Revision notes - MA3269

Ma Hongqiang

October 16, 2017

Contents

0 Preliminary Result

0.1 Summation of series

- $\sum_{i=0}^{n} y^i = \frac{1-y^{n+1}}{1-y}$
- $\sum_{i=0}^{\infty} y^i = \frac{1}{1-y}$ provided $|r| < 1$.
- $\sum_{i=1}^{n} iy^{i-1} = \frac{1-y^n(1+n-ny)}{(1-y)^2}$
- $\sum_{i=1}^{\infty} iy^{i-1} = \frac{1}{(1-y)^2}$ provided $|r| < 1$.
- $\sum_{i=1}^{n} iy^i = \frac{y(1-y^n)-ny^{n+1}(1-y)}{(1-y)^2}$
- $\sum_{i=1}^{\infty} iy^i = \frac{y}{(1-y)^2}$ provided $|r| < 1$.
- $\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$
- $\sum_{i=1}^{n} i^2 = \frac{1}{6}n(n+1)(2n+1)$
- $\sum_{i=1}^{n} i^3 = \frac{1}{4}n^2(n+1)^2$

0.2 Newton Rhapson Method

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

0.3 Force of Interest

Accumulation function $a(s,t)$ can be derived from **force of interest** as such:

$$a(s,t) = e^{\int_s^t \delta(r)\,\mathrm{d}\,r}$$

where $\delta(r)$ is the force of interest with $s \leq r \leq t$.

0.4 Standard Derivatives

1 Theory of Interest

1.1 Interest

Definition 1.1 (Accumulation Function)**.**
When a principal of 1 dollar is deposited in an interest-paying account at time $t = 0$, it earns some interest over the time interval $[0, t]$.
The accumulated value of 1 dollar at time $t \geq 0$, denoted by $a(t)$, is known as the **accumulation function**. Clearly, $a(0) = 1$.

Definition 1.2 (Simple and Compound Interest)**.**
Let r be the annual rate of interest.
Based on the **simple-interest** method of calculating interest,

$$a(t) = 1 + rt \quad \text{for } t \geq 0$$

If the **compound interest** method is used,

$$a(t) = (1 + r)^t \quad \text{for } t \geq 0$$

Suppose the interest rate is r_i for the period $[\sum_{k=0}^{i-1} t_i, \sum_{k=1}^{i} t_i]$, where $t_0 = 0$,

$$a(t_j) = 1 + \sum_{i=1}^{j} r_i t_i \quad \text{when simple interest is used;}$$

$$a(t_j) = \prod_{i=1}^{j} (1 + r_i)^{t_i} \quad \text{when compound interest is used;}$$

Definition 1.3 (Frequency of Compounding)**.**
When an interest of $r = r^{(p)}$ is paid p times a year (or equivalently, $r^{(p)}$ is **convertible** p**thly** or $r^{(p)}$ is compounded p times a year), we call p the **frequency of compounding** and $r^{(p)}$ the **nominal** rate of interest.

The interest to be paid over the period, is $\frac{r^{(p)}}{p}$. Effectively, \$1 invested at time $t = 0$ will grow to $\left(1 + \frac{r^{(p)}}{p}\right)$ over a period of length $\frac{1}{p}$, so that the accumulated amount after one year is$\left(1 + \frac{r^{(p)}}{p}\right)^p$.
Remarks

1. We write the superscript (p) for $r^{(p)}$ to indicate the frequency of compounding p.
2. We can drop the superscript (p) when $p = 1$.
3. $p = 2, 4, 12$ correspond to semi-annual, quarterly and monthly compounding respectively,

Definition 1.4 (Equivalent Interest Rates)**.**
Two nominal interest rates are said to be **equivalent** if and only if they yield same accumulation amount over a year. Hence, the nominal rates $r^{(p)}$ and $r^{(q)}$ are equivalent if and only if

$$\left(1+\frac{r^{(p)}}{p}\right)^p = \left(1+\frac{r^{(q)}}{q}\right)^q$$

In particular, the **effective** annual interest rate (when $p = 1$), denoted by r_e, is given by

$$1 + r_e = \left(1+\frac{r^{(p)}}{p}\right)^p$$

The corresponding accumulation function is

$$a(t) = (1 + r_e)^t = \left(1+\frac{r^{(p)}}{p}\right)^{pt}$$

It can be shown that $r_e \geq r^{(p)}$ for $p > 1$.

Definition 1.5 (Continuous Compounding)**.**
The interest is **compounded continuously** when the frequency of compounding tends to infinity.
Let $r^{(\infty)}$ denote the nominal rate of interest under continuous compounding. Then,

$$a(1) = \lim_{p\to\infty}\left(1+\frac{r^{(\infty)}}{p}\right)^p = e^{r^{(\infty)}}$$

The number $r^{(\infty)}$ is known as the **continuously compounded** rate of interest. The corresponding accumulatio function is

$$a(t) = e^{r^{(\infty)}t}, \quad t \geq 0$$

Note that $e^{r^{(\infty)}} = 1 + r_e$.
It can be shown that

$$e^r > \left(1+\frac{r}{p}\right)^p$$

for any $r > 0$ and for any $p \in \mathbb{Z}^+$.

1.2 Present Value

Definition 1.6 (Present Value, Time Value)**.**
Let $a(t)$ be the accumulation function. Let X be the amount that must be invested at time $t = 0$ to accumulate to 1 dollar at $t = T$. Then

$$X \cdot a(T) = 1$$

or equivalently, $X = \frac{1}{a(T)}$.
The amount $X = \frac{1}{a(T)}$ is the **present value** of 1 paid at time T.

