Lecture 6. Mean-Variance Portfolio Theory

1. Asset Return

Definition 1. The *total return*, denoted by R, is defined as

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

where X_0 is the amount of money invested at time 0 and X_1 is the amount of money received at time 1. The *return*, denoted by r, is defined as

$$r = \frac{X_1 - X_0}{X_0} = \frac{X_1}{X_0} - 1.$$

Remark 2. Since $R = 1 + r$, one knows $X_1 = (1+r)X_0 = RX_0$.

1.1. Short Sales

You can sell an asset that you do not own through *short selling*. In the short selling, you borrow the asset from a broker and then sell it at X_0. At a later date, you close out the short position by purchasing the same asset at X_1, and return the asset to the broker.

- If $X_0 > X_1$, you make a profit of $X_0 - X_1$. If $X_0 < X_1$, the loss is $X_1 - X_0$. Hence, the short selling is profitable if the asset price declines.

1.2. Portfolio Return

Given n different assets, you form a *portfolio* of these n assets by apportioning an initial amount X_0 among the n assets. The amount invested in ith asset, denoted by X_{0i}, is expressed as a fraction of the total investment or

$$X_{0i} = w_i X_0$$

initial amount of ith asset.

for $i = 1, \ldots n$, where w_i is the weight of asset i in the portfolio.

- Since $\sum_{i=1}^{n} X_{0i} = X_0$, it holds $\sum_{i=1}^{n} w_i = 1$. If the short selling is allowed for the ith asset, w_i is negative.

Theorem 3. *Both the total return, denoted by R_p, and the return, denoted by r_p, of a portfolio of n assets are equal to the weighed average of the corresponding individual asset returns, with the weight of an asset being its relative weight in the portfolio:*

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

개별 자산들의 total return의 가중평균

$$r_p = \sum_{i=1}^{n} w_i r_i,$$

〃 return 〃

where R_i and r_i are the total return and the return of asset i, respectively.

$(1+r_i)X_{0i}$ $\quad X_{0i} = w_i X_0$

Proof. The amount of money generated at $t=1$ by the ith asset is $X_{1i} = R_i X_{0i} = R_i w_i X_0$, and the total amount received at $t=1$ is $X_1 = \sum_{i=1}^n X_{1i} = \sum_{i=1}^n R_i w_i X_0$. Therefore, the total return of the portfolio is given by

$$R_p = \frac{X_1}{X_0} = \frac{\sum_{i=1}^n R_i w_i X_0}{X_0} = \sum_{i=1}^n R_i w_i.$$

Using $R_p = 1 + r_p$, one obtains

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

$$r_p = R_p - 1 = \sum_{i=1}^n R_i w_i - 1 = \sum_{i=1}^n R_i w_i - \sum_{i=1}^n w_i$$ □

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

2. Random Variables

2.1. *Random Variables*

A variable for which values are not known until an experiment is carried out is called a *random variable*. $\Pr(X = x)$ represents the probability that X takes a particular value of x. The set of all possible values of a random variable, with their associated probabilities, is called the *probability distribution* of the random variable. A random variable X is *discrete* if x is countable, and is *continuous* if x is not countable.

Example 4. Let X be the outcome of rolling a dice. There are six possible outcomes; i.e., $X = 1,2,3,4,5,6$. The outcome is a random variable because the actual value is not known before rolling the dice. If it is a fair dice, each outcome has an equal probability of 1/6; i.e., $\Pr(X = i) = 1/6$ for $i = 1,\ldots,6$.

Remark 5. The amount of money to be received when selling an asset at $t=1$ is uncertain at $t=0$. So, the return is random and can be described in probabilistic terms.

2.2. *Expected Value*

Definition 6. For a discrete random variable X, the *expected value* of a random variable, denoted by either $E[X]$ or μ_X, is defined as

$$E[X] = \sum_{i=1}^n p_i x_i,$$

where $p_i = \Pr(X = x_i)$ for $i = 1,\ldots,n$ and n is the number of possible outcomes.

- The expected value is the average value of the random variable in an infinite number of repetitions of the experiment. This gives a measure of the *center* of the data on a random variable.

Remark 7. If c is a constant (i.e., not random), then $E[c] = c$. If X and Y are random, then $E[a+bX] = a + bE[X]$ and $E[aX+bY] = aE[X] + bE[Y]$ for any constants a and b.

Example 8. The expected value of the outcome of rolling of a dice is given by

$$\begin{aligned} E[X] &= \sum_{i=1}^{6} p_i x_i \\ &= \Pr(X = x_1)x_1 + \Pr(X = x_2)x_2 + \Pr(X = x_3)x_3 \\ &\quad + \Pr(X = x_4)x_4 + \Pr(X = x_5)x_5 + \Pr(X = x_6)x_6 \\ &= \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 + \frac{1}{6} \times 6 \\ &= 3.5. \end{aligned}$$

Example 9. We are given the following information on two stocks, A and B, and the various weather conditions that can occur:

State of Weather	Probability (p_i)	Stock A's Return (r_A)	Stock B's Return (r_B)
Extremely Cold	0.1	-15%	35%
Cold	0.3	-5%	15%
Average	0.4	10%	5%
Hot	0.2	30%	-5%

The expected returns on each stock are given by

$$E[r_A] = 0.1(-15\%) + 0.3(-5\%) + 0.4(10\%) + 0.2(30\%) = 7\%$$

$$E[r_B] = 0.1(35\%) + 0.3(15\%) + 0.4(5\%) + 0.2(-5\%) = 9\%$$

2.3. Variance

Definition 10. The *variance* of a random variable X, denoted by either $Var[X]$ or σ_X^2, is defined as

$$\begin{aligned} Var[X] &= E[(X - \mu_X)^2] \\ &= \sum_{i=1}^{n} p_i (x_i - \mu_X)^2. \end{aligned}$$

The positive square root of σ_X^2 is called the *standard deviation* and dented by σ_X.

- The variance gives a measure of the degree of *spread* of a distribution around its mean; that is, small values imply that X is very likely to close to $E[X]$, while large values mean that X deviates greatly from $E[X]$.
- Alternatively, the variance is given by

$$Var[X] = E[X^2] - \mu_X^2,$$

since

$$\begin{aligned}\sigma_X^2 &= E[(X-\mu_X)^2] \\ &= E[X^2 - 2\mu_X X + \mu_X^2] \\ &= E[X^2] - 2\mu_X E[X] + \mu_X^2 \\ &= E[X^2] - \mu_X^2.\end{aligned}$$

$$\begin{aligned}* \ Var(aX+bY) &= Var(aX) + Var(bY) + 2Cov(aX, bY) \\ &= a^2 Var(X) + b^2 Var(Y) + 2ab\, Cov(X,Y) \\ &= a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2\end{aligned}$$

Remark 11. If c is a constant, then $Var[c] = 0$. If X is random and a and b are constants, then $Var[aX + b] = a^2 Var[X]$.

Example 12. The variances of r_A and r_B are computed as

$$\begin{aligned}Var[r_A] &= \left[0.1(-15\%)^2 + 0.3(-5\%)^2 + 0.4(10\%)^2 + 0.2(30\%)^2\right] - (7\%)^2 \\ &= 201\%^2 \\ &= 0.0201\end{aligned}$$

0.00225
0.00075
0.004
0.018
$- 0.0049 = 0.0201$

$$\begin{aligned}Var[r_B] &= \left[0.1(35\%)^2 + 0.3(15\%)^2 + 0.4(5\%)^2 + 0.2(-5\%)^2\right] - (9\%)^2 \\ &= 124\%^2 \\ &= 0.0124.\end{aligned}$$

2.4. *Covariance*

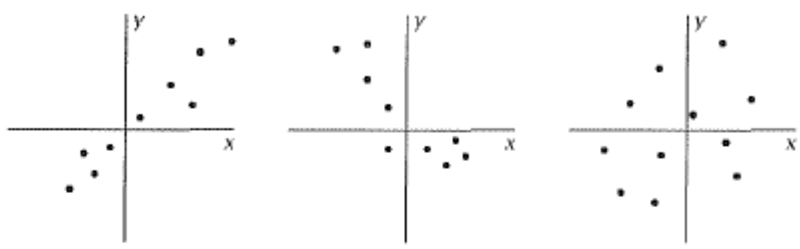
Definition 13. The *covariance* of X and Y, denoted by either $Cov[X,Y]$ or σ_{XY}, is the number defined by

$$\begin{aligned}Cov[X,Y] &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= \sum_{i=1}^{n} p_i (x_i - \mu_X)(y_i - \mu_Y),\end{aligned}$$

where $p_i = Pr(X = x_i, Y = y_i)$ represents the *joint probability* that X takes x_i and Y takes y_i.
$f_{XY}(x,y)$

- The sign of σ_{XY} gives information regarding the direction of of covariation of X and Y; that is, positive values indicate that X and Y tend to move together, while negative values imply that X tends to move in the opposite direction to Y and vice versa. When σ_{XY} is zero, then X and Y are said to be uncorrelated.

$COV = 0 \not\Rightarrow$ indep.
$\Leftarrow$



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

- Alternatively, the covariance is given by

$$Cov[X,Y] = E[XY] - E[X]E[Y],$$

since

$$\begin{aligned}\sigma_{XY} &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= E[XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y] \\ &= E[XY] - \mu_Y E[X] - \mu_X E[Y] + \mu_X\mu_Y \\ &= E[XY] - \mu_X\mu_Y.\end{aligned}$$

- By definition, it shows $Cov[X,X] = Var[X]$.

