Mathematical Finance II (6CCM338A)–Utility Theory

Purba Das
Department of Mathematics, King's College London

Contents

[1]

Acknowledgement: This note is heavily influenced by previous notes of Dr Ryan Donnelly and various other online available lecture notes on utility theory.

I would like to thank Martin Dattge for proofreading the lecture notes and for his valuable feedback.

[1] You should follow the class notes as main materials; this is additional material, and may not cover every detail discussed in class.
This is also a draft version of topics covered in class; this is not a textbook on utility theory.

Decision-Making Under Risk

Let's consider the following 'fair' game:
You flip a coin. If it comes down heads, you get £10. If it comes down tails, you get £0. What is the maximum amount x that you would be willing to pay to play this game?

Mathematically speaking, you have the following two options:

- Do not play the game and get £0 for sure.
- Play the game and get $-\pounds x$ with probability 50% and $\pounds 10 - x$ with probability 50%.

Approach 1: Expected Value: The expected amount that you would earn from playing the game is

$$0.5(-x) + 0.5(10 - x).$$

This is bigger than 0 if

$$\begin{aligned} &0.5(-x) + 0.5(10 - x) \geq 0 \\ \iff &5 \geq x. \end{aligned}$$

So does this suggest we should pay at most £5 to play the game? **NOT REALLY.** Let's look at the following example:

The St. Petersburg Paradox

This 'expected payoff' approach was the accepted approach until *Daniel Bernoulli* (1738) suggested the following modification of the game.
Game description: Flip a coin repeatedly until it lands heads. If the first head appears on the n-th toss, you receive £2^n.

Hence, the probability of winning 2^n is

$$\mathbb{P}(\text{winning in n-th toss}) = \mathbb{P}(\text{All T for first n-1 tosses} + H \text{ in n-th toss}) = \frac{1}{2^n}.$$

How much would you pay to play this game with the expected value approach?

The expected value of this game is

$$\begin{aligned} \mathbb{E}[Game] &= \frac{1}{2}\pounds 2 + \frac{1}{4}\pounds 4 + \frac{1}{8}\pounds 8 + \frac{1}{16}\pounds 16 + \cdots \\ &= \pounds 1 + \pounds 1 + \pounds 1 + \pounds 1 + \cdots \\ &= \infty. \end{aligned}$$

So you should pay an infinite amount of money to play this game. Which does not make sense with our usual intuition. This is why it is called the St. Petersburg paradox.
So what is going wrong here?
Bernoulli resolved the problem by suggesting that people should be **maximizing expected utility**, not the expected value/payoff of the game.
Let us assume $U(w)$ denotes the utility of an amount w (can be thought of as wealth/income/bank balance). Then people consider

$$\mathbb{E}[U(w)]$$

and not $\mathbb{E}[w]$. Moreover, 'marginal utility' decreases with increasing wealth — the psychological value of an extra pound is smaller for a millionaire than for someone with £100 bank balance. This is mathematically ensured by assuming $U(w)$ is a concave increasing function of wealth w.

As an example, you can think of

$$\begin{aligned} U(£0) &= 0, \\ U(£499,999) &= 10, \\ U(£1,000,000) &= 17, \\ U(£1,500,000) &= 22, \\ &\vdots \end{aligned}$$

This gap between the expected value and actual human behaviour motivates utility theory.

1 Utility Functions

Utility Maximization Framework
We will denote a utility function by U. An investor's wealth at the beginning of the period is W_0, and the wealth at the end of the period is W or W_1. Note that W_0 is known to the investor, but W is a random variable (and depends on the investor's choices).

A utility function is a mapping

$$U : \mathbb{R}_+ \to \mathbb{R},$$

that assigns a numerical value $U(w)$ representing the level of satisfaction or preference associated with wealth w.

Formally, an investor's goal is to maximize her expected utility

$$\sup_{\Theta} \mathbb{E}[U(W)],$$

where Θ can be thought of as a feasible strategy for the investor.
HW: Suppose an investor's terminal wealth W is uncertain and is uniformly distributed on the interval $[0, 100]$, i.e. $W \sim Unif(0, 100)$. Assume further the investor has utility function $U(w) = \sqrt{w}$. Waht is $\mathbb{E}U(W)$?

This concept is best illustrated with a simple example:

Example 1. *Which game would you rather play?*

- *Game A: Win £5 or £2 with equal probability.*
- *Game B: Win £10 or lose £3 with equal probability.*

Generally, the first game is favoured, even though they have the same expected payoff.
Now, think of a situation where you have £1 Million pounds in your bank account and your friend has £100 in their bank account. Both games cost the same amount of money to enter. Does the preference for Game A/B change in terms of current wealth?

By this example, and many others, one can see that human behaviour is not consistent with the idea of *maximizing expectation.* One attempt at mathematically describing the observed behaviour is the notion of **(expected) utility maximization**. The mathematical foundations of utility theory are based on axioms developed by Von Neumann and Morgenstern.

Axioms of Utility functions

1. **Monotonicity (More-is-Better)** Investors prefer more wealth at least as much as they prefer less wealth. Assume an individual prefers consumption of bundle A of goods to bundle B. For $\alpha > 1$, this assumption says that individuals prefer αA to A, which in turn is preferred to B, but also A itself. For our example, if one week of food is preferred to one week of clothing, then two weeks of food is a preferred package to one week of food. Mathematically, the more-is-better assumption is called the monotonicity assumption on preferences:

$$w_2 > w_1 \implies U(w_2) \geq U(w_1).$$

   Remark: One can always argue that this assumption breaks down frequently. It is not difficult to imagine that a person whose stomach is full would turn down additional food. However, this situation is easily resolved. Suppose an individual is given the option of disposing the additional food to another person or a charity of his or her choice. In this case, the individual will still prefer more food even if he or she has eaten enough. Thus, under the monotonicity assumption, a hidden property allows costless disposal of excess quantities of any bundle.

2. **Strict monotonicity** Investors are unsatiated (having more is always preferred to staying the same). Mathematically speaking:

$$w_2 > w_1 \implies U(w_2) > U(w_1).$$

3. **Completeness** Individuals can rank order all possible bundles. Rank ordering implies that, no matter how many combinations of consumption bundles are placed in front of the individual, each individual can always rank them in some order based on preferences. This, in turn, means that individuals can somehow compare any bundle with any other bundle and rank them in order of the satisfaction each bundle provides. Mathematically, this can be captured by assuming U is continuous on its domain (wealth or something more general), ensuring preferences are always well-defined, i.e., for any two bundles A and B either

$$A \succsim B \text{ or } B \succsim A \text{ or } A \sim B \text{ (indifferent)}.$$

   For example, half a week of food and clothing can be compared to one week of food alone, one week of clothing alone, or any such combination.

4. **Convexity (Mix-is-better)** Suppose an individual is indifferent to the choice between one week of clothing alone and one week of food. Thus, either choice by itself is not preferred over the other. The "mix-is-better" assumption on preferences says that a mix of the two, say half-week of food mixed with half-week of clothing, will be preferred to both stand-alone choices. The mix-is-better assumption is called the "convexity" assumption on preferences.

5. **Transitivity (Rationality)** This is the most important and controversial assumption that underlies all of utility theory. Under the assumption of rationality, individuals' preferences avoid any kind of circularity; that is, if bundle A is preferred to B, and bundle B is preferred to C, then A is also preferred to C. Under no circumstances will the individual prefer C to A. Mathematically:

$$A \succsim B, B \succsim C \implies A \succsim C.$$

   One can likely see why this assumption is controversial. It assumes that the innate preferences (rank orderings of bundles of goods) are fixed, regardless of the context and time (which is often not the case in real-life situations).

6. If for two bundles $A \succsim B$, then for any other bundle C and $\alpha \in (0, 1)$:

$$\alpha A + (1 - \alpha)C \succsim \alpha B + (1 - \alpha)C.$$

This ensures consistent treatment of probabilistic mixtures.

Von Neumann–Morgenstern Theorem

If preferences satisfy these axioms, then there exists a real-valued function U (unique up to an affine transformation; i.e. $aU + b$, for $a > 0$) defined on the domain of all possible bundles such that for any two bundles (or lotteries) A and B:

$$A \succsim B \iff \mathbb{E}[U(A)] \geq \mathbb{E}[U(B)].$$

Thus, utility theory provides a quantitative representation of qualitative preferences under uncertainty.

Remark 1. *Utility functions are ordinal, not cardinal: If $U_2 = aU_1 + b$ with $a > 0$ then U_1 and U_2 represent the same preferences; i.e. only the ranking of outcomes matters.*

2 Utility Theory- as a function of wealth

Notes on Utility functions (on wealth):

- A utility function is increasing over wealth.
- Utility maximization from a utility function is only used to rank preferences. It does not provide how much one strategy is preferred over another.
- Utility functions are invariant under affine transformations, i.e., suppose $U_2 = aU_1 + b$ with $a > 0$. Then U_1 and U_2 are equivalent and we write $U_1 \sim U_2$.
- Utility functions do not have monetary units. However, there are various ways of converting an expected utility into monetary units (will be discussed later).

2.1 Examples of Utility functions

2.1.1 Quadratic Utility

The most basic example of a utility function is a quadratic utility function:

$$U(w) = w - \frac{b}{2}w^2;\ b > 0.$$

Strictly speaking, this is not a utility function because it is not monotonic. But for small values of w (precisely, for $w \in [0, \frac{1}{b}]$) it behaves like a utility function. This is often used in the literature due to its mathematical simplicity.

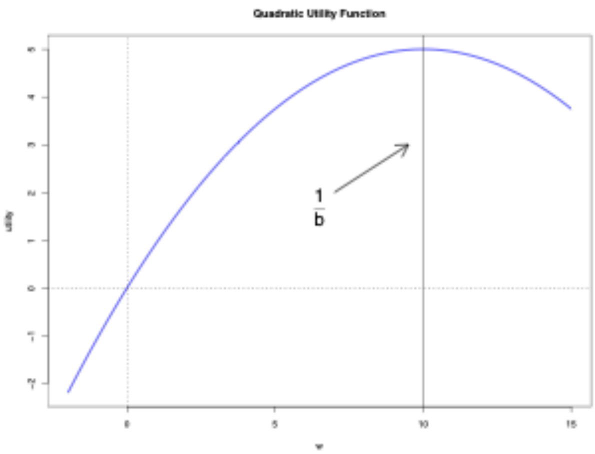


Figure 1: Quadratic utility function

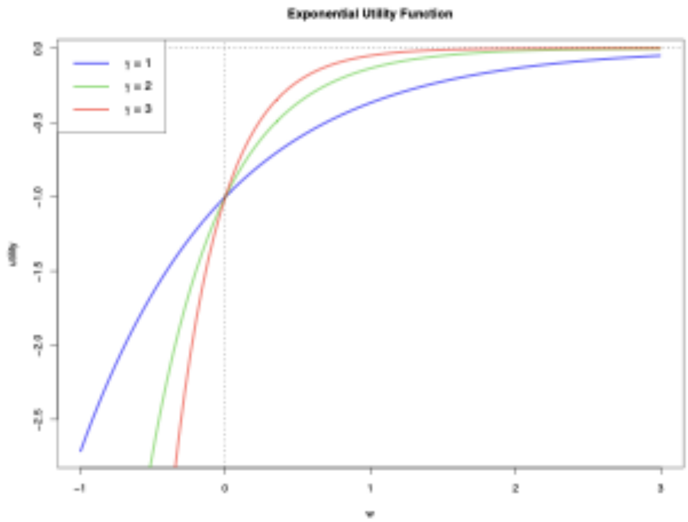


Figure 2: Exponential Utility

2.1.2 Exponential Utility

The exponential utility function is given as follows:

$$U(w) = -e^{-\gamma w};\ \gamma > 0.$$

Unlike the quadratic utility, this is a true utility function as it is strictly increasing. The exponential utility is popular because it allows for convenient calculations when combined with normal distributions. The value of γ is chosen to reflect the investor's level of risk-aversion (later we will also allow $\gamma < 0$).
Sometimes it is also written as

$$U(w) = \frac{1 - e^{-\gamma w}}{\gamma}.$$

HW: Show that these two definitions are equivalent.

Example 2. *If the end of period wealth $W \sim \mathcal{N}(\mu, \sigma)$ then*

$$\mathbb{E}[U(W)] = -\exp(-\gamma\mu + \frac{1}{2}\gamma^2\sigma^2),$$

which has a closed form and thus makes exponential utility analytically convenient.

2.1.3 Power Utility (sometimes also known as isoelastic utility)

Another common example of a utility function is the power utility (frequently used for practical applications in economics and finance):

$$U(w) = \frac{\omega^{1-\gamma} - 1}{1 - \gamma} \;\; , \gamma > 0; \; \gamma \neq 1.$$

We can consider $\gamma = 1$ by taking a limit $\gamma \uparrow 1$ which gives

$$U(w) = \log(w); \; w > 0.$$

This utility is somewhat more difficult to work with computationally, but often considered more

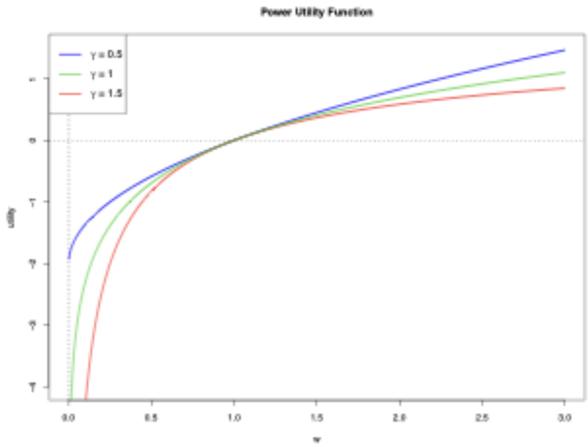


Figure 3: Power utility

realistic than the exponential utility. Some properties of this utility function are as follows:

1. Captures realistic behaviour: For $\gamma > 0$, the investor is risk-averse. Furthermore, as wealth grows, the marginal utility of extra wealth declines.

2. Scale invariance: Relative risk aversion is constant at γ, independent of the level of wealth. This makes the power utility especially suitable for portfolio problems — rich and poor investors put the same fraction of wealth into risky assets (although rich investors take more in absolute terms).

3. Special case–log utility: This corresponds to risk aversion that exactly offsets the proportional growth risk, leading to a natural connection to the Kelly criterion in gambling/investment[2].

2.2 Risk-Aversion and Utility concavity

A strictly risk-averse investor is unwilling to accept a fair game. Let's first describe what is meant by a fair game.

[2]See the Wikipedia page for more details: https://en.wikipedia.org/wiki/Kelly_criterion

Definition 1 (Fair game)**.** *A lottery/game is called a fair game if*

$$\mathbb{E}(W) = W_0,$$

where W_0 denotes the initial wealth and W is the end-of-period wealth.

Example 3. *Consider a lottery/game L where the player/investor receives a random payoff, denoted by H, which takes values h_0 and h_1 with probabilities p and $1-p$. If the player/investor begins with wealth W_0, then their end-of-period wealth is denoted as*

$$W = W_0 + H = \begin{cases} W_0 + h_0 & \text{with probability } p \\ W_0 + h_1 & \text{with probability } 1-p. \end{cases}$$

The expected payoff for this game is $h_0 p + h_1(1-p)$. So L is a fair game if and only if

$$\mathbb{E}[H] = h_0 p + h_1(1-p) = 0.$$

The expected utility of this lottery/game is:

$$\mathbb{E}[U(W)] = pU(W_0 + h_0) + (1-p)U(W_0 + h_1).$$

Furthermore, if L is indeed a fair game, we can write:

$$U(W_0) = \mathbb{E}[U(W_0)] = U\left[p(W_0 + h_0) + (1-p)(W_0 + h_1)\right].$$

This is coming from the fact that for a fair game $\mathbb{E}(H) = 0 \iff \mathbb{E}(W) = \mathbb{E}(W_0) = W_0$. And, $\mathbb{E}[U(W_0)] = \mathbb{E}[U(\mathbb{E}(W))]$.

What is risk-aversion?

Intuitively, risk-aversion is the notion that people generally prefer to be exposed to less randomness (particularly with respect to financial investments) if all the other parameters/variables are kept the same.

So what kind of utility function will a risk-averse investor use?

Lemma 2.1 (Characterization)**.** *For a strictly risk-averse investor (i.e. an investor who is risk-averse with a non-affine utility function), the utility function is strictly concave.*

Proof. A strictly risk-averse investor is unwilling to accept a fair game. If the risk-averse investor is maximizing their expected utility and they reject the fair game, then:

$$\begin{aligned} & \mathbb{E}[U(W_0)] > \mathbb{E}[U(W)] \\ \implies & U(W_0) > \mathbb{E}[U(W)] \\ \implies & U(\mathbb{E}W) > \mathbb{E}[U(W)] \\ \implies & U\left(p(W_0 + h_0) + (1-p)(W_0 + h_1)\right) > pU(W_0 + h_0) + (1-p)U(W_0 + h_1). \end{aligned} \tag{1}$$

But this must be true for any fair game/lottery L. That is the very last line is true for any $h_0, h_1 \in \mathbb{R}$ and $p \in [0,1]$ with $ph_0 + (1-p)h_1 = 0$.

Therefore, U must be a strictly concave function. If the investor is indifferent to some fair game (rather than rejecting it) then, the inequalities are not strict. This can also be seen from the fact that $U(\mathbb{E}[W]) > \mathbb{E}[U(W)]$ is true if and only U is concave (Jensen's inequality). □

The degree of concavity quantifies how strongly the agent dislikes risk.

Definition 2 (Marginal utility)**.** *Marginal utility at any given wealth level is the slope of the utility function at that wealth level. A diminishing marginal utility means the utility is always increasing, although at a decreasing rate. Similarly, increasing marginal utility means the utility is always increasing at an increasing rate.*

2.3 Risk-neutral and risk-seeking investors

Most common utility functions are smooth, so we can refer to their derivatives. We just showed that if an investor is risk-averse, the utility function is concave. This is a restriction on its second derivative:

$$\frac{d^2U}{dw^2} \leq 0.$$

Equality in the above relation defines a **risk-neutral** investor. If the above inequality holds in the opposite direction (convex function), an investor is called **risk-seeking** or **risk-loving**.

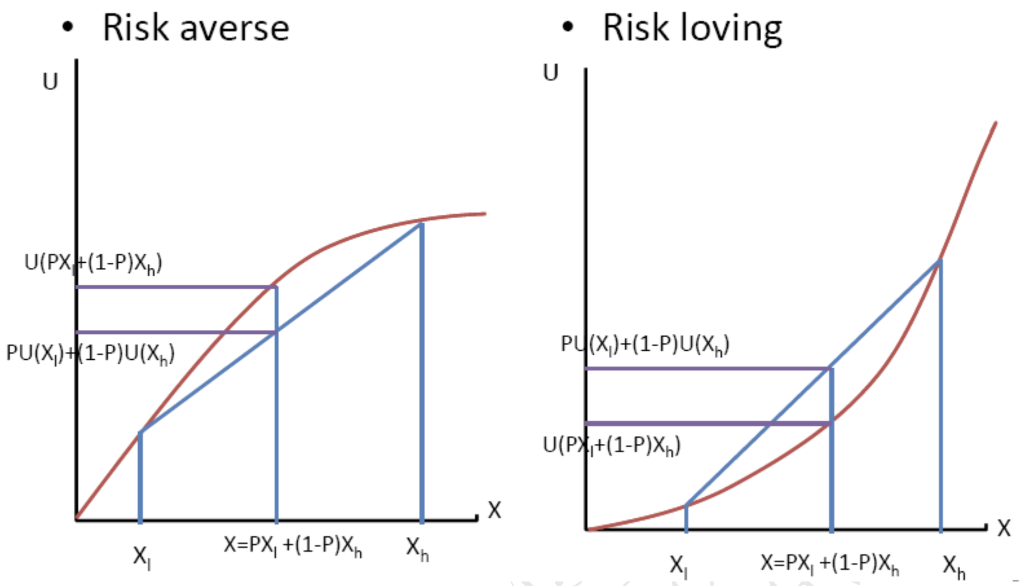


Figure 4: Utility function for a risk-averse (left) and risk-loving (right) investor.

Remark 2. *For any utility function, we always have* $\frac{d}{dw}U(w) \geq 0$*. This follows from the monotonicity assumption on the utility function.*

2.4 Certainty Equivalent

Motivation: Suppose an investor wants to compare two options, Choice (1) and Choice (2), for investing. Denote the end-of-period wealth by $W^{(1)}$ and $W^{(2)}$ depending on which choice is made. If an investor believes in utility U, they may make a choice depending on which is larger:

$$\mathbb{E}[U(W^{(1)})] \text{ or } \mathbb{E}[U(W^{(2)})]$$

Since U is unitless, this comparison does not tell the investor how much better one investment is compared to the other. The notion of a certainty equivalent is one way of quantifying, in monetary terms, the value of an investment.

Definition 3 (Certainty Equivalent)**.** *For* ***any*** *distribution of wealth W and for any valid utility function U, it is possible to find a unique constant W_C such that*

$$U(W_C) = \mathbb{E}[U(W)].$$

W_C is called the certainty equivalent corresponding to the investor's utility U.

Different investments can be compared in terms of their certainty equivalents. If

$$\mathbb{E}[U(W^{(1)})] \leq \mathbb{E}[U(W^{(2)})] \text{ then } W_C^{(1)} \leq W_C^{(2)}.$$

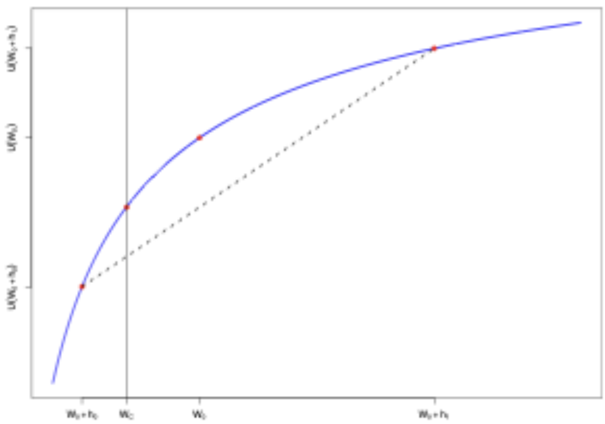


Figure 5: Certainty Equivalent

Risk-averse investors place less monetary value on an investment than the expected value of the investment.

Lemma 2.2. *For a risk-averse investor* $W_C \leq \mathbb{E}[W]$.

Proof. Jensen's inequality for a concave function f takes the following form:

$$\mathbb{E}[f(t)] \leq f(\mathbb{E}[t]).$$

If f is strictly concave then the above inequality is strict. Since utility functions for a risk-averse investor are concave, this translates to

$$\mathbb{E}[U(W)] \leq U(\mathbb{E}[W]).$$

By the definition of the certainty equivalent this implies

$$U(W_C) \leq U(\mathbb{E}[W]).$$

A utility function is strictly increasing. Therefore the previous relation implies

$$W_C \leq \mathbb{E}[W].$$

□

Intution: A risk-averse investor is willing to accept a (guaranteed) value smaller than $\mathbb{E}[W]$ *if this allows them to avoid the randomness associated with the game/lottery* W.

Furthermore, since a fair game is defined as $\mathbb{E}[W] = W_0$, it also holds for a risk-averse investor in a fair game W that $W_C \leq W_0$. (we draw a detailed picture in class to explain this).

Definition 4 (Risk Premium)**.** *The difference between the expected monetary value of a game/lottery and a risk-averse investor's certainty equivalent of the game/lottery is called the investor's risk premium, i.e.,*

$$R_p = \mathbb{E}(W) - W_C.$$

Intuition: The risk premium is the amount that a risk-averse investor will pay to avoid taking a risk (or, in other words, to eliminate risk).

Remark 3. *Note that for a fair game with initial wealth W_0 and end of period wealth W:*

$$\text{Risk-averse} \iff R_p \geq 0$$
$$\text{Risk-seeking} \iff R_p \leq 0$$
$$\text{Risk-neutral} \iff R_p = 0.$$

2.5 Absolute and Relative Risk-Aversion

Let $U \in C^2(\mathbb{R}_+, \mathbb{R})$.

Definition 5 (ARA)**.** *The coefficient of absolute risk-aversion is defined as*

$$\gamma^a(w) = -\frac{U''(w)}{U'(w)}.$$

In accordance with Definition 2, the first derivative $U'(w)$ is called the marginal utility of wealth. The concavity (measured by the second derivative) measures how much the investor likes to avoid randomness, but it should be normalised by the marginal utility of wealth. Hence, the function γ^a shows how an investor's preference to risk changes as a function of their current wealth.

