds, $\boldsymbol{w}_g$ and $\boldsymbol{w}_d$, their covariance is given by

$$\begin{aligned}
\sigma_{gd} &= \text{cov}(r_P^g, r_P^d) = \text{cov}\left(\sum_{i=1}^N w_i^g r_i, \sum_{j=1}^N w_j^d r_j\right) \\
&= \sum_{i=1}^N \sum_{j=1}^N w_i^g w_j^d \text{cov}(r_i, r_j) \qquad \text{(bilinear property of covariance)} \\
&= \boldsymbol{w}_g^T \Omega \boldsymbol{w}_d = \left(\frac{\Omega^{-1}\mathbf{1}}{a}\right)^T \Omega \left(\frac{\Omega^{-1}\boldsymbol{\mu}}{b}\right) \\
&= \frac{\mathbf{1}^T \Omega^{-1}\boldsymbol{\mu}}{ab} = \frac{1}{a} \quad \text{since} \quad b = \mathbf{1}^T \Omega^{-1}\boldsymbol{\mu}.
\end{aligned}$$

In general, consider the two portfolios parametrized by u and v:

$$\boldsymbol{w}_u = (1-u)\boldsymbol{w}_g + u\boldsymbol{w}_d \text{ and } \boldsymbol{w}_v = (1-v)\boldsymbol{w}_g + v\boldsymbol{w}_d$$

so that

$$r_u = (1-u)r_g + ur_d \quad \text{and} \quad r_v = (1-v)r_g + vr_d.$$

The covariance of their random rates of portfolio return is given by

$$\begin{aligned}
\text{cov}(r_P^u, r_P^v) &= \text{cov}((1-u)r_g + ur_d, (1-v)r_g + vr_d) \\
&= (1-u)(1-v)\sigma_g^2 + uv\sigma_d^2 + [u(1-v) + v(1-u)]\sigma_{gd} \\
&= \frac{(1-u)(1-v)}{a} + \frac{uvc}{b^2} + \frac{u+v-2uv}{a} \\
&= \frac{1}{a} + \frac{uv\Delta}{ab^2}.
\end{aligned}$$

For any portfolio $\boldsymbol{w}_P$, we always have

$$\text{cov}(r_g, r_P) = \boldsymbol{w}_g^T \Omega \boldsymbol{w}_P = \frac{\mathbf{1}^T \Omega^{-1} \Omega \boldsymbol{w}_P}{a} = \frac{1}{a} = \text{var}(r_g).$$

Minimum variance portfolio and its uncorrelated counterpart

For any frontier portfolio u, we can find another frontier portfolio v such that these two portfolios are uncorrelated. This can be done by setting

$$\frac{1}{a} + \frac{uv\Delta}{ab^2} = 0,$$

and solve for v, provided that $u \neq 0$. Portfolio v is the uncorrelated counterpart of portfolio u.

The case $u = 0$ corresponds to $\boldsymbol{w}_g$. We cannot solve for v when $u = 0$, indicating that the uncorrelated counterpart of the global minimum variance portfolio does not exist. This observation is consistent with the result that $\text{cov}(r_g, r_P) = \text{var}(r_g) = 1/a \neq 0$, indicating that the uncorrelated counterpart of $\boldsymbol{w}_g$ does not exist.

2.4 Inclusion of the risk free asset: One-fund Theorem

Consider a portfolio with weight α for the risk free asset (say, US Treasury bonds) and $1-\alpha$ for a risky asset. The risk free asset has the deterministic rate of return r_f. The expected rate of portfolio return is

$$\overline{r}_P = \alpha \overline{r}_f + (1-\alpha)\overline{r}_j \quad \text{(note that } r_f = \overline{r}_f\text{)}.$$

The covariance σ_{fj} between the risk free asset and any risky asset j is zero since

$$E[(r_j - \overline{r}_j)\underbrace{(r_f - \overline{r}_f)}_{\text{zero}}] = 0.$$

Therefore, the variance of portfolio return σ_P^2 is

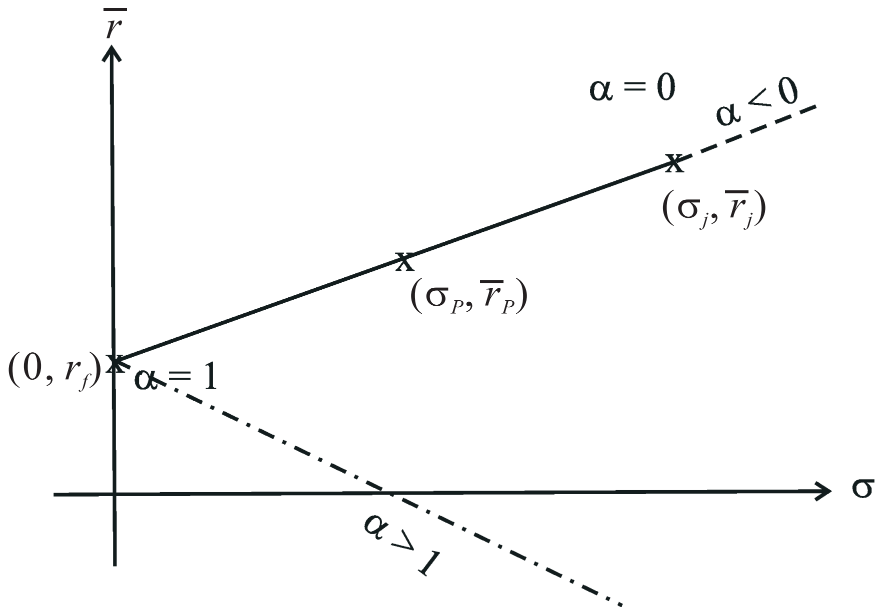
$$\sigma_P^2 = \alpha^2 \underbrace{\sigma_f^2}_{\text{zero}} + (1-\alpha)^2\sigma_j^2 + 2\alpha(1-\alpha)\underbrace{\sigma_{fj}}_{\text{zero}}$$

so that

$$\sigma_P = |1-\alpha|\sigma_j.$$

Since both $\overline{r}_P$ and σ_P are linear functions of α, so $(\sigma_P, \overline{r}_P)$ lies on a pair of line segments in the σ-$\overline{r}$ diagram. Normally, we expect $\overline{r}_j > r_f$ since an investor should expect to have expected rate of return of a risky asset higher than r_f to compensate for the risk.

1. For $0 < \alpha < 1$, the points representing $(\sigma_P, \overline{r}_P)$ for varying values of α lie on the straight line segment joining $(0, r_f)$ and $(\sigma_j, \overline{r}_j)$.



2. If borrowing of the risk free asset is allowed, then α can be negative. In this case, the line extends beyond the right side of $(\sigma_j, \overline{r}_j)$ (possibly up to infinity).

