Chapter 1

Vector spaces and linear mappings

Vector spaces are probably the most useful mathematical structure in economics and finance, as in many other scientific fields. Elements of vector spaces are called vectors and the reader already knows this mathematical object, at least in an intuitive way. In fact, we live in a 3-dimensional vector space and, as a good approximation, the page you are now reading is part of a 2-dimensional vector space.

In finance and economics, vectors are generally characterized by more than 2 or 3 coordinates, and in some cases they are elements of infinite-dimensional vector spaces. Whatever the case, it is fundamental to master these mathematical tools because they are important in a number of applications like arbitrage pricing, portfolio choice, and empirical studies in general.

Section I presents the definition of a vector space and its elementary properties. In the beginning of the chapter, we restrict the presentation to finite-dimensional spaces that are natural generalizations of the 2 and 3-dimensional spaces we are used to. The mathematical concept of a *vector space* is illustrated by means of the economic concept of a *complete market*.

The second section of the chapter develops the properties of linear map-

pings. Linear mappings are fundamental components of arbitrage pricing models. Representation of linear mappings by matrices is developed in section 3 and the special case of square matrices is addressed in more details. In particular, we present the diagonalization of square matrices and the notions of eigenvalues and eigenvectors.

Norms and inner products, arising naturally in valuation models, are developed in section 1.4 and their properties are discussed in the general framework of Hilbert spaces in section 1.5. Finally, section 1.6 presents separation theorems and Farkas lemma. In financial theory, these results allow to link the no-arbitrage assumption to the existence of a risk-neutral probability measure[1] in an economy with a finite number of states of nature.

1.1   Vector spaces : definitions and general properties

1.1.1   Definition and examples of vector spaces

Definition 1 *A **vector space** is a set E of elements, called **vectors**, that can be added (addition is denoted "+" as usual) and multiplied by real numbers (multiplication by a number is denoted "."). E satisfies the following properties:*

1) $(E, +)$ is a commutative group[2]

2) $\forall (\alpha, \beta) \in \mathbb{R}^2, \quad \forall u \in E$

$$\alpha(\beta.u) = (\alpha\beta).u \qquad (associativity)$$

---

[1]See Roger, P., Probability for Finance, 2010.

[2]It means that $+$ is associative, has an identity element denoted **0,** and any element u has an inverse denoted $-u$ satisfying $u + (-u) = \mathbf{0}$

3) $\forall (\alpha, \beta) \in \mathbb{R}^2, \quad \forall u \in E,$

$(\alpha + \beta).u = \alpha.u + \beta.u$ *(distributivity with respect to the addition in $\mathbb{R}$)*

4) $\forall \alpha \in \mathbb{R}, \quad \forall (u, v) \in E^2,$

$\alpha.(u + v) = \alpha.u + \alpha.v$ *(distributivity with respect to the addition in E.)*

5) $\forall u \in E, 1.u = u$

Remark : The identity element for addition is the null vector denoted **0** (in bold characters for the moment to avoid possible confusion with the real number 0).

Example 2 $\mathbb{R}$ *is a vector space, endowed with the usual addition and the usual multiplication. The above remark concerning the notation of **0** appears to be important here because the number 0 is simultaneously the real number 0, and the identity element of the addition for the vector space $\mathbb{R}$. We let the reader check that $\mathbb{R}$ satisfies the statements of definition 1.*

Example 3 *Let $\mathbb{R}^n$ denote the set of $n-uples$[3] $x' = (x_1, x_2, ..., x_n)$ where $x_i \in \mathbb{R}$ for any i; $\mathbb{R}^n$ is a vector space if addition is defined by:*

$$x + y = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ ... \\ x_n + y_n \end{pmatrix}$$

and the product by a scalar is defined by:

$$\alpha.x = \alpha. \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ ... \\ \alpha x_n \end{pmatrix}$$

where $\alpha \in \mathbb{R}$.

The vector space $\mathbb{R}^n$ is the natural generalization of the usual 2 and 3-dimensional spaces.

Example 4 *Let E be a vector space and $\mathcal{A}(E)$ be the set of mappings[4] from E to $\mathbb{R}$; $\mathcal{A}(E)$ is a vector space if addition and product by a scalar are defined as follows:*

$$\forall (f, g) \in \mathcal{A}(E), \forall u \in E, \quad \begin{cases} (f + g)(u) = f(u) + g(u) \\ (\alpha.f)(u) = \alpha f(u) \end{cases} \qquad (1.1)$$

Though these definitions seem intuitive, the space $\mathcal{A}(E)$ is much more complex than the vector space $\mathbb{R}^n$; in particular, a vector $f \in \mathcal{A}(E)$ cannot

---

[3]Without precision, x denotes a column vector. x' (with a prime) is the corresponding row vector called the transpose of x. These notations are consistent with the notations for matrices in part I of the book.

[4]Mappings have been defined in chapter 1 of part I of the book.

be described by a finite set of real numbers because it is a mapping from E to $\mathbb{R}$.

1.1.2   Vector subspaces

Definition 5 *Let E be a vector space and F a subset of E; F is a **vector subspace** of E if, for any $\alpha \in \mathbb{R}$ and any $v \in F$, $\alpha.v \in F$, and if conditions (2) to (5) of definition 1 are satisfied for F, when addition and multiplication by a real number ("+" and ".") are restricted to F.*

Definition 5 looks complex but its meaning is simple. F is a vector subspace of E if F is itself a vector space when it is endowed with the same addition of vectors and the same product by a real number (meaning that $\alpha.v$ should stay in F if $v \in F$ and $u + v$ is in F if u and v are in F).

The following proposition provides a simple criterion to check if a subset of E is a vector subspace.

Proposition 6 *Let E be a vector space and F a subset of E; F is a vector subspace of E if and only if:*

$$\forall (\alpha, \beta) \in \mathbb{R}^2, \quad \forall (u, v) \in F^2, \alpha.u + \beta.v \in F$$

Example 7 *Let $E = \mathbb{R}^3$ and F_1 a subset of E defined by:*

$$F_1 = \{x \in E \ / \ x_1 + x_2 + x_3 = 0\}$$

It is obvious to prove that if two vectors x and y in E satisfy the condition "the sum of their coordinates is zero", any combination $\alpha.x + \beta.y$ also satisfies the condition, (α, β) being a couple of real numbers. Therefore F_1 is a vector subspace of E. In the same way, consider the subset $F_2 = \{\mathbf{0}\}$ which contains only the vector $\mathbf{0}$. It is the smallest vector subspace of E and the only one containing a single vector.

Therefore, $\mathbf{0}$ *belongs to any vector subspace of* E *and any intersection of vector subspaces contains at least the vector* $\mathbf{0}$. *This remark is generalized in the following proposition.*

Proposition 8 *Any intersection of vector subspaces of* E *is a vector subspace of* E.

To illustrate this proposition, let $F_3 = \{x \in E \ / \ x_1 - 2x_2 + 3x_3 = 0\}$ and show that $F_1 \bigcap F_3$ is a vector subspace de $\mathbb{R}^3$. You can use proposition 6. On the opposite, show that $F_1 \bigcup F_3$ is not a vector subspace (Hint. choose $u^1 \in F_1$ and $u^3 \in F_3$ satisfying $u^1 + u^3 \notin F_1 \bigcup F_3$).

This latter question shows that, in general, the union of two vector subspaces is not a vector subspace (denoted V.S hereafter)

On the contrary, if we define F_{13} as follows:

$$F_{13} = \left\{ x \in \mathbb{R}^3 \ / \ x = y + z \text{ with } y \in F_1 \text{ and } z \in F_3 \right\}$$

then F_{13} is a vector subspace of $E = \mathbb{R}^3$. This remark is generalized in the proposition below.

Proposition 9 *Let* $F_1, ..., F_k$ *be* k *vector subspaces of a vector space* E *and* F *be defined by:*

$$F = \left\{ \begin{array}{c} x \in E \ / \ \exists (\alpha_1, .., \alpha_k) \in \mathbb{R}^k \text{ and } u^1 \in F_1, ..., u^k \in F_k \text{ such that} \\ x = \sum_{i=1}^{k} \alpha_i u^i \end{array} \right\}$$

F *is a V.S of* E, *called the **sum** of* $F_1, F_2, ..., F_k$. *We write:*

$$F = F_1 + ... + F_k \tag{1.2}$$

Proposition 9 does not say that $\alpha_1, .., \alpha_k$ and the vectors $u^1 \in F_1, ..., u^k \in F_k$ are uniquely defined for a given x. In general it is not the case and an easy counter-example is given by assuming $k = 2$ and $F_1 = F_2$.

If the decomposition is unique, we use the word "direct sum" as defined below.

Definition 10 *a) The **direct sum** of k vector subspaces $F_1, ..., F_k$ of E (if it exists), is a V.S such that any x in F can be written in a unique way as $x = \sum_{i=1}^{k} \alpha_i u^i$ where $(\alpha_1, .., \alpha_k) \in \mathbb{R}^k$ et $u^1 \in F_1, ..., u^k \in F_k$. We then note:*

$$F = \bigoplus_{i=1}^{k} F_i$$

*b) If E is the direct sum of two V.S F_1 and F_2, the two subspaces are said **supplementary**.*

The direct sum does not always exist because the decomposition of vectors is not always unique. Some subspaces F_i may have common vectors different from **0.** This intuition is formalized in the following proposition.

Proposition 11 *The direct sum of vector subspaces $F_i, i = 1, ..., k$ is properly defined when for any pair (i, j), $F_i \bigcap F_j = \{\mathbf{0}\}$.*

Example : Completing a financial market with option contracts

Let $E = \mathbb{R}^n$ and $x \in E$ defined by:

$$\forall i = 1, ..., n, \;\; x_i = i$$

Let $\mathbf{1}$ denote the vector in $\mathbb{R}^n$ with all coordinates equal to 1 and $y^k \in E$ defined by:

$$y^k = (y_i^k, i = 1, .., n) \text{ where } y_i^k = \max(x_i - k \; ; \; 0), k = 1, ..., n - 1$$

or equivalently:
$$y^k = \max(x - k\mathbf{1} \; ; \; 0)$$

F_k is the V.S containing all the vectors proportional to y_k. We then have[5]:

$$E = \bigoplus_{k=0}^{n-1} F_k$$

where, by convention $F_0 = \{\beta x, \beta \in \mathbb{R}\}$

This relation says that any vector z in E can be uniquely decomposed as follows:
$$z = \sum_{k=0}^{n-1} \alpha_k \max(x - k\mathbf{1} \; ; \; 0) \tag{1.3}$$

The financial interpretation of this example is the following. x denotes the payoffs of a financial security (a stock or an index for example) which pays $1, 2, ...$or n depending on the state of nature that occurs at the final date[6] (there is only one future date T). The vectors y^k are payoffs of **call**

---

[5]The proof is left as an exercise.

[6]Our reasoning is valid as soon as payoffs x_i are different in different states. Choosing

options on x with a strike price k. A call option gives the right (and not the obligation) to its holder to buy asset x at a price k at date T.

Of course, the holder of the option buys the asset x if $x > k$ and the final net cash-flow is $x - k$. But if $x < k$ the holder of the option contract does not exercise the contract and no cash-flow is exchanged at date T. The payoff is then equal to 0.

Remark that for $k = 0$, F_0 is the V.S of vectors proportional to x. In fact there are only $n - 1$ option contracts with exercise prices $k = 1, ..., n - 1$. The relationship 1.3 shows that any financial security can be written as a portfolio composed of x and the $n - 1$ option contracts. A financial market satisfying this property is said complete. More details on this financial example can be found in our companion book *Probability for Finance*[7].