Example 4. *Consider an exponential utility function:*

$$U(w) = -e^{-\gamma w}, \quad U'(w) = \gamma e^{-\gamma w}, \quad U''(w) = -\gamma^2 e^{-\gamma w}.$$

Thus, $\gamma^a(w) = \gamma$.

This is why exponential utility is called *constant absolute risk-aversion (CARA).* The amount of risk an investor is willing to take (in an absolute sense) does not depend on their current level of wealth.

Definition 6 (RRA)**.** *The coefficient of relative risk aversion is defined as*

$$\gamma^r(w) = -w\frac{U''(w)}{U'(w)} = w\gamma^a(w).$$

This quantity measures how sensitive an investor is to risk when expressed in a relative sense. If γ^a is constant, then the investor becomes more risk-sensitive in a relative sense as their wealth increases.

Example 5. *Consider a power utility function:*

$$U(w) = \frac{w^{1-\gamma} - 1}{1-\gamma}, \quad U'(w) = w^{-\gamma}, \quad U''(w) = -\gamma w^{-(1+\gamma)}.$$

Thus $\gamma^r(w) = \gamma$.

This is why power utility is sometimes called *constant relative risk-aversion (CRRA).*
The amount of risk an investor is willing to take (in a relative sense) does not depend on their current level of wealth. If an individual has constant absolute risk-aversion, then they have the same reluctance to participate in a fair game regardless of their wealth. This is not a typical behaviour we see in reality, as an individual is usually more willing to participate in a fair game if they hold more initial wealth. This is why, although computationally convenient, exponential utility is often considered less realistic than power utility.

2.6 Expected Utility maximization: An example

We assume the goal of an investor is to maximize their expected utility of wealth. The parameters over which the maximum is chosen vary from case to case, e.g., percentage of wealth in each asset, spending and savings plans, production plans of a firm. If we denote the investment choice by θ then the optimization problem is

$$\sup_{\theta} \mathbb{E}[U(W)],$$

where it is implied that W somehow depends on θ.

Example 6. *Consider an investor with the option of buying shares of an asset. The investor's initial wealth is W_0 and each share costs P_0. If the investor buys θ shares, the end-of-period wealth is $W = W_0 - \theta P_0 + \theta P_1$. Assume further $P_1 \sim \mathcal{N}(P_0 + \mu, \sigma)$ and that the investor uses exponential utility. What is the optimal number of shares to purchase?*

First, we compute expected utility. Let $Z \sim \mathcal{N}(0, 1)$, then

$$\begin{aligned}
\mathbb{E}[U(W)] &= \mathbb{E}[-e^{-\gamma(W_0 - \theta P_0 + \theta(P_0 + \mu + \sigma Z))}] \\
&= -e^{-\gamma W_0} e^{-\gamma\theta\mu} \mathbb{E}[e^{-\gamma\theta\sigma Z}] \\
&= -e^{-\gamma W_0} e^{-\gamma\theta\mu} e^{\frac{\gamma^2\theta^2\sigma^2}{2}}.
\end{aligned}$$

Now we maximize with respect to θ:

$$\begin{aligned}
\theta^* =& arg \max_{\theta} \left[-e^{-\gamma W_0} e^{-\gamma\theta\mu} e^{\frac{\gamma^2\theta^2\sigma^2}{2}} \right] \\
=& arg \min_{\theta} \left[e^{-\gamma\theta\mu} e^{\frac{\gamma^2\theta^2\sigma^2}{2}} \right] \\
=& arg \min_{\theta} \left[\frac{\gamma^2\theta^2\sigma^2}{2} - \gamma\theta\mu \right] \\
=& \frac{\mu}{\gamma\sigma^2}.
\end{aligned}$$

What about the intuition regarding what the investor should do?

1. If μ is positive, the investor buys the asset.
2. If μ is negative, the investor sells the asset.
3. If $\mu = 0$, the investor does not take any position.
4. More randomness (larger σ) gives a position closer to zero.
5. More risk-aversion (larger γ) gives a position closer to zero.

3 Utility Functions for Multiple Goods

Up until now, we have considered utility as a function of wealth only. However, in consumer theory and portfolio optimization, utility typically depends on several goods or consumption categories. This section develops the mathematical framework for multi-good utility, the associated optimization problem, and comparative results.
Let there be n goods with:

- Prices $p_1, p_2, \cdots, p_n > 0$, respectively,
- Consumed quantities $x_1, x_2, \cdots, x_n \geq 0$, respectively,
- Consumer income (or total budget) $m > 0$.

A consumption bundle is the vector:

$$x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n.$$

A utility function is a mapping:

$$U : \mathbb{R}^n_+ \to \mathbb{R},$$

assigning a satisfaction level to each bundle in $\mathbb{R}^n_+$. For example, $U(2,1,0)$ represents the utility from consuming 2 units of good 1, 1 unit of good 2, and 0 units of good 3.
Then the **budget constraint** is written as

$$\boxed{p_1x_1 + p_2x_2 + \cdots + p_nx_n \leq m.}$$

The **affordable consumption** bundles are bundles that do not cost more than the consumer's income or total budget. The **set of affordable consumption** bundles is the budget set of the consumer given by:

$$\left\{ x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n_+ : \sum_{i=1}^{n} p_i x_i \leq m \right\}.$$

The **budget line** is given by

$$\boxed{p_1x_1 + p_2x_2 + \cdots p_nx_n = m.}$$

Example 7 (Two goods)**.** *The budget line for two goods can be written as*

$$x_2 = \frac{m}{p_2} - \frac{p_1}{p_2}x_1,$$

i.e., slope of budget line = opportunity cost of good 1 wrt good 2 $:= \frac{p_1}{p_2}$. *Increasing* m *makes a parallel*

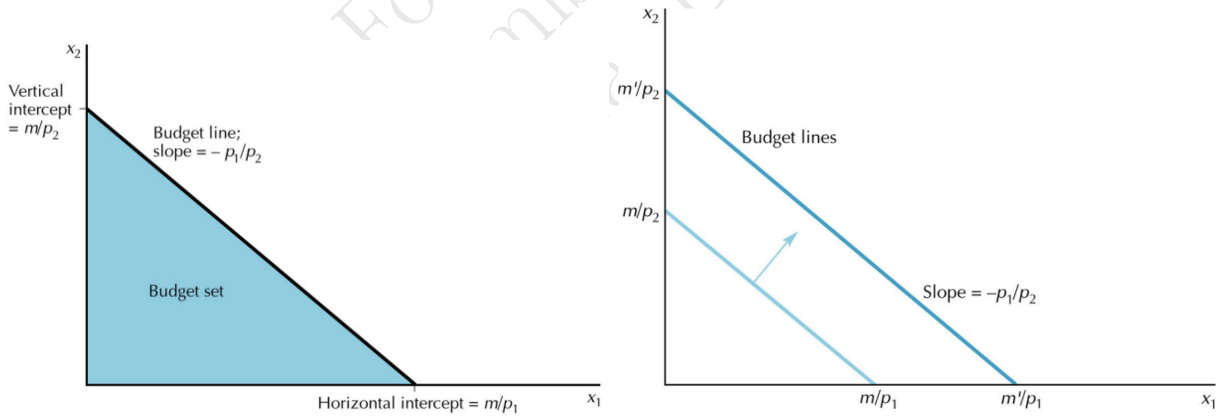


Figure 6: Budget set (left) and increase in total budget makes a parallel shift out (right).

shift outwards. The vertical/horizontal intercept increases, while the slope remains the same. Increasing p_1 *makes the budget line steeper. The vertical intercept remains the same, and the slope changes.*

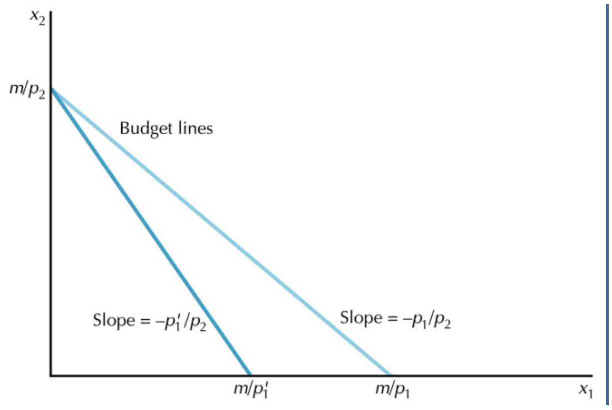


Figure 7: Increase of price of good 1 only

Proportional inflation: Multiplying all prices by t is just like dividing income by t:

$$tp_1x_1 + tp_2x_2 + \cdots + tp_nx_n = m \iff p_1x_1 + p_2x_2 + \cdots + p_nx_n = \frac{m}{t}.$$

Multiplying all prices and income by t doesn't change the budget line:

$$tp_1x_1 + tp_2x_2 + \cdots + tp_nx_n = tm \iff p_1x_1 + p_2x_2 + \cdots + p_nx_n = m.$$

A perfectly balanced inflation does not change consumption possibilities.

3.1 Nonlinear budget set

While we often assume linear budget constraints, real-world pricing schemes can be nonlinear — e.g., quantity discounts, rationing, or progressive tariffs.

Example 8 (Quantity discount)**.** *The agent has income $m = 30$. Good 1 has a per–unit price of $p_1 = 2$ for $x_1 < 10$, and a per–unit price of $p_1 = 1$ for $x_1 \geq 10$. Good 2 has a constant price of $p_2 = 2$. Let's consider two cases:*
First, when the agent buys $x_1 < 10$, the price of good 1 is $p_1 = 2$ and the equation of the budget line is therefore $2x_1 + 2x_2 = 30$. For example, if the agent spends all her money on good 2, she can afford $x_2 = 15$.
Second, when $x_1 \geq 10$, the agent spends 20 on the first 10 units of x_1 and 1 per unit thereafter. Hence, her budget constraint is

$$20 + (x_1 - 10) + 2x_2 = 30.$$

Example 9 (Rationing)**.** *The agent has income $m = 30$. Good 1 has a per–unit price of $p_1 = 2$ for $x_1 \leq 10$, but she is only allowed to purchase 10 units. Good 2 has a constant price of $p_2 = 2$. When the agent buys $x_1 \leq 10$, the price of good 1 is $p_1 = 2$ and the budget line is*

$$2x_1 + 2x_2 = 30.$$

For example, when the agent spends all her money on good 2, she can afford $x_2 = 15$. The agent is unable to buy more than 10 units of x_1, so the budget set is cut off at $x_1 = 10$.

Figure 9 characterizes the agent's optimal choice. Graphically, one can imagine the **indifference curves** (set of all bundles for which the consumer is indifferent to) flying in from the top right corner

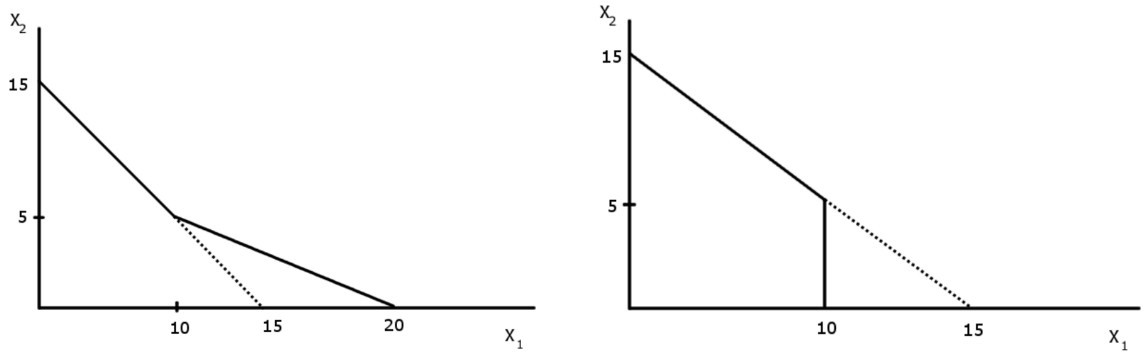


Figure 8: Quantity discounts (left) and rationing (right)

(where utility is highest) and stopping when it touches the budget set. To understand this further, consider Figure 9 (Right). There are 3 indifference curves. I_1 yields the highest utility, but never intersects with the budget set. I_2 corresponds to the agent's optimal choice (point A). I_3 yields a lower level of utility which is attainable but not desirable (please refer to the pictures we have drawn in class for more details).

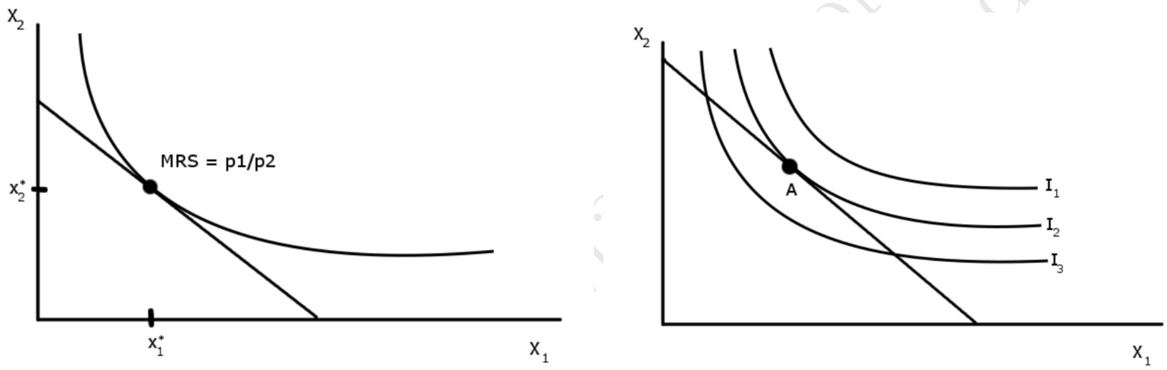


Figure 9: Optimal demand for goods 1 and 2

Remark 4. *Two indifferent curves can never intersect. (See class notes for proof)*

Assume the optimal bundle belongs to the interior of the budget set. At the optimal point, the budget line is tangential to the indifference curve. As a result, the budget line and the indifference curve have the same slope. This **tangency condition** means that

$$MRS(x_1^*, x_2^*) = \frac{p_1}{p_2}$$

where the marginal rate of substitution (MRS)[3] is evaluated at the optimal choice, (x_1^*, x_2^*). The MRS is defined as

$$MRS = \frac{\frac{dU}{dx_1}}{\frac{dU}{dx_2}},$$

[3]MRS of Good 1 with respect to Good 2

the slope of the utility function U. This generalises to the n goods case. At the optimum, the marginal utility per pound spent is the same across all goods:

$$\frac{1}{p_1}\frac{\partial U}{\partial x_1} = \frac{1}{p_2}\frac{\partial U}{\partial x_2} = \cdots = \frac{1}{p_n}\frac{\partial U}{\partial x_n} = \lambda.$$

Recall, $\frac{\partial U}{\partial x_i}$ is the marginal utility of good i. λ represents the **marginal utility of income**, sometimes called the shadow price of the budget constraint:

$$\lambda = \frac{dU^*}{dm} := \frac{dU(x_1^*, \cdots, x_n^*)}{dm}.$$

It measures the extra utility obtainable by increasing income by one unit.

A typical example of a utility function of multiple goods is the Cobb-Douglas utility function. For two goods with respective consumption x_1 and x_2, the Cobb-Douglas utility is given by:

$$U(x_1, x_2) = Ax_1^{\alpha}x_2^{(1-\alpha)} \qquad \text{where, } \alpha \in (0,1);\ A > 0.$$

3.2 Mathematical formulation of the Consumer's (primal) Problem

The primal optimization problem is written as follows. Assume the consumer chooses $x_1, \ldots, x_n$ (representing the units of good $1, \ldots, n$), to maximize utility

$$\begin{aligned} &\max_{(x_1,\ldots,x_n)} U(x_1, \ldots, x_n) \\ &\text{s.t. } p_1x_1 + \cdots + p_nx_n \leq m, \\ &x_i \geq 0. \end{aligned} \tag{2}$$

Lagrangian Formulation:

$$\mathcal{L}(x_1, \cdots, x_n, \lambda) = U(x_1, \cdots, x_n) + \lambda\,(m - p_1x_1 - \cdots - p_nx_n)$$

First-order (necessary) conditions for an interior maximum are:

$$(1)\quad \frac{\partial \mathcal{L}}{\partial x_i} = U_{x_i} - \lambda p_i = 0,$$

$$(2)\quad \frac{\partial \mathcal{L}}{\partial \lambda} = m - \sum_{i=1}^{n} p_i x_i = 0.$$

Solving the FOC, one gets the optimal consumer bundle $(x_1^*, x_2^*, \cdots, x_n^*)$.

Rearranging (1) for different i, we get:

$$\frac{U_{x_i}}{U_{x_j}} = \frac{p_i}{p_j}.$$

This means that the marginal utility trade-off is equal to the monetary trade-off between the two goods. Equation (2) states that the budget is exhausted (non-satiation). At the solution of the consumer's problem (more specifically, an interior solution/optimum), for n many goods $(x_1, x_2, \ldots, x_n)$ the following conditions will hold:

$$\frac{\partial U/\partial x_1}{p_1} = \frac{\partial U/\partial x_2}{p_2} = \ldots = \frac{\partial U/\partial x_n}{p_n} = \lambda.$$

This expression says that **at the utility-maximising point (optimum)**, the next unit (e.g., one pound/dollar ...) spent on each good yields the same marginal utility.

What is the meaning of the artificial variable λ – the Lagrange multiplier?

What about the quantity $\frac{dU\left(x_1^*,\cdots,x_n^*\right)}{dm}$, where $x_1^*,\cdots,x_n^*$ are the consumer's optimal consumption choices subject to her budget constraint? It turns out that

$$\frac{dU\left(x_1^*,\cdots,x_n^*\right)}{dm} = \lambda.$$

λ represents the marginal utility of income, sometimes called the shadow price of the budget constraint, i.e. *it expresses the quantity of utils that could be obtained with the next unit of consumption.* Note that this expression only holds when $x_i = x_i^*$ for all i. If x_i's were not at their optimal values, then the total derivative of $\mathcal{L}$ with respect to m would also include additional cross-partial terms. These cross-partials are ONLY zero when $x_i = x_i^*$ for all i.
What does the "shadow price" mean?
It's the additional "utility value" of relaxing the budget constraint by one unit.

Corner solutions and non-negativity

We say the consumer problem has a corner solution if the consumer buys zero quantity of some good and spends the entire budget on other good(s).

What problem does this create for us when we try to solve the Lagrangian? The problem above

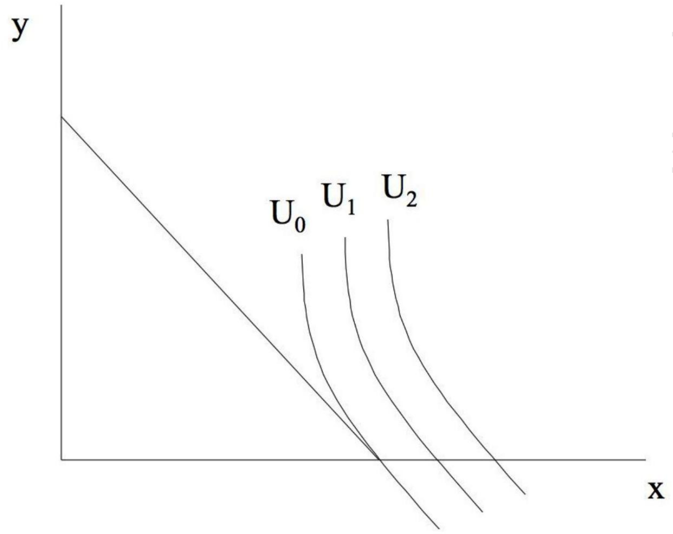


Figure 10: Corner solution for goods 1 and 2

is that a point of tangency does not exist for positive values of all x_i. Hence, we also need to impose "non-negativity constraints" $x_i \geq 0$. These are known as Kuhn–Tucker conditions.

For the case of two goods, the consumer's optimization problem can be written as follows:

$$\begin{aligned} &\max_{x,y} U(x,y) \\ &\text{s.t. } \; p_x x + p_y y \leq m \cdot \\ &\qquad y \geq 0. \end{aligned}$$

So the Lagrange problem becomes

$$\mathcal{L}(x,y,\lambda,\mu) = U(x,y) + \lambda\left(m - p_x x - p_y y\right) + \mu(y-0)$$

The corresponding first-order conditions are:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} =& U_x - \lambda p_x = 0,\\ \frac{\partial \mathcal{L}}{\partial y} =& U_y - \lambda p_y + \mu = 0,\\ \mu y =& 0.\end{aligned}$$

3.3 Indirect Utility Function & Marshallian demand function

We learnt how to obtain a set of optimally chosen quantities using the Lagrange multiplier:

$$\begin{aligned}x_1^* =& x_1 (p_1, p_2, \ldots, p_n, m)\\ &\vdots\\ x_n^* =& x_n (p_1, p_2, \ldots, p_n, m).\end{aligned}$$

So when we say

$$\max U (x_1, \ldots, x_n) \text{ s.t. } p_1 x_1 + \ldots + p_n x_n \leq m,$$

we get as a result the optimal utility:

$$U (x_1^* (p_1, \ldots, p_n, m), \ldots, x_n^* (p_1, \ldots, p_n, m)) \equiv V (p_1, \ldots, p_n, m).$$

The function/quantity $V(\cdot)$ is known as the "Indirect Utility Function." **This is the value of maximized utility under given prices and income** (See class notes for details).
Properties of indirect utility:

1. Homogeneous of degree 0 in $(p_1, \cdots, p_n, m)$, i.e., $V(tp_1, \cdots, tp_n, tm) = V(p_1, \cdots, p_n, m)$.
2. Increasing in m, decreasing in each p_i.
3. Quasi-concave in $(p_1, \cdots, p_n, m)$.

A consumer's **Marshallian demand function**, denoted by M_i, is the quantity consumer demand of a particular good as a function of price vector $(p_1, \cdots p_n)$ and total income m.

$$M_i(p_1, \cdots p_n, m) = x_i^* (p_1, \ldots, p_n, m)$$

Example 10 (Cobb-Douglas Utility)**.**

$$\begin{aligned}\max U(x_1, x_2) = x_1^{0.5} x_2^{0.5}\\ s.t.\ p_1 x_1 + p_2 x_2 \leq m.\end{aligned}$$

The Lagrange and the FOC are as follows:

$$\begin{aligned}\mathcal{L}(x_1, x_2, \lambda) &= x_1^{0.5} x_2^{0.5} + \lambda (m - p_1 x_1 - p_2 x_2)\\ \frac{\partial \mathcal{L}}{\partial x_1} &= 0.5 x_1^{-0.5} x_2^{0.5} - \lambda p_1 = 0\\ \frac{\partial \mathcal{L}}{\partial x_2} &= 0.5 x_1^{0.5} x_2^{-0.5} - \lambda p_2 = 0\\ \frac{\partial \mathcal{L}}{\partial \lambda} &= m - p_1 x_1 - p_2 x_2 = 0\end{aligned}$$

By solving the above equations, one obtains the shadow price λ as follows:

$$\lambda = \frac{0.5x_1^{-0.5}x_2^{0.5}}{p_1} = \frac{0.5x_1^{0.5}x_2^{-0.5}}{p_2}.$$

This simplifies to:

$$x_1 = \frac{p_2 x_2}{p_1}.$$

Substituting x_1 into the budget constraint gives us:

$$m - p_1\frac{p_2 x_2}{p_1} - p_2 x_2 = 0,$$

$$p_2 x_2 = \frac{1}{2}m, \quad p_1 x_1 = \frac{1}{2}m.$$

Hence,

$$x_1^* = \frac{m}{2p_1}, \quad x_2^* = \frac{m}{2p_2}.$$

So in the optimal bundle, half of the budget goes to each good.

Thus, a consumer with utility function $U(x_1, x_2) = x_1^{0.5}x_2^{0.5}$, total budget m, and facing prices p_1 and p_2, will choose x_1^* and x_2^* and obtain utility:

$$U(x_1^*, x_2^*) = \left(\frac{m}{2p_1}\right)^{0.5}\left(\frac{m}{2p_2}\right)^{0.5}.$$

Hence, the indirect utility for this consumer is

$$V(p_1, p_2, m) = U(x_1^*(p_1, p_2, m), x_2^*(p_1, p_2, m)) = \left(\frac{m}{2p_1}\right)^{0.5}\left(\frac{m}{2p_2}\right)^{0.5}.$$

The Marshallian demand function is given by:

$$M_i(p_1, \cdots p_n, m) = x_i^*(p_1, \ldots, p_n, m) = \frac{m}{2P_i} \qquad \forall\, i \in \{1, 2\}.$$

Why do we even calculate the indirect utility function?
It saves us time. Instead of recalculating the utility level for every set of prices and budget constraints, we can plug in prices and income to get the consumer's optimal utility. This comes in handy when working with individual demand functions. Demand functions give the quantity of goods purchased by a given consumer as a function of prices and income.