It follows that the present value of a single payment of C at time $t+T$ is $\frac{C}{a(T)}$.
More generally, for a cash flow $\mathbf{C} = \{(c_1,t_1),(c_2,t_2),\ldots,(c_n,t_n)\}$ consisting of a series of payments, with c_i received at time t_i, for $i=1,2,3,\ldots,n$, where $t_1 \geq 0$ and $t_i < t_j$ for $i<j$, the present value of this cash flow, denoted by $\mathrm{PV}(\mathbf{C})$, is defined by

$$\mathrm{PV}(\mathbf{C}) = \sum_{i=1}^{n} \frac{c_i}{a(t_i)}$$

Definition 1.7 (Time Value)**.**
The **time value** of the cash flow $\mathbf{C}$ at time $t \geq 0$, denoted by $\mathrm{TV}(\mathbf{C},t)$, is given by

$$\mathrm{TV}(\mathbf{C},t) = \mathrm{PV}(\mathbf{C}) \times a(t)$$

A consequence of the above definition is that for $0<s<t$,

$$\mathrm{TV}(\mathbf{C},t) = \frac{a(t)}{a(s)} \times \mathrm{TV}(\mathbf{C},s)$$

Definition 1.8 (Principle of Equivalence)**.**
In an environment where both the *interest rate* and its *method of accumulation* remain the same over any time period, two cash flows streams are **equivalent** if and only if they have the same present value.
(Alternatively, if and only if they have the same time value at $t=T$ for any $T \geq 0$).

It follows that the cash flow $\mathbf{C} = \{(c_1,t_1),(c_2,t_2),\ldots,(c_n,t_n)\}$ is equivalent to a single payment of $\mathrm{PV}(\mathbf{C}) = \sum_{i=1}^{n} \frac{c_i}{a(t_i)}$ at time $t=0$.

Definition 1.9 (Deferred Cash Flow)**.**
Let $k>0$ and define the cash flow $\mathbf{C}_{(k)} = \{(c_1,t_1+k),(c_2,t_2+k),\ldots,(c_n,t_n+k)\}$ which is essentially the cash flow $\mathbf{C} = \{(c_1,t_1),(c_2,t_2),\ldots,(c_n,t_n)\}$ deferred by k years.
If the accumulation function is $a(t)$, then

$$\frac{\mathrm{PV}(\mathbf{C})}{\mathrm{PV}(\mathbf{C}_{(k)})} = a(k)$$

Notations:
For the special case when $t_i = i-1$,

$$\mathbf{C} = \{(c_1,0),(c_2,1),\ldots,(c_n,n-1)\}$$

can be written as $(c_1,c_2,\ldots,c_n)$.

Definition 1.10 (Equation of Value)**.**
Consider the cash flow stream $\mathbf{C} = \{(c_1,t_1),)c_2,t_2),\ldots,(c_n,t_n)\}$. The equation

$$\mathrm{PV}(\mathbf{C}) = \sum_{i=1}^{n} \frac{c_i}{(1+r)^{t_i}} = 0$$

is known as the **equation of value**.

Definition 1.11 (Internal Rate of Return(IRR))**.**
Any non-negative root, r of the equation of value is called the **yield** or **internal rate of return (IRR)**, of the cash flow stream.

1.3 Annuities

Definition 1.12 (Annuities Immediate and Annuities Due)**.**
An annuity is a series of payment made at regular intervals.
An **annuity-due** is one for which payments are made at the *beginning* of each period.
An **annuity-immediate** is one for which payments are made at the *beginning* of each period.

Definition 1.13 (Perpetuity)**.**
A **perpetuity** is an annuity with an infinite number of payments.

Definition 1.14 (Loans)**.**
Loans are normally repaid by a series of installment payments made at *periodic* intervals.
The size of each installment can be determined using present-value analysis.
Specifically, if we let L be the amount of loan taken at time $t = 0$ and let $\mathbf{C} = \{(c_1, t_1), (c_2, t_2), \ldots, (c_n, t_n)\}$ be the series of repayments, then

$$L := \mathrm{PV}(\mathbf{C})$$

We can also compute the balance of the loan at any point in time.

Definition 1.15 (Loan Balance)**.**
The **loan balance** L_m^{Balance} immediately after the mth installment has been paid is the **time value** at $t = m$ of the remaining $(n - m)$ installment payments.
Suppose installment is paid annually with effectively annual rate r and each repayment of value c_i for year $m + i$, the loan balance

$$L_m^{\text{Balance}} = \sum_{i=1}^{n-m} \frac{c_i}{(1+r)^i}$$

Suppose each annual repayment is of value A. In reality, the loan is usually fully paid with n repayment of A plus a final payment B made at time $t \geq n$, where B is determined from the equation

$$L = \mathrm{PV}(0, \underbrace{A, A, \ldots, A}_{n\text{payments}}) + \mathrm{PV}(\{(B, t)\})$$

2 Bonds and Term Structure

2.1 Bond Terminology

Definition 2.1 (Bond)**.**
A **bond** is a written contract between the issuers(borrowers) and the investers(lenders) which specifies the following:

- **Face value**, F, of the bond: the amount based on which periodic interest payments are computed
- **Redemption/maturity value**, R, of the bond: the amount to be repaid at the end of the loan
- **Maturity date** of the bond: the date on which the loan will be fully repaid
- **Coupon rate**, c, (for coupon-paying bonds): the bond's interest payments, as a percentage of the par value, to be made to investors at regular intervals during the term of the loan

2.2 Bond Valuations

We use the following notations in connection with the bond pricing formula that follows.

