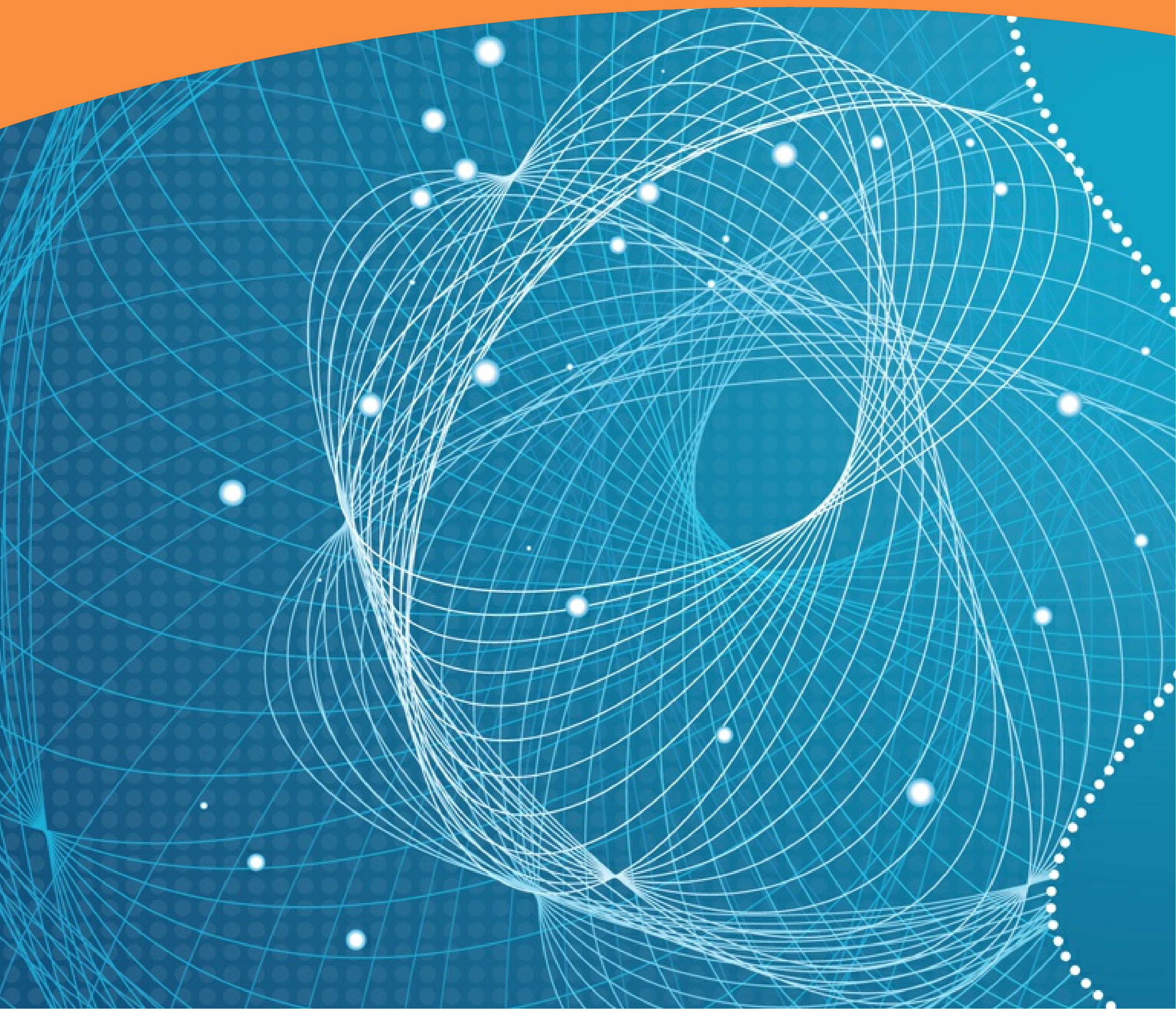
Mathematical Models in Portfolio Analysis

Farida Kachapova



bookboon
The eBook company

Farida Kachapova

Mathematical Models in Portfolio Analysis

Mathematical Models in Portfolio Analysis
1st edition

ISBN 978-87-403-0370-4

Contents

Click on the ad to read more

Preface

Portfolio analysis is the part of financial mathematics that is covered in existing textbooks mainly from the financial point of view without focussing on mathematical foundations of the theory. The aim of this book is to explain the foundations of portfolio analysis as a consistent mathematical theory, where assumptions are stated, steps are justified and theorems are proved. However, we left out details of the assumptions for equilibrium market and capital asset pricing model in order to keep the focus on mathematics.

Part 1 of the book is a general mathematical introduction with topics in matrix algebra, random variables and regression, which are necessary for understanding the financial chapters. The mathematical concepts and theorems in Part 1 are widely known, so we explain them briefly and mostly without proofs.

The topics in Part 2 include portfolio analysis and capital market theory from the mathematical point of view. The book contains many practical examples with solutions and exercises.

The book will be useful for lecturers and students who can use it as a textbook and for anyone who is interested in mathematical models of financial theory and their justification. The book grew out of a course in financial mathematics at the Auckland University of Technology, New Zealand.

Dr. Farida Kachapova

Part 1:
Mathematical Introduction

In Chapters 1–4 we briefly describe some basic mathematical facts necessary for understanding of the book.

1 Matrices and Applications

1.1 Terminology

- A **matrix** is a rectangular array of numbers.
- A matrix with m rows and n columns is called an $m\times n$-**matrix** (m by n matrix).
- An $n\times n$-matrix is called a **square** matrix.
- A $1\times n$-matrix is called a **row** matrix.
- An $m\times 1$-matrix is called a **column** matrix.

Example 1.1.

1) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$ is a 2×4-matrix.

2) $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ is a 3×3-matrix, a square matrix.

3) $\begin{bmatrix} 2 & 4 & 6 \end{bmatrix}$ is a 1×3-matrix, a row.

4) $\begin{bmatrix} 5 \\ 3 \\ 2 \\ 1 \end{bmatrix}$ is a 4×1-matrix, a column. ☐

For a matrix A the element in row i, column j is denoted a_{ij}. So $A = [a_{ij}]$ is the short notation for the matrix A.

Denote **0** a column of all zeroes (the length is usually obvious from context).

A square matrix $A = [a_{ij}]$ is called **symmetric** if $a_{ij} = a_{ji}$ for any i, j.

An $n\times n$-matrix is called **identity** matrix and is denoted $\boldsymbol{I_n}$ if its elements are $a_{ij} = \begin{cases} 1, \textit{if} & i = j, \\ 0, \textit{if} & i \neq j. \end{cases}$

So the identity matrix has the form $I_n = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$.

1.2 Matrix Operations

1.2.1 Multiplication by a Number.

For a number c and matrix $A = [a_{ij}]$ the product $c \cdot A = [ca_{ij}]$. Each element of the matrix A is multiplied by the number c.

Example 1.2. $-2 \times \begin{bmatrix} 3 & 4 \\ 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -8 \\ -2 & 0 \\ 2 & -2 \end{bmatrix}$. □

1.2.2 Transposition.

This operation turns the rows of a matrix into columns. The result of transposition of matrix A is called the **transpose** matrix and is denoted A^T. For $A = [a_{ij}]$, $A^T = [a_{ji}]$.

Example 1.3. For $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$, $A^T = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$. □

1.2.3 Matrix Addition.

For $m \times n$-matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ the **sum** is defined by: $A + B = [a_{ij} + b_{ij}]$.

Example 1.4.

1) $\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 1 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 6 \\ 0 & 8 & 9 \end{bmatrix}$.
2) For $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}$, $A + B$ is not defined, since A and B have different dimensions. □

1.2.4 Matrix Multiplication.

For an $m \times n$-matrix $A = [a_{ij}]$ and an $n \times p$-matrix $B = [b_{ij}]$ the **product** $C = A \cdot B = [c_{ij}]$ is defined by: $c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$ $(1 \le i \le m,\ 1 \le j \le p)$.

So the product $A \cdot B$ is defined only when the number of columns in A equals the number of rows in B.

Example 1.5.

1) $\begin{bmatrix} 1 & 2 & 4 \\ 3 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 5 & -2 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 7 & 2 & 2 \\ -4 & 15 & -7 & 0 \end{bmatrix}.$

2) For $A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 & 7 \\ 1 & 0 & -1 \end{bmatrix}$, $A \cdot B$ is not defined, since the number of columns in A is 3 and the number of rows in B is 2 (different). □

> **Theorem 1.1.** 1) For a symmetric matrix A, $A^T = A$.
>
> 2) If $A{\cdot}B$ is defined, then $B^T{\cdot}A^T$ is defined and $(A{\cdot}B)^T = B^T{\cdot}A^T$.

1.2.5 Inverse Matrix.

Suppose A and B are $n{\times}n$-matrices. B is called the **inverse** of A if $A{\cdot}B = B{\cdot}A = I_n$.
If matrix A has an inverse, then A is called an **invertible** matrix.
If matrix A is invertible, then the inverse is unique and is denoted A^{-1}.

1.2. Exercises

1. If A is an $m{\times}n$-matrix, what is the dimension of its transpose A^T?
2. If A is an $m{\times}n$-matrix and B is an $n{\times}p$-matrix, what is the dimension of their product $A{\cdot}B$?
3. When a column of length n is multiplied by a row of length n, what is the dimension of their product?
4. When a row of length n is multiplied by a column of length n, what is the dimension of their product?
5. Prove that for any $m{\times}n$-matrix A: $AI_n = A$ and $I_m A = A$.
6. For $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix}$ show that $AB \neq BA$.
7. Suppose A is an $m{\times}n$-matrix, B is an $n{\times}k$-matrix and C is a $k{\times}p$-matrix. Prove that $(A{\cdot}B)\cdot C = A\cdot(B{\cdot}C)$.
8. Suppose $x = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$, $y = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}$ and $S = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -2 & 0 \\ 3 & 0 & 1 \end{bmatrix}$. Calculate the following:

   1) $x^T y$, 2) $x^T S\, y$, 3) $x^T S\, x$. *Answers*: 1) 3, 2) 4, 3) 23.
9. Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -2 & 0 \\ 3 & 0 & 2 \end{bmatrix}$ and $B = \frac{1}{6}\begin{bmatrix} -4 & -4 & 6 \\ -4 & -7 & 6 \\ 6 & 6 & -6 \end{bmatrix}$. Show that $B = A^{-1}$.

1.3 Determinants

We will define the **determinant** *det A* for any $n\times n$-matrix A using induction by n.

1) For a 1×1-matrix A (a number) $det\ A = A$.

2) For a 2×2-matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ the determinant is defined by: $det\ A = ad - bc$.

3) Now suppose $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ is an $n\times n$-matrix, $n \geq 3$.

- For each element a_{ij} the corresponding **minor** M_{ij} is the determinant of the matrix obtained from A by removing row i and column j, and the corresponding **cofactor** $A_{ij} = (-1)^{i+j} M_{ij}$.
- The **determinant** of A is defined by $det\ A = \sum_{j=1}^{n} a_{kj} A_{kj}$, where k is any integer between 1 and n. The result does not depend on the choice of k.

The determinant of A is also denoted $\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$.

Example 1.6.

1) $\begin{vmatrix} 5 & 3 \\ 2 & 1 \end{vmatrix} = 5\times 1 - 2\times 3 = -1.$

2) $\begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{vmatrix} = 1\cdot\begin{vmatrix} 5 & 8 \\ 6 & 9 \end{vmatrix} - 4\cdot\begin{vmatrix} 2 & 8 \\ 3 & 9 \end{vmatrix} + 7\cdot\begin{vmatrix} 2 & 5 \\ 3 & 6 \end{vmatrix} = 1\cdot(-3) - 4\cdot(-6) + 7\cdot(-3) = 0.$ ☐

Theorem 1.2. A square matrix A is invertible if and only if $det\ A \neq 0$.

1.3. Exercises

1. For $n\times n$ matrices A and B prove the following.
   1) $det\ (A^T) = det\ A$.
   2) $det\ (A\cdot B) = det(A) \cdot det(B)$.
   3) $det\ (I_n) = 1$.
   4) If A^{-1} exists, then $det\ (A^{-1}) = \dfrac{1}{det\ A}$.

2. For any invertible matrix A prove the following.
   1) $(A^{-1})^T = (A^T)^{-1}$. 2) If A is symmetric, then A^{-1} is symmetric.

3. Find the determinant of the matrix A. Is A invertible?
   1) $A = \begin{bmatrix} 2 & -5 \\ 4 & 3 \end{bmatrix}$. 2) $A = \begin{bmatrix} 2 & -1 & 2 \\ 5 & 0 & -2 \\ -3 & 3 & 1 \end{bmatrix}$.

Answers: 1) 26, invertible, 2) 41, invertible.

1.4 Systems of Linear Equations

Consider a system of m linear equations with n unknowns:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{1}$$

It can be written in matrix form $AX = B$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},\ X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \text{ and } B = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}.$$

A system of the form (1) is called **homogeneous**, if $B = \mathbf{0}$.

Cramer's Rule. If $m = n$ and $\det A \neq 0$, then the system (1) has a unique solution given by:

$x_i = \dfrac{\Delta_i}{\Delta}$ ($i = 1, \dots, n$), where $\Delta = det\ A$ and Δ_i is the determinant obtained from *det A* by replacing the i-th column by the column B.

Theorem 1.3. If $m < n$, then a homogeneous system of m linear equations with n unknowns has a non-zero solution (that is a solution different from $\mathbf{0}$).

Theorem 1.4. Suppose X_0 is a solution of system (1). Then
X is a solution of the system (1) $\Leftrightarrow X = X_0 + Y$ for some solution Y of the corresponding homogeneous system $AX = \mathbf{0}$, where all $b_1, b_2, \dots, b_m$ are replaced by zeroes.

1.5 Positive Definite Matrices

A symmetric $n \times n$-matrix S is called **positive definite** if for any $n \times 1$-matrix $x \neq \mathbf{0}$,

$$x^T S\, x > 0.$$

A symmetric $n \times n$ -matrix S is called **non-negative definite** if for any $n \times 1$-matrix x,

$$x^T S\, x \geq 0.$$

A symmetric matrix S is called **negative definite** if the matrix $-S$ is positive definite.

For a square matrix A, a **principal leading minor** of A is the determinant of an upper left corner of A.

So for the matrix $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ the principal leading minors are:

$$\Delta_1 = a_{11},\quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\quad \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \dots,\ \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Sylvester Criterion. A symmetric matrix S is positive definite if and only if each principal leading minor of S is positive.

Example 1.7. Determine whether the matrix S is positive definite, negative definite or neither.

$$1)\ S=\begin{bmatrix}5 & 2\\ 2 & 1\end{bmatrix},\quad 2)\ S=\begin{bmatrix}3 & -1 & 2\\ -1 & 2 & 1\\ 2 & 1 & 5\end{bmatrix},\quad 3)\ S=\begin{bmatrix}-9 & 5 & 6\\ 5 & -5 & 0\\ 6 & 0 & -10\end{bmatrix},\quad 4)\ S=\begin{bmatrix}6 & 0 & 2\\ 0 & 3 & 1\\ 2 & 1 & 1\end{bmatrix}.$$

Solution

We will use the Sylvester criterion.

1) The principal leading minors of S are: $\Delta_1 = 5 > 0$ and $\Delta_2 = \begin{vmatrix}5 & 2\\ 2 & 1\end{vmatrix} = 1 > 0$. They are both positive, hence the matrix S is positive definite.

2) The principal leading minors of S are: $\Delta_1 = 3 > 0$, $\Delta_2 = \begin{vmatrix}3 & -1\\ -1 & 2\end{vmatrix} = 5 > 0$, and

$$\Delta_3 = \begin{vmatrix}3 & -1 & 2\\ -1 & 2 & 1\\ 2 & 1 & 5\end{vmatrix} = 10 > 0.$$ They are all positive, hence the matrix S is positive definite.

3) The first leading minor is $-9 < 0$, so the matrix S is not positive definite.

To check whether it is negative definite, consider the matrix $-S = \begin{bmatrix} 9 & -5 & -6 \\ -5 & 5 & 0 \\ -6 & 0 & 10 \end{bmatrix}$.

For this matrix $\Delta_1 = 9 > 0$, $\Delta_2 = \begin{vmatrix} 9 & -5 \\ -5 & 5 \end{vmatrix} = 20 > 0$, and $\Delta_3 = \begin{vmatrix} 9 & -5 & -6 \\ -5 & 5 & 0 \\ -6 & 0 & 10 \end{vmatrix} = 20 > 0$.

The matrix $-S$ is positive definite, so S is negative definite.

4) $\Delta_3 = \begin{vmatrix} 6 & 0 & 2 \\ 0 & 3 & 1 \\ 2 & 1 & 1 \end{vmatrix} = 0$. Hence the matrix S is neither positive definite, nor negative definite.

One can also check that for $x = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}$, $x^{\mathrm{T}} S\, x = 0$. ☐

Theorem 1.5. If an $n \times n$-matrix S is positive definite, then it is invertible and S^{-1} is also positive definite.

Proof

By the Sylvester criterion $\det S > 0$, so S is invertible by Theorem 1.2.

Consider an $n\times 1$-matrix $x \neq \mathbf{0}$ and denote $y = S^{-1} x$. Then y is also an $n\times 1$-matrix. If $y = \mathbf{0}$, then $S^{-1} x = \mathbf{0}$, $S\,(S^{-1}x) = \mathbf{0}$ and $x = \mathbf{0}$. Contradiction. Hence $y \neq \mathbf{0}$.

S is symmetric, so S^{-1} is also symmetric. $y^T S\, y = (S^{-1}x)^T S\, (S^{-1}x) = x^T (S^{-1})^T I_n\, x = x^T S^{-1} x$. So $x^T S^{-1} x = y^T S\, y > 0$ because S is positive definite. Therefore S^{-1} is positive definite. ☐

1.6 Hyperbola

Standard hyperbola is the curve on (x, y)-plane given by an equation of the form: $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

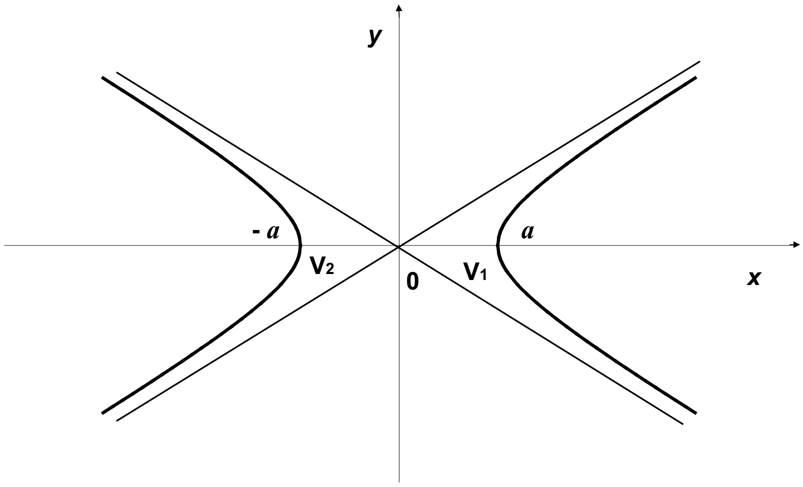


Figure 1.1. Standard hyperbola

- The parameters of the hyperbola are a^2 and b^2.
- The centre is at the point (0, 0).
- The vertices are v_1 $(a, 0)$ and v_2 $(-a, 0)$.
- The asymptotes of the hyperbola are given by the equations: $y = \pm \frac{b}{a} x$.

Vertically translated hyperbola is given by an equation of the form: $\frac{x^2}{a^2} - \frac{(y - y_0)^2}{b^2} = 1$.

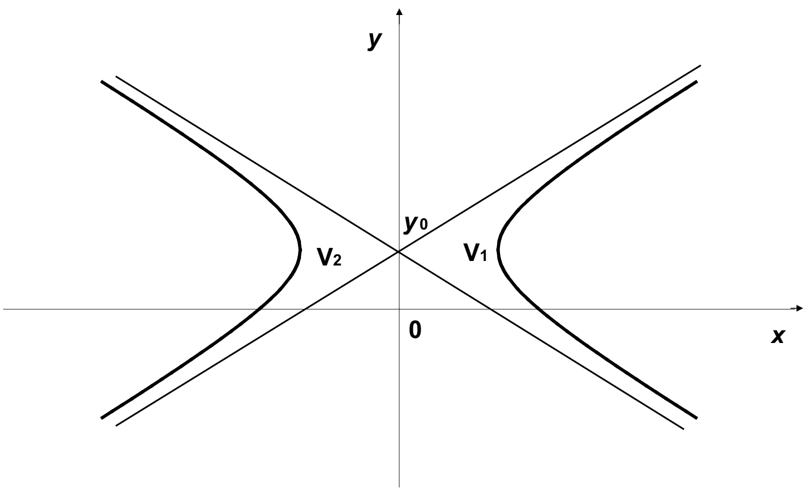


Figure 1.2. Vertically translated hyperbola

- The parameters of the hyperbola are a^2 and b^2.
- The centre is at the point $(0, y_0)$.
- The vertices are $v_1\,(a, y_0)$ and $v_2\,(-a, y_0)$.
- The asymptotes of the hyperbola are given by the equations: $y - y_0 = \pm \frac{b}{a} x$.

More details on hyperbola and curves of second degree can be found in textbooks on analytic geometry; see, for example, Riddle (1995), and Il'in and Poznyak (1985).

2 Orthogonal Projection

2.1 Orthogonal Projection onto a Subspace

Denote $\boldsymbol{R}$ the set of all real numbers. Denote $\boldsymbol{R}^n$ the set of all ordered sequences of real numbers of length n.

A non-empty set L with operations of addition and multiplication by a real number is called a **linear space** if it satisfies the following 10 axioms:

for any $x, y, z \in L$ and $\lambda, \mu \in \boldsymbol{R}$:

1) $(x + y) \in L$;

2) $\lambda x \in L$;

3) $x + y = y + x$;

4) $(x + y) + z = x + (y + z)$;

5) there exists an element $\boldsymbol{0} \in L$ such that $(\forall x \in L)(\boldsymbol{0} + x = x)$;

6) for any $x \in L$ there exists $-x \in L$ such that $-x + x = \boldsymbol{0}$;

7) $1 \cdot x = x$;

8) $(\lambda\mu)\, x = \lambda(\mu x)$;

9) $(\lambda + \mu)\, x = \lambda x + \mu x$;

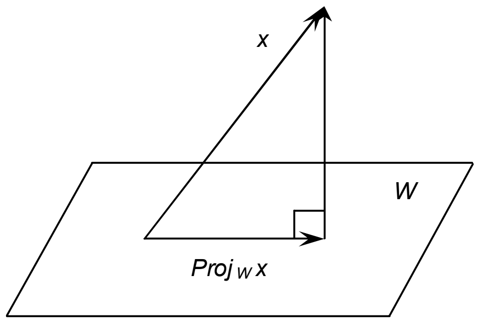
10) $\lambda(x + y) = \lambda x + \lambda y$.

In this chapter we consider a Euclidean space L, that is a linear space with scalar product. In this space (x, y) denotes the scalar product of vectors x and y; $\| x \| = \sqrt{(x,x)}$ is the norm of the vector x.

Vectors x and y are called **orthogonal** ($x \perp y$) if the scalar product $(x, y) = 0$.

Suppose x is a vector in L and W is a linear subspace of L. A vector z is called the **orthogonal projection** of x onto W if $z \in W$ and $(x - z) \perp W$.

Then z is denoted $Proj_W\, x$.



Theorem 2.1.

1) $Proj_W x$ is the closest to x vector in W and it is the only vector with this property.

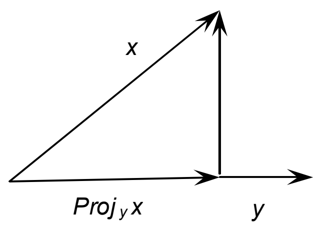
2) If $v_1, \ldots, v_n$ is an orthogonal basis in W, then

$$Proj_W x = \frac{(x, v_1)}{(v_1, v_1)} v_1 + \ldots + \frac{(x, v_n)}{(v_n, v_n)} v_n .$$

2.2 Orthogonal Projection onto a Vector

The **orthogonal projection** of a vector x onto a vector y is $Proj_W x$, where $W = \{ty \mid t \in \mathbf{R}\}$. This projection is denoted $Proj_y x$.

The length of $Proj_y x$ is called the **orthogonal scalar projection** of x onto y and is denoted $proj_y x$.



$$proj_y x = \frac{(x, y)}{\| y \|}; \qquad Proj_y x = proj_y x \cdot \frac{y}{\| y \|} .$$

2.3 Minimal Property of Orthogonal Projection

A subset Q of a linear space B is called an **affine subspace** of B if there is $q \in Q$ and a linear subspace W of B such that $Q = \{q + w \mid w \in W\}$. Then W is called the corresponding linear subspace.