1.1.3   Basis and dimension of a vector space

In the previous example, we have shown that it is possible to construct any vector of $\mathbb{R}^n$ by combining the reference vectors x and y^k, $k = 1, .., n - 1$. It is time to properly define what means "combining" and to specify the conditions under which a subset of vectors generates a given vector space.

Spanning sets of vectors

Definition 12 *Let $u^1, u^2, .., u^k$ be vectors in E and $\alpha_1, ..., \alpha_k$ be real numbers; a **linear combination** of the $u^j, j = 1, ..., k$ with coefficients α_j is the vector v defined by:*

$$v = \sum_{j=1}^{k} \alpha_j u^j$$

---

payoffs equal to $1, 2, ..., n$ is not crucial but simplifies the example.

[7]The idea of completing a market by traded options was initially developed by Steve Ross in a paper entitled "Options and Efficiency", published in the Quarterly Journal of Economics in 1976.

Linear combinations are really fundamental tools in financial models because, as we saw in the previous example, v is the payoff of a portfolio when the payoffs of individual securities are the vectors u^j and the α_j denote the quantities of assets.

Proposition 13 *Let $u^1, u^2, .., u^k$ be vectors in E, and F be the set of linear combinations of vectors $u^j, j = 1, ...k$, that is:*

$$F = \left\{ x \in E /\ \exists \alpha \in \mathbb{R}^k,\ x = \sum_{j=1}^{k} \alpha_j u^j \right\}$$

then F is a vector subspace of E.

In financial terms, F is the subspace of portfolios that can be built with primary securities $u^1, ..., u^k$. This result means that, using proposition 6, a linear combination of two portfolios is a portfolio.

Definition 14 *Let $u^1, u^2, .., u^k$ be vectors in E; they are **linearly dependent** if there exist coefficients $\alpha' = (\alpha_1, ..., \alpha_k)$ with $\alpha \neq \mathbf{0}$ such that:*

$$\sum_{j=1}^{k} \alpha_j u^j = 0 \tag{1.4}$$

The set $u^1, u^2, .., u^k$ is called a linearly dependent family.

This definition says that any vector u^j in the family with a weight $\alpha_j \neq 0$, can be written as a linear combination of the other $k - 1$ vectors.

Moreover, if a family of k vectors is linearly dependent, one can add any number of new vectors to the family, it stays linearly dependent. In fact, it is sufficient to give null weights to the new vectors to find the same kind of linear combination.

Remark 15 *If a vector in E represents the payoffs of a financial security in the different states of nature, a linear combination is then a vector of*

portfolio payoffs. If a family of vectors is linearly dependent it means that you can build a portfolio generating a 0 payoff in each state. In financial terms, one of the assets is a hedge for a portfolio of the other assets. The reader can easily imagine that such a situation has some consequences on the prices of these assets, the intuitive idea being: "a portfolio that pays nothing (in all states of nature) should cost nothing". We come back to this approach of arbitrage pricing at the end of the chapter.

Linearly independent vectors and basis of a vector space

Definition 16 *Let $u^1, u^2, .., u^k$ be a set of vectors in E; they are **linearly independent** if they are not linearly dependent. The following implication is then true.*

$$\sum_{j=1}^{k} \alpha_j u^j = 0 \Rightarrow \alpha = 0 \qquad (1.5)$$

In particular, two vectors u^1 and u^2 are linearly independent if there does not exist a real number β satisfying $u^2 = \beta u^1$. The two vectors cannot be colinear if they are linearly independent.

Example 17 *Let $E = \mathbb{R}^3$ and u^1, u^2, u^3 be three vectors defined by:*

$$
u^1 = \begin{pmatrix} 1 \\ 1 \\ a \end{pmatrix} ; u^2 = \begin{pmatrix} 2 \\ a \\ 3 \end{pmatrix} ; u^3 = \begin{pmatrix} a \\ 4 \\ -1 \end{pmatrix}
$$

What are the conditions on the number a under which these three vectors are linearly independent?

We need to solve the following equations:

$$
\begin{aligned}
\alpha_1 + 2\alpha_2 + a\alpha_3 &= 0 \\
\alpha_1 + a\alpha_2 + 4\alpha_3 &= 0 \\
a\alpha_1 + 3\alpha_2 - \alpha_3 &= 0
\end{aligned}
$$

and find if there are non zero solutions for $\alpha' = (\alpha_1; \alpha_2; \alpha_3)$

The first equation leads to

$$
\alpha_1 = -2\alpha_2 - a\alpha_3 \tag{1.6}
$$

We replace α_1 by its expression in the two other equations. It writes:

$$
\begin{aligned}
(a - 2)\alpha_2 + (4 - a)\alpha_3 &= 0 \\
(3 - 2a)\alpha_2 - (1 + a^2)\alpha_3 &= 0
\end{aligned} \tag{1.7}
$$

We can now write α_2 as a function of α_3 to obtain:

$$
\alpha_2 = \frac{(a - 4)\alpha_3}{(a - 2)} \tag{1.8}
$$

In equation (1.8) a must be different from 2. If $a = 2$, it is obvious that $\alpha_3 = 0$ in the first equation of system (1.7). It implies $\alpha_2 = 0$ in the second equation and finally $\alpha_1 = 0$, showing that the three vectors are linearly independent.

Assume now that $a \neq 2$; using equation (1.8) and replacing α_2 by its value in the second equation of system (1.7) gives:

$$(3 - 2a)\frac{(a - 4)\alpha_3}{(a - 2)} - (1 + a^2)\alpha_3 = 0$$

For this equation to be satisfied with $\alpha_3 \neq 0$, we need:

$$(3 - 2a)\frac{(a - 4)}{(a - 2)} - (1 + a^2) = 0$$

or, equivalently:

$$
\begin{aligned}
(3 - 2a)(a - 4) - (a - 2)(1 + a^2) &= 0 \\
-a^3 + 10a - 10 &= 0
\end{aligned}
$$

This equation has, at least, one solution[8]; the three vectors are then not linearly independent.

Remark 18 *For linearly independent families, we have a property similar (or more precisely, symmetric) to the one obtained for linearly dependent families. If k vectors are linearly independent, any subset of these k vectors is also a linearly independent family.*

Definition 19 *A family $(u^1, u^2, .., u^k)$ of vectors in E is a **spanning family** if any $x \in E$ can be written as a linear combination of $(u^1, u^2, .., u^k)$.*

$$\forall x \in E, \ \exists \alpha \in \mathbb{R}^k \ \ tel \ que \ x = \sum_{j=1}^{k} \alpha_j u^j$$

---

[8]In chapter 2 of part I of the book (devoted to limits and continuity), we saw how this result can be obtained. Intuitively, we observe that if a is positive and large, the left hand side (LHS) of the equation is negative due to the term $-a^3$. On the opposite, if a is negative and large in absolute value, this same LHS is positive. Therefore, there is at least an a for which this LHS is equal to 0 because this third-degree polynomial is a continuous function of a.

In example of subsection 1.1.2 dealing with option contracts, we showed that x and call options on x denoted y^k, $k = 1, ..., n-1$ constitute a spanning family of $\mathbb{R}^n$. When a family $\mathcal{U}$ of vectors is a spanning family of a vector space E, it is clear that any family $\mathcal{U}^*$ containing $\mathcal{U}$ is also a spanning family of E. However, if $\mathcal{U}^* \supset \mathcal{U}$ and $\mathcal{U}^* \neq \mathcal{U}$ then $\mathcal{U}^*$ is a linearly dependent family.

The natural question appearing now is: what is the "smallest" spanning family of a given vector space?

Definition 20 *A family $\mathcal{U}$ of vectors in E is a **basis** of E if $\mathcal{U}$ is a spanning and linearly independent family of E.*

When $\mathcal{U}$ is linearly dependent and spans E, it is always possible to find, for a given vector x, several linear combinations in $\mathcal{U}$ that are equal to x. In fact, assume that:

$$x = \sum_{i=1}^{n} \alpha_i u^i \tag{1.9}$$

with $\mathcal{U} = \{u^1, ..., u^n\}$. If $\mathcal{U}$ is a linearly dependent family, we can write $u^1 = \sum_{i=2}^{n} \beta_i u^i$. Replacing u^1 by its value in equation 1.9 leads to:

$$x = \sum_{i=2}^{n} (\alpha_i + \alpha_1 \beta_i) u^i$$

It is a second linear combination of vectors of $\mathcal{U}$ which is equal to x.

But if $\mathcal{U}$ is linearly independent, the decomposition of x is unique. This leads to the following proposition

Proposition 21 *A family $\mathcal{U} = \{u^1, ..., u^n\}$ is a basis of a vector space E if and only if any vector $x \in E$ can be decomposed in a unique way as a linear combination of vectors of $\mathcal{U}$.*

Proof. If $\mathcal{U}$ is a basis, any vector x can be written as a linear combination of vectors of $\mathcal{U}$. Assume that there exist two decompositions as follows:

$$x = \sum_{i=1}^{n} \alpha_i u^i$$
$$x = \sum_{i=1}^{n} \gamma_i u^i$$

Substracting the second equation from the first one gives:

$$0 = x = \sum_{i=1}^{n} (\alpha_i - \beta_i) u^i$$

Equation 1.5 then implies $\alpha_i = \gamma_i$ for all i.

To prove the sufficient condition, proceed as follows. If any $x \in E$ can be written as a linear combination of vectors of $\mathcal{U}$, it means that $\mathcal{U}$ spans E. But we showed that if $\mathcal{U}$ is linearly dependent, there exist several linear combinations to obtain x. Consequently, $\mathcal{U}$ is linearly independent if the combination is unique. Therefore, $\mathcal{U}$ is linearly independent and spans E, it is then a basis. ∎

For any vector x and any basis $\mathcal{U}$, x is characterized by coefficients $\alpha' = (\alpha_1; ..; \alpha_n)$ satisfying $x = \sum_{i=1}^{n} \alpha_i u^i$. These coefficients do depend on the considered basis $\mathcal{U}$. The most simple basis in $E = \mathbb{R}^n$ is called the canonical basis, denoted $e^1, ..., e^n$, where the vectors e^i are defined by :

$$e^1 = \begin{pmatrix} 1 \\ 0 \\ ... \\ 0 \end{pmatrix} ; e^2 = \begin{pmatrix} 0 \\ 1 \\ ... \\ 0 \end{pmatrix} ; ... e^n = \begin{pmatrix} 0 \\ 0 \\ ... \\ 1 \end{pmatrix}$$

e^i has all its components equal to 0 except the i-th which is equal to 1. Therefore, for any n-tuple $x' = (x_1, ..., x_n)$, we obtain the decomposition:

$$x = \sum_{i=1}^{n} x_i e^i$$

When E is interpreted as the set of all possible portfolio payoffs, the vectors $e^1, ..., e^n$ are financial securities called **Arrow-Debreu securities**, or pure contingent securities. They pay one unit in a given state of nature and nothing in all the other states.

Definition 22 *A vector space E is finite-dimensional if there exists a spanning family composed of a finite number of vectors. In this case, the dimension of E is the number of vectors[9] in a basis of E.*

This definition characterizes properly the dimension of a vector space only if all the bases of a given space have the same number of vectors. The proof of this statement is left to the reader as an exercise (hint: assume it is not true and exhibit a contradiction). From this remark, we can also deduce the following proposition.

Proposition 23 *Let F denote a V.S of a finite dimensional E with $F \neq E$. Then $\dim(F) < \dim(E)$.*

Definition 22 shows that the dimension of a vector space cannot be identified by the number of vectors in a spanning family but only by the number of elements in a basis (linearly independent spanning family). Therefore, in a family of vectors included in a finite-dimensional space E, there exists a maximum number of linearly independent vectors which is equal to the dimension of the space.