- P = the current price of a bond
- F = face value of the bond
- R = redemption/maturity value of bond
- c = nominal coupon rate
- m = number of coupon payments per year
- n = total number of coupon payments (number of years $\times m$)
- λ = nomial yield

Theorem 2.1 (Price of a Bond)**.**
The price of a bond equals to the present value of the cash flow consisting of all coupon payments and the redemption value at maturity, calculated at yield λ.
For the case when the cash flow is made up of:

- coupon payments of $\frac{cF}{m}$ at time $t = \frac{1}{m}, \frac{2}{m}, \ldots, \frac{n}{m}$ (a total of n payments)
- redemption value R at $t = \frac{n}{m}$

We have

$$P = \frac{R}{\left(1 + \frac{\lambda}{m}\right)^n} + \sum_{i=1}^{n} \frac{\frac{cF}{m}}{\left(1 + \frac{\lambda}{m}\right)^i}$$

When $F = R$,

$$P = F + F\left(\frac{c-\lambda}{\lambda}\right)\left[1 - \frac{1}{\left(1+\frac{\lambda}{m}\right)^n}\right]$$

A bond is said to be priced

- at a **premium** if $P > F$
- at **par** if $P = F$
- at a **discount** if $P < F$

From the proceding bond pricing formula, it is clear

- $P > F$ if and only if $c > \lambda$
- $P = F$ if and only if $c = \lambda$
- $P < F$ if and only if $c < \lambda$

Theorem 2.2 (Makeham Formula)**.**
Let $K = \frac{F}{\left(1+\frac{\lambda}{m}\right)^n}$, we have

$$P = K + \frac{c}{\lambda}(F - K)$$

Theorem 2.3.
Let P_k be the price immediately after the k the coupon payment. Then

$$P_{k+1} = P_k\left(1 + \frac{\lambda}{m}\right) - \frac{cF}{m}$$

Definition 2.2 (Zero Coupon Bonds)**.**
Zero coupon bonds are bonds that pay no coupons. The cash flow for a N-year zero-coupon bond is the maturity value, R at $t = N$. Hence, at an annual yield of λ,

$$P = \frac{R}{(1+\lambda)^N}$$

Definition 2.3 (Perpetual Bonds)**.**
A bond that never matures (i.e., $n \to \infty$) is called a **perpetual bond**. Clearly,

$$P = \frac{cF}{\lambda}$$

Definition 2.4 (Bond Price Between Coupon Payments)**.**
The price of a bond traded in $t = \frac{k+\varepsilon}{m}$, $(0 \leq \varepsilon < 1$, which is between kth and $k+1$th coupon payment dates is

$$P_{k+\varepsilon} = (1+\mu)^{\varepsilon} P_k$$

where μ is the effective annual yield of the bond over the period $[k, k+1)$.

2.3 Macaulay Duration and Modified Duration

Definition 2.5 (Macaulay Duration)**.**
The **Macaulay duration** is one of the commonly used measures of bond's price sensitivity to changes in interest rate.
For cash flow stream $\mathbf{C} = \{(c_i, t_i) \mid i = 1, 2, \ldots, n\}$, the Macaulay duration, D. is defined by

$$D = \frac{\sum_{i=1}^{n} t_i \cdot \mathrm{PV}(c_i)}{\sum_{i=1}^{n} \mathrm{PV}(c_i)}$$

Equivalently, the Macaulay duration can be defined by the weighted average time to maturity of the cash flow stream:

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

where weight $w_i = \frac{\mathrm{PV}(c_i)}{\sum_{j=1}^{n} \mathrm{PV}(c_j)}$.

Theorem 2.4 (Properties of Macaulay Duration)**.**

- If $c_i \geq 0$ for all i, then $t_0 \leq D \leq t_n$.
- For a zero-coupon bond, $D = t_n$.

We can extend definition of Macaulay duration D to any infinite cash flow stream $\mathbf{C} = \{(c_i, t_i) \mid i = 1, 2, \ldots\}$

$$D = \frac{\sum_{i=1}^{\infty} t_i \cdot \mathrm{PV}(c_i)}{\sum_{i=1}^{\infty} \mathrm{PV}(c_i)}$$

Theorem 2.5 (Macaulay Duration of bonds)**.**
For a bond that pays a total of n coupons at a frequency of m payments a year. Let the nominal bond yield be λ and nominal coupon rate be c respectively. The cash flow stream in this case is

$$\mathbf{C} = \{(\frac{cF}{m}, t_1), \ldots, (\frac{cF}{m}, t_{n-1}), (\frac{cF}{m} + F, t_n)\}$$

as $t_i = \frac{i}{m}$, so that

$$D = \frac{1}{P}\left[\sum_{i=1}^{n} \frac{i}{m} \frac{\frac{cF}{m}}{\left(1 + \frac{\lambda}{m}\right)^i} + \frac{n}{m} \frac{F}{\left(1 + \frac{\lambda}{m}\right)^n}\right]$$

where

$$P = \sum_{i=1}^{n} \frac{\frac{cF}{m}}{\left(1 + \frac{\lambda}{m}\right)^i} + \frac{F}{\left(1 + \frac{\lambda}{m}\right)^n}$$

Let $\mu = \frac{\lambda}{m}$ and $\gamma = \frac{c}{m}$, then

$$D = \frac{\sum_{i=1}^{n} \frac{i}{m} \frac{\gamma}{(1+\mu)^i} + \frac{n}{m} \frac{1}{(1+\mu)^n}}{\sum_{i=1}^{n} \frac{\gamma}{(1+\mu)^i} + \frac{1}{(1+\mu)^n}}$$

It can be shown that

$$D = \frac{1+\mu}{m\mu} - \frac{1 + \mu + n(\gamma - \mu)}{m\mu + m\gamma\left[(1+\mu)^n - 1)\right]}$$

As the time to maturity tends to infinity, i.e. $n \to \infty$, for a perpetual bond,

$$D = \frac{1+\mu}{m\mu}$$

Macalay duration measures the sensitivity of bond prices to interest rates.
To see this, differentiate the pricing formula, we will have

$$\frac{\mathrm{d}\,P}{\mathrm{d}\,\lambda} = \left(-\frac{1}{1+\frac{\lambda}{m}}D\right)P$$

Definition 2.6 (Modified duration)**.**
The term $\frac{1}{1+\frac{\lambda}{m}}D$ is defined as the **modified duration** and is denoted by D_{M}.

In general, for a cash flow $\mathbf{C} = \{(c_i, t_i) \mid i = 1, 2, \ldots, n\}$ at an effective annual rate of r, the relation

$$\frac{\mathrm{d}\,P}{\mathrm{d}\,r} = -D_{\mathrm{M}}P$$

still holds.