Example 14. The covariance between r_A and r_B is computed as

$$\begin{aligned}Cov[r_A, r_B] &= [0.1(-15\%)(35\%) + 0.3(-5\%)(15\%) + 0.4(10\%)(5\%) + 0.2(30\%)(-5\%)] \\ &\quad -(7\%)(9\%) \\ &= -148\%^2 \\ &= -0.0148.\end{aligned}$$

Definition 15. The *correlation* of X and Y, denoted by ρ_{XY}, is the number defined by

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}.$$

Theorem 16. *For any random variables X and Y, it shows that*

1. $-1 \le \rho_{XY} \le 1$,
2. $|\rho_{XY}| = 1$ if and only if there exists numbers $a \neq 0$ and b such that $\Pr(Y = aX + b) = 1$. If $\rho_{XY} = 1$, then $a > 0$, and if $\rho_{XY} = -1$, then $a < 0$.

Proof. Consider the function

$$\begin{aligned}h(t) &= E[((X-\mu_X)t + (Y-\mu_Y))^2] \\ &= E[(X-\mu_X)^2 t^2 + 2(X-\mu_X)(Y-\mu_Y)t + (Y-\mu_Y)^2] \\ &= \underline{\sigma_X^2 t^2 + 2\sigma_{XY} t + \sigma_Y^2.}\end{aligned} \tag{1}$$

Since $h(t)$ is the expected value of a "nonnegative" random variable, it should be greater than or equal to 0 for all t. This means that the function $h(t)$ must have a nonpositive discriminant; i.e.,

$$(2\sigma_{XY})^2 - 4\sigma_X^2\sigma_Y^2 \le 0,$$

$\sigma_{XY}^2 - \sigma_X^2 \cdot \sigma_Y^2 \le 0$

$\sigma_{XY}^2 \le \sigma_X^2\sigma_Y^2$

which is equivalent to

$$-\sigma_X\sigma_Y \le \sigma_{XY} \le \sigma_X\sigma_Y$$

$-\sigma_X\sigma_Y \le \sigma_{XY} \le \sigma_X\sigma_Y$

$-1 \le \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \le 1$

or

$$-1 \leq \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \leq 1.$$

Notice that $|\rho_{XY}| = 1$ if and only if the discriminant is equal to zero, or equivalently, $h(t)$ has a single root. Since $((X - \mu_X)t + (Y - \mu_Y))^2 \geq 0$ for all t, it shows that $h(t) = 0$ if and only if

$$\Pr(((X - \mu_X)t + (Y - \mu_Y))^2 = 0) = 1$$

or

$$\Pr((X - \mu_X)t + (Y - \mu_Y) = 0) = 1 \tag{2}$$

for all t. By writing (2) as

$$\Pr(Y = -tX + \mu_X t + \mu_Y) = 1,$$

one knows that $a = -t$ and $b = \mu_X t + \mu_Y$, where t is the root of $h(t)$. Applying the quadratic formula to (1), one obtains the root as

$-t = \frac{\sigma_{XY}}{\sigma_X^2}$ $\quad sgn(a) = \sigma_{XY}.$

$> 0.$

$$t = \frac{-2\sigma_{XY} \pm \sqrt{(2\sigma_{XY})^2 - 4\sigma_X^2\sigma_Y^2}}{2\sigma_X^2} = -\frac{\sigma_{XY}}{\sigma_X^2} (\because (2\sigma_{XY})^2 = 4\sigma_X^2\sigma_Y^2 \text{ if } |\rho_{XY}| = 1),$$

thereby proving that a has the same sign of σ_{XY}. □

- The correlation is always between -1 and 1, with the values -1 and 1 indicating a perfect *linear* relationship between X and Y. For instance, $\rho_{XY} = 1$ implies that X and Y are perfectly linearly positively correlated; i.e., X and Y differ only by some multiple and/or constant, or $Y = aX + b$, where $a > 0$ and b are constants. In this case, knowing the value of X will *exactly* reveal the value of Y. When $\sigma_{XY} = 0$, or equivalently, $\rho_{XY} = 0$, X and Y are said to be *independent*. In this case, knowledge of the value of one variable gives no information about the other.

Example 17. The correlation of r_A and r_B is computed as

$$\rho_{R_A R_B} = \frac{-148\%^2}{\sqrt{201\%^2} \times \sqrt{124\%^2}} = -0.937.$$

Theorem 18. *If X and Y are any two random variables and a and b are any two constants, then*

$$Var[aX + bY] = a^2 Var[X] + b^2 Var[Y] + 2abCov[X, Y].$$

Proof. By definition, we have

$$\begin{aligned} Var[aX + bY] &= E\left[(aX + bY)^2\right] - (a\mu_X + b\mu_Y)^2 \\ &= E[a^2X^2 + 2abXY + b^2Y^2] - (a^2\mu_X^2 + 2ab\mu_X\mu_Y + b^2\mu_Y^2) \\ &= a^2\left(E[X^2] - \mu_X^2\right) + 2ab\left(E[XY] - \mu_X\mu_Y\right) + b^2\left(E[Y^2] - \mu_Y^2\right) \\ &= a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2. \end{aligned}$$

□

- If X and Y are independent random variables, then indep $\Rightarrow \sigma_{XY}=0$.

$$Var[aX+bY] = a^2Var[X]+b^2Var[Y].$$

Example 19. The variance of r_A+r_B is given by

$$\begin{aligned} Var[r_A+r_B] &= Var[r_A]+2Cov[r_A,r_B]+Var[r_B] \\ &= 0.0201+2(-0.0148)+0.0124 \\ &= 0.0029. \end{aligned}$$

3. Multivariate Distributions

The term *random vector* applies to a vector whose elements are random variables. Let x be an $n\times 1$ random vector:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Definition 20. A *mean vector* of x, denoted by either $E[x]$ or μ, is defined as

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} = \begin{bmatrix} E[x_1] \\ E[x_2] \\ \vdots \\ E[x_n] \end{bmatrix}.$$

$\underbrace{E[x]}_{\text{vector}} = \mu$

$\underbrace{Var[x]}_{\text{vector}} = \underbrace{\Sigma}_{\text{matrix}}$

covariance matrix.

The *variance matrix* of x, denoted by either $Var[x]$ or Σ, is defined as

$$\begin{aligned} \Sigma &= E[(x-\mu)(x-\mu)'] \\ \begin{bmatrix} x_1-\mu_1 \\ \vdots \\ x_n-\mu_n \end{bmatrix}\begin{bmatrix} x_1-\mu_1 & \cdots & x_n-\mu_n \end{bmatrix} &= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}, \end{aligned}$$

$\underbrace{E[a'x]}_{\text{scalar}} = a'\mu$

$\underbrace{Var[a'x]}_{\text{scalar}} = a'\Sigma a$

$\underbrace{E[Ax]}_{\text{vector}} = A\mu$

$\underbrace{Var[Ax]}_{\text{vector}} = \underbrace{A\Sigma A'}_{\text{matrix}}$

where $\sigma_{ij} = Cov[x_i,x_j] = E[(x_i-\mu_i)(x_j-\mu_j)]$ and $\sigma_{ii} = \sigma_i^2$.

- Let a be an $n\times 1$ vector or $a = [a_1,a_2,\ldots,a_n]'$. It shows

$a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

$a' = [a_1 \cdots a_n]$

$$\begin{aligned} \underbrace{E[a'x]}_{\text{scalar}} &= E[a_1x_1+\cdots+a_nx_n] \\ &= a_1E[x_1]+\cdots+a_nE[x_n] \\ &= a_1\mu_1+\cdots+a_n\mu_n \\ &= \underbrace{a'\mu}_{\text{scalar}} \end{aligned}$$

* scalar → scalar

vector → vector (μ)

→ matrix (Σ)

and

scalar $\quad E[a'x] = a'\mu$

$$
\begin{aligned}
Var[a'x] &= E[(a'x - E[a'\mu])(a'x - E[a'\mu])'] \\
&= E[(a'x - a'\mu)(a'x - a'\mu)'] \\
&= E[a'(x-\mu)(x-\mu)'a] \\
&= a'E[(x-\mu)(x-\mu)']a \\
&= a'\Sigma a \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \sigma_{ij}.
\end{aligned}
$$

$(a'x - a'\mu)' = (a'(x-\mu))'$
$= (x-\mu)'a$
$(\because (ab)^T = b^T a^T)$

$[a_1 \cdots a_n]\begin{bmatrix}\sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn}\end{bmatrix}\begin{bmatrix}a_1 \\ \vdots \\ a_n\end{bmatrix}$ scalar

Remark 21. Given an $n \times n$ matrix A, it shows

$$
\begin{aligned}
E[Ax] &= A\mu \\
Var[Ax] &= A\Sigma A'.
\end{aligned}
$$

$Var[Ax] = E[(Ax - E[Ax])(Ax - E[Ax])']$
$= E[(Ax - A\mu)(Ax - A\mu)']$
$= E[A(x-\mu)(x-\mu)'A']$
$= AE[(x-\mu)(x-\mu)']A'$
$= A\Sigma A'$

$Ax = \begin{bmatrix}a_{11} & a_{12} \\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}x_1 \\ x_2\end{bmatrix} = \begin{bmatrix}a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2\end{bmatrix}$

$E[Ax] = \begin{bmatrix}E[a_{11}x_1 + a_{12}x_2] \\ E[a_{21}x_1 + a_{22}x_2]\end{bmatrix}$
$= \begin{bmatrix}a_{11}E[x_1] + a_{12}E[x_2] \\ a_{21}E[x_1] + a_{22}E[x_2]\end{bmatrix} = A\mu$

Definition 22. Let the vector x be the set of n random variables, μ their mean vector and Σ their variance matrix. For $x \sim N(\mu, \Sigma)$, the general form of the joint density is given by

joint pdf

$$f(x) = (2\pi)^{-n/2}|\Sigma|^{-1/2}e^{(-1/2)(x-\mu)'\Sigma(x-\mu)}.$$

$\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}e^{-\frac{(x-\mu)'\Sigma(x-\mu)}{2}}$