---

[9]By convention, the vector space containing only the null vector is 0-dimensional.

Proposition 24 *Let $\mathcal{U} = (u^1, ..., u^n)$ be a basis of E and x denote a vector in E; the family $\mathcal{U}_x = (u^1, ..., u^n, x)$ is linearly dependent.*

Proof. x can be written $\sum_{i=1}^{n} x_i u^i$ since $\mathcal{U}$ is a basis. Therefore, we can find $(\alpha_1, ..., \alpha_n, \alpha_{n+1})$ with at least one of these cefficients different from 0 such that:

$$\sum_{i=1}^{n} \alpha_i u^i + \alpha_{n+1} x = 0$$

It is enough that $\alpha_i = x_i$ and $\alpha_{n+1} = -1$. Definition 14 implies that $\mathcal{U}_x$ is linearly dependent. ∎

Let $\mathcal{U}$ be a set of financial securities and x a new financial contract introduced on the market. The above proposition shows that the payoffs of the new asset x can be replicated by a portfolio of securities [10] of $\mathcal{U}$. In this situation, x is called a redundant asset. Later on, we analyze the consequences of this remark on the evaluation of financial securities.

---

[10]Replication means that the future payoffs of x are identical to the future payoffs of the "replicating" portfolio.

Definition 25 *Let $\mathcal{U}$ be a family of vectors in E; the **rank** of $\mathcal{U}$, denoted $rk(\mathcal{U})$, is the maximum number of vectors in $\mathcal{U}$ that are linearly independent.*

In proposition 13 we showed that the set of linear combinations of a subset of vectors of E is a V.S of E. Moreover, the rank of $\mathcal{U}$ when we add to $\mathcal{U}$ a linear combination of vectors of $\mathcal{U}$ does not change. Consequently, we obtain the following proposition.

Proposition 26 *1)Let $\mathcal{U}$ be a family of vectors in E; the set of linear combinations of vectors in $\mathcal{U}$ is a V.S of dimension $p = rk(\mathcal{U})$.*

2) Let v denote a vector which is not a linear combination of vectors of $\mathcal{U}$; we then have:

$$rk(\mathcal{U} \bigcup \{v\}) = rk(\mathcal{U}) + 1$$

To illustrate the second part of the proposition, consider a vector $x \in E$ defined by :

$$x = \begin{pmatrix} 1 \\ 2 \\ ... \\ n \end{pmatrix}$$

Denote y the vector defined by $y = \max(x - k\mathbf{1};0)$ with $k \in \mathbb{N}$, and $\mathbf{1}$ the vector in E with all coordinates equal to 1. As far as $n > k > 0$, y and x are linearly independent, that is not colinear[11]. Therefore $rk(\{x, y\}) = rk(\{x\}) + 1 = 2$. One more time, if x is interpreted as the possible future prices of a stock, y is the vector of future payoffs of a call option on x with exercise price k. We observe that portfolios (linear combinations) based on x and y span a two-dimensional space when asset x only spans a 1-dimensional space. It explains why (at least in theory) options are able to improve the allocation of risk and resources in the economy. This property had been already mentioned in example of subsection 1.1.2.

---

[11] When two vectors are linearly dependent, they are said "colinear".

1.2   Linear mappings

In finance models, linear mappings are fundamental because of the large number of applications in which they are involved. It is especially the case in the theory of valuation based on the no-arbitrage assumption. In fact, a fundamental result of this approach is that when the market is free from arbitrage opportunities, the mapping linking future cash-flows to current prices is linear.

1.2.1   Definitions and notations

Definition 27 *Let E_1 and E_2 be two vector spaces; a mapping f from E_1 to E_2 is **linear** if:*

1) $\forall (u, v) \in E_1 \times E_1, f(u + v) = f(u) + f(v)$

2) $\forall \alpha \in \mathbb{R}, \ \forall u \in E, \ \ f(\alpha.u) = \alpha.f(u)$

As before, it is worth to notice that, in $f(u + v)$, the "+" sign refers the addition of vectors in the space E_1 but the same sign "+" in $f(u) + f(v)$ refers to the addition in E_2. Remember that the two additions may be quite different, depending on the characterizations of E_1 and E_2. The same remark should be done for the multiplication by a real number, even if the difference is less striking.

Linearity of a mapping f means that the image of a linear combination of vectors in E_1 is the linear combination of images in E_2 with the same coefficients. The following proposition formalizes this remark. It is sometimes used as the definition of a linear mapping.

Proposition 28 *A mapping $f : E_1 \to E_2$ is linear if and only if for any family $(u^1, ..., u^p)$ of vectors in E_1 and any p-tuple $(\alpha_1, ...\alpha_p) \in \mathbb{R}^p$, the following equality is satisfied:*

$$f\left(\sum_{i=1}^{p} \alpha_i u^i\right) = \sum_{i=1}^{p} \alpha_i f(u^i)$$

Remark 29 *a)* *The definition of linearity implies immediately* $f(\mathbf{0}_{E_1}) = \mathbf{0}_{E_2}$ *if* $\mathbf{0}_{E_1}$ *and* $\mathbf{0}_{E_2}$ *denote the null vectors of the two spaces. Moreover, for any* $u \in E_1$, $f(-u) = -f(u)$.

b) If $\mathcal{U} = \{u^1, ..., u^p\}$ *denotes a linearly dependent family in* E_1 *then* $f(\mathcal{U}) = \{f(u^1), ..., f(u^p)\}$ *is a linearly dependent family in* E_2. *On the opposite, if* $\mathcal{U} = \{u^1, ..., u^p\}$ *is a linearly independent family in* E_1, $f(\mathcal{U})$ *is not always a linearly independent family in* E_2 *(the proof is left to the reader; it is sufficient to consider the mapping* $u \rightarrow f(u) = u_1 \mathbf{1}$ *where* u_1 *is the first coordinate of* u *and* $\mathbf{1}$ *is, as usual, the vector with all coordinates equal to* 1*).*

1.2.2 Kernel and image of a linear mapping

Definition 30 *1)Let* f *denote a linear mapping from* E_1 *to* E_2; *the* **kernel** *of* f *denoted* $Ker(f)$ *is the set of vectors* $u \in E_1$ *satisfying* $f(u) = 0$.

2) The image of f, *denoted* $\mathrm{Im}(f)$ *is the subset of* E_2 *defined by :*

$$\mathrm{Im}(f) = \{y \in E_2 \ / \ \exists \ x \in E_1 \ such \ that \ y = f(x)\}$$

$Ker(f)$ is then a subset of E_1 equal to the reciprocal image of the null vector in E_2 (sometimes written $f^{-1}(\mathbf{0}_{E_2})$). On the contrary $\mathrm{Im}(f)$ is a subset of E_2, sometimes written $f(E_1)$ because it contains all vectors in E_2 that can be written $f(u)$ with $u \in E_1$.

Proposition 31 $Ker(f)$ *and* $\mathrm{Im}(f)$ *are V.S of* E_1 *and* E_2 *respectively.*

Proof. Using proposition 6, it is enough to show, for the kernel $Ker(f)$:

$$\forall (u, v) \in Ker(f) \times Ker(f), \forall (\alpha, \beta) \in \mathbb{R}^2, \alpha.u + \beta.v \in Ker(f)$$

The linearity of f implies that[12]:

$$f(\alpha.u + \beta.v) = \alpha.f(u) + \beta.f(v) = \alpha.0 + \beta.0 = 0$$

For the image, we have to prove that:

$$\forall (x,y) \in \mathrm{Im}(f) \times \mathrm{Im}(f), \forall(\alpha,\beta) \in \mathbb{R}^2, \alpha.x + \beta.y \in \mathrm{Im}(f)$$

Let u and v be two vectors in E_1 such that $f(u) = x$ and $f(v) = y$. We can write:

$$\alpha.x + \beta.y = \alpha.f(u) + \beta.f(v) = f(\alpha.u + \beta.v)$$

Therefore $\alpha.x + \beta.y$ is the image of $\alpha.u + \beta.v$ through f, implying $\alpha.x + \beta.y \in \mathrm{Im}(f)$. ∎

This proposition shows in particular that the image of E by f is a V.S of E_2. With the same technique of proof as before, we can demonstrate the following proposition.

---

[12]We come back here to standard notations where 0 denotes the null vector in either space, assuming that the reader is now able to identify the reference space if necessary.

Proposition 32 *Let F_1 be a V.S of E_1; $f(F_1)$ is a V.S of E_2.*

Remmember that $\{0\}$ is a V.S of E_2 and the reciprocal image of $\{0\}$ by f is the V.S $Ker(f)$. This remark can be generalized as follows.

Proposition 33 *If F_2 is a V.S of E_2, $f^{-1}(F_2)$ is a V.S of E_1.*

An important property is to characterize the relationship between the kernel dimension and the properties of f. In particular, the question is to know if a vector $u \neq 0$ can satisfy $f(u) = 0$

Proposition 34 *If f is injective then $Ker(f) = \{0\}$*

Proof. f injective means that $y \neq x \implies f(y) \neq f(x)$. As f is a linear mapping, this implication writes:

$$y - x \neq 0 \Rightarrow f(y - x) \neq 0$$

and therefore $Ker(f) = \{0\}$. The reciprocal goes as follows. If $Ker(f) = \{0\}$ and if there exist x and y, $x \neq y$ satisfying $f(x) = f(y)$, an obvious contradiction arises because $f(x - y) = 0$, meaning that $x - y \in Ker(f)$. ∎

Proposition 35 *If f is surjective then $\mathrm{Im}(f) = E_2$*

Proof. This result is obvious because f surjective means that any vector in E_2 has a reciprocal image in E_1 ∎

Definition 36 *A bijective linear mapping from E_1 to E_2 is called an **isomorphism**.*

This notion of isomorphism is fundamental when it comes to associate a space of linear mappings to a space of matrices, or a vector space to its dual space, as we will see in the next section.

Proposition 37 *Two vector spaces E_1 and E_2 are **isomorphic**[13] if and only if their dimensions are equal.*

Proof. Denote n and p the respective dimensions of E_1 and E_2; let $\mathcal{U}$ and $\mathcal{V}$ be bases of these two spaces and f be a bijective linear mapping from E_1 to E_2.

We first show that "f injective" is equivalent to "$f(u^1), ..., f(u^n)$ are linearly independent".

If f is injective, $Ker(f) = \{0\}$. Therefore, any linear combination $\sum_{i=1}^{n} x_i u^i$ satisfies:

$$\sum_{i=1}^{n} x_i u^i = 0 \Leftrightarrow x_i = 0 \text{ for any } i \tag{1.10}$$

In this case,

$$f\left(\sum_{i=1}^{n} x_i u^i\right) = \sum_{i=1}^{n} x_i f(u^i) = 0$$

which shows that the vectors $f(u^i)$ are linearly independent. The reciprocal goes as follows: if the $f(u^i), i = 1, ..., n$ are linearly independent, we can write:

$$\sum_{i=1}^{n} x_i f(u^i) = 0 \Leftrightarrow x_i = 0 \text{ for any } i \tag{1.11}$$

but the linearity of f implies that $\sum_{i=1}^{n} x_i u^i \in Ker(f)$. Relation (1.11) then implies $Ker(f) = 0$ and f is injective. It follows directly that $n \leq p$.

As f is also surjective, $\text{Im}(f) = E_2$ and the rank of the family of vectors $f(u^1), ..., f(u^n)$ is p, meaning that $n \geq p$.

We show now that if E_1 and E_2 have equal dimensions, they are isomorphic. The basis $\mathcal{V}$ with p vectors defines a linear mapping f from E_1 to E_2 such that $\mathcal{V}$ is the image of a family $\mathcal{W}$ of E_1. We then have $rk(\mathcal{W}) = n = \dim(E_1)$ proving that f is injective.