3. When $\alpha > 1$, this corresponds to short selling of the risky asset. In this case, the portfolios are represented by a line with slope negative to that of the line segment joining $(0, r_f)$ and $(\sigma_j, \overline{r}_j)$ (see the lower dotted-dashed line).

- The lower dotted-dashed line can be seen as the mirror image with respect to the vertical $\overline{r}$-axis of the upper solid line segment that would have been extended beyond the left side of $(0, r_f)$. This is due to the swapping in sign in $|1 - \alpha|\sigma_j$ when $\alpha > 1$.

- The holder bears the same risk, like long holding of the risky asset, while μ_P falls below r_f. This is highly insensible for the investor. An investor would short sell a risky asset when $\overline{r}_j < r_f$.

Consider a portfolio that starts with N risky assets originally, what is the impact of the inclusion of a risk free asset on the feasible region?

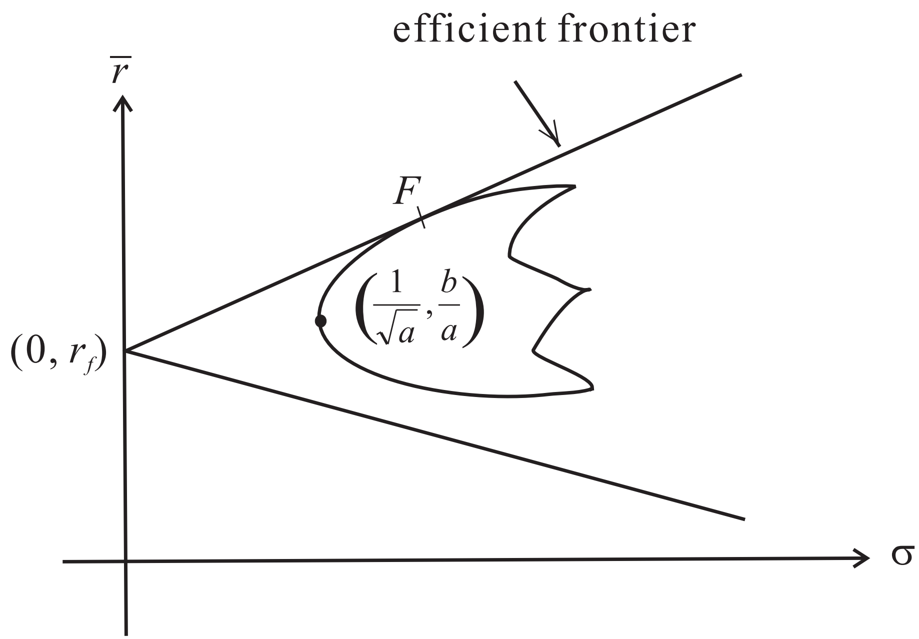
Lending and borrowing of the risk free asset is allowed

For each portfolio formed using the N risky assets, the new combinations with the inclusion of the risk free asset trace out the pair of symmetric half-lines originating from the risk free point and passing through the point representing the original portfolio.

The totality of these lines forms an infinite triangular feasible region bounded by a pair of symmetric half-lines through the risk free point, one line is tangent to the original feasible region while the other line is the mirror image about the horizontal line: $\overline{r} = r_f$. The infinite triangular wedge contains the original feasible region.

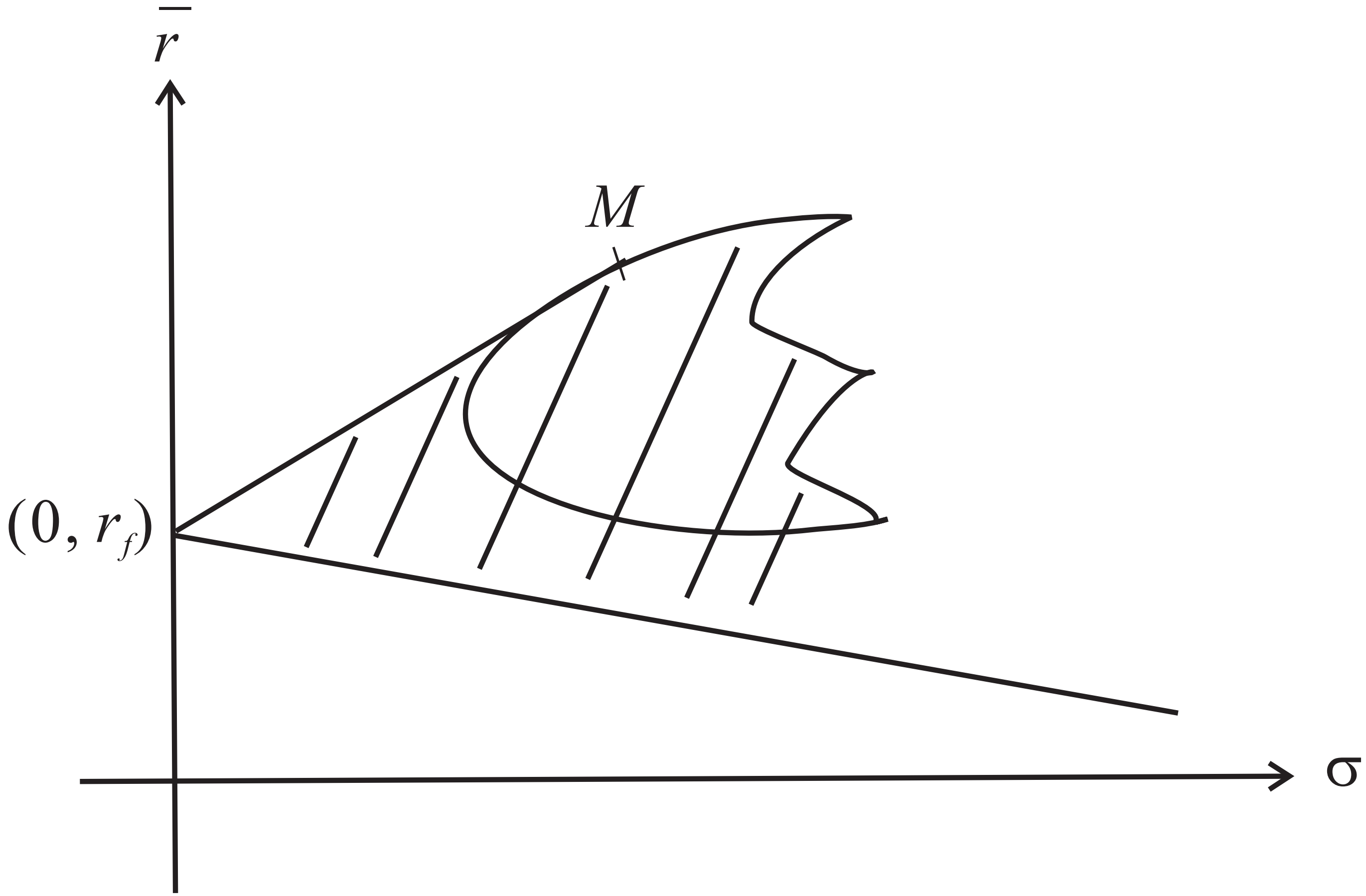
We consider the more realistic case where $r_f < \mu_g$ (a risky portfolio should demand an expected rate of return high than r_f). For $r_f < \frac{b}{a}$, the upper line of the symmetric double line pair touches the original feasible region.