Remark 23. If $x \sim N(\mu, \Sigma)$, each element of x is also normally distributed; i.e.,

mgf로 증명,

$$x_k \sim N(\mu_k, \Sigma_{kk}) \quad \sigma_k^2$$

$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

for $k = 1, \ldots, n$, where Σ_{kk} is the kth diagonal element of Σ. If $x \sim N(\mu, \Sigma)$, it shows

matrix $\quad$ vector

$$Ax + b \sim N(A\mu + b, A\Sigma A').$$

4. Portfolio Mean and Variance

4.1. Mean and Variance

Suppose that there are n assets with random returns $r_1, \ldots, r_n$. Let r be an $n \times 1$ vector of the returns r_i and w be an $n \times 1$ vector of the weights w_i, $i = 1, \ldots, n$. Then, the return on a portfolio of these n assets is given by

$$r_p = \sum_{i=1}^{n} w_i r_i = w'r.$$

- The expected return of the portfolio is given by

$E[r] = \begin{bmatrix}E[r_1] \\ \vdots \\ E[r_n]\end{bmatrix} = \begin{bmatrix}\mu_1 \\ \vdots \\ \mu_n\end{bmatrix} = \mu$

$$
\begin{aligned}
E[r_p] &= E[w'r] \\
&= w'\mu \\
&= \sum_{i=1}^{n} w_i \mu_i,
\end{aligned}
$$

$E[r_p] = [w_1 \cdots w_n]\begin{bmatrix}\mu_1 \\ \vdots \\ \mu_n\end{bmatrix}$

where μ represents an $n \times 1$ vector of $\mu_i = E[r_i]$, $i = 1, \ldots, n$.

Def 22. Multivariate Normal Distribution.

$$x = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \sim MN\left(\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \underset{n\times n}{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & & \sigma_{2n} \\ \vdots & & \ddots & \vdots \\ \sigma_{n1} & & & \sigma_n^2 \end{bmatrix} \right)$$

covariance matrix.

joint pdf.

$$f_{X_1 \cdots X_n}(x_1, \cdots, x_n) = \frac{1}{\sqrt{|2\pi\Sigma|}} \cdot \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right).$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right).$$

$$x_k \sim N(\mu_k, \Sigma_{kk}), \quad Ax+b \sim N(A\mu+b, A\Sigma A').$$

- The variance of the return of the portfolio is given by

$$
\begin{aligned}
Var[r_p] &= Var[w'r] = E[(w'r - E[w'r])(w'r - E[w'r])'] \\
&= w'\Sigma w = E[w'(r-\mu)(r-\mu)'w] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}, \quad = w'E[(r-\mu)(r-\mu)']w \\
&\qquad = w'Var(r)w \\
&\qquad = w'\Sigma w
\end{aligned}
$$

where Σ is an $n \times n$ matrix of $\sigma_{ij} = Cov[r_i, r_j]$, $i, j = 1, \ldots, n$.

Example 24. Suppose that there are two assets, r_1 and r_2. Given w_1 and w_2, a portfolio is formed as $r_p = w_1 r_1 + w_2 r_2$. The portfolio mean is given by

$$
\begin{aligned}
E[r_p] &= E[w_1 r_1 + w_2 r_2] \\
&= w_1\mu_1 + w_2\mu_2,
\end{aligned}
$$

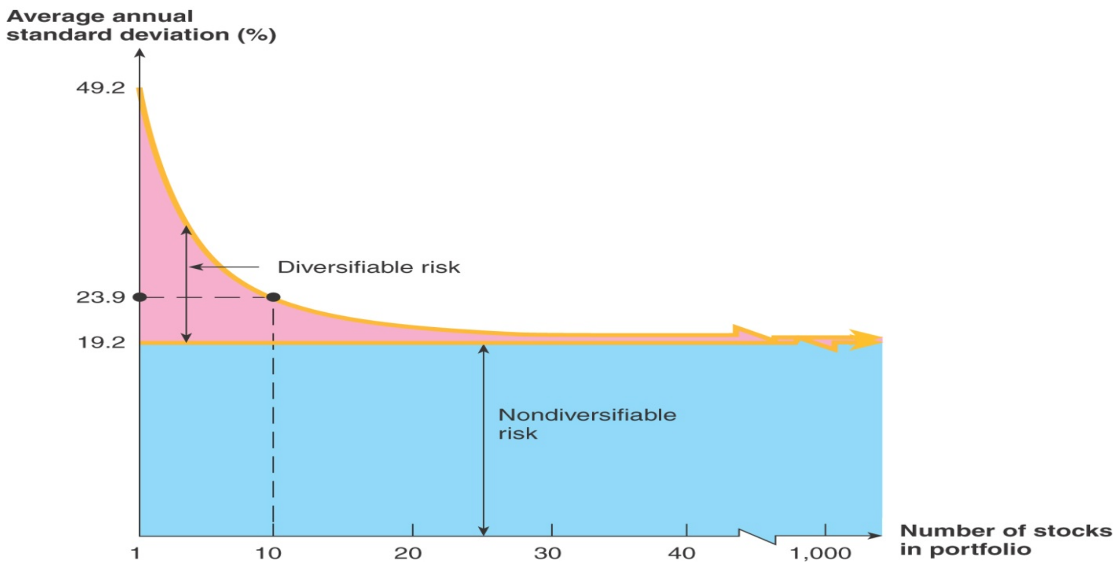
$Var(aX+bY)$
$= a^2\sigma_X^2 + 2ab\,\sigma_{XY} + b^2\sigma_Y^2$

and the portfolio variance is given by

$$
\begin{aligned}
Var[r_p] &= Var[w_1 r_1 + w_2 r_2] \\
&= w_1^2\sigma_1^2 + 2w_1 w_2 \sigma_{12} + w_2^2\sigma_2^2 \\
&= w_1^2\sigma_1^2 + 2w_1 w_2 \sigma_1\sigma_2\rho_{12} + w_2^2\sigma_2^2.
\end{aligned}
$$

4.2. Diversification $\sigma(r_p) = w_1\sigma_1 + w_2\sigma_2$ (if $\rho_{12} = 1$).

The variance of the portfolio return can be reduced by including additional assets in the portfolio, a process referred to as *diversification*. The eliminated portion of risk is called *nonsystematic* or *diversifiable* risk. The minimum level of risk that cannot be diversified away is called *systematic* or *nondiversifiable* risk.



Suppose that a portfolio is constructed by taking equal portions of n assets; i.e., $w_i = 1/n$ for $i =$

$1, \ldots, n$. Then, the variance of the portfolio return is

$$\begin{aligned} Var[r_p] &= w' \Sigma w \\ &= \left(\frac{1}{n} i\right)' \Sigma \left(\frac{1}{n} i\right) \\ &= \frac{1}{n^2} i' \Sigma i, \quad = \frac{1}{n^2} \underbrace{\sum_i \sum_j \sigma_{ij}}_{\text{모든 원소 더한 값.}} \qquad (3) \end{aligned}$$

n이 커질수록 $Var[rp]$가 낮아짐.

where $i = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}'$. For simplicity, assume that $\sigma_i^2 = \sigma^2$ for $i = 1, \ldots, n$.

for $\forall i$, $Var(r_i) = \sigma^2$ $\Rightarrow \Sigma = \begin{bmatrix} \sigma^2 & & \sigma_{ij} \\ & \ddots \ \sigma^2 & \\ \sigma_{ij} & & \ddots \ \sigma^2 \end{bmatrix}$

- If all assets are uncorrelated (i.e., $\sigma_{ij} = 0$ for all $i \neq j$), then (3) is written as

$$\Sigma = \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix}$$

$$\begin{aligned} Var[r_p] &= \frac{1}{n^2}(n\sigma^2) \\ &= \frac{\sigma^2}{n}. \end{aligned}$$

So, the portfolio variance diminishes to zero as n increases; namely, it is possible through diversification to reduce the portfolio variance to zero by taking n large.

- If each return pair has the same positive covariance (i.e., $Cov[r_i, r_j] = \delta\sigma^2 > 0$ for all $i \neq j$), then (3) is written as

$$\Rightarrow \Sigma = \begin{bmatrix} \sigma^2 & & \delta\sigma^2 \\ & \ddots & \\ \delta\sigma^2 & & \sigma^2 \end{bmatrix}$$

$$\begin{aligned} Var[r_p] &= \frac{1}{n^2}\left(n\sigma^2 + (n^2 - n)\delta\sigma^2\right) \\ &= \frac{\sigma^2}{n} + \left(1 - \frac{1}{n}\right)\delta\sigma^2. \end{aligned}$$

So, the portfolio variance diminishes to $\delta\sigma^2$, not zero, as n increases; namely, diversification cannot fully reduce the portfolio variance and there is a minimum level of risk that cannot be diversified away.

4.3. Diagram of a Portfolio

Different return-standard deviation combinations are available for a portfolio by changing the weights in the assets. All of such combinations are called the *portfolio opportunity set*. In a *mean-standard deviation diagram*, the horizon axis is used for the standard deviation, and the vertical axis is used for the mean.