We sometimes calculate the indirect utility function because it tells us how well off a consumer is as a function of prices and income, not choices. For example; If policy A gives higher indirect utility than policy B, consumers prefer A – No need to look at any bundles.

Takeaway message: Marshallian demand tells us what consumers choose; indirect utility tells us how well off they are.

3.4 The Dual Optimization Problem

The utility maximization problem (the *primal problem*) has an associated *dual problem*, known as the *expenditure minimization problem*. Instead of maximizing utility subject to a budget constraint m, the consumer minimizes total expenditure subject to achieving a given level of utility.

Expenditure Minimization Problem

Fix a target utility level $\bar{u} \in \mathbb{R}$. The dual problem is

$$\begin{aligned} \min_{(x_1,\dots,x_n)} \; & p_1 x_1 + \cdots + p_n x_n \\ \text{s.t. } & U(x_1,\dots,x_n) \geq \bar{u}. \end{aligned} \tag{3}$$

This problem asks for the least costly bundle of goods that yields utility of at least $\bar{u}$.

Lagrangian Formulation

The Lagrangian associated with the expenditure minimization problem is

$$\mathcal{L}(x_1,\dots,x_n,\mu) = p_1 x_1 + \cdots + p_n x_n + \mu\left(\bar{u} - U(x_1,\dots,x_n)\right),$$

where $\mu \geq 0$ is the Lagrange multiplier associated with the utility constraint.

First-Order Conditions

For an interior solution, the necessary first-order conditions are

$$\begin{aligned} (1) \quad & \frac{\partial \mathcal{L}}{\partial x_i} = p_i - \mu U_{x_i} = 0, \quad i = 1,\dots,n, \\ (2) \quad & \frac{\partial \mathcal{L}}{\partial \mu} = \bar{u} - U(x_1,\dots,x_n) = 0. \end{aligned}$$

From condition (1), we obtain

$$\frac{U_{x_i}}{U_{x_j}} = \frac{p_i}{p_j}, \quad \forall i, j,$$

which coincides with the optimality condition from the primal problem.
Notation: Let us call the optimal solution of the dual problem $(x_1^h, \cdots, x_n^h)$.

The **expenditure function** gives the minimum level of income required to achieve a given level of utility at a given vector of prices. Formally, for a price vector $(p_1,\dots,p_n)$ and a target utility level $\bar{u}$, the expenditure function is defined as

$$e(p_1,\dots,p_n,\bar{u}) = \sum_{i=1}^{n} p_i x_i^h.$$

Thus, $e(p,\bar{u})$ represents the least amount of expenditure necessary for the consumer to attain utility level $\bar{u}$ given prices $p_1,\dots,p_n$.

Hicksian Demand

The solution to the dual problem defines the *Hicksian (compensated) demand functions.* The Hicksian (or compensated) demand function, denoted by H_i, is defined as the quantity of good i that minimizes total expenditure for a given price vector $(p_1,\cdots,p_n)$ while achieving a fixed level of utility $\bar{u}$. Formally,

$$H_i(p_1,\dots,p_n,\bar{u}) = x_i^h(p_1,\dots,p_n,\bar{u}),$$

where $x^h(p_1, \dots, p_n, \bar{u})$ solves the expenditure minimization problem (3). Thus, while the Marshallian demand $M_i(p, m)$ describes optimal consumption as a function of prices and income, the Hicksian demand $H_i(p, \bar{u})$ describes optimal consumption as a function of prices and a fixed utility level.

These demand functions describe how consumption responds to changes in prices while keeping utility fixed.

Example 11 (Cont. of Example 10)**.**

$$U(x_1, x_2) = \sqrt{x_1 x_2}$$

Expenditure minimization problem:

$$\min_{x_1, x_2 \geq 0} \; p_1 x_1 + p_2 x_2 \quad s.t. \quad \sqrt{x_1 x_2} \geq \bar{u}$$

Equivalently,

$$x_1 x_2 \geq \bar{u}^2.$$

Lagrangian:

$$\mathcal{L} = p_1 x_1 + p_2 x_2 + \lambda \left(\bar{u}^2 - x_1 x_2 \right).$$

First-order conditions:

$$\frac{\partial \mathcal{L}}{\partial x_1} = p_1 - \lambda x_2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = p_2 - \lambda x_1 = 0.$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \bar{u}^2 - x_1 x_2 = 0$$

Solving the FOCs:

$$\frac{p_1}{p_2} = \frac{x_2}{x_1} \quad \Rightarrow \quad x_2 = \frac{p_1}{p_2} x_1.$$

Substitute into the constraint:

$$x_1 \left(\frac{p_1}{p_2} x_1 \right) = \bar{u}^2$$

$$x_1^2 = \bar{u}^2 \frac{p_2}{p_1}$$

Hicksian Demand function:

$$H_1(p_1, p_2, \bar{u}) = x_1^h(p_1, p_2, \bar{u}) = \bar{u} \sqrt{\frac{p_2}{p_1}}$$

$$H_2(p_1, p_2, \bar{u}) = x_2^h(p_1, p_2, \bar{u}) = \bar{u} \sqrt{\frac{p_1}{p_2}}$$

Expenditure function:

$$e(p_1, p_2, \bar{u}) = p_1 x_1^h + p_2 x_2^h$$

$$= p_1 \bar{u} \sqrt{\frac{p_2}{p_1}} + p_2 \bar{u} \sqrt{\frac{p_1}{p_2}} = 2\bar{u}\sqrt{p_1 p_2}.$$

Primal–Dual Relationship (For the case of an interior solution)

The utility maximization problem (primal problem) and the expenditure minimization problem (dual problem) are closely related and provide two equivalent descriptions of consumer behavior:

- Solving the primal problem yields the Marshallian demand $x_i^*(p_1, \cdots, p_n, m)$,
- solving the dual problem yields the Hicksian demand $x_i^h(p_1, \cdots, p_n, \bar{u})$,
- these two demand functions are linked through the indirect utility function $V(p_1, \ldots, p_n, m)$ and the expenditure function $e(p_1, \ldots, p_n, \bar{u})$. In particular, at the optimum,

  $$\bar{u} = V(p_1, \ldots, p_n, m) \quad \text{and} \quad m = e(p_1, \ldots, p_n, \bar{u}).$$

  Consequently, Marshallian and Hicksian demands are related by

  $$M_i(p, m) = H_i\big(p, V(p, m)\big), \quad \text{and} \quad H_i(p, \bar{u}) = M_i\big(p, e(p, \bar{u})\big).$$

  Thus, the primal and dual problems generate the same optimal consumption bundle, differing only in whether income or utility is taken as given.

Thus, the primal and dual problems represent two equivalent ways of describing consumer choice.

3.5 A practical problem in expected utility

While expected utility theory provides a valuable tool for analyzing how rational people should make decisions under uncertainty, observed behaviour may not always confirm this. Daniel Kahneman and Amos Tversky (1974) were the first to provide evidence that $\mathbb{E}(U)$ theory does not provide a complete description of how people actually decide under uncertain conditions. The authors conducted experiments that demonstrate this variance from the $\mathbb{E}(U)$ theory, and these experiments have withstood the test of time. It turns out that individual behaviour under some circumstances violates the axioms of rational choice of $\mathbb{E}(U)$ theory.

Kahneman and Tversky (1981) provide the following example: Suppose the country is going to be struck by the avian influenza (bird flu) pandemic. Two programs are available to tackle the pandemic, A and B. Two groups of physicians, X and Y, are set with the task of containing the disease. Each group has the outcomes that the two programs will generate. However, the outcomes have different phrasings for each group. Group X is told about the efficacy of the programs in the following words:

Program A: If adopted, it will save exactly 200 out of 600 patients.

Program B: If adopted, the probability that 600 people will be saved is 1/3, while the probability that no one will be saved is 2/3.

76% of the doctors in group X chose to administer program A.

Group Y, on the other hand, is told about the efficacy of the programs in these words:

Program A: If adopted, exactly 400 out of 600 patients will die.

Program B: If adopted, the probability that nobody will die is 1/3, while the probability that all 600 will die is 2/3.

Only 13 % of the doctors in this group chose to administer program A.

The only difference between the two sets presented to groups X and Y is the description of the outcomes. Every outcome to group X is defined in terms of "saving lives," while for group Y it is in terms of how many will "die." Doctors, being who they are, have a bias towards "saving" lives.

This experiment has been repeated several times with different subjects and the outcome has always been the same, even if the numbers differ. Other experiments with different groups of people also showed that the way alternatives are worded results in different choices among groups. The coding of alternatives that makes individuals vary from $\mathbb{E}(U)$ maximizing behavior is called the framing effect.

In order to explain these deviations from $\mathbb{E}(U)$, Kahneman and Tversky suggest that individuals use a value function to assess alternatives. This is a mathematical formulation that seeks to explain observed behavior without making any assumptions about preferences. The nature of the value function is such that it is much steeper in losses than in gains. The authors insist that it is a purely descriptive device and is not derived from axioms like the $\mathbb{E}(U)$ theory. In the language of mathematics, we say the value function is convex in losses and concave in gains. For the same concept, economists will say that the function is risk-seeking in losses and risk-averse in gains.

Mathematical Finance II (6CCM338A)
Portfolio management–Two securities

Purba Das

March 27, 2026

Contents

[1]

1 Portfolio Management: Introduction & Assumptions

Markowitz Portfolio Theory, also called Modern Portfolio Theory (MPT), was introduced by Harry Markowitz in 1952. It is a framework for constructing a portfolio of assets to maximize expected return for a given level of risk, or equivalently, minimize risk for a given level of expected return.

The key insight is that risk and return should be considered together, and diversification can reduce risk without sacrificing expected return.

The main question in portfolio theory is the following:
Given an initial capital $V(0)$, and opportunities (buy or sell) in N securities for investment, how would you allocate the capital $V(0)$ so that the return on the portfolio is optimal in a certain way?

More specifically, what we are looking for is a collection of weights $w_1, w_2, \cdots, w_N$ with $w_1 + w_2 + w_3 + \cdots + w_N = 1$, and $w_i V(0)$ invested in security i for $i = 1, 2, \cdots, N$, such that the return of the portfolio is optimal in a certain sense. Note that these securities can include both risky and risk-free assets. This leads to the concept of efficient portfolios.

General assumptions for this chapter:

- We only consider a single-period model, i.e., the investor only observes the market at two times $t = 0$ (initial time) and $t = 1$ (terminal time).
- Money is invested at the initial time, and the payoff is attained at the end of the period.
- Each investment is measured by two measures only: mean and variance, ignoring the more informative shape of the distribution.

[1] This chapter is based on (Capinski and Zastawniak, 2003, Chapter 5).

- The risk is measured via standard deviation.
Note: The standard deviation $\sigma_K = \sqrt{Var(K)}$ of the return is a more convenient measure of risk than the variance for the following reason: If a quantity is measured in certain units, then the standard deviation will be expressed in the same units, so it can be related directly to the original quantity, in contrast to variance, which will be expressed in squared units.

Some additional assumptions are:

- Investors are rational.
- There are no arbitrage opportunities.
- Access to information is available to all participants.
- The market is liquid.
- There is no transaction cost.
- There are no taxes.
- Everyone has the same opportunity to borrow and lend.

Definition 1 (Short selling)**.** *Sometimes it is possible to sell an asset that you do not own. This process is called short selling, or shorting the asset.*

Short selling is considered quite risky – even dangerous– by many investors. This is because the potential loss is unlimited. If the asset value increases, the loss is $S_1 - S_0$. Since S_1 can increase arbitrarily, so can the loss (S_0= current price of the asset, S_1= price of the asset after 1 year/one time step).

Example 1. *Simple case: Consider two stocks and two time points. (We will generalize later to N stocks.)*

	Stock 1		*Stock 2*	
	Position	*Price*	*Position*	*Price*
$t = 0$	x_1	$S_1(0)$	x_2	$S_2(0)$
$t = 1$	x_1	$S_1(1)$	x_2	$S_2(1)$

Definition 2 (Weights)**.** *We define the* ***weights*** *of the two stocks in the portfolio with initial value $V(0)$ by*

$$w_1 := \frac{x_1 S_1(0)}{V(0)}, \qquad w_2 := \frac{x_2 S_2(0)}{V(0)},$$

i.e., w_i denotes ***the portion/percentage of our initial wealth invested in Stock/Asset*** *i.*

- The value of the initial portfolio thus is $V(0) = x_1 S_1(0) + x_2 S_2(0)$.
- $x_i S_i(0) = w_i V(0)$ is the amount of money invested in stock i today ($t = 0$);
- Trivially, it holds $w_1 + w_2 = 1$.
- $w_1, w_2 \in \mathbb{R}$ could be positive, zero or negative. What does negative w_1 mean? Short selling on stock 1.

- The rate of return of stock i is defined as $K_i := \frac{S_i(1)-S_i(0)}{S_i(0)}$. Often, the rate of return is simply called return.

Proposition 1.1. *For a portfolio consisting of two stocks, it holds that*

$$K_V = w_1 K_1 + w_2 K_2,$$

where K_V, K_1 and K_2 (i.e. $K_V = \frac{V(1)-V(0)}{V(0)}$, $K_i = \frac{S_i(1)-S_i(0)}{S_i(0)}$) are the rate of return of the portfolio, and w_i, $i = 1,2$ are the weights of security i.

"The return of the portfolio is the weighted return of the individual securities."

Proof. Let us initially compute $V(1)$ in terms of the returns K_i:

$$V(1) = x_1 S_1(1) + x_2 S_2(1) = x_1 S_1(0)(1 + K_1) + x_2 S_2(0)(1 + K_2)$$

So,

$$V(1) - V(0) = x_1 S_1(0) K_1 + x_2 S_2(0) K_2.$$

Therefore,

$$\begin{aligned} K_V &= \frac{V(1) - V(0)}{V(0)} = \frac{x_1 S_1(0) K_1 + x_2 S_2(0) K_2}{V(0)} \\ &= \frac{w_1 V(0) K_1 + w_2 V(0) K_2}{V(0)} = w_1 K_1 + w_2 K_2. \end{aligned}$$ □

Proposition 1.2. *For a portfolio consisting of two securities, it holds that*

$$e^{k_V} = w_1 e^{k_1} + w_2 e^{k_2},$$

where $k_V := \ln \frac{V(1)}{V(0)}$, resp. $k_i = \ln \frac{S_i(1)}{S_i(0)}$, $i = 1,2$, is the log return of the portfolio, resp. of stock i, over the time period $[0, 1]$.

Proof. Homework. □

Example 2. *Suppose $V(0) = 100$, $w_1 = 30\%$, $S_1(0) = 55$, $S_2(0) = 42$. If the price of stock 1 drops to $S_1(1) = 45$, then what is the minimal value for $S_2(1)$ to reach to guarantee a portfolio return of at least 10%.*

Proof. Homework. □

1.1 Risk and Expected Return on a Portfolio

We will now extend the concepts of risk and expected return from individual assets to a two-asset portfolio setup.

Definition 3 (Expected return, Risk)**.** *The **expected return** of the asset/security X is defined by $\mu_X := \mathbb{E}[K_X]$, e.g., μ_i for $i = 1, 2$. The **risk** of the asset X is defined as the standard deviation of the return K_X and will be denoted by σ_X. For asset $i = 1, 2$, $\sigma_i := \sqrt{var(K_i)}$.*

For a bond $B, \sigma_B = 0$ (risk-free asset, as we exactly know what money we will get in a year).

*For two assets X, Y, we define their **correlation coefficient**, denoted by ρ_{XY}, as the correlation between K_X and K_Y.*

Remark 1. *Recall that* $\rho_{XY} = \frac{\mathrm{Cov}(K_X, K_Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(K_X - \mathbb{E}[K_X])(K_Y - \mathbb{E}[K_Y])]}{\sigma_X \sigma_Y}$ *for* $\sigma_X \sigma_Y > 0$. *Moreover, if* $|\rho_{XY}| = 1$, *then the returns* K_X, K_Y *are affinely related, i.e., there exist* α, β *s.t.*

$$K_X = \alpha K_Y + \beta,$$

where $sign(\alpha) = sign(\rho_{XY})$. *If* $\rho_{XY} = 0$, *they are said to be uncorrelated (i.e. the variable* X *roughly contains no 'linear' information of the other variable* Y*).*

Proposition 1.3. *For a portfolio consisting of two securities,*

$$\begin{aligned} \mu_V &= w_1 \mu_1 + w_2 \mu_2, \\ \sigma_V^2 &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \mathrm{Cov}(K_1, K_2) \\ &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \rho_{12} \sigma_1 \sigma_2. \end{aligned}$$

Proof: These are known formulae from probability using the formula from Proposition 1.1. Please check if not clear.

Example 3. *Use the following data to find the risk of the portfolio* V *if* $\mu_V = 5.96\%$.

Scenario	*Proba.*	K_1	K_2
A	0.3	−8%	15%
B	0.7	10%	5%

Proof: Let us initially determine the respective expected returns:

$$\mu_1 = 0.3 \cdot (-0.08) + 0.7 \cdot 0.10 = 0.046 \text{ and } \mu_2 = 0.3 \cdot 0.15 + 0.7 \cdot 0.05 = 0.08.$$

$$\begin{cases} w_1 + w_2 = 1 \\ w_1 \mu_1 + w_2 \mu_2 = \mu_V \end{cases} \Leftrightarrow \begin{cases} w_1 = 0.6 \\ w_2 = 0.4 \end{cases}.$$

$$\sigma_V = \left[w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \mathrm{Cov}(K_1, K_2) \right]^{\frac{1}{2}} = \sqrt{0.00097104} = 0.03116.$$

1.2 Risk bounds and Risk reduction: Diversification

Proposition 1.4. *If short selling is not allowed, then* $\sigma_V \le \max\{\sigma_1, \sigma_2\}$.

Proof. The restriction "short selling is prohibited" is equivalent to $x_1, x_2 \ge 0$, which further implies that $w_1, w_2 \ge 0$. Assume, w.l.o.g., that $\sigma_1 > \sigma_2$ or equivalently $\sigma_1 = \max\{\sigma_1, \sigma_2\}$. Then,

$$\begin{aligned} \sigma_V^2 &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 \underbrace{w_1}_{>0} \underbrace{w_2}_{>0} \sigma_1 \sigma_2 \underbrace{\rho_{12}}_{\in[-1,1]} \\ &\le w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \sigma_1 \sigma_2 = (w_1 \sigma_1 + w_2 \sigma_2)^2 \\ &\le (w_1 \sigma_1 + w_2 \sigma_1)^2 = \sigma_1^2. \end{aligned}$$

□

This shows that having two securities/assets in a portfolio can never increase the risk. In fact, it can reduce the portfolio's risk compared to having a single risky asset. This is known as **diversification**.
This process reflects the maxim "Don't put all your eggs in one basket". The effects of diversification can be quantified using the formulas for combining variances.

Example 4. *Suppose that there are n assets, all of which are mutually uncorrelated, i.e., the return of each asset K_i is uncorrelated with that of any other asset K_j in the group. Suppose that the expected rate of return and variance of each of these assets is μ and σ^2, respectively. A portfolio is created by taking equal portions of each of these assets; that is, $w_i = 1/n$ for each i. Calculate the expected rate of return and the risk of this portfolio.*

The rate of return K_V of this portfolio is:

$$K_V = \frac{1}{n}\sum_{i=1}^{n} K_i.$$

The expected rate of return of individual assets is $\mu_i = \mathbb{E}K_i = \mu$, which is independent of i. So the expected rate of return of the portfolio is

$$\mu_V := \mathbb{E}(K_V) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu.$$

The corresponding risk of the portfolio is

$$\sigma_V = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n} \sigma^2} = \frac{\sigma}{\sqrt{n}}.$$

(Note, the above formula uses the fact that the individual returns are uncorrelated.)
So the variance decreases rapidly as n increases.

Homework: (Effect of diversification with uncorrelated assets) In the context of the above example, how many assets do you need to reduce the portfolio risk to below or equal to 5%?

Homework: Show that when all the assets are positively correlated and at least two of the assets are not perfectly correlated, the risk of the portfolio is still lower than σ but higher than the uncorrelated case. Also, show that if the assets are negatively correlated, then the risk decreases even further from the uncorrelated case.

Type of Risk

Risk comes in two types:

- Diversifiable risk: also known as 'unsystematic/ idiosyncratic risk'.

  Some examples are:
  Business Risk: The risk associated with the operational efficiency and management of a company. For example, poor management decisions, production failures, or supply chain disruptions can adversely affect a company's profitability.
  Internal Business Risk: Risks related to the internal operations of the company, such as operational inefficiencies or poor decision-making.

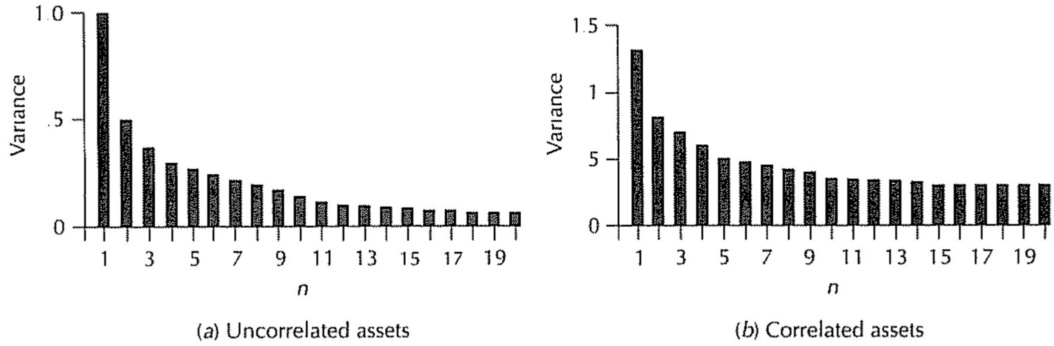


FIGURE 6.7 Effects of diversification. If assets are uncorrelated, the variance of a portfolio can be made very small If assets are positively correlated, there is likely to be a lower limit to the variance that can be achieved

External Business Risk: Risks related to external factors, such as competition, changes in consumer preferences, or economic conditions affecting the industry.
Financial Risk: The risk associated with a company's capital structure, particularly the use of debt. High levels of debt increase the risk of financial distress if the company cannot meet its interest obligations.

- Non-diversifiable (market, systematic) risk due to macro (business cycle, inflation, etc.) / market conditions (liquidity): also known as 'systematic risk'.

  Some examples are:
  Market Risk: The risk of losses due to overall market movements. For example, during a market downturn, the prices of most stocks tend to fall, regardless of the company's individual performance.
  Interest Rate Risk: The risk that changes in interest rates will affect the value of investments, particularly bonds. When interest rates rise, bond prices typically fall.
  Inflation Risk: The risk that inflation will erode the purchasing power of returns. High inflation can reduce the real value of future cash flows from investments.
  Political Risk: The risk of political instability or changes in government policy that could negatively impact the financial markets.

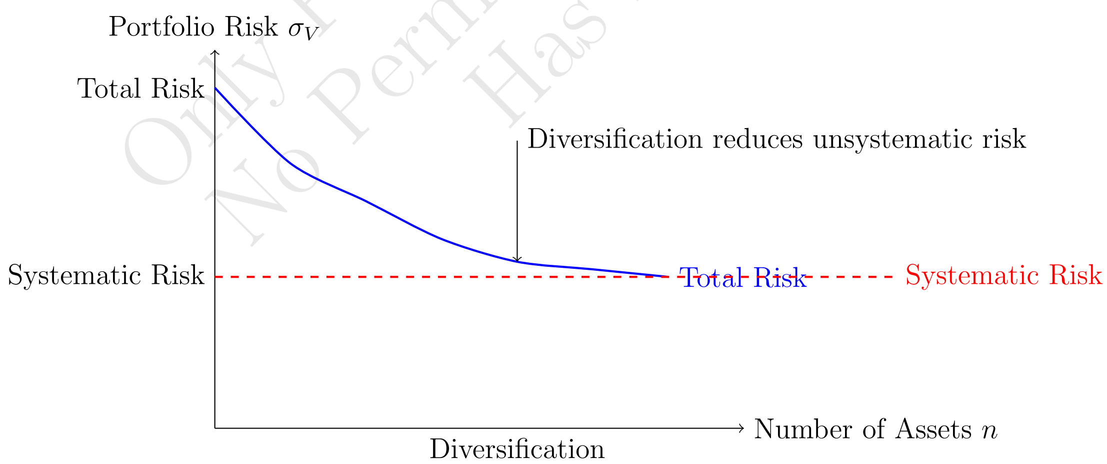


Figure 1: Effect of diversification on portfolio risk. As the number of assets n increases, unsystematic risk decreases, leaving only systematic risk.