---

[13] "isomorphic" means that there exists an isomorphism between E_1 and E_2.

But we also know that $\dim(E_1) = \dim(E_2)$; therefore f is also surjective. As a result, f is bijective and E_1 and E_2 are isomorphic. $\blacksquare$

1.2.3 The space of linear mappings

The set of linear mappings defined on a vector space E_1 and taking values in a vector space E_2 is denoted by $\mathcal{L}(E_1, E_2)$, or simply $\mathcal{L}$ when no confusion is possible.

In example 4, we showed that the set of mappings defined on a vector space E and taking values in $\mathbb{R}$ is a vector space. The property is still valid if $\mathbb{R}$ is replaced by another vector space. Therefore, $\mathcal{L}(E_1, E_2)$ is a subset of $\mathcal{A}(E_1, E_2)$. As the elements in $\mathcal{L}$ are linear mappings, we have the following property.

Proposition 38 $\mathcal{L}$ *is a V.S of* $\mathcal{A}(E_1, E_2)$

Proof. Let $(f, g) \in \mathcal{L}^2$ and $(\alpha, \beta) \in \mathbb{R}$; for any couple of vectors (x, y) of $E_1 \times E_1$, we have:

$$
\begin{aligned}
(\alpha f + \beta g)(x + y) &= \alpha f(x + y) + \beta g(x + y) \\
&= \alpha f(x) + \alpha f(y) + \beta g(x) + \beta g(y) \\
&= (\alpha f + \beta g)(x) + (\alpha f + \beta g)(y)
\end{aligned}
$$

For any $x \in E_1$ and $\gamma \in \mathbb{R}$, we also have:

$$
\begin{aligned}
(\alpha.f + \beta.g)(\gamma x) &= \alpha f(\gamma x) + \beta g(\gamma x) \\
&= \gamma \alpha f(x) + \gamma \beta g(x) \\
&= \gamma (\alpha f + \beta g)(x)
\end{aligned}
$$

$\blacksquare$

One of the most common situations appears if E_1 is a general vector

space[14] and E_2 is $\mathbb{R}$. In this situation, the elements of $\mathcal{L}(E_1, \mathbb{R})$ are called **linear functionals (or one-forms)**; we will see later on that the mapping linking a function to its integral is a linear functional. In the same way, the expectation operator is a linear functional defined on a space on random variables. In finance models, the no-arbitrage assumption implies that the valuation operator mapping future cash-flows and today prices is a linear functional[15].

Definition 39 *Let E be a vector space; the set $\mathcal{L}(E, \mathbb{R})$ of linear functionals defined on E is called the **dual space** of E.*

When E is a finite-dimensional space of dimension n, its dual space satisfies the following property.

Proposition 40 *If $\dim(E) = n < +\infty$; $\dim\left(\mathcal{L}(E, \mathbb{R})\right) = \dim(E)$*

Proof. Let $(u^1, .., u^n)$ be a basis of E and $x \in E$ written as:

$$x = \sum_{i=1}^{n} x_i u^i \tag{1.12}$$

Consider $f^1, ..., f^n$, a set of linear functionals defined by:

$$\forall i = 1, ..., n; \forall x \in E, f^i(x) = x_i$$

The family $\mathcal{F} = (f^1, ..., f^n)$ spans $\mathcal{L}(E, \mathbb{R})$. In fact, for any given linear functional g, we have:

$$g(x) = \sum_{i=1}^{n} x^i g(u^i) = \sum_{i=1}^{n} g(u^i) f^i(x) = \left(\sum_{i=1}^{n} g(u^i) f^i \right)(x) \tag{1.13}$$

---

[14]In particular, E_1 may be an infinite-dimensional space of random variables in probability frameworks.

[15]A financial security is defined by the future cash-flows it generates. In general, financial securities can be represented as elements of a general vector space.

It shows that g can be written as a linear combination of the f^i.

We now show that $\mathcal{F}$ is a linearly independent family. The equality $\sum_{i=1}^{n} \alpha_i f^i = 0$ means:

$$\forall x \in E, \sum_{i=1}^{n} \alpha_i f^i(x) = 0$$

but, according to the definition of f^i, this equality is equivalent to:

$$\sum_{i=1}^{n} \alpha_i x_i = 0$$

For this equality to be satisfied by any x, it is necessary that all the α_i are equal to zero[16], and this ends the proof. ■

---

[16]The linear mapping f^i is called the **projection** on u^i.

1.3 Finite-dimensional spaces and matrices

In this section, we assume that the two spaces E_1 and E_2 we refer to are finite-dimensional. The notations are unchanged. E_1 (E_2) denotes a vector space of dimension n (p). $\mathcal{L}(E_1, E_2)$ is the space of linear mappings from E_1 to E_2.

1.3.1 Representation of a linear mapping by a matrix

Proposition 41 *Let $\mathcal{U} = (u^1, .., u^n)$ be a basis of E_1 and $\mathcal{V} = (v^1, v^2, ..., v^p)$ be a basis of E_2; any mapping $f \in \mathcal{L}(E_1, E_2)$ is completely defined by the family of vectors $f(u^1), ..., f(u^n)$ expressed in the basis $\mathcal{V}$.*

Proof. Let $x \in E_1$ such that $x = \sum_{i=1}^n x_i u^i$; $f(x)$ can be written:

$$f(x) = f\left(\sum_{i=1}^n x_i u^i\right) = \sum_{i=1}^n x_i f(u^i)$$

Therefore, if the images $f(u^i)$ of the vectors of the family $\mathcal{U}$ are known, it is possible to characterize the image of any vector x. Each vector $f(u^i)$ belongs to E_2, it is then a p-dimensional vector. The linear mapping f is then completely specified by $n \times p$ numbers equal to the coordinates of the n vectors $f(u^i)$, $i = 1, ..., n$. ■

If we denote $f(u^i)$ as follows:

$$f(u^i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ ... \\ a_{pi} \end{pmatrix}$$

we can introduce the following definition.

Definition 42 *The **matrix of the linear mapping** f, denoted $M_f(\mathcal{U}, \mathcal{V})$ is a $p \times n$ matrix whose columns are the vectors $f(u^i), i = 1, ..., n$.*

According to the above notation for $f(u^i)$, we have:

$$M_f\left(\mathcal{U}, \mathcal{V}\right) = \begin{bmatrix} a_{11} & a_{12} & .... & a_{1n} \\ a_{21} & ... & ... & ... \\ ... & ... & ... & ... \\ a_{p1} & ... & ... & a_{pn} \end{bmatrix} \tag{1.14}$$

The notation $M_f(\mathcal{U}, \mathcal{V})$ is cumbersome but it is used here to emphazise that the matrix representing f depends on the two bases on E_1 and E_2. Of course, in the next sections, we will simply write M_f when no confusion can be made.

The above remarks show that being given $\mathcal{U}$ and $\mathcal{V}$, the matrix M_f is linked to f. But more generally any $p \times n$ matrix defines a linear mapping from E_1 (of dimension n) to E_2 (of dimension p). The most usual case is the one where $\mathcal{U}$ and $\mathcal{V}$ are the canonical[17] bases of E_1 and E_2.

1.3.2   Compounding linear mappings

Consider three vector spaces E_1, E_2, E_3 with dimensions n, p, m and bases $\mathcal{U}, \mathcal{V}, \mathcal{W}$; let f denote a linear mapping from E_1 to E_2 and g a linear mapping from E_2 to E_3. In general we describe this sequence as follows:

$$E_1 \xrightarrow{\ f\ } E_2 \xrightarrow{\ g\ } E_3 \tag{1.15}$$

Compounding the mappings f and g aims at defining a new mapping from E_1 to E_3.

---

[17]Remember that the canonical basis is the basis for which vectors have all their coordinates equal to 0, except one which is equal to 1. For example in $\mathbb{R}^3$, the canonical basis is

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} ; u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} ; u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Definition 43 *The **compound mapping** of f and g is the mapping denoted $g \circ f$ from E_1 to E_3 defined by:*

$$\forall x \in E_1; \ g \circ f(x) = g\left[f(x)\right]$$

x writes $\sum_{i=1}^{n} x_i u^i$; but f and g are linear, so we have:

$$g \circ f(x) = g\left[f(x)\right] = g\left[\sum_{i=1}^{n} x_i f(u^i)\right] = \sum_{i=1}^{n} x_i g \circ f(u^i)$$

This equality shows that $g \circ f$ is a linear mapping.

The compounding of linear mappings is linked to the product of matrices. Denote M_f and M_g the matrices associated to the mappings f and g, defined as before. For any $x \in E_1$, $f(x) = M_f x$. The vector $f(x)$ belongs to E_2; as such it has p coordinates. Therefore, the image of $f(x)$ by g is obtained by a premultiplication of $f(x)$ by M_g. This leads to:

$$g \circ f(x) = M_g\left(M_f x\right) = M_g M_f x = M_{g \circ f}(x)$$

To calculate $M_{g \circ f}$, we apply successively f and g. Denote for example:

$$M_f = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & ... & ... & ... \\ ... & ... & ... & ... \\ a_{p1} & ... & ... & a_{pn} \end{pmatrix}$$

$$M_g = \begin{pmatrix} b_{11} & b_{12} & ... & b_{1p} \\ b_{21} & ... & ... & ... \\ ... & ... & ... & ... \\ b_{mp} & ... & ... & b_{mp} \end{pmatrix}$$

The generic element of $M_{g \circ f}$ is c_{ik} defined as:

$$c_{ik} = \sum_{j=1}^{p} b_{kj} a_{ji}$$

We then observe that $M_{g \circ f} = M_g M_f$, the product of M_g and M_f defined in part I of the book (chapter 4).

The following proposition is a special case of this relationship.

Proposition 44 *A matrix A associated to a linear mapping f is invertible if and only if f is a bijection. The matrix representing f^{-1} is the inverse of A denoted A^{-1}.*

In fact, if B is associated to f^{-1}, the relationship $AB = I_n$ means that $B = A^{-1}$. We obtain the following corollary.

Corollary 45 *1) Let A denote a (n, n) invertible matrix. For any $u \in \mathbb{R}^n$, the system of equations $Ax = u$ has a solution $x \in \mathbb{R}^n$ defined by $x = A^{-1}u$.*

2) If a matrix A is invertible, its columns are linearly independent vectors of $\mathbb{R}^n$.

We mention these two properties as corollaries but they are equivalent to proposition 44. When the columns of A are future payoffs of financial assets in the n states of nature, the columns of A^{-1} are the quantities of securities to be held to duplicate the Arrow-Debreu securities because $AA^{-1} = I_E$.

Example 46 *Discounting*

In chapter 4 of part I, we showed that a bank can create contracts paying a single future cash-flow by combining bonds of different maturities. We calculated the prices of the contracts using the prices of the bonds.

Suppose that the cash-flows of the three bonds are stored in a matrix M as follows:

$$M = \begin{pmatrix} 104 & 6 & 4 \\ 0 & 106 & 4 \\ 0 & 0 & 104 \end{pmatrix}$$

The price vector π is:

$$\pi = \begin{pmatrix} 99,5 \\ 100,4 \\ 99,6 \end{pmatrix}$$

Denote now f the linear mapping defined from $\mathbb{R}^3$ to $\mathbb{R}$ such that $f(M^j) = \pi_j$, were M^j is the j-th column of M (the cash-flows generated by the j-th bond) and π_j is the price of bond j[18]. The mapping f is represented by a vector denoted A :

$$A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

So we have:

$$100 \times a_1 + 0 \times a_2 + 0 \times a_3 = 99,5$$
$$6 \times a_1 + 106 \times a_2 + 0 \times a_3 = 100,4$$
$$4 \times a_1 + 4 \times a_2 + 104 \times a_3 = 99,6$$

This system can be written as:

$$M'A = \pi$$

But M is triangular with non-zero diagonal terms, it is then invertible. Consequently:

$$A = (M')^{-1}\pi$$

The elements in A are in fact the prices of zero-coupon bonds with respective maturities 1, 2 and 3 years.