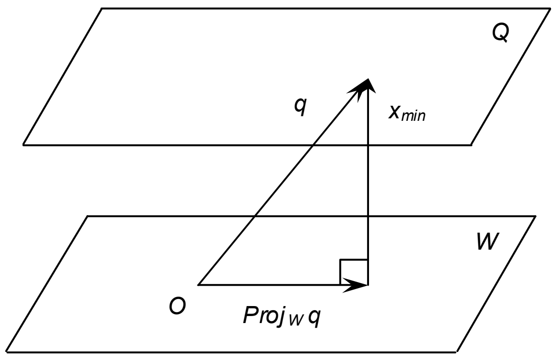
It is easy to check that any vector in Q can be taken as q.

Lemma 2.1. Consider a consistent system of m linear equations with n unknowns in its matrix form: $AX = B$. The set of all solutions of the system $AX = B$ is an affine subspace of $\boldsymbol{R}^n$ and its corresponding linear subspace is the set of all solutions of the homogeneous system $AX = \mathbf{0}$.

Theorem 2.2. Let $Q = \{q + w \mid w \in W\}$ be an affine subspace of L. Then the vector in Q with smallest length is unique and is given by the formula:

$$x_{min} = q - Proj_{W}\, q.$$

Proof



Denote $z = Proj_{W}\, q$, then $x_{min} = q - z$.

Consider any vector $y \in Q$. For some $w \in W$, $y = q + w$. By Theorem 2.1.1), z is the vector in W closest to q and $-w \in W$, so we have

$$\| y \| = \| q - (-w) \| \geq \| q - z \| = \| x_{min} \|.$$

The equality holds only when $-w = z$, that is when $y = q + w = q - z = x_{min}$.

Since x_{min} is unique, it does not depend on the choice of q. ☐

3 Random Variables

3.1 Numerical Characteristics of a Random Variable

Consider a **probability space** $(\Omega, \Im, P)$ where Ω is a sample space of elementary events (outcomes), $\Im$ is a σ-field of events and P is a probability measure on the pair $(\Omega, \Im)$. We will fix the probability space for the rest of the chapter.

- A function $X: \Omega \to \boldsymbol{R}$ is called a **random variable** if for any real number x, $\{\omega \in \Omega \mid X(\omega) \le x\} \in \Im$.
- The **distribution function** F of a random variable X is defined by $F(x) = P\{X \le x\}$ for any real number x.

- A random variable X is called **discrete** if the set of its possible values is finite or countable.

- A function f is called the **density function** of a random variable X if for any real number x:

  $f(x) \geq 0$ and $F(x) = \int_{-\infty}^{x} f(t)dt$ for the distribution function F of X.
- A random variable X is called **continuous** if it has a density function.

The **distribution table** of the discrete variable X is the table

x_1	x_2	x_3	…
p_1	p_2	p_3	…

(2)

where $x_1, x_2, x_3,\ldots$ are all possible values of X and $p_i = P(X = x_i)$, $i = 1, 2, 3, \ldots$

Example 3.1. A player rolls a fair die. He wins \$1 if a three turns up, he wins \$5 if a four turns up and he wins nothing otherwise. Denote X the value of a win.

Here the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. $\Im$ is the set of all subsets of Ω. The function X is defined by: $X(1) = X(2) = X(5) = X(6) = 0$, $X(3) = 1$, $X(4) = 5$.

Clearly X is a discrete random variable.

Since the die is fair, the probability of getting any of the numbers 1, 2, 3, 4, 5, 6 equals $\frac{1}{6}$. This is the distribution table of X:

x_i	0	1	5
p_i	$\frac{4}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

□

Define a binary relation ~ for random variables: $X \sim Y$ if $P\{\omega \mid X(\omega) \neq Y(\omega)\} = 0$. Next two lemmas are about this binary relation.

Lemma 3.1. The defined relation ~ is an equivalence relation on random variables.

Proof

a) For any random variable X, $X \sim X$ because $\{\omega \mid X(\omega) \neq X(\omega)\} = \varnothing$, the empty set. So ~ is reflexive.
b) For any random variables X and Y, $X \sim Y$ implies $Y \sim X$ because $\{\omega \mid X(\omega) \neq Y(\omega)\} =$
   $= \{\omega \mid Y(\omega) \neq X(\omega)\}$. So ~ is symmetric.

c) Assume that for random variables X, Y, Z, $X \sim Y$ and $Y \sim Z$. Then

$\{\omega \mid X(\omega) \neq Z(\omega)\} \subseteq \{\omega \mid X(\omega) \neq Y(\omega)\} \cup \{\omega \mid Y(\omega) \neq Z(\omega)\}$ and

$0 \leq P\{\omega \mid X(\omega) \neq Z(\omega)\} \leq P\{\omega \mid X(\omega) \neq Y(\omega)\} + P\{\omega \mid Y(\omega) \neq Z(\omega)\} = 0 + 0 = 0$. So $\sim$ is transitive. ☐

In other words, two random variables X and Y are **equivalent** ($X \sim Y$) if they are equal with probability 1.

Lemma 3.2.

1) For any random variables X_1, X_2 and $\lambda \in \mathbf{R}$: $X_1 \sim X_2 \Rightarrow (\lambda X_1) \sim (\lambda X_2)$.
2) For any random variables X_1, X_2, Y: $X_1 \sim X_2 \Rightarrow (X_1 + Y) \sim (X_2 + Y)$.
3) For any random variables X_1, X_2, Y_1, Y_2: $X_1 \sim X_2$ & $Y_1 \sim Y_2 \Rightarrow (X_1 + Y_1) \sim (X_2 + Y_2)$.

Proof

1) is obvious.
2) follows from the equality $\{ \omega \mid X_1(\omega) + Y(\omega) \neq X_2(\omega) + Y(\omega)\} = \{ \omega \mid X_1(\omega) \neq X_2(\omega)\}$.
3) follows from 2) and the fact that $\sim$ is an equivalence relation. ☐

Denote $[X]$ the equivalence class of a random variable X.

Operations of addition and multiplication by a real number on equivalence classes are given by the following: $[X] + [Y] = [X + Y]$ and $\lambda \cdot [X] = [\lambda X]$.

Lemma 3.2 makes these definitions valid.

In the rest of the book we will use the notation X instead of $[X]$ for brevity remembering that equivalent random variables are considered equal.

- The **expected value** of a discrete random variable X with possible values $x_1, x_2, x_3, \ldots$ is $E(X) = \sum_i x_i P(X = x_i)$.
- The **expected value** of a continuous random variable X with density function f is $E(X) = \int_{-\infty}^{\infty} x f(x) dx$.
- Expected value is also called **expectation** or **mean value**.
- $E(X)$ is also denoted μ_X or μ.

- The **variance** of the random variable X is $Var(X) = E[(X - \mu_X)^2]$.
- The **standard deviation** of the random variable X is $\sigma_X = \sqrt{Var(X)}$. It is also denoted $\sigma(X)$.

Thus, $Var(X) = \sigma_X^2$. For the discrete random variable X with distribution table (2):

$$Var(X) = \sum_i (x_i - \mu_X)^2 p_i .$$

Both variance and standard deviation are measures of spread of the random variable.

Example 3.2. Find the expected value, variance and standard deviation of the random variable from Example 3.1.

Solution

$$E(X) = \sum_i x_i P(X = x_i) = 0 \cdot \frac{4}{6} + 1 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} = 1.$$

$$Var(X) = (0-1)^2 \cdot \frac{4}{6} + (1-1)^2 \cdot \frac{1}{6} + (5-1)^2 \cdot \frac{1}{6} = \frac{10}{3} \cdot \sigma_X = \sqrt{\frac{10}{3}} . \ \square$$

Properties of expectation. For any random variables *X*, *Y* and real number *c*:

- $E(cX) = cE(X)$;
- $E(c) = c$;
- $E(X+Y) = E(X) + E(Y)$;
- $X \geq 0 \Rightarrow E(X) \geq 0$;
- $X \leq Y \Rightarrow E(X) \leq E(Y)$.

Properties of variance. For any random variable *X* and real number *c*:

- $Var(X) \geq 0$;
- $Var(c) = 0$;
- $Var(X+c) = Var(X)$;
- $Var(cX) = c^2 Var(X)$;
- $Var(X) = E(X^2) - (EX)^2$.

Properties of standard deviaton. For any random variable *X* and real number *c*:

- $\sigma(X) \geq 0$;
- $\sigma(c) = 0$;
- $\sigma(X+c) = \sigma(X)$;
- $\sigma(cX) = |c|\sigma(X)$.

Example 3.3. The random variable *X* has expectation of –3 and standard deviation of 2. Calculate the following:

1) $E(2X)$, 2) $E(-3X)$, 3) $E(-X)$, 4) $Var(2X)$, 5) $Var(-3X)$, 6) $Var(-X)$, 7) $\sigma(2X)$, 8) $\sigma(-3X)$, 9) $\sigma(-X)$.

Solution

$E(X) = -3$ and $\sigma = \sigma_X = 2$. $Var(X) = \sigma^2 = 4$.

1) $E(2X) = 2\,E(X) = 2 \cdot (-3) = -6$.
2) $E(-3X) = -3\,E(X) = -3 \cdot (-3) = 9$.
3) $E(-X) = -\,E(X) = 3$.
4) $Var(2X) = 2^2 \cdot Var(X) = 16$.
5) $Var(-3X) = (-3)^2 \cdot Var(X) = 36$.
6) $Var(-X) = (-1)^2 \cdot Var(X) = 4$.
7) $\sigma(2X) = \sqrt{Var(2X)} = \sqrt{16} = 4$.
8) $\sigma(-3X) = \sqrt{Var(-3X)} = \sqrt{36} = 6$.
9) $\sigma(-X) = \sqrt{Var(-X)} = \sqrt{4} = 2$. ☐

3.2 Covariance and Correlation Coefficient

The **covariance** of random variables X and Y is $Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$.

The **correlation coefficient** of random variables X and Y is $\rho_{X,Y} = \dfrac{Cov(X,Y)}{\sigma_X \sigma_Y}$.

Correlation coefficient is the normalised covariance.

Random variables X and Y are called **independent** if for any $x, y \in \mathbf{R}$:

$$P(X \le x \text{ and } Y \le y) = P(X \le x) \cdot P(Y \le y).$$

Properties of covariance. For any random variables X, Y, Z and real number c:

- $Cov(Y,X) = Cov(X,Y)$;
- $Cov(X,X) = Var(X)$;
- $Cov(X,c) = 0$;
- $Cov(X,Y+c) = Cov(X,Y)$;
- $Cov(X,cY) = c\,Cov(X,Y)$;
- $Cov(X,Y+Z) = Cov(X,Y) + Cov(X,Z)$;
- $|Cov(X,Y)| \le \sigma_X \cdot \sigma_Y$;
- $Cov(X,Y) = E(XY) - E(X) \cdot E(Y)$;
- $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$;
- if X and Y are independent variables, then $Cov(X,Y) = 0$ and $Var(X+Y) = Var(X) + Var(Y)$.

Properties of correlation coefficient. For any random variables X and Y:

- $\rho_{Y,X} = \rho_{X,Y}$;
- $\rho_{X,X} = 1$;
- $-1 \le \rho_{X,Y} \le 1$;
- if X and Y are independent, then $\rho_{X,Y} = 0$.

Example 3.4. Random variables X and Y have the following parameters:

$$E(X) = 15,\ \sigma(X) = 3,\ E(Y) = -10,\ \sigma(Y) = 2,\ Cov(X, Y) = 1.$$

For $Z = 2X + 5Y$ calculate the following: 1) $E(Z)$, 2) $Var(Z)$, 3) $\sigma(Z)$.

Solution

$Var(X) = \sigma^2(X) = 9$, $Var(Y) = \sigma^2(Y) = 4$.

1) $E(Z) = 2E(X) + 5E(Y) = 2 \cdot 15 - 5 \cdot 10 = -20$.
2) $Var(Z) = Var(2X) + Var(5Y) + 2\,Cov(2X, 5Y) = 2^2\,Var(X) + 5^2\,Var(Y) + 2{\cdot}2{\cdot}5\,Cov(X, Y) =$
$= 2^2 \cdot 9 + 5^2 \cdot 4 + 20 \cdot 1 = 156$.
3) $\sigma(Z) = \sqrt{Var(Z)} = \sqrt{156}$. □

Example 3.5. Random variables X and Y have the following parameters:

$$E(X) = 15,\ \sigma(X) = 3,\ E(Y) = -10,\ \sigma(Y) = 2,\ Cov(X, Y) = 1.$$

For $Z = X - Y$ calculate the following: 1) $E(Z)$, 2) $Var(Z)$, 3) $\sigma(Z)$.

Solution

1) $E(Z) = E(X) - E(Y) = 15 + 10 = 25$.
2) $Var(Z) = Var(X + (-Y)) = Var(X) + Var(-Y) + 2Cov(X, -Y) = Var(X) + Var(Y) - 2Cov(X, Y)$
$= 9 + 4 - 2 \cdot 1 = 11$.
3) $\sigma(Z) = \sqrt{Var(Z)} = \sqrt{11}$. □

Example 3.6. Random variables X and Y have the following parameters:

$$E(X) = 15,\ \sigma(X) = 3,\ E(Y) = -10,\ \sigma(Y) = 2,\ Cov(X, Y) = 1.$$

For $Z = 2X - 5Y$ calculate the following: 1) $E(Z)$, 2) $Var(Z)$, 3) $\sigma(Z)$.

Solution

1) $E(Z) = 2E(X) - 5E(Y) = 2 \cdot 15 + 5 \cdot 10 = 80.$
2) $Var(Z) = Var(2X + (-5\ Y)) = Var(2X) + Var(-5Y) + 2\ Cov(2X, -5Y) =$
   $= 2^2\ Var(X) + (5)^2\ Var(Y) - 2{\cdot}2{\cdot}5\ Cov(X, Y) = 2^2 \cdot 9 + 5^2 \cdot 4 - 20 \cdot 1 = 116.$
3) $\sigma(Z) = \sqrt{Var(Z)} = \sqrt{116}$. ☐

3.3 Covariance Matrix

- A set of real numbers $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ is called **trivial** if $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.
- A group of random variables $X_1, X_2, \dots, X_n$ is called **linearly dependent** if for some non-trivial set of real numbers $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$,

  $$(\lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_n X_n) \text{ is constant.}$$

- A group of random variables $X_1, X_2, \dots, X_n$ is called **linearly independent** if it is not linearly dependent.

Lemma 3.3.

1) If random variables X and Y are independent, then they are linearly independent.
2) The inverse is not true.
3) X and Y are linearly dependent $\Leftrightarrow |\ Cov(X, Y)\ | = \sigma_X \cdot \sigma_Y$.

For random variables $X_1, X_2, \dots, X_n$, denote $\sigma_{ij} = Cov(X_i, X_j)$. The matrix

$$S = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix}$$

is called the **covariance matrix** of $X_1, X_2, \dots, X_n$.

Properties of covariance matrix.

Suppose S is the covariance matrix of random variables $X_1, X_2,\dots, X_n$. Then

- S is symmetric;
- S is non-negative definite;
- $X_1, X_2,\dots, X_n$ are linearly dependent $\Leftrightarrow$ $det\ S = 0$;
- $X_1, X_2,\dots, X_n$ are linearly independent $\Leftrightarrow$ $det\ S > 0 \Leftrightarrow$ S is positive definite (see the definition in Section 1.5).

4 Regression

4.1 Euclidean Space of Random Variables

Define $H = \{X \mid X$ is a random variable on $(\Omega, \Im, P)$ and $E(X^2) < \infty\}$. Thus, H is the set of all random variables on the probability space, whose squares have finite expectations. A similar approach is used in the textbook by Grimmett and Stirzaker (2004).

Lemma 4.1.

1) For any $X \in H$ and $\lambda \in \boldsymbol{R}$: $E(X^2) < \infty \quad \Rightarrow \quad E[(\lambda X)^2] < \infty$.
2) For any $X, Y \in H$: $E(X^2) < \infty$ & $E(Y^2) < \infty \quad \Rightarrow \quad E[(X + Y)^2] < \infty$.

Proof

1) It follows from the fact that $E[(\lambda X)^2] = \lambda^2 E(X^2)$.
2) Since $2XY \le X^2 + Y^2$, we have $0 \le (X + Y)^2 = X^2 + Y^2 + 2XY \le 2X^2 + 2Y^2$ and this implies $E[(X + Y)^2] < \infty$. ☐

Lemma 4.1 shows that the set H is closed under the operations of addition and multiplication by a real number.

Theorem 4.1. The set H (where equivalent random variables are considered equal) with the operations of addition and multiplication by a real number is a linear space.

Proof

H is a linear space means it satisfies the following 10 axioms: for any $X, Y, Z \in H$ and $\lambda, \mu \in \boldsymbol{R}$,

1) $(X + Y) \in H$;
2) $\lambda X \in H$;
3) $X + Y = Y + X$;
4) $(X + Y) + Z = X + (Y + Z)$;
5) there exists an element $\boldsymbol{0} \in H$ such that $(\forall X \in H)(\boldsymbol{0} + X = X)$;
6) for any $X \in H$ there exists $-X \in H$ such that $-X + X = \boldsymbol{0}$;
7) $1 \cdot X = X$;
8) $(\lambda\mu) X = \lambda(\mu X)$;
9) $(\lambda + \mu) X = \lambda X + \mu X$;
10) $\lambda(X + Y) = \lambda X + \lambda Y$.

In probability theory it is proven that for any random variables X and Y their sum $X + Y$ is a random variable, and for any real number λ the product λX is also a random variable. Together with Lemma 4.1 this proves the conditions 1) and 2). $\boldsymbol{0}$ in condition 5) is the random variable that always equals 0. The remaining conditions are quite obvious. ❑

For any $X, Y \in H$, define $\varphi(X, Y) = E(XY)$. We have: $E(X^2) < \infty$ and $E(Y^2) < \infty$.

Since $|XY| \leq \frac{1}{2}X^2 + \frac{1}{2}Y^2$, then $E(XY) < \infty$. So the definition $\varphi(X, Y)$ is valid for any $X, Y \in H$.

Theorem 4.2. For any $X, Y \in H$ and $\lambda \in \boldsymbol{R}$,

1. $\varphi(Y, X) = \varphi(X, Y)$;
2. $\varphi(\lambda X, Y) = \lambda\varphi(X, Y)$;
3. $\varphi(X + Y, Z) = \varphi(X, Z) + \varphi(Y, Z)$;
4. $\varphi(X, X) \geq 0$;
5. $\varphi(X, X) = 0 \Rightarrow X = \boldsymbol{0}$.

Proof

Properties 1–4 follow directly from the definition of φ and properties of expectation.

5. If $\varphi(X, X) = 0$, then $E(X^2) = 0$ and $X = \boldsymbol{0}$ with probability 1. ❑

Theorem 4.2 implies the following.

$(X, Y) = E(XY)$ defines a **scalar product** on the linear space H.
H with this scalar product is a Euclidean space.

In simple cases we can construct a basis of the linear space H. The following example illustrates that.

Example 4.1. Consider a finite sample space $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ with the probabilities of the outcomes $p_i = P(\omega_i) > 0$, $i = 1, \ldots, n$. In this case we can introduce a finite orthogonal basis in the Euclidean space H.

For each i define a random variable F_i as follows: $F_i(\omega_j) = \begin{cases} 1 & if \quad j = i, \\ 0 & if \quad j \neq i. \end{cases}$

Then for any random variable X in H,

$$X = \sum_{i=1}^{n} x_i F_i \,, \tag{3}$$

where $x_i = X(\omega_i)$.

For any $i \neq j$, $F_i \cdot F_j = 0$ and $(F_i, F_j) = E(F_i \cdot F_j) = 0$, so

$$F_i \perp F_j. \qquad (4)$$

(3) and (4) mean that $F_1,\ldots, F_n$ make an orthogonal basis in H and the dimension of H is n.

For any $X, Y \in H$, their scalar product equals $(X, Y) = \sum_{i=1}^{n} p_i x_i y_i$, where $y_i = Y(\omega_i)$. □

Define **norm** on H by the following: $\| X \| = \sqrt{(X, X)}$ for any $X \in H$.

Define **distance** in H by the following: $d(X, Y) = \| X - Y \|$ for any $X, Y \in H$.

Since $d(X, Y) = \sqrt{\left(E(X - Y)^2\right)}$, the distance between two random variables X and Y is the average difference between their values.

Denote I the random variable that equals 1 with probability 1:

$$I(\omega) = 1 \text{ for any } \omega \in \Omega.$$

We will call I the **unit variable**.

Since $I^2 = I$ and $E(I) = 1$, we have $I \in H$.

Lemma 4.2. For any $X \in H$, $E(X)$ and $Var(X)$ are defined.

Proof

Consider $X \in H$. The scalar product (X, I) is defined, since $I \in H$. $(X, I) = E(XI) = E(X)$, so $E(X)$ is defined. Since $X \in H$, $E(X^2)$ is defined. Therefore $Var(X) = E(X^2) - E(X)^2$ is defined. □

Lemma 4.3. Let $X \in H$ have parameters $E(X) = \mu$ and $Var(X) = \sigma^2$, $\sigma \geq 0$. Then

1) $\| X \| = \sqrt{\sigma^2 + \mu^2}$; 2) $\| X - \mu \| = \sigma$.

Proof

1) $\| X \|^2 = (X, X) = E(X^2) = Var(X) + E(X)^2 = \sigma^2 + \mu^2$, so $\| X \| = \sqrt{\sigma^2 + \mu^2}$.
2) follows from 1) because $E(X - \mu) = 0$ and $Var(X - \mu) = \sigma^2$. □

Example 4.2. Suppose $X \in H$, $E(X) = -2$ and $Var(X) = 5$. Then by Lemma 4.3:

$\| X \| = \sqrt{\sigma^2 + \mu^2} = \sqrt{5 + (-2)^2} = 3$ and $\| X + 2 \| = \| X - \mu \| = \sigma = \sqrt{5}$. □

$\angle(X, Y)$ denotes the **angle** between random variables X and Y.

X and Y are called **orthogonal** ($X \perp Y$) if $\angle(X, Y) = 90°$.

Lemma 4.4. Suppose $X, Y \in H$ and they have the following parameters:

$$E(X) = \mu_1,\ Var(X) = \sigma_1^2,\ E(Y) = \mu_2,\ Var(Y) = \sigma_2^2.$$

Then:

1) $(X, Y) = Cov(X, Y) + \mu_1 \mu_2$;

2) if $X \neq \mathbf{0}$ and $Y \neq \mathbf{0}$, then $Cos\,\angle(X,Y) = \dfrac{Cov(X,Y) + \mu_1\mu_2}{\sqrt{\mu_1^2 + \sigma_1^2} \cdot \sqrt{\mu_2^2 + \sigma_2^2}}$;

3) $Cos\angle(X - \mu_1, Y - \mu_2) = \rho_{X,Y}$, the correlation coefficient of X and Y;

4) $X \perp Y \Leftrightarrow Cov(X, Y) = -\mu_1 \mu_2$;

5) if X and Y are independent, then $(X, Y) = \mu_1 \mu_2$.

Proof

1) follows from a property of covariance: $Cov(X, Y) = E(XY) - E(X)\,E(Y) = (X, Y) - \mu_1 \mu_2$.

2) By the definition $Cos \angle(X,Y) = \dfrac{(X,Y)}{|X| \cdot |Y|}$ and the rest follows from 1) and Lemma 4.3.1).

3) $Cos\angle(X - \mu_1, Y - \mu_2) = \dfrac{(X-\mu_1, Y-\mu_2)}{|X-\mu_1| \cdot |Y-\mu_2|} = \dfrac{E((X-\mu_1)(Y-\mu_2))}{\sigma_1\sigma_2} = \dfrac{Cov(X,Y)}{\sigma_1\sigma_2} = \rho_{X,Y}$ by Lemma 4.3.2).

4) $X \perp Y \Leftrightarrow (X, Y) = 0 \Leftrightarrow Cov(X, Y) + \mu_1 \mu_2 = 0$ by 1).

5) follows from 1) because the covariance of independent random variables equals 0. ❒

Example 4.3. Suppose $X, Y \in H$ and they have the following parameters:

$$E(X) = 2,\ Var(X) = 4,\ E(Y) = 4,\ Var(Y) = 9,\ Cov(X, Y) = -2.$$

Then by Lemmas 4.3 and 4.4:

1) $(X, Y) = Cov(X, Y) + \mu_1 \mu_2 = -2 + 2 \cdot 4$, $(X, Y) = 6$;

2) $\|X\| = \sqrt{4+2^2}$, $\|X\| = \sqrt{8}$; $\|Y\| = \sqrt{9+4^2}$, $\|Y\| = 5$;

3) $Cos\angle(X, Y) = \dfrac{(X,Y)}{\|X\| \cdot \|Y\|} = \dfrac{6}{\sqrt{8} \cdot 5} = \dfrac{3}{5\sqrt{2}}$, $\angle(X, Y) = arccos\left(\dfrac{3}{5\sqrt{2}}\right) \approx 64.9°$;

4) $Cos\angle(X - 2, Y - 4) = \rho_{X,Y} = \dfrac{Cov(X,Y)}{\sigma_1\sigma_2} = \dfrac{-2}{\sqrt{4} \cdot \sqrt{9}} = \dfrac{-2}{2 \cdot 3} = -\dfrac{1}{3}$,

$\angle(X - 2, Y - 4) = arccos\left(-\dfrac{1}{3}\right) \approx 109.5°$. ❒

Example 4.4. Suppose $X, Y \in H$, $E(X) = 2$, $E(Y) = 5$ and $X \perp Y$. Find the covariance of X and Y.