Diversification only reduces unsystematic risk. For better understanding, let us look at the following example:

Example 5. *Consider an equally-weighted portfolio of n assets (i.e. $w_i = 1/n$ for all $i = 1, \cdots n$). We denote the variance of the return of asset i as $\sigma_{ii} = \sigma_i^2$, and the covariance of the return of asset i and asset j by $\sigma_{ij} := Cov(K_i, K_j)$. Then, the variance of the return of the portfolio can be written as*

$$\begin{aligned}\sigma_V^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \\ &= \sum_{i=1}^{n}\left(\frac{1}{n}\right)^2 \sigma_i^2 + \sum_{i=1}^{n}\sum_{i\neq j=1}^{n}\left(\frac{1}{n}\right)^2 \sigma_{ij} \\ &= \left(\frac{1}{n}\right) \times \underbrace{\left[\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\right]}_{\text{average variance}} + \left(\frac{n^2-n}{n^2}\right) \times \underbrace{\left[\frac{1}{n^2-n}\sum_{i=1}^{n}\sum_{j(\neq i)=1}^{n}\sigma_{ij}\right]}_{\text{average covariance}}.\end{aligned}$$

As $n \to \infty$:
- *Contribution of variance terms goes to zero. –'Unsystematic risk'*
- *Contribution of covariance terms goes to "average covariance". – 'Systematic risk'*

2 Expected return and risk interplay

Recall once again that $-1 \leq \rho_{12} \leq 1$. In the following, we will assume $\sigma_1, \sigma_2 \neq 0$, ρ_{12} to be fixed in order to examine the behavior of the function

$$w_1 \mapsto \sigma_V = [w_1^2\sigma_1^2 + w_2{}^2\sigma_2^2 + 2w_1 w_2 \sigma_1\sigma_2\rho_{12}]^{\frac{1}{2}}.$$

Replacing $w_2 = 1 - w_1$ we get:

$$w_1 \mapsto \sigma_V = [w_1^2\sigma_1^2 + (1-w_1)^2\sigma_2^2 + 2w_1(1-w_1)\sigma_1\sigma_2\rho_{12}]^{\frac{1}{2}}.$$

In this case, we will discuss the perfect positive or negative correlation between the **securities in the portfolio**.

Risk Reduction (minimize σ_V) - The case $\rho_{12} = 1$ or $\rho_{12} = -1$

Proposition 2.1. *i) If $\rho_{12} = 1$, then $\sigma_V = |w_1\sigma_1 + w_2\sigma_2|$. In particular,*

$$\boxed{\sigma_V = 0} \quad \textit{iff} \quad \boxed{\sigma_1 \neq \sigma_2,\ w_1 = \frac{-\sigma_2}{\sigma_1 - \sigma_2} \quad \text{and} \quad w_2 = \frac{\sigma_1}{\sigma_1 - \sigma_2}}.$$

Observe that either w_1 or w_2 is negative. Therefore, short selling has to be allowed; otherwise, we cannot attain 0 risk case.

ii) If $\rho_{12} = -1$, then $\sigma_V = |w_1\sigma_1 - w_2\sigma_2|$. In particular,

$$\boxed{\sigma_V = 0} \quad \textit{iff} \quad \boxed{w_1 = \frac{\sigma_2}{\sigma_1 + \sigma_2} \quad \text{and} \quad w_2 = \frac{\sigma_1}{\sigma_1 + \sigma_2}}.$$

Observe that both weights are positive. Therefore, no short selling is required to achieve zero risk case.

Can both w_1 and w_2 be negative? No, since $w_1 + w_2 = 1$.

Proof. i) Since $\rho_{12} = 1$, we immediately get $\sigma_V = |w_1\sigma_1 + w_2\sigma_2|$. So, $\sigma_V = 0$ if and only if

$$\begin{cases} w_1 + w_2 = 1 \\ w_1\sigma_1 + w_2\sigma_2 = 0 \end{cases} \Leftrightarrow \begin{cases} w_1 = -\frac{\sigma_2}{\sigma_1 - \sigma_2} \\ w_2 = \frac{\sigma_1}{\sigma_1 - \sigma_2} \end{cases}$$

Observe that in the case where $\sigma_1 = \sigma_2 \neq 0$, the system has no solution.

ii) Analogous computation leads to the result; verify it on your own!

□

Definition 4 (Affine combination, convex combination)**.** *Let A, B be points on the plane.*

- *Every point on the plane that can be written as $xA + (1-x)B$ for some $x \in \mathbb{R}$, is called **affine combination** of A and B.*
- *Every point on the plane that can be written as $xA + (1-x)B$ for some $x \in [0, 1]$, is called **convex combination** of A and B.*

An asset (with deterministic/ non-deterministic rate of returns) can be represented on a two-dimensional diagram. An asset with an expected rate of return μ and standard deviation σ can be represented (uniquely) as a point in a $\sigma - \mu$ diagram. Conventionally, the horizontal axis is used for the risk, and the vertical axis is used for the expected return. This diagram is called a standard deviation-mean diagram, or simply σ/μ diagram.

Each portfolio can be represented by a point with coordinates σ_V and μ_V on the σ, μ plane.

Risk Reduction (minimize σ_V) - The case $-1 < \rho_{12} < 1$

Theorem 2.2. *Let ρ_{12} be fixed and s.t. $-1 < \rho_{12} < 1$. Then, the function $w_1 \mapsto \sigma_V(w_1)$ attains its minimum value (may not be 0) at (it is probably easier to just remember the matrix formula we do later in the course)*

$$w_1^{\min} = \frac{\sigma_2^2 - \rho_{12}\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2} = \frac{\sigma_2^2 - \mathrm{Cov}(K_1, K_2)}{\sigma_1^2 + \sigma_2^2 - 2\mathrm{Cov}(K_1, K_2)}.$$

If short selling is not allowed, the smallest risk is attained when

$$(w_1, w_2) = \begin{cases} (0, 1), & \text{if } w_1^{\min} < 0 \\ (w_1^{\min}, 1 - w_1^{\min}), & \text{if } 0 \leq w_1^{\min} \leq 1 \\ (1, 0) & \text{if } 1 < w_1^{\min} \end{cases}.$$

Proof. Instead of minimizing $\sigma_V(w_1)$ we can equivalently minimize $\sigma_V^2(w_1)$, since the function mapping $x \mapsto x^2$ is strictly increasing on $[0, \infty)$. It is easy to verify that

i) $\frac{\partial \sigma_V^2}{\partial w_1}(w_1^{\min}) = 0$ and

ii) $\frac{\partial^2 \sigma_V^2}{\partial w_1^2}(w_1^{\min}) > 0$.

The cases where no short selling is allowed follow immediately. □

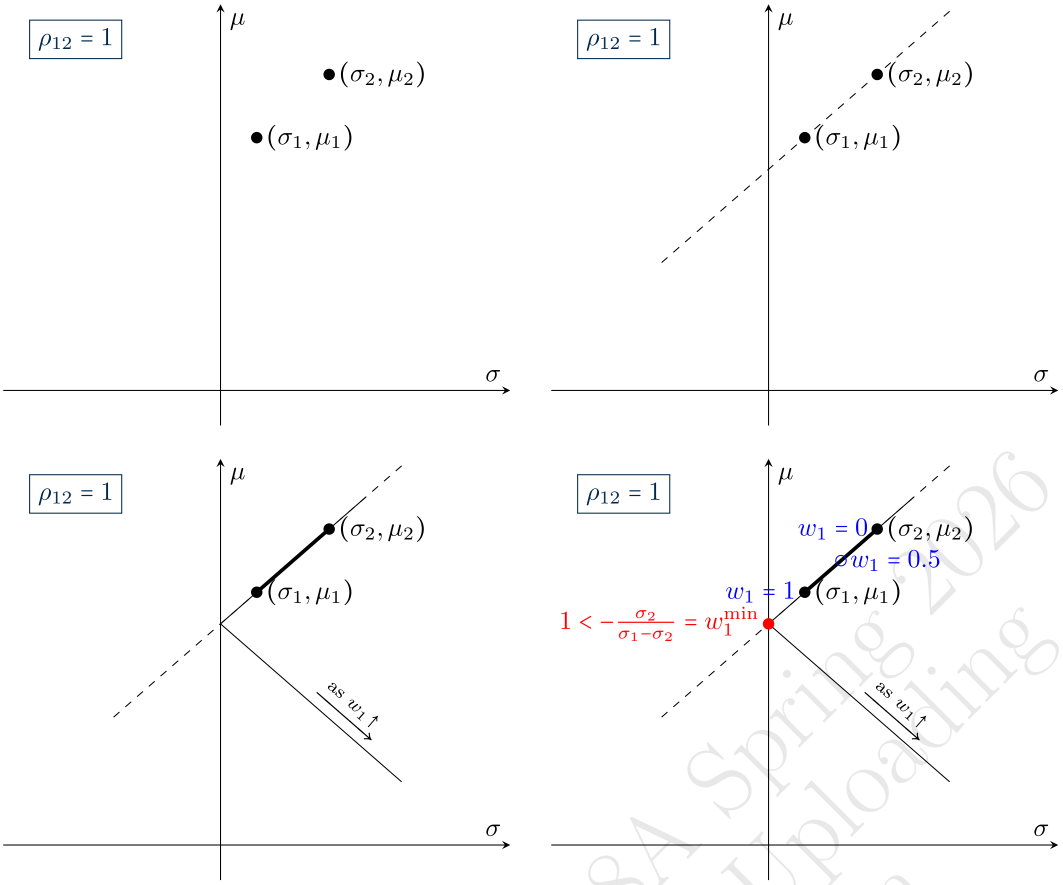


Figure 2: Positive correlation

Remark 2. *The minimum-variance portfolio depends only on the covariance structure of asset returns and not on expected returns.*

Remark 3. *Suppose $\sigma_1 \leq \sigma_2$, then we summarize the different cases as follows:*

Cases	*Is there a portfolio with $\sigma_V < \sigma_1 (\leq \sigma_2)$?*	*short selling needed?*
1) $-1 \leq \rho_{12} < \frac{\sigma_1}{\sigma_2}$	*yes*	*no*
2) $\rho_{12} = \frac{\sigma_1}{\sigma_2}$	*no*	
3) $\frac{\sigma_1}{\sigma_2} < \rho_{12} \leq 1$	*yes*	*yes*

Definition 5. *The collection of all portfolios that can be constructed by investing in given assets is called the **feasible (or attainable)** set.*

Remark 4. *When a portfolio consists of two assets, each point on the feasible set can be uniquely determined by the pair (w_1, w_2).*

Remark 5 (Common pitfalls)**.** *Some common pitfalls are:*

- *Low individual asset risk does not imply low portfolio risk.*
- *High expected return does not guarantee efficiency.*
- *Diversification fails when assets are highly correlated.*

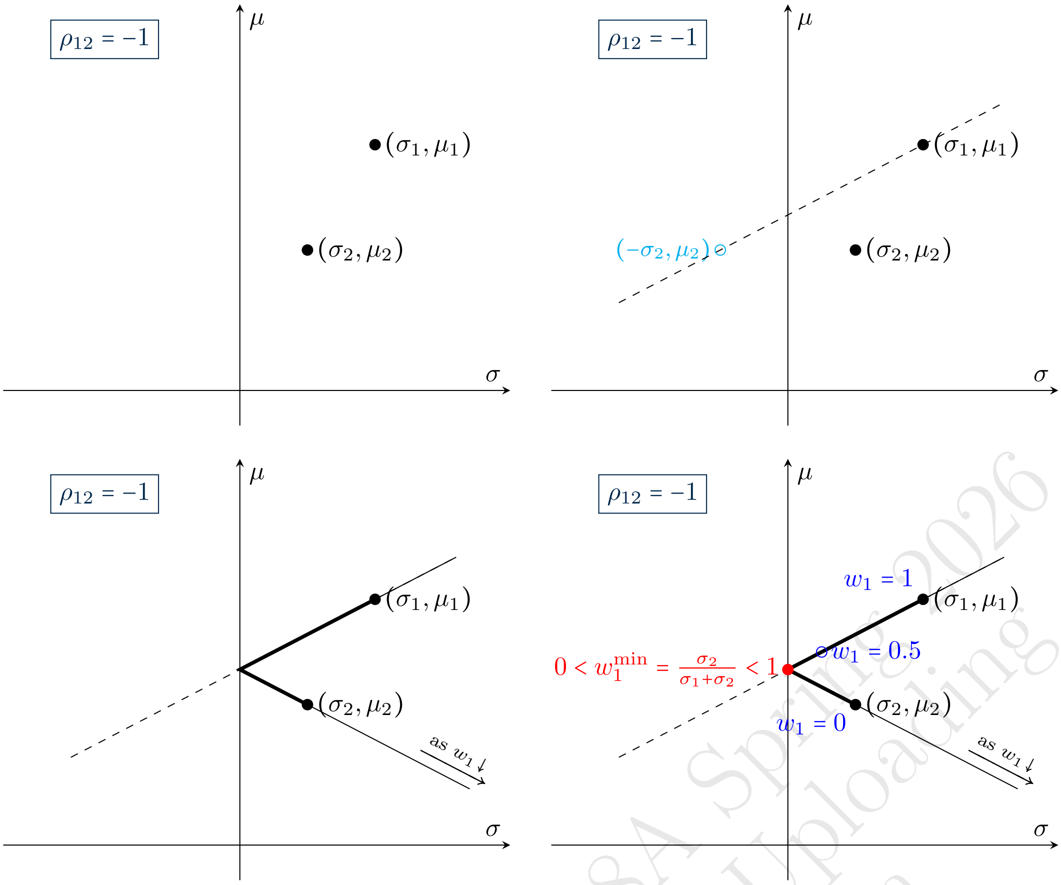
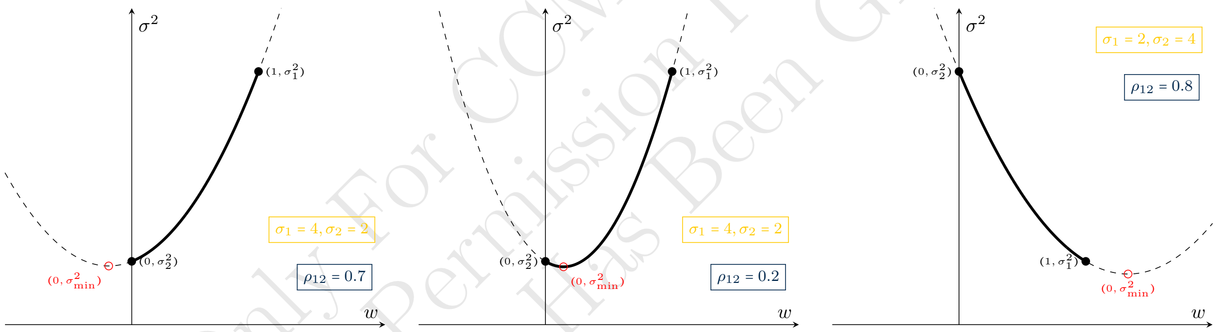
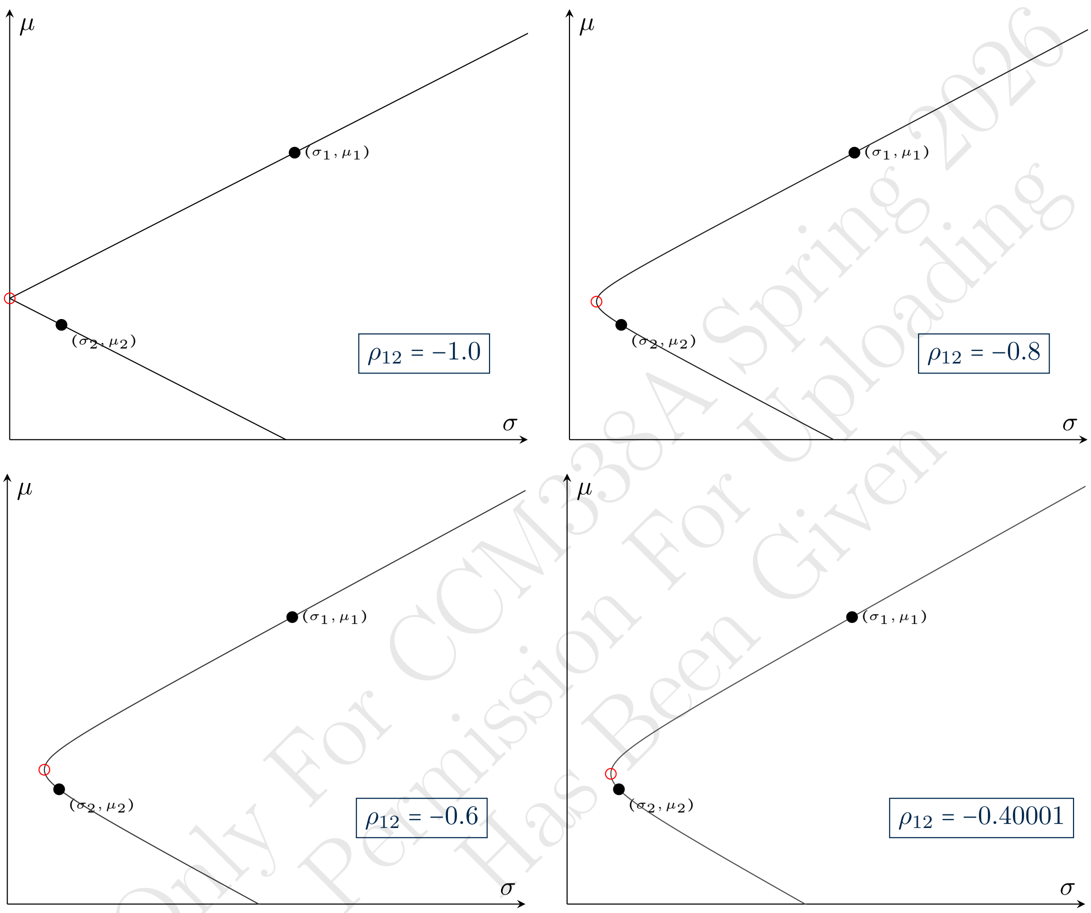


Figure 3: Negative correlation



2.1 Diagram of feasible portfolio



μ
σ
(σ_1, μ_1)
(σ_2, μ_2)
$\rho_{12} = -1.0$
μ
σ
(σ_1, μ_1)
(σ_2, μ_2)
$\rho_{12} = -0.8$
μ
σ
(σ_1, μ_1)
(σ_2, μ_2)
$\rho_{12} = -0.6$
μ
σ
(σ_1, μ_1)
(σ_2, μ_2)
$\rho_{12} = -0.40001$

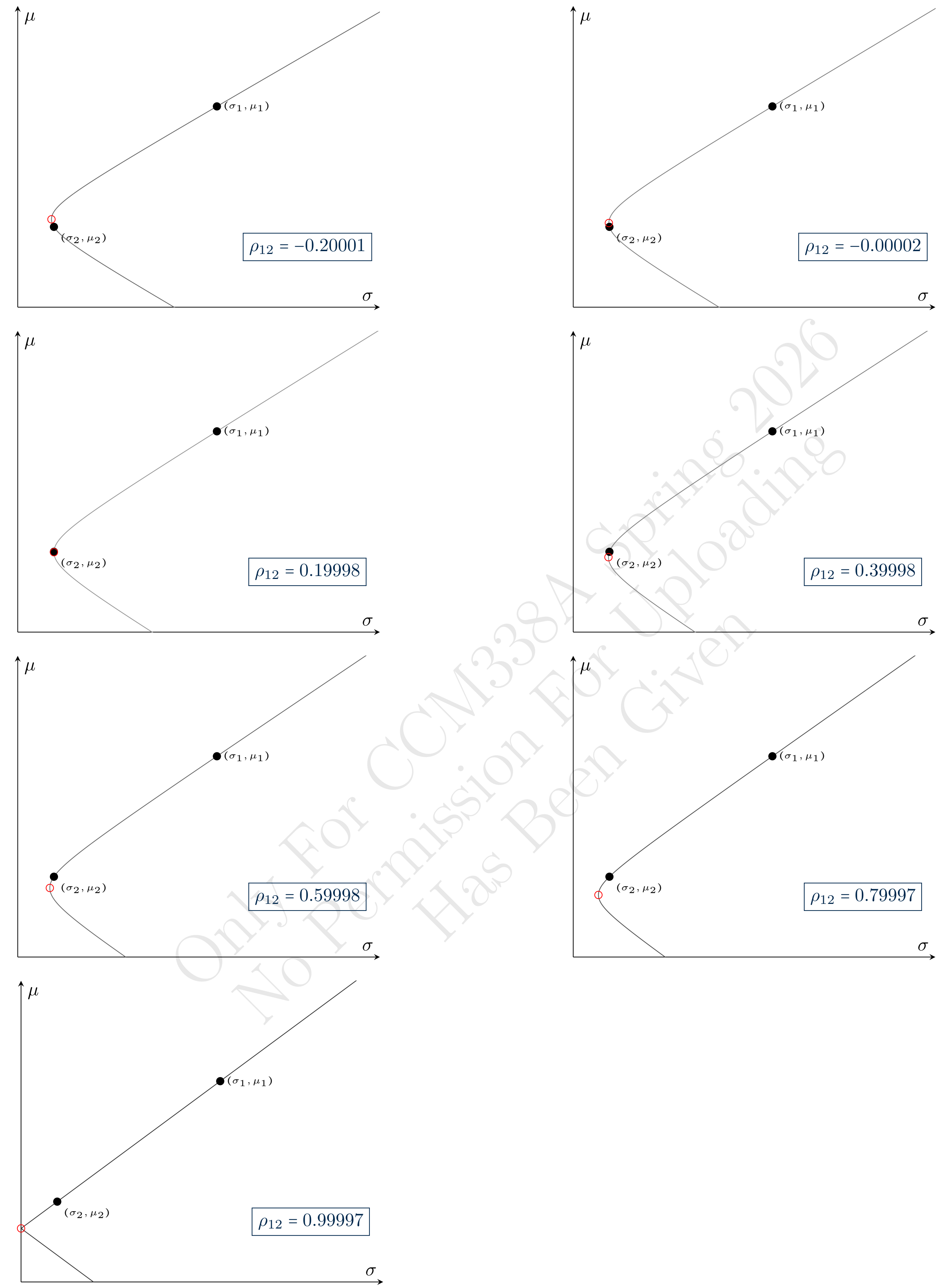


Figure 4: Shape of the feasible set for different values of the correlation coefficient between the two securities.

2.2 Shape of the feasible/attainable portfolios

Theorem 2.3 (Expected return and risk for MVP)**.** *If $\sigma_1 \neq \sigma_2$, the expected return and risk for MVP*[2] *are given by (It is probably easier to just remember the matrix formula we do later in the course)*

$$\mu_{MVP} = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2 - (\mu_1+\mu_2)\rho_{12}\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2-2\rho_{12}\sigma_1\sigma_2}$$

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

Proof. This can be seen using the results from Theorem 2.2. □

Remark 6. *If both assets are risky, then the variance for the minimal-variance portfolio is strictly smaller than either of the individual asset variances,*

$$\sigma_{MVP}^2 < \min\{\sigma_1^2, \sigma_2^2\},$$

unless $\rho = \frac{\min\{\sigma_1,\sigma_2\}}{\max\{\sigma_1,\sigma_2\}}$*, in which case*

$$\sigma_{MVP}^2 = \min\{\sigma_1^2, \sigma_2^2\}.$$

(This includes the case $\sigma_1^2 = \sigma_2^2$ and $\rho = 1$). In particular,

$$\sigma_{MVP}^2 < \min\{\sigma_1^2, \sigma_2^2\} \quad \textit{whenever} \quad \rho < 0.$$

Proof: Homework.
Hint: Assume wlog $\sigma_1 \leq \sigma_2$. Then try to show $\sigma_1^2 - \sigma_{MVP}^2 \geq 0$ and > 0 respectively.

Theorem 2.4 (Shape of the feasible set)**.** *Assume that $-1 < \rho_{12} < 1$ and $\mu_1 \neq \mu_2$. Then each portfolio V on a feasible set $x = \sigma_V$ and $y = \mu_V$ satisfy an equation of a* ***hyperbola****:*

$$x^2 - A^2(y - \mu_{MVP})^2 = \sigma_{MVP}^2$$

with $A^2 = \frac{\sigma_1^2+\sigma_2^2-2\rho_{12}\sigma_1\sigma_2}{(\mu_1-\mu_2)^2} > 0.$

Acknowledgement: I would like to thank Martin Dattge for proofreading the lecture notes and for his valuable feedback.