Theorem 2.6 (Linear Approximation of Price Change)**.**
If $\Delta\lambda$ is a small change in λ, then

$$\Delta P = -D_{\mathrm{M}}P\Delta\lambda$$

Definition 2.7 (Duration of Bond Portfolio)**.**
Consider a bond portfolio consisting of α_i units of bond i, $i = 1, 2, \ldots, n$, assuming that the bonds have a *common* effective annual yield to maturity.
Let P_i and D_i be respectively the price and duration of bond i. Then, the **duration D_p of a portfolio** of n bonds of equal yield to maturity, λ is given by

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

where the **portfolio weight** $w_i = \frac{\alpha_i P_i}{\sum_{i=1}^{n}\alpha_i P_i}$

Definition 2.8 (Convexity C)**.**
Convexity of the bond C, is defined as the second derivative of the bond price with respect to bond yield, divided by the price of the bond.

$$C := \frac{\frac{\mathrm{d}^2\,P}{\mathrm{d}\,\lambda^2}}{P}$$

By Taylor series, it can be show that

$$\Delta P \approx -D_M P\Delta\lambda + \frac{1}{2}\frac{\mathrm{d}^2\,P}{\mathrm{d}\,\lambda^2}(\Delta\lambda)^2$$

Therefore,

$$\Delta P \approx P\left[-D_M \Delta\lambda + \frac{1}{2}C(\Delta\lambda)^2\right]$$

This obtains a better approximation of the change in price.
Also, from the bond pricing formula $P = \sum_{i=1}^{n} \frac{c_i}{\left(1+\frac{\lambda}{m}\right)^i}$, we have

$$\begin{aligned}
C &= \frac{\frac{\mathrm{d}^2 P}{\mathrm{d}\lambda^2}}{P} \\
&= \frac{1}{Pm^2\left(1+\frac{\lambda}{m}\right)^2}\sum_{i=1}^{n} i(i+1)\frac{c_i}{\left(1+\frac{\lambda}{m}\right)^i} \\
&= \frac{F}{P}\left\{\frac{2c}{\lambda^3}\left(1-\frac{1}{\left(1+\frac{\lambda}{m}\right)^n}\right) - \frac{2nc}{m\lambda^2\left(1+\frac{\lambda}{m}\right)^{n+1}} - \frac{n(n+1)(c-\lambda)}{\lambda m^2\left(1+\frac{\lambda}{m}\right)^{n+2}}\right\}
\end{aligned}$$

2.4 Yield curves and Term Structure of Interest Rates

Definition 2.9 (Spot Rates)**.**
A **spot rate** is the *annual* interest rate that begins today ($t = 0$) and lasts until some future time t. We denote this rate by s_t.
In effect the spot rate s_t is the yield to maturity of a zero-coupon bond that matures at t.

Definition 2.10 (Forward Rate)**.**
The interest rate observed at some future time $t_1 > 0$ and lasts until a time $t_2 > t_1$ is called a **forward rate**, denoted by f_{t_1,t_2}.
Note that $f_{0,t} = s_t$

Theorem 2.7.
In general,

$$(1+s_k)^k = (1+s_j)^j(1+f_{j,k})^{k-j}$$

and

$$(1+s_n)^n = (1+s_1)(1+f_{1,2})(1+f_{2,3})\cdots(1+f_{n-1,n})$$

3 Expected Utility Theory

3.1 Expected Utility and Risk Attitude

Definition 3.1 (Expected Utility).
An individual with an initial wealth of w_0 is considering a **risky prospect**with a random payoff X. He is assumed to have a **utility function** that is real-valued, continuous and **increasing**. He will make his investment decision based on the **expected utility** of his final wealth $W := X + w_0$, defined as follows.

- **Discrete** X
  If the risky investment has n possible mutually exclusive payoffs $(x_1, x_2, \ldots, x_n)$ with associated probabilities $p_1, p_2, \ldots, p_n)$, where $\sum_{i=1}^n p_i = 1$, then the **expected utility** of the individual's final wealth W, is given by

$$\mathrm{E}[U(W)] = \mathrm{E}[U(X + w_0)] := \sum_{i=1}^{n} p_i U(x_i + w_0)$$

- If X is a continuous random variable having a density function $f : (a, b) \to (0, \infty)$, then

$$\mathrm{E}[U(X + w_0)] := \int_a^b f(x) U(x + w_0) \,\mathrm{d}\, x$$

Definition 3.2 (Utility-based Decision).
Under **utility-based decision**, he individual will

- invest in the risky prospect if $\mathrm{E}[U(X + w_0)] > U(w_0)$.
- avoid the risky prospect if $\mathrm{E}[U(X + w_0)] < U(w_0)$.
- be indifferent to the risky prospect if $\mathrm{E}[U(X + w_0)] = U(w_0)$.

Given a set of risky prospects, an individual will *most* favour the one that maximises the expected utility of his final wealth.

Definition 3.3 (Characterisation of Risk Attitude).
An individual with utility function U is said to be

- risk averse if U is strictly concave.[1]
- risk neutral if U is linear.
- risk loving if U is strictly convex.

By Jensen Inequality, we deduce

Theorem 3.1 (Equivalent condition for Risk Attitude Characterisation).

[1] A function U is strictly concave on I if $U'' < 0$ on I.

- risk averse if $\mathrm{E}[U(W)] < U[\mathrm{E}(W)]$.
- risk neutral if $\mathrm{E}[U(W)] = U[\mathrm{E}(W)]$
- risk loving if $\mathrm{E}[U(W)] > U[\mathrm{E}(W)]$

for **any** risky investment that yields a final wealth of W.

Definition 3.4 (Positive Affine Transformation)**.**
Let U be an utility function. For any $\alpha > 0, \beta \in \mathbb{R}$, the function $\alpha U + \beta$ is a **positive affine transformation** of U.

Obviously, both function have the same attitude towards risks.