The new *efficient set* is the single straight line on the top of the new triangular feasible region. This tangent line touches the original feasible region at a point F, where F lies on the efficient frontier of the original feasible set.



No shorting of the risk free asset $(r_f < \mu_g)$

The line originating from the risk free point cannot be extended beyond the points in the original feasible region (otherwise entails borrowing of the risk free asset). The upper half line is extended up to the tangency point only while the lower half line can be extended to infinity.



One-fund Theorem

Any efficient portfolio (represented by a point on the upper tangent line) can be expressed as a combination of the risk free asset and the portfolio (or fund) represented by M.

"There is a single fund M of risky assets such that any efficient portfolio can be constructed as a combination of the fund M and the risk free asset."

The One-fund Theorem is based on the assumptions that

- every investor is a mean-variance optimizer
- they all agree on the probabilistic structure of asset returns
- a unique risk free asset exists.

Then everyone purchases a single fund, which then becomes the *market portfolio*.

The proportion of wealth invested in the risk free asset is $1 - \sum_{i=1}^{N} w_i$.
Write r as the constant rate of return of the risk free asset.

Modified Lagrangian formulation

$$\text{minimize} \qquad \frac{\sigma_P^2}{2} = \frac{1}{2}\boldsymbol{w}^T \Omega \boldsymbol{w}$$

$$\text{subject to} \qquad \boldsymbol{\mu}^T \boldsymbol{w} + (1 - \boldsymbol{1}^T \boldsymbol{w}) r = \mu_P.$$

Define the Lagrangian: $L = \frac{1}{2}\boldsymbol{w}^T \Omega \boldsymbol{w} + \lambda[\mu_P - r - (\boldsymbol{\mu} - r\boldsymbol{1})^T \boldsymbol{w}]$

$$\frac{\partial L}{\partial w_i} = \sum_{j=1}^{N} \sigma_{ij} w_j - \lambda(\mu_i - r) = 0, \quad i = 1, 2, \cdots, N \tag{1}$$

$$\frac{\partial L}{\partial \lambda} = 0 \quad \text{giving} \quad (\boldsymbol{\mu} - r\boldsymbol{1})^T \boldsymbol{w} = \mu_P - r. \tag{2}$$

$(\boldsymbol{\mu} - r\boldsymbol{1})^T \boldsymbol{w}$ is interpreted as the weighted sum of the expected excess rate of return above the risk free rate r.

Remark

In the earlier mean-variance model without the risk free asset, we have

$$\sum_{j=1}^{N} w_j \overline{r}_j = \mu_P.$$

However, with the inclusion of the risk free asset, the corresponding relation is modified to become

$$\sum_{j=1}^{N} w_j (\overline{r}_j - r) = \mu_P - r.$$

In the new formulation, we now consider $\overline{r}_j - r$, which is the excess expected rate of return of asset j above the riskfree rate of return r. This is more convenient since the contribution of the riskfree asset to this excess expected rate of return is zero so that the weight of the riskfree asset becomes immaterial in the new formulation. Hence, in the current context, it is not necessary to impose the constraint that sum of weights of the risky assets equals one.

Solution to the constrained optimization model

Comparing to the earlier Markowitz model without the riskfree asset, the new formulation considers the expected rate of return above the riskfree rate of the risky assets. There is no constraint on "sum of weights of the risky assets" equals one.

The governing systems of algebraic equations are given by

$$(1):\ \Omega \boldsymbol{w}^* = \lambda(\boldsymbol{\mu} - r\mathbf{1}) \quad \text{and} \quad (2):\ \boldsymbol{w}^T(\boldsymbol{\mu} - r\mathbf{1}) = \mu_P - r.$$

Solving (1): $\boldsymbol{w}^* = \lambda \Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1})$. There is only one Lagrangian multiplier λ. As usual, we substitute into eq.(2) (constraint equation) to determine λ. This gives

$$\mu_P - r = \lambda(\boldsymbol{\mu} - r\mathbf{1})^T \Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1}) = \lambda(c - 2br + ar^2).$$

We would like to relate the target expected portfolio rate of return μ_P set by the investor and the resulting portfolio variance σ_P^2. By eliminating λ, the relation between μ_P and σ_P is given by the following pair of half lines ending at the risk free asset point $(0, r)$:

$$\begin{aligned} \sigma_P^2 &= \boldsymbol{w}^{*T} \Omega \boldsymbol{w}^* = \lambda(\boldsymbol{w}^{*T} \boldsymbol{\mu} - r\boldsymbol{w}^{*T} \mathbf{1}) \\ &= \lambda(\mu_P - r) = (\mu_P - r)^2/(c - 2br + ar^2), \end{aligned}$$

or

$$\sigma_P = \pm \frac{\mu_P - r}{\sqrt{ar^2 - 2br + c}}.$$

With the inclusion of the risk free asset, the set of minimum variance portfolios are represented by portfolios on the two half lines

$$L_{up} : \mu_P - r = \sigma_P \sqrt{ar^2 - 2br + c} \tag{3a}$$

$$L_{low} : \mu_P - r = -\sigma_P \sqrt{ar^2 - 2br + c}. \tag{3b}$$

Recall that $ar^2 - 2br + c > 0$ for all values of r since $\Delta = ac - b^2 > 0$.

