The Geometry of Payoff Spaces

Marcel Hendrickx

June 4, 1998

Contents

Chapter 1

Introduction

This thesis contains an overview of *payoff spaces*, as they are studied in the finance literature. We show how many results from classical *mean-variance analysis* can be translated into the formalism of payoff spaces. We also establish the link between mean-variance analysis and *minimum-variance estimators* from statistics.

Mean-variance analysis

Mean-variance analysis is one of the pillars of modern finance. It was developed by Markowitz in 1952 ([Mar52]). In the 1960's mean-variance analysis was integrated into a more general economic equilibrium model, the Capital Asset Pricing Model (CAPM) of Sharpe and Lintner ([Sha64, Lin65]), which, despite the many attacks and counterattacks, still survives today.

In mean-variance analysis, returns on assets (stocks, bonds) are modeled as random variables. We start with a given set of assets. Linear combinations of (the returns on) these assets are called portfolios. For each return we have a mean (expected value) and a variance. The variance characterizes the riskiness of the portfolio. Different portfolios with the same expected return may have different variances. By suitably changing the composition of a portfolio (diversification), we may reduce the variance without changing the mean. The problem of mean-variance analysis is to find the portfolio with minimum variance for a given mean.

Portfolios can be plotted in a plane as points (E, σ), where E is the expected return of the portfolio and σ the standard deviation. This plane is called the E-σ plane. For a given mean there exists a unique portfolio that has minimum variance among all portfolios with that mean. The minimum-variance portfolios lie on a hyperbola, called the *minimum-variance frontier*.

Payoff spaces

Returns of assets are modeled as random variables. Whereas in classical mean-variance analysis returns are studied in the E-σ plane using their means and variances, the payoff space approach studies the random variables as elements of a linear space[1]. The payoff space approach to mean-variance analysis is simpler and more intuitive than the classical approach.

It is well known that the random variables on a given sample space form a linear space and certain aspects of this linear space structure have been studied in different contexts. But there seem to be few systematic studies. [PDN90] contains a systematic and extensive study of linear spaces of observations (as in regression analysis). Even though the elements are observations, that is, realisations of random variables, rather than random variables themselves, there is an intimate relation between the two kinds of linear spaces. The topics discussed in the book are classical topics like regression, principal components, analysis of variance, canonical analysis, discriminant analysis and correspondance analysis. But it contains nothing on mean-variance analysis.

In the finance literature, the linear space of random variables is called a payoff space. Even though payoffs spaces have been used since the mid 1980's, there still exists no systematic treatment. The most extensive treatment is probably Cochrane's manuscript on asset pricing ([Coc97][2]). Cochrane uses payoff spaces as a framework but the scope of his manuscript (asset pricing) is larger and his treatment of payoff spaces is not altogether systematic and somewhat unsatisfactory.

Minimum-variance estimators

Mean-variance analysis is concerned with minimizing variance for a given mean. The same problem is the subject of the study of minimum-variance estimators in statistics. Among a class of estimators with a given mean we look for the estimator with minimum variance. The question then imposes itself as to the relation between mean-variance analysis and minimum-variance estimators.

Even though this question is a natural one, the relation between mean-variance analysis and minimum-variance estimators has apparently not yet been the subject of study. In this thesis it is shown that mean-variance analysis is a natural tool for the study of minimum-variance estimators. A geometric proof of Barankin's theorem is given. This theorem, although less well known, comprises some well known theorems as special cases, such as the Cramér-Rao theorem.

Acknowledgements

I am indebted to Prof. M. Verbeek for pointing me to Cochrane's manuscript and for his comments on the text.

---

[1] That is, a vector space. In the non-mathematical world, the term 'vector space' seems to be less frequently used.

[2] Cochrane's manuscript is available at the Graduate School of Business, University of Chicago, at ftp-finance-gsb.uchicago.edu in pub/cochrane/papers/finbook#.ps, where # is the draft number.

Chapter 2

Payoff spaces

2.1 Introduction

In this chapter we give a formal definition of payoff spaces and introduce the necessary tools for a geometric study of these spaces.

2.2 Definition of payoff space

We start with a finite sample space Ω. The elements of a sample space are variously called outcomes, elementary events or sample points. We will call them *states* . Subsets of the sample space are called *events*.