Solution

By Lemma 4.4.4), $Cov(X, Y) = -\mu_1 \mu_2 = -2 \cdot 5 = -10$. ❒

Lemma 4.5. Let $X, Y \in H$. If X and Y are independent random variables, then $\|XY\| = \|X\| \cdot \|Y\|$.

Proof

X^2 and Y^2 are also independent, so $E(X^2\, Y^2) = E(X^2)\, E(Y^2)$ and $\|\, XY \,\|^2 = E(X^2\, Y^2) =$

$= E(X^2)\, E(Y^2) = \|\, X \,\|^2 \cdot \|\, Y \,\|^2$. Hence $\|\, XY \,\| = \|\, X \,\| \cdot \|\, Y \,\|$. □

Lemma 4.6. 1) $E(I) = 1$. 2) $Var(I) = 0$. 3) $(I, I) = 1$. 4) $\|\, I \,\| = 1$.

For any $X \in H$ with $E(X) = \mu$:

5) $Cov(X, I) = 0$, 6) $(X, I) = \mu$, 7) $proj_I X = \mu$, 8) $Proj_I X = \mu I$.

Proof

1) and 2) are obvious.

3), 4) Since $I^2 = I$, we have $(I, I) = E(I^2) = 1$ and $\|\, I \,\| = 1$.

5) follows from a property of covariance.

6) $(X, I) = E(X{\cdot}I) = E(X)$.

7) The scalar projection of X onto I is $proj_I X = \frac{(X, I)}{\| I \|} = \frac{\mu}{1} = \mu$.

8) The vector projection of X onto I is $Proj_I X = proj_I X \cdot I = \mu I$. ☐

4.2 Regression

Regression means "estimating an inaccessible random variable Y in terms of an accessible random variable X" (Hsu, 1997), that is finding a function $f(X)$ "closest" to Y. $f(X)$ can be restricted to a certain class of functions, the most common being the class of linear functions. We describe "closest" in terms of the distance d defined in Section 4.1.

Theorem 2.1 shows that $Proj_W Y$ is the vector in subspace W that minimizes distance $d(Y, U)$ from the fixed vector Y to vector U in W. In statistical terms, $Proj_W Y$ minimizes the **mean square error** $E((Y-U)^2) = d^2(Y, U)$ for vector U in W.

Theorem 4.3. The conditional expectation $E(Y \mid X)$ is the function of X closest to Y.

Proof

It is based on the following fact:

$$E(Y \mid X) = Proj_W Y$$

for $W = \{ f(X) \mid f\colon \boldsymbol{R} \to \boldsymbol{R} \text{ and } f(X) \in H \}$.

Grimmett & Stirzaker (2004) prove this fact by showing that $E(Y \mid X) \in W$ and that for any $h(X) \in W$, $E[(Y - E(Y \mid X)) \cdot h(X)] = 0$, that is $(Y - E(Y \mid X)) \perp h(X)$. ☐

By choosing different W in Theorem 2.1 we can get different types of regression: simple linear, multiple linear, quadratic, polynomial, etc.

4.3 Regression to a Constant

When we want to estimate a random variable Y by a constant, we use a subspace $W = \{aI \mid a \in \boldsymbol{R}\}$ of the space H.

Theorem 4.4. For any $Y \in H$ with $E(Y) = \mu$:

1) $Proj_W Y = \mu I$, we denote μI as μ;
2) μ is the constant closest to Y.

Proof

1) By Lemma 4.6.8), $Proj_I Y = \mu I$. So $(Y - \mu I) \perp I$ and $(Y - \mu I, I) = 0$.

   For any vector $aI \in W$, $(Y - \mu I, aI) = a\,(Y - \mu I, I) = 0$, so $(Y - \mu I) \perp aI$. By the definition of orthogonal projection, $\mu I = Proj_W Y$.

2) By Theorem 2.1, $Proj_W Y = \mu I$ is the vector in W closest to Y, and W is the set of constant random variables. So μI is the constant random variable closest to Y. ☐

Theorem 4.4 shows that the expectation $E(Y)$ is the best constant estimator for the random variable Y.

4.4 Simple Linear Regression

Theorem 4.5. If $\sigma_X \neq 0$, then the linear function of X closest to Y is given by

$$\alpha + \beta X, \text{ where } \begin{cases} \beta = \dfrac{Cov(Y,X)}{\sigma_X^2}, \\ \alpha = \mu_Y - \beta\mu_X. \end{cases} \qquad (5)$$

Proof

Denote $W = \{a + b\,X \mid a, b \in \boldsymbol{R}\}$. Since $Proj_W Y \in W$, we have $Proj_W Y = \alpha + \beta X$ for some $\alpha, \beta \in \boldsymbol{R}$. We just need to show that α and β are given by the formula (5).

For $\varepsilon = Y - Proj_W Y = Y - (\alpha + \beta X)$, we have $\varepsilon \perp 1$ and $\varepsilon \perp X$, since $1, X \in W$.

So $(\varepsilon, 1) = 0$ and $(\varepsilon, X) = 0$, $(\alpha + \beta X, 1) = (Y, 1)$ and $(\alpha + \beta X, X) = (Y, X)$, which leads to a system of linear equations:

$$\begin{cases} E(\alpha + \beta X) = E(Y) \\ E(\alpha X + \beta X \cdot X) = E(Y \cdot X) \end{cases} \text{ and } \begin{cases} \alpha + \beta\mu_X = \mu_Y \\ \alpha\mu_X + \beta E(X^2) = E(YX) \end{cases}.$$

Subtract the first equation multiplied by μ_X from the second equation:

$$\begin{cases} \alpha + \beta\mu_X = \mu_Y \\ \beta\left[(X,X) - \mu_X^2\right] = (Y,X) - \mu_Y\mu_X \end{cases}.$$

Since $(X, X) - \mu_X^{\,2} = \mathrm{E}(X^2) - \mu_X^{\,2} = \sigma_X^{\,2}$ and $(Y, X) - \mu_Y\mu_X = Cov(Y, X)$, we get

$$\begin{cases} \alpha + \mu_X\,\beta = \mu_Y \\ \sigma_X^2\,\beta = Cov(Y,X) \end{cases}.$$

The solution of this system is given by (5). ☐

Corollary. Denote $\hat{Y} = \alpha + \beta X$ the best linear estimator of Y from Theorem 4.5. The corresponding residual $\varepsilon = Y - \hat{Y}$ has the following properties:

$$1)\ \mu_{\varepsilon} = 0,\ 2)\ Cov\,(\varepsilon, X) = 0.$$

Proof

1) $\varepsilon \perp 1$, so $E(\varepsilon) = 0$.
2) $\varepsilon \perp X$, so $E(\varepsilon X) = 0$ and $Cov\,(\varepsilon, X) = E(\varepsilon X) - E(\varepsilon) \cdot E(X) = 0$. ❐

According to the Corollary, the residuals (estimation errors) equal 0 on average and are uncorrelated with the predictor X; this is another evidence that $\hat{Y}$ is the best linear estimator of Y.

Example 4.5. Create a linear regression model for a response variable Y versus a predictor variable X if $\mu_X = 2$, $\sigma_X = 1$, $\mu_Y = 3$ and $Cov(Y, X) = -2$.

Solution

According to Theorem 4.5, $\beta = \frac{-2}{1^2}$, $\beta = -2$ and $\alpha = 3 - 2\cdot(-2)$, $\alpha = 7$.

Hence $\hat{Y} = 7 - 2X$ is the linear function of X closest to Y; $Y = 7 - 2X + \varepsilon$. ❐

4.4. Exercise

1. Create a linear regression model for a response variable Y versus a predictor variable X if $\mu_X = 5$, $\sigma_X = 0.5$, $\mu_Y = 3$, $\sigma_Y = 0.2$ and $\rho_{X,Y} = 0.8$.

   Answer: $Y = 1.4 + 0.32\,X + \varepsilon$.

Part 2:
Portfolio Analysis

5 Portfolio Modelling

5.1 Terminology

- An **asset** is any possession that has value in an exchange, e.g. a financial asset or a physical asset like a car or a house.
- **Investment** is the creation of more money through the use of capital.
- A **security** is a contract with underlying value that can be bought or sold.
- A **bond** is a security issued or sold by a corporation or a government in order to borrow money from the public on a long-term basis.
- A **share** is a certificate entry representing ownership in a corporation or similar entity.
- **Stock** is ownership of a corporation indicated by shares, which represent a piece of the corporation's assets or earnings.

Example 5.1. The value of an investment increased from \$200 to \$260 in one year.

1) What is this investment's return in dollars?
2) What is this investment's return in percents?

Solution

1) 260 – 200 = \$60. The \$ return is \$60.
2) $\frac{60}{200}$ = 0.3 = 30%. The % return is 30%. ☐

A **portfolio** is a group of financial assets held by an investor.

Consider a portfolio consisting of assets $A_1, A_2, \ldots, A_N$.

The **proportion** (or **weight**) of asset A_k ($k = 1, 2, \ldots, N$) in the portfolio equals

$$\frac{\$ \text{ value invested in asset } A_k}{\text{total } \$ \text{ value invested in the portfolio}}.$$

We will represent a portfolio of assets $A_1, A_2, \ldots, A_N$ as a column $x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_N \end{bmatrix}$, where x_k is the proportion of asset A_k in the portfolio ($k = 1, 2, \ldots, N$). Clearly, $x_1 + x_2 + \ldots + x_N = 1$.

Example 5.2. Ben invested \$20 in stock A_1, \$30 in stock A_2 and \$50 in stock A_3. What are the stocks' proportions in this portfolio?

Solution

The total \$ value of the portfolio is \$20 + \$30 + \$50 = \$100. The proportions are:

for stock A_1: $x_1 = \frac{20}{100} = 0.2$, for stock A_2: $x_2 = \frac{30}{100} = 0.3$, for stock A_3: $x_3 = \frac{50}{100} = 0.5$.

So this portfolio is represented with the column $x = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}$. Note that 0.2 + 0.3 + 0.5 = 1. ☐

Example 5.3. Alan invested \$50 in stock A_1 and \$30 in stock A_2. A_1 has the return of 2% and A_2 has the return of 10% in one year.

1) What are the stocks' proportions in this portfolio?
2) What is the % return of the portfolio in one year?

Solution

1) The total \$ value of the portfolio is \$50 + \$30 = \$80.

   The proportions are, for stock A_1: $x_1 = \frac{50}{80} = \frac{5}{8} = 0.625$ and for stock A_2: $x_2 \frac{30}{80} = \frac{3}{8} 0.375$.

So the portfolio is represented with the column $x = \begin{bmatrix} 0.625 \\ 0.375 \end{bmatrix}$.

2) The \$ return for A_1 equals \$50×0.02 = \$1, the \$ return for A_2 equals \$30×0.10 = \$3, and the \$ return for portfolio x equals \$1 + \$3 = \$4. The % return for x equals $\frac{4}{80} = 0.05 = 5\%$.

Another way to calculate % return: 0.625×2% + 0.375×10% = 5%. ❑

> The % return r from an investment is treated as a random variable.
>
> The **expected return** from the investment is $\mu = E(r)$.
>
> The variance $Var(r)$ and standard deviation σ_r are used as measures of the investment's **risk**.

Example 5.4. The following table shows all possible returns from asset G:

Return from asset G (%)	20	12	4
Probability	0.3	0.5	0.2

For asset G calculate: 1) the expected return, 2) variance, 3) standard deviation.

Solution

Denote r the return from asset G; r is a random variable.

1) The expected return is $\mu = E(r) = 20 \times 0.3 + 12 \times 0.5 + 4 \times 0.2 = 12.8\%$.
2) The variance is $\sigma^2 = Var(r) = (20-12.8)^2 \times 0.3 + (12-12.8)^2 \times 0.5 + (4-12.8)^2 \times 0.2 = 31.36$.
3) The standard deviation is $\sigma = \sqrt{Var(r)} = \sqrt{31.36} = 5.6\%$. ❑

Next we fix a group of assets $A_1, A_2, \ldots, A_N$ and consider only portfolios of these assets. For $k = 1, \ldots, N$ denote

- v_k = \$ value of asset A_k in portfolio x;
- r_k = % return of asset A_k in one year.

The total value of portfolio x equals $\sum_{i=1}^{N} v_i$ and $x = \begin{bmatrix} x_1 \\ \ldots \\ x_N \end{bmatrix}$, where $x_k = \frac{v_k}{\sum_{i=1}^{N} v_i}$ $(k = 1, \ldots, N)$.

In one year the \$ return of portfolio x equals $\sum_{k=1}^{N} r_k v_k$ and the % return of x equals:

$$\frac{\sum_{k=1}^{N} r_k v_k}{\sum_{i=1}^{N} v_i} = \sum_{k=1}^{N} r_k \frac{v_k}{\sum_{i=1}^{N} v_i} = \sum_{k=1}^{N} r_k x_k .$$

In the rest of the book we will use the term return instead of % return for brevity. We will identify any portfolio x with its return and also denote the return by x. Thus, we have the following theorem.

Theorem 5.1. For any portfolio x of assets A_1, A_2,…, A_N,

1) the return $x = \sum_{k=1}^{N} x_k r_k$; 2) $x_1 + \ldots + x_N = 1$.

5.2 Short Sales

In the previous examples all proportions in the portfolios are positive. However, the case of negative proportions is also possible; it means short sales.

A **short sale** (or **short position**) represents the ability to sell a security that the seller does not own.

The purpose of it can be profiting from an expected price decline of the asset. In this book we will ignore all transaction costs.

In contrast to a short position, a **long position** involves simply buying and holding an asset. In this case the proportion of the asset in the portfolio is positive.

Example 5.5. Margaret invested \$30 in stock A_1, \$40 in stock A_2 and uses someone else's stock A_3, which is worth \$20 (a short sale). What are the stocks' proportions in this portfolio?

Solution

The total \$ value of the portfolio is 30 + 40 – 20 = \$50. The proportions are:

for stock A_1: $x_1 = \frac{30}{50} = 0.6$, for stock A_2: $x_2 = \frac{40}{50} = 0.8$, for stock A_3: $x_3 = \frac{-20}{50} = -0.4$.

So this portfolio is represented with the column $x = \begin{bmatrix} 0.6 \\ 0.8 \\ -0.4 \end{bmatrix}$. Note that 0.6 + 0.8 – 0.4 = 1. ☐

Example 5.6. John invested \$50 in stock A_1 and uses someone else's stock A_2, which is worth \$10 (a short sale). What are the stocks' proportions in this portfolio?

Solution

The total \$ value of the portfolio is 50 – 10 = \$40. The proportions are,

for stock A_1: $x_1 = \frac{50}{40} = 1.25$, for stock A_2: $x_2 = \frac{10}{40} = -0.25$.

So this portfolio is represented with the column $x = \begin{bmatrix} 1.25 \\ -0.25 \end{bmatrix}$. Note that 1.25 > 1. In case of short sales proportions over 1 are possible. ☐

5.3 Minimizing Risk

To measure the risk of an investment we will use the standard deviation of its return, since standard deviation expresses the variability of the return values of the investment.

Example 5.7. The following tables show possible returns from assets A_1 and A_2:

Return from asset A_1	3	8
Probability	0.4	0.6

Return from asset A_2	5	10
Probability	0.8	0.2

Which asset should an investor choose?

Solution

The expected returns are,
for asset A_1: $\mu_1 = 3 \times 0.4 + 8 \times 0.6 = 6$, for asset A_2: $\mu_2 = 5 \times 0.8 + 10 \times 0.2 = 6$.

The variances are,
for asset A_1: $\sigma_1^2 = (3 - 6)^2 \times 0.4 + (8 - 6)^2 \times 0.6 = 6$,
for asset A_2: $\sigma_2^2 = (5 - 6)^2 \times 0.8 + (10 - 6)^2 \times 0.2 = 4$.

Thus, the assets have the same expected returns. The variance is lower for the asset A_2 and so is the standard deviation. The investor should choose asset A_2, since it has less risk. ☐

Example 5.8. Suppose the returns of assets A_1 and A_2 from Example 5.7 have the covariance $\sigma_{12} = -1$. Find the expected returns and variances of the following portfolios:

1) $x = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$, 2) $y = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$, 3) $z = \begin{bmatrix} -0.2 \\ 1.2 \end{bmatrix}$.

4) Which portfolio is the best?

Solution

We have found parameters of the assets in Example 5.7.
For asset A_1: $\mu_1 = 6$, $\sigma_1^2 = 6$. For asset A_2: $\mu_2 = 6$, $\sigma_2^2 = 4$.

1) $x = 0.5r_1 + 0.5r_2$,
$\mu_x = 0.5\mu_1 + 0.5\mu_2 = 0.5 \times 6 + 0.5 \times 6 = 6$,
$\sigma_x^2 = Var(0.5r_1 + 0.5r_2) = 0.5^2\sigma_1^2 + 0.5^2\sigma_2^2 + 2 \times 0.5 \times 0.5 \times Cov(r_1, r_2) =$
$= 0.25 \times 6 + 0.25 \times 4 + 0.5 \times (-1) = 2$.

2) $y = 0.4r_1 + 0.6r_2$,
$\mu_y = 0.4\mu_1 + 0.6\mu_2 = 0.4 \times 6 + 0.6 \times 6 = 6$,
$\sigma_y^2 = Var(0.4r_1 + 0.6r_2) = 0.4^2 \times 6 + 0.6^2 \times 4 + 2 \times 0.4 \times 0.6 \times (-1) = 1.92$.

3) $z = -0.2r_1 + 1.2r_2$,
$\mu_z = -0.2\mu_1 + 1.2\mu_2 = -0.2 \times 6 + 1.2 \times 6 = 6$,

$\sigma_z^2 = Var(0.2r_1 + 1.2r_2) = 0.2^2 \times 6 + 1.2^2 \times 4 - 2 \times 0.2 \times 1.2 \times (-1) = 6.48.$

4) All three portfolios have the same expected return of 6. Portfolio y has the lowest variance $\sigma_y^2 = 1.92$. Portfolio y is the best, since it has the lowest risk. ☐

Example 5.9. Assets A_1 and A_2 have the covariance matrix $S = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$. Find:

1) the portfolio of these assets with lowest risk and 2) its variance.

Solution

Consider any portfolio $x = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}$ of the assets A_1 and A_2. Its return (denoted by the same letter) is $x = \gamma r_1 + (1-\gamma) r_2$ and its variance is $\sigma^2 = \gamma^2 \sigma_1^2 + (1-\gamma)^2 \sigma_2^2 + 2\gamma(1-\gamma)\, Cov(r_1, r_2)$.

From the covariance matrix S: $\sigma_1^2 = 2$, $\sigma_2^2 = 3$, $Cov(r_1, r_2) = \sigma_{12} = -1$.

So $\sigma^2 = 2\gamma^2 + 3(1-\gamma)^2 - 2\gamma(1-\gamma) = 7\gamma^2 - 8\gamma + 3$.

1) Denote $f(\gamma) = \sigma^2 = 7\gamma^2 - 8\gamma + 3$. We need to minimize $f(\gamma)$.

$f'(\gamma) = 14\gamma - 8$. The necessary condition is $f'(\gamma) = 0$.

$14\gamma - 8 = 0$, $\gamma = \frac{4}{7}$. $f''(\gamma) = 14 > 0$, so $\gamma = \frac{4}{7}$ is the point of minimum.

The portfolio with lowest risk is $x = \begin{bmatrix} \frac{4}{7} \\ \frac{3}{7} \end{bmatrix}$.

2) Its variance is $\sigma^2 = 7\gamma^2 - 8\gamma + 3 = 7 \cdot \left(\frac{4}{7}\right)^2 - 8 \cdot \frac{4}{7} + 3 = \frac{5}{7}$. □

5.3. Exercise

1. Two assets have expected returns of 20% and 10%, respectively, and the covariance matrix $S = \begin{bmatrix} 6 & 4 \\ 4 & 3 \end{bmatrix}$. Find the portfolio v of these assets with lowest risk, and find its expected return and variance.

   Answer: $v = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mu_v = 0$, $\sigma_v^2 = 2$.

5.4 Statistical Parameters of an N-Asset Portfolio

We continue to consider a fixed group of assets $A_1, \ldots, A_N$ and portfolios of these assets.

Notations

r_k denotes the return of asset A_k; $\mu_k = E(r_k)$, $\sigma_k^2 = Var(r_k)$ and $\sigma_{kj} = Cov(r_k, r_j)$.

$U = [1 \ \ldots \ 1]$ is the row of ones of length N.

$M = [\mu_1 \ \ldots \ \mu_N]$ is the row of the expected asset returns, where not all μ_k are the same. Then the vectors M and U are not proportional.

$S = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \cdots & \cdots & \cdots \\ \sigma_{N1} & \cdots & \sigma_{NN} \end{bmatrix}$ is the covariance matrix of the asset returns.

We will assume that all expected returns $\mu_1, \ldots, \mu_N$ are defined and the covariance matrix S exists with $det\, S > 0$. The positivity of the determinant of S is equivalent to the fact that the random variables $r_1, \ldots, r_N$ are linearly independent (see properties of covariance matrix in Section 3.3). This also implies that $r_1, \ldots, r_N$ are linearly independent as vectors in the Euclidean space H described in Section 4.1.

Let us consider the set K of all linear combinations of $r_1,\dots, r_N$. Clearly K is a linear subspace of H.

Theorem 5.2. 1) $r_1,\dots, r_N$ is a basis in K. 2) The dimension of K equals N.

According to Theorems 5.1 and 5.2, any portfolio x of assets $A_1,\dots, A_N$ is a vector in the N-dimensional Euclidean space K and can be represented as a column of its coordinates in the basis $r_1,\dots, r_N$:

$$x = \begin{bmatrix} x_1 \\ \dots \\ x_N \end{bmatrix}.$$

The portfolio's return is $x = \sum_{i=1}^{N} x_i\ r_i$ and its expected return is $E(x) = \mu_x = \sum_{i=1}^{N} x_i\ \mu_i = Mx$.

If $y = \begin{bmatrix} y_1 \\ \dots \\ y_N \end{bmatrix}$ is another portfolio of assets $A_1,\dots, A_N$, then $Cov(x,y) = Cov\left(\sum_{i=1}^{N} x_i r_i, \sum_{j=1}^{N} y_j r_j\right) =$

$= \sum_{i,j=1}^{N} x_i y_j Cov(r_i, r_j) = \sum_{i,j=1}^{N} x_i\ y_j \sigma_{ij} = x^T S\ y$, where x^T is the transpose of x.

If $y = x$, then $Var(x) = Cov(x,x) = x^T S\ x$. Thus, we have the following formulas for portfolios x and y.

Matrix formulas for portfolios' statistical parameters:

$$\mu_x = Mx,$$

$$\sigma_x^2 = x^T Sx,$$

$$Cov(x, y) = x^T Sy.$$

5.5 Envelope of Financial Assets

Generally portfolios with higher expected returns carry higher risks. However, it is possible to identify, among all portfolios with the same expected return, a portfolio with lowest risk, i.e., a portfolio with lowest variance.

A portfolio is called an **envelope portfolio** if it has the lowest variance among all portfolios with the same expected return.

An envelope portfolio minimizes risk for a given targeted return.

The set of all envelope portfolios is called the **envelope** of the assets $A_1, \dots, A_N$ and is denoted $Env(A_1, \dots, A_N)$.

A vector $x = \begin{bmatrix} x_1 \\ \ldots \\ x_N \end{bmatrix}$ is an envelope portfolio if it is a solution of the following minimization problem:

$$\begin{cases} Var(x) \to min \\ x_1 + \ldots + x_N = 1 \\ E(x) = \mu \end{cases} \tag{6}$$

for some fixed real number μ.

Portfolios are combinations of at least two assets, so only the case $N \geq 2$ is meaningful.

The problem (6) can be solved as a calculus problem in coordinate form using partial derivatives and Lagrange multipliers. Instead we will apply orthogonal projection in the Euclidean space K to produce an invariant geometric solution.

Since $E(x) = Mx$, the last two equations in (6) have this matrix form:

$$\begin{cases} Ux = 1 \\ Mx = \mu \end{cases} \tag{7}$$

This system has a solution, since M and U are not proportional. This is the corresponding homogeneous system:

$$\begin{cases} Ux = 0 \\ Mx = 0 \end{cases} \tag{8}$$

In case $N = 2$ for any $\mu \in \boldsymbol{R}$ there is only one portfolio with expected return μ. We will consider the case $N \geq 3$.

Denote Q and W the sets of all solutions of the systems (7) and (8), respectively. By Lemma 2.1, Q is an affine subspace of K and W is its corresponding linear subspace with dimension $N–2$ (since U and M are independent). The affine subspace Q can be written as $Q = \{q + w \mid w \in W\}$ for any solution q of the system (7).