3.2 Certainty Equivalent

Definition 3.5 (Certainty Equivalent)**.**
Let U be the utility function of an individual. Given a risky prospect with payoff X ,the **certainty equivalent** of X with respect to U, is defined to be the real number $c =]CE(X;U)$ for which

$$U(c) = \mathrm{E}(U(w_0 + X))$$

It follows that an individual

- invests in the risky prospect if $\mathrm{CE}(X;U) > w_0$
- avoids the risky prospect if $\mathrm{CE}(X;U) < w_0$
- is indifferent if $\mathrm{CE}(X;U) = w_0$

For positive affine transformation $\alpha U + \beta$ where $\alpha > 0$, we have

$$\mathrm{CE}(X, \alpha U + \beta) = \mathrm{CE}(X, U)$$

Definition 3.6 (Risk Premium)**.**
The **risk premium** of a risky prospect with respect to an utility function U is the real number $r = \mathrm{RP}(X;U)$ for which

$$U(w_0 - r) = \mathrm{E}(U(w_0 + X))$$

where X is the payoff.

Clearly,

$$r = w_0 - c$$

and hence, an individual

- invests in the risky prospect if $\mathrm{RP}(X;U) < 0$
- avoids the risky prospect if $\mathrm{RP}(X;U) > 0$
- is indifferent if $\mathrm{RP}(X;U) = 0$

3.3 Arrow-Pratt Measures of Risk Aversion

Definition 3.7 (Absolute Risk Aversion)**.**
For a *risk averse* individual whose utility function is U, his **Arrow-Pratt absolute risk aversion Coefficient**(ARA) at wealth level w is

$$-\frac{U''(w)}{U'(w)}$$

Theorem 3.2 (ARA of positive affine transformation)**.**
$U_{\text{ARA}} = V_{\text{ARA}}$ if and only if U and V are positive affine transformation of each other.

We can say that two utility functions are **equivalent** if and only if they have the same ARA. Suppose two individuals with utility functions U and V admits the following condition:

$$U_{\text{ARA}}(w) > V_{\text{ARA}}(w)$$

at **all** wealth level, w, we say the individual with utility function U is **globally more risk averse** than the individual with utility function V.

Theorem 3.3.
More generally, an individual with utility function U is **globally more risk averse** than an individual with utility function V if and only if there is an increasing and strictly concave function g such that

$$U(w) = g(V(w))$$

3.4 Portfolio Selection

An individual with an initial wealth of w_0 can invest a portion (say αw_0, where $\alpha \in [0, 1]$) of his money in a risky investment X that has a random **rate of return**, R. The expected utility of his final wealth is

$$\mathrm{E}[U(W)] = \mathrm{E}[U(w_0(1 + \alpha R))]$$

Note that $\frac{\mathrm{d}^2}{\mathrm{d}\,\alpha^2}\,\mathrm{E}[U(W)] < 0$, hence setting the first order derivative to 0 always yields maxima α^*, although there is no guarantee that such $\alpha^* \in [0, 1]$.

4 Mean-Variance Analysis

In this chapter, the accumulation function is constant 1. Therefore, we do not consider time value.

4.1 Return and Risk of Asset

Definition 4.1 (Rate of Return)**.**
Asset is a tradable financial instruments. We denote that each asset is traded over one time period, from $t = 0$(initial) to $t = 1$(end-of-period).
If W_0 invested in an asset at time $t = 0$ is worth a **random** amount of W_1 at time $t = 1$, then the **rate of return** of the asset, denoted by r, is a **random variable** given by

$$r = \frac{W_1 - W_0}{W_0} = \frac{W_1}{W_0} - 1$$

Equivalently, $W_1 = W_0(1 + r)$.
The rate of return can also be defined in terms of the initial and end-of-period prices of the asset. Let P_0 be the price at $t = 0$ and P_1 be the **random** price at $t = 1$. Then

$$r = \frac{P_1 - P_0}{P_0} = \frac{P_1}{P_0} - 1$$

Equivalently, $P_1 = P_0(1 + r)$,

Definition 4.2 (Risk of Asset)**.**
The standard deviation, $\sigma_i = \sqrt{\mathrm{Var}(r_i)}$, of the rate of return of asset i, is a measure of the risk of asset i.

$$\sigma_i = \sqrt{\mathrm{Var}(r_i)} = \sqrt{\mathrm{E}[(r_i - \mathrm{E}(r_i))^2]} = \sqrt{\mathrm{E}[r_i^2 - E(r_i)^2]}$$

Definition 4.3 (Correlation of Returns)**.**
A statistical measure of the association of the returns of two assets, i and j, is the covariance $\sigma_{i,j} = \mathrm{Cov}(r_i, r_j)$.

$$\sigma_{i,j} = \mathrm{Cov}(r_i, r_j) = \mathrm{E}[(r_i - \mathrm{E}(r_i))(r_j - \mathrm{E}(r_j))] = \mathrm{E}[r_i r_j] - \mathrm{E}(r_i)\,\mathrm{E}(r_j) = \mathrm{E}[r_i(r_j - \mathrm{E}(r_j))]$$

A standardised measure is the correlation coefficient defined by

$$\rho_{i,j} = \frac{\sigma_{i,j}}{\sigma_i \sigma_j}$$

It can be shown that $|\rho_{i,j}| \leq 1$.

Definition 4.4 (Short Selling)**.**
Short selling of an asset refers to one borrowing a certain number of units of the asset from the lender at $t = 0$ and seems them immediately to receive an amount W_0. At some preagreed date $t = 1$, the short seller will buy the same number of units of the asset for an amount W_1 and return the asset to the lender.
The borrower will make a profit of $W_0 - W_1$ which is positive if and only if the value of the asset falls.

Obviously, the loss can be unlimited but the gain is bounded above by W_0.