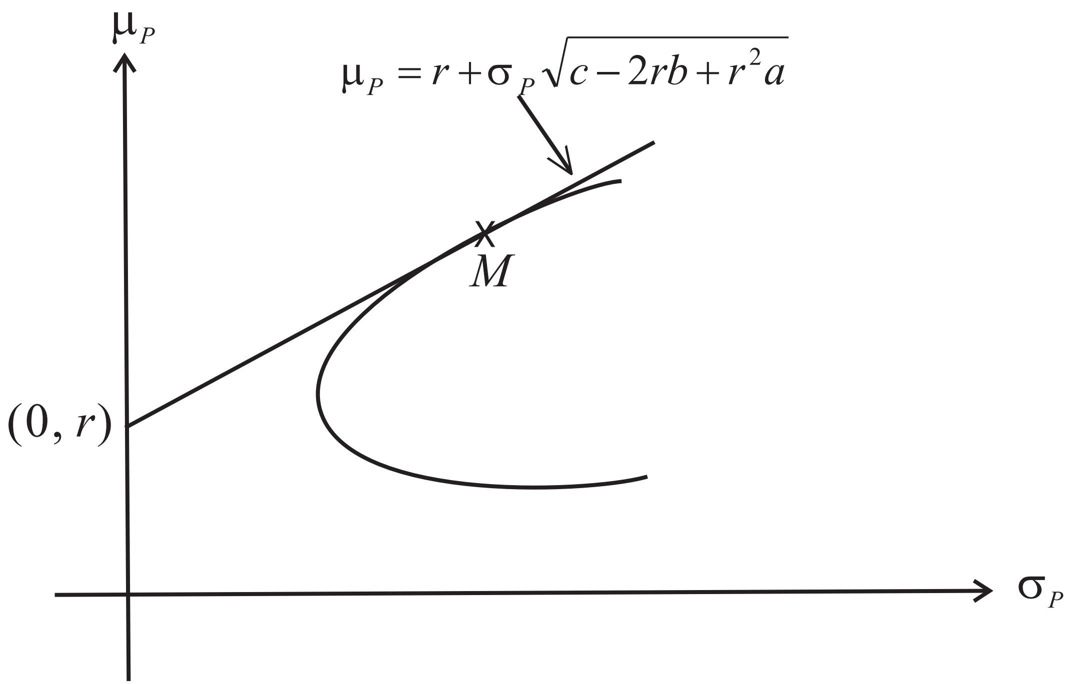
The pair of half lines give the frontier of the feasible region of the risky assets plus the risk free asset?

The minimum variance portfolios without the risk free asset lie on the hyperbola in the (σ_P, μ_P)-plane

$$\sigma_P^2 = \frac{a\mu_P^2 - 2b\mu_P + c}{\Delta}.$$

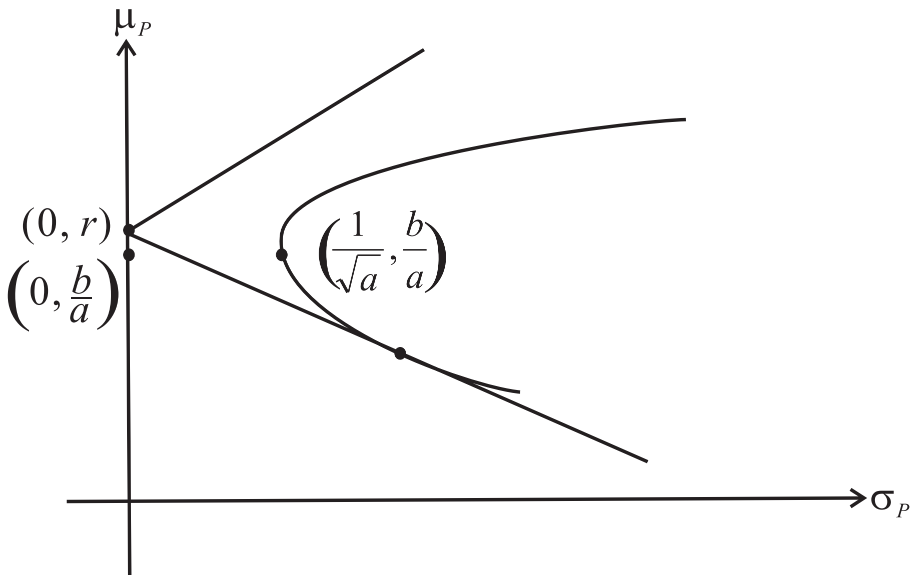
We expect $r < \mu_g$ since a risk averse investor should demand the expected rate of return from a risky portfolio to be higher than the risk free rate of return.

When $r < \mu_g = \dfrac{b}{a}$, one can show geometrically that the upper half line is a tangent to the hyperbola. The tangency portfolio is the tangent point to the efficient frontier (upper part of the hyperbolic curve) through the point $(0, r)$.



What happen when $r > \frac{b}{a}$?

The lower half line touches the feasible region with risky assets only.



- Any portfolio on the upper half line involves short selling of the tangency portfolio and investing the proceeds in the risk free asset. It makes good sense to short sell the tangency portfolio since it has an expected rate of return that is lower than the risk free asset.

Solution of the tangency portfolio when $r < \mu_g$

The tangency portfolio M is represented by the point $(\sigma_{P,M}, \mu_P^M)$, and the solution to $\sigma_{P,M}$ and μ_P^M are obtained by solving simultaneously

$$\begin{aligned} \sigma_P^2 &= \frac{a\mu_P^2 - 2b\mu_P + c}{\Delta} \\ \mu_P &= r + \sigma_P\sqrt{ar^2 - 2br + c}. \end{aligned}$$

From the first order conditions that are obtained by differentiating the Lagrangian by the control variables $\boldsymbol{w}$, we obtain

$$\boldsymbol{w}^* = \lambda\Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1}), \tag{a}$$

where λ is then determined by the constraint condition:

$$\mu_P - r = (\boldsymbol{\mu} - r\mathbf{1})^T\boldsymbol{w}. \tag{b}$$

Recall that the tangency portfolio M lies in the feasible region that corresponds to the absence of the riskfree asset, so $\mathbf{1}^T \boldsymbol{w}_M = 1$. Note that $\boldsymbol{w}_M$ should satisfy eq. (a) but eq. (b) has less relevance since μ_p^M is not yet known (not to be set as target return but has to be determined as part of the solution).

This crucial observation that $\boldsymbol{w}_M$ has zero weight on the risk free asset leads to

$$1 = \lambda_M[\mathbf{1}^T \Omega^{-1}\boldsymbol{\mu} - r\mathbf{1}^T \Omega^{-1}\mathbf{1}]$$

so that $\lambda_M = \frac{1}{b-ar}$ (provided that $r \neq \frac{b}{a}$). The corresponding μ_P^M and $\sigma_{P,M}^2$ can be determined as follows:

$$\begin{aligned}
\mu_P^M &= \boldsymbol{\mu}^T \boldsymbol{w}_M^* = \frac{1}{b-ar}(\boldsymbol{\mu}^T \Omega^{-1}\boldsymbol{\mu} - r\boldsymbol{\mu}^T \Omega^{-1}\mathbf{1}) = \frac{c-br}{b-ar}, \\
\sigma_{P,M}^2 &= \boldsymbol{w}_M^{*T} \Omega \boldsymbol{w}_M^* = \frac{1}{(b-ar)^2}(\boldsymbol{\mu} - r\mathbf{1})^T \Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1}) \\
&= \frac{ar^2 - 2br + c}{(b-ar)^2}, \\
\text{or} \quad \sigma_{P,M} &= \frac{\sqrt{ar^2 - 2br + c}}{|b-ar|}.
\end{aligned}$$