Theorem 5.3. For any $\mu \in \boldsymbol{R}$, there is a unique envelope portfolio with expected return μ. It equals:

$$x_\mu = q - Proj_W q,$$

where q is any solution of (7).

Proof

An envelope portfolio x with expected return μ is a solution of the system (7) with smallest variance and

$$\| x \|^2 = (x, x) = E(x^2) = Var(x) + [E(x)]^2 = Var(x) + \mu^2.$$

So the smallest variance means the smallest length, since $E(x) = \mu$ is fixed. Next apply Theorem 2.2. ☐

Theorem 5.4. Let $v_1, \ldots, v_{N-2}$ be an orthogonal system of solutions of the system (8). Then

1) for any $y \in H$, $(y, v_k) = Cov(y, v_k)$, $k = 1, \ldots, N–2$;

2) the envelope portfolio with expected return μ equals:

$$x_\mu = q - \sum_{k=1}^{N-2} \frac{(q, v_k)}{(v_k, v_k)} v_k ,$$

where q is a solution of (7);

3) the envelope is the set $\{x_\mu \mid \mu \in \boldsymbol{R}\}$, where x_μ is given by the previous formula.

Proof

1) Each v_k is a solution of (8), so $E(v_k) = M v_k = 0$.

For any $y \in H$, $(y, v_k) = E(y \cdot v_k) = Cov(y, v_k) + E(y) \cdot E(v_k) = Cov(y, v_k) + E(y) \cdot 0 = Cov(y, v_k)$.

2) The system $v_1,..., v_{N-2}$ is an orthogonal basis in W, so the result follows from Theorems 5.3 and 2.1.2).
3) is obvious. ☐

Example 5.10. Three assets have expected returns of 3%, 1% and 3%, respectively, and covariance matrix

$S = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & -2 \\ -1 & -2 & 2 \end{bmatrix}$. For a targeted return of 5% find:

1) the portfolio of these assets with lowest risk and
2) its variance.

Solution

1) $M = [3\ 1\ 3]$, $\mu = 5$. The system (7) has the form $\begin{cases} x_1 + x_2 + x_3 = 1 \\ 3x_1 + x_2 + 3x_3 = 5 \end{cases}$. $q = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$ is a solution of this system.

Next we find a non-zero solution v of the homogeneous system $\begin{cases} x_1 + x_2 + x_3 = 0 \\ 3x_1 + x_2 + 3x_3 = 0 \end{cases}$: $v = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

By Theorem 5.4, $x_\mu = q - \dfrac{(q,v)}{(v,v)} v$, $(q, v) = Cov(q, v)$ and $(v, v) = Var(v)$.

So $(q, v) = q^T S\, v = 1$; $(v, v) = v^T S\, v = 5$; $x_\mu = q - \dfrac{(q,v)}{(v,v)} v = \begin{bmatrix} 1.8 \\ -1 \\ 0.2 \end{bmatrix}$.

2) The variance of x_μ equals $x_\mu^T S\, x_\mu = 2.8$.

Thus, an investor can expect a higher return (5%) from the right combination of the assets than from each asset, with the lowest possible risk of 2.8. ☐

Example 5.11. Four assets have expected returns of 1%, 3%, 1% and 1%, respectively, and covariance

matrix $S = \begin{bmatrix} 3 & 0 & 1 & 1 \\ 0 & 2 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$. For a targeted return of 3% find:

1) the portfolio of these assets with lowest risk and 2) its variance.

Solution

1) $M = [1\ 3\ 1\ 1]$, $\mu = 3$. The system (7) has the form $\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ x_1 + 3x_2 + x_3 + x_4 = 3 \end{cases}$. $q = \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}$ is a solution of this system.

Next we find a non-zero solution v_1 of the homogeneous system $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + 3x_2 + x_3 + x_4 = 0 \end{cases}$: $v_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$.

We also need to find a non-zero solution $v_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ of the same system that is orthogonal to v_1.

Thus, $0 = (v_1, v_2) = Cov(v_1, v_2) = v_1^T S\, v_2 = -x_2 + x_3 - 2x_4$.

We find v_2 as a solution of the system $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + 3x_2 + x_3 + x_4 = 0 \\ -x_2 + x_3 - 2x_4 = 0 \end{cases}$: $v_2 = \begin{bmatrix} -3 \\ 0 \\ 2 \\ 1 \end{bmatrix}$.

By Theorem 5.4, $x_\mu = q - \dfrac{(q, v_1)}{(v_1, v_1)} v_1 - \dfrac{(q, v_2)}{(v_2, v_2)} v_2$,

$(q, v_1) = Cov(q, v_1) = q^T S\, v_1 = 2$; $\qquad (q, v_2) = Cov(q, v_2) = q^T S\, v_2 = -2$;

$(v_1, v_1) = Var(v_1) = v_1^T S v_1 = 3;$ $(v_2, v_2) = Var(v_2) = v_2^T S v_2 = 15.$

So $x_\mu = q - \frac{(q, v_1)}{(v_1, v_1)} v_1 - \frac{(q, v_2)}{(v_2, v_2)} v_2 = \begin{bmatrix} -0.4 \\ 1 \\ 0.6 \\ -0.2 \end{bmatrix}.$

2) The variance of x_μ equals $x_\mu^T S x_\mu = 1.4.$

Thus, the portfolio x_μ has the same expected return of 3% as the second asset, but the lower risk of 1.4 versus 2. ☐

5.5 Exercises

1. Three assets have expected returns of 2%, 2% and 4%, respectively, and covariance matrix $S = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. For a targeted return of 3% find the portfolio of these assets with lowest risk and its variance.

   Answer: $x = \begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix}$, $\sigma_x^2 = 1.$

2. Three assets have expected returns of 4%, 2% and 4%, respectively, and covariance matrix $S = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 3 & 2 \\ -1 & 2 & 2 \end{bmatrix}$. For a targeted return of 3% find the portfolio of these assets with lowest risk and its variance.

   Answer: $x = \begin{bmatrix} 0.6 \\ 0.5 \\ -0.1 \end{bmatrix}$, $\sigma_x^2 = 0.45.$

3. Three assets have expected returns of 4%, 2% and 0%, respectively, and covariance matrix $S = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. For a targeted return of 5% find the portfolio of these assets with lowest risk and its variance.

   Answer: $x = \frac{1}{6}\begin{bmatrix} 4 \\ 7 \\ -5 \end{bmatrix}$, $\sigma_x^2 = \frac{7}{12}.$

4. Three assets have expected returns of 2%, 1% and 2%, respectively, and covariance matrix $S = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 6 \end{bmatrix}$. For a targeted return of 3% find the portfolio of these assets with lowest risk and its variance.

   Answer: $x = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, $\sigma_x^2 = 5$.

6 Mean-Variance Analysis

6.1 Principle of Two Fund Separation

Principle of Two Fund Separation. Let x and y be two envelope portfolios, $x \neq y$. Then the envelope equals the set $\{\gamma x + (1-\gamma) y \mid \gamma \in \boldsymbol{R}\}$.

Proof

We need to prove that $Env(A_1, \ldots, A_N) = \{\gamma x + (1-\gamma) y \mid \gamma \in \boldsymbol{R}\}$.

Denote $\mu_1 = E(x)$ and $\mu_2 = E(y)$. Fix an orthogonal system $v_1,\ldots, v_{N-2}$ of solutions of the system:

$$\begin{cases} Ux = 0 \\ Mx = 0 \end{cases},$$

where, as in Section 5.4, U is the row of ones and M is the row of the expected returns of the assets $A_1, \ldots, A_N$.

Since x and y are envelope portfolios, by Theorem 5.4.2), for some numbers μ_1, μ_2 and vectors q_1, q_2:

$$x = x_{\mu_1} = q_1 - \sum_{k=1}^{N-2} (q_1, v_k)\frac{v_k}{(v_k, v_k)}, \quad \begin{cases} Uq_1 = 1 \\ Mq_1 = \mu_1 \end{cases} \tag{9}$$

and

$$y = x_{\mu_2} = q_2 - \sum_{k=1}^{N-2} (q_2, v_k)\frac{v_k}{(v_k, v_k)}, \quad \begin{cases} Uq_2 = 1 \\ Mq_2 = \mu_2 \end{cases}. \tag{10}$$

For any $\gamma \in \boldsymbol{R}$, $U(\gamma q_1 + (1-\gamma) q_2) = \gamma Uq_1 + (1-\gamma) Uq_2 = \gamma + (1-\gamma) = 1$ by (9) and (10). So

$$\text{for any } \gamma \in \boldsymbol{R},\ U(\gamma q_1 + (1-\gamma) q_2) = 1. \tag{11}$$

$\Leftarrow$ Suppose $z = \gamma x + (1-\gamma) y$ for some $\gamma \in \boldsymbol{R}$. Denote $q = \gamma q_1 + (1-\gamma) q_2$ and $\mu = \gamma \mu_1 + (1-\gamma) \mu_2$. By (9) and (10),

$$z = \gamma q_1 + (1-\gamma) q_2 - \sum_{k=1}^{N-2} \left[\gamma (q_1, v_k) + (1-\gamma)(q_2, v_k)\right]\frac{v_k}{(v_k, v_k)} \text{ and } z = q - \sum_{k=1}^{N-2} (q, v_k)\frac{v_k}{(v_k, v_k)}.$$

By (11), $Uq = 1$ and $Mq = \gamma Mq_1 + (1-\gamma) Mq_2 = \gamma \mu_1 + (1-\gamma) \mu_2 = \mu$. So $\begin{cases} Uq = 1 \\ Mq = \mu \end{cases}$, $z = x_\mu$ and by Theorems 5.3 and 5.4.2), $z \in Env(A_1, \ldots, A_N)$.

$\Rightarrow$ Suppose $z \in Env(A_1, \ldots, A_N)$. Denote $\mu = E(z)$. First we solve the equation:

$$\gamma \mu_1 + (1-\gamma)\mu_2 = \mu. \qquad (12)$$

$\gamma(\mu_1 - \mu_2) = \mu - \mu_2$. If $\mu_1 = \mu_2$, then by Theorem 5.3, $x = x_{\mu_1} = x_{\mu_2} = y$. Therefore $\mu_1 \neq \mu_2$ and $\gamma = \frac{\mu - \mu_2}{\mu_1 - \mu_2}$ is the solution of the equation (12).

Denote $q = q_1 + (1-\gamma) q_2$. By (11), $Uq = 1$. $Mq = \gamma Mq_1 + (1-\gamma) Mq_2 = \gamma \mu_1 + (1-\gamma)\mu_2 = \mu$ by (12).

So $\begin{cases} Uq = 1 \\ Mq = \mu \end{cases}$.

Consider the envelope portfolio $x_\mu = q - \sum_{k=1}^{N-2} \frac{(q, v_k)}{(v_k, v_k)} v_k$. Clearly, $x_\mu = \gamma x + (1-\gamma) y$. By Theorems 5.3 and 5.4.2), x_μ is the only envelope portfolio with expected return μ, so $z = x_\mu = \gamma x + (1-\gamma) y$. ☐

The Principle of Two Fund Separation can be rephrased in financial terms: any portfolio on the envelope can be formed as a weighted average of any two funds on the envelope.

Example 6.1. In the world of assets A_1, A_2 and A_3 portfolios $x = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}$ and $y = \begin{bmatrix} -0.2 \\ 0.8 \\ 0.4 \end{bmatrix}$ are on the envelope. Find the envelope of these assets.

Solution

By the Principle of Two Fund Separation, $Env(A_1, A_2, A_3) = \{\gamma x + (1-\gamma) y \mid \gamma \in \boldsymbol{R}\} =$

$$= \left\{ \gamma \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix} + (1-\gamma) \begin{bmatrix} -0.2 \\ 0.8 \\ 0.4 \end{bmatrix} \middle| \gamma \in R \right\} = \left\{ \begin{bmatrix} 0.4\gamma - 0.2 \\ -0.5\gamma + 0.8 \\ 0.1\gamma + 0.4 \end{bmatrix} \middle| \gamma \in R \right\}. \ ☐$$

Example 6.2. Three assets have expected returns of 1%, 5% and 4%, respectively, and covariance matrix $S = \begin{bmatrix} 9 & 2 & 2 \\ 2 & 4 & 1 \\ 2 & 1 & 1 \end{bmatrix}$. Do all these assets lie on the envelope?

Solution

Denote the assets x, y and z. Suppose all these assets are on the envelope. By the Principle of Two Fund Separation, $z = \gamma x + (1-\gamma) y$ for some $\gamma \in \boldsymbol{R}$.

Then $\mu_z = \gamma \mu_x + (1-\gamma)\mu_y$, so $4 = \gamma \cdot 1 + (1-\gamma) \cdot 5$, and $\gamma = \frac{1}{4}$.

$$\sigma_z^{\,2} = Var\,[\gamma\, x + (1-\gamma)\, y] = Var\left(\frac{1}{4}x + \frac{3}{4}y\right) = \left(\frac{1}{4}\right)^2 \cdot \sigma_x^2 + \left(\frac{3}{4}\right)^2 \cdot \sigma_y^2 + 2\cdot\frac{1}{4}\cdot\frac{3}{4}\cdot Cov(x,y) =$$

$=\left(\frac{1}{4}\right)^2 \cdot 9 + \left(\frac{3}{4}\right)^2 \cdot 4 + 2\cdot\frac{1}{4}\cdot\frac{3}{4}\cdot 2 = \frac{57}{16}$. This contradicts the fact that $\sigma_z^{\,2} = 1$. Hence not all three assets lie on the envelope. ☐

6.1. Exercise

1. In the world of assets A_1, A_2, A_3 and A_4 the following portfolios x and y are on the envelope:

$x = \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \\ 0.4 \end{bmatrix}$ and $y = \begin{bmatrix} 0.3 \\ 0.3 \\ 0.5 \\ -0.1 \end{bmatrix}$. Find the envelope.

Answer: $Env(A_1, A_2, A_3, A_4) = \left\{ \begin{bmatrix} -0.2\gamma + 0.3 \\ -0.1\gamma + 0.3 \\ -0.2\gamma + 0.5 \\ 0.5\gamma - 0.1 \end{bmatrix} \middle|\; \gamma \in R \right\}$.

6.2 Efficient Frontier

A portfolio is called an **efficient portfolio** if it has the highest expected return among all portfolios with the same variance.

An efficient portfolio maximizes expected return for a given risk.

The set of all efficient portfolios is called the **efficient frontier** of the assets $A_1, \dots, A_N$ and is denoted $EF(A_1, \dots, A_N)$.

6.3 Mean-Variance Relation

To each portfolio of assets $A_1,\dots, A_N$ we assign a pair (σ, μ) of its standard deviation and expected return. The set of all such pairs is called the **feasible set** of the assets $A_1, \dots, A_N$. The feasible set is represented by a figure on (σ, μ)-plane. There is no one-to-one correspondence between portfolios and points of the feasible set because two different portfolios can have equal means and equal standard deviations.

Theorem 6.1.

1) The envelope is represented on the (σ, μ)-plane by the right branch of the hyperbola:

$$\frac{\sigma^2}{A} - \frac{(\mu - \mu_0)^2}{B} = 1 \quad (\sigma > 0). \tag{13}$$

for some constants $A > 0$, $B > 0$ and μ_0.

2) For $N \geq 3$, the feasible region is the region to the right of the curve (13) including the curve itself.
3) The portfolio x_{min} with lowest risk corresponds to the vertex of the curve (13); it has the mean μ_0 and the variance A.
4) The efficient frontier is represented on the (σ, μ)-plane by the top half of the curve (13):

$$\begin{cases} \dfrac{\sigma^2}{A} - \dfrac{(\mu - \mu_0)^2}{B} = 1 \\ \sigma > 0 \\ \mu \geq \mu_0 \end{cases}.$$

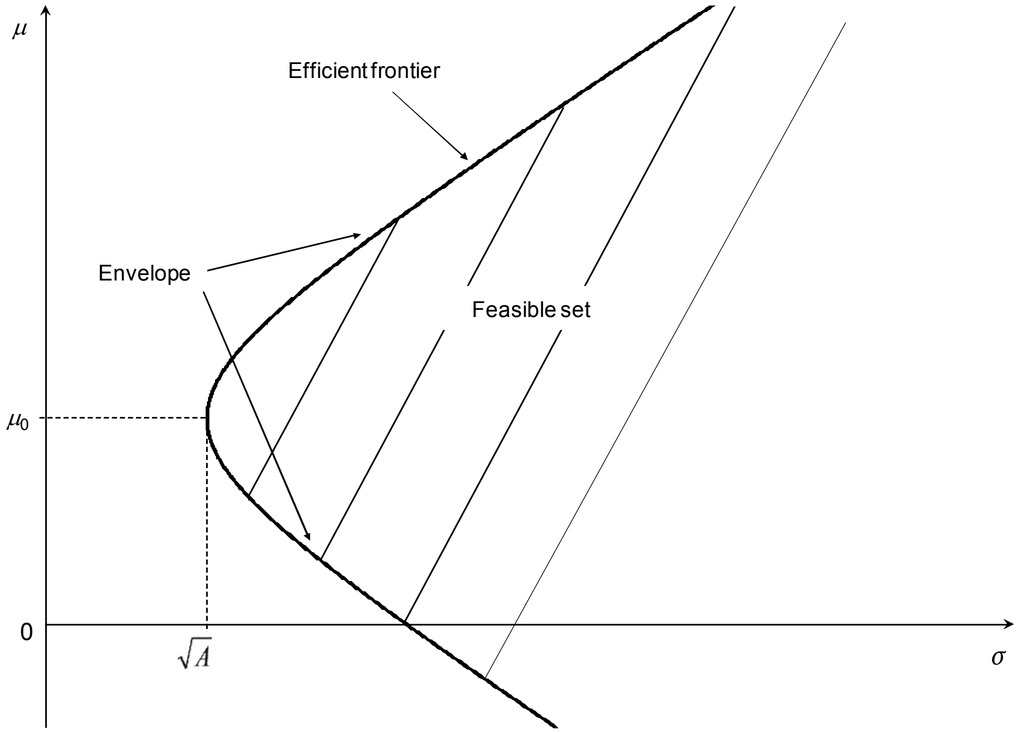


Figure 6.1.

Proof

1) Denote x_μ the envelope portfolio with expected return μ and denote σ^2 its variance. We will use Theorem 5.4.

The system (7) of linear equations from Section 5.5 can be solved using Gauss-Jordan elimination, so any solution q is a linear function of the parameter μ. From Theorem 5.4.2) it follows that the envelope portfolio x_μ is also a linear function of μ: $x_\mu = c + b\mu$, where c and b are some vectors independent of μ, $b \neq \mathbf{0}$.

So $\sigma^2 = Var(x_\mu) = x_\mu^T Sx_\mu = (c + b\mu)^T S (c + b\mu) = c^T S c + \mu^2 b^T Sb + \mu (c^T Sb + b^T Sc)$.

The relation between μ and σ for envelope portfolios is given by this second-degree equation:

$$\mu^2 b^T Sb + \mu (c^T Sb + b^T Sc) - \sigma^2 + c^T Sc = 0. \tag{14}$$

Comparing this with the general equation of a second-degree curve:

$$a_{11} u^2 + 2a_{12} uv + a_{22} v^2 + 2a_{13} u + 2a_{23} v + a_{33} = 0$$

we get $a_{11} = b^T S\, b$, $a_{22} = -1$, $a_{33} = c^T Sc$, $a_{13} = \frac{1}{2}(c^T Sb + b^T Sc)$, $a_{12} = a_{23} = 0$.

Consider two invariants I_2 and I_3.

$$I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}^2 = -b^T Sb.\ I_2 < 0, \text{ since the matrix } S \text{ is positive definite.}$$

$$I_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & a_{13} \\ 0 & -1 & 0 \\ a_{13} & 0 & a_{33} \end{vmatrix} = -\begin{vmatrix} a_{11} & a_{13} \\ a_{13} & a_{33} \end{vmatrix} = \frac{1}{4}\left[\left(c^T Sb + b^T Sc\right)^2 - 4\left(b^T Sb\right)\left(c^T Sc\right)\right].$$

Since the matrix S is positive definite, for any real number t: $(c + bt)^T\, S\,(c + bt) > 0$. Hence for any t,

$t^2\, b^T Sb + t\,(c^T Sb + b^T Sc) + c^T Sc > 0$. So the discriminant of this quadratic is negative:

$$(c^T Sb + b^T Sc)^2 - 4\,(b^T Sb)\,(c^T Sc) < 0.$$

This implies $I_3 < 0$. Also we have $I_2 < 0$, therefore the equation (14) defines a non-degenerate hyperbola with the real axis parallel to Oσ. This follows from properties of second-degree curves (see, for example, Il'in and Poznyak, 1985). The equation (14) can be transformed to a canonical form by completing the square for μ, so the result has the form (13).

2) Since the curve (13) represents the envelope portfolios, all other portfolios have higher variances, so their corresponding (σ, μ)-points lie to the right of the curve.

3) From the equation (13) we see that the right vertex of the hyperbola has coordinates $\left(\mu_0, \sqrt{A}\right)$; it corresponds to the portfolio with lowest risk.

4) As Figure 6.1 shows, each feasible value of σ (except the vertex value) corresponds to two points on the envelope and two values of μ. The point with a higher μ is on the efficient frontier. ☐

Example 6.3. Three assets have expected returns of 2%, 1% and 1%, respectively, and covariance matrix

$$S = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 6 \end{bmatrix}.$$

1) Find the envelope of these assets.
2) On the mean-variance plane find: a) the envelope and b) the feasible region.
3) Find the portfolio with lowest risk.
4) Find the efficient frontier of the assets and its (σ, μ)-representation.

Solution

1) $N = 3$ and $M = [2\ 1\ 1]$. The system (7) from Section 5.5 can be written as:

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + x_3 = \mu \end{cases} \text{ and } q = \begin{bmatrix} \mu - 1 \\ 2 \\ -\mu \end{bmatrix} \text{ is one of its solutions.}$$

This is the corresponding homogeneous system: $\begin{cases} x_1 + x_2 + x_3 = 0 \\ 2x_1 + x_2 + x_3 = 0 \end{cases}$ and one of its non-trivial solutions:

$v_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. By Theorem 5.4.2), $x_\mu = q - \dfrac{(q, v_1)}{(v_1, v_1)} v_1$.

$$(q, v_1) = Cov(q, v_1) = q^T S\, v_1 = \begin{bmatrix} \mu - 1 & 2 & -\mu \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = 4\mu - 2.$$

Similarly, $(v_1, v_1) = Var(v_1) = v_1^T S\, v_1 = 3$.

$$\text{So } x_\mu = q - \frac{(q, v_1)}{(v_1, v_1)} v_1 = q - \frac{4\mu - 2}{3} v_1 = \begin{bmatrix} \mu - 1 \\ 2 \\ -\mu \end{bmatrix} - \frac{4\mu - 2}{3} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

$$x_\mu = \frac{1}{3}\begin{bmatrix} 3\mu - 3 \\ -4\mu + 8 \\ \mu - 2 \end{bmatrix}. \tag{15}$$

$$Env(A_1, A_2, A_3) = \left\{ \frac{1}{3}\begin{bmatrix} 3\mu - 3 \\ -4\mu + 8 \\ \mu - 2 \end{bmatrix} \;\middle|\; \mu \in R \right\}.$$

2) The variance of x_μ is $\sigma^2 = x_\mu^{\ T} S\, x_\mu = \frac{2}{3}(\mu^2 - \mu + 1)$.

The envelope is represented on the mean-variance plane by the curve: $2(\mu^2 - \mu + 1) = 3\sigma^2$. After completing the square we have:

$$3\sigma^2 - 2\left(\mu - \frac{1}{2}\right)^2 = \frac{3}{2}.$$

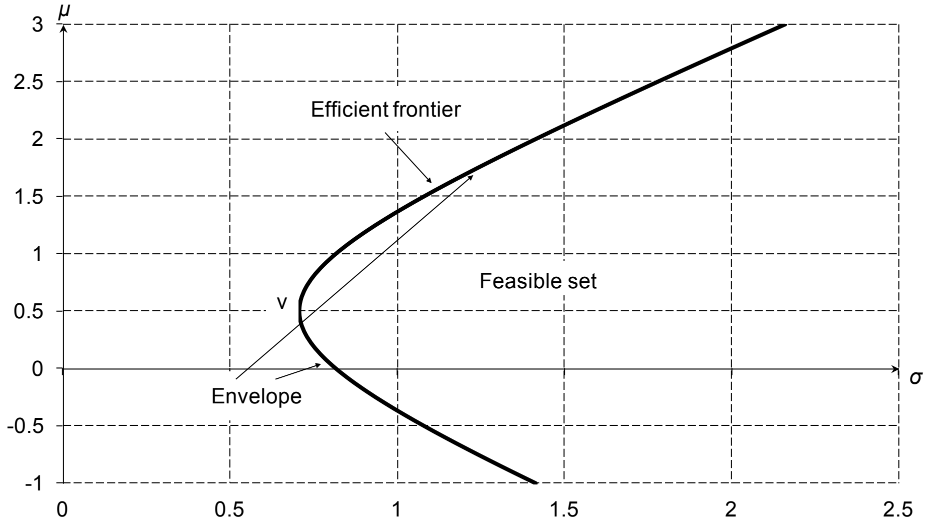
a) Changing this to canonical form we get this equation for the envelope:

$$\frac{\sigma^2}{0.5} - \frac{(\mu - 0.5)^2}{0.75} = 1 \;\; (\sigma > 0). \tag{16}$$

b) The feasible region is the region to the right of the curve (16) including the curve, so it is given by:

$$\sigma \geq \sqrt{\frac{2}{3}\left(\mu - \frac{1}{2}\right)^2 + \frac{1}{2}}\,.$$

3) The portfolio x_{min} with lowest risk corresponds to the right vertex **v** of the hyperbola (16), which has the coordinates $\left(\sqrt{0.5}, 0.5\right)$. So the parameters of x_{min} are $\sqrt{0.5}$ and 0.5. Substituting $\mu = 0.5$ into (15) we get $x_{min} = \begin{bmatrix} -0.5 \\ 2 \\ -0.5 \end{bmatrix}$, the portfolio with the lowest risk.