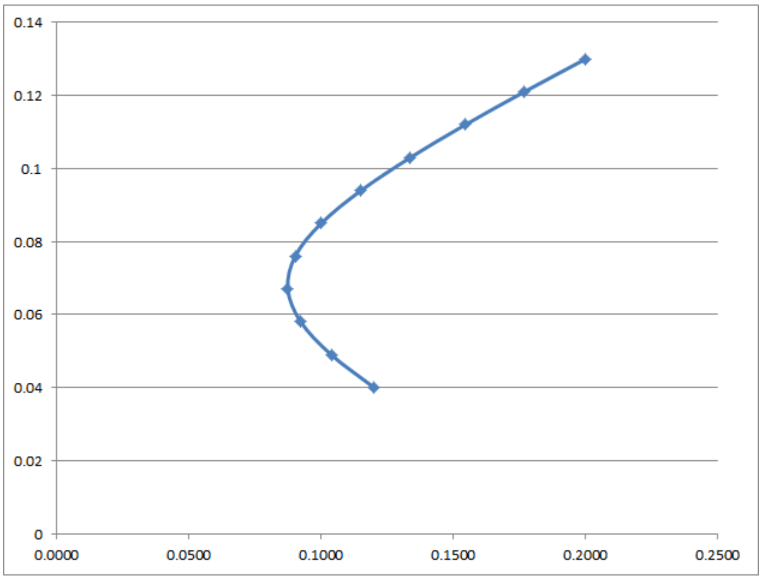
Example 25. Consider a portfolio of two assets with $\mu_1 = 0.04$, $\mu_2 = 0.13$, $\sigma_1 = 0.12$, $\sigma_2 = 0.2$, and $\rho_{12} = -0.3$.

μ는 주어짐.

→ w에 대한 함수

w_1	w_2	$w_1\mu_1$	$w_2\mu_2$	$E[r_p]$	$w_1^2\sigma_1^2$	$w_2^2\sigma_2^2$	$2w_1w_2\sigma_1\sigma_2\rho_{12}$	$Var[r_p]$	σ_p
0	1	0.0000	0.1300	0.1300	0.0000	0.0400	0.0000	0.0400	0.2000
0.1	0.9	0.0040	0.1170	0.1210	0.0001	0.0324	-0.0013	0.0312	0.1768
0.2	0.8	0.0080	0.1040	0.1120	0.0006	0.0256	-0.0023	0.0239	0.1545
0.3	0.7	0.0120	0.0910	0.1030	0.0013	0.0196	-0.0030	0.0179	0.1337
0.4	0.6	0.0160	0.0780	0.0940	0.0023	0.0144	-0.0035	0.0132	0.1151
0.5	0.5	0.0200	0.0650	0.0850	0.0036	0.0100	-0.0036	0.0100	0.1000
0.6	0.4	0.0240	0.0520	0.0760	0.0052	0.0064	-0.0035	0.0081	0.0902
0.7	0.3	0.0280	0.0390	0.0670	0.0071	0.0036	-0.0030	0.0076	0.0874
0.8	0.2	0.0320	0.0260	0.0580	0.0092	0.0016	-0.0023	0.0085	0.0923
0.9	0.1	0.0360	0.0130	0.0490	0.0117	0.0004	-0.0013	0.0108	0.1038
1	0	0.0400	0.0000	0.0400	0.0144	0.0000	0.0000	0.0144	0.1200

$E(r_p)$



$Var(r_p)$.

4.4. Math for the Portfolio Opportunity Set

σ_{12} $\sigma_{12}/\sigma_1\cdot\sigma_2$

Since the covariance (or correlation) is included in the portfolio variance, there are many the portfolio opportunity sets available, depending on the covariance of the asset returns.

- Given a portfolio of two assets, the portfolio return is $r_p = wr_1 + (1-w)r_2$, so that the standard deviation of the portfolio return, denoted by σ_p, is stated as

$$\sigma_p = \sqrt{w^2\sigma_1^2 + 2w(1-w)\sigma_1\sigma_2\rho_{12} + (1-w)^2\sigma_2^2}.$$

$Var[r_p] = Var[wr_1 + (1-w)r_2]$
$= w^2\sigma_1^2 + 2w(1-w)\sigma_{12} + (1-w)^2\sigma_2^2$
$= w^2\sigma_1^2 + 2w(1-w)\sigma_1\sigma_2\rho_{12} + (1-w)^2\sigma_2^2$

- When $\rho_{12} = 1$, we have

$$\begin{aligned}\sigma_p &= \sqrt{w^2\sigma_1^2 + 2w(1-w)\sigma_1\sigma_2 + (1-w)^2\sigma_2^2} \\ &= \sqrt{(w\sigma_1 + (1-w)\sigma_2)^2} \\ &= w\sigma_1 + (1-w)\sigma_2.\end{aligned}$$

$\sqrt{x^2} = |x|$
$(\because x$가 양수$)$.
x.

In this case, the standard deviation of the portfolio equals the weighted average of the component standard deviations, thereby meaning that no diversification occurs.

σ_p 가 각 σ_i의 가중평균으로 결정되어 number of asset (diversification)과 관계X.

- When $\rho_{12} = -1$, we have

$$\begin{aligned}\sigma_p &= \sqrt{w^2\sigma_1^2 - 2w(1-w)\sigma_1\sigma_2 + (1-w)^2\sigma_2^2} \\ &= \sqrt{(w\sigma_1 - (1-w)\sigma_2)^2} \\ &= |w\sigma_1 - (1-w)\sigma_2|.\end{aligned}$$

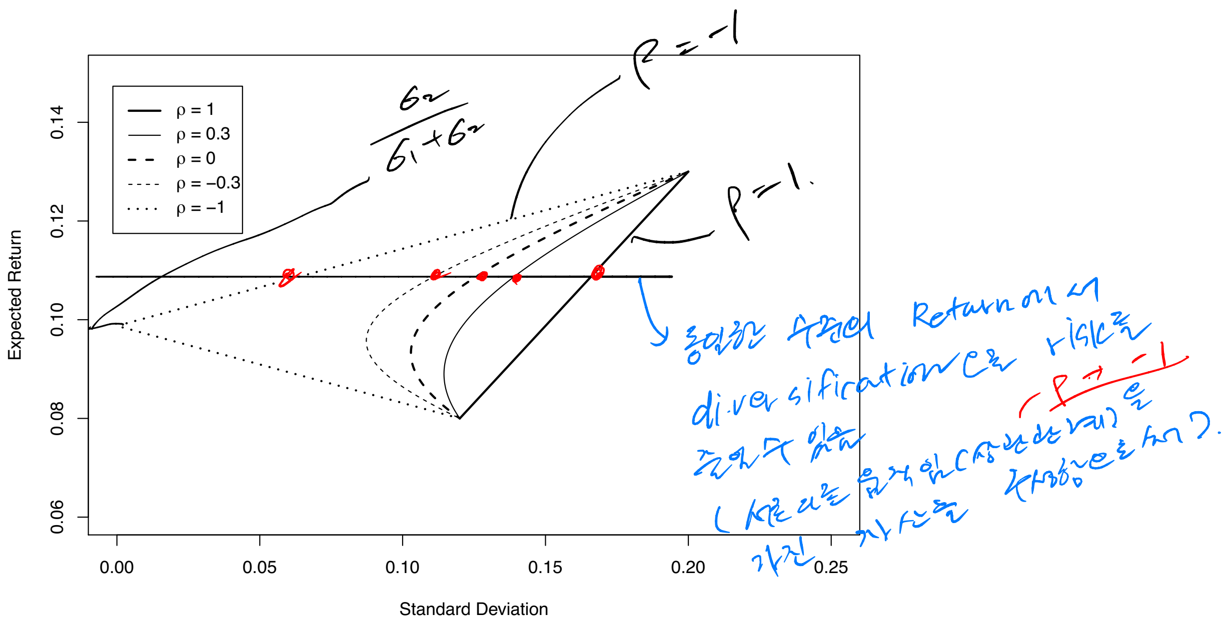
In this case, the portfolio standard deviation is less than the weighted average of the component standard deviations:

$$\sigma_p < w\sigma_1 + (1-w)\sigma_2,$$

thereby meaning that diversification occurs. $|w\sigma_1 - (1-w)\sigma_2| \leq w\sigma_1 + (1-w)\sigma_2$

- When $-1 \leq \rho_{12} < 1$, the portfolio standard deviation is *less than or equal to* the weighted average of component standard deviations. Thus, the portfolio of less than perfectly positively correlated ($\rho = 1$) assets always offers better risk-return opportunities than the individual component securities on their own. This is the benefit from diversification.

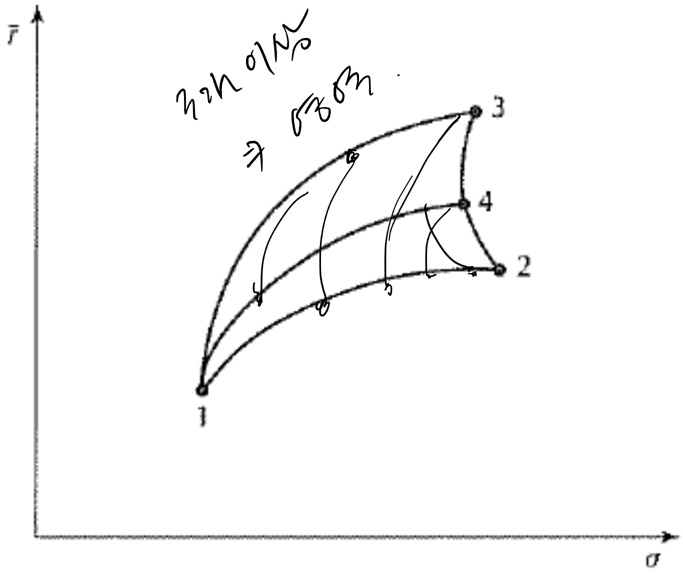
Example 26. Consider a portfolio of two assets with $\mu_1 = 0.08$, $\mu_2 = 0.13$, $\sigma_1 = 0.12$ and $\sigma_2 = 0.2$.