Recall $\mu_g = \dfrac{b}{a}$. When $r < \dfrac{b}{a}$, we can establish $\mu_P^M > \mu_g$ as follows:

$$\begin{aligned}\left(\mu_P^M - \frac{b}{a}\right)\left(\frac{b}{a} - r\right) &= \left(\frac{c-br}{b-ar} - \frac{b}{a}\right)\frac{b-ar}{a} \\ &= \frac{c-br}{a} - \frac{b^2}{a^2} + \frac{br}{a} \\ &= \frac{ac-b^2}{a^2} = \frac{\Delta}{a^2} > 0,\end{aligned}$$

so we deduce that $\mu_P^M > \dfrac{b}{a} > r$.

Similarly, when $r > \frac{b}{a}$, we have $\mu_p^M < \frac{b}{a} < r$.

Also, we can deduce that $\sigma_{P,M} > \sigma_g$ as expected. This is because both Portfolio M and Portfolio g are portfolios generated by the same set of risky assets (with no inclusion of the riskfree asset), and g is the global minimum variance portfolio.

Example (5 risky assets and one riskfree asset)

Data of the 5 risky assets are given in the earlier example, and $r = 10\%$.

The system of linear equations to be solved is

$$\sum_{j=1}^{5} \sigma_{ij} v_j = \overline{r}_i - r = 1 \times \overline{r}_i - r \times 1, \quad i = 1, 2, \cdots, 5.$$

Recall that $\boldsymbol{v}^1$ and $\boldsymbol{v}^2$ in the earlier example are solutions to

$$\sum_{j=1}^{5} \sigma_{ij} v_j^1 = 1 \quad \text{and} \quad \sum_{j=1}^{5} \sigma_{ij} v_j^2 = \overline{r}_i, \text{ respectively,} \quad i = 1, 2, \ldots, 5.$$

Hence, $v_j = v_j^2 - r v_j^1, j = 1, 2, \cdots, 5$ (numerically, we take $r = 10\%$).

In matrix representation, we have

$$\boldsymbol{v}_1 = \Omega^{-1}\mathbf{1} \quad \text{and} \quad \boldsymbol{v}_2 = \Omega^{-1}\boldsymbol{\mu}.$$

Now, we have obtained $\boldsymbol{v}$ where

$$\boldsymbol{v} = \Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1}) = \boldsymbol{v}_1 - r\boldsymbol{v}_2$$

Note that the optimal weight vector for the 5 risky assets satisfies

$$\boldsymbol{w} = \lambda \boldsymbol{v} \quad \text{for some scalar } \lambda.$$

We determine λ by enforcing $(\boldsymbol{\mu} - r\mathbf{1})^T \boldsymbol{w} = \mu_P - r$, or equivalently,

$$\lambda(\boldsymbol{\mu} - r\mathbf{1})^T \boldsymbol{v} = \lambda(c - 2br + ar^2) = \mu_P - r,$$

where μ_P is the target rate of return of the portfolio.

Recall $a = \mathbf{1}^T \Omega^{-1} \mathbf{1} = \sum_{j=1}^{5} v_j^1, b = \mathbf{1}^T \Omega^{-1} \boldsymbol{\mu} = \sum_{j=1}^{5} v_j^2$, and $c = \boldsymbol{\mu}^T \Omega^{-1} \boldsymbol{\mu} = \sum_{j=1}^{5} \overline{r}_j v_j^2$. We find λ by setting

$$\lambda = \frac{\mu_P - r}{ar^2 - 2br + c}.$$

The weight of the risk free asset is then given by $1 - \sum_{j=1}^{5} w_j$.

Properties of the minimum variance portfolios for $r < b/a$

1. *Efficient portfolios*

   Any portfolio on the upper half line

   $$\mu_P = r + \sigma_P \sqrt{ar^2 - 2br + c}$$

   within the segment FM joining the two points $F(0, r)$ and M involves long holding of the market portfolio M and the risk free asset F, while those outside FM involves short selling of the risk free asset and long holding of the market portfolio.

2. Any portfolio on the lower half line

   $$\mu_P = r - \sigma_P \sqrt{ar^2 - 2br + c}$$

   involves short selling of the market portfolio and investing the proceeds in the risk free asset. This represents a non-optimal investment strategy since the investor faces risk but gains no extra expected return above r.

Location of the tangency portfolio with regard to $r < b/a$ or $r > b/a$

Note that $\mu_P^M - r = \frac{ar^2-2br+c}{b-ar}$ and $\sigma_{P,M} = \frac{\sqrt{ar^2-2br+c}}{|b-ar|}$. One can show that

(i) when $r < \frac{b}{a}$, we obtain

$$\mu_P^M - r = \sigma_{P,M}\sqrt{ar^2 - 2br + c} \quad \text{(equation of } L_{up}\text{);}$$

(ii) when $r > \frac{b}{a}$, we have $|b - ar| = ar - b$ and obtain

$$\mu_P^M - r = -\sigma_{P,M}\sqrt{ar^2 - 2br + c} \quad \text{(equation of } L_{low}\text{).}$$

Interestingly, the flip of sign in $b - ar$ with respect to $r < \mu_g$ or $r > \mu_g$ would dictate whether the point $(\sigma_{P,M}, \mu_P^M)$ representing the tangency portfolio lies in the upper or lower half line, respectively.

Degenerate case occurs when $\mu_g = \dfrac{b}{a} = r$

- What happens when $r = b/a$? The pair of half lines become

$$\mu_P = r \pm \sigma_P \sqrt{c - 2\left(\frac{b}{a}\right)b + \frac{b^2}{a}} = r \pm \sigma_P \sqrt{\frac{\Delta}{a}},$$

which correspond to the asymptotes of the hyperbolic left boundary of the feasible region with risky assets only. The tangency portfolio does not exist, consistent with the mathematical result that $\lambda_M = \dfrac{1}{b - ar}$ is not defined when $r = \dfrac{b}{a}$. The tangency point $\left(\sigma_{P,M}, \mu_P^M\right) = \left(\dfrac{\sqrt{ar^2 - 2br + c}}{b - ar}, \dfrac{c - br}{b - ar}\right)$ tends to infinity when $r = \dfrac{b}{a}$, consistent with the property of the half lines being asymptotes.

- Under the scenario: $r = \dfrac{b}{a}$, efficient funds still lie on the upper half line, though the tangency portfolio does not exist.