4) The efficient frontier is the top half of the curve (16). On the mean-variance plane it is given by:

$$\begin{cases} \dfrac{\sigma^2}{0.5} - \dfrac{(\mu-0.5)^2}{0.75} = 1 \\ \sigma > 0 \\ \mu \geq 0.5 \end{cases}. \text{ Also } EF(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3) = \left\{ \frac{1}{3}\begin{bmatrix} 3\mu - 3 \\ -4\mu + 8 \\ \mu - 2 \end{bmatrix} \middle| \mu \geq 0.5 \right\}. \ \square$$

Example 6.4. Four assets have expected returns of 1%, 1%, 2% and 1%, respectively, and covariance matrix $S = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{bmatrix}$.

1) Find the envelope of these assets.
2) On the mean-variance plane find: a) the envelope and b) the feasible region.
3) Find the portfolio with lowest risk.
4) Find the efficient frontier of the assets and its (σ, μ)-representation.

Solution

1) $N = 4$ and $M = [1\ 1\ 2\ 1]$. The system (7) from Section 5.5 can be written as:

$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ x_1 + x_2 + 2x_3 + x_4 = \mu \end{cases}$ and $q = \begin{bmatrix} 0 \\ 0 \\ \mu - 1 \\ 2 - \mu \end{bmatrix}$ is one of its solutions.

This is the corresponding homogeneous system: $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + 2x_3 + x_4 = 0 \end{cases}$ and one of its non-trivial solutions: $v_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$. Consider its another non-trivial solution $v_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ orthogonal to v_1.

Then $0 = (v_1, v_2)$ = [by Theorem 5.4.1)] = $Cov(v_1, v_2) = v_1^T S\, v_2 =$

$$= \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = -x_2 - x_3.$$

Thus, v_2 should satisfy these three equations: $\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + 2x_3 + x_4 = 0 \\ -x_2 - x_3 = 0 \end{cases}$, $v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$.

By Theorem 5.4.2), $x_\mu = q - \dfrac{(q, v_1)}{(v_1, v_1)} v_1 - \dfrac{(q, v_2)}{(v_2, v_2)} v_2$.

$(q, v_1) = Cov(q, v_1) = q^T S\, v_1 = \begin{bmatrix} 0 & 0 & \mu - 1 & 2 - \mu \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} = 1 - \mu.$

Similarly, $(q, v_2) = Cov(q, v_2) = q^T S v_2 = 3\mu - 6$; $(v_1, v_1) = Var(v_1) = v_1^T S v_1 = 1$;

$(v_2, v_2) = Var(v_2) = v_2^T S v_2 = 5$.

So $x_\mu = q - \dfrac{(q, v_1)}{(v_1, v_1)} v_1 - \dfrac{(q, v_2)}{(v_2, v_2)} v_2 = \begin{bmatrix} 0 \\ 0 \\ \mu - 1 \\ 2-\mu \end{bmatrix} + (\mu - 1)\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} - \dfrac{3\mu - 6}{5}\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$,

$$x_\mu = \frac{1}{5}\begin{bmatrix} 2\mu + 1 \\ -5\mu + 5 \\ 5\mu - 5 \\ -2\mu + 4 \end{bmatrix}. \qquad (17)$$

$$Env(A_1, A_2, A_3, A_4) = \left\{ \frac{1}{5}\begin{bmatrix} 2\mu + 1 \\ -5\mu + 5 \\ 5\mu - 5 \\ -2\mu + 4 \end{bmatrix} \;\middle|\; \mu \in R \right\}.$$

2) The variance of x_μ is $\sigma^2 = x_\mu^T S x_\mu = \frac{1}{5}(6\mu^2 - 14\mu + 9)$.

The envelope is represented on the mean-variance plane by the curve: $6\mu^2 - 14\mu + 9 = 5\sigma^2$. After completing the square we have:

$$5\sigma^2 - 6\left(\mu - \frac{7}{6}\right)^2 = \frac{5}{6}.$$

a) Changing this to canonical form we get this equation for the envelope:

$$\frac{\sigma^2}{\frac{1}{6}} - \frac{\left(\mu - \frac{7}{6}\right)^2}{\frac{5}{36}} = 1 \quad (\sigma > 0). \qquad (18)$$

b) The feasible region is the region to the right of the curve (18) including the curve, so it is given by:

$$\sigma \geq \sqrt{\frac{36\left(\mu - \frac{7}{6}\right)^2 + 5}{30}}.$$

3) The portfolio x_{min} with lowest risk corresponds to the right vertex of the hyperbola (18), which has the coordinates $\left(\frac{1}{\sqrt{6}}, \frac{7}{6}\right)$. So the parameters of x_{min} are $\frac{1}{\sqrt{6}}$ and $\frac{7}{6}$.

Substituting $\mu = \dfrac{7}{6}$ into (17) we get $x_{min} = \dfrac{1}{6}\begin{bmatrix} 4 \\ -1 \\ 1 \\ 2 \end{bmatrix}$, the portfolio with the lowest risk.

4) The efficient frontier is the top half of the curve (18). On the mean-variance plane it is given by:

$$\begin{cases} \dfrac{\sigma^2}{\frac{1}{6}} - \dfrac{\left(\mu - \frac{7}{6}\right)^2}{\frac{5}{36}} = 1 \\ \sigma > 0 \\ \mu \geq \dfrac{7}{6} \end{cases}$$

Also $EF(A_1, A_2, A_3, A_4) = \left\{ \dfrac{1}{5}\begin{bmatrix} 2\mu + 1 \\ -5\mu + 5 \\ 5\mu - 5 \\ -2\mu + 4 \end{bmatrix} \middle| \mu \geq \dfrac{7}{6} \right\}$. □

6.3. Exercises

Consider assets with given expected returns and covariance matrix *S*. In each case find the following.

1) Find the envelope of these assets.
2) On the mean-variance plane find: a) the envelope and b) the feasible region.
3) Find the portfolio with lowest risk.
4) Find the efficient frontier of the assets and its (σ, μ)-representation.

1. Expected returns: 2%, 2% and 4%, $S = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 3 \end{bmatrix}$.

Answers:

1) $Env(A_1, A_2, A_3) = \left\{ \frac{1}{2}\begin{bmatrix} -\mu+5 \\ -1 \\ \mu-2 \end{bmatrix} \middle| \mu \in R \right\}$; 2a) $\frac{\sigma^2}{0.5} - \frac{(\mu-2)^2}{1} = 1 \ (\sigma > 0)$,

2b) $\sigma \geq \sqrt{0.5 + 0.5\,(\mu-2)^2}$; 3) $x_{min} = \begin{bmatrix} 1.5 \\ -0.5 \\ 0 \end{bmatrix}$;

4) $EF(A_1, A_2, A_3) = \left\{ \frac{1}{2}\begin{bmatrix} -\mu+5 \\ -1 \\ \mu-2 \end{bmatrix} \middle| \mu \geq 2 \right\}$, its ($\sigma$, μ)-representation is $\begin{cases} \frac{\sigma^2}{0.5} - \frac{(\mu-2)^2}{1} = 1 \\ \sigma > 0 \\ \mu \geq 2 \end{cases}$.

2. Expected returns: 4%, 2% and 4%, $S = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 3 & 2 \\ -1 & 2 & 2 \end{bmatrix}$.

Answers:

1) $Env(A_1, A_2, A_3) = \left\{ \begin{bmatrix} 0.6 \\ -0.5\mu+2 \\ 0.5\mu-1.6 \end{bmatrix} \middle| \mu \in R \right\}$; 2a) $\frac{\sigma^2}{0.2} - \frac{(\mu-4)^2}{0.8} = 1 \ (\sigma > 0)$,

2b) $\sigma \geq \sqrt{0.2 + 0.25(\mu-4)^2}$; 3) $x_{min} = \begin{bmatrix} 0.6 \\ 0 \\ 0.4 \end{bmatrix}$;

4) $EF(A_1, A_2, A_3) = \left\{ \begin{bmatrix} 0.6 \\ -0.5\mu+2 \\ 0.5\mu-1.6 \end{bmatrix} \middle| \mu \geq 4 \right\}$, its ($\sigma$, μ)-representation is $\begin{cases} \frac{\sigma^2}{0.2} - \frac{(\mu-4)^2}{0.8} = 1 \\ \sigma > 0 \\ \mu \geq 4 \end{cases}$.

3. Expected returns: 4%, 2% and 0%, $S=\begin{bmatrix}2&0&1\\0&1&1\\1&1&2\end{bmatrix}$.

Answers:

1) $Env(A_1, A_2, A_3) = \left\{ \frac{1}{6}\begin{bmatrix}\mu-1\\\mu+2\\-2\mu+5\end{bmatrix} \middle| \mu\in R \right\}$; 2a) $\frac{\sigma^2}{0.5} - \frac{(\mu-4)^2}{6} = 1 \ (\sigma>0)$,

2b) $\sigma \geq \sqrt{0.5+\frac{(\mu-4)^2}{12}}$; 3) $x_{min} = \begin{bmatrix}0.5\\1\\-0.5\end{bmatrix}$;

4) $EF(A_1, A_2, A_3) = \left\{ \frac{1}{6}\begin{bmatrix}\mu-1\\\mu+2\\-2\mu+5\end{bmatrix} \middle| \mu\geq 4 \right\}$, its (σ, μ)-representation is $\begin{cases}\frac{\sigma^2}{0.5} - \frac{(\mu-4)^2}{6} = 1\\ \sigma>0\\ \mu\geq 4\end{cases}$.

4. Expected returns: 2%, 1% and 2%, $S=\begin{bmatrix}2&1&1\\1&1&-1\\1&-1&6\end{bmatrix}$.

Answers:

1) $Env(A_1, A_2, A_3) = \left\{ \frac{1}{6}\begin{bmatrix}7\mu-9\\-6\mu+12\\-\mu+3\end{bmatrix} \middle| \mu\in R \right\}$; 2a) $\frac{\sigma^2}{0.2} - \frac{(\mu-0.6)^2}{0.24} = 1 \ (\sigma>0)$,

2b) $\sigma \geq \sqrt{0.2+\frac{5(\mu-0.6)^2}{6}}$; 3) $x_{min} = \begin{bmatrix}-0.8\\1.4\\0.4\end{bmatrix}$;

4) $EF(A_1, A_2, A_3) = \left\{ \frac{1}{6}\begin{bmatrix}7\mu-9\\-6\mu+12\\-\mu+3\end{bmatrix} \middle| \mu\geq 0.6 \right\}$, its (σ, μ)-representation is $\begin{cases}\frac{\sigma^2}{0.2} - \frac{(\mu-0.6)^2}{0.24} = 1\\ \sigma>0\\ \mu\geq 0.6\end{cases}$.

6.4 Mean-Variance Relation for 2-Asset Portfolios

Previously we assumed that at least two values of the expectations $\mu_1,\ldots, \mu_N$ are different. In case $N=2$ this means $\mu_1 \neq \mu_2$. We will keep this assumption but not the assumption $det\ S > 0$. For $N = 2$ we will consider the case $det\ S > 0$ and then the case $det\ S = 0$. We will assume $\sigma_1 > 0$ and $\sigma_2 > 0$, which means the returns r_1 and r_2 of the assets A_1 and A_2 are not constant.

Denote ρ the coefficient of correlation between the returns of assets A_1 and A_2. As stated in Section 3.3, $det\ S = 0 \Leftrightarrow r_1$ and r_2 are linearly dependent $\Leftrightarrow |\ Cov(r_1, r_2)\ | = \sigma_1\ \sigma_2 \Leftrightarrow |\ \rho\ | = 1$. Thus,

$$det\ S = 0 \Leftrightarrow |\ \rho\ | = 1. \tag{19}$$

In case $N = 2$ for any $\mu \in \boldsymbol{R}$ there is only one portfolio with expected return μ, so the envelope is the same as the feasible set.

The feasible set of the assets A_1 and A_2 depends on the value of ρ as the following theorems show.

Theorem 6.2. Suppose $N = 2$, $\mu_1 \neq \mu_2$ and $\rho \neq \pm 1$. The feasible set and envelope of A_1, A_2 are both represented on the (σ, μ)-plane by the right branch of the hyperbola:

$$\frac{\sigma^2}{A} - \frac{(\mu - \mu_0)^2}{B} = 1 \quad (\sigma > 0).$$

for some constants $A > 0$, $B > 0$ and μ_0.

Click on the ad to read more

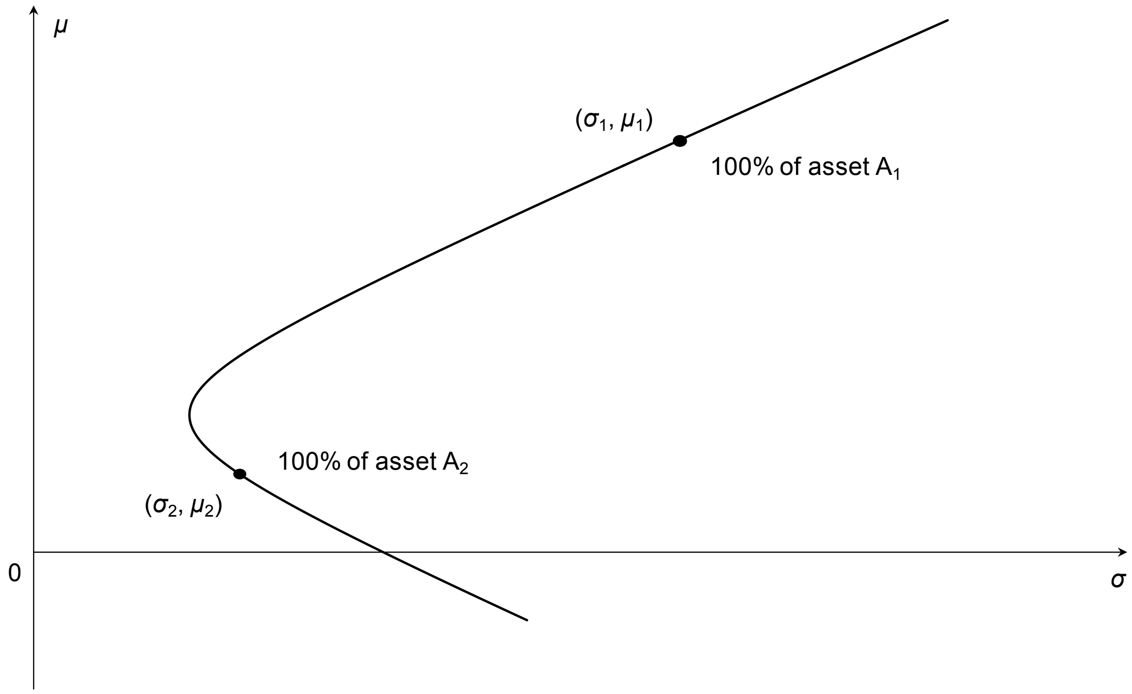


Figure 6.2 Feasible set of two assets ($\rho \neq \pm 1$)

Proof

The proof is similar to the proof of Theorem 6.1, since the feasible set and envelope both equal the set of solutions x_μ of the system $\begin{cases} Ux = 1 \\ Mx = \mu \end{cases}$ and each solution x_μ can be written in the form $x_\mu = c + b\mu$. ☐

Theorem 6.3. Suppose $N = 2$, $\mu_1 \neq \mu_2$ and $\rho = -1$.

The feasible set and envelope of A_1, A_2 are both represented on the (σ, μ)-plane by two rays coming out of one point on the vertical μ-axis.

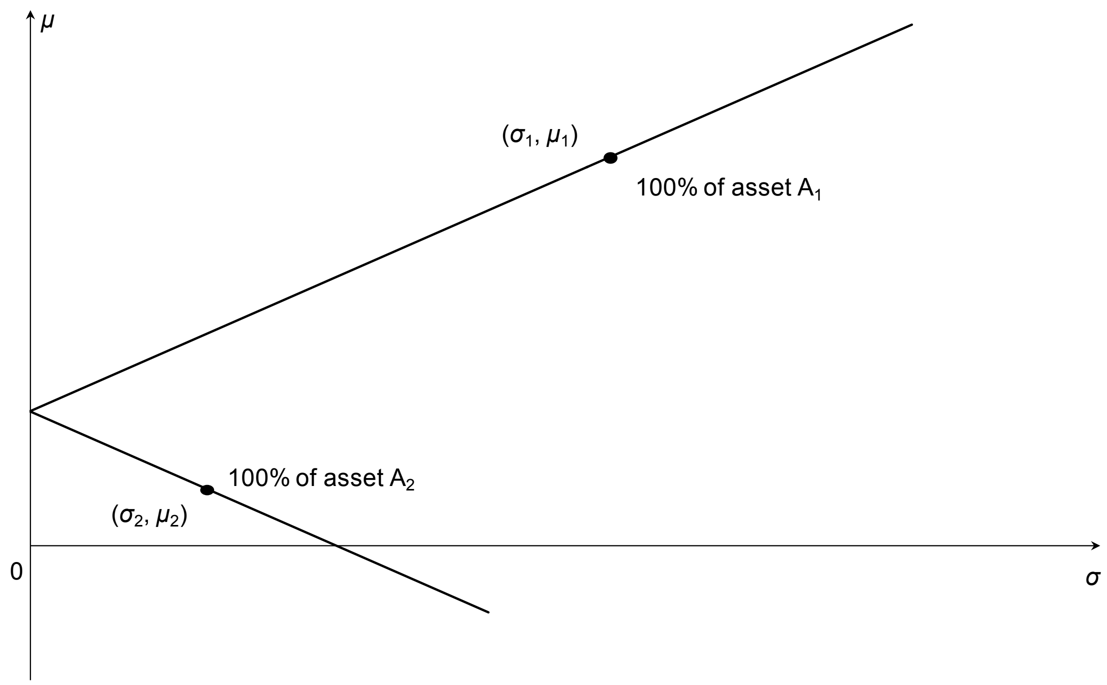


Figure 6.3 Feasible set of two assets ($\rho = -1$)

Proof

Any 2-asset portfolio can be represented in the form $x = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}$. Denote $\mu = \mu_x$ and $\sigma = \sigma_x$.

$x = \gamma\, r_1 + (1-\gamma)\, r_2$, so $\begin{cases} \mu = \gamma\,\mu_1 + (1-\gamma)\mu_2 \\ \sigma^2 = \gamma^2\sigma_1^2 + (1-\gamma)^2\sigma_2^2 + 2\gamma(1-\gamma)\sigma_{12} \end{cases}$.

Since $\rho = -1$, we have $\sigma_{12} = \rho\sigma_1\,\sigma_2 = -\,\sigma_1\,\sigma_2$ and $\sigma^2 = \gamma^2\sigma_1^{\,2} + (1-\gamma)^2\sigma_2^{\,2} - 2\gamma(1-\gamma)\,\sigma_1\,\sigma_2 = \left[\gamma\sigma_1 - (1-\gamma)\sigma_2\right]^2$.

Standard deviation $\sigma \geq 0$, so $\begin{cases} \mu = \gamma\,\mu_1 + (1-\gamma)\mu_2 & I \\ \sigma = \left|\gamma\sigma_1 - (1-\gamma)\sigma_2\right| & II \end{cases}$

From equation I: $\gamma = \dfrac{\mu - \mu_2}{\mu_1 - \mu_2}$ and $1-\gamma = \dfrac{\mu_1 - \mu}{\mu_1 - \mu_2}$. Substitute this into equation II:

$$\sigma = \left|\frac{(\mu-\mu_2)\sigma_1 - (\mu_1 - \mu)\sigma_2}{\mu_1 - \mu_2}\right| = \left|\frac{\mu(\sigma_1 + \sigma_2) - \mu_2\sigma_1 - \mu_1\sigma_2}{\mu_1 - \mu_2}\right| = \frac{\sigma_1 + \sigma_2}{\left|\mu_1 - \mu_2\right|}\cdot\left|\mu - \frac{\mu_2\sigma_1 + \mu_1\sigma_2}{\sigma_1 + \sigma_2}\right|.$$

Denote $k = \dfrac{\sigma_1 + \sigma_2}{|\mu_1 - \mu_2|} > 0$ and $\mu_0 = \dfrac{\mu_2\sigma_1 + \mu_1\sigma_2}{\sigma_1 + \sigma_2}$. Then we have the equation $\sigma = k\,|\mu - \mu_0|$, which defines two rays coming out of the point $(0, \mu_0)$ on the vertical μ-axis. □

> **Theorem 6.4.** Suppose $N = 2$, $\rho = 1$, $\mu_1 \neq \mu_2$ and $\sigma_1 \neq \sigma_2$.
>
> The feasible set and envelope of A_1, A_2 are both represented on the (σ, μ)-plane by two rays coming out of one point on the vertical μ-axis.

Proof

The diagram looks similar to the diagram in the previous theorem.

Since $\rho = 1$, we have $\sigma_{12} = \rho\sigma_1\,\sigma_2 = \sigma_1\,\sigma_2$. As in the proof of Theorem 6.3, for any portfolio $x = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}$ we have $\sigma^2 = \gamma^2\sigma_1^2 + (1-\gamma)^2\sigma_2^2 + 2\,(1-\gamma)\,\sigma_1\,\sigma_2 = [\gamma\sigma_1 + (1-\gamma)\sigma_2]^2$.

$$\text{So} \begin{cases} \mu = \gamma\mu_1 + (1-\gamma)\mu_2 & I \\ \sigma = |\gamma\sigma_1 + (1-\gamma)\sigma_2| & II \end{cases}$$

Click on the ad to read more

Substitute $\gamma = \dfrac{\mu - \mu_2}{\mu_1 - \mu_2}$ and $1 - \gamma = \dfrac{\mu_1 - \mu}{\mu_1 - \mu_2}$ into equation *II*:

$$\sigma = \left| \frac{(\mu - \mu_2)\sigma_1 + (\mu_1 - \mu)\sigma_2}{\mu_1 - \mu_2} \right| = \left| \frac{\mu(\sigma_1 - \sigma_2) - \mu_2\sigma_1 + \mu_1\sigma_2}{\mu_1 - \mu_2} \right| = \left| \frac{\sigma_1 - \sigma_2}{\mu_1 - \mu_2} \right| \cdot \left| \mu - \frac{\mu_2\sigma_1 - \mu_1\sigma_2}{\sigma_1 - \sigma_2} \right|.$$

Denote $k = \left| \dfrac{\sigma_1 - \sigma_2}{\mu_1 - \mu_2} \right| > 0$ and $\mu_0 = \dfrac{\mu_2\sigma_1 - \mu_1\sigma_2}{\sigma_1 - \sigma_2}$. Then we have the equation $\sigma = k\,|\mu - \mu_0|$, which defines two rays coming out of the point $(0, \mu_0)$ on the vertical μ-axis. ☐

Theorem 6.5. Suppose $N = 2$, $\rho = 1$, $\mu_1 \neq \mu_2$ and $\sigma_1 = \sigma_2$.

The feasible set and envelope of A_1, A_2 are both represented on the (σ, μ)-plane by the vertical straight line given by $\sigma = \sigma_1$.