References

Capinski, M. and Zastawniak, T. (2003). Mathematics for finance. *An Introduction*, pages 118–124.

[2] MVP = Minimum variance portfolio

Mathematical Finance II (6CCM338A)-Portfolio management- the case of several securities

Purba Das

March 27, 2026

Contents

[1]

Acknowledgement: I would like to thank Martin Dattge for proofreading the lecture notes and for his valuable feedback.

[1] Some of the materials are taken from previous lecture notes of Dr Ryan Donnelly.

1 Recap – Covariance and correlation

Please revise the basics of matrix algebra, variance, covariance and correlation coefficients from previous courses.

Definition 1 (Covariance)**.** *For two random variables X and Y with finite variance, their covariance is given by*

$$\mathbb{C}[X,Y] = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)].$$

Note: From the definition $\mathbb{C}[X,X] = Var[X]$. Covariance can also be written as

$$\mathbb{C}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Sample covariance is given by

$$\bar{\sigma}_{X,Y} := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{\mu}_X)(Y_i - \bar{\mu}_Y).$$

Definition 2 (Correlation)**.** *For two random variables X and Y with finite variance, their correlation (also known as correlation coefficient) is given by*

$$\rho_{XY} = \rho[X,Y] = \frac{\mathbb{C}[X,Y]}{\sqrt{Var[X]Var[Y]}}.$$

Given a collection of finite variance random variables $\{X_i\}_{i=1}^n$, pairwise covariances can be put into an $n \times n$ matrix Σ with (i,j)-th component given as $\Sigma_{i,j} = \mathbb{C}[X_i, X_j]$. i.e.

$$\Sigma := \begin{pmatrix} \mathbb{C}(X_1,X_1) & \cdots & \mathbb{C}(X_1,X_n) \\ & \ddots & \\ \mathbb{C}(X_n,X_1) & \cdots & \mathbb{C}(X_n,X_n) \end{pmatrix}.$$

The matrix Σ is positive semi-definite. One can construct a correlation matrix in the same way:

$$\rho = \left(\rho_{i,j}\right)_{i,j=1}^{n} \text{ with } \rho_{i,j} = \rho[X_i, X_j].$$

The following relation is easy to check using matrix algebra (HW):

$$\rho = diag(\Sigma)^{-\frac{1}{2}}\;\Sigma\; diag(\Sigma)^{-\frac{1}{2}}.$$

The matrix ρ is also positive semi-definite and its diagonal consists of 1's. *We often use $\mathbb{C}$ to use the covariance matrix instead of using the notation Σ. Some books/notes use Σ to represent the covariance matrix.*

1.1 Properties of correlation coefficient

Correlation (and covariance) is a measurement of how two variables tend to move together. More precisely, how linearly two random variables are linked. It is the most commonly used dependence measure.

1. If X and Y have positive (negative) correlation, then large values of X tend to happen simultaneously with large (small) values of Y.

2. The Correlation is bounded: $\rho \in [-1, 1]$. This can be easily verified using the Cauchy-Schwarz inequality.

3. Independence of two random variables implies $\rho = 0$.

4. The converse is false in general, but sometimes true if additional information is given (HW: construct an example when the converse is false, and another example when it is true).

5. If two variables have $\rho = 1$, then they must be linearly related, i.e.,

$$X = aY + b, \quad a > 0.$$

6. Likewise, if $\rho = -1$, then we still have a linear relation like before but with $a < 0$.

2 Mean-Variance portfolio – n assets with $n \geq 2$

Suppose that the market has n different assets. For simplicity, you can think of all of these representing different stocks. Then, a portfolio V consisting of only these n assets can be described by the corresponding **weights** (in these stocks)

$$w_i = \frac{x_i S_i(0)}{V(0)}, \qquad i = 1, \ldots, n,$$

where x_i denotes the position in stock i and portfolio value at the time $t = 0$ is denoted as $V(0) = \sum_{i=1}^n x_i S_i(0)$. Note that, $\sum_{i=1}^n w_i = w_1 + \cdots + w_n = 1$ (recall for two securities: $w_1 + w_2 = 1$). Let K_i be the return of stock i, then we have

$$K_V = w_1 K_1 + \cdots + w_n K_n = \sum_{i=1}^n w_i K_i = \begin{pmatrix} w_1 & \cdots & w_n \end{pmatrix} \begin{pmatrix} K_1 \\ \vdots \\ K_n \end{pmatrix}.$$

(recall for two securities: $K_V = w_1 K_1 + w_2 K_2$)

Definition 3 (Notations/ Terminology)**.** *We define the following vectors/matrices:*

- *Define the **weight (row) vector** $\boldsymbol{w} := (w_1, \ldots, w_n)_{1\times n}$ and $u = (1, \ldots, 1)_{1\times n}$, thus we get $u\boldsymbol{w}^\top = (1, \ldots, 1)(w_1, \cdots, w_n)^\top = 1$;*
- *Define the **mean vector** $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)_{1\times n}$ where $\mu_i = \mathbb{E}(K_i)$.*
- *Define the $n \times n$ covariance matrix $\mathbb{C} = (c_{ij})_{n\times n}$, where $c_{ij} = \mathrm{Cov}(K_i, K_j)$.*

Remark 1.

- *$\mathbb{C}$ is symmetric (the transpose $\mathbb{C}^T = \mathbb{C}$) and positive semi-definite (recall this means $x\mathbb{C}x^T \geq 0$, for any real vector $x = (x_1, \cdots, x_n)_{1\times n} \in \mathbb{R}^n$).*
- *The diagonal elements are the same as variance, i.e., $c_{ii} = \mathrm{Cov}(K_i, K_i) = \mathrm{Var}(K_i) = \sigma_i^2$.*

Theorem 2.1. *The expected return $\mu_V = E(K_V)$ and variance $\sigma_V^2 = \mathrm{Var}(K_V)$ of a portfolio with weights w are given by*

$$\mu_V = \mu w^\top = (\mu_1, \cdots, \mu_n)(w_1, \cdots, w_n)^\top = \sum_{i=1}^n w_i \mu_i;$$

$$\sigma_V^2 = w\mathbb{C}w^\top = \sum_{i,j=1}^n w_i w_j Cov(K_i, K_j) = \sum_{i,j=1}^n w_i w_j \rho_{i,j} \sigma_i \sigma_j.$$

Note: The risk of the portfolio is given as $\sigma_V = \sqrt{w\mathbb{C}w^\top}$

Example 1. *Given* $w_1 = 50\%$, $w_2 = -25\%$, $w_3 = 75\%$; $\mu_1 = 10\%$, $\mu_2 = 7\%$, $\mu_3 = 12\%$; $\sigma_1 = 1$, $\sigma_2 = 0.5$, $\sigma_3 = 2$; *and* $\rho_{12} = -0.5$, $\rho_{13} = 0.5$, $\rho_{23} = 0$, *find the expected return* μ_V *and the risk* σ_V *of the portfolio.*

Proof. Writing in matrix notation:

$$w = (w_1, w_2, w_3) = (50\%, -25\%, 75\%), \quad \mu = (\mu_1, \mu_2, \mu_3) = (10\%, 7\%, 12\%),$$
$$\mathbb{C} = \begin{pmatrix} 1 & -0.25 & 1 \\ -0.25 & 0.25 & 0 \\ 1 & 0 & 4 \end{pmatrix}$$

$$\mu_V = E(K_V) = \mu w^T = 0.1225 (= 12.25\%),$$
$$\sigma_V^2 = w\mathbb{C}w^T = \frac{207}{64} \Rightarrow \sigma_V = 1.7984.$$

□

Another potential reformulation of the above example:

Example 2. *An investor holds a portfolio of three assets with weights* $w_1 = 0.5, w_2 = -0.25$ *and* $w_3 = 0.75$. *The assets have expected returns and standard deviations*

$$\mu_1 = 10\%, \ \mu_2 = 7\%, \ \mu_3 = 12\%, \quad \sigma_1 = 1, \ \sigma_2 = 0.5, \ \sigma_3 = 2,$$

and correlations $\rho_{13} = 0.5$, $\rho_{23} = 0$, *while* ρ_{12} *is unknown. The portfolio standard deviation is observed to be* $\sigma_V = 1.824$.

(a) Find the covariance $\mathrm{Cov}(K_1, K_2)$ *and correlation coefficient* $\rho_{1,2}$.

(b) Compute the expected return μ_V *of the portfolio.*

Definition 4 (Minimum Variance Portfolio (MVP))**.** *Among all feasible portfolios, the portfolio with the smallest risk (aka variance) is called the minimum variance portfolio (MVP). Its weight vector, expected return and risk are denoted by* $w_{\mathrm{MVP}}, \mu_{\mathrm{MVP}}, \sigma_{\mathrm{MVP}}$ *respectively.*

Remark 2. *Recall that, in the case of* $n = 2$, *we determined (Follow the class note for details)*

$$w_1^{\min} = \frac{\sigma_2^2 - \rho_{12}\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2} \quad \text{and} \quad w_2^{\min} = 1 - w_1^{\min}.$$

So,

$$w_1^{\min} = \frac{\sigma_{22} - \sigma_{12}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} \quad \text{and} \quad w_2^{\min} = \frac{\sigma_{11} - \sigma_{12}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}}.$$

Rewriting this gives us (under the assumption $\det(\mathbb{C}) \neq 0$, *i.e.,* $\mathbb{C}^{-1}$ *exists):*

$$w_{\mathrm{MVP}} := (w_1^{\min}, w_2^{\min}) = \frac{u\mathbb{C}^{-1}}{u\mathbb{C}^{-1}u^\top} \quad \text{with,} \ \ u w_{\mathrm{MVP}}^\top = w_{\mathrm{MVP}} u^\top = 1.$$

In order to write the above formula in a compact form through matrices, the following observations are useful:

1. $|\mathbb{C}| = \det(\mathbb{C}) := \begin{vmatrix} \sigma_1^2 & \mathrm{Cov}(K_1, K_2) \\ \mathrm{Cov}(K_1, K_2) & \sigma_2^2 \end{vmatrix} = \sigma_1^2\sigma_2^2(1 - \rho_{12}^2)$. *Note that,* $\det(\mathbb{C}) \neq 0$ *as long as the assets are not perfectly correlated.*

2. *Then,*
$$\mathbb{C}^{-1} := \frac{1}{\det(\mathbb{C})} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix}$$

3. *Therefore, for* $u := (1, 1)$, *we get* $\det(\mathbb{C}) \cdot u\mathbb{C}^{-1} = (\sigma_{22} - \sigma_{12}, \sigma_{11} - \sigma_{12})$ *and consequentially,* $\det(\mathbb{C}) \cdot u\mathbb{C}^{-1}u^\top = \sigma_{22} + \sigma_{11} - 2\sigma_{12}$.

HW: Show that for any covariance matrix $\mathbb{C}_{n\times n}$ with inverse, we have $(\mathbb{C}^{-1})^\top = \mathbb{C}^{-1}$.

Theorem 2.2 (n securities)**.** *Assume that the determinant of the covariance matrix is not zero, i.e.,* $\det(\mathbb{C}) \neq 0$. *Then the weights of the minimum variance portfolio are given by the vector (assuming* $u = (1, \dots, 1)$ *is the* $1 \times n$ *row vector)*

$$w_{\mathrm{MVP}} = \frac{u\mathbb{C}^{-1}}{u\mathbb{C}^{-1}u^\top}.$$

Proof. We want to minimize the risk/standard division/variance of the portfolio. This means we want to solve the problem

$$\min_w \underbrace{w\mathbb{C}w^\top}_{\text{var of portfolio}} \quad \text{subject to the constraint } \sum_{i=1}^n w_i = uw^\top = 1.$$

Since we have a minimization problem with constraints, we will use the Lagrange multipliers method. To this end, define

$$\mathcal{L}(w, \lambda) := w\mathbb{C}w^\top - \lambda(uw^\top - 1).$$

Then, for $\nabla_w := (\frac{\partial}{\partial w_1}, \dots, \frac{\partial}{\partial w_N})$, we get

$$\nabla_w\mathcal{L} = 0 \Leftrightarrow 2w\mathbb{C} - \lambda u = 0 \Leftrightarrow w = \frac{\lambda}{2}u\mathbb{C}^{-1}. \tag{1}$$

Substitute, now, the above value of $w = w(\lambda)$ in the constraint in order to determine the λ. This provides

$$1 = uw^\top = u\big(\frac{\lambda}{2}u\mathbb{C}^{-1}\big)^\top = \frac{\lambda}{2}u\mathbb{C}^{-1}u^\top \Leftrightarrow \lambda = \frac{2}{u\mathbb{C}^{-1}u^\top}. \tag{2}$$

Finally, use (1) & (2) in order to derive the weights

$$w_{\mathrm{MVP}} = \frac{u\mathbb{C}^{-1}}{u\mathbb{C}^{-1}u^\top}. \tag{3}$$

Since Hessian$(\mathcal{L}) = \mathbb{C}$ is positive definite, we know that the unique critical point determines the unique constrained minimum. □

Remark 3. *Sometimes the* w_{MVP} *is also written as (assuming* u *is a row vector as before)*

$$w_{\mathrm{MVP}} = \frac{\mathbb{C}^{-1}u^\top}{u\mathbb{C}^{-1}u^\top}.$$

Theorem 2.3. *The MVP has the following mean and variance (Expected return and risk) (assuming u, μ are row vectors).*

$$\mu_{\text{MVP}} = \frac{u\mathbb{C}^{-1}\mu^\top}{u\mathbb{C}^{-1}u^\top}. \tag{4}$$

$$\sigma^2_{\text{MVP}} = \frac{1}{u\mathbb{C}^{-1}u^\top}. \tag{5}$$

Remark: Note all the u's in (3), (4) and (5) are row vectors.

Proof. This is a simple consequence of the weights for MVP calculated in the previous theorem. □

2.1 The feasible set of portfolio: Markowitz bullet / Minimum Variance Line (MVL)

Definition 5 (Feasible Set)**.** *The set of points (μ_V, σ_V) corresponding to a portfolio that can be constructed by the initial n assets/securities $(\mu_1, \sigma_1), (\mu_2, \sigma_2), \dots, (\mu_n, \sigma_n)$ are called **feasible set or feasible region**.*

Remark 4. *1. The bold line(s) of the Figure(s) 1 is(are) called the Markowitz bullet; it was named after Harry Markowitz in conjunction with its shape ("bullet").*

2. This is also the left boundary of a feasible set.

*3. An important feature of the Markowitz bullet is that each point on it has **the smallest risk among all portfolios** with the same expected return at the point.*

*4. **This is why the Markowitz bullet is also called the minimum variance line/set.***

5. If there are at least 3 risky assets (not perfectly correlated and with different μ), then the feasible set is a solid two-dimensional region.

6. The feasible region is convex to the left. That is, given any two points in the region, the straight line connecting them does not cross the left boundary of the feasible region. This is because the minimum variance curve in the mean-variance plot is a parabolic curve.

Definition 6 (MVL)**.** *The set of the portfolios with minimum risk for a given level of expected return is called the Minimum Variance Line (MVL).*

2.2 Link to Risk-averse and Risk-seeking

Any rational investor will prefer the portfolio corresponding to the leftmost point on the line, that is, the point with the smallest standard deviation for the given mean. An investor who agrees with this viewpoint is said to be risk-averse, since they seek to minimise risk (as measured by standard deviation). An investor who would select a point apart from the one of minimum standard deviation is said to be risk-seeking (The Markowitz theory only deals with risk-averse investors).

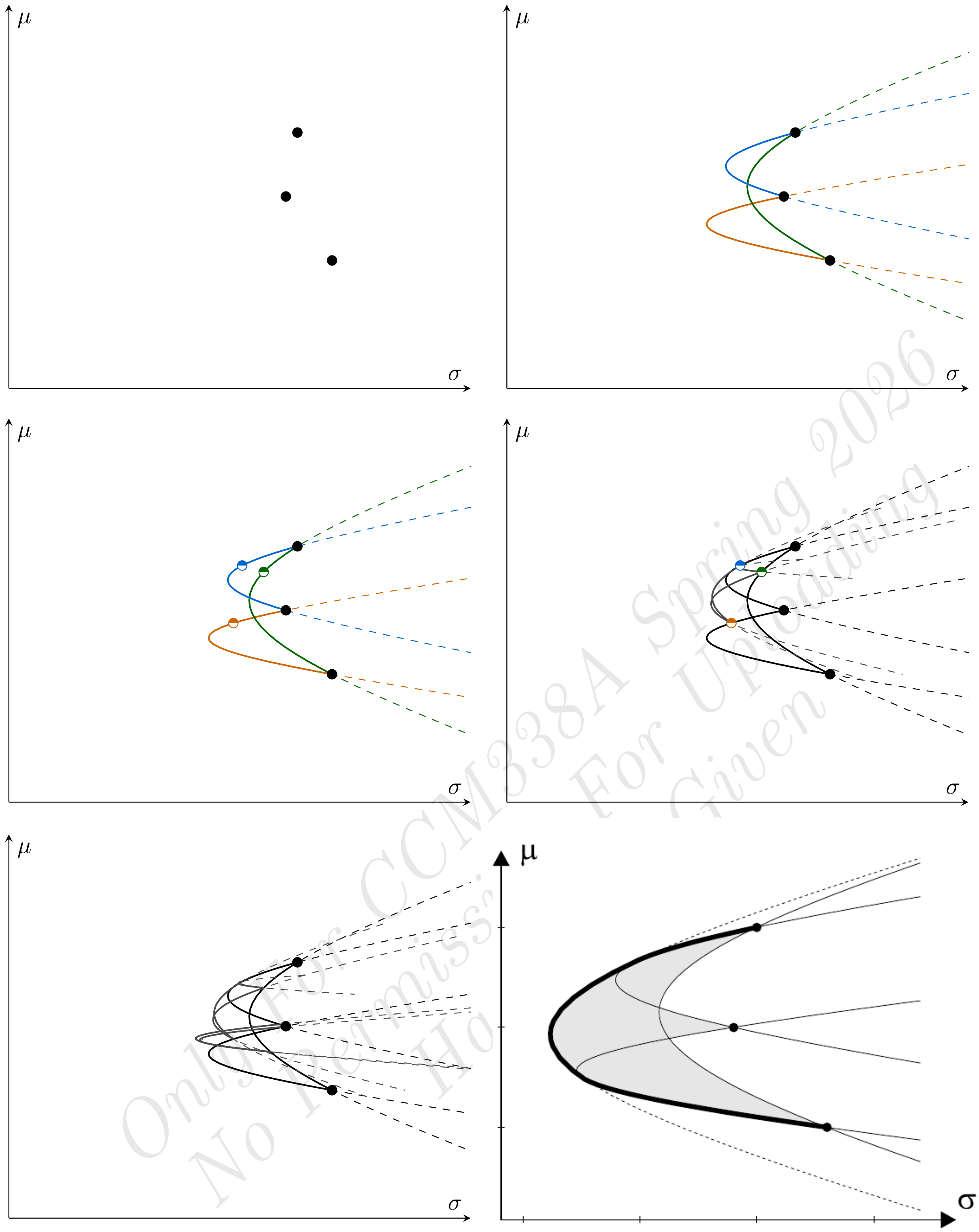


Figure 1: The feasible set and the Markowitz bullet **without short selling**

Theorem 2.4 (Weights of portfolios on the minimal variance line)**.** *Assume that* $\det \mathbb{C} \neq 0$. *Fix a level of expected return* μ_V. *Then a portfolio has the smallest risk among all portfolios with expected* *return* μ_V *if and only if it has the weights*

$$w_{\mu_V}^{\min} = \frac{\begin{vmatrix} 1 & u\mathbb{C}^{-1}\mu^\top \\ \mu_V & \mu\mathbb{C}^{-1}\mu^\top \end{vmatrix} \mathbb{C}^{-1}u^\top + \begin{vmatrix} u\mathbb{C}^{-1}u^\top & 1 \\ \mu\mathbb{C}^{-1}u^\top & \mu_V \end{vmatrix} \mathbb{C}^{-1}\mu^\top}{\begin{vmatrix} u\mathbb{C}^{-1}u^\top & u\mathbb{C}^{-1}\mu^\top \\ \mu\mathbb{C}^{-1}u^\top & \mu\mathbb{C}^{-1}\mu^\top \end{vmatrix}}, \tag{6}$$

where we have used the usual notation for the determinant of a 2×2 *matrix.*

Proof. We want to solve the problem

$$\min_w w\mathbb{C}w^\top \text{ subject to the constraints } uw^\top = 1 \ \& \ \mu w^\top = \mu_V.$$

Similarly to the previous case, we will use the Lagrange multipliers method. To this end, define

$$\mathcal{L}(w, \lambda) := w\mathbb{C}w^\top - \lambda_1(uw^\top - 1) - \lambda_2(\mu w^\top - \mu_V).$$

Then,

$$\nabla_w \mathcal{L} = 0 \Leftrightarrow 2w\mathbb{C} - \lambda_1 u - \lambda_2 \mu = 0 \Leftrightarrow w = \frac{1}{2}(\lambda_1 u + \lambda_2 \mu)\mathbb{C}^{-1}, \tag{7}$$

which, in conjunction with the constraints ($uw^\top = 1$ and $\mu w^\top = \mu_V$), leads to the system with unknowns only involving λ:

$$\begin{cases} \lambda_1 u\mathbb{C}^{-1}u^\top + \lambda_2 u\mathbb{C}^{-1}\mu^\top = 2 \\ \lambda_1 \mu\mathbb{C}^{-1}u^\top + \lambda_2 \mu\mathbb{C}^{-1}\mu^\top = 2\mu_V \end{cases}, \tag{8}$$

whose solution can be immediately determined, *e.g.*, by Cramer's rule. Then, we derive immediately $w_{\mu_V}^{\min}$ by substituting in (7). □

Remark 5 (Finding a second efficient portfolio). *For notational convenience, assume the following notations:*

$$A = u\mathbb{C}^{-1}u^\top > 0,$$
$$B = \mu\mathbb{C}^{-1}u^\top = u\,\mathbb{C}^{-1}\mu^\top,$$
$$C = \mu\mathbb{C}^{-1}\mu^\top > 0,$$

with $AC - B^2 > 0$*. Then,* $w_{\mu_V}^{min}$ *in* (6) *can be rewritten as*

$$w_{\mu_V}^{min} = \frac{C - \mu_V B}{AC - B^2}\mathbb{C}^{-1}u^\top + \frac{\mu_V A - B}{AC - B^2}\mathbb{C}^{-1}\mu^\top.$$

Note that by choosing $\mu_V = \frac{B}{A}$*, the second term in* $w_{\mu_V}^{min}$ *disappears. This recovers the MVP without constraints. i.e.*

$$w_{\mu_V}^{min} = \frac{C - \mu_V B}{AC - B^2}\mathbb{C}^{-1}u^\top = \frac{C - \frac{B^2}{A}}{AC - B^2}\mathbb{C}^{-1}u^\top = \frac{1}{A}\mathbb{C}^{-1}u^\top = \frac{\mathbb{C}^{-1}u^\top}{u\mathbb{C}^{-1}u^\top} = w_{\text{MVP}}$$

In order to construct a second efficient portfolio from the Efficient frontier, one can use a similar trick as before. We will now simply choose $\mu_V = \frac{C}{B}$*, so that the first term in* $w_{\mu_V}^{min}$ *disappears. Hence, a second efficient portfolio (of course, this is not the unique choice, there are infinitely many such choices) can be constructed with the weights:*

$$w_{\mu_V = C/B}^{min} = \frac{\mu_V A - B}{AC - B^2}\mathbb{C}^{-1}\mu^\top = \frac{\frac{AC}{B} - B}{AC - B^2}\mathbb{C}^{-1}\mu^\top = \frac{1}{B}\mathbb{C}^{-1}\mu^\top = \frac{\mathbb{C}^{-1}\mu^\top}{\mu\mathbb{C}^{-1}u^\top}$$

Remark 6. *The weights* $w^{\min}_{\mu_V}$ *depend affinely on* μ_V, i.e., *there exist vectors* $\bar{A}, \bar{B}$ *such that the weights are given by* $w^{\min}_{\mu_V} = \bar{A}\mu_V + \bar{B}$. *Indeed, from* (6), *we have*

$$w^{\min}_{\mu_V} = \bar{A}\mu_V + \bar{B}. \tag{9}$$

The property essentially says that one can easily determine the weights of any portfolio on the MVL once we know the required expected return.