4.2 Portfolio Mean and Variance

At time $t = 0$, an individual invests in n assets in such a way that a fraction w_i of his investment capital is invested in asset i. It is possible that $w_i < 0$, which means the individual short sells asset i.
We call the vector $\mathbf{w} = (w_1, w_2, \ldots, w_n)^T$ the individual's **portfolio weight vector**, or simply **portfolio**.
It is assumed that

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

We will then have its final wealth $W_1 = \sum_{i=1}^{n} w_i W_0(1 + r_i)$.
The rate of return r_p of the portfolio is related to the rate of return of individual assets, r_1, by

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

It follows that the expected rate of return of the portfolio, or **portfolio mean**, is

$$\mu_p = \mathrm{E}(r_p) = \sum_{i=1}^{n} w_i \mu_i = \mathbf{w}^T \mu$$

where

$$\mu = (\mu_1, \ldots, \mu_n)^T$$

is the vector of expected rates of return of the assets $(r_1, \ldots, r_n)$ respectively. This vector is called **mean vector** for simplicity.
The variance of rate of return of portfolio $\mathrm{Var}(r_p)$, or simply **portfolio variance**, of $\mathbf{w}$, is

$$\begin{aligned}
\sigma_p^2 = \mathrm{Var}(r_p) &= \mathrm{Cov}(\sum_{i=1}^{n} w_i r_i, \sum_{j=1}^{n} w_j r_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \,\mathrm{Cov}(r_i, r_j) \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} \sigma_{ij} \\
&= \mathbf{w}^T \mathbf{C} \mathbf{w} \text{ in matrix notation}
\end{aligned}$$

where

$$\mathbf{C} = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma nn \end{pmatrix}$$

is known as the **covariance matrix** of the random vector $\mathbf{r} = (r_1, \cdots, r_n)$. We also have, by noting $\sigma_{ii} = \mathrm{Var}(r_i) := \sigma_i^2$ and $\sigma_{ij} = \sigma_{ji}$,

$$\mathrm{Var}(r_p) = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{j<i} w_i w_j \sigma_{ij}$$

4.2.1 Diversification

Let $\overline{\sigma}^2$ and $\overline{\phi}$ be the average variance and average covariance of an n assets, that is

$$\overline{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 \quad \text{and} \quad \overline{\phi} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}$$

Suppose that $\overline{\sigma}^2 \to \sigma^2$ and $\overline{\phi} \to \phi$ as $n \to \infty$, then for an equally weighted portfolio, we have

$$\sigma_p^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2 + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\sigma_{ij} \to \phi$$

Therefore, there is limitation of diversification as a tool to reduce portfolio risk.
While the asset specific risk $\overline{\sigma}^2$ can be driven to zero, the market wide risk, $\overline{\phi}$ cannot be eliminated even if one holds infinitely many assets.

4.3 Portfolio of Two Assets

Consider a portfolio with weight vector $\mathbf{w} = \begin{pmatrix}\alpha & 1-\alpha\end{pmatrix}^t$ of two assets. The portfolio mean is

$$\mu_p = \alpha\mu_1 + (1-\alpha)\mu_2$$

and

$$\begin{aligned}\sigma_p^2 &= \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 + 2\alpha(1-\alpha)\sigma_{12} = \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 + 2\alpha(1-\alpha)\rho_{12}\sigma_1\sigma_2\\ &= (\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2)\alpha^2 + 2\sigma_2(\rho_{12}\sigma_1 - 1)\alpha + \sigma_2^2 \qquad (\#)\end{aligned}$$

4.3.1 Global Minimum-variance Portfolio

A risk averse individual seeks a portfolio with the *smallest* risk. He will thus seek the optimal value of α that minimises σ_p^2.
From the above equation (#),, σ_p^2 admits a parabola concave upwards, and the minimum portfolio variance σ_p^2 occurs when

$$\alpha = \alpha^* = \frac{\sigma_2(\sigma_2 - \rho_{12}\sigma_1)}{\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2}$$

and the minimum portfolio variance is

$$(\sigma_p^2)^* = \frac{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2}$$

The corresponding portfolio mean can then be determined from

$$\mu_p^* = \alpha^*\mu_1 + (1-\alpha^*)\mu_2$$

We call the portfolio with minimum variance the global minimum-variance portfolio.

4.3.2 Portfolio Graph

Definition 4.5 (Portfolio Graph)**.**
Portfolio graph is the graph of portfolio mean μ_p against portfolio risk σ_p.

From $\mu_p = \alpha\mu_1 + (1-\alpha)\mu_2$, we have

$$\alpha = \frac{\mu_p - \mu_2}{\mu_1 - \mu_2}$$

and by substituting the above equation to (#), we obtain an equation of the form

$$\sigma_p^2 = A\mu_p^2 + B\mu_p + C$$

for some constants A, B and C, with $A > 0$.[2]
This is an equation of a hyperbola. Rearranging

$$\sigma_p^2 = A(\mu_p - \frac{B}{2A})^2 + (C - \frac{B^2}{4A})$$

Therefore, $\min \sigma_p^2 = C - \frac{B^2}{4A}$ at $\mu_p = -\frac{B}{2A}$. This corresponds to the global minimum-variance portfolio.
The asymptotes of this graph are

$$\sigma_p = \pm\sqrt{A}(\mu_p + \frac{B}{2A})$$

or more naturally,

$$\mu_p = \pm\frac{1}{\sqrt{A}}\sigma_p - \frac{B}{2A}$$

When $\rho_{12} = 1, -1$, this hyperbola degenerate into a pair of lines.

4.4 Feasible Sets

4.4.1 Feasible Sets for two assets

Definition 4.6 (Feasible Sets for Two Assets)**.**
Given any two risky assets 1 and 2, it can be shown that the feasible set

1. is a **straight line** joining the (σ_1, μ_1) and (σ_2, μ_2) when $\rho_{12} = 1$.(Perfect positive correlation)

2. is a **V-shaped graph** comprising two straight lines, each joining the (σ, μ) point of one asset to a point with **zero portfolio variance**, when $\rho_{12} = -1$.(Perfect negative correlation)

3. is a curve passing through the (ρ, μ) points of the two assets when $|\rho_{12}| < 1$, where ρ_{12} denotes the correlation of the rates of return.

The feasible set will be extended by extending the corresponding line segments beyond the end points if short selling *is* allowed.

[2]This is due to the quadratic coefficient of (#) is greater than 0.