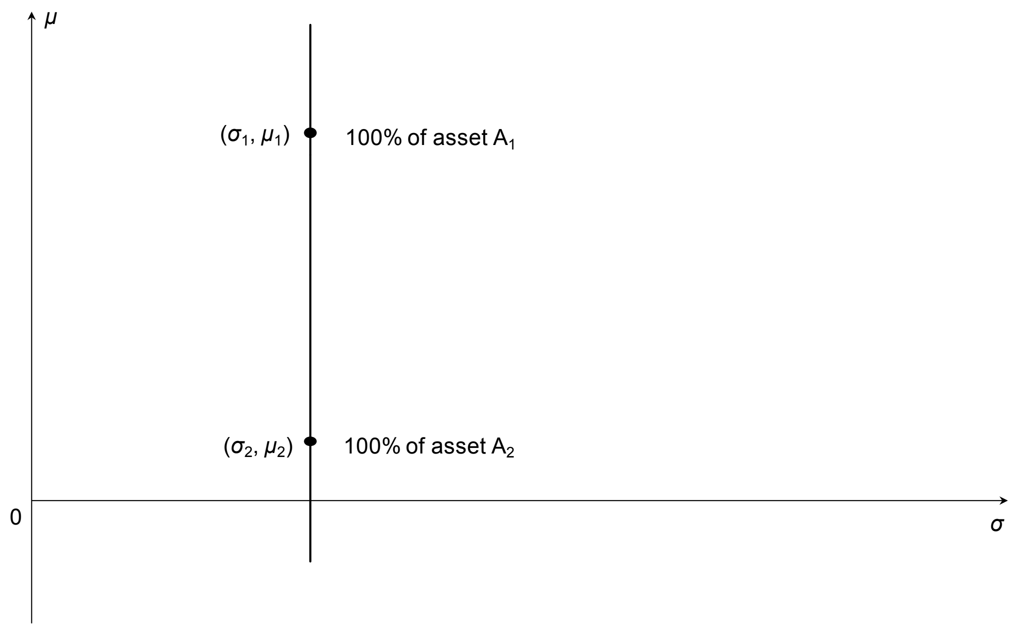


Figure 6.4 Feasible set of two assets ($\rho = 1$ & $\sigma_1 = \sigma_2$)

Proof

$\sigma_{12} = \rho\sigma_1\,\sigma_2 = \sigma_1^2$. As in the proof of Theorem 6.4, for any portfolio $x = \begin{bmatrix} \gamma \\ 1-\gamma \end{bmatrix}$ we have
$\sigma^2 = \gamma^2\sigma_1^2 + (1-\gamma)^2\sigma_2^2 + 2\gamma(1-\gamma)\,\sigma_1\,\sigma_2 = \sigma_1^2\,(\gamma^2 + (1-\gamma)^2 + 2\gamma(1-\gamma)\,) = \sigma_1^2\,(\gamma + 1-\gamma)^2 = \sigma_1^2$.

So $\begin{cases} \mu = \gamma\mu_1 + (1-\gamma)\mu_2 \\ \sigma = \sigma_1 \end{cases}$, which defines the vertical straight line $\sigma = \sigma_1$. ☐

Example 6.5. Asset A_1 has expected return of 3 and standard deviation of 2. Asset A_2 has expected return of 1 and standard deviation of 1. For the following values of correlation coefficient ρ find the feasible sets of assets A_1, A_2 and sketch them on the same graph.

1) $\rho = -1$, 2) $\rho = 0$, 3) $\rho = 1$.

Solution

$\mu_1 = 3$, $\sigma_1 = 2$, $\mu_2 = 1$, $\sigma_2 = 1$. For any portfolio x:

$\mu = \gamma\mu_1 + (1-\gamma)\mu_2$, $\mu = 3\gamma + (1-\gamma)$, $\gamma = \dfrac{\mu-1}{2}$, $1-\gamma = \dfrac{3-\mu}{2}$; $\sigma^2 = \gamma^2\sigma_1^{\,2} + (1-\gamma)^2\sigma_1^{\,2} + 2\gamma(1-\gamma)\,\sigma_{12}$.

1) $\rho = -1$. $\sigma_{12} = \rho\sigma_1\sigma_2 = -2$. So $\sigma^2 = 4\gamma^2 + (1-\gamma)^2 - 4\gamma(1-\gamma) = \dfrac{4(\mu-1)^2 + (3-\mu)^2 - 4(\mu-1)(3-\mu)}{4} = \left(\dfrac{3\mu-5}{2}\right)^2$,

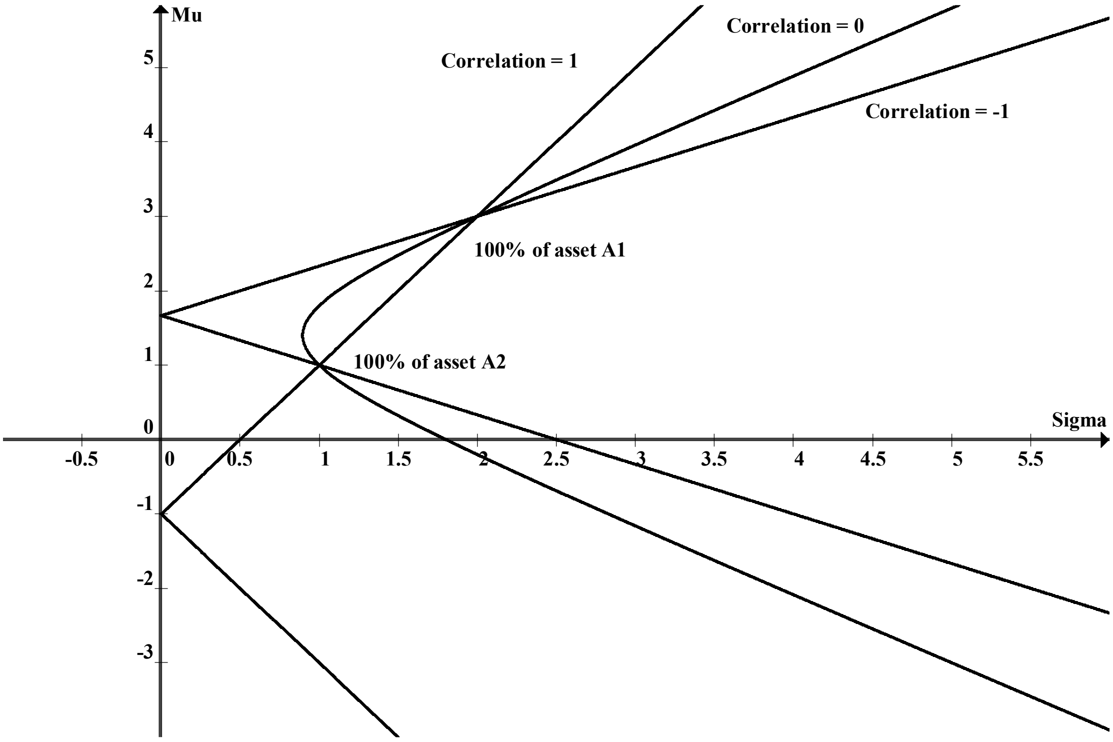
$\sigma = \dfrac{3}{2}\left|\mu - \dfrac{5}{3}\right|$, which defines two rays coming out of the point $\left(0, \dfrac{5}{3}\right)$.

2) $\rho = 0$. $\sigma_{12} = \rho\sigma_1\sigma_2 = 0$. $\sigma^2 = 4\gamma^2 + (1-\gamma)^2 = \dfrac{4(\mu-1)^2 + (3-\mu)^2}{4} = \dfrac{5\mu^2 - 14\mu + 13}{4} = \dfrac{5}{4}\left(\mu - \dfrac{7}{5}\right)^2 + \dfrac{4}{5}$.

So $\sigma^2 = \dfrac{5}{4}\left(\mu - \dfrac{7}{5}\right)^2 + \dfrac{4}{5}$ and $\dfrac{\sigma^2}{0.8} - \dfrac{(\mu-1.4)^2}{0.64} = 1\ (\sigma > 0)$, which defines the right branch of the hyperbola with centre (0, 1.4) and vertex ($\sqrt{0.8}$, 1.4).

3) $\rho = 1$. $\sigma_{12} = \rho\sigma_1\sigma_2 = 2$. $\sigma^2 = 4\gamma^2 + (1-\gamma)^2 + 4\gamma(1-\gamma) = \dfrac{4(\mu-1)^2 + (3-\mu)^2 + 4(\mu-1)(3-\mu)}{4} = \left(\dfrac{\mu+1}{2}\right)^2$,

$\sigma = \dfrac{1}{2}|\mu+1|$, which defines two rays coming out of the point (0, –1).


Mu
Correlation = 1
Correlation = 0
Correlation = -1
100% of asset A1
100% of asset A2
Sigma
5
4
3
2
1
0
-1
-2
-3
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5

□

Example 6.6. Two assets have expected returns of 1% and 3%, respectively, and covariance matrix $S = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$.

1) Find the envelope of these assets.
2) On the mean-variance plane find the envelope and feasible region.
3) Find the portfolio with lowest risk.
4) Find the efficient frontier of the assets and its (σ, μ)-representation.

Solution

1) $N = 2$ and $M = [1\ 3]$. The system (7) from Section 5.5 can be written as:

$\begin{cases} x_1 + x_2 = 1 \\ x_1 + 3x_2 = \mu \end{cases}$ and this is its only solution:

$$x_\mu = \frac{1}{2}\begin{bmatrix} -\mu + 3 \\ \mu - 1 \end{bmatrix}. \qquad (20)$$

$Env(A_1, A_2) = \left\{ \frac{1}{2}\begin{bmatrix} -\mu + 3 \\ \mu - 1 \end{bmatrix} \ \middle|\ \mu \in R \right\}$.

2) The variance of x_μ is $\sigma^2 = x_\mu^{\ T} S\, x_\mu = \frac{1}{4}\left(10\mu^2 - 32\mu + 26\right)$.

The envelope is represented on the mean-variance plane by the curve: $10\mu^2 - 32\mu + 32 = 4\sigma^2$. After completing the square we have: $4\sigma^2 - 10(\mu - 1.6)^2 = 0.4$.

Changing this to canonical form we get this equation for the envelope and the feasible region:

$$\frac{\sigma^2}{0.1} - \frac{(\mu - 1.6)^2}{0.04} = 1 \ \ (\sigma > 0). \qquad (21)$$

3) The portfolio x_{min} with lowest risk corresponds to the right vertex of the hyperbola (21), which has the coordinates $\left(\sqrt{0.1}, 1.6\right)$. So the parameters of x_{min} are $\sqrt{0.1}$ and 1.6. Substituting $\mu = 1.6$ into (20) we get $x_{min} = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$, the portfolio with lowest risk.

4) The efficient frontier is the top half of the curve (21). On the mean-variance plane it is given by:

$$\begin{cases} \dfrac{\sigma^2}{0.1} - \dfrac{(\mu - 1.6)^2}{0.04} = 1 \\ \sigma > 0 \\ \mu \geq 1.6 \end{cases}.$$

Also $EF(A_1, A_2, A_3, A_4) = \left\{ \frac{1}{2}\begin{bmatrix} -\mu+3 \\ \mu-1 \end{bmatrix} \;\middle|\; \mu \geq 1.6 \right\}$. □

6.4 Exercises

1. Asset A_1 has expected return of 12 and standard deviation of 1. Asset A_2 has expected return of 15 and standard deviation of 2. For the following values of correlation coefficient ρ find the feasible sets of assets A_1, A_2 and sketch them on the same graph.

1) $\rho = -1$, 2) $\rho = -\frac{1}{4}$, 3) $\rho = 0$, 4) $\rho = \frac{1}{2}$, 5) $\rho = 1$.

Answers: 1) $\sigma = |\mu - 13|$, 2) $\frac{\sigma^2}{\frac{5}{8}} - \frac{\left(\mu - \frac{51}{4}\right)^2}{\frac{15}{16}} = 1 \ (\sigma > 0)$, 3) $\frac{\sigma^2}{0.8} - \frac{(\mu - 12.6)^2}{1.44} = 1 \ (\sigma > 0)$,

4) $\frac{\sigma^2}{1} - \frac{(\mu - 12)^2}{3} = 1 \ (\sigma > 0)$, 5) $\sigma = \frac{1}{3}|\mu - 9|$.

2. Asset A_1 has expected return of 10 and standard deviation of 2. Asset A_2 has expected return of 12 and standard deviation of 2. For the following values of correlation coefficient ρ find the feasible sets of assets A_1, A_2 and sketch them on the same graph.

1) $\rho = -1$, 2) $\rho = -0.6$, 3) $\rho = 0$, 4) $\rho = 0.6$, 5) $\rho = 1$.

Answers: 1) $\sigma = 2\,|\mu - 11|$, 2) $\frac{\sigma^2}{0.8} - \frac{(\mu - 11)^2}{0.25} = 1 \ (\sigma > 0)$, 3) $\frac{\sigma^2}{2} - \frac{(\mu - 11)^2}{1} = 1 \ (\sigma > 0)$,

4) $\frac{\sigma^2}{3.2} - \frac{(\mu - 11)^2}{4} = 1 \ (\sigma > 0)$, 5) $\sigma = 2$.

3. Two assets have expected returns of 20% and 10%, respectively, and covariance matrix $S = \begin{bmatrix} 6 & 4 \\ 4 & 3 \end{bmatrix}$.

1) Find the envelope of these assets.
2) On the mean-variance plane find the envelope and feasible region.
3) Find the portfolio with lowest risk.
4) Find the efficient frontier of the assets and its (σ, μ)-representation.

Answers: 1) $\left\{ \begin{bmatrix} 0.1\mu - 1 \\ -0.1\mu + 2 \end{bmatrix} \;\middle|\; \mu \in R \right\}$, 2) $\frac{\sigma^2}{2} - \frac{\mu^2}{200} = 1 \ (\sigma > 0)$, 3) $x_{min} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, 4) $\begin{cases} \frac{\sigma^2}{2} - \frac{\mu^2}{200} = 1 \\ \sigma > 0 \\ \mu \geq 0 \end{cases}$.

7 Capital Market Theory

7.1 Risk-Free Asset

A **risk-free asset** is an asset with a constant return; it is denoted F.

The return of the asset F is denoted f.

For any portfolio x the difference $(\mu_x - f)$ is called the **risk premium** of x.

Thus, $E(f) = f$, $Var(f) = 0$.

For example, government bonds are considered risk-free assets because the government can always raise taxes to pay its bills.

As in the previous chapters, we will consider portfolios of assets $A_1,\ldots, A_N$, which are represented by vectors in the N-dimensional space K with the basis $r_1,\ldots, r_N$ of the asset returns. Portfolios of assets $A_1,\ldots, A_N$, F are represented by vectors in an $(N+1)$-dimensional space, where the asset returns $r_1,\ldots, r_N$, f make a basis. In this basis $f = \begin{bmatrix} 0 \\ \ldots \\ 0 \\ 1 \end{bmatrix}$.

Lemma 7.1. $f \in Env(A_1, A_2,\ldots, A_N, F)$.

Proof

$\sigma_f = 0$ and 0 is the lowest possible value for standard deviation, hence f is on the envelope. □

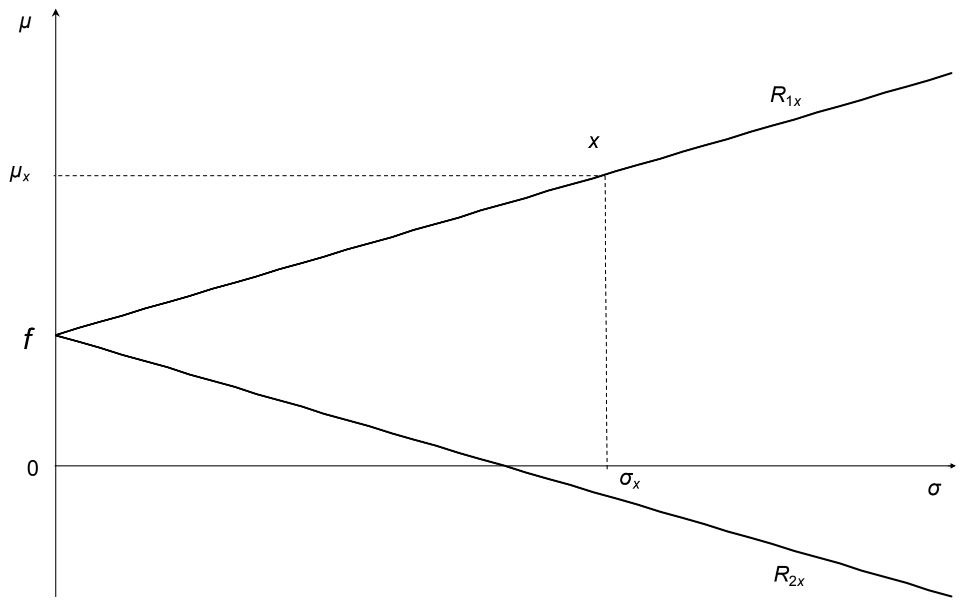
For any portfolio x of the assets $A_1,\ldots, A_N$ define two rays R_{1x} and R_{2x} on the (σ, μ)-plane starting at point $(0, f)$: R_{1x} is passing through point (σ_x, μ_x) and R_{2x} is passing through point $(\sigma_x, 2f-\mu_x)$.

Clearly, they are defined by the following equations,

$$R_{1x}:\ \mu = f + \frac{\mu_x - f}{\sigma_x}\sigma\ (\sigma \ge 0),\quad R_{2x}:\ \mu = f - \frac{\mu_x - f}{\sigma_x}\sigma\ (\sigma \ge 0). \tag{22}$$

Lemma 7.2. 1) x is a portfolio of $A_1,\ldots, A_N$ and $y = \lambda x + (1-\lambda)f$ for some $\lambda \in \mathbf{R}$ $\Leftrightarrow$ $(\sigma_y, \mu_y) \in R_{1x} \cup R_{2x}$.

2) The feasible region of $A_1,\ldots, A_N$, F equals $\{R_{1x} \cup R_{2x} \mid x$ is a portfolio of $A_1,\ldots, A_N\}$.



Proof

1) $\Rightarrow$ Suppose x is a portfolio of $A_1,\dots, A_N$ and $y = \lambda x + (1-\lambda)f$. Denote $\sigma = \sigma_y$ and $\mu = \mu_y$. Then

$$\mu = \lambda\mu_x + (1-\lambda)f,\ \ \sigma^2 = \lambda^2\sigma_x^2 \ \text{ and } \sigma = |\lambda|\ \sigma_x.$$

We will consider 2 cases: $\lambda \geq 0$ and $\lambda < 0$.

a) $\lambda \geq 0$.

Then $\sigma = \lambda\sigma_x$, $\lambda = \dfrac{\sigma}{\sigma_x}$. Substitute: $\mu = \dfrac{\sigma}{\sigma_x}\mu_x + (1-\dfrac{\sigma}{\sigma_x})f$, $\mu - f = \dfrac{\sigma}{\sigma_x}(\mu_x - f)$, $\dfrac{\mu - f}{\mu_x - f} = \dfrac{\sigma}{\sigma_x}$ $(\sigma \geq 0)$.

This defines the ray R_{1x}.

b) $\lambda < 0$.

Then $\sigma = -\lambda\sigma_x$, $\lambda = -\dfrac{\sigma}{\sigma_x}$. Substitute: $\mu = -\dfrac{\sigma}{\sigma_x}\mu_x + (1+\dfrac{\sigma}{\sigma_x})f$, $\mu - f = -\dfrac{\sigma}{\sigma_x}(\mu_x - f)$, $\dfrac{\mu - f}{\mu_x - f} = -\dfrac{\sigma}{\sigma_x}$ $(\sigma \geq 0)$.

This defines the ray R_{2x}.

$\Leftarrow$ This part is proven by reversing calculations in the previous part with $\lambda = \pm\dfrac{\sigma}{\sigma_x}$.

2) Due to (1), it is sufficient to prove:

y is a portfolio of $A_1,\dots, A_N$, F $\Leftrightarrow$ $y = \lambda x + (1-\lambda)f$ for some $\lambda \in \boldsymbol{R}$ and portfolio x of $A_1,\dots, A_N$ (23)

Proof of (23)

$\Rightarrow$ Suppose y is a portfolio of $A_1,\dots, A_N$, F. Then $y = kf + y_1r_1 + \dots + y_N r_N$, where $k + y_1 + \dots + y_N = 1$.

So $y_1 + \dots + y_N = 1-k$. Denote $\lambda = 1-k$ and $x = \dfrac{1}{\lambda}\left(y_1r_1 + \dots + y_Nr_N\right)$. Then x is a portfolio of $A_1,\dots, A_N$ and $\lambda x + (1-\lambda)f = y$.

$\Leftarrow$ Suppose $y = \lambda x + (1-\lambda)f$ and x is a portfolio of $A_1,\dots, A_N$. Then $x = x_1r_1 + \dots + x_N r_N$ and $\sum_{k=1}^{N} x_k = 1$.

So $y = \lambda x_1r_1 + \dots + \lambda x_N r_N + (1-\lambda)f$ and $\lambda x_1 + \dots + \lambda x_N + (1-\lambda) = \lambda\sum_{k=1}^{N} x_k + (1-\lambda) = \lambda + (1-\lambda) = 1$. Therefore y is a portfolio of $A_1,\dots, A_N$, F. ☐

7.2 Tangency Portfolio

Lemma 7.3. Consider the right branch of the hyperbola representing *Env*($A_1,\dots, A_N$), as in Theorem 6.1:

$$\frac{\sigma^2}{A} - \frac{(\mu - \mu_0)^2}{B} = 1 \quad (\sigma > 0).$$

Suppose $f \neq \mu_0$. Then there is one tangent line to this curve from point $(0, f)$ and it is given by the equation:

$$\mu = f + \frac{\mu_t - f}{\sigma_t}\sigma,$$

where (σ_t, μ_t) are the coordinates of tangency point t:

$$\sigma_t = \sqrt{A + \frac{AB}{(\mu_0 - f)^2}}, \quad \mu_t = \mu_0 + \frac{B}{\mu_0 - f}.$$

Proof

For a standard hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ the tangent line through tangency point (X, Y) is given by the equation: $\frac{Xx}{a^2} - \frac{Yy}{b^2} = 1$ (see, for example, Il'in and Poznyak, 1985). So the tangent line to our hyperbola through tangency point (σ_t, μ_t) has the form:

$$\frac{\sigma_t}{A}\sigma - \frac{\mu_t - \mu_0}{B}(\mu - \mu_0) = 1.$$

Click on the ad to read more

For point $(0, f)$ we have: $\sigma = 0$ and $\mu = f$:

$$-\frac{\mu_t - \mu_0}{B}(f - \mu_0) = 1, \quad \mu_t - \mu_0 = \frac{B}{\mu_0 - f}, \quad \mu_t = \mu_0 + \frac{B}{\mu_0 - f}.$$

Substituting this into the equation of the hyperbola we find σ_t. ☐

Since for each value of expected return there is exactly one envelope portfolio, the tangency point t from Lemma 7.3 corresponds to a unique portfolio in $Env(A_1, \ldots, A_N)$; this portfolio will also be denoted t.

> **Theorem 7.1.** Suppose $f \neq \mu_0$ and t is the tangency point from Lemma 7.3.
>
> 1) On (σ, μ)-plane $Env(A_1, \ldots, A_N, F)$ is represented by the set
>
> $$R_{1t} \cup R_{2t} = \left\{ (\sigma, \mu) \middle| \ \sigma \geq 0, \ \mu = f \pm \frac{\mu_t - f}{\sigma_t}\sigma \right\}.$$
>
> 2) The feasible region is the angle between the rays R_{1t}: $\mu = f + \frac{\mu_t - f}{\sigma_t}\sigma$ and R_{2t}: $\mu = f - \frac{\mu_t - f}{\sigma_t}\sigma$, with the boundaries included.

Proof

We use Lemma 7.2. Since point t is on the hyperbola, the portfolio t is a portfolio of $A_1, \ldots, A_N$. The rays R_{1t} and R_{2t} are given by:

$$R_{1t}\text{: } \mu = f + \frac{\mu_t - f}{\sigma_t}\sigma \text{ and } R_{2t}\text{: } \mu = f - \frac{\mu_t - f}{\sigma_t}\sigma.$$

The first one is the tangent line to hyperbola. For any portfolio x of $A_1, \ldots, A_N$, the rays R_{1x} and R_{2x} are between R_{1t} and R_{2t}. This proves the theorem.

See detailed pictures for cases $f > \mu_0$ and $f < \mu_0$ in following Lemma 7.5 and Theorem 7.2. ☐

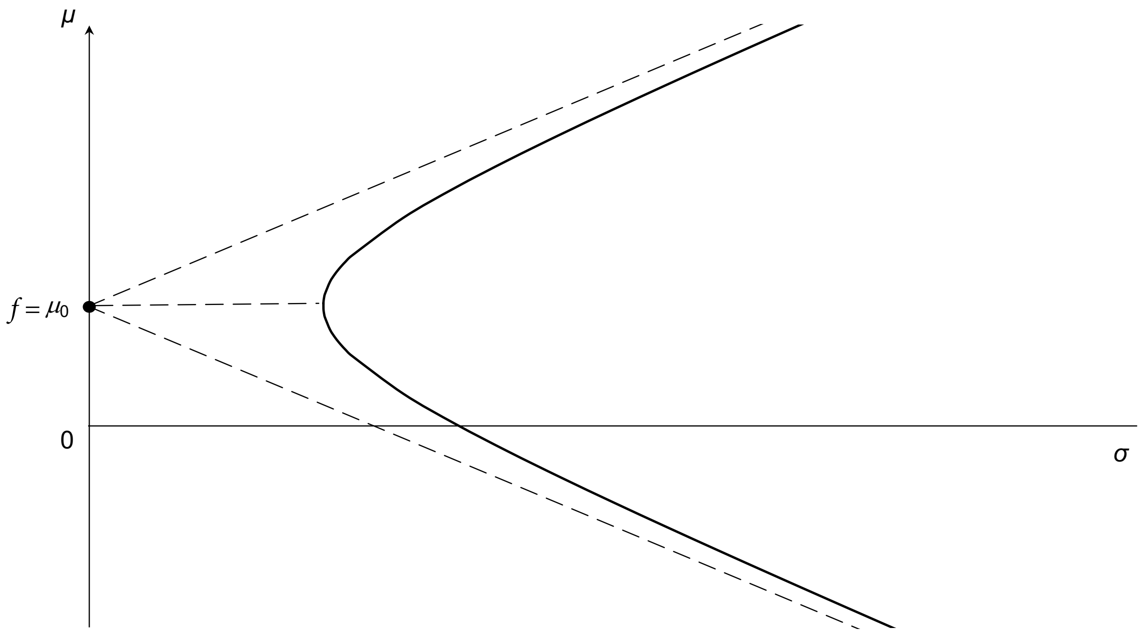
Lemma 7.4. Suppose $f = \mu_0$.