Indeed, once we compute the vectors $\bar{A}$ *and* $\bar{B}$ (e.g., *by the data of the market), it is then straightforward to evaluate an affine function. Playing around, we can make the story even simpler.*

Theorem 2.5 (Risk associated to the best portfolio given expected return μ_V)**.** *Assume that* $\det \mathbb{C} \neq 0$. *Fix a level of expected return* μ_V. *Then, the portfolio that has the smallest risk among all portfolios with expected return* μ_V *has risk/variance as follows:*

$$\sigma^2_{w^{min}_{\mu_V}} = \sigma^2_{\text{MVP}} + \frac{(\mu_V - \mu_{\text{MVP}})^2}{\psi\sigma^2_{\text{MVP}}}, \quad \textit{with,}$$
$$\psi = AC - B^2 = u^\top\mathbb{C}^{-1}u\mu^\top\mathbb{C}^{-1}\mu - (u^\top\mathbb{C}^{-1}\mu)^2.$$

Proof. We will skip the proof, as this is computationally heavy. □

By inspecting this relation, we can draw some conclusions (which we already somewhat know):

- The minimum value of $\sigma^2_{w^{min}_{\mu_V}}$ is σ^2_{MVP} (since ψ is positive, which follows from the fact that $\mathbb{C}$ is positive definite.)
- This minimum is achieved only when $\mu_{w^{min}_{\mu_V}} = \mu_{\text{MVP}}$.

In addition, having a simple relation between $\sigma_{w^{min}_{\mu_V}}$ and $\mu_{w^{min}_{\mu_V}}$ will allow us to derive lines which are tangent to the frontier boundary.
The portfolios corresponding to $w^{min}_{\mu_V}$ are also known as **target return portfolios**.

2.3 Two fund theorem

Assume that you have determined two (arbitrary, but distinct) portfolios on the MVL. Then you can simply "properly split" your wealth and invest on the aforementioned portfolios in order to obtain a portfolio with a specific level of risk-return (σ, μ) and which lies on MVL.

Key idea: Although investors may have different attitudes toward risk, all efficient portfolios can be generated by combining just two fixed portfolios "funds"). Thus, portfolio choice separates into:

1. Security selection (choosing the two funds), and
2. Risk preference (choosing how to combine them).

Theorem 2.6 (Two fund theorem)**.** *Choose arbitrarily two distinct portfolios on the minimum variance line (MVL), say* V_1^{MVL} *and* V_2^{MVL}, *with associated weights* w_1 *and* w_2 *and expected returns* $\mu_1 \neq \mu_2$. *Then, a portfolio* V *lies on the MVL if and only if its weight vector* w *can be written as an affine combination of* w_1 *and* w_2, i.e.

$$w = \alpha w_1 + (1 - \alpha)w_2, \quad \textit{for some } \alpha \in \mathbb{R}.$$

In other words, the above theorem can also be reformulated as follows.

Two efficient funds (portfolios) can be established so that any efficient portfolio can be duplicated, in terms of mean and variance, as a combination of these two. In other words, all investors seeking efficient portfolios need only invest in combinations of these two portfolios/funds.

Question: Based on what criterion should we make our choice among all the attainable portfolios?
Answer: Let V_i, $i = 1, 2$ be two risky assets, *e.g.*, stock price, portfolio *etc.*, with associated data (σ_i, μ_i), $i = 1, 2$. Then,

1. If $\mu_1 = \mu_2$, then we "naturally" prefer the asset with the lower risk.
2. If $\sigma_1 = \sigma_2$, then we "naturally" prefer the asset with the higher expected return.
3. If $\sigma_1 \le \sigma_2$ and $\mu_1 \ge \mu_2$, then we "naturally" prefer the asset V_1 –lower risk & higher expected return.
4. If $\sigma_1 \le \sigma_2$ and $\mu_1 \le \mu_2$, *i.e.*, high risk-high returns vs. low risk-low returns, then the choice depends on the investor's individual risk preference –no obvious best choice.

Question: If Security 1 dominates Security 2, is Security 2 redundant? Or, can Security 2 still be useful?
Answer: In general, we can use Security 2 in order to build a portfolio with risk even smaller than $\sigma_1 \le \sigma_2$ (diversification).

Question : If we fix the level of the expected return μ_V, resp. the risk σ_V, how can we uniquely choose a portfolio on the attainable set?
Answer: If we fix the level of the expected return, μ_V then the portfolios can only vary at the horizontal level, i.e., at the right half-line inside the attainable set. We see that the minimum risk portfolio we can get appears at the intersection of the Markowitz Bullet and the horizontal level line $\mu = \mu_V$.

If we fix the level of the risk σ_V, i.e., portfolios can only vary at a vertical segment inside the attainable set, we see that the highest expected return portfolio we can get appears at the intersection of the Markowitz Bullet and the upper point of the vertical segment $\sigma = \sigma_V$.

2.4 Efficient Frontier

We will start with a few relevant definitions.

Definition 7 (Dominating Asset)**.** *We say that an asset with expected return μ_1 and risk σ_1 dominates another asset with associated data (μ_2, σ_2) if*

$$\mu_1 \ge \mu_2 \text{ and } \sigma_1 \le \sigma_2.$$

Definition 8 (Efficient Portfolio, Efficient Frontier)**.** *A portfolio is called efficient if there is no other portfolio, except itself, that dominates it.*

In a few weeks, you will see that this definition is similar to Nash equilibrium.

The set of efficient portfolios among all attainable portfolios is called the efficient frontier.

Remark 7. *In other words, each efficient portfolio has*

1. the highest expected return among all attainable portfolios with the same risk;

2. the lowest risk among all attainable portfolios with the same expected return.

So, the efficient frontier is the "upper part" of the Markowitz bullet.

In terms of mathematics, efficient portfolios are those with associated data

$$(\sigma(w_{\mu_V}^{\min}), \mu_V) \text{ for } \mu_V \geq \mu_{\text{MVP}},$$

where μ_{MVP} is the expected return of the minimum variance portfolio.

Example 3. *What is the efficient frontier for Example 4 from the two security notes?*

In this case, the efficient set becomes a single point. This single point is represented by $(\frac{\sigma}{\sqrt{n}}, \mu)$.

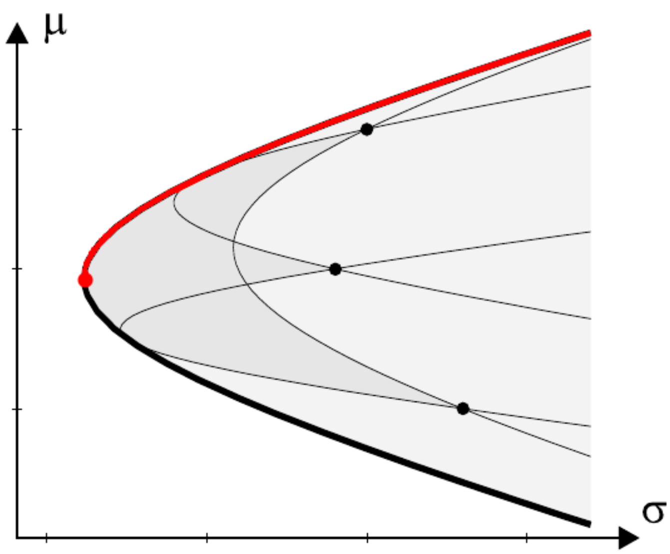


Figure 2: The Efficient Frontier in the example we considered before.

Notation: For convenience, we will denote the weights of the portfolios lying on the efficient frontier by w_{EF}, instead of $w_{\mu_V}^{\min}$ for $\mu_V \geq \mu_{\text{MVP}}$.

Remark 8. *The top half of the frontier boundary is called the efficient frontier. An investor utilizing mean-variance portfolio theory should never choose a portfolio that is not on the efficient frontier.*

The next Lemma provides a property of the efficient frontier, which will be proven useful in the Capital Asset Pricing Model (CAPM).

Lemma 2.7. *The weights w of any portfolio on the efficient frontier (except for the minimum variance portfolio) satisfy the condition*

$$\gamma w_{EF} \mathbb{C} = \mu - \mu_{EF} u$$

for some real number $\gamma > 0$.

In other words, the efficient frontier is the set of points whose weights w_{EF} are given by

$$w_{\text{EF}} = \bar{A}\mu_{EF} + \bar{B}, \quad \text{for } \mu \geq \mu_{\text{MVP}},$$

where $\bar{A}, \bar{B}$ as introduced in the previous Remark (see Identity (9)) and μ_{MVP} is the expected return of the Minimum Variance Portfolio (MVP).

3 Performance of a portfolio

Recall (from utility theory) that quadratic utility takes the following form:

$$U(w) = w - \frac{b}{2}w^2;\ b > 0.$$

In utility maximization theory, the investor would choose their investment W to maximize $\mathbb{E}[U(W)]$. We will now use a similar criteria for evaluating the performance of a portfolio selection. (Note that his is the simplest possible situation we can think of. There are much more complicated portfolio performance criteria, which are beyond the scope of the course):

$$\mathbb{E}[K_V] - \frac{\gamma}{2}Var[K_V], \quad \gamma > 0.$$

When we have portfolio weights denoted by w, a vector of mean returns μ, and a covariance matrix of returns $\mathbb{C}$ this performance criteria becomes

$$w^\top \mu - \frac{\gamma}{2} w^\top \mathbb{C} w.$$

Theorem 3.1. *The weights associated to the portfolio with the best portfolio performance criteria is given by(No need to remember the formulas):*

$$w_{\textit{best-performence}} = \frac{\mathbb{C}^{-1}u}{u^\top \mathbb{C}^{-1}u} + \frac{1}{\gamma}\left(\frac{u^\top \mathbb{C}^{-1}u\mathbb{C}^{-1}\mu - \mu^\top \mathbb{C}^{-1}u\mathbb{C}^{-1}u}{u^\top \mathbb{C}^{-1}u}\right).$$

Furthermore, the expected return and the variance of this optimal portfolio are given by

$$\begin{aligned}
\mu_{\textit{best-performence}} &= \frac{u^\top \mathbb{C}^{-1}\mu}{u^\top \mathbb{C}^{-1}u} + \frac{1}{\gamma}\left(\frac{\mu^\top \mathbb{C}^{-1}\mu u^\top \mathbb{C}^{-1}u - (u^\top \mathbb{C}^{-1}\mu)^2}{u^\top \mathbb{C}^{-1}u}\right)\\
&= \mu_{MVP} + \frac{1}{\gamma}\psi\sigma^2_{MVP},\\
\sigma^2_{\textit{best-performence}} &= \sigma^2_{MVP} + \frac{1}{\gamma^2}\left(\frac{\mu^\top \mathbb{C}^{-1}\mu u^\top \mathbb{C}^{-1}u - (u^\top \mathbb{C}^{-1}\mu)^2}{u^\top \mathbb{C}^{-1}u}\right)\\
&= \sigma^2_{MVP} + \frac{1}{\gamma^2}\psi\sigma^2_{MVP},
\end{aligned}$$

where $\psi := \mu^\top \mathbb{C}^{-1}\mu u^\top \mathbb{C}^{-1}u - (u^\top \mathbb{C}^{-1}\mu)^2$ is a scalar value. Note that this ψ is the same constant which appears in Theorem 2.5. This is a consequence of $\psi = \psi^T$ (since ψ is scalar) and, since $\mathbb{C}$ is symmetric, $\mathbb{C}^{-1} = (\mathbb{C}^{-1})^T$.

Proof. Proof using the Lagrange multiplier. □

4 Indifference Curves

We already calculated the collection of optimal portfolios using quadratic utility is equivalent to the portfolios which minimize variance with a target return.

Definition 9 (Indifference curve)**.** *An indifference curve is a collection of points of the form (μ_V, σ_V) in which the investor does not prefer one point over any other.*

That is, the investor is indifferent to which pair of mean and variance the portfolio achieves. For quadratic utility, an indifference curve is found by setting the performance criteria equal to a constant, i.e.,

$$\mu_V - \frac{\gamma}{2}\sigma_V^2 = c.$$

The set of points (μ_V, σ_V) is one of many indifference curves, and it will depend on the values of c and γ.

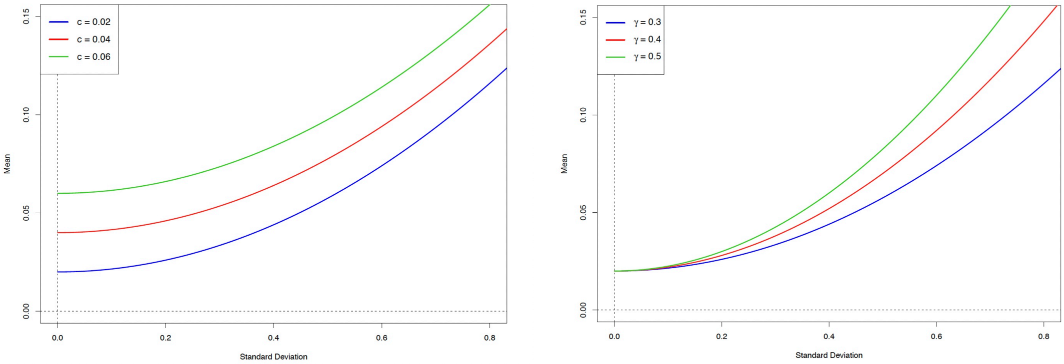


Figure 3: Indifference Curves

For a fixed level of risk-aversion, an investor prefers to have a portfolio which lies on a higher curve (in the sense that the corresponding performance criteria is higher). If we overlay several indifference curves (different values of c) with the mean-variance frontier, we want to find the portfolio which puts us on the highest curve.

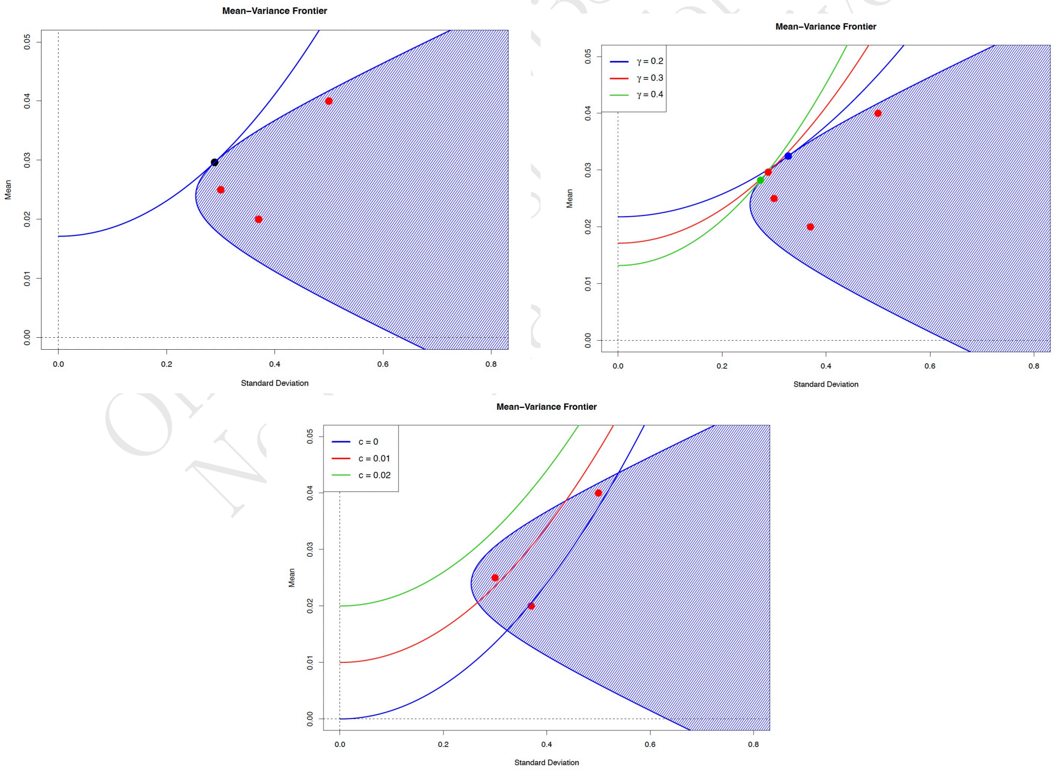


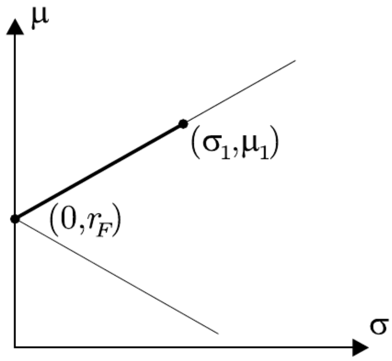
Figure 4: Indifference Curves

5 Market Portfolio – Adding a risk-free asset

Assumption: From now on we extend the market of the n risky assets by introducing a risk-free asset with return $\mu_{\text{RF}}/R/R_{rf}/R_f$. Recall that risk-free assets have a return that is deterministic. So expected return and rate of return are the exact same quantity in this case. Moreover, we will assume that $\mu_{\text{RF}} < \mu_{\text{MVP}}$. (Why? Since otherwise, just investing in the Risk-free asset would be logical...So this assumption is in line with our intuition)

Let's try to draw the feasible set of portfolios that we can construct based on a risk-free security and a single risky security.

Let us assume a (generic) portfolio V with $w = (w_f, w_1)$. Then for the portfolio V:



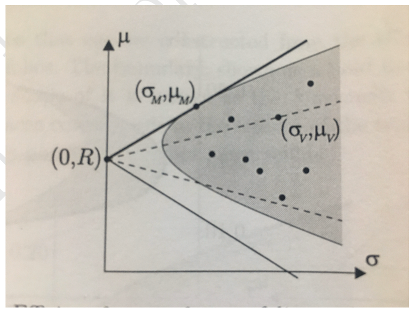
$$\mu_V = w_f r_f + w_1 \mu_1$$
$$\sigma_V^2 = w_1^2 \sigma_1^2 \iff \sigma_V = |w_1 \sigma_1|,$$

since the risk-free asset has 0 risk and no covariance with the risky asset (HW: mathematically show that $Cov(K_{rf}, K_P) = 0$ for any assete/ security/ portfolio P). As a consequence of the above two equation one can write μ_V as a linear function of σ_V.

Now let's think of the situation with N-risky security and 1 risk-free asset (Note that a market can never have 2 risk-free assets with different R at the same time, as this will create arbitrage).

Question:

1. What is the graph of the set of feasible portfolios that consists of one risk-free security and N-risky assets?
   Step 1: What is the feasible region for 1-RF + 1 portfolio V with (σ_V, μ_V) anywhere in the feasible set of just N-riskty asset (shadow region)?
   Step 2: What about another portfolio V_1? $\cdots$



2. In this new feasible set, which portfolio(s) should a rational investor choose? (Why?)

Definition 10 (CML, Market Portfolio)**.** *The half-line that starts at the risk-free asset and runs through the market portfolio M is called the Capital Market Line (CML). The portfolio M corresponding to the tangency point (σ_M, μ_M) is called the market portfolio.*

Theorem 5.1 (One-fund theorem)**.** *Assume the market has N risky assets and a risk-free asset. Then there exists a single fund M (the Market portfolio) such that any efficient portfolio can be constructed as a combination of M and the risk-free asset.*

Theorem 5.2. *Consider a market that consists of one risk-free security with return R, and n risky securities with expected return $\mu = (\mu_1, ..., \mu_n)$ and covariance matrix $\mathbb{C}$. Assume that $R < \mu_{\text{MVP}}$, and if $\mathbb{C}$ is invertible, then the market portfolio M exists and its weights are given by*

$$w_M = \frac{(\mu - Ru)\mathbb{C}^{-1}}{(\mu - Ru)\mathbb{C}^{-1}u^\top}.$$

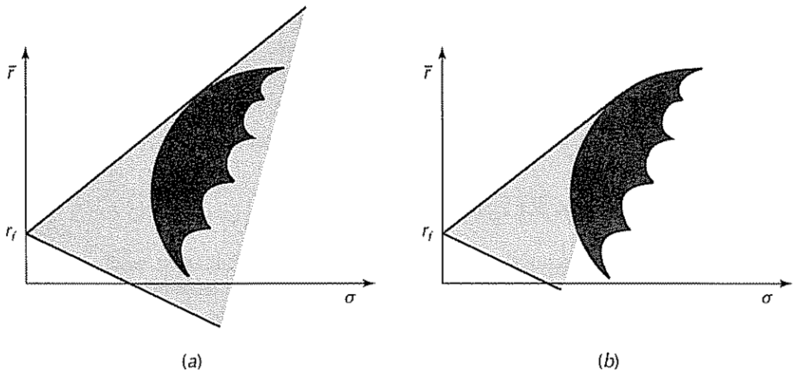


FIGURE 6.13 Effect of a risk-free asset. Inclusion of a risk-free asset adds lines to the feasible region (*a*) If both borrowing and lending are allowed, a complete infinite triangular region is obtained (*b*) If only lending is allowed, the region will have a triangular front end, but will curve for larger σ

Proof. We look for the V (or equivalently, w) that satisfies:

$$\max_{w} \frac{\mu w^\top - R}{\sqrt{w \mathbb{C} w^\top}} \quad \text{subject to the constraint} \quad u w^\top = 1.$$

To this end, we define the Lagrangian $\mathcal{L}(w, \lambda) := \frac{\mu w^\top - R}{\sqrt{w \mathbb{C} w^\top}} - \lambda(u w^\top - 1)$. By calculating the FOC, we derive

$$\begin{aligned} \nabla \mathcal{L} = 0 &\Leftrightarrow \frac{\mu [w \mathbb{C} w^\top]^{\frac{1}{2}} - (\mu w^\top - R)[w \mathbb{C} w^\top]^{-\frac{1}{2}} w \mathbb{C}}{w \mathbb{C} w^\top} - \lambda u = 0 \\ &\Leftrightarrow \mu - \lambda [w \mathbb{C} w^\top]^{\frac{1}{2}} u = \frac{\mu w^\top - R}{w \mathbb{C} w^\top} w \mathbb{C}. \end{aligned} \tag{10}$$

By multiplying with $w^\top$ on the right and using the constraint, we get

$$\lambda = \frac{R}{[w \mathbb{C} w^\top]^{\frac{1}{2}}}.$$

Plug in back to (10), we have

$$\mu - R u = \gamma w \mathbb{C}, , \tag{11}$$

where $\gamma := \dfrac{\mu w^\top - R}{w \mathbb{C} w^\top}$ (observe that γ depends on w).

Multiplying by $\mathbb{C}^{-1} u^\top$ on the right, we get

$$\gamma = (\mu - R u) \mathbb{C}^{-1} u^\top. \tag{12}$$

Note that the RHS is independent of w now.

To solve w from (11), we have to make sure that $\gamma \neq 0$. Recall that

$$\begin{aligned} \mu_{\text{MVP}} = w_{\text{MVP}} \mu^\top = \frac{u \mathbb{C}^{-1} \mu^\top}{u \mathbb{C}^{-1} u^\top} = \frac{\mu \mathbb{C}^{-1} u^\top}{u \mathbb{C}^{-1} u^\top} &\overset{\text{Assumption}}{>} R \\ \Leftrightarrow \mu \mathbb{C}^{-1} u^\top - R u \mathbb{C}^{-1} u^\top > 0 \Leftrightarrow \gamma > 0. \end{aligned}$$

Combine (11) and (12) to get the required form. □

Existence of Market Portfolio: $R < \mu_{\text{MVP}}$ and $\mathbb{C}$ is invertible.

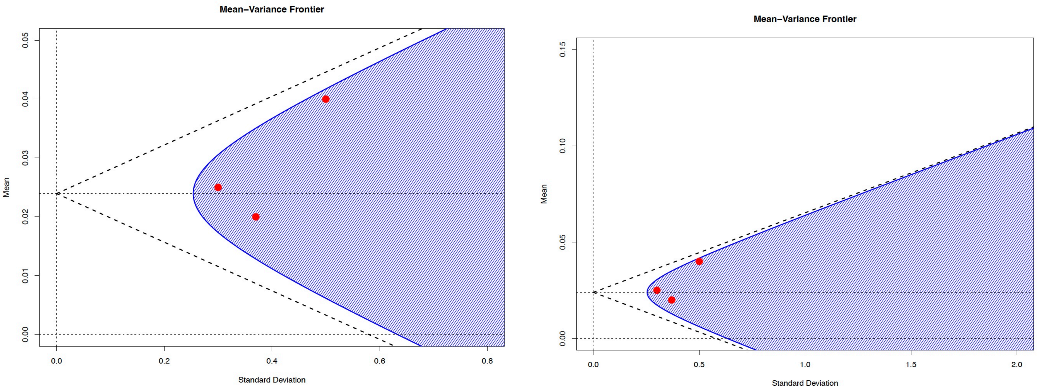


Figure 5: Left/ Right: Non-existence of tangency portfolio.