---

[18]We assume here that this mapping is linear. In fact it is true when the market is arbitrage-free.

1.3.3   The vector space of matrices

Being given two bases $\mathcal{U}$ and $\mathcal{V}$ on the vector spaces E_1 and E_2 (of dimensions n and p), the set $\mathcal{M}_{pn}$ of matrices with p rows and n columns must have the same structure as the set $\mathcal{L}(E_1, E_2)$ of linear mappings from E_1 to E_2, that is a structure of vector space. This is the only way to make coherent the structure of vector space of $\mathcal{L}(E_1, E_2)$ with the operators (addition and product by a real number) on $\mathcal{M}_{pn}$. This relationship is formalized in the following proposition.

Proposition 47 $\mathcal{M}_{pn}$ and $\mathcal{L}(E_1, E_2)$ are isomorphic and the dimension of these two spaces is $p \times n$.

Proof.   If we prove these two spaces are isomorphic, we will be able to conclude that the dimensions are equal because of proposition 37.

The mapping which links $f \in \mathcal{L}(E_1, E_2)$ to M_f is bijective because M_f is defined by the images of the basis vectors of E_1.

Let us denote A_{ij} the matrix with all null elements, except the one on i-th row and the j-th column which is equal to 1. Any matrix $A = (a_{ij}, i = 1, ..., p; j = 1, .., n)$ can be decomposed in a unique manner as:

$$A = \sum_{i=1}^{p} \sum_{j=1}^{n} a_{ij} A_{ij}$$

Therefore the family $(A_{ij}, i = 1, ..., p; j = 1, .., n)$ is a basis of $\mathcal{M}_{pn}$ and has $p \times n$ elements. This ends the proof. ■

Example 48 *Without entering into technical details, consider a set of mutual funds, each fund being a portfolio of individual securities. In a model with one period and n states of nature, the future payoffs of the individual securities (let K be the number of traded securities) can be summarized in a matrix D with n rows and K columns. The mapping "portfolio", denoted f, from $\mathbb{R}^K$ to $\mathbb{R}^n$ associates a vector θ to a vector $f(\theta) = D\theta$ where the*

vector θ contains the quantities of securities in the portfolio and the vector $D\theta$ represents the future payoffs of the portfolio θ. Let now g be the mapping from $\mathbb{R}^n$ to $\mathbb{R}$ which links to any vector x a number $g(x) = \sum p_i x_i$ where $P = (p_i, i = 1, ..., n)$ is the vector of probabilities of occurrence of states $i = 1, ..., n$.

If x represents the payoffs of a portfolio in the different states of nature, $g(x)$ is the expected payoff of the portfolio. Consequently, the product $PD\theta$ is the expected payoff of the mutual fund θ. It can also be written $gof(\theta)$.

1.3.4   The special case of square matrices

A square matrix has, by definition, the same number of rows and columns. In particular, the product of two (n, n) square matrices is still a (n, n) square matrix. In other words, the result of the product stays in the same space $\mathcal{M}_n$. Consider three vector spaces E_1, E_2 and E_3, all of dimension n. The product $M_g M_f$ is a (n, n) matrix if f is a mapping from E_1 to E_2 and g a mapping from E_2 to E_3.

The most important case addressed in what follows is $E_1 = E_2 = E_3 = E$.

Definition 49 *A linear mapping from E to E is called an* **endomorphism.**

We just described the links between matrices and linear mappings. Of course, an endomorphism is represented by a matrix in $\mathcal{M}_n$. But we also know by chapter 2 of part I that if a mapping f is bijective, it has an inverse denoted f^{-1} and satisfying $f \circ f^{-1} = f^{-1} \circ f = i_E$ where i_E is the identity mapping of E defined by $i_E(x) = x$ for any $x \in E$.

i_E is obviously linear and it is easy to see that the matrix I_n represents i_E where I_n is defined by:

$$
I_n = \begin{bmatrix}
1 & 0 & ... & ... \\
0 & 1 & & \\
... & & ... & \\
... & & & 1
\end{bmatrix}
$$

The reader can check that $I_n x = x$ for any $x \in \mathbb{R}^n$. The matrix I_n is called the **identity matrix**. Finally, we also know that the matrix associated to the compound of two mappings is the product of the matrices of the two mappings involved in the compounding. From all these remarks, we deduce that if A and B are matrices representing f and f^{-1}, we have:

$$
AB = I_n
$$

Proposition 50 *If f is a bijective endomorphism of E represented by a matrix M_f, f^{-1} is represented by the inverse matrix M_f^{-1}.*

Determinants

Knowing if the determinant of a square matrix is zero allows to know if this matrix is invertible. In a more geometric approach, a zero determinant means that the columns (or rows) of a given matrix are linearly dependent.

Determinant of a (2,2) matrix   Denote x and y two vectors in $\mathbb{R}^2$. They are colinear if there exists $\alpha \in \mathbb{R}$ satisfying $y = \alpha x$. This equality is equivalent to:

$$
\begin{aligned}
y_1 &= \alpha x_1 \\
y_2 &= \alpha x_2
\end{aligned}
$$

From these two equations we deduce $x_1 y_2 - x_2 y_1 = 0$; on the contrary, if $x_1 y_2 - x_2 y_1 \neq 0$, the vectors x and y are linearly independent. If x and y are the two columns of a square matrix A, A is invertible if $x_1 y_2 - x_2 y_1 \neq 0$. This remark justifies the definition of the determinant of a $(2,2)$ matrix.

Definition 51 *Let A be a $(2,2)$ matrix with generic term $a_{ij}, i, j = 1, 2$. The **determinant** of A (denoted $\det(A)$) the number:*

$$
\det(A) = a_{11}a_{22} - a_{12}a_{21}
$$

Determinants of larger matrices are defined by induction. The determinant of a (n, n) matrix is a function of determinants of $(n-1, n-1)$ matrices.

The general case

Definition 52 *1) Let A be a (n, n) matrix. Let D_{ij} be the determinant of the matrix deduced from A by deleting the i-th row and the j-th column of A. The (i, j)-th **cofactor** of A is the number $C_{ij} = (-1)^{i+j} D_{ij}$.*

*2) The i-th **principal minor** of A is the (i, i) matrix obtained by deleting the last $n - i$ rows and columns of A.*

Part (1) helps in calculating the determinant of A as shown in the following definition. Part (2) will prove useful to characterize positive (negative) definite matrices later on in this chapter

Definition 53 *Let A be a (n, n) matrix. The **determinant** of A is defined as follows:*

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij}$$

for any i between 1 and n.

The determinant of a matrix A is also usually denoted as follows (the matrix is placed between two vertical bars):

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & & & a_{2n} \\ ... & & & ... \\ a_{n1} & ... & ... & a_{nn} \end{vmatrix}$$

Example 54 *Let A be defined by:*

$$A = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 4 & 2 \\ 2 & 6 & 4 \end{pmatrix}$$

Definition 52 for $i = 1$ gives the following development.

$$\begin{aligned} \det(A) &= 3 \begin{vmatrix} 4 & 2 \\ 6 & 4 \end{vmatrix} - 1 \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} + 2 \begin{vmatrix} 1 & 4 \\ 2 & 6 \end{vmatrix} \\ &= 3 \times (4 \times 4 - 6 \times 2) - 1 \times (1 \times 4 - 2 \times 2) + 2(1 \times 6 - 2 \times 4) \\ &= 8 \end{aligned}$$

Suppose now $i = 2$. We obtain:

$$
\begin{aligned}
\det(A) &= -1 \begin{vmatrix} 1 & 2 \\ 6 & 4 \end{vmatrix} + 4 \begin{vmatrix} 3 & 2 \\ 2 & 4 \end{vmatrix} - 2 \begin{vmatrix} 3 & 1 \\ 2 & 6 \end{vmatrix} \\
&= -1 \times (4 - 12) + 4 \times (12 - 4) - 2 \times (18 - 2) \\
&= 8
\end{aligned}
$$

Of course, the result is the same. It is not a proof, but the proof itself is cumbersome and uninteresting on a practical point of view, so we omit it.

One of the main results concerning determinants is related to products and transposition of matrices.

Proposition 55 *1) Let A and B two (n,n) matrices. We have:*

$$\det(AB) \ = \ \det(A)\det(B)$$
$$\det(A^T) \ = \ \det(A)$$

2) If a matrix B is deduced from a (n,n) matrix A by swapping two rows or two columns, the determinants of the two matrices satisfy:

$$\det(B) = -\det(A)$$

3) A square matrix A is invertible if and only if its determinant is different from zero. If $\det(A) \neq 0$, the inverse matrix A^{-1} writes:

$$A^{-1} = \frac{1}{\det(A)}C^T \tag{1.16}$$

where $C = (C_{ij}, i, j = 1, .., n)$ is the cofactor matrix. Moreover, $\det(A^{-1}) = \frac{1}{\det(A)}$.

Proposition 55 gives a way to calculate determinants but this method is not the most numerically efficient.

Equality 1.16 is illustrated in the following example 56.

Example 56 *Let A denote the $(3,3)$ matrix:*

$$A = \begin{pmatrix} 3 & 2 & 5 \\ 4 & 6 & 1 \\ 2 & 3 & 2 \end{pmatrix}$$

The determinant of A is calculated as follows:

$$\det(A) = 3\begin{vmatrix} 6 & 1 \\ 3 & 2 \end{vmatrix} - 2\begin{vmatrix} 4 & 1 \\ 2 & 2 \end{vmatrix} + 5\begin{vmatrix} 4 & 6 \\ 2 & 3 \end{vmatrix} = 15$$

Applying relation 1.16 leads to:

$$A^{-1} = \frac{1}{15} \begin{pmatrix} 9 & 11 & -28 \\ -6 & -4 & 17 \\ 0 & -5 & 10 \end{pmatrix}$$

If we apply the same technique to A^T, the cofactor matrix of A^T is the transpose of the cofactor matrix of A. Therefore we obtain:

$$\left(A^T\right)^{-1} = \frac{1}{15} \begin{pmatrix} 9 & -6 & 0 \\ 11 & -4 & -5 \\ -28 & 17 & 10 \end{pmatrix}$$

1.3.5 Changing the basis

Matrix of a linear mapping after a basis change

A linear mapping f defined on $\mathbb{R}^n$, endowed with a basis $\mathcal{U} = (u_1, ..., u_n)$, and taking values in $\mathbb{R}^m$, endowed with a basis $\mathcal{V} = (v_1, ..., v_m)$, is represented by a matrix $M_f(\mathcal{U}, \mathcal{V})$. As mentioned before in definition 42, the notation $M_f(\mathcal{U}, \mathcal{V})$ recalls that M_f depends on the two bases. In particular, the columns of M_f are the images of vectors of $\mathcal{U}$ by f, expressed in the basis $\mathcal{V}$.

It turns out that a modification of one of the two bases changes the matrix M_f. Denote $\mathcal{W}$ a second basis of E and P the matrix having in columns the vectors of $\mathcal{W}$, expressed in the initial basis $\mathcal{U}$. This matrix will be called a **change-of-basis matrix** from basis $\mathcal{U}$ to basis $\mathcal{W}$.

We can show the following proposition.