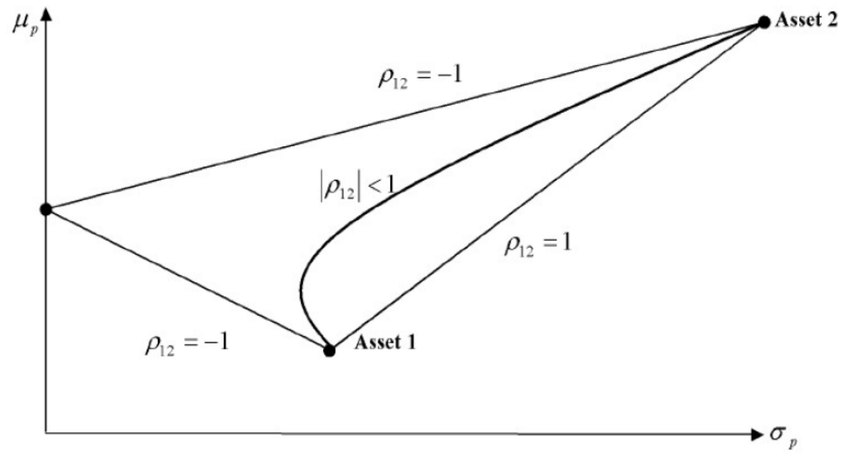

Figure 1: $0 \leq \alpha \leq 1$, when short selling is not allowed.

4.4.2 Feasible Sets of Portfolio of Three or More Assets

We construct the feasible sets intuitively by considering only the two-asset combination first, which gives rises to a finite set of hyperbolas and line segments. Next, we consider the combination of any two assets represented as two distinct points of these curves, which give rises to infinite number of hyperbolas and line segments. All the points on this infinite set will be inside the feasible set.
We will show that the feasible set of a portfolio with $n(> 2)$ assets has the following property.

Theorem 4.1 (Properties of Feasible Set)**.**

1. For any fixed $\mu \in \mathbb{R}$, $\exists \sigma > 0$ such that $(\sigma, \mu) \in F$.
2. For each $(\sigma, \mu) \in F$, $(\sigma', \mu) \in F$ for all $\sigma' > \sigma$.
3. For each pair of points (σ, μ) and (σ', μ') in the feasible set F, and for any $\lambda \in [0, 1]$, the point $\lambda(\sigma, \mu) + (1 - \lambda)(\sigma', \mu')$ lies in the set F.
   Equivalently, F is a **convex set**.
4. For any fixed $\mu \in \mathbb{R}$, there exists $\sigma^* > 0$ such that
   (a) $(\sigma^*, \mu) \in F$
   (b) if $(\sigma, \mu) \in F$, then $\sigma^* \leq \sigma$.

   We call this point (σ^*, μ) the **minimum-variance point** with mean μ.

Definition 4.7 (Minimum-Variance Frontier)**.**
The theorem above suggests there is a minimum-variance point for any porfolio mean. The set of all minimum-variance points is called the **minimum-variance frontier**.

It will be shown later that this minimum variance frontier is a **hyperbolic** curve.

Definition 4.8 (Global Minimum Variance Point)**.**
The extreme left point on this frontier is called the **global minimum variance point**.

Definition 4.9 (Efficient Frontier)**.**
The minimum variance frontier *above* the global minimum variance point is called the **efficient frontier**.

5 Portfolio Theory & Capital Asset Pricing Model

5.1 Markowitz's Portfolio Theory

Given n risky assets with mean vector $\mu = \begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{pmatrix}^{\mathrm{T}}$ and covariance matrix $\mathbf{C} = (\rho_{ij})_{n\times n}$, we seek a portfolio $\mathbf{w}$ with the smallest portfolio risk.
Let $\mu_p = \mathbf{w}^{\mathrm{T}}\boldsymbol{\mu}$ denote the portfolio mean and $\sigma_p^2 = \mathbf{w}^{\mathrm{T}}\mathbf{C}\mathbf{w}$ denote the portfolio variance.
We have two further assumptions:

Assumption 1 $\boldsymbol{\mu} = \begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{pmatrix}^{\mathrm{T}}$ is *not* a scalar multiple of $\mathbf{1}$.[3]

Assumption 2 $\mathbf{C} = (\sigma_{ij})_{n\times n}$ is *positive definite*. Equivalently, $\forall \mathbf{x} \neq \mathbf{0}, \mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x} > 0$.[4]

5.1.1 Minimisation of Portfolio Risk

The first problem is to minimise the portfolio risk σ_p^2 given a portfolio mean $\mu_p = \mu$.
It can be shown that the partial derivative of $\sigma_p^2 = \sum_{j=1}^n \sum_{i=1}^n w_i w_j \sigma_{ij} = \mathbf{w}^{\mathrm{T}}\mathbf{C}\mathbf{w}$, with respect to w_k is

$$\frac{\partial}{\partial w_k}(\sigma_p^2) = 2\sum_{j=1}^n \sigma_{kj} w_j$$

for $k = 1, \ldots, n$. Therefore, in matrix notation, we have

$$\frac{\partial}{\partial \mathbf{w}}(\mathbf{w}^{\mathrm{T}}\mathbf{C}\mathbf{w}) = 2\mathbf{C}\mathbf{w}$$

This result will be used in Lagrange multiplier, together with constraint $\sum_{i=1}^n w_i = 1$ and $\sum_{i=1}^n \mu_i w_i = \mu$.
Lagrangian function L is

$$L = \frac{1}{2}\sigma_p^2 - \lambda_1\left(\sum_{i=1}^n w_i - 1\right) - \lambda_2\left(\sum_{i=1}^n w_i\mu_i - \mu\right)$$

So the equations to solve are

$$\frac{\partial L}{\partial w_i} = \sum_{j=1}^n \sigma_{ij} w_j - \lambda_1 - \lambda_2\mu_i = 0 \quad \text{for } i = 1, 2, \ldots, n$$

with two constraints

$$\sum_{i=1}^n w_i - 1 = 0$$
$$\sum_{i=1}^n \mu_i w_i - \mu = 0$$

[3]If so, we have $\mu_p = \mathbf{w}^{\mathrm{T}}(k\mathbf{1}) = k$ a constant, which is degenerate.