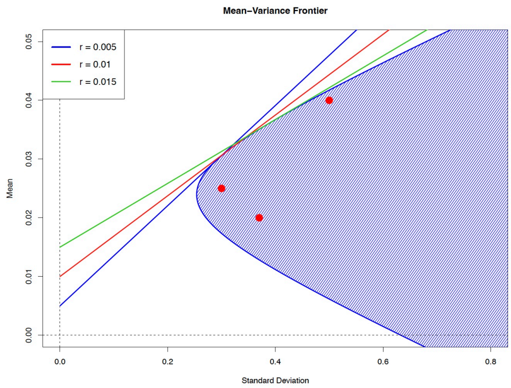


Figure 6: Market portfolio for various risk-free returns R.

Why the name Market portfolio?

Market portfolio M contains all risky securities in consideration (i.e. in our market). And the weights of each risky securities is exactly equal to their relative share in the whole market. Because of this property, it is called the market portfolio.
In practice, the market portfolio is approximated by a suitable stock exchange index.

6 Capital Market Line

Assume there are no transaction costs.
Everyone has the same risk-free asset and the same n risky asset available to them to invest.
Further, assume everyone uses the same means, variances, and covariances for all the n available risky assets.
With these assumptions, what can we conclude? From the **One-fund theorem**, we know that everyone will purchase a single fund/portfolio of risky assets, and they may, in addition, borrow or lend at a risk-free rate. Furthermore, since everyone uses the same means, variances, and covariances, everyone will use the same risky fund. The mix of these assets, the risky fund and the risk-free asset, will likely vary across individuals according to their individual tastes for risk. Some will seek to avoid risk and will, accordingly, have a high percentage of the risk-free assets in their portfolios; others, who are more aggressive, will have a high percentage of the risky fund. **However, every individual will form a portfolio that is a mix of risk-free asset and a single, risky**

fund. Hence, the One-fund theorem 5.1 is really the only fund that is used.

Theorem 6.1. *The Capital Market Line (CML) satisfies the equation*

$$\mu = R + \frac{\mu_M - R}{\sigma_M}\sigma.$$

Sometimes we will write $R = R_f$. The quantity $\frac{\mu_M - R}{\sigma_M}\sigma$ is called the risk premium. The slope of the capital market line is $\frac{\mu_M - R}{\sigma_M}$, and this value is frequently called the ***price of risk****. It tells by how much the expected rate of return of a portfolio must increase if the risk (standard deviation) of that rate increases by one unit.*

Remark 9.

1. *The risk premium can be seen as the additional return above the risk-free return R, which compensates for exposure to risk. Of course, for risk free asset, the risk-premium is 0.*
2. *The risk-premium of the Market portfolio is $\mu_M - R$, independent of the risk of M.*
3. *If every investor uses the same μ and $\mathbb{C}$ and shares the same principles as the one we have described,* e.g., *the efficient portfolios are preferable, then every investor will construct their portfolio on the CML.*
4. *This line shows the relation between the expected rate of return and the risk of return (as measured by the standard deviation) for efficient assets or portfolios of assets.* ***It is also referred to as a pricing line, since prices should adjust so that efficient assets fall on this line.***
5. *CML states that as risk increases, the corresponding expected rate of return must also increase. Furthermore, this relationship can be described by a straight line if risk is measured by standard deviation.*

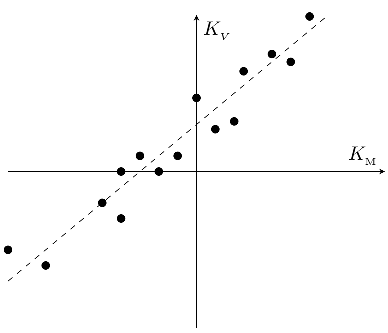
(Look at the Example 7.1 and 7.2 from the Investment Science book)

7 The pricing model: Capital Asset Pricing Model (CAPM) & Beta Factor

The capital market line relates the expected rate of return of an efficient portfolio to its standard deviation, but it does not show how the expected rate of return of an individual asset (not necessarily efficient) relates to its individual risk. This relation is expressed by the capital asset pricing model.

Goal: Given any portfolio/asset V in the feasible set, we want to understand how the return K_V of this portfolio will react to trends affecting the whole market, i.e., the return K_M of the market portfolio.

Recall that the CML is tangent to the efficient frontier at the point (σ_M, μ_M) representing the market portfolio M. The slope of CML is maximal. *In other words, "the market will be in equilibrium" and the relative proportions of risky assets held in each portfolio will be the same for all investors.*



Question: All portfolios constructed from M and V form a hyperbola. Is CML also the tangent line of it? If so, where is the tangent point?

Answer: Yes! The tangent point is still (σ_M, μ_M).

Question: Concider a market with N-risky asset and 1-risk free with $R < \mu_{\text{MVP(N-risky asset)}}$, where the $\mu_{\text{MVP(N-risky asset)}}$ represents the MVP constructed just from the N-risky asstes (not including the Risk free).

- What is the efficient frontier and feasiable set if the market was consists of only N-risky asset?
- What is the efficient frontier and feasiable set if the market was consists of N-risky asset along with the 1-Risk free asset?
- What is the efficient frontier and feasiable set if the market was consists of N-risky asset along with the 1-Risk free asset, but the investor only wants to invest in the risky asset?

Answer: HW

Definition 11 (Beta factor)**.** *The beta factor of a portfolio/security/asset V is defined by*

$$\beta_V = \frac{Cov(K_V, K_M)}{\sigma_M^2}.$$

Consider a portfolio P constructed by M and V, denote the weights of V and M in the portfolio P by $(x, 1-x)$, then the risk and return of portfolio P are

$$\sigma_P = \left(x^2\sigma_V^2 + (1-x)^2\sigma_M^2 + 2x(1-x)Cov(K_V, K_M)\right)^{\frac{1}{2}},$$

$$\mu_P = x\mu_V + (1-x)\mu_M.$$

We compute the derivatives with respect to x at $x = 0$:

$$\left.\frac{\partial \sigma_P}{\partial x}\right|_{x=0} = \frac{Cov(K_V, K_M) - \sigma_M^2}{\sigma_M},$$

$$\left.\frac{\partial \mu_P}{\partial x}\right|_{x=0} = \mu_V - \mu_M.$$

7.1 Beta Factor & CAPM

The slope of the tangent line at market portfolio M is

$$\left.\frac{\partial \mu_P}{\partial \sigma_P}\right|_{x=0} = \left.\frac{\frac{\partial \mu_P}{\partial x}}{\frac{\partial \sigma_P}{\partial x}}\right|_{x=0} = \frac{\mu_V - \mu_M}{\frac{Cov(K_V, K_M) - \sigma_M^2}{\sigma_M}} = \text{the slope of CML} = \frac{\mu_M - R}{\sigma_M}.$$

Solving for μ_V, we can get

$$\mu_V = R + \frac{Cov(K_V, K_M)}{\sigma_M^2}(\mu_M - R).$$

Theorem 7.1. *Suppose $R < \mu_{\text{MVP}}$. Then under the assumption that the market portfolio is efficient, the expected return μ_V for a feasible portfolio V is given by the CAPM*

$$\mu_V = R + \beta_V(\mu_{\text{M}} - R).$$

Proof. A formal proof is provided in (Luenberger, 1998, page 177). Please read the proof. □

We can think of this as '**reward is proportional to the risk**'. CAPM can give an educated guess for the value of assets that has not been priced in the market. Although many assumptions are overly simplistic, the model still gives insights and is still used in some applications.

Corollary 7.2. *The Beta of Market portfolio is given by* $\beta_M = 1$. *The Beta of a risk-free asset is given by* $\beta_{rf} = 0$.

Summarizing assumptions of CAPM:

- All investors select portfolios by mean-variance optimization.
- All investors have the same investment horizon.
- All investors use the same mean return vector and covariance matrix.
- Individual investors may have different levels of risk aversion.
- Assets are infinitely divisible.
- No taxes and transaction costs.
- The market is in equilibrium.
- No constraints on investing.

Proposition 7.3. *Suppose a portfolio* P *is constructed with* n *assets with corresponding weights* $w_1, w_2, \cdots, w_n$. *Then:*

$$\beta_P = \sum_{i=1}^{n} w_i \beta_i.$$

8 Security Market Line

It is clear from the CAPM that on the $(\sigma, \mu)-$plane, all efficient portfolios lie on the straight line described by CAPM (in fact, the CML). But in a $(\sigma, \mu)-$plane, the set of all feasible portfolios forms a 2-dimensional region, not a straight line.

Question: In what plane does the set of all feasible portfolios become a straight line? Recall, CAPM:

$$\mu_{_V} = R + \beta_V(\mu_{_M} - R).$$

Answer: in the (β, μ)-plane.
The graph of this function in the (β, μ)-plane is called the Security Market Line (SML). Note that $\beta_M = 1$ (Why?–Hint: use the definition).
Two remarks regarding the SML:

1. It is easy to determine the the SML since we know two points $(0, R)$ and $(1, \mu_{_M})$.
2. Despite some securities/portfolios have small returns and large risks, they remain attractive to the investors. The reason is that these securities have negative beta, or equivalently a negative correlation. This means that the prices of these securities tend to move on the opposite direction to the market.

   A standard example of an asset with a negative beta is gold, which acts as insurance in a financial crisis.

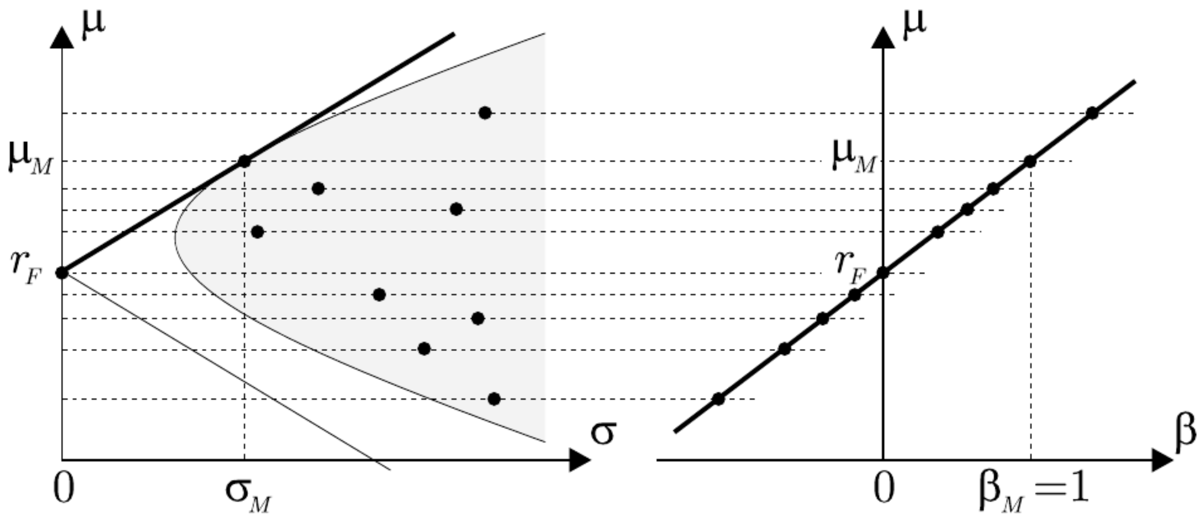


Figure 7: The Capital Market Line and the Security Market Line

9 Limitations of Mean–Variance Theory

Mean–variance portfolio theory provides a powerful and elegant framework for understanding the trade-off between risk and expected return. However, its practical applicability relies on several strong assumptions, which may not hold in real financial markets. We briefly summarize the main limitations below.

- **Normality of returns.** Mean–variance analysis is fully descriptive only when asset returns are normally distributed (or, more generally, elliptically distributed). In such cases, the mean and variance completely characterize the return distribution. In practice, asset returns often exhibit skewness, excess kurtosis, and fat tails, which are not captured by variance alone.

- **Variance as the sole measure of risk.** Variance penalizes both upside and downside deviations from the mean equally. However, investors are typically more concerned with downside risk. This motivates alternative risk measures such as Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), or downside semivariance.

- **Estimation error.** The inputs of mean–variance optimization, namely expected returns and the covariance matrix, must be estimated from historical data. Expected returns are particularly difficult to estimate accurately, and small estimation errors can lead to large changes in the optimal portfolio weights.

- **Instability of optimal portfolios.** Mean–variance optimization often produces extreme portfolio weights, especially when short selling is allowed. As a result, the optimal portfolio may be highly sensitive to small changes in the input parameters, leading to poor out-of-sample performance.

- **Single-period framework.** The theory is based on a single-period investment horizon and does not explicitly account for dynamic rebalancing, transaction costs, or intermediate consumption, which are relevant in multi-period investment settings.

- **Homogeneous investor assumptions.** The model assumes that all investors have identical expectations regarding returns, variances, and correlations. In reality, investors differ in information, beliefs, preferences, and constraints.

- **Market frictions ignored.** Mean–variance theory abstracts from transaction costs, taxes, liquidity constraints, and borrowing limits. These frictions can significantly affect optimal portfolio choices in practice.

Remark 10. *Despite these limitations, mean–variance theory remains a foundational benchmark in finance. Many modern portfolio optimization techniques can be viewed as extensions or refinements of the mean–variance framework, designed to address one or more of the above shortcomings.*

References

Luenberger, D. G. (1998). *Investment science.* Oxford university press.

Petters, A. O. and Dong, X. (2016). *An introduction to mathematical finance with applications.* Springer.

Mathematical Finance II (6CCM338A)-Game Theory

Purba Das
Department of Mathematics, King's College London

Contents

[1] [2]

1 Strategic form game

A game in strategic form (or normal form) is a tuple $G = \left(N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\right)$ where

- N is the set of players.
- S_i is the set of all (pure) actions or strategies available to player i. Denote $S = \prod_i S_i$ as the set of strategy profiles.
- The function $u_i : S \to \mathbb{R}$ is player i's utility (or payoff) for each strategy profile.

We will assume that players have obvious preferences over utility: more is better (as we also learnt in Week-1).

Definition 1 (Finite game)**.** *We say that G is finite if the cardinality of both sets N and S is finite.*

[1] You should follow the class notes as main material; this is additional material and may not cover every detail discussed in class. This is also a draft version of topics covered in class; this is not a textbook on Game theory. Some of the contents are taken from various online materials/ lecture notes.

[2] This is not a math book but gives good intuition, not a replacement of class notes:
Game theory : a very short introduction Binmore, K. G., 1940- 2007

2 (Pure) Nash equilibria

Notation: For a strategy profile $s = \Pi_i s_i \in S$, the notation $s_{-i} \in S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_N$ represents the strategy profile where, except for the i^{th} player, all other players have the same strategy as s. It is the tuple of strategies of everyone other than Player-i.

Definition 2 (Profitable deviation)**.** *Given a strategy profile $s = \Pi_i s_i \in S$, a profitable deviation for player i is a strategy t_i such that*

$$u_i\left(s_{-i}, t_i\right) > u_i\left(s_{-i}, s_i\right).$$

A strategy profile s is a Nash equilibrium if no player has a profitable deviation. These are also called pure Nash equilibria, for reasons that we will see later. They are often just called (pure) equilibria/equilibrium.

Another way to see the same definition is as follows:

Definition 3 (Nash equilibrium (NE))**.** *The strategy profile s^* is a Nash equilibrium if*

$$\underbrace{u_i(s_i^*, s_{-i}^*)}_{u_i(s^*)} \geq u_i(s_i, s_{-i}^*) \textit{ for all } s_i \in S_i.$$

A game can have more than one (pure) Nash equilibrium or no (pure) Nash equilibrium. Since the equality is $\geq$, a player might be indifferent among several strategies given the other players' choices.

There is a stronger notion of NE that people use sometimes, which is called the 'strict Nash equilibrium'. This is the case when the above inequality is strict, i.e.,

$$u_i(s_i^*, s_{-i}^*) > u_i(s_i, s_{-i}^*) \text{ for all } s_i \in S_i \backslash \{s_i^*\}.$$

For this course, we will not focus much on the strict NE.

2.1 Classical examples

Typically a two player game is represented by a matrix (as we discussed in class). For all the examples we follow the matrix notation.

Matching pennies

In this game (and throughout the course), there will only be two players: a row player (R) and a column player (C). We will represent the game as a payoff matrix, showing for each strategy profile $s = (s_R, s_C)$ the payoffs $u_R(s), u_C(s)$ of the row player and the column player, respectively. In this case, the strategy profile of row player is $s_R \in \{H, T\}$ and similarly for the column player $s_C \in \{H, T\}$.

	H	T
H	1,0	0,1
T	0,1	1,0

In this game, each player has to choose either heads (H) or tails (T). The row player wants the choices to match, while the column player wants them to mismatch.

Exercise: Show that matching pennies has no pure Nash equilibria.

Prisoners' dilemma

A crime was committed in a prison cell, and there are only two people who could have done it. The two prisoners are faced with a dilemma. Each prisoner has to choose between testifying to police against the other (and thus betraying the other) or keeping his/her mouth shut. In the former case, we say that the prisoner cooperated (with the police and betrayed the other prisoner), and in the latter, (s)he not cooperated (with the police, not with the other prisoner).

If both not cooperate with police (i.e., keep their mouths shut), they will have to serve the remainder of their sentences, which are 2 years each. If both cooperate (i.e., agree to testify against each other), each will serve 3 years (additional penalty for lying to police and putting fingure to each other). If one cooperated and the other does not cooperates, the person cooperated will be released immediately, and the person who has not cooperator will serve 10 years for the crime.

Assuming that a player's goal is to minimise the number of years served, the payoff matrix is the following.

	cooporated	not cooporated
cooporated	-3,-3	0,-10
not cooporated	-10,0	-2,-2

Exercise. Show that the unique pure Nash equilibrium is (D, D).

Split or steal

Split or steal is a game played in a British game show called Golden Balls.

	St	Sp
St	0,0	1,0
Sp	0,1	$\frac{1}{2}, \frac{1}{2}$

Exercise. What are the pure equilibria of this game? (0,0) (1,0) (0,1)

Stag hunt

Two friends go on a hunt. Each has to decide whether to try and catch a hare-which she can do by herself-or to cooperate on trying to catch a stag. The payoff matrix is the following.

	S	H
S	2,2	0,1
H	1,0	1,1

Exercise. What are the pure Nash equilibria? (2,2) & (1,1)

Battle of the sexes

Adam and Steve are a married couple. They have to decide whether to spend the night at the monster truck show (M) or at the opera (O). Adam (the row player) prefers the truck show, while Steve prefers the opera. Both would rather go out than do nothing, which is what would happen if they could not agree. Their payoff matrix is the following.

	M	O
M	2,1	0,0
O	0,0	1,2

Exercise. Find all the (pure) equilibria of this game. (M,M) and (O,O).

2.2 Dominated strategies

A strategy s_i of player i in $G = \left(N, \{S_i\}_{i\in N}, \{u_i\}_{i\in\mathbb{N}}\right)$ is **strictly dominated** (this is also known as 'dominated') if there exists another strategy t_i such that, for all choices of strategy $s_{-i} \in S_{-i}$ of the other players, it holds that

$$u_i(s_{-i}, t_i) > u_i(s_{-i}, s_i).$$

That is, regardless of what the other players do, t_i is a better choice for i than s_i.

We say that s_i is **weakly dominated** if there exists another strategy t_i such that for all s_{-i}

$$u_i(s_{-i}, t_i) \geq u_i(s_{-i}, s_i)$$

and furthermore for some s_{-i}

$$u_i(s_{-i}, t_i) > u_i(s_{-i}, s_i).$$

Exercise. Does matching pennies have strictly dominated strategies? No Weakly dominated strategies? No How about the prisoners' dilemma? Yes, C for row player and C for column player.

3 Repeated elimination of dominated strategies

It seems unreasonable that a rational person would choose a strictly dominated strategy, because she has an obviously better choice.

Theorem 3.1 (Dominance principal)**.** *Let $G = \left(N, \{S_i\}_{i\in N}, \{u_i\}_{i\in\mathbb{N}}\right)$ be a finite game, let $d_j \in S_j$ be a dominated strategy of player j, and let $G' = \left(N, \{S'_i\}_{i\in N}, \{u'_i\}_{i\in\mathbb{N}}\right)$, where*

$$S'_i = \begin{cases} S_i & \text{for } i \neq j \\ S_j \backslash \{d_j\} & \text{for } i = j \end{cases}$$

and u'_i is the restriction of u_i to S'.

Then every Nash equilibrium $s \in S$ of G is in S'. Furthermore, $s \in S'$ is a Nash equilibrium of G if and only if it is a Nash equilibrium of G'.

The proof of this Theorem is straightforward (we won't cover it in class).

The following theorem of Gilboa, Kalai, and Zemel(1990) shows that the order of elimination of dominated strategies does not matter, as long as one continues eliminating until there are no more dominated strategies.

Note that this is not true for weakly dominated strategies.

Theorem 3.2 (Gilboa, Kalai and Zemel, 1990)**.** *Fix a finite game G, and let G_1 be a game that*

- *is the result of repeated elimination of dominated strategies from G, and*
- *has no dominated strategies.*

Let G_2 be a game with the same properties. Then $G_1 = G_2$. Furthermore, $G_1 = G_2 \equiv G$. The equality is in terms of NE and value of the game.

Note that a game may not always be reduced to 1×1 game. For example, the following game can only be reduced to 2×2, not further by row/column elimination.

	B_1	B_2	B_3
A_1	$3, -3$	$2, -2$	$5, -5$
A_2	$4, -4$	$3, -3$	$6, -6$
A_3	$1, -1$	$4, -4$	$3, -3$

3.1 Dominant strategies

A strategy s_i of player i in $G = \left(N, \{S_i\}_{i\in\mathbb{N}}, \{u_i\}_{i\in\mathbb{N}}\right)$ is strictly dominant if for every other strategy t_i it holds that

$$u_i\left(s_{-i}, s_i\right) > u_i\left(s_{-i}, t_i\right).$$

That is, regardless of what the other players do, s_i is a better choice for i than t_i.

Exercise. Which of the games above has a strictly dominant strategy? D is a dominant strategy for Player 1 in the Prisoners' dilemma.

Theorem 3.3. *If player i has a strictly dominant strategy d_i then, for every pure Nash equilibrium s^* it holds that $s^*_i = d_i$.*

4 Mixed equilibria and Nash's Theorem

Consider this game:

	a	b
A	2,-2	-3,3
B	0,0	3,-3

.

In this case neither player would want to play a single strategy with certainty, as in this case the other player could take advantage of such a choice. The only sensible plan is to use some random device to decide which strategy to play. For example, Player-1 might flip a coin to decide between A and B. Such a plan, which involves playing a mixture of strategies according to certain fixed probabilities, is called a mixed strategy.

Given a finite set X, denote by ΔX the set of probability distributions over X.

Definition 4. *Let $G = \left(N, \{S_i\}_{i\in\mathbb{N}}, \{u_i\}_{i\in\mathbb{N}}\right)$ be a finite game. A mixed strategy σ_i (for player i) is an element of ΔS_i. Given a mixed strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$, we overload notation and let σ be the element of ΔS given by the product $\prod_i \sigma_i$. That is, σ is the distribution over $\prod_i S_i$ in which we pick independently from each S_i, with distribution σ_i.*

For clarity, we sometimes refer to a member of S_i as a pure strategy. As an example, a mixed strategy for Matching pennies is $\sigma = (s_R, s_C) = \{(\frac{1}{3}, \frac{2}{3}), (0, 1)\}$. An example of a pure strategy in matching pennies is $H \times T$, i.e. $\{(1,0), (0,1)\}$.

Mixed extension of a game:

The mixed extension of the strategic game $G = \left(N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N}\right)$ is another strategic game $\hat{G} = \left(N, \left\{\hat{S}_i\right\}_{i \in \mathbb{N}}, \{\hat{u}_i\}_{i \in \mathbb{N}}\right)$ in which,

- For all $i \in \{1, 2, \cdots, N\}$, $\hat{S}_i = \Delta S_i$, the set of probability distributions over S_i.
- $\hat{u}_i : \Delta S \to \mathbb{R}$ where, each mixed strategy σ is assigned to the expected value under u_i of the lottery over $S = S_1 \times \cdots \times S_N$ that is induced by σ, mathematically $\hat{u}_i(\sigma) = \mathbb{E}_{s \sim \sigma}[u_i(s)]$. For a finite game, this is simply

$$\hat{u}_i(\sigma) = \underbrace{\sum_{s_1 \in S_1} \sum_{s_2 \in S_2} \cdots \sum_{s_N \in S_N} [\sigma_1(s_1) \cdot \sigma_2(s_2) \cdots \sigma_N(s_N) \cdot u_i(s_1, s_2, \ldots, s_N)]}_{= \sum_{s \in S} \left[\prod_{j=1}^{N} \sigma_j(s_j)\right] u_i(s)}$$

Intuition: Take the random pure profile $s \in S$ drawn according to the joint distribution induced by σ, compute $u_i(s)$, and average (expectation) over all possible s.