Proposition 57 *Let v be a vector of $\mathbb{R}^n$ with coordinates $x^T = (x_1, x_2, ..., x_n)$ in basis $\mathcal{U}$ and $y^T = (y_1, y_2, ..., y_n)$ in basis $\mathcal{W}$. We then have:*

$$x = Py \text{ and } y = P^{-1}x$$

$M_f(\mathcal{W}, \mathcal{V})$ *is given as follows:*

$$M_f(\mathcal{W}, \mathcal{V}) = P^{-1} M_f(\mathcal{U}, \mathcal{V}) P$$

Example 58 *Let $M_f(\mathcal{U}, \mathcal{V})$ and x be defined by:*

$$M_f(\mathcal{U}, \mathcal{V}) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 4 \\ 2 & 3 & 1 \end{pmatrix} \qquad x = \begin{pmatrix} 4 \\ 6 \\ 1 \end{pmatrix}$$

Assume that $\mathcal{U}$ is the canonical basis of $\mathbb{R}^3$ and define $\mathcal{W}$ by:

$$w^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad w^2 = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} \qquad w^3 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

The image $f(x)$ of vector x (in basis $\mathcal{U}$) is given by:

$$f(x) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 4 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \\ 1 \end{pmatrix} = \begin{pmatrix} 17 \\ 10 \\ 27 \end{pmatrix}$$

The matrix P writes:

$$P = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

The inverse of P is calculated using the cofactors (equation 1.16). We obtain:

$$P^{-1} = \frac{1}{7} \begin{pmatrix} 3 & 4 & -6 \\ -1 & 1 & 2 \\ 2 & -2 & 3 \end{pmatrix}$$

We deduce from this formulation of P^{-1}:

$$P^{-1}M_f(\mathcal{U}, \mathcal{V}) \;=\; \frac{1}{7}\begin{pmatrix} 3 & 4 & -6 \\ -1 & 1 & 2 \\ 2 & -2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 4 \\ 2 & 3 & 1 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} -9 & -8 & 13 \\ 3 & 5 & 5 \\ 8 & 11 & -3 \end{pmatrix}$$

$$P^{-1}M_f(\mathcal{U}, \mathcal{V})P \;=\; \frac{1}{7}\begin{pmatrix} -9 & -8 & 13 \\ 3 & 5 & 5 \\ 8 & 11 & -3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \\ 1 & 3 & 0 \\ 0 & 2 & 1 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} -17 & 2 & -5 \\ 8 & 25 & 11 \\ 19 & 27 & 13 \end{pmatrix}$$

Therefore, in the new basis $\mathcal{W}$, x writes :

$$x = \frac{1}{7}\begin{pmatrix} 3 & 4 & -6 \\ -1 & 1 & 2 \\ 2 & -2 & 3 \end{pmatrix}\begin{pmatrix} 4 \\ 6 \\ 1 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 30 \\ 4 \\ -1 \end{pmatrix}$$

The following section studies the case where $M_f(\mathcal{W}, \mathcal{V})$ is diagonal when $\mathcal{W}$ and $\mathcal{V}$ have the same dimension.

Trace of a square matrix

Definition 59 *The **trace** of a (n, n) square matrix A is the sum of its diagonal terms and is denoted $Tr(A)$.*

$$Tr(A) = \sum_{i=1}^{n} a_{ii}$$

The elementary properties of the trace of a matrix are summarized in the following proposition.

Proposition 60 *Let A and B be two (n, n) matrices and $c \in \mathbb{R}$:*

$$\begin{aligned} Tr(cA + B) &= cTr(A) + Tr(B) \\ Tr(AB) &= Tr(BA) \end{aligned}$$

Denote A a matrix representing an endomorphism f of $\mathbb{R}^n$ when the bases are $\mathcal{U}$ and $\mathcal{V}$. The columns of A are the basis vectors of $\mathbb{R}^n$ transformed by f. If $\mathbb{R}^n$ is endowed with a new basis $\mathcal{W}$ the matrix representing f is modified (denoted B) but the trace does not change.

Proposition 61 $Tr(A) = Tr(B)$.

The reader can check that the proposition is true in example 58. The matrices with respect to the two bases were:

$$M_f(\mathcal{U}, \mathcal{V}) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 4 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{et} \quad M_f(\mathcal{W}, \mathcal{V}) = \frac{1}{7} \begin{pmatrix} -17 & 2 & -5 \\ 8 & 25 & 11 \\ 19 & 27 & 13 \end{pmatrix}$$

The trace of the two matrices is equal to 3.

Diagonalization of square matrices

A question coming naturally to mind when dealing with changes of bases is the following: can we transform the matrix of an endomorphism in a more simple one, more precisely in a diagonal matrix, by a change of basis?

Methods of data analysis like Principal Component Analysis of Factor Analysis are based on such transformations. Even if these methods are not addressed in the present book, the reader should know that they are used in multifactor models, especially the Arbitrage Pricing Theory[19].

Eigenvalues and eigenvectors   Eigenvalues and eigenvectors are the mathematical tools allowing to formalize a change of basis in such a way that the resulting matrix becomes diagonal.

Let f be an endomorphism of $\mathbb{R}^n$ and let M denote the matrix of f in basis $\mathcal{U}$.

Definition 62 *An **eigenvalue** of f (or equivalently of M) is a real number λ such that there exists a non zero vector $u \in \mathbb{R}^n$ satisfying:*

$$f(u) = Mu = \lambda u$$

*u is then called an **eigenvector** of f (of M) associated to the eigenvalue λ.*

Several linearly independent vectors can satisfy $Mu = \lambda u$. But if two linearly independent vectors u and v satisfy $Mu = \lambda u$ and $Mv = \lambda v$, then

---

[19]The seminal paper is Ross, S. (1976),"The Arbitrage Theory of Capital Asset Pricing". *Journal of Economic Theory* 13 (3): 341–360.

Two examples of papers using data analysis methods are Roll, R. and Ross, S. (1980). "An Empirical Investigation of the Arbitrage Pricing Theory". *Journal of Finance* 35 (5): 1073–1103.

Chamberlain, G. and Rothschild, M. (1983), "Arbitrage, Factor Structure, and Mean Variance Analysis on Large Asset Markets." *Econometrica* 51, 1281-1304.

the same relationship is true for any linear combination of u and v:

$$\forall (a,b) \in \mathbb{R} \times \mathbb{R}, M(au+bv) = \lambda(au+bv)$$

Of course, this relation is satisfied because f is a linear mapping:

$$
\begin{aligned}
f(au+bv) &= af(u) + bf(v) = M(au+bv) \\
&= a\lambda u + b\lambda v = \lambda(au+bv)
\end{aligned}
$$

We then obtain the following definition.

Definition 63 *The **eigenspace** of the eigenvalue λ is the vector subspace F_λ defined by:*

$$F_\lambda = \{u \in \mathbb{R}^n \text{ such that } f(u) = Mu = \lambda u\}$$

To determine the eigenvalues of a linear mapping f, we use the characteristic polynomial.

The characteristic polynomial

Definition 64 *Denote $M - \lambda I_n$ the following matrix:*

$$
M - \lambda I_n = \begin{pmatrix}
m_{11} - \lambda & m_{12} & \ldots & m_{1n} \\
m_{21} & m_{22} - \lambda & & m_{2n} \\
\ldots & & m_{33} - \lambda & \ldots \\
m_{n1} & m_{n2} & & m_{44} - \lambda
\end{pmatrix}
$$

*The **characteristic polynomial** of f is the polynomial $Q(\lambda)$ defined by:*

$$Q(\lambda) = \det(M - \lambda I_n)$$

$M - \lambda I_n$ is obtained by substracting λ times the identity matrix I_n to M. Solving $Q(\lambda) = 0$ provides all the eigenvalues of M.

Proposition 65 *The eigenvalues of f are the solutions of $Q(\lambda) = 0$.*

By definition, λ is an eigenvalue of M associated to the eigenvector u if:

$$Mu = \lambda u$$

This equality is equivalent to:

$$(M - \lambda I_n)\, u = 0$$

The matrix $M - \lambda I_n$ is then not invertible because $u \neq 0$. Therefore its determinant $\det(M - \lambda I_n) = Q(\lambda)$ is equal to 0.

Example 66 *Let f be a linear mapping represented by M in the canonical basis of $\mathbb{R}^3$:*

$$M = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Calculating the determinant of $M - \lambda I_n$ along the first line leads to:

$$
\begin{aligned}
\det(M - \lambda I_n) &= (1 - \lambda)\begin{vmatrix} 2 - \lambda & 2 \\ 1 & 1 - \lambda \end{vmatrix} - 0 \begin{vmatrix} 0 & 2 \\ 0 & 1 - \lambda \end{vmatrix} + 2 \begin{vmatrix} 0 & 2 - \lambda \\ 0 & 1 \end{vmatrix} \\
&= (1 - \lambda)\left((2 - \lambda)(1 - \lambda) - 2\right) \\
&= (1 - \lambda)(\lambda^2 - 3\lambda) \\
&= \lambda(1 - \lambda)(\lambda - 3)
\end{aligned}
$$

Equation $Q(\lambda) = 0$ has three solutions that are $\lambda_1 = 3$; $\lambda_2 = 0$ and $\lambda_3 = 1$.

What are the corresponding eigenvectors u^1, u^2, u^3? First, we need to

solve $(M - \lambda_1 I_3) u^1 = 0$, *that is:*

$$\begin{pmatrix} -2 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} u_1^1 \\ u_2^1 \\ u_3^1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The solution satisfies $u_1^1 = u_3^1$ *and* $u_2^1 = 2u_3^1$. *The following vector is an example of solution:*

$$u^1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

With the same approach for u^2 *and* u^3, *we obtain:*

$$u^2 = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} \quad and \quad u^3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

What is the matrix of f *in the basis* (u^1, u^2, u^3)?

The change-of-basis matrix P *is the matrix built with* u^1, u^2, u^3 *as columns because the initial basis was the canonical basis:*

$$P = \begin{pmatrix} 1 & -2 & 1 \\ 2 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

The inverse of P *is equal to:*

$$P^{-1} = \begin{pmatrix} 0 & 1/3 & 1/3 \\ 0 & -1/3 & 2/3 \\ 1 & -1 & 1 \end{pmatrix}$$

As a consequence, we have:

$$P^{-1}M = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

$$P^{-1}MP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Example 66 illustrates how M becomes a diagonal matrix when M is written in the basis of eigenvectors. Moreover, the elements on the diagonal are exactly the eigenvalues. We let the reader check the result of proposition 61, that is $Tr(M) = Tr(P^{-1}MP) = 4$.

When can we diagonalize a matrix?

Definition 67 *A (n, n) matrix M is **diagonalizable** if it has n eigenvalues and n linearly independent eigenvectors.*

An equivalent definition could be: M is diagonalizable if there exists a diagonal matrix D (contenant les eigenvalues) and an invertible matrix P satisfying:

$$M = P^{-1}DP$$

Of course, the diagonal elements of D are the eigenvalues of M and the columns of P are the corresponding eigenvectors of M.

It may happen that two eigenvalues are equal, for example if the characteristic polynomial is:

$$Q(\lambda) = (\lambda - 1)(\lambda - 3)^2$$

In this situation, M is diagonalizable if the dimension of the eigenspace F_3 (associated to $\lambda = 3$) is equal to 2.

On the contrary if $\dim(F_3) = 1$, M is not diagonalizable. We cannot find an invertible change-of-basis matrix P.

Symmetric matrices   A non negligible part of financial theory deals with portfolio choice and portfolio management. In this framework, an important piece of information is the covariance matrix of returns which is a symmetric matrix. These matrices are special because of the following proposition.