[4]Note C as the covariance matrix is positive semidefinite. Should $\exists \mathbf{x} \neq \mathbf{0}$, such that $\mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x} = 0$, we can normalise $\mathbf{x}$ using 1-norm, and the left hand side become some constant times some arbitrary porfolio variance, which equals 0. This suggests the portfolio variance induced by $\mathbf{C}$ is 0 for some portfolio, which is degenerate. Therefore, we exclude the degenerate case from consideration.

In matrix notation,

$$\begin{cases} \mathbf{C}\mathbf{w} &= \lambda_1 \mathbf{1} + \lambda_2 \mu \\ \mathbf{1}^{\mathrm{T}}\mathbf{w} &= 1 \\ \boldsymbol{\mu}^{\mathrm{T}}\mathbf{w} &= \mu \end{cases}$$

Solving $\mathbf{w}$, we have

$$\mathbf{w} = \lambda_1 \mathbf{C}^{-1}\mathbf{1} + \lambda_2 \mathbf{C}^{-}1\boldsymbol{\mu}$$

Substituting back to the rest two equations, we have

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \mu \end{pmatrix}$$

where

$$a = \mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}$$
$$b = \mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\boldsymbol{\mu}$$
$$c = \boldsymbol{\mu}^{T}\mathbf{C}^{-1}\boldsymbol{\mu}$$

It can be easily seen, from Cauchy-Schwartz Inequality on Inner product, $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is invertible. Hence, solving the above matrix equation,

$$\begin{cases} \lambda_1 = \frac{c-b\mu}{ac-b^2} \\ \lambda_2 = \frac{a\mu-b}{ac-b^2} \end{cases}$$

Therefore, by substituting back λ_1 and λ_2 into equation of $\mathbf{w}$, we have the optimal weight vector $\mathbf{w} = \mathbf{w}_\mu^*$

$$\mathbf{w}_\mu^* = \left(\frac{c-b\mu}{ac-b^2}\right)\mathbf{C}^{-1}\mathbf{1} + \left(\frac{a\mu-b}{ac-b^2}\right)\mathbf{C}^{-1}\boldsymbol{\mu}$$

And variance of this optimal portfolio[5] is

$$\sigma_\mu^2 = \frac{a\mu^2 - 2b\mu + c}{ac-b^2}$$

The above equation gives the minimum-variance frontier.
By rearranging the above equation in the form

$$a\sigma^2 - \frac{a^2}{\Delta}(\mu - \frac{b}{a})^2 = 1$$

where $\Delta = ac - b^2 = \det \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, we see the frontier is indeed a **hyperbola**.
From the parabola equation, we deduce the GMVP mean and variance as

$$\sigma_{\mathrm{GMVP}} = \sqrt{\frac{1}{a}}$$
$$\mu_{\mathrm{GMVP}} = \frac{b}{a}$$

[5] $\sigma_\mu^2 = \mathbf{w}^{\mathrm{T}}\mathbf{C}\mathbf{w} = \mathbf{w}_\mu^{*\mathrm{T}}(\lambda_1\mathbf{1} + \lambda_2\boldsymbol{\mu}) = \lambda_1 + \lambda_2\mu$

and the portfolio weight vector of GVMP as

$$\mathbf{w}_{\text{GMVP}} = \frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}$$

An important consequence about the test of efficient frontier is that, a portfolio (σ_p, μ_p) lying *on* the frontier is efficient *if and only if* $\mu_p > \frac{b}{a} = \mu_{\text{GMVP}}$.

5.2 Two Fund Theorem

The vector vector for portfolios on the minimum-variance frontier is

$$\mathbf{w}_\mu^* = \left(\frac{c - b\mu_p}{ac - b^2}\right)\mathbf{C}^{-1}\mathbf{1} + \left(\frac{a\mu_p - b}{ac - b^2}\right)\mathbf{C}^{-1}\boldsymbol{\mu}$$

The two vectors, after being normalised, leads to

$$\mathbf{w}_\mu^* = \alpha\left(\frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}\right) + (1-\alpha)\left(\frac{\mathbf{C}^{-1}\boldsymbol{\mu}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\boldsymbol{\mu}}\right)$$

where $\alpha = a\left(\frac{c-b\mu_p}{ac-b^2}\right)$.
Note that the first normalised vector is the weight vector of the GMVP.

Definition 5.1 (Fund 1, Fund 2)**.**
We denote the two normalised, therefore weight vector by **fund 1** and **fund 2**.

$$\mathbf{w}_{\text{Fund 1}} = \frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}$$
$$\mathbf{w}_{\text{Fund 2}} = \frac{\mathbf{C}^{-1}\boldsymbol{\mu}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\boldsymbol{\mu}}$$

Theorem 5.1. The set of minimum-variance portfolio is **spanned/generated** by the two minimum-variance portfolios, $\mathbf{w}_{\text{Fund 1}}$ and $\mathbf{w}_{\text{Fund 2}}$.

In fact, this theorem can be generalised.

Theorem 5.2 (Two Fund Theorem)**.**
Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be the weight vectors of any two *distinct* portfolios **on the minimum-variance frontier**.
The minimum-variance set of portfolios is the set

$$\{\alpha\mathbf{w}_1 + (1-\alpha)\mathbf{w}_2 : \alpha \in \mathbb{R}\}$$

Theorem 5.3 (Two Corollaries)**.**

1. Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be two distinct efficient porfolios. Then any **convex sum**, $\alpha\mathbf{w}_1 + (1-\alpha)\mathbf{w}_2, \alpha \in [0,1]$ of $\mathbf{w}_1$ and $\mathbf{w}_2$ is an **efficient** portfolio.

2. If $b = \mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\boldsymbol{\mu} > 0$, then Fund 2 is **efficient**. In this case, any portfolio of the form $\alpha\mathbf{w}_{\text{Fund 1}} + (1-\alpha)\mathbf{w}_{\text{Fund 2}}$, with the **restriction** $\alpha < 1$, is also an efficient portfolio.