1) There is no tangent line to the hyperbola from point $(0, f)$.

2) The feasible region of $A_1, \ldots, A_N$, F is the angle between the hyperbola asymptotes $\mu = \mu_0 \pm \sqrt{\frac{B}{A}}\sigma$ (the asymptotes are not included).

3) $Env(A_1,\ldots, A_N, F)$ does not exist, that is, among portfolios with a given expected return the lower bound of risk is not attained.

4) $EF(A_1,\ldots, A_N, F)$ does not exist, that is, among portfolios with a given risk the upper bound of expected returns is not attained.



Proof

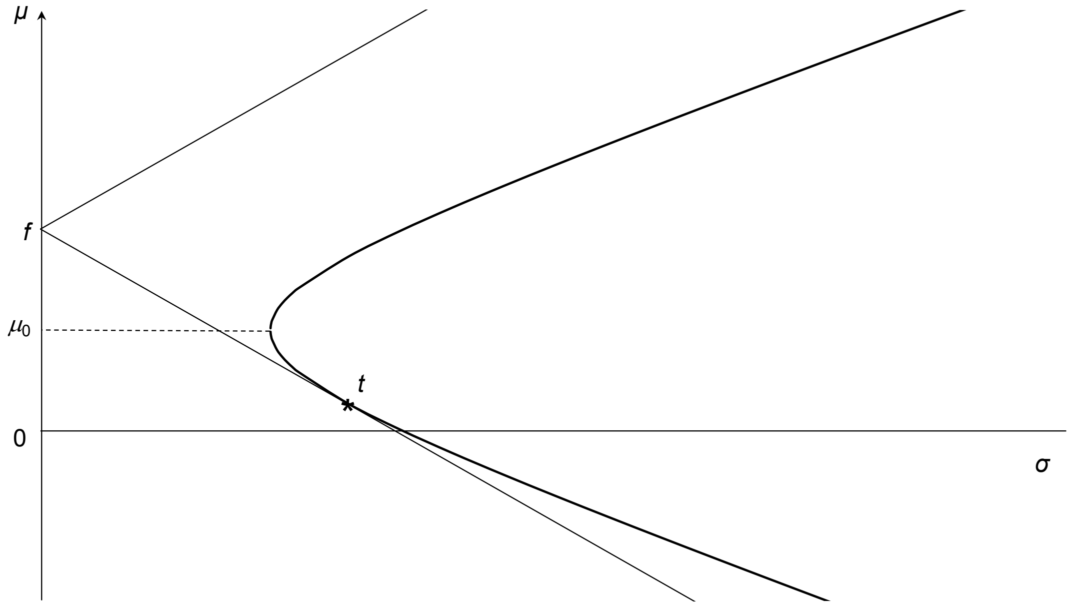
1) follows from a general property of hyperbola.
2) follows from Lemma 7.2.
3) and 4) follow from 2). ☐

Lemma 7.5. Suppose $f > \mu_0$.

1) The tangent line to the hyperbola from point $(0, f)$ is below the hyperbola.

2) The feasible region of $A_1,\ldots, A_N$, F is the set $\left\{(\sigma,\mu) \middle| \ \sigma \geq 0,\ f + \frac{\mu_t - f}{\sigma_t}\sigma \leq \mu \leq f - \frac{\mu_t - f}{\sigma_t}\sigma \right\}$.

3) On (σ, μ)-plane $EF(A_1,\ldots, A_N, F)$ is represented by the set $R_{2t} = \left\{(\sigma,\mu) \middle| \ \sigma \geq 0,\ \mu = f - \frac{\mu_t - f}{\sigma_t}\sigma \right\}$.



Proof

1) By Lemma 7.3, the second coordinate of the tangency point is $\mu_t = \mu_0 + \dfrac{B}{\mu_0 - f}$ and $f > \mu_0$. So $\mu_t < \mu_0$, the tangency point is below the vertex and the tangent line is below the hyperbola.

2) and 3) follow from 1) and Theorem 7.1. ☐

Only the case $f < \mu_0$ is realistic because in real world the risk-free return is lower than the expected return of any efficient portfolio of assets $A_1, \ldots, A_N$, including μ_0, the expected return of the portfolio with lowest risk. Further we will consider only the case $f < \mu_0$.

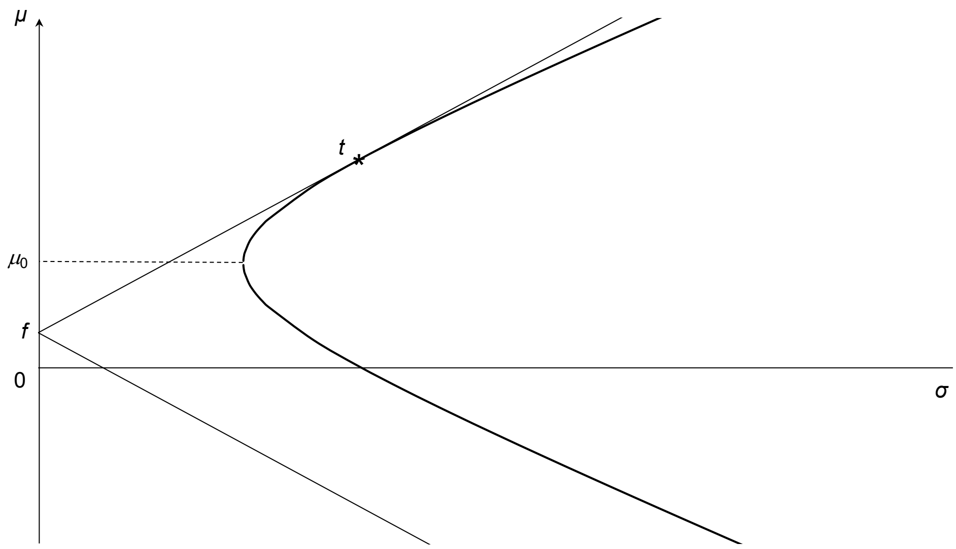
Theorem 7.2. Suppose $f < \mu_0$.

1) The tangent line to the hyperbola from point $(0, f)$ is above the hyperbola.

2) The feasible region of $A_1, A_2, \ldots, A_N$, F is the set

$$\left\{(\sigma,\mu) \mid \sigma \geq 0,\ f - \frac{\mu_t - f}{\sigma_t}\sigma \leq \mu \leq f + \frac{\mu_t - f}{\sigma_t}\sigma\right\}.$$

3) On (σ, μ)-plane $EF(A_1, \ldots, A_N, F)$ is represented by the ray R_{1t}.



Proof

1) By Lemma 7.3, the second coordinate of the tangency point is $\mu_t = \mu_0 + \dfrac{B}{\mu_0 - f}$ and $f < \mu_0$. So $\mu_t > \mu_0$, the tangency point is above the vertex and the tangent line is above the hyperbola.

2) and 3) follow from 1) and Theorem 7.1. ☐

Since for each value of expected return there is exactly one envelope portfolio, the tangency point *t* corresponds to a unique portfolio (also denoted *t*) that belongs to both envelopes $Env(A_1,\ldots, A_N)$ and $Env(A_1,\ldots, A_N, F)$.

Tangency portfolio *t* is the unique portfolio that belongs to both envelopes $Env(A_1,\ldots, A_N)$ and $Env(A_1,\ldots, A_N, F)$.

7.3 Market Portfolio. Capital Market Line

The **market portfolio** is the portfolio containing every asset in a given market with the weight proportional to its market value. It is denoted *m*.

Thus, $m = \begin{bmatrix} m_1 \\ \cdots \\ m_N \end{bmatrix}$, where each weight $m_k = \dfrac{market\ value\ of\ asset\ A_k}{total\ market\ value\ of\ all\ assets}$.

Though it is not obvious from the definition, the market portfolio depends on the value of the risk-free return in a market, since the market values of risky assets depend on it.

Market portfolio is a theoretical concept. In practice they often use proxies such as S&P 500 in the US or FTSE100 in the UK to represent the market portfolio. S&P 500 stands for Standard and Poor 500, the index consisting of 500 largest US stocks. FTSE100 is a stock market index made up of the 100 largest UK registered companies.

In this book we consider only an **equilibrium model** of a financial market. This model assumes that investors can freely buy and sell assets, and the prices of financial securities at any time rapidly reflect all available relevant information; so the market is balanced and functions well.

In case $f < \mu_0$, in the equilibrium model the market portfolio *m* and the tangency portfolio *t* are the same.

Example 7.1. Suppose the risk-free return is 2% and the market portfolio has expected return of 10% and standard deviation of 15%. If you invested 0.2 of your wealth into the market portfolio and the rest into the risk-free asset, what are the expected return and risk of your investment?

Solution

$\mu_m = 10$, $\sigma_m = 15$, $f = 2$, $x = 0.2m + 0.8f$.

So $\mu_x = 0.2\mu_m + 0.8f = 0.2\times 10 + 0.8\times 2 = 3.6\%$.

Since *f* is constant, we have $\sigma_x = 0.2\sigma_m = 0.2\times 15 = 3\%$. ☐

The tangent line to the hyperbola from point $(0, f)$ is called the **Capital Market Line** (CML).

The equation of CML is $\mu = f + \frac{\mu_m - f}{\sigma_m}\sigma\ (\sigma \geq 0)$, as was shown in Lemma 7.3. The CML passes through points f and m. The CML represents $EF(A_1,\ldots, A_N, F)$.

7.3. Exercises

1. Suppose the risk-free return is 10% and the market portfolio has expected return of 20% and standard deviation of 25%. If you invested 0.4 of your wealth in the market portfolio and the rest in the risk-free asset, what are the expected return and risk of your investment?

Answer: $\mu_x = 14\%$ and $\sigma_x = 10\%$.

2. Assume the same parameters for the risk-free asset and market portfolio as in Exercise 1. If you invested 1.2 of your wealth into the market portfolio and the rest into the risk-free asset (a short sale), what are the expected return and risk of your investment?

Answer: $\mu_x = 22\%$ and $\sigma_x = 30\%$.

7.4 Finding Market Portfolio and CML Equation

Theorem 7.3. 1) $Env(A_1,\ldots, A_N, F) = \{\lambda m + (1-\lambda)f \mid \lambda \in \boldsymbol{R}\}$.

2) $EF(A_1,\ldots, A_N, F) = \{\lambda m + (1-\lambda)f \mid \lambda \geq 0\}$.

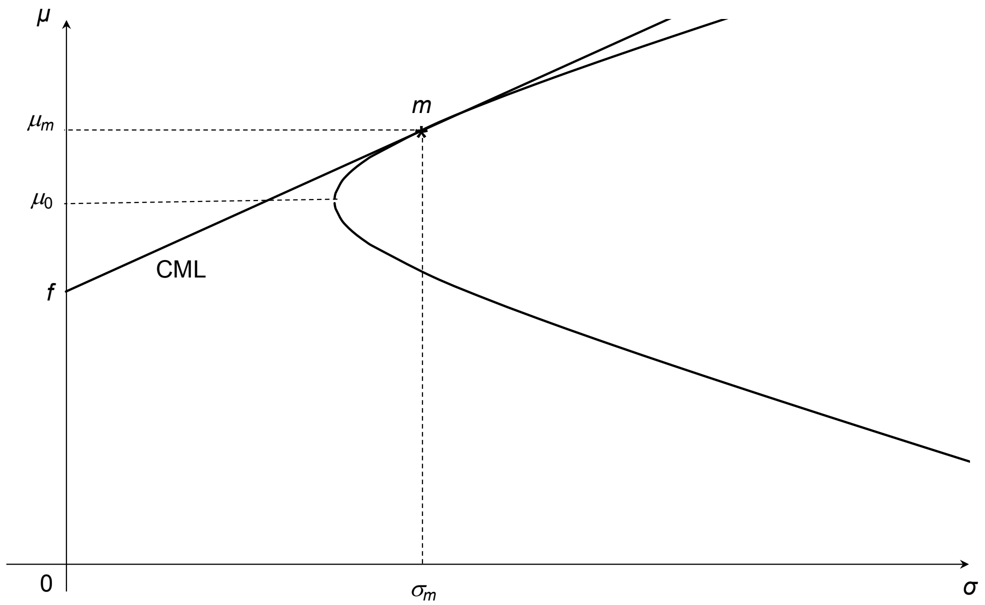
Proof

1) $y = \lambda m + (1-\lambda)f \Leftrightarrow (\sigma_y, \mu_y) \in R_{1m} \cup R_{2m} \Leftrightarrow y \in EF(A_1,\ldots, A_N, F)$.

The first equivalence follows from lemma 7.2.1) and the second equivalence follows from Theorem 7.1.1).

2) follows from 1) because $EF(A_1,\ldots, A_N, F)$ is represented by CML and points on CML correspond to portfolios $\lambda m + (1-\lambda)f$ with $\lambda \geq 0$. ☐

The CML is the ray R_{1m} passing through the points $(0, f)$ and (σ_m, μ_m): $\dfrac{\sigma}{\sigma_m} = \dfrac{\mu - f}{\mu_m - f}$ $(\sigma \geq 0)$.



Here the top half of the right branch of the hyperbola represents the efficient frontier of the N risky assets. CML represents the efficient frontier of the N risky assets combined with the risk-free asset F (see Theorem 7.2.3). For best expected returns investors should invest in the portfolios along CML, that is the portfolios of the form $\lambda m + (1-\lambda) f$. The investors who tolerate risk will invest in the market portfolio m or the portfolios above m (with $\lambda > 1$). The risk averse investors will have a higher proportion of the risk-free asset f, which means investing in the portfolios below m (with $0 \leq \lambda < 1$); so these investors will have a lower risk but also a lower expected return.

Example 7.2. The market portfolio has expected return of 20% and standard deviation of 8%. The risk-free return is 4%.

1) Write the equation of the CML.
2) Efficient portfolio x has expected return of 16%.

a) Find its standard deviation.
b) If the total value of portfolio x is \$1000, how is this value allocated?

Solution

$\mu_m = 20$, $\sigma_m = 8$, $f = 4$.

1) The CML passes through the points $(0, f)$ and (σ_m, μ_m), which are (0, 4) and (8, 20). So

CML: $\frac{\sigma}{8} = \frac{\mu - 4}{20 - 4}$, CML: $\mu = 2\sigma + 4$ $(\sigma \geq 0)$.

2) a) Portfolio x is efficient, therefore it is on the CML. So $\mu_x = 2\sigma_x + 4$, $16 = 2\sigma_x + 4$, $\sigma_x = 6\%$.

b) Since x is on the envelope, it has the form $x = m + (1-\lambda) f$ by Theorem 7.3.1). Its expected return equals $\mu_x = \lambda\mu_m + (1-\lambda) f$, $16 = \lambda \cdot 20 + (1-\lambda) \cdot 4$, $\lambda = 0.75$ and $x = 0.75m + 0.25f$. So \$750 is invested in the market portfolio and \$250 is invested in the risk-free asset. ❑

Theorem 7.4. The market portfolio m has the following numerical characteristics:

$$\sigma_m = \sqrt{A + \frac{AB}{(\mu_0 - f)^2}}, \quad \mu_m = \mu_0 + \frac{B}{\mu_0 - f}.$$

Proof

It follows from Lemma 7.3. ❑

Theorem 7.4 can be used to find the market portfolio. Next theorem provides another method of calculating the market portfolio. The theorem is stated without proof; the proof can be found in (Kachapova and Kachapov, 2006).

Theorem 7.5. The market portfolio m equals kz, where $z = S^{-1}(M^T - fU^T)$ and the coefficient of proportionality $k = \dfrac{1}{\sum_{i=1}^{N} z_i}$.

Theorem 7.5 gives a method for calculating the market portfolio in two steps:

step 1: calculate $z = S^{-1}(M^T - fU^T)$; step 2: calculate $m = \dfrac{z}{\sum_{i=1}^{N} z_i}$.

Example 7.3. Two assets from Example 6.6 have expected returns of 1% and 3%, respectively, and covariance matrix $S = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$. The risk-free return is 1.2%.

1) Find the market portfolio using Theorem 7.4.
2) Find the equation of CML.
3) Find the envelope and efficient frontier of the two risky assets combined with the risk-free asset.

Solution

1) $f = 1.2$. In Example 6.6 we found the envelope and its representation on (σ, μ)-plane:

$$Env(A_1, A_2) = \left\{ \frac{1}{2}\begin{bmatrix} -\mu+3 \\ \mu-1 \end{bmatrix} \;\middle|\; \mu \in R \right\} \text{ and } \frac{\sigma^2}{0.1} - \frac{(\mu-1.6)^2}{0.04} = 1 \;(\sigma > 0).$$

So $\mu_0 = 1.6$, $B = 0.04$. By Theorem 7.4, $\mu_m = \mu_0 + \dfrac{B}{\mu_0 - f} = 1.6 + \dfrac{0.04}{1.6-1.2} = 1.7.$

Substitute $\mu = 1.7$ into the expression for envelope portfolio to find the market portfolio:

$$m = \frac{1}{2}\begin{bmatrix} -1.7+3 \\ 1.7-1 \end{bmatrix}, \quad m = \begin{bmatrix} 0.65 \\ 0.35 \end{bmatrix}.$$

2) $A = 0.1$. By Theorem 7.4, $\sigma_m = \sqrt{A + \dfrac{AB}{(\mu_0 - f)^2}} = \sqrt{0.1 + \dfrac{0.1 \times 0.04}{(1.6-1.2)^2}}, \quad \sigma_m = \sqrt{0.125}.$

CML passes through the points $f\,(0, 1.2)$ and $m\left(\sqrt{0.125}, 1.7\right)$. Its equation is $\dfrac{\sigma}{\sqrt{0.125}} = \dfrac{\mu - 1.2}{1.7-1.2}$.

CML: $\mu = \dfrac{\sigma}{\sqrt{0.5}} + 1.2 \;(\sigma \ge 0).$

3) By Theorem 7.3.1), $Env(A_1, A_2, F) = \{\lambda m + (1-\lambda) f \mid \lambda \in \boldsymbol{R}\}$.

$$\lambda m + (1-\lambda) f = \lambda \begin{bmatrix} 0.65 \\ 0.35 \\ 0 \end{bmatrix} + (1-\lambda)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.65\lambda \\ 0.35\lambda \\ 1-\lambda \end{bmatrix}, \; Env(A_1, A_2, F) = \left\{ \begin{bmatrix} 0.65\lambda \\ 0.35\lambda \\ 1-\lambda \end{bmatrix} \;\middle|\; \lambda \in R \right\}.$$

By Theorem 7.3.2), $EF(A_1, A_2, F) = \{\lambda m + (1-\lambda) f \mid \lambda \ge 0\}$. So $EF(A_1, A_2, F) = \left\{ \begin{bmatrix} 0.65\lambda \\ 0.35\lambda \\ 1-\lambda \end{bmatrix} \;\middle|\; \lambda \ge 0 \right\}$. □

Example 7.4. Three assets from Example 6.3 have expected returns of 2%, 1% and 1%, respectively and covariance matrix $S = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 6 \end{bmatrix}$. The risk-free return is 0.4%.

1) Find the market portfolio using Theorem 7.4.
2) Find the equation of CML.

Solution

1) $f = 1.2$. In Example 6.3 we found the envelope and its representation on (σ, μ)-plane:

$$Env(A_1, A_2, A_3) = \left\{ \frac{1}{3}\begin{bmatrix} 3\mu-3 \\ -4\mu+8 \\ \mu-2 \end{bmatrix} \middle| \ \mu \in R \right\} \text{ and } \frac{\sigma^2}{0.5} - \frac{(\mu-0.5)^2}{0.75} = 1 \ (\sigma > 0).$$

So $\mu_0 = 0.5$, $B = 0.75$. By Theorem 7.4, $\mu_m = \mu_0 + \dfrac{B}{\mu_0 - f} = 0.5 + \dfrac{0.75}{0.5-0.4} = 8.$

Substitute $\mu = 8$ into the expression for envelope portfolio to find the market portfolio:

$$m = \frac{1}{3}\begin{bmatrix} 3\times 8-3 \\ -4\times 8+8 \\ 8-2 \end{bmatrix}, \ m = \begin{bmatrix} 7 \\ -8 \\ 2 \end{bmatrix}.$$

2) $A = 0.5$. By Theorem 7.4, $\sigma_m = \sqrt{A + \dfrac{AB}{(\mu_0 - f)^2}} = \sqrt{0.5 + \dfrac{0.5\times 0.75}{(0.5-0.4)^2}}, \quad \sigma_m = \sqrt{38}.$

CML passes through the points $f\,(0, 0.4)$ and $m\left(\sqrt{38}, 8\right)$. Its equation is $\dfrac{\sigma}{\sqrt{38}} = \dfrac{\mu - 0.4}{8-0.4}$.

CML: $\mu = \sqrt{1.52}\,\sigma + 0.4 \ (\sigma \geq 0)$. ☐

Example 7.5. There are two assets A_1 and A_2 on a market with expected returns of 8% and 6%, respectively, and covariance matrix $S = \begin{bmatrix} 5 & 1 \\ 1 & 1 \end{bmatrix}$. The risk-free return is 3%.

1) Find the market portfolio using Theorem 7.5.
2) If the market value of company A_1 is \$3 mln, what is the market value of company A_2?

Solution

1) $f = 3$, $M = [8\ 6]$, $S^{-1} = \dfrac{1}{4}\begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}$.

By Theorem 7.5, $z = S^{-1}(M^T - fU^T) = \dfrac{1}{4}\begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}\left(\begin{bmatrix} 8 \\ 6 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right), \qquad z = \dfrac{1}{2}\begin{bmatrix} 1 \\ 5 \end{bmatrix}.$

$\sum_{i=1}^{2} z_i = 3. \qquad m = \dfrac{z}{3}, \qquad m = \dfrac{1}{6}\begin{bmatrix} 1 \\ 5 \end{bmatrix}.$

2) The proportion of asset A_1 in the market portfolio is $\frac{1}{6}$, which equals the proportion of the market value of company A_1 on the market. Hence the total value of both stocks equals $3 \div \frac{1}{6} = \$18$ mln and the market value of company A_2 equals $18 \times \frac{5}{6} = \$15$ mln. ☐

Example 7.6. Three assets have expected returns of 5%, 6% and 3%, respectively, and covariance matrix $S = \begin{bmatrix} 3 & 3 & 1 \\ 3 & 5 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. The risk-free return is 1%.

1) Find the market portfolio using Theorem 7.5.
2) Find the equation of CML.
3) Find the envelope and efficient frontier of the three risky assets combined with the risk-free asset.

Solution

1) $f = 3$, $M = [5\ 6\ 3]$, $S^{-1} = \frac{1}{2}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 3 \end{bmatrix}$.

By Theorem 7.5, $z = S^{-1}(M^T - fU^T) = \frac{1}{2}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 3 \end{bmatrix}\left(\begin{bmatrix} 5 \\ 6 \\ 3 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$, $\quad z = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

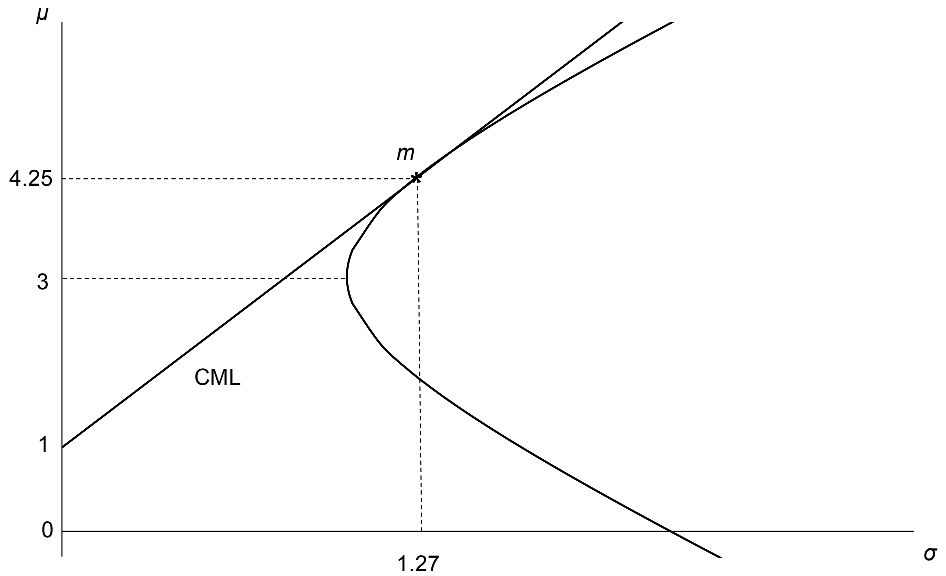
$\sum_{i=1}^{3} z_i = 2. \qquad m = \frac{z}{2}, \qquad m = \frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

2) $\mu_m = Mm = [5\ 6\ 3]\frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, $\mu_m = \frac{17}{4}$.