Recall that

$$\boldsymbol{w}^* = \lambda\Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1})$$

so that the sum of weights of the risky assets is

$$\mathbf{1}^T\boldsymbol{w}^* = \lambda(\mathbf{1}^T\Omega^{-1}\boldsymbol{\mu} - r\mathbf{1}^T\Omega^{-1}\mathbf{1}) = \lambda(b - ra).$$

When $r = b/a$, sum of weights of the risky assets $= \mathbf{1}^T\boldsymbol{w}^* = 0$ as λ is finite. Since the portfolio weights are proportional dollar amounts, "sum of weight being zero" means the sum of values of risky asset held in the portfolio is zero. Any minimum variance portfolio involves investing everything in the riskfree asset and holding a zero-value portfolio of risky assets.

Suppose we specify μ_P to be the target expected rate of return of the efficient portfolio, then the multiplier λ is determined by (see p.76)

$$\lambda = \left.\frac{\mu_P - r}{c - 2br + ar^2}\right|_{r=b/a} = \frac{\mu_P - r}{c - 2\left(\frac{b}{a}\right)b + \frac{b^2}{a}} = \frac{a(\mu_P - r)}{\Delta}.$$

Financial interpretation

Given the target expected rate of portfolio return μ_P, the corresponding optimal portfolio is to hold 100% on the riskfree asset and w_j on the j^{th} risky asset, $j = 1, 2, \cdots, N$, where w_j is given by the j^{th} component of $\frac{a(\mu_P - r)}{\Delta}\Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1})$.

One should check whether the expected rate of return of the whole portfolio equals μ_P.

The expected rate of return from all the risky assets is

$$\frac{a(\mu_P - r)}{\Delta}\boldsymbol{\mu}^T[\Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1})] = \frac{a(\mu_P - r)}{ac - b^2}\left(c - \frac{b^2}{a}\right) = \mu_P - r.$$

The overall expected rate of return of the portfolio is

$$w_0 r + \sum_{j=1}^{N} w_j \overline{r}_j = r + (\mu_P - r) = \mu_P, \text{ where } w_0 = 1.$$

One-fund Theorem under $r = \mu_g = b/a$

In this degenerate case, $r = b/a$, the tangency fund does not exist. The universe of risky assets just provide an expected rate of return that is the same as the riskfree return r at its global minimum variance portfolio g. Since the global minimum variance portfolio of risky assets g has the same expected rate of return as that of the riskfree asset, a sensible investor would place 100% weight on the riskfree asset to generate the level of expected rate of return equals r.

The optimal portfolio is to invest 100% on the riskfree asset and a scalar multiple λ of the fund z whose weight vector is

$$\boldsymbol{w}_z = \Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1}).$$

The scalar λ is determined by the investor's target expected rate of return μ_P, where $\lambda = \dfrac{a(\mu_P - r)}{\Delta}$. The value of the portfolio $\boldsymbol{w}_z$ is zero. The role of the tangency fund is replaced by the z-fund.

Nature of the portfolio z: $\boldsymbol{w}_z = \Omega^{-1}(\boldsymbol{\mu} - r\mathbf{1})$, where $r = b/a$

1. Recall $\boldsymbol{w}_d = \Omega^{-1}\boldsymbol{\mu}/b$ and $\boldsymbol{w}_g = \Omega^{-1}\mathbf{1}/a$, so

$$\boldsymbol{w}_z = b\boldsymbol{w}_d - \frac{b}{a}(a\boldsymbol{w}_g) = b(\boldsymbol{w}_d - \boldsymbol{w}_g).$$

   Its sum of weights is seen to be zero since it longs b units of $\boldsymbol{w}_d$ and short the same number of units of $\boldsymbol{w}_g$.

2. Location of the portfolio z in the mean-variance plot

$$\begin{aligned}
\sigma_z^2 &= \boldsymbol{w}_z^T \Omega \boldsymbol{w}_z = b^2(\boldsymbol{w}_d - \boldsymbol{w}_g)^T \Omega (\boldsymbol{w}_d - \boldsymbol{w}_g) \\
&= b^2(\sigma_d^2 - 2\sigma_{g_d} + \sigma_g^2) = b^2\left(\frac{c}{b^2} - \frac{2}{a} + \frac{1}{a}\right) = b^2\frac{\Delta}{ab^2} = \frac{\Delta}{a}; \\
\mu_z &= \boldsymbol{\mu}^T \boldsymbol{w}_z = b(\mu_d - \mu_g) = b\left(\frac{c}{b} - \frac{b}{a}\right) = \frac{\Delta}{a}.
\end{aligned}$$

   We have $\sigma_z = \sqrt{\frac{\Delta}{a}}$ and $\mu_z = \frac{\Delta}{a}$; so any scalar multiple of z lies on the line: $\mu_P = \sqrt{\frac{\Delta}{a}}\ \sigma_P$.