5. Feasible Set

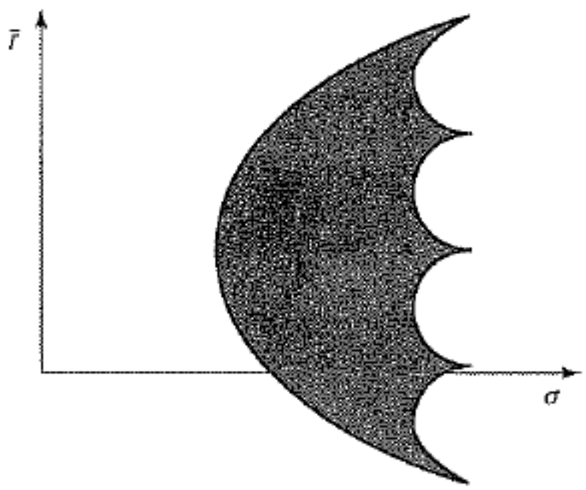
Suppose that there are n assets and you form portfolios using every possible weighting scheme. The resulting set of points that correspond to the constructed portfolios is called the *feasible set*.

- When $n \geq 3$, the feasible set will be a solid two-dimensional region.

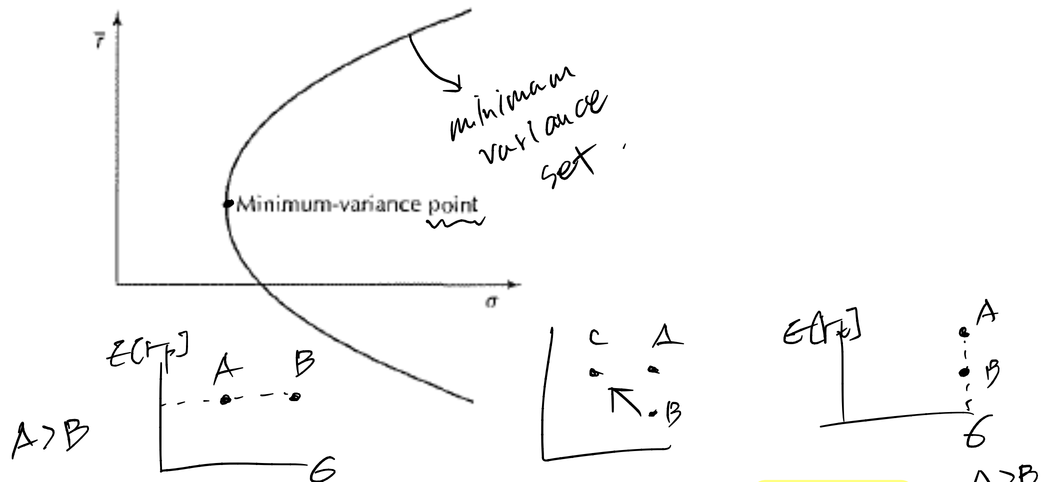


convex set은 이유

- The feasible set is convex to the left, meaning that the straight line connecting any two points does not cross the left boundary of the feasible set.

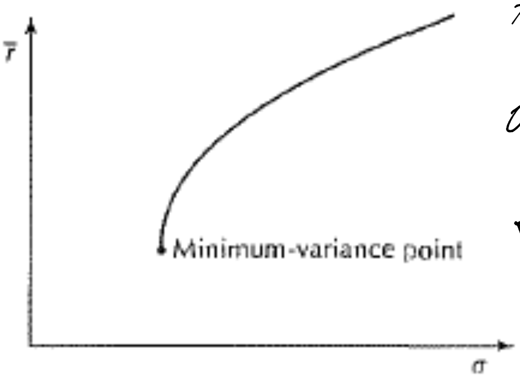


- The left boundary of a feasible set is called the *minimum-variance set* which contains the point with lowest possible variance for any value of the portfolio mean return. The *minimum-variance point* refers to a point on the minimum-variance set having minimum variance.



- It is assumed that an investor is *risk averse* (i.e., prefer less risk to more risk) and has a *non-satiated* utility function (i.e, prefer more money to less money). Then, she chooses (a) the portfolio with the smallest standard deviation for the given mean or (b) the portfolio with the largest mean for the given standard deviation. Put differently, the upper part of the minimum-variance set will be of interest to investors who are risk averse and satisfy non-satiation. This upper part of the minimum-variance set is termed the *efficient frontier* of the feasible region.

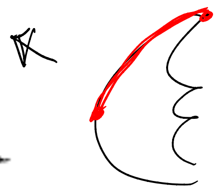
수익이 같을 때 작은 Risk 선호.



가치이 없어야 c를 선택.

대부분은 가치에 부합

choice의 문제는 이 가정 위에서 전개.



6. Optimization with Equality Constraints

Given the following function

$$z = f(x_1, x_2),$$

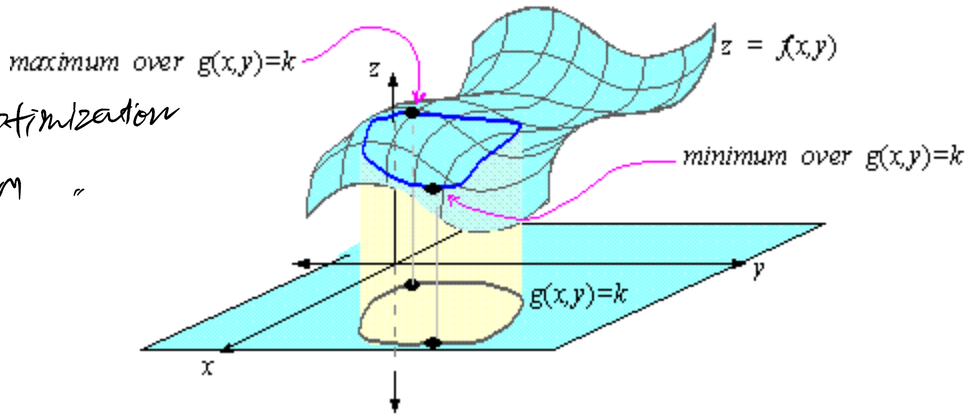
a local maximum occurs at (x_1^*, x_2^*) satisfying (a) $f_1 = f_2 = 0$ and (b) $d^2z < 0$. A local extremum is found by treating that x_1 is *independent* of x_2 and vice versa. Now one imposes an *equality* constraint on x_1 and x_2,

$$g(x_1, x_2) = c. \quad \rightarrow \text{제약조건.}$$

Then, the choices of x_1 and x_2 become *mutually dependent* since the constraint restricts the domain and narrows the range of the objective function.

제약조건이 없을 때 x_1, x_2 선택에 서로 독립적이였다면,
〃 있을 경우 (선택에) $g(x_1,x_2)=c$ 를 만족해야 하므로 서로 dependent 하다.

Example 27. Local extrema of $z = f(x,y)$ subject to $g(x,y) = k$:



포트폴리오 optimization
→ 제약조건 하에 〃

— f, g : differentiable and f′, g′ continuous.

Theorem 28. *Let $f, g \in C^1$. Suppose that $(x_1^*, \ldots, x_n^*)$ is a critical point of $z = f(x_1, \ldots, x_n)$ subject to $g(x_1, \ldots, x_n) = c$ and $g_i \neq 0$ for $i = 1, \ldots, n$ at $(x_1^*, \ldots, x_n^*)$. Then there exists a number λ^* such that $(x_1^*, \ldots, x_n^*, \lambda^*)$ is a critical point of*

$$\mathscr{L} = f(x_1, \ldots, x_n) + \lambda \left[c - g(x_1, \ldots, x_n) \right].$$

- The theorem implies that at $(x_1^*, \ldots, x_n^*, \lambda^*)$ we have

$$\begin{aligned} \mathscr{L}_\lambda &= c - g(x_1, \ldots, x_m) = 0 \\ \mathscr{L}_i &= f_i - \lambda g_i = 0 \text{ for } i = 1, \ldots, n. \end{aligned}$$

Example 29. Find the extremum of $z = x_1^2 + x_2^2$ subject to $x_1 + 4x_2 = 2$. Since the Lagrange function is given by

$$\mathscr{L} = x_1^2 + x_2^2 + \lambda(2 - x_1 - 4x_2),$$

then there exists $(x_1^*, x_2^*, \lambda^*)$ satisfying a system of equations such as

$$\begin{aligned} 2 - x_1 - 4x_2 &= 0 \\ 2x_1 - \lambda &= 0 \\ 2x_2 - 4\lambda &= 0. \end{aligned}$$

Solving the equations, we have

$$\lambda^* = \frac{4}{17},\ x_1^* = \frac{2}{17},\ x_2^* = \frac{8}{17}$$

and the extremum of z is calculated as

$$z^* = \left(\frac{2}{17}\right)^2 + \left(\frac{8}{17}\right)^2 = \frac{4}{17}.$$

Theorem 30. *Let $f, g^j \in C^1$ for $j = 1, \ldots, m$. Suppose that $(x_1^*, \ldots, x_n^*)$ is a critical point of $z = f(x_1, \ldots, x_n)$ subject to $g^1(x_1, \ldots, x_n) = c_1, \ldots, g^m(x_1, \ldots, x_n) = c_m$ and $g_i^j \neq 0$ for $i = 1, \ldots, n$ and $j =$*

$1,\ldots,m$ at $(x_1^*,\ldots,x_n^*)$. Then there exists m numbers $\lambda_1^*,\ldots,\lambda_m^*$ such that $(x_1^*,\ldots,x_n^*,\lambda_1^*,\ldots,\lambda_m^*)$ is a critical point of

$$\mathscr{L} = f(x_1,\ldots,x_n) + \sum_{j=1}^{m} \lambda_j \left[c_j - g^j(x_1,\ldots,x_n)\right].$$

- The theorem implies that at $(x_1^*,\ldots,x_n^*,\lambda_1^*,\ldots,\lambda_m^*)$ we have

$$\begin{aligned}\frac{\partial \mathscr{L}}{\partial \lambda_j} &= c_j - g^j(x_1,\ldots,x_m) = 0 \text{ for } j = 1,\ldots,m \\ \frac{\partial \mathscr{L}}{\partial x_i} &= \frac{\partial f}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g^j}{\partial x_i} = 0 \text{ for } i = 1,\ldots,n.\end{aligned}$$

Remark 31. The constraint qualification requires $m < n$; in words, the number of constraints should be less than the number of choice variables. Otherwise, the optimization problem becomes trivial.