That is, $\hat{G}$ is a game whose strategies are the mixed strategies of G, and whose utilities are the expected utilities of G, taken with respect to the given mixed strategies. A pure Nash equilibrium of the mixed extension $\hat{G}$ is called a mixed Nash equilibrium of the original strategic game G. That is, a mixed strategy profile $\sigma \in \prod_i \Delta S_i$ is a mixed Nash equilibrium if no player can improve her expected utility by deviating to another mixed strategy.

We will often just say "Nash equilibrium" when referring to mixed equilibria. We will use u_i to also mean $\hat{u}_i$; that is, we will extend u_i from a function $S \to \mathbb{R}$ to a function $\Delta S \to \mathbb{R}$.

For example: for Matching pennies with mixed strategy $\sigma = (s_R, s_C) = \{(\frac{1}{3}, \frac{2}{3}), (0, 1)\}$ we have:

$$u_1(\sigma) = 1 * \frac{1}{3} * 0 + 0 * \frac{1}{3} * 1 + 0 * \frac{2}{3} * 0 + 1 * \frac{2}{3} * 1.$$

Similarly,

$$u_2(\sigma) = 0 * \frac{1}{3} * 0 + 1 * \frac{1}{3} * 1 + 1 * \frac{2}{3} * 0 + 0 * \frac{2}{3} * 1.$$

Remark 1. *Fix any σ_{-i}, and note that as a function $u_i(\sigma_{-i}, \cdot) : \Delta S_i \to \mathbb{R}$ is linear. That is,*

$$u_i(\sigma_{-i}, \alpha\sigma_i + (1-\alpha)\tau_i) = \alpha u_i(\sigma_{-i}, \sigma_i) + (1-\alpha) u_i(\sigma_{-i}, \tau_i)$$

Nash's celebrated theorem states that every finite game has a mixed equilibrium.

Theorem 4.1 (Nash, 1950)**.** *Every finite strategic game has a mixed Nash equilibrium.*

To prove Nash's Theorem, one needs Brouwer's Fixed Point Theorem (sometimes other fixed point theorems are also used). Essential to this proof is the assumption that the set of actions of each player is finite. Glicksberg (1952)[3] shows that a game in which each strategy set is a convex compact subset of a Euclidean space and each payoff function is continuous has a mixed strategy Nash equilibrium. (If each player's payoff function is also quasi-concave in his/her own action, then one can also show that such a game has a pure strategy Nash equilibrium.) For the detailed proof (not examinable), please refer to [1, Proposition 33.1].

[3] "A Further Generalisation of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium Points", Proceedings of the American Mathematical Society 3, 170–174.

Theorem 4.2 (Brouwer's Fixed Point Theorem)**.** *Let X be a compact convex subset of $\mathbb{R}^d$. Let $T : X \to X$ be continuous. Then T has a fixed point. i.e., there exists an $x \in X$ such that $T(x) = x$.*

A simple application: If you are in a room and hold a map of the room horizontally, then there is a point in the map that is exactly above the point it represents.

4.1 Best responses

Let $G = \left(N, \{S_i\}_{i\in\mathbb{N}}, \{u_i\}_{i\in\mathbb{N}}\right)$ be a finite game, and let σ be a mixed strategy profile in G. We say that $s_i \in S_i$ is a best response to σ_{-i} if, for all $t_i \in S_i$,

$$u_i\left(\sigma_{-i}, s_i\right) \geq u_i\left(\sigma_{-i}, t_i\right)$$

This notion can be naturally extended to mixed strategies i.e. $s_i, t_i \in \Delta S$.

The following proposition is helpful for understanding/ calculating mixed Nash equilibria.

Proposition 4.3. *Let $G = \left(N, \{S_i\}_{i\in\mathbb{N}}, \{u_i\}_{i\in\mathbb{N}}\right)$ be a finite strategic game. Then, $\sigma^* = (\sigma_1^*, \cdots, \sigma_N^*)$ is a mixed Nash equilibrium of G if and only if for every player i, every pure strategy s_i in the support of σ_i^* (i.e., any s_i to which σ_i^* assigns positive probability) is a best response to σ_{-i}^*.*

Proof. Suppose $s_i \in S_i$ is in the support of σ_i^*, but is not a best response to σ_{-i}^*, and let $t_i \in S_i$ be some best response to σ_{-i}^*. We will prove the claim by showing that t_i is a profitable deviation for i.

Let $C = u_i\left(\sigma_{-i}^*, t_i\right)$. Then $u_i\left(\sigma_{-i}^*, r_i\right) \leq C$ for any $r_i \in S_i$, and $u_i\left(\sigma_{-i}^*, s_i\right) < C$. It follows that

$$\begin{aligned} u_i\left(\sigma^*\right) &= \sum_{r_i \in S_i} \sigma_i^*\left(r_i\right) u_i\left(\sigma_{-i}^*, r_i\right) \\ &< C \end{aligned}$$

and so t_i is indeed a profitable deviation, since it yields utility C for i.

It follows that if σ^* is an equilibrium then $u_i\left(\sigma_{-i}^*, s_i\right)$ is the same for every s_i in the support of σ_i^*. That is, i is indifferent between all the pure strategies in support of her mixed strategy. □

Definition 5. *Let $G = \left(N, \{S_i\}_{i\in\mathbb{N}}, \{u_i\}_{i\in\mathbb{N}}\right)$ be a finite game, and let σ be a mixed strategy profile in G. We say that σ_i is* ***completely mixed*** *if it assigns (strictly) positive probability to each $s_i \in S_i$.*

Mixed strategies that are not pure are often called **'randomised strategies'**. A randomised strategy does not have to be a completely mixed strategy. For example,

$$\sigma = \{(1/2, 0, 1/2); (1/3, 1/3, 1/3)\}$$

is not a completely mixed strategy but is a randomised strategy (i.e., not a pure strategy).

Further reading: A very nice interpretation of mixed strategy Nash Equilibrium is discussed in [1, Section 3.2].

Useful tricks for computing Nash Equilibrium for the most general setup.

Assume a game G (this technique can be applied to any finite game, but after Step 1 below, if the game is not 2×2 game, it's not very practical to calculate without a computer). We will demonstrate this using an example. Consider the following two-player game:

G =

	A	B
a	1,1	1,0
b	0,1	4,4

To solve the set of all NE of the above game G, one can follow the following steps.
Step 1: Check the row/ column domination and see if the game can be reduced further. In this particular case, no row/column domination exists, and hence the game can not be reduced further.
Step 2 (Not necessary but can be helpful for step 3): Check for pure NE in the game by looking at the best response for all the individual coordinates. For this particular game:

- the strategy (a, A) is a best response. $\rightarrow$ This is a (pure) NE.
- the strategy (b, A) is NOT a best response. $\rightarrow$ This is NOT a (pure) NE.
- the strategy (a, B) is NOT a best response. $\rightarrow$ This is NOT a (pure) NE.
- the strategy (b, B) is a best response. $\rightarrow$ This is a (pure) NE.

Step 3: Find all NE using best response (this should include the NE we have found in Step 2 as well). Using Theorem 4.1, one can ensure the above game has a NE. Assume $\sigma^* = (x^*, y^*)$ is **a** NE in the above game. So in terms of notation, we can write

$$\sigma^* = (x^*, y^*) \text{ with,}$$
$$x^* = (p^*, 1-p^*) \qquad \text{and,} \qquad y^* = (q^*, 1-q^*) \text{ for some } p^*, q^* \in [0,1].$$

The definition of NE suggests that it must be a best response for all players. Using the fact that σ^* is a best response of Player-1, we can write the following:

$$\begin{aligned}
& u_1(\sigma^*) \geq u_1(\sigma^*_{-1}, x) \qquad \forall\, x \in \Delta S_1 \\
\iff & \begin{pmatrix} p^* & 1-p^* \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} q^* \\ 1-q^* \end{pmatrix} \geq \begin{pmatrix} p & 1-p \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} q^* \\ 1-q^* \end{pmatrix} \; \forall p \in [0,1] \\
\iff & \underbrace{p^*q^* + p^*(1-q^*)}_{p^*} + 4(1-p^*)(1-q^*) \; \forall p \in [0,1] \geq \underbrace{pq^* + p(1-q^*)}_{p} + 4(1-p)(1-q^*) \\
\iff & (p^* - p)(1 - 4(1-q^*)) \geq 0 \qquad \forall p \in [0,1] \\
\iff & \begin{cases} \text{Either } p^* = 1 \text{ and } 1 - 4(1-q^*)) \geq 0 \text{ i.e. } q^* \geq \frac{3}{4} \\ \text{Or } p^* = 0 \text{ and } 1 - 4(1-q^*)) \leq 0 \text{ i.e. } q^* \leq \frac{3}{4} \\ \text{Or } 1 - 4(1-q^*)) = 0 \text{ i.e. } q^* = \frac{3}{4} \text{ and } p^* \in [0,1] \end{cases}
\end{aligned}$$

In a similar line of arguments, one can also use Player 2's best response. i.e.

$$u_2(\sigma^*) \geq u_2(\sigma^*_{-2}, y) \qquad \forall\, y \in \Delta S_2$$
$$\vdots$$

$$\iff \begin{cases} \text{Either } q^* = 1 \text{ and } 1 - 4(1-p^*)) \geq 0 \text{ i.e. } p^* \geq \frac{3}{4} \\ \text{Or } q^* = 0 \text{ and } 1 - 4(1-p^*)) \leq 0 \text{ i.e. } p^* \leq \frac{3}{4} \\ \text{Or } 1 - 4(1-p^*)) = 0 \text{ i.e. } p^* = \frac{3}{4} \text{ and } q^* \in [0,1] \end{cases}$$

From the definition, for a strategy to be an NE, it must satisfy both sets of inequalities obtained above. Combining the above two sets of 3 inequalities, we can conclude that the only possibilities are the following:

- If $p^* = 1$, then only the first inequality from the second set is feasible, i.e. $q^* = 1$, yielding the NE $\{(1,0);(1,0)\}$ (Already found in step 2).
- If $p^* = 0$, then only the second inequality from the second set is feasible, i.e. $q^* = 0$, yielding the NE $\{(0,1);(0,1)\}$ (Already found in step 2).
- If $q^* = \frac{3}{4}$, then only the third inequality from the second set is feasible, i.e. $p^* = \frac{3}{4}$, yielding the NE $\{(\frac{3}{4},\frac{1}{4});(\frac{3}{4},\frac{1}{4})\}$.

Since there is no other σ^* satisfying both sets of inequalities, the above 3 is the set of all NE of this game.

Another useful trick to find Completely Mixed strategies–The Indifference Principle

Proposition 4.4 (Support Property). *In any mixed equilibrium σ^*, all pure strategies s_i (for both players) that receive positive probability under σ^* yield equal expected payoffs:*

$$u_i(\sigma_{-i}, s_i) = u_i(\sigma_{-i}, s_i') \quad \textit{for all } s_i, s_i' \in Supp(\sigma_i^*).$$

Otherwise, the player could shift probability towards the more profitable pure strategy.

This leads to the so-called Indifference Principle, often used to compute mixed equilibria in 2×2 games.

Example 1. *Consider the Matching Pennies game. Let p = probability the row player plays H, and q = probability the column player plays H. Row player's expected payoff:*

$$u_R(H) = q * 1 + (1-q) * 0; \qquad u_R(T) = q * 0 + (1-q) * 1$$

Using the above support property we get $q = 1 - q \implies q = \frac{1}{2}$.

Similarly, using the column player's expected utility argument, we can show that $p = \frac{1}{2}$.

Note: In this calculation we assume both players play the strategy with non-zero probability. This method can not detect pure strategies.

5 Two-player zero-sum games

A zero-sum game is a strategic environment in which two players (decision-makers) have opposed objectives.

Motivating example of 'Min-Max'

Consider the following game G. We only write the utility values of the row player. The utilities of the column player are always minus the utility value of the row player.

	J	1	2	3
J	+1	-1	-1	-1
1	-1	-1	+1	+1
2	-1	+1	-1	+1
3	-1	+1	+1	-1

Consider the point of view of the row player, and assume that she uses the mixed strategy σ_R in which J is played with some probability q and each of 1,2, and 3 is played with probability $(1-q)/3$. Then,

$$u_R(\sigma_R, J) = 2q - 1 \quad \text{and} \quad u_R(\sigma_R, 1) = u_R(\sigma_R, 2) = u_R(\sigma_R, 3) = \frac{1}{3} - \frac{4q}{3}.$$

If we assume that the column player knows what q is and would do what is best for her (which here happens to be what is worst for the row player), then the row player would like to choose q that maximizes the minimum of $\{2q-1, \frac{1}{3} - \frac{4q}{3}\}$. This happens when $q = 2/5$, in which case the row player's utility is $-1/5$, regardless of which (including pure and mixed) strategy the column player chooses.

The same reasoning can be used to show that the column player will also choose $q = 2/5$. Her expected utility will be $+1/5$. Note that this strategy profile $\{(\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}); (\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})\}$ for the above game is a mixed Nash equilibrium. This leads to the following theorem:

Theorem 5.1 (Expected value principal)**.** *If you know that your opponent is playing a given mixed strategy, and will continue to play it regardless of what you do, you should play your strategy which has the largest expected utility/payoff/value.*

5.1 Definition and results

A two player game $G = (N = \{1,2\}, \{S_i\}, \{u_i\})$ is called zero-sum if $u_1 + u_2 = 0$. For such games, we drop the subscript on the utility functions and use $u := u_1$.

Remark 2. *A general Strategic/normal form game may have cooperative components (see, e.g., the battle of the sexes game), in the sense that moving from one strategy profile to another can benefit both players. However, zero-sum games are competitive: whatever one player gains exactly the other player loses. Hence, a rational player always wants to prepare herself for the worst possible outcome, namely one in which, given her own strategy, her opponent will choose the strategy yielding her the minimal utility/payoff. Hence, an interesting quantity for a strategy σ_1 for player 1 is the guaranteed utility u_g when playing σ_1:*

$$u_g(\sigma_1) = \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2).$$

Continuing with this line of reasoning, player 1 will choose a strategy that maximizes her guaranteed utility. She would thus choose an action in

$$\underset{\sigma_1 \in \Delta S_1}{\text{argmax}} u_g(\sigma_1) = \underset{\sigma_1 \in \Delta S_1}{\text{argmax}} \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2).$$

Any such strategy is called a max-min strategy for player 1. It gives her the best possible guaranteed utility (highest achievable payoff) when using a worst-case model, i.e., under the assumption that Player-2 is rational and maximizing their own payoff, which is

$$\underline{v} := \max_{\sigma_1 \in \Delta S_1} u_g(\sigma_1) = \max_{\sigma_1 \in \Delta S_1} \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2).$$

A similar line of argument can be written for player 2 using the worst-case model from player 2's perspective.

Note that $\max\min(u)$ represents the payoff that the row player can guarantee if she goes first, and correspondingly $\min\max(u)$ represents the payoff that she (row player) can guarantee if the column player goes first.

It is intuitively true (we will not do the formal proof; this links to the primal-dual problem we have learnt in Utility theory) that playing second can only be an advantage: your strategy space is not limited by the first player's action, and you have more information. In particular, this implies that for any game with player 1's utility denoted as u, one has

$$\bar{v} := \min\max(u) \geq \max\min(u) =: \underline{v}$$

Theorem 5.2 (The Minimax Theorem–Von-Neumann (1928))**.** *In any zero-sum game G with player 1's utility represented by u, one has*

$$\bar{v} = \underline{v}.$$

Furthermore, in this case the unique value $\bar{v} = \underline{v}$ is known as the value of the game.

Notation

From now onwards, the utility function u for a zero-sum game will always be represented in a matrix form $A = (a_{ij})_{m \times n}$, where

- player 1 is the row player and has m many actions (pure strategies) $\{1, 2, \cdots, m\}$.
- player 2 is the column player and has n many actions (pure strategies) $\{1, 2, \cdots, n\}$.
- $a_{i,j}$ represents player 1's utility when choosing pure strategy/action i and player 2 chooses pure strategy/action j.
- as a result of a zero-sum game, player 2's utility will be $-a_{ij}$.
- player 1 wants to maximize its utility/payoff, and player 2 wants to minimize its utility/payoff.
- player 1's (mixed) strategy is denoted by x such that $x = (x_1, x_2, \cdots, x_m)$, with $\sum_{i=1}^m x_i = 1$ and $x_i \geq 0$.
- player 2's (mixed) strategy is denoted by y such that $y = (y_1, y_2, \cdots, y_n)$, with $\sum_{i=1}^n y_i = 1$ and $y_i \geq 0$.
- the utility/payoff of the game $G = G_A$, when player 1 and 2 play strategy x, y, respectively, is

$$A(x, y) := \sum_{i=1}^{m} \sum_{j=1}^{n} x_i a_{ij} y_j = xAy^\top$$

The next theorem shows that maxmin strategies and mixed Nash equilibria are closely related in zero-sum games.

Theorem 5.3 (Borel, 1921, von Neumann, 1928)**.** *Let G be a finite zero-sum game with payoff matrix given by $A_{m\times n}$. Then:*

1. In every mixed Nash equilibrium $\sigma^* = (x^*, y^*)$ each strategy is a maxmin strategy for that player.

2. There is a unique real number $v \in \mathbb{R}$ such that

$$max_x min_y u(x, y) = v,$$
$$max_y min_x - u(x, y) = -v.$$

The quantity v is called the value of $G = G_A$. It follows that $u(\sigma^*) = u(x^*, y^*) = v$ for any equilibrium $\sigma^* = (x^*, y^*)$.

Furthermore, the optimal strategy/value of the game can always be found from a $k \times k$ square sub game of the original $m \times n$ game.

Remark 3. *Another way to interpret value of the game is 'For any matrix game, if there is a number $v \in \mathbb{R}$ such that row player has a strategy which guarantees that she will win at least v, and the column player has a strategy which guarantees that the row player will win no more than v, then this number is called the value of the game.'*

Remark 4 (Minimax)**.** *For any matrix game $G = G_A$, the value of the game with payoff matrix A can be computed in the following two equivalent ways:*

$$v(G) = \max_x \min_y xAy^\top = \min_y \max_x xAy^\top.$$

Theorem 5.4. *Let A be the payoff matrix of a game. Then two stochastic vectors x^* and y^* are optimal strategies for player 1 and player 2 respectively if and only if*

$$\textit{Minimum entry of } x^*A = \textit{ Maximum entry of } Ay^{*\top}.$$

In this case, the common value of those expressions is the value of the game.

Proposition 5.5. *Let $A_{n\times n}$ be the payoff matrix of a game. Assume $x^* = (x_1^*, x_2^*, \cdots, x_n^*)$ and $y^* = (y_1^*, y_2^*, \cdots, y_n^*)$ are optimal for player 1 and player 2 respectively. Then:*

$$x^*A \geq v \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix},$$
$$A(y^*)^\top \leq v \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Calculating Nash Equilibria–from a computational prospective

- Can one practically compute a max min strategy (and thus a Nash equilibrium) of a general game?
- How difficult is it computationally for 2-player games?

In particular, for a two-player zero-sum game, **a** Nash equilibrium (not all) can be computed in polynomial time using linear programming. This is a polynomial in the total number of strategies of the two players: $|S_1|$ and $|S_2|$. It can be formulated as a primal convex optimization problem as follows:

$$
\begin{aligned}
\text{(Primal)} \quad \max_{x\in\mathbb{R}^m,\, v\in\mathbb{R}} \quad & v \\
\text{s.t.} \quad & \sum_{i=1}^{m} x_i A_{ij} \geq v, \quad j = 1,\dots,n, \\
& \sum_{i=1}^{m} x_i = 1, \\
& x_i \geq 0, \quad i = 1,\dots,m.
\end{aligned}
$$

The dual formulation (again a convex optimization problem) of the above problem is as follows:

$$
\begin{aligned}
\text{(Dual)} \quad \min_{y\in\mathbb{R}^n,\, w\in\mathbb{R}} \quad & w \\
\text{s.t.} \quad & \sum_{j=1}^{n} y_j A_{ij} \leq w, \quad i = 1,\dots,m, \\
& \sum_{j=1}^{n} y_j = 1, \\
& y_j \geq 0, \quad j = 1,\dots,n.
\end{aligned}
$$

5.2 Symmetric games are fair

A game is called fair if its value is zero. This means that, on average, the wins and losses for each player balance out. Remember that a game is symmetric if "its rules are the same for player one and player two", or more formally, if its payoff matrix is antisymmetric, i.e., $A = -A^\top$.

It seems obvious that a symmetric game is fair: if the rules are the same for both players, neither player should have an advantage! The formal proof is as follows:

Lemma 5.6. *For any payoff matrix A one has*

$$v(A) = -v(-A^\top).$$

Proof.

$$
\begin{aligned}
v(-A^\top) &= \max_x \min_y x(-A^\top)y^\top \\
&= \max_x \min_y \left(x(-A^\top)y^\top\right)^\top \\
&= \max_x \min_y -yAx^\top \\
&= \max_x -(\max_y yAx^\top) \\
&= -\min_x \max_y yAx^\top \\
&= -\min_y \max_x xAy^\top = -v(A).
\end{aligned}
$$

□

As a consequence of the above lemma, one has the following proposition.

Proposition 5.7. *If the game $G = G_A$ is symmetric, that is, if $A = -A^\top$, then*

$$v(G_A) = 0.$$

Furthermore, if x^* is optimal for player 1, x^* will also be optimal for player 2.

This implies every symmetric game is a fair game.

Example 2 (Solving for Nash-equilibrium-(Rock-Paper-Scissors))**.** *This is a two-player game with payoff matrix A given by*

$A=$

	R	P	S
R	0	-1	2
P	1	0	-3
S	-2	3	0

.

A classical R-P-S game has a payoff matrix $\bar{A} =$

	R	P	S
R	0	-1	1
P	1	0	-1
S	-1	1	0

.

Since $A = -A^\top$, we know that $v(G_A) = 0$. So, if we assume $x = (x_1, x_2, x_3)$ to be an optimal strategy for player 1, then

$$(x_1, x_2, x_3)\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 3 & 0 \end{pmatrix} \geq 0$$

$$\iff \begin{cases} x_2 - 2x_3 \geq 0 \\ -x_1 + 3x_3 \geq 0 \\ 2x_1 - 3x_2 \geq 0. \end{cases}$$

If we assume equality in the above set of equations, we get:

$$x_1 = \frac{3}{6}, x_2 = \frac{2}{6}, x_3 = \frac{1}{6}.$$

So indeed $\{(\frac{3}{6}, \frac{2}{6}, \frac{1}{6}), (\frac{3}{6}, \frac{2}{6}, \frac{1}{6})\}$ is the NE for the game.
HW: *In a similar line of arguments, calculate the NE for the classical R-P-S game denoted by the payoff matrix $\bar{A}$.*

Definition 6 (Saddle point)**.** *A saddle point of a pay-off matrix A is the position(s) where the maximum of the row-minima coincides with the minimum of the column-maxima.*

So for a payoff matrix A, max min of player 1 = min max of player 2 is the saddle point of the game.

Example 3. *Take the payoff matrix as*

$$A = \begin{pmatrix} 6 & 8 & 6 \\ 4 & 12 & 2 \end{pmatrix}.$$

The saddle points of this matrix game are given by the points $(1, 1)$ and $(1, 3)$.

Example 4. *Find all the saddle points in the following game*

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

No saddle point exists for this game.

Remark 5. *Saddle points may or may not exist for a game. There may be more than one saddle point.*

Theorem 5.8 (Saddle point principal)**.** *The saddle point corresponds to the the value of the game. Saddle point also provides a pure NE of the game.*

So in the above example, the value of the game is 6. Generally, in zero-sum games, every NE $\iff$ Saddle Point.
HW: If multiple saddle points exist, do they have to be equal?

Remark 6. *For a Non-zero-sum game one can still define saddle points, but the above theorems break down. For example, in 'Battle of the sexes', the $Maxmin$ value of the row player is* 0*, and the $Minmax$ value of the column player is* 1*. So, no saddle point exists in pure strategies. But is has two pure NE! In non-zero-sum games, NE does not necessarily correspond to a saddle point!*

References

[1] M. J. Osborne and A. Rubinstein, *A course in game theory*, MIT press, 1994.