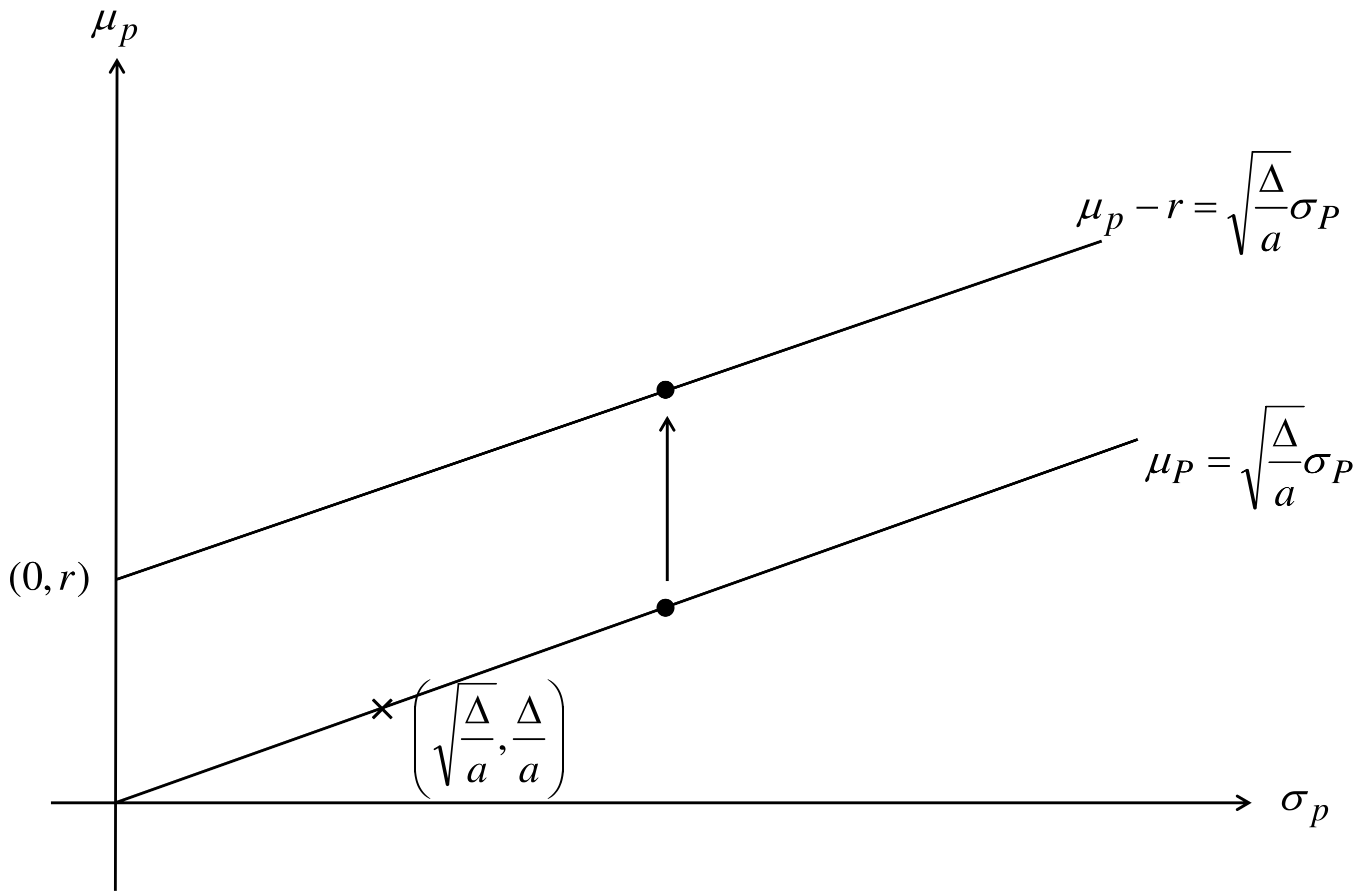
3. *Location of efficient portfolios in the mean-variance diagram*

Suppose the investor specifies her target rate of return to be μ_P. The target expected rate of portfolio return above r is produced by longing λ units of portfolio z whose sum of weights in this risky portfolio equals zero. The scalar λ is determined by setting

$$\mu_P = r + \lambda\mu_z = r + \lambda\frac{\Delta}{a} \text{ giving } \lambda = \frac{a(\mu_P - r)}{\Delta}.$$

Also, the standard deviation of the optimal portfolio's return arises only from the portfolio z, where $\sigma_P = \lambda\sigma_z = \lambda\sqrt{\frac{\Delta}{a}}$. By eliminating λ, we observe that the efficient portfolio lies on the line:

$$\mu_P - r = \sqrt{\frac{\Delta}{a}}\sigma_P.$$



The upper line represents the set of frontier funds generated by investing 100% on the riskfree asset and $\dfrac{a(\mu_P - r)}{\Delta}$ units of the z-fund. The point $\left(\sqrt{\dfrac{\Delta}{a}}, \dfrac{\Delta}{a}\right)$ on the lower line represents the z-fund, $\boldsymbol{w}_z$.

“Riskfree” portfolio of risky assets

So far, we have assumed the existence of Ω^{-1}. The corresponding global minimum portfolio has expected rate of return $\mu_g = \dfrac{b}{a}$ and portfolio variance $\sigma_g^2 = \dfrac{1}{a}$. What would happen when Ω^{-1} does not exist (or Ω is singular)?

When the covariance matrix Ω is singular, then $\det \Omega = 0$. Accordingly, there exists a non-zero vector $\boldsymbol{w}_F$ that satisfies the homogeneous system of equations:

$$\Omega \boldsymbol{w}_F = \boldsymbol{0}, \quad \boldsymbol{w}_F \neq \boldsymbol{0}.$$

Write $r_F = \boldsymbol{w}_F^T \boldsymbol{r}$, where $\boldsymbol{r} = (r_1 \ldots r_N)^T$, as the random rate of return of this F-fund. This F-fund would have zero portfolio variance since

$$\text{var}(r_F) = \boldsymbol{w}_F^T \Omega \boldsymbol{w}_F = 0.$$

This zero-variance fund can be used as a proxy of the riskfree asset. The corresponding riskfree point in the σ_P-$\overline{r}_P$ diagram would be $(0, \overline{r}_F)$, where $\overline{r}_F = \boldsymbol{w}_F^T \boldsymbol{\mu}$. We would expect to have the paradoxical scenario where $\overline{r}_F$ may not be the same as the observed riskfree rate r in the market. Assuming market efficiency where investors can take arbitrage on the difference, the two rates $\overline{r}_F$ and r would tend to each other under market equilibrium.

The Two-fund Theorem for risky assets can be interpreted as the one-fund version with this F-fund as the (proxy) riskfree asset. When Ω is singular, the parabolic arc representing the frontier now becomes a pair of half lines.

How to generate an efficient fund that lies on the upper half line? It can be done by choosing a combination of two efficient funds or combination of this F-fund with another efficient fund. In this sense, the two-fund theorem remains valid.

Tangency portfolio under One-fund Theorem and market portfolio

- The One-fund Theorem states that everyone purchases a single fund (tangency portfolio) of risky assets and borrow or lend at the riskfree rate.

- If everyone purchases the same tangency portfolio (applying the same weights on all risky assets in the market), what must that fund be? This fund is simply the *market portfolio*. In other words, if everyone buys just one fund, and their purchases add up to the market, then the proportional weights in the tangency fund must be the same as those of the market portfolio.

- In the situation where everyone follows the mean-variance methodology with the same estimates of parameters, the tangency fund of risky assets will be the market portfolio.

How can this happen? The answer is based on the equilibrium argument.

- If everyone else (or at least a large number of people) solves the problem, we do not need to. The return on an asset depends on both its initial price and its final price. The other investors solve the mean-variance portfolio problem using their common estimates, and they place orders in the market to acquire their portfolios.

- If orders placed do not match with what is available, the prices must change. The prices of the assets under heavy demand will increase while the prices of the assets under light demand will decrease. These price changes affect the estimates of asset returns directly, and hence investors will recalculate their optimal portfolio. This process continues until demand exactly matches supply, that is, it continues until an equilibrium prevails.

Summary

- In the idealized world, where every investor is a mean-variance investor and all have the same estimates, everyone buys the same portfolio and that would be the market portfolio.

- Prices adjust to drive the market to efficiency. Then after other people have made the adjustments, we can be sure that the single efficient portfolio is the market portfolio.