7. Matrix and Calculus

A system of m linear equations in n variables $x_{1,}x_{2,}\ldots,x_n$ can be arranged into such a format

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n. \end{aligned}$$

This system of equations is simplified as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

or

$$y = Ax.$$

- The matrix A is of dimension $m \times n$, and can be written as

$$A = \begin{bmatrix} a_1^{'} \\ a_2^{'} \\ \vdots \\ a_m^{'} \end{bmatrix},$$

  where the $1 \times n$ vector of $a_i^{'}$ is the ith row of A.

Definition 32. For a function $y = f(x) = f(x_1, x_2, \ldots, x_n)$, a *gradient* is defined as

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \partial y/\partial x_1 \\ \partial y/\partial x_2 \\ \vdots \\ \partial y/\partial x_n \end{bmatrix}$$

- Since each element y_i of y is

$$y_i = a_i^{'} x = \sum_{j=1}^{n} a_{ij} x_j,$$

the gradient of y_i is computed as

$$\frac{\partial a_i^{'} x}{\partial x} = a_i$$

$$\frac{\partial y_i}{\partial x} = \frac{\partial (a_i^{'} x)}{\partial x} = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{bmatrix} = a_i.$$

So, the gradient of y_i is the transpose of ith row of A.

Theorem 33. *For a quadratic form $x^{'}Ax$, it holds that*

$$\frac{\partial x^{'}Ax}{\partial x} = \begin{cases} 2Ax & \text{if } A \text{ is symmetric} \\ (A + A^{'})x & \text{if } A \text{ is not symmetric.} \end{cases}$$

- Collecting $\partial y_i/\partial x$ for $i = 1, \ldots, n$, one obtains

$$\begin{aligned} \frac{\partial Ax}{\partial x} &= \begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} & \cdots & \frac{\partial y_m}{\partial x} \end{bmatrix} \\ &= \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} \\ &= A^{'}. \end{aligned}$$

Example 34. Consider

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix}$$

and

$$x^{'}Ax = x_1^2 + 4x_2^2 + 6x_1x_2.$$

Then

$$\frac{\partial x^{'}Ax}{x} = \begin{bmatrix} 2x_1 + 6x_2 \\ 6x_1 + 8x_2 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2Ax.$$

시험범위.

8. Markowitz Model

Suppose that there are n assets for forming a portfolio. Let the n risky assets have the $n \times 1$ mean vector μ and the $n \times n$ variance matrix Σ:

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}.$$

pf of Theorem 33).

quadratic form : $x^T A x$.

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

$$x^T A x = (x_1 \ x_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (a_{11}x_1 + a_{21}x_2)x_1 + (a_{12}x_1 + a_{22}x_2)x_2$$

$$= a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2$$

$$\frac{\partial(x^T A x)}{\partial x} = \begin{pmatrix} \frac{\partial(x^T A x)}{\partial x_1} \\ \frac{\partial(x^T A x)}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2a_{11}x_1 + (a_{12} + a_{21})x_2 \\ (a_{12} + a_{21})x_1 + 2a_{22}x_2 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_1 + a_{11}x_1 + a_{12}x_2 + a_{21}x_2 \\ a_{12}x_1 + a_{21}x_1 + a_{22}x_2 + a_{22}x_2 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} + \begin{pmatrix} a_{11}x_1 + a_{21}x_2 \\ a_{12}x_1 + a_{22}x_2 \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= Ax + A^T x = \underline{(A + A^T)x}$$

If A is symmetric matrix,
then $a_{12} = a_{21}$. (i.e. $A = A^T$).

$$\therefore \frac{\partial(x^T A x)}{\partial x} = \underline{2Ax}.$$

Given the $n \times 1$ vector of portfolio weights $w = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}'$, the portfolio p has mean return $w'\mu$ and variance $w'\Sigma w$.

Definition 35. When short selling is allowed, the portfolio p is the *minimum-variance portfolio* of all portfolios with mean μ_p if its portfolio weight vector is the solution to

$$\min_w w'\Sigma w$$

subject to

$$w'\mu = \mu_p$$
$$w'i = 1. \qquad \sum_i w_i = 1$$

To solve this minimization problem, write the Lagrangian function as

$$\mathscr{L} = w'\Sigma w + 2\lambda_1(\mu_p - w'\mu) + 2\lambda_2(1 - w'i).$$

The first-order condition $\mathscr{L}_w = 0$ is

$$2\Sigma w - 2\lambda_1\mu - 2\lambda_2 i = 0 \qquad \frac{\partial(w'\Sigma w)}{\partial w} = 2\Sigma w$$

$$\Sigma w = \lambda_1\mu + \lambda_2 i$$

or

$$w = \Sigma^{-1}(\lambda_1\mu + \lambda_2 i). \qquad \frac{\partial(w'\mu)}{\partial w} = \mu \tag{4}$$

The conditions $\mathscr{L}_{\lambda_1} = 0$ and $\mathscr{L}_{\lambda_2} = 0$ suggest

$$\mu'w = \mu_p \tag{5}$$
$$i'w = 1. \tag{6}$$

Using (4), one writes (5) and (6) as

$$\begin{aligned} \mu'w &= \mu'\Sigma^{-1}(\lambda_1\mu + \lambda_2 i) \\ &= \mu'\Sigma^{-1}\mu\lambda_1 + \mu'\Sigma^{-1}i\lambda_2 \\ &= \mu_p \end{aligned} \tag{7}$$

$$\mu'\Sigma^{-1}(\lambda_1\mu + \lambda_2 i) = \mu'\Sigma^{-1}\mu\lambda_1 + \mu'\Sigma^{-1}i\lambda_2 = \mu_p$$

and

$$i'\Sigma^{-1}(\lambda_1\mu + \lambda_2 i) = i'\Sigma^{-1}\mu\lambda_1 + i'\Sigma^{-1}i\lambda_2 = 1$$

$$\begin{aligned} i'w &= i'\Sigma^{-1}(\lambda_1\mu + \lambda_2 i) \\ &= i'\Sigma^{-1}\mu\lambda_1 + i'\Sigma^{-1}i\lambda_2 \\ &= 1, \end{aligned} \tag{8}$$

$$\begin{bmatrix} \mu_p \\ 1 \end{bmatrix} = \begin{bmatrix} \mu'\Sigma^{-1}\mu & \mu'\Sigma^{-1}i \\ i'\Sigma^{-1}\mu & i'\Sigma^{-1}i \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

respectively. This system of equation (7) and (8) is simplified as

$$\begin{array}{cc} A & B \\ \end{array}$$

$$\begin{bmatrix} \mu'\Sigma^{-1}\mu & \mu'\Sigma^{-1}i \\ i'\Sigma^{-1}\mu & i'\Sigma^{-1}i \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix}.$$

$$\begin{array}{cc} B' & C \\ \end{array}$$

$\mu'\Sigma^{-1}i$

$= (\mu_1\ \mu_2)\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$= \mu_1\sigma_{11} + \mu_2\sigma_{21} + \mu_1\sigma_{12} + \mu_2\sigma_{22}$

Therefore, (7) gives

$$\begin{aligned} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} &= \begin{bmatrix} A & B \\ B & C \end{bmatrix}^{-1} \begin{bmatrix} \mu_p \\ 1 \end{bmatrix} \\ &= \frac{1}{AC - B^2} \begin{bmatrix} C & -B \\ -B & A \end{bmatrix} \begin{bmatrix} \mu_p \\ 1 \end{bmatrix} \\ &= \frac{1}{AC - B^2} \begin{bmatrix} C\mu_p - B \\ -B\mu_p + A \end{bmatrix}, \end{aligned}$$

$i'\Sigma^{-1}\mu$

$= (1\ 1)\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$

$= \mu_1\sigma_{11} + \mu_1\sigma_{21} + \mu_2\sigma_{12} + \mu_2\sigma_{22}$

Σ is invertible symmetric matrix, $\therefore\ \sigma_{12} = \sigma_{21}$.

where $A = \mu'\Sigma^{-1}\mu$, $B = \mu'\Sigma^{-1}i$, and $C = i'\Sigma^{-1}i$. Finally, the optimal weight is computed as

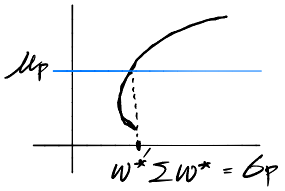
$$\begin{aligned} w &= \Sigma^{-1}(\lambda_1\mu + \lambda_2 i) \\ &= \Sigma^{-1}\left(\frac{\mu(C\mu_p - B) + i(A - B\mu_p)}{AC - B^2}\right). \end{aligned}$$

Remark 36. With the optimal weight w^*, the variance of the minimum-variance portfolio is computed as $\sigma_p = w^{*'}\Sigma w^*$ for the given value of mean μ_p. One computes all possible combinations of (σ_p, μ_p) by taking different values of μ_p, which consequently leads to the minimum-variance set.

주어진 μ_p에서 최소의 $w'\Sigma w$를 찾는다.
모든 μ_p에서 찾아주면 minimum-variance set이 된다.