Proposition 68 *Any square symmetric matrix M is diagonalizable and we have:*

$$\begin{aligned} P^{-1} &= P' \\ M &= PDP' \end{aligned}$$

where P denotes the matrix of eigenvectors.

The first result says that the inverse of P is its transpose P'. Such a property characterizes orhogonal[20] matrices.

1.4　Norms and inner products

1.4.1　Normed vector spaces

In the section devoted to topology in chapter 1 of Part I, we defined the concepts of distance (metric) and metric spaces. The Euclidean distance on $\mathbb{R}^n$ was defined by:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \qquad (1.17)$$

If $n = 2$, $d(x,y)$ is the length of the straight line joining $x' = (x_1, x_2)$ to $y' = (y_1, y_2)$. More generally, in a finite-dimensional space, the concept of "length" of a vector is defined through a norm on the vector space under consideration.

Definition 69 *Let E be a vector space; a **norm** on E is a mapping, denoted $\|.\|$, defined on E and taking values in $\mathbb{R}^+$ satisfying:*

$$
\begin{aligned}
\|x\| &= 0 \Leftrightarrow x = 0 \\
\forall (x,y) &\in E, \|x+y\| \leq \|x\| + \|y\| \\
\forall x &\in E, \forall c \in \mathbb{R}^+, \|c.x\| = |c|\,\|x\|
\end{aligned}
$$

It appears that a norm $\|.\|$ on a vector space induces a metric d on the same space if the metric is defined by :

$$d(x,y) = \|x-y\|$$

---

[20]In the next section, we justify the word "orthogonal".

The Euclidean metric on $\mathbb{R}^n$ defined in relation 1.17 is induced by the following Euclidean norm on $\mathbb{R}^n$:

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

As for metrics, many different norms can be defined on a vector space. For example the mapping $\|x\|_{\max} = \max_i |x_i|$ can be used as a norm.

In finance, norms are associated to risk measures. For example if x_i denotes the future value of a portfolio in state i, the two abovementioned norms are interpreted differently.

Let $\mathbf{1}$ denote as usual the vector with all coordinates equal to 1 and $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ the average payoff. A usual measure of risk is the empirical variance calculated as:

$$\sigma^2(x) = \frac{1}{n} \|x - \overline{x}\mathbf{1}\|^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

But in a different approach called *Value at Risk*[21], we could use $\|x - \overline{x}\mathbf{1}\|_{\max}$ as a measure of risk[22]; risk is then evaluated as the maximum difference with respect to the average payoff.

The second important tool to structure a vector space is the concept of inner product. In finite-dimensional spaces, norms and inner products are closely related. It is not always the case in infinite-dimensional spaces.

1.4.2   Inner products in vector spaces

Definition 70 *An **inner product** on a vector space E is a mapping, denoted $< ., . >$, defined on $E \times E$ and taking values in $\mathbb{R}$, symmetric, bilinear*

---

[21] For a detailed presentation of Value-at-Risk, see Jorion (2006), *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill Professional.

[22] This measure is not exactly what is called *Value at Risk* in the financial literature but it is in the same spirit.

and positive, that is satisfying:

1. $< x, y > = < y, x >$

2. $\forall (a, b, c, d) \in \mathbb{R}^4, \forall (x, y, z, t) \in E^4,$

   $< ax + by, cz + dt > = ac < x, z > + ad < x, t > + bc < y, z > + bd < y, t >$

3. $< x, x > = 0 \Leftrightarrow x = 0 \ otherwise \ < x, x > \ > 0.$

Part 1 defines symmetry, part 2 bilinearity and part 3 positivity. The usual inner product on $\mathbb{R}^n$ is defined by:

$$< x, y > = \sum_{i=1}^{n} x_i y_i$$

Alternative notations of the inner product of two vectors x and y are (x, y) or $x'y$. The latter is consistent with the rules used to multiply matrices (see part I, chapter 4). The reason is that a column-vector is a matrix with n rows and 1 column. Consequently, x' is a matrix with 1 row and n columns. The product $x'y$ is then a matrix with 1 row and 1 column, that is a number.

Definition 70 allows for general inner products. However, we need to recall what is a positive-definite matrix to generalize inner products beyond the usual Euclidean ones.

Definition 71 *A square matrix A of dimension n is **positive (negative) semi-definite** if:*

$$\forall x \in \mathbb{R}^n, \ x'Ax \geq (\leq) \, 0$$

*A square matrix A of dimension n is **positive (negative) definite** if:*

$$\forall x \in \mathbb{R}^n, \ x \neq 0 \Rightarrow x'Ax > (<)0$$

This definition allows a general characterization of inner products on $\mathbb{R}^n$.

Proposition 72 *Let A be a square symmetric positive definite matrix; the mapping $(x, y) \rightarrow x'Ay$ associating any pair of vectors of $\mathbb{R}^n$ to the product $x'Ay$ is an inner product on $\mathbb{R}^n$ denoted $< ., . >_A$. The norm associated to this inner product, denoted $\|.\|_A$ is defined by $\|x\|_A = \sqrt{x'Ax}$*

Without entering into the details of the proof, remark that the condition $< x, x >_A > 0$ for $x \neq 0$ is satisfied because A is positive definite. In the same way, A is symmetric, property ensuring that the inner product is symmetric. Moreover, if A is the identity matrix, we are back to the definition of the usual inner product.

Definition 73 *Two vectors of E are **orthogonal** if their inner product is equal to 0.*

This definition of orthogonality refers to "right angles" when the usual two-dimensional space is endowed with the standard inner product. But the definition also shows that orthogonality is a much more general concept and, mainly, that being orthogonal for a pair of vectors depends on the inner product the vector space is endowed with. For example, if A is a diagonal matrix with strictly positive numbers on the diagonal satisfying $\sum a_{ii} = 1$, A defines an inner product allowing to calculate the expectation of the product of two random variables because the diagonal terms of A define a probability measure. In this example, orthogonality is far from the usual geometric interpretation[23].

Geometric interpretation

To elaborate on geometric aspects, consider the space $\mathbb{R}^n$ endowed with the usual norm and inner product. Let x and y denote two vectors in $\mathbb{R}^n$; the norm of the normalized vector $x^* = \frac{x}{\|x\|}$ is equal to 1 by construction. Let y^o be the projection of y on the line Δ_x generated by x. y^o is proportional to x, and more precisely we have the following equality:

$$y^o = < y, x^* > x^*$$

In other words, the inner product of x and y is equal to the coordinate of the projection of y on Δ_x (apart from the standardisation factor $\|x\|$), as illustrated by figure 1.1.

Let a denote the angle between x and y, we can establish the following

---

[23]For example the expectation of a random variable X can be written as $\mathbf{1}'AX = \sum_{i=1}^{n} a_{ii}X_i$ where X_i is the value of X in state i and a_{ii} is the probability of occurrence of state i.
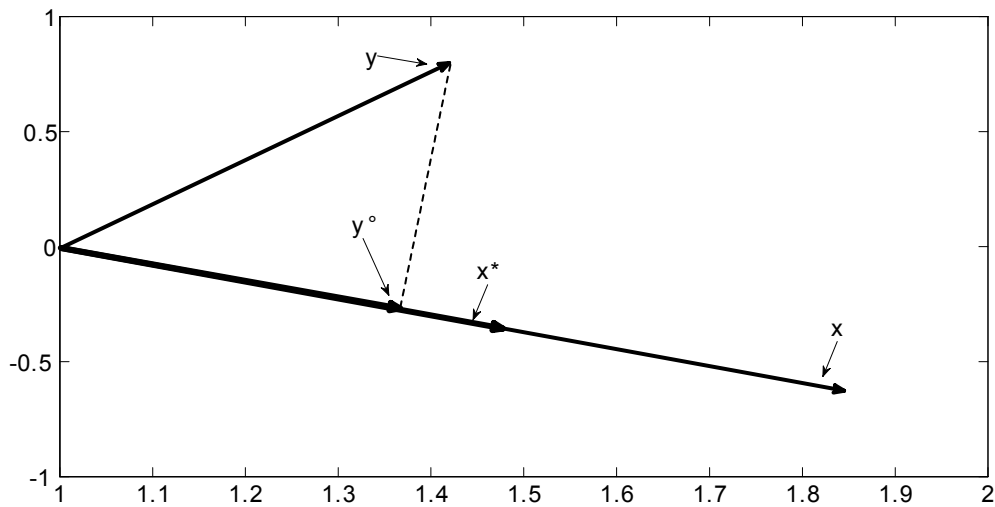
Figure 1.1: Geometric interpretation of the inner product

relationship between the inner product $< x, y >$ and the cosine of a :

$$< x, y >= \cos(a). \, \|x\| \, \|y\|$$

We are back to the well-known relationship saying that the cosine is equal to the ratio of the inner product divided by the product of norms.

1.4.3   Quadratic forms

Definition 74 *A **quadratic form** f, defined on an open subset $D \subset \mathbb{R}^n$, taking values in $\mathbb{R}$, is defined by :*

$$\forall x \in D, f(x) = x'Ax$$

where A is a symmetric square matrix.

Proposition 75 *A quadratic form is convex (concave) if and only if A*

is positive(negative) semi-definite. If A is positive(negative) definite, f is strictly convex (concave).

Quadratic forms are naturally present in portfolio management because the variance of the return of a portfolio x containing n stocks writes $x'\mathbf{V}x$ where $\mathbf{V}$ is the covariance matrix of returns of the n stocks.

In finance models, $\mathbf{V}$ is generally assumed positive definite, meaning that it is not possible to build a zero-variance portfolio (that is a risk-free portfolio) by combining n risky assets. It is an assumption but what is sure is that $\mathbf{V}$ is positive semi-definite because a variance of return $x'\mathbf{V}x$ cannot be negative.

The other domain where quadratic forms arise naturally is non linear optimization. We will see later on that it is easy to find the maximum (minimum) of a quadratic form when it is concave (convex).

1.5   Hilbert spaces

1.5.1   Definition

We mentioned several times that mathematical properties satisfied in finite-dimensional spaces could be false in more general spaces. However, there exists a category of infinite-dimensional vector spaces for which important properties remain valid. These spaces are called Hilbert spaces and they are well fitted to study financial problems, as it is illustrated in *Probability for Finance.*

Definition 76 *A **Hilbert space** is a vector space E whose norm is deduced from an inner product and that is complete as a metric space*[24].

---

[24]Remember that a metric space is complete if any Cauchy sequence converges. Therefore, speaking about a complete normed vector space is not really correct because completeness is a notion defined in metric spaces. This expression simply means that the metric d deduced from the norm on E makes (E, d) a complete metric space. This metric is defined by $d(x, y) = \|x - y\|$.

Of course, finite-dimensional spaces like $\mathbb{R}^n$ are Hilbert spaces when they are endowed with an inner product like the one defined in proposition 72.

The two essential properties for financial applications are the projection theorem and the Riesz representation theorem. Before presenting these results we first recall what is a convex set in a vector space.

Definition 77 *Let E denote a vector space and C a subset of E; C is **convex** if:*

$$\forall (x, y) \in C \times C, \forall \alpha \in [0; 1], \alpha x + (1 - \alpha)y \in C$$

First, it is important to notice that convexity can only make sense in vector spaces because, in the definition, there is a linear combination of vectors, $\alpha x + (1 - \alpha)y$. It is then necessary that this combination belongs to the vector space for the definition to make sense. As a consequence, in preceding chapters or in part 1 of the book, we could not have used convexity in a general framework. Nevertheless, the geometric interpretation of the convexity of a set is similar to what we proposed in $\mathbb{R}$ for intervals. A set is convex if, as soon as it contains two elements x and y, it also contains the segment joining these two elements.