Market Portfolio is a portfolio consisting of a weighted sum of every asset in the market, with weights in the proportions that they exist in the market (under the assumption that these assets are infinitely divisible).

- The Hang Seng index may be considered as a proxy of the market portfolio of the Hong Kong stock market.

Appendix: Mathematical properties of covariance matrix Ω

1. It is known that the eigenvalues of a symmetric matrix are real. We would like to show that all eigenvalues of Ω are non-negative.

   If otherwise, suppose λ is a negative eigenvalue of Ω and $\boldsymbol{x}$ is the corresponding eigenvector. We have

   $$\Omega\boldsymbol{x} = \lambda\boldsymbol{x}, \quad \boldsymbol{x} \neq \boldsymbol{0},$$

   so that

   $$\boldsymbol{x}^T\Omega\boldsymbol{x} = \lambda\boldsymbol{x}^T\boldsymbol{x} < 0,$$

   a contradiction to the semi-positive definite property of Ω.

2. Ω is non-singular (Ω^{-1} exists) if and only if all eigenvalues are positive

First, we recall:

$$\det\,\Omega = \text{product of eigenvalues.}$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the n eigenvalues of Ω. Note that

$$\det\,(\Omega - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda).$$

Putting $\lambda = 0$ on both sides, we obtain $\det\, A = \lambda_1\lambda_2\cdots\lambda_n$. To show the main result, note that

$$\Omega \text{ is non-singular} \Leftrightarrow \det\,\Omega \neq 0.$$

3. Decomposition of Ω and representation of Ω^{-1} when Ω is non-singular

Let λ_i, $i = 1, 2, \ldots, n$, be the eigenvalues of Ω (allowing multiplicities) and $\boldsymbol{x}_i$ be the corresponding eigenvector of eigenvalue λ_i. Since Ω is symmetric, it has a full set of eigenvectors. We then have

$$\Omega S = S\Lambda,$$

where Λ is the diagonal matrix whose entries are the eigenvalues of Ω and S is the matrix whose columns are the eigenvectors of Ω (arranged in the corresponding sequential order). The eigenvectors are orthogonal to each other since Ω is symmetric and we can always normalize the eigenvectors to be unit length.

That is, S can be constructed to be an orthonormal matrix so that $S^{-1} = S^T$. We then have

$$\Omega = S\Lambda S^{-1} = S\Lambda S^T.$$

Provided that all eigenvalues of Ω are non-zero so that Λ^{-1} exists, we then have

$$\Omega^{-1} = (S\Lambda S^T)^{-1} = (S^T)^{-1}\Lambda^{-1}S^{-1} = S\Lambda^{-1}S^T.$$

4. $\Delta = ac - b^2 > 0$, where $\boldsymbol{\mu} \neq h\mathbf{1}$
Note that

$$\begin{aligned} a &= \mathbf{1}^T\Omega^{-1}\mathbf{1} = (\mathbf{1}^T S\Lambda^{-1/2})(\Lambda^{-1/2}S^T\mathbf{1}) \\ b &= \boldsymbol{\mu}^T\Omega^{-1}\mathbf{1} = (\boldsymbol{\mu}^T S\Lambda^{-1/2})(\Lambda^{-1/2}S^T\mathbf{1}) \\ c &= \boldsymbol{\mu}^T\Omega^{-1}\boldsymbol{\mu} = (\boldsymbol{\mu}^T S\Lambda^{-1/2})(\Lambda^{-1/2}S^T\boldsymbol{\mu}). \end{aligned}$$

We write $\boldsymbol{x} = \Lambda^{-1/2}S^T\mathbf{1}$ and $\boldsymbol{y} = \Lambda^{-1/2}S^T\boldsymbol{\mu}$ so that $a = \boldsymbol{x}^T\boldsymbol{x}$, $b = \boldsymbol{y}^T\boldsymbol{x}$ and $c = \boldsymbol{y}^T\boldsymbol{y}$.

The Cauchy-Schwarz inequality gives

$$|\boldsymbol{y}^T\boldsymbol{x}|^2 \leq (\boldsymbol{x}^T\boldsymbol{x})(\boldsymbol{y}^T\boldsymbol{y}).$$

We have equality if and only if $\boldsymbol{x}$ and $\boldsymbol{y}$ are dependent. For $\boldsymbol{\mu} \neq h\mathbf{1}$, $\boldsymbol{x}$ and $\boldsymbol{y}$ are then linearly independent, we have

$$\Delta = ac - b^2 > 0.$$

5. Singular covariance matrix. Recall that

Ω is singular $\Leftrightarrow$ the set of eigenvalues of Ω contains "zero".

That is, there exists non-zero vector $\boldsymbol{w}_0$ such that

$$\Omega \boldsymbol{w}_0 = \mathbf{0}.$$

As an example, consider the two-asset portfolio with $\rho = -1$, the corresponding covariance matrix is

$$\Omega = \begin{pmatrix} \sigma_1^2 & -\sigma_1\sigma_2 \\ -\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Obviously, Ω is singular since the columns are dependent. Accordingly, we obtain

$$\boldsymbol{w}_0 = \begin{pmatrix} \frac{\sigma_2}{\sigma_1+\sigma_2} \\ \frac{\sigma_1}{\sigma_1+\sigma_2} \end{pmatrix},$$

where $\Omega \boldsymbol{w}_0 = \mathbf{0}$ and $\boldsymbol{w}_0^T \mathbf{1} = 1$.

6. Ω^{-1} is symmetric and positive definite

Given that Ω is symmetric, where $\Omega^T = \Omega$. Consider

$$I = (\Omega^{-1}\Omega)^T = \Omega^T(\Omega^{-1})^T = \Omega(\Omega^{-1})^T,$$

implying that Ω has $(\Omega^{-1})^T$ as its inverse. Since inverse of a square matrix is unique, so $(\Omega^{-1})^T = \Omega^{-1}$.

To show the positive definite property of Ω^{-1}, it suffices to show that all eigenvalues of Ω^{-1} are all positive. Let λ be an eigenvalue of Ω, then $\lambda \boldsymbol{v} = \Omega \boldsymbol{v}$, where $\boldsymbol{v}$ is the corresponding eigenvector. We then have $\Omega^{-1}\boldsymbol{v} = \dfrac{1}{\lambda}\boldsymbol{v}$, so $\dfrac{1}{\lambda}$ is an eigenvalue of Ω^{-1}. Since all eigenvalues of Ω are positive, so do those of Ω^{-1}. Therefore, Ω^{-1} is positive definite. As a result, we observe

$$\mathbf{1}^T\Omega^{-1}\mathbf{1} = a > 0 \quad \text{and} \quad \boldsymbol{\mu}^T\Omega^{-1}\boldsymbol{\mu} = c > 0.$$