9. Inclusion of a Risk-Free Asset

A risk-free asset has a deterministic (확정된) return; that is, its variance is zero. Consider the portfolio of one risky stock and one risk-free asset. In this case, the portfolio return is



$$\begin{aligned} r_p &= wr_A + (1 - w)r_f \\ &= r_f + w(r_A - r_f), \end{aligned}$$

where r_A is the risky stock return and r_f is the risk-free return.

- Inclusion of the risk-free asset in a portfolio implies *lending* (빌려주다) or *borrowing* (빌리다) cash at the risk-free rate. Lending corresponds to the risk-free asset having a positive weight, while borrowing corresponds to having a negative weight.

  risk-free asset을 갖고 있는 것.

Definition 37. The *risk premium* (RP) is defined by the excess return required for an investment in a risky asset A over that required rate for a risk-free investment:

$$RP_A = E[r_A] - r_f.$$

- The RP is a measure of the reward of bearing *additional risk*; i.e., the greater the risk, the higher the RP will be.

$r_p = r_f + w(r_A - r_f)$

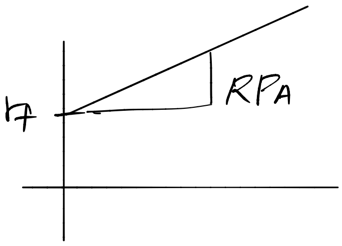
The portfolio expected return is given by

$E[r_p] = r_f + w(E[r_A] - r_f)$

$E[r_p]$

$$
\begin{aligned}
E[r_p] &= r_f + w(E[r_A] - r_f) \qquad = r_f + w \cdot RP_A \\
&= r_f + wRP_A.
\end{aligned}
\tag{9}
$$

- The base rate of return is r_f when $w = 0$ (i.e., all money is invested in the risk-free asset). Over the base rate, the portfolio is expected to earn wRP_A that depends on (a) a risk premium of the risky stock, RP_A, and (b) the investor's position in the risky stock, w.



The portfolio variance is given by

$$
\begin{aligned}
Var[r_p] &= Var[r_f + w(r_A - r_f)] \\
&= w^2 Var[r_A],
\end{aligned}
$$

$\sigma_p = w\,\sigma_A$

so that the standard deviation σ_p is

σ_p

$$
\sigma_p = w\sigma_A. \qquad \sigma_p^2 = w^2 \sigma_A^2 \tag{10}
$$

- The risk of portfolio in terms of σ_p is determined by (a) the risky stock, σ_A, and (b) the investor's position in the risky stock, w.

w에는 의지가 반영.
여기에는 X. 포트폴리오에
직접적 반영 X.

Using $w = \sigma_p/\sigma_A$ in (10), the portfolio expected return in (9) can be rewritten as

σp가 증가할때
E[rp]가 얼마나 증가?

$E[r_p] = r_f + S_A \cdot \sigma_p$.

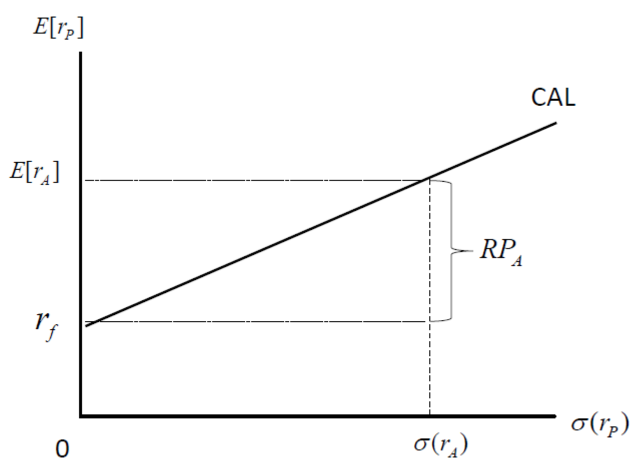
$w = \dfrac{\sigma_p}{\sigma_A}$

$$
\begin{aligned}
E[r_p] &= r_f + \left(\frac{E[r_A] - r_f}{\sigma_A}\right)\sigma_p \\
&= r_f + S_A \cdot \sigma_p.
\end{aligned}
\qquad S_A = \frac{E[r_A] - r_f}{\sigma_A}
\tag{11}
$$

$= \dfrac{RP_A}{\sigma_A}$

- The *reward-to-variability ratio* S_A represents the increase in the expected portfolio return per unit of additional standard deviation of the portfolio. This suggests the expected return-standard deviation trade-off; i.e., as σ_p increases by 1%, $E[r_p]$ increases by S_A% (since $S_A > 0$ theoretically).

From (11), the feasible set of the portfolio of one risky asset and one risk-free asset is depicted as follows:



- The *capital allocation line* (CAL) depicts all the risk-return combinations available to investors. Depending on the value of w, there are four cases; (a) if $w = 0$, $E[r_p] = r_f$, (b) if $w = 1$, $E[r_p] = r_A$, (c) if $0 < w < 1$, $E[r_p] = wE[r_A] + (1-w)r_f$, and (d) if $w > 1$, leverage (i.e., borrowing) occurs. The leverage with the CAL means that one borrows a certain amount at the risk-free rate and invests that in the risky stock. For instance, suppose you have \$100 and $w = 1.2$. This means that you borrow 20% of \$100 at the risk-free rate and invest \$120 of total investment funds in the risky stock A.

It is common that investors lend at the rate lower than the rate at which they can borrow; i.e., $r_f^L < r_f^B$, where r_f^L is the lending rate and r_f^B is the borrowing rate.

- For lending (i.e., $0 \le w < 1$),

$$\begin{aligned} E[r_p] &= r_f^L + \left(\frac{E[r_A] - r_f^L}{\sigma_A}\right)\sigma_p \\ &= r_f^L + S_A^L \cdot \sigma_p. \end{aligned}$$

For borrowing (i.e., $w > 1$),

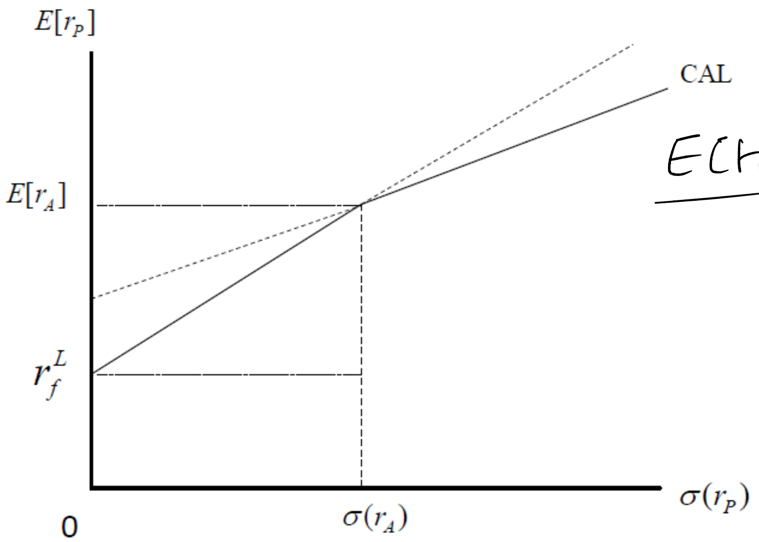
$$\begin{aligned} E[r_p] &= r_f^B + \left(\frac{E[r_A] - r_f^B}{\sigma_A}\right)\sigma_p \\ &= r_f^B + S_A^B \cdot \sigma_p. \end{aligned}$$

$0 \le w < 1$ $\quad w > 1$

$r_f^L < r_f^B.$

$S_A^L > S_A^B$

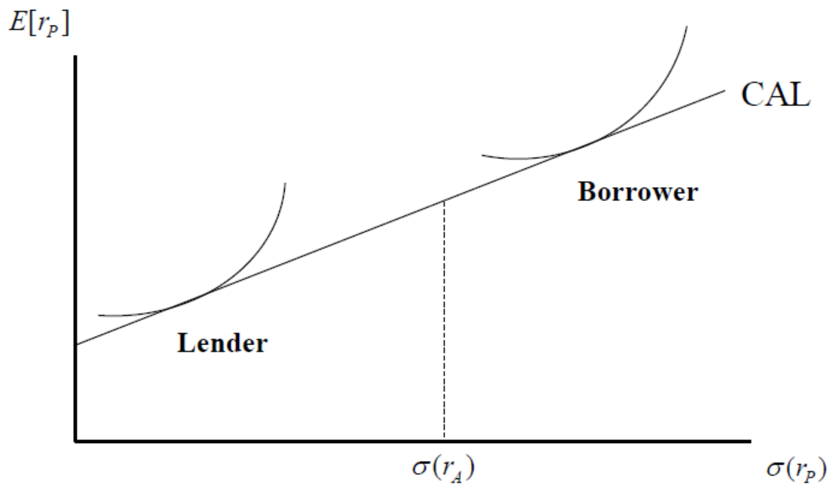
It is straightforward to see that $S_A^L > S_A^B$.



$E[r_p] = r_f + S_A \cdot \sigma_p.$

Investors confronting the CAL choose the optimal portfolio from the set of feasible choices. Individual investors with different degree of risk-aversion make different decisions.

- Greater levels of risk-aversion lead to holding larger proportions of the risk-free rate, while lower levels of risk-aversion lead to holding larger proportions of the portfolio of risky assets.