Convexity is a standard assumption for consumption sets in microeconomics textbooks. It only means that goods are divisible. The same assumption on a set of portfolios would mean that portfolios and stocks can be combined in non integer quantities.

1.5.2   The projection theorem

Proposition 78 *Let E denote a Hilbert space and C a non empty convex set in E; any vector x in E has a unique projection on C, denoted x^* and satisfying:*

$$\forall y \in C, \langle x - x^*, y - x^* \rangle \leq 0$$

$x - x^*$ is orthogonal to the tangent to C at x^*. Consequently, the angle between $y - x^*$ and $x - x^*$ lies between $\frac{\pi}{2}$ and $\frac{3\pi}{2}$. The cosine of this angle is

then negative; but it is proportional to the inner product of the two vectors, meaning that this inner product is also negative. These remarks do not prove the proposition but they provide the geometric intuition for this proposition. One of the fundamental applications of the projection theorem consists in considering the case where C is a vector subspace of E. It is exactly the proposition allowing to define the conditional expectation of a random variable as a projection on a subspace of the vector space of square integrable random variables (see *Probability for Finance*).

The representation theorem presented below can also be interpreted with the same geometric approach.

1.5.3   The Riesz representation theorem

We saw in a preceding section of this chapter that a linear mapping defined on $\mathbb{R}^n$ and taking values in $\mathbb{R}^m$ can be associated to a matrix. In particular if $m = 1$, this matrix is a vector and the mapping is a linear form. In other words, if this mapping is denoted f, we write:

$$f(x) = \sum_{i=1}^{n} a_i x_i$$

and the vector $a' = (a_1, ..., a_n)$ represents f. This result may be generalized in the framework of Hilbert spaces, provided that f is continuous.

Proposition 79 *Let E be a Hilbert space and f be a continuous linear form defined on E; there exists a unique vector $y_f \in E$ such that:*

$$\forall x \in E, \ \ f(x) = \langle x, y_f \rangle$$

The vector y_f represents the linear mapping f and the important result is that y_f belongs to E.

In a financial framework the vector y_f has a natural interpretation if f is a valuation operator linking the future payoffs of a financial security x to its date-0 price $f(x)$. The coordinates of y_f are linked to the prices of the Arrow-Debreu securities. We already mentioned this characteristic in finite-dimensional spaces.

1.6   Separation theorems and Farkas lemma

1.6.1   Introductive example

Let f be a linear form defined on $\mathbb{R}^2$, characterized by the relation:

$$f(x) = a_1 x_1 + a_2 x_2$$

where $x' = (x_1, x_2)$ and a_1, a_2 are real numbers. The equation $f(x) = a_1 x_1 + a_2 x_2 = 0$ defines a line D in $\mathbb{R}^2$. Therefore, for any linear form f, the space $\mathbb{R}^2$ is divided in three regions denoted R_1, R_2 and D. These regions are characterized by:

$$\begin{aligned}
\forall x \quad &\in \quad R_1, f(x) > 0 \\
\forall x \quad &\in \quad R_2, f(x) < 0 \\
\forall x \quad &\in \quad D, f(x) = 0
\end{aligned}$$

Let C denote a non empty convex set not containing 0; there exists a linear form f, that is coefficients a_1 and a_2 satisfying:

$$\forall x \in C, \ f(x) > 0$$

In other words, the convex set C is entirely in R_1. This result is intuitive because any tangent to C induces a separation such that C is on one side of the tangent. For a given tangent Δ, separing 0 and C, consider the parallel to Δ containing 0. C is entirely on one side of this parallel to Δ. This line is defined by an equation like $a_1 x_1 + a_2 x_2 = 0$. If the elements $x \in C$ satisfy $f(x) > 0$, the desired result is obtained. If $f(x) < 0$ for $x \in C$, it is enough to choose the linear form g defined by $g(x) = -f(x) = -a_1 x_1 - a_2 x_2$.

1.6.2 Separation theorems and Farkas lemma

The following proposition is a generalization of the approach illustrated in the introductive example.

Proposition 80 *Separation theorem*

Let E be an Euclidean vector space, C be a non empty convex subset of E that does not contain the null vector. There exists a linear form f defined on E such that for any x in C, $f(x) \geq 0$.

The matrix expression of this separation theorem is called Farkas lemma. We propose hereafter two versions of this lemma, the second being usually called lemma of the alternative. This result has a beautiful financial interpretation, as we will illustrate later on.

Proposition 81 *Farkas lemma*

Let A be a matrix with m rows and n columns; a vector $x \in \mathbb{R}^n$ satisfies $x'y \geq 0$ for any $y \in \mathbb{R}^n$ such that $Ay \geq 0$ if and only if there exists a vector $z \in \left(\mathbb{R}_+^*\right)^m$ satisfying $x' = z'A$.

Proposition 82 *Lemma of the alternative*

Let A be a matrix with m rows and n columns; one and only one of the two following properties is true.

1. The equation $Ax = 0$ has a solution in $\mathbb{R}^n$ with all strictly positive coordinates.

2. Inequality $y'A > 0$ has a solution in $\mathbb{R}^m$.

1.6.3   Application to no-arbitrage pricing

Consider a one-period financial market on which investors trade securities at date 0, these securities providing random payoffs at date T. Recall that an arbitrage opportunity is a portfolio that costs nothing at date 0 (or maybe the cost is negative) and pays a positive amount at date T in all states of nature. Assume there are n possible states of nature and K securities traded on the market. The date-T payoffs are stored in a matrix D with n rows and K columns, each column corresponding to a security and each row to a state of nature.

Let $D = (d_{jk}, j = 1, .., n; k = 1, .., K)$, that is:

$$D = \begin{bmatrix} d_{11} & ... & d_{1K} \\ ... & d_{jk} & ... \\ ... & ... & ... \\ d_{n1} & ... & d_{nK} \end{bmatrix}$$

Denote $\Pi' = (\pi_1, ... \pi_K)$ the date-0 price vector; an arbitrage opportunity is a portfolio $\theta \in \mathbb{R}^K$ that satisfies one of the two following properties:

$$a) \ D\theta \ \geq \ 0 \text{ and } \Pi'\theta < 0$$
$$b) D\theta \ > \ 0 \text{ and } \Pi'\theta \leq 0$$

The first case (a) means that the portfolio has a strictly negative cost ($\Pi'\theta < 0$). Moreover, $D\theta \geq 0$ means that final payoffs are never negative. Consequently, an investor characterized by a strictly increasing utility function would be ready to buy an infinite quantity of this portfolio because holding this portfolio increases date-0 utility without decreasing date-T expected utility.

Case (b) is a little bit more subtle. Remember that $D\theta > 0$.[25] Therefore, the date-T expected utility increases by holding portfolio θ. But, at the same time, $\Pi'\theta \leq 0$, meaning that date-0 utility does not decrease when buying portfolio θ. As in case (a), an investor with a strictly increasing utility function would ask an infinite quantity of θ. In a well-functioning market, arbitrage opportunities should disappear very quickly by price adjustments due to excess supply or excess demand.

At a first glance, it may be difficult to see the relationship between the definition of an arbitrage opportunity and the lemma of the alternative...except that the two use matrix notations! The difficulty comes from the fact that

---

[25]Being given a matrix A, writing $A \geq 0$ means that all elements of A are positive, $A > 0$ means $A \geq 0$ and at least one element is strictly positive, and finally, $A >> 0$ means that all elements of A are strictly positive.

the definition of an arbitrage opportunity takes simultaneously into account date 0 and date T.

We are going to "forget" this specificity by defining a matrix D^* which is the concatenation of D and of $-\Pi'$(minus the price vector).

$$D^* = \begin{bmatrix} d_{11} & ... & d_{1K} \\ ... & d_{jk} & ... \\ ... & ... & ... \\ d_{n1} & ... & d_{nK} \\ -\pi_1 & ... & -\pi_K \end{bmatrix}$$

This notation allows to define an arbitrage opportunity as a vector $\theta \in \mathbb{R}^K$ satisfying:

$$D^*\theta > 0 \tag{1.18}$$

In fact, $D^*\theta > 0$ means that portfolio θ generates non negative cash-flows in all states and has a non positive cost. Moreover, at least one of the components of $D^*\theta$ is strictly positive. If the last component is strictly positive, we have a type (a) arbitrage opportunity. If this last component is zero, one of the other components is strictly positive and we face a type (b) arbitrage opportunity.

After transposing the two sides of inequality 1.18, we obtain:

$$\theta' D^{*\prime} > 0$$

This inequality corresponds to part (2) of lemma 82 when applied to D^*. We just showed that arbitrage opportunities are incompatible with equilibrium prices. We then have to assume that $D^*\theta > 0$ has no solution in θ.

As a consequence, lemma 82 implies there exists $\beta \in \mathbb{R}^{n+1}$ the components of which being all strictly positive, such that:

$$D^{*\prime}\beta = 0 \tag{1.19}$$

This equality must be true because part (b) of the lemma is false....then (a) is true!

The financial interpretation of relation 1.19 goes as follows. For the sake of clarity, focus on the first term of $D^{*\prime}\beta$; it is the inner product of β and of the first row of $D^{*\prime}$ (which corresponds to the first security). This inner product writes:

$$\sum_{j=1}^{n+1} \beta_j d_{j1}^* = 0$$

with $d_{j1}^* = d_{j1}$ if $j \leq n$ and $d_{j1}^* = -\pi_1$ if $j = n+1$. The above equality is then equivalent to:

$$\sum_{j=1}^{n} \beta_j d_{j1} = \beta_{n+1}\pi_1 \tag{1.20}$$

Denote $\gamma_j = \frac{\beta_j}{\beta_{n+1}}$ (these coefficients are well-defined because $\beta_{n+1} > 0$); relation 1.20 can be transformed in:

$$\sum_{j=1}^{n} \gamma_j d_{j1} = \pi_1 \qquad (1.21)$$

In the financial approach, this equality is very important because the left-hand side contains future cash-flows and the right-hand side contains the initial price. This equality is a typical valuation model (cash-flows on one side, price on the other side). However, the economic interpretation of equation 1.21 is difficult. But if we define $\gamma_j^* = \frac{\gamma_j}{\sum_{k=1}^{n} \gamma_k}$, we obtain:

$$\left(\sum_{k=1}^{n} \gamma_k \right) \sum_{j=1}^{n} \gamma_j^* d_{j1} = \pi_1$$

In this formula, the γ_j^* are positive numbers between 0 and 1 and satisfying $\sum_{j=1}^{n} \gamma_j^* = 1$. They define a probability measure on the set of states of nature. It is also remarkable that this probability measure does not depend on the asset we considered (here we selected the first but it does not matter). It remains to give an economic interpretation to $\sum_{k=1}^{n} \gamma_k$.

To simplify this interpretation, assume that asset numbered 1 is a risk-free asset paying 1 in each state, that is a zero-coupon bond. From equation 1.21 we deduce :

$$\sum_{j=1}^{n} \gamma_j = \pi_1$$

The quantity $\sum_{j=1}^{n} \gamma_j$ is the price of a security paying 1 at date T in each state of nature. If r denotes the risk-free rate, that is the return on the risk-free asset, we can write:

$$\sum_{j=1}^{n} \gamma_j = \frac{1}{1 + r}$$

It turns out that the valuation of any asset k writes:

$$\pi_k = \frac{1}{1+r} \sum_{j=1}^{n} \gamma_j^* d_{jk}$$

The price is then equal to the weighted average (an expectation using probabilistic vocabulary) of future cash-flows, discounted at the risk-free rate. Of course this interpretation can only be done when there is a risk-free asset traded on the market. But it can be generalized if there exists a portfolio generating strictly positive payoffs in any state of nature. Such a portfolio is named a **numéraire**.