$\sigma_m{}^2 = m^T S m = \frac{1}{16}\begin{bmatrix} 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 3 & 1 \\ 3 & 5 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, $\sigma_m{}^2 = \frac{13}{8}$, $\sigma_m = \sqrt{\frac{13}{8}}$.

CML passes through the points $f\,(0, 1)$ and $m\left(\sqrt{\frac{13}{8}}, \frac{17}{4}\right)$.

Its equation is $\dfrac{\sigma}{\sqrt{\frac{13}{8}}} = \dfrac{\mu - 1}{\frac{17}{4} - 1}$, $\quad$ CML: $\mu = \frac{\sqrt{26}}{2}\sigma + 1$ $(\sigma \geq 0)$.



3) By Theorem 7.3.1), $Env(A_1, A_2, F) = \{\lambda m + (1-\lambda) f \mid \lambda \in \boldsymbol{R}\}$.

$$\lambda m + (1-\lambda) f = \lambda \begin{bmatrix} 0.25 \\ 0.25 \\ 0.5 \\ 0 \end{bmatrix} + (1-\lambda) \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.25\lambda \\ 0.25\lambda \\ 0.5\lambda \\ 1-\lambda \end{bmatrix}, \ Env(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} 0.25\lambda \\ 0.25\lambda \\ 0.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}.$$

By Theorem 7.3.2), $EF(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} 0.25\lambda \\ 0.25\lambda \\ 0.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$. ☐

7.4. Exercises

In each of the following exercises find:

1) the market portfolio;
2) the equation of CML;
3) the envelope and efficient frontier of the three risky assets combined with the risk-free asset.

1. Three assets have expected returns of 2%, 2% and 4%, respectively, and covariance matrix $S = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. The risk-free return is 0%.

Answer: 1) $m = \frac{1}{4}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$, $\mu_m = 2.5$, $\sigma_m^2 = \frac{5}{8}$. 2) CML: $\mu = 2\sigma + 1$ ($\sigma \geq 0$).

3) $Env(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} \lambda \\ -0.5\lambda \\ 0.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} \lambda \\ -0.5\lambda \\ 0.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

2. The same risky assets as in exercise 1 with risk-free return of 1%.

Answer: 1) $m = \begin{bmatrix} 1 \\ -0.5 \\ 0.5 \end{bmatrix}$, $\mu_m = 3$, $\sigma_m^2 = 1$. 2) CML: $\mu = 2\sigma + 1$ ($\sigma \geq 0$).

3) $Env(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} \lambda \\ -0.5\lambda \\ 0.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} \lambda \\ -0.5\lambda \\ 0.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

3. Three assets have expected returns of 4%, 2% and 4%,respectively, and covariance matrix $S = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 3 & 2 \\ -1 & 2 & 2 \end{bmatrix}$. The risk-free return is 2%.

Answer: 1) $m = \frac{1}{5}\begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$, $\mu_m = 4.4$, $\sigma_m{}^2 = \frac{6}{25}$. 2) CML: $\mu = 2\left(\sqrt{6}\,\sigma + 1\right)$ $(\sigma \geq 0)$.

3) $Env(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} 0.6\lambda \\ -0.2\lambda \\ 0.6\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} 0.6\lambda \\ -0.2\lambda \\ 0.6\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

4. The same risky assets as in exercise 3 with risk-free return of 3%.

Answer: 1) $m = \frac{1}{5}\begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix}$, $\mu_m = 4.8$, $\sigma_m{}^2 = \frac{9}{25}$. 2) CML: $\mu = 3\sigma + 3$ $(\sigma \geq 0)$.

3) $Env(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} 0.6\lambda \\ -0.4\lambda \\ 0.8\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} 0.6\lambda \\ -0.4\lambda \\ 0.8\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

5. Three assets have expected returns of 4%, 2% and 0%, respectively, and covariance matrix $S = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. The risk-free return is 2%.

Answer: 1) $m = \frac{1}{2}\begin{bmatrix} 2 \\ 3 \\ -3 \end{bmatrix}$, $\mu_m = 7$, $\sigma_m^{\ 2} = \frac{5}{4}$. 2) CML: $\mu = 2\sqrt{5}\,\sigma + 2$ ($\sigma \geq 0$).

3) $Env(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} \lambda \\ 1.5\lambda \\ -1.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} \lambda \\ 1.5\lambda \\ -1.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

6. The same risky assets as in exercise 5 with risk-free return of 3%.

1) $m = \frac{1}{2}\begin{bmatrix} 3 \\ 4 \\ -5 \end{bmatrix}$, $\mu_m = 10$, $\sigma_m^{\ 2} = \frac{7}{2}$. 2) CML: $\mu = \sqrt{14}\,\sigma + 3$ ($\sigma \geq 0$).

3) $Env(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} 1.5\lambda \\ 2\lambda \\ -2.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \begin{bmatrix} 1.5\lambda \\ 2\lambda \\ -2.5\lambda \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

7. Three assets have expected returns of 2%, 1% and 2%, respectively, and covariance matrix $S = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 6 \end{bmatrix}$. The risk-free return is 0%.

Answer: 1) $m = \frac{1}{3}\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$, $\mu_m = 1$, $\sigma_m^{\ 2} = \frac{1}{3}$. 2) CML: $\mu = \sqrt{3}\,\sigma$ ($\sigma \geq 0$).

3) $Env(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \frac{1}{3}\begin{bmatrix} -\lambda \\ 3\lambda \\ \lambda \\ 3-3\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{F}) = \left\{ \frac{1}{3}\begin{bmatrix} -\lambda \\ 3\lambda \\ \lambda \\ 3-3\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

8. The same risky assets as in exercise 7 with risk-free return of 0.5%.

1) $m = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, $\mu_m = 3$, $\sigma_m^{\ 2} = 5$. 2) CML: $\mu = \frac{\sqrt{5}\sigma + 1}{2}$ ($\sigma \geq 0$).

3) $Env(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} 2\lambda \\ -\lambda \\ 0 \\ 1-\lambda \end{bmatrix} \middle| \lambda \in R \right\}$, $EF(A_1, A_2, A_3, F) = \left\{ \begin{bmatrix} 2\lambda \\ -\lambda \\ 0 \\ 1-\lambda \end{bmatrix} \middle| \lambda \geq 0 \right\}$.

7.5 Regression in Finance. Beta Coefficient

Linear regression was covered in Section 4.4. In the world of N risky assets $A_1, \ldots, A_N$ and a risk-free asset F all portfolios are regressed to the market portfolio m. The market portfolio is not risk-free, so $\sigma_m \neq 0$.

Theorem 7.6. For any portfolio x the following holds.

1) $x = \alpha_x + \beta_x m + \varepsilon$, where $\begin{cases} \beta_x = \dfrac{Cov(x,m)}{\sigma_m^2}, \\ \alpha_x = \mu_x - \mu_m \beta_x. \end{cases}$

The regression line $\alpha_x + \beta_x m$ is the linear function of m closest to x. So the coefficient β_x is the average rate of change of x's return with respect to the market return.

2) For the residual ε, $\mu_\varepsilon = 0$ and $Cov(\varepsilon, m) = 0$.

3) The variance σ_x^2 represents the total risk of portfolio x;

$\beta_x^2 \sigma_m^2$ is the **systematic risk** (or **market risk**) of x, the risk that affects most investments;

σ_ε^2 is the **unsystematic risk** (or **asset-specific risk**) of x, the risk that affects only a small number of investments.

The total risk of x is a sum of the systematic risk and unsystematic risk:

$$\sigma_x^2 = \beta_x^2 \sigma_m^2 + \sigma_\varepsilon^2 .$$

Proof

1) and 2 follow directly from Theorem 4.5 and its Corollary.

3) $\sigma_x^2 = Var(x) = Var(\alpha_x + \beta_x m + \varepsilon) = Var(\beta_x m + \varepsilon) = \beta_x^2 Var(m) + Var(\varepsilon) + 2\beta_x Cov(\varepsilon, m) =$

$= \beta_x^2 \sigma_m^2 + \sigma_\varepsilon^2$, since $Cov(\varepsilon, m) = 0$. ☐

These are some examples of factors of systematic risk:

- unexpected drop in interest rates;
- inflation;
- release of government figures on GDP (gross domestic product).

These are some examples of factors of unsystematic risk for a particular company:

- appointment of a new president of the company;
- news about the company's sales figures.

> The regression line $x = \alpha_x + \beta_x\, m$ is called the **characteristic line** of portfolio x.
>
> β_x is called the **beta coefficient** (or just **beta**) of portfolio x.

Thus, β_x is the slope and α_x is the y-intercept of the characteristic line.

Beta measures the response of the portfolio's return to a change in the market return. For example, if a portfolio has $\beta = 2$ and the market goes up by 10%, then this portfolio's return will increase by 20% on average.

Since $\mu_\varepsilon = 0$, we have $\mu_x = \alpha_x + \beta_x \mu_m$. So the expected return μ_x depends only on β_x, not σ_ε; that is μ_x depends only on systematic risk.

For a portfolio x denote $\rho_{x,m}$ the coefficient of correlation between x and m.

Properties of Beta and Correlation Coefficients.

1) The beta for the market portfolio $\beta_m = 1$.

2) The beta for the risk-free asset $\beta_f = 0$.

3) For portfolio $x = \begin{bmatrix} x_1 \\ \dots \\ x_N \\ x_{N+1} \end{bmatrix}$ the beta equals $\beta_x = \sum_{i=1}^{N} x_i \beta_i$, where β_i is the beta coefficient for asset A_i ($i = 1, \dots, N$).

4) $\beta_x = \rho_{x,m} \dfrac{\sigma_x}{\sigma_m}$.

5) $\rho_{x,m}^2$ is the proportion of systematic risk of portfolio x and $1-\rho_{x,m}^2$ is the proportion of unsystematic risk of x.

Proof

1) $\beta_m = \dfrac{Cov(m,m)}{\sigma_m^2} = \dfrac{\sigma_m^2}{\sigma_m^2} = 1$.

2) $Cov(f, m) = 0$, so $\beta_f = 0$.

3) $x = \sum_{i=1}^{N} x_i r_i + x_{N+1} f$, $Cov(x,m) = \sum_{i=1}^{N} x_i Cov(r_i, m)$, since $Cov(f, m) = 0$.

$$\beta_x = \frac{Cov(x,m)}{\sigma_m^2} = \sum_{i=1}^{N} x_i \frac{Cov(r_i,m)}{\sigma_m^2} = \sum_{i=1}^{N} x_i \beta_i .$$

4) $\rho_{x,m} = \dfrac{Cov(x,m)}{\sigma_x \sigma_m}$. So $\rho_{x,m} \cdot \dfrac{\sigma_x}{\sigma_m} = \dfrac{Cov(x,m)}{\sigma_x \sigma_m} \cdot \dfrac{\sigma_x}{\sigma_m} = \dfrac{Cov(x,m)}{\sigma_m^2} = \beta_x$.

5) $\dfrac{systematic\ \ risk\ \ of\ \ x}{total\ \ risk\ \ of\ \ x} = \dfrac{\beta_x^2 \sigma_m^2}{\sigma_x^2} = \beta_x^2 \dfrac{\sigma_m^2}{\sigma_x^2} =$ (by Theorem 7.6.1) $= \left(\dfrac{Cov(x,m)}{\sigma_m^2}\right)^2 \cdot \dfrac{\sigma_m^2}{\sigma_x^2} =$

$$= \left(\frac{Cov(x,m)}{\sigma_m \sigma_x}\right)^2 = \rho_{x,m}^2 .$$

The remaining proportion is $1-\rho_{x,m}^2$, which is the proportion of unsystematic risk of portfolio x. ☐

Theorem 7.7. For any portfolio x, β_x^2 is a ratio of the systematic risk of x to the risk of the market portfolio.

Proof

$$\beta_x^2 = \frac{\beta_x^2 \sigma_m^2}{\sigma_m^2} = \frac{systematic\ risk\ of\ x}{risk\ of\ the\ market\ portfolio}.$$ ☐

Beta coefficient is used for ordinal ranking of assets according to their systematic risk, due to the previous theorem.

An asset with $\beta > 1$ is called an **aggressive** asset (it is more volatile than the market portfolio).

An asset with $\beta < 1$ is called a **defensive** asset (it is less volatile than the market portfolio).

The following table shows the estimated beta coefficients of some companies in January 2013:

Company Name	**Beta coefficient**
McDonald's Corporation	0.39
The Procter & Gamble company	0.46
PepsiCo, Inc.	0.48
Coca-Cola Company	0.51
Exxon Mobil Corporation	0.51
Vodafone Group plc (ADR)	0.70
Toyota Motor Corporation	0.75
Chevron Corporation	0.79
Yahoo! Inc.	0.83
Honda Motor Company	0.85
Amazon.com, Inc	0.93
Microsoft Corporation	0.97
Intel Corporation	1.07
Hewlett-Packard Company	1.07
Google Inc	1.08
Oracle Corporation	1.11
Mitsubishi Financial Corporation	1.12
The Walt Disney Company	1.19
Starbucks Corporation	1.20
Apple Inc	1.21
Boeing Company	1.21
Dell Inc.	1.37
eBay Inc.	1.46
Nissan Motor Co	1.51
Sony Corporation	1.54
Nokia	1.55
Xerox Corporation	1.60
General Electric Company	1.60
Harley-Davidson	2.16
Ford Motor Company	2.28
Quantum Corp	3.23

This range of betas is typical for stocks of large corporations.

7.6 Capital Asset Pricing Model and Security Market Line

Theorem 7.8. For any portfolio x of assets $A_1, \dots, A_N$:

1) $\beta_x = \dfrac{\mu_x - f}{\mu_m - f}$;

2) $\mu_x = f + \beta_x(\mu_m - f)$. This equation is called **Capital Asset Pricing Model** (CAPM).

Proof

1) By Theorem 7.5 the market portfolio $m = kz$, where $z = S^{-1}(M^T - fU^T)$ and $k = \dfrac{1}{\sum_{i=1}^{N} z_i}$.

$Cov(x, m) = x^T S m = k\, x^T S\, S^{-1}(M^T - fU^T) = k\,(x^T M^T - f x^T U^T) = k\,(Mx - fUx)^T = k\,(\mu_x - f)$.

When $x = m$, we have $\sigma_m^{\ 2} = Cov(m, m) = k\,(\mu_m - f)$. So $\beta_x = \dfrac{Cov(x,m)}{\sigma_m^2} = \dfrac{k(\mu_x - f)}{k(\mu_m - f)} = \dfrac{\mu_x - f}{\mu_m - f}$.

2) follows directly from 1). □

The graph of CAPM equation on the (β, μ)-plane is called the **Security Market Line** (**SML**).

The slope of SML is $(\mu_m - f)$, which is called the **market risk premium**. It is the risk premium of the market portfolio.

The CAPM estimates how expected return depends on systematic risk. The CAPM model says that in a balanced market (the equilibrium model) all portfolios of the N risky assets have their (β, μ)-points on the Security Market Line. Mean-variance analysis and CAPM are used by brokers, pension-fund managers and consultants when formulating investment strategies and giving financial advice.

Example 7.7. The market portfolio has a standard deviation of 10%. Portfolio x has a standard deviation of 20% and a correlation coefficient of 0.65 with the market portfolio. Find the beta of the portfolio x.

Solution

By property 4) of beta, $\beta_x = \rho_{x,m} \dfrac{\sigma_x}{\sigma_m} = 0.65 \times \dfrac{20}{10} = 1.3$. □

Example 7.8. Portfolios x, y and z have the following characteristic lines and coefficients of correlation with the market portfolio:

$x = 2 + 1.5m$, $\rho_{x,m} = 0.9$; $y = 3 + 2.1m$, $\rho_{y,m} = 0.7$; $z = 4 + 0.8m$, $\rho_{z,m} = 0.6$.

1) Which portfolio has the most systematic risk?
2) For each portfolio find the percentage of systematic risk and unsystematic risk.
3) For each portfolio determine whether it is aggressive or defensive.

Solution

1) $\beta_x = 1.5$, $\beta_y = 2.1$, $\beta_z = 0.8$. Portfolio y has the highest beta, hence y has the most systematic risk.

2) By property 5), the proportion of systematic risk of portfolio x equals $\rho^2_{x,m} = 0.9^2 = 0.81 =$ 81%.

   The percentage of unsystematic risk for x is 100 – 81 = 19%.

   Similarly $\rho^2_{y,m} = 0.7^2 = 0.49 =$ 49%. So portfolio y has 49% of systematic risk and 51% of unsystematic risk.

   $\rho^2_{z,m} = 0.6^2 = 0.36 =$ 36%. So portfolio z has 36% of systematic risk and 64% of unsystematic risk.

3) $\beta_x = 1.5 > 1$, $\beta_y = 2.1 > 1$, so portfolios x and y are aggressive. $\beta_z = 0.8 < 1$, so portfolio z is defensive. □

Example 7.9. Portfolio x consists of three stocks with the following market values and betas:

Stock	*Market Value ($)*	*Beta*
A_1	2,000	1.3
A_2	3,000	0.8
A_3	5,000	0.6

1) Find the proportion of each stock in portfolio x.
2) Find the beta coefficient for x.
3) If the market portfolio has expected return of 7% and the risk-free return is 2%, what is the expected return of portfolio x? Use CAPM model.

Solution

1) The total value of the portfolio is 2000 + 3000 + 5000 = $10,000. The proportions are:

$$x_1 = \frac{2000}{10000} = 0.2,\ x_2 = \frac{3000}{10000} = 0.3,\ x_3 = \frac{5000}{10000} = 0.5.\quad x = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}.$$

2) By property 3) of beta, $\beta_x = \sum_{i=1}^{3} x_i\beta_i = 0.2\times1.3 + 0.3\times0.8 + 0.5\times0.6$, $\beta_x = 0.8$.

3) $\mu_m = 7, f = 2$. CAPM: $\mu_x = f + \beta_x(\mu_m - f)$, $\mu_x = 2 + 0.8\times(7-2) = 6\%$. □

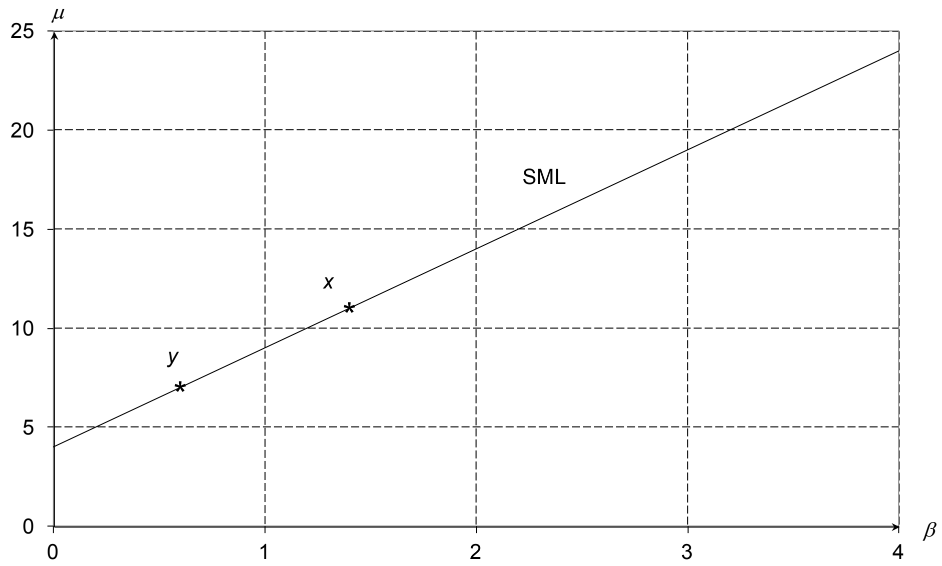
Example 7.10. The market portfolio has expected return of 9% and the risk-free return is 4%.

1) Write the equation of CAPM.
2) Draw the SML.
3) Use the CAPM to find the following:
   a) the expected return for portfolio x with beta of 1.4;
   b) the beta for portfolio y with expected return of 7%.

Solution

$\mu_m = 9, f = 4$.

1) $\mu_x = f + \beta_x(\mu_m - f)$, $\mu_x = 4 + 5\beta_x$.
2)



3) a) $\mu_x = 4 + 5\times1.4 = 11\%$. b) $7 = 4 + 5\times\beta_y$, $\beta_y = 0.6$. □

7.6. Exercises

1. An equilibrium market has two risky assets A_1 and A_2 and a risk-free asset. The expected returns of A_1 and A_2 are 6% and 4%, respectively, and the risk-free return is 2.5%. The market value of company A_1 equals \$6 mln and the market value of company A_2 equals \$2 mln.

1) Find the market portfolio and its expected return.
2) Find the equation of CAPM.

Answer: 1) $m = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$, $\mu_m = 5.5\%$. 2) CAPM: $\mu_x = 2.5 + 3\beta_x$.

2. Assets A_1, A_2 and A_3 have betas of 0.8, 1.5 and 2.0, respectively.

1) Which asset has the most systematic risk?
2) For each asset determine whether it is aggressive or defensive.

Answer: 1) A_3 . 2) A_1 is defensive, A_2 and A_3 are aggressive.

3. $x = \begin{bmatrix} 0.5 \\ 0.2 \\ 0.3 \end{bmatrix}$ is a portfolio of the assets from exercise 2.

1) Find the beta of portfolio *x*.

2) The risk-free return is 5% and the expected return of the market portfolio is 25%. Use CAPM to find the expected return of x.

Answer: 1) $\beta_x = 1.3$. 2) $\mu_x = 31\%$.

4. The market portfolio has a standard deviation of 14%. Portfolio x has a standard deviation of 35% and a beta of 1.75. Find the percentage of systematic risk and unsystematic risk for portfolio x.

Answer: 49% of systematic risk and 51% of unsystematic risk.

5. The market portfolio has expected return of 40% and the risk-free return is 15%.

1) Find the equation of CAPM. 2) If the expected return of portfolio x is 55%, what is its beta?

Answer: 1) $\mu_x = 15 + 25\beta_x$. 2) $\beta_x = 1.6$.

8 Bibliography

Adams, A., Bloomfield, D., Booth, P., England, P. (1993). *Investment mathematics and statistics*. Graham and Trotman, London.

Benninga, S. (2001). *Financial Modeling* (2nd edition). The MIT Press, Cambridge, London.

Brealey, R. and Myers, S. (1985). *Principles of corporate finance* (2nd edition). McGraw-Hill.

Francis, J.C. and Taylor, R.W. (2000). *Theory and problems of investments* (2nd edition). McGraw-Hill.

Gitman, L.J. (1997). *Principles of managerial finance* (8th edition). Addison-Wesley.

Grimmett, G.R. and Stirzaker, D.R. (2004). *Probability and random processes* (3rd edition). Oxford University Press, Oxford.

Grinblatt, M. and Titman, S. (2002). *Financial markets and corporate strategy* (2nd edition). McGraw-Hill, Irwin.

Hsu, H.P. (1997). *Probability, random variables, and random processes.* New York: McGraw-Hill.

Il'in, V.A. and Poznyak, E.G. (1985). *Analytic Geometry*. Imported Pubn.

Kachapova, F. and Kachapov, I. (2005). Calculating optimal portfolios of three financial assets. *Mathematics and Informatics Quarterly*. Vol. 15, No. 3/4, pp. 100–107.

Kachapova, F. and Kachapov, I. (2006). *Mathematical approach to portfolio analysis*. Maths Ken.

Kachapova, F. and Kachapov, I. (2010). Orthogonal projection in teaching regression and financial mathematics. *Journal of Statistics Education* (*JSE*). ISSN 1069-1898, Vol. 18, No. 1, pp. 1–18. URL: http://www.amstat.org/publications/jse/v18n1/kachapova.pdf.

Kachapova, F. and Kachapov, I. (2012). Finding the envelope and efficient frontier of financial assets. *Journal of Mathematics and Statistics*, ISSN 1549-3644, Vol. 8, No. 3, pp. 323–329. URL: http:///thescipub.com/abstract/10.3844/jmssp.2012.323.329.

Kremer, N.Sh. (2004). *Probability theory and mathematical statistics* (2nd edition). Unity, Moscow.

Lay, D.C. (1997). *Linear algebra and its applications* (2nd edition). Addison-Wesley.

Riddle, D. (1995). *Analytic Geometry*, 6^{th} edition. Brooks Cole.

Ross, S.A., Westerfield, R.W., Jordan, B.D. (1998). *Fundamentals of corporate finance* (4th edition). Irwin, McGraw-Hill.

Teall, J.L. and Hasan, I. (2002). *Quantitative methods for finance and investments.* Blackwell Publishing.