10. Optimal Portfolios of n Risky Assets and One Risk-Free Asset

Suppose that there is a portfolio of n stocks, denoted by p. You plan to form a new portfolio, denoted by c, using (a) the portfolio of n risky stocks and (b) a risk-free asset. In this case, the portfolio return is given by

$$\begin{aligned} r_c &= (1-w)r_f + wr_p \\ &= r_f + w(r_p - r_f). \end{aligned}$$

- It shows that

$$\begin{aligned} E[r_c] &= r_f + w(E[r_p] - r_f) \\ \sigma_c &= w\sigma_p. \end{aligned}$$

Using $w = \sigma_c/\sigma_p$, the expected return on the new portfolio c is given by

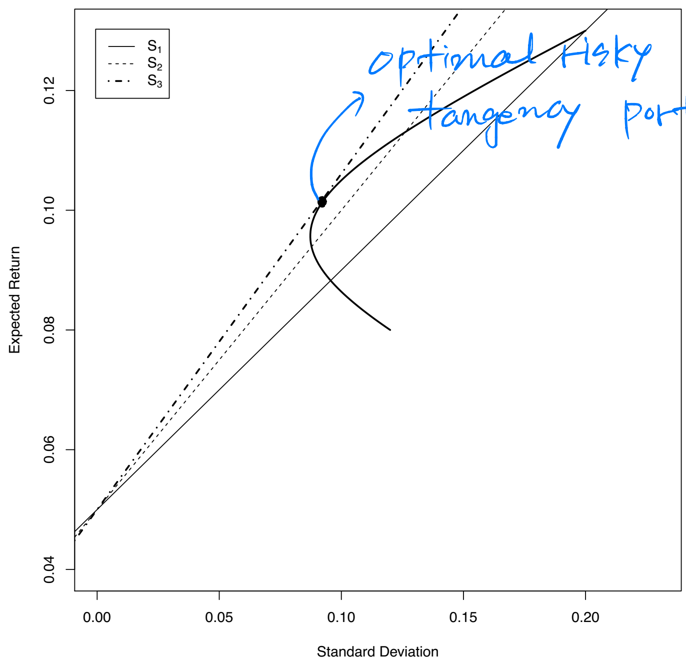
$$\begin{aligned} E[r_c] &= r_f + \left(\frac{E[r_P] - r_f}{\sigma_p}\right)\sigma_c \\ &= r_f + S_p \cdot \sigma_c. \end{aligned}$$

Example 38. Consider the portfolio of two risky assets with $\mu_1 = 0.08$, $\mu_2 = 0.13$, $\sigma_1 = 0.12$, and $\sigma_2 = 0.2$. Now, let's introduce a risk-free asset with $r_f = 0.05$.

$$\begin{aligned} r_c &= (1-w) r_f + w r_p \\ &= r_f + w(r_p - r_f) \end{aligned}$$

$$\begin{aligned} E[r_c] &= r_f + w(E[r_p] - r_f) \\ &= r_f + \underbrace{\left(\frac{E[r_p] - r_f}{\sigma_p}\right)}_{S_p} \cdot \sigma_c \end{aligned}$$

$$\sigma_c = w \cdot \sigma_p$$

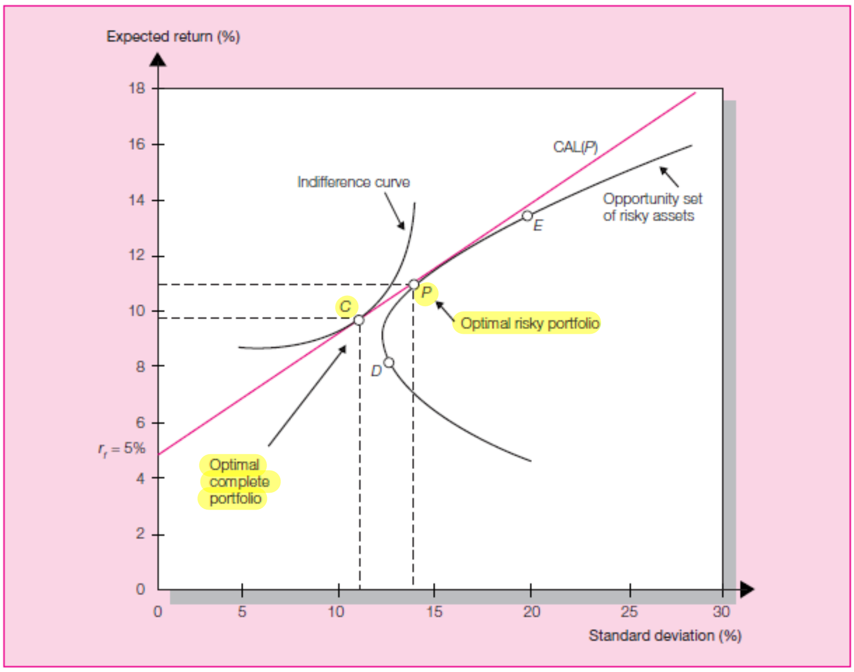


optimal risky portfolio
tangency portfolio.

- As the reward-to-variability ratio S_p increases, the CAL is ratcheted upward and finally reaches the point of tangency with the investment opportunity set of two stocks. The tangency portfolio, denoted by p, is the *optimal risky portfolio* to mix with the risk-free asset in that the CAL has the highest feasible reward-to-variability ratio. $S_p = \frac{E[r_p]-r_f}{\sigma_p}$

가장 큰 Sp

Given the optimal risky portfolio p, one uses the individual investor's degree of risk-aversion, represented in terms of indifference curves, to calculate the optimal proportion of the *complete portfolio c* to invest in the risky component.



- Suppose that one picks c as the optimal complete portfolio. To obtain the optimal complete portfolio, w_f of your funds is invested in the risk-free asset and $1-w_f$ is invested in the optimal risky portfolio p.

According to the *separation property*, the portfolio choice problem is separated into two *independent* tasks, such as (a) determination of the optimal risky portfolio p, performed by *portfolio managers* and (b) allocation of the complete optimal portfolio c to the risk-free asset r_f and the optimal risky portfolio p, decided by *clients*.

- Importantly, a portfolio manager will offer the same optimal risky portfolio p to all clients regardless of their degree of risk aversion. The degree of risk aversion of the client comes into play only in the selection of the desired points along the CAL. Thus, the first task and the second task are effectively separated.

11. Markowitz Model with a Risk-Free Asset

Consider forming a portfolio of n risky assets and a risk-free asset. Given a risk-free asset return r_f, the minimum-variance portfolio for the given expected return μ_p is the solution to

$$\min_w w'\Sigma w$$

subject to

$$-2\lambda(1-w'i)r_f$$

$$w'\mu + (1-w'i)r_f = \mu_p.$$

$$2\lambda w' i r_f$$

- With the risk-free asset, the portfolio weights of the risky assets are "not" constrained to sum to 1, since $(1-w'i)$ can be invested in the risk-free asset.

The Lagrangian function is

$$\mathcal{L}_w = 2\Sigma w - 2\lambda\mu + 2\lambda i r_f = 0$$
$$= 2\Sigma w + 2\lambda(-\mu + r_f i)$$

$$\mathcal{L} = w'\Sigma w + 2\lambda(\mu_p - w'\mu - (1-w'i)r_f),$$

and the first-order conditions of $\mathcal{L}_w$ and $\mathcal{L}_\lambda$ are given by

$$w'\mu + r_f - w' i r_f = \mu_p.$$

$$2\Sigma w + 2\lambda(-\mu + r_f i) = 0 \quad (12)$$

$$w'(\mu - r_f i) = \mu_p - r_f$$

$$w'\mu + (1-w'i)r_f = \mu_p, \quad (13)$$

respectively.

$$\Sigma w = -\lambda(-\mu + r_f i)$$

- One writes (12) as

$$w = \lambda\Sigma^{-1}(\mu - r_f i)$$

$$w = \lambda\Sigma^{-1}(\mu - r_f i) \quad (14)$$

$$w' = \lambda(\mu - r_f i)'\Sigma^{-1}$$

 and (13) as

$$w'(\mu - r_f i) = \mu_p - r_f. \quad (15)$$

Plugging (14) into (15) leads to

$$\lambda(\mu - r_f i)'\Sigma^{-1}(\mu - r_f i) = \mu_p - r_f,$$

which gives

$$\lambda = \frac{\mu_p - r_f}{(\mu - r_f i)'\Sigma^{-1}(\mu - r_f i)}.$$

Finally, one obtains w in (14) as a function of μ_p: i.e.,

$$w = \frac{(\mu_p - r_f)}{(\mu - r_f i)'\Sigma^{-1}(\mu - r_f i)}\Sigma^{-1}(\mu - r_f i). \tag{16}$$

The portfolio weights w produce all the portfolios for the given mean μ_p, which lie on the steepest CAL, in the sense that w of funds is invested in n risky assets and the remaining $(1-w)$ is invested in the risk-free asset.

The portfolio weights w in (16) is decomposed into (a) a scalar, denoted by c_p, which depends on the risky portfolio mean, μ_p, and (b) a weight vector, denoted by $\overline{w}$, which do not depend on μ_p:

$$w = \underbrace{\frac{(\mu_p - r_f)}{(\mu - r_f i)'\Sigma^{-1}(\mu - r_f i)}}_{c_p}\underbrace{\Sigma^{-1}(\mu - r_f i)}_{\overline{w}}. \tag{17}$$

So, all minimum-variance portfolios are a combination of (a) the risky asset portfolio with weights proportional to $\overline{w}$ and (b) the risk-free asset.

- When the targeted mean of the portfolio, μ_p is r_f, then $w = 0$, since $c_p = 0$. In this case, all funds are invested in the risk-free asset and nothing is invested in the portfolio of risky assets.

- For the tangency portfolio p, all funds are invested in the portfolio of risky assets: i.e.,

$$i'w = 1$$

or

$$c_p = \frac{1}{i'\bar{w}}. \tag{18}$$

Plugging (18) into (17), one obtains the weight vector of the optimal risky portfolio as

$$\begin{aligned} w &= \frac{1}{i'\bar{w}}\Sigma^{-1}(\mu - r_f i) \\ &= \frac{\Sigma^{-1}(\mu - r_f i)}{i'\Sigma^{-1}(\mu - r_f i)}. \end{aligned}$$