With each state s we associate a *probability* $\pi(s)$ with

$$0 < \pi(s) < 1 \text{ and } \sum_{s \in \Omega} \pi(s) = 1. \tag{2.1}$$

The random variables defined on Ω form a linear space. This linear space will be called the *payoff space*[1] of Ω. The elements of the linear space are called *payoffs*. As the sample space is finite, the linear space is finite dimensional.

Each payoff x has a *mean*

$$E(x) = \sum_{s \in \Omega} \pi(s) x(s) \tag{2.2}$$

and a *variance*

$$\sigma^2(x) = E[(x - E(x))^2]. \tag{2.3}$$

The probability π defines an *inner product*

$$x.y = E(xy).$$

We will use $E(xy)$ rather than $x.y$ to denote the inner product.

[1] We use *payoff space* to denote the whole linear space, not the subspaces.

Definition 2.1 *The* payoff space *of a finite sample space* Ω *is the linear space of random variables defined on* Ω *with the inner product defined by the probability* π.

Two payoffs x and y are *orthogonal* ($x \perp y$) if $E(xy) = 0$. The *norm* or *length* of a payoff is denoted by $\|x\|$. We have

$$\|x\| = \sqrt{E(xx)}.$$

The *null payoff* is the payoff $0(s) = 0$ for every state s. The *unit payoff* is the payoff $1(s) = 1$. Any scalar multiple of the unit payoff is called a *fixed payoff* . A *contingent claim* is a payoff that assigns 1 to a given state s and 0 to all other states:

$$c_s(s') = \delta_{ss'}.$$

where $\delta_{ss'}$ is equal to one if $s = s'$ and zero otherwise.

Lemma 2.2 *The contingent claims form an orthogonal basis.*

Proof. We have:

$$\begin{aligned} E(c_{s'}c_{s''}) &= \sum_{s\in\Omega} \pi(s)c_{s'}(s)c_{s''}(s) \\ &= 0 \text{ if } s' \neq s'' \\ &= \pi(s) \text{ if } s' = s'' = s \end{aligned}$$

□

A *price function* is a linear map $p(x)$ from the payoff space to the real numbers. $p(x)$ is called the *price* of x. A payoff with $p(x) = 1$ is called a *return* . A payoff with $p(x) = 0$ is called an *excess return* .

2.3 Linear and affine subspaces

A *linear subspace* is a subset of the payoff space that is itself a linear space with respect to the addition and scalar multiplication of the payoff space and that has the same null payoff. The linear subspace inherits the probability and the inner product of the payoff space.

By definition a linear subspace contains the null payoff. That means that the returns do not form a linear subspace. However, if we substract from each return a given return then we obtain the set of excess returns. This set does form a linear subspace. We thus see that we can obtain the set of returns by adding an arbitrary return to all excess return. Geometrically, adding a payoff to each element of a set of payoffs corresponds to a *translation* of the given set. The translation of a linear subspace is called an *affine subspace*.

The affine subspaces can be turned into linear subspaces by arbitrarily fixing an origin in the affine subspace. We thus can define affine subspaces of affine subspaces, and this definition is independent of the choice of the origin.

By *subspace* we mean affine subspace. A linear subspace is also an affine subspace and the payoff space can be considered as an affine subspace of itself. The affine subspace formed by the returns is called the *return space* .

A *hyperplane* of an n-dimensional subspace is a $(n-1)$-dimensional subspace of the n-dimensional subspace. A *line* is a one-dimensional subspace. If x_1 and x_2 are different payoffs, then there exists a unique line containing both. This line is denoted $line(x_1, x_2)$. We write $line(x)$ for $line(0, x)$. A *hypersphere* [2] of an n-dimensional subspace is a $(n-1)$-dimensional surface consisting of all payoffs with a fixed distance to a given origin. A *hyperboloid* of an n-dimensional subspace is a $(n-1)$-dimensional surface that generalizes to n dimensions the two-dimensional hyperbola.

Let $x_1, \ldots, x_n$ be n payoffs. Then $span(x_1, \ldots, x_n)$ denotes the linear subspace generated by the payoffs. That is, the set of all linear combinations

$$x = \lambda_1 x_1 + \ldots + \lambda_n x_n.$$

The linear combinations

$$x = \sum_i \lambda_i x_i \text{ with } \sum_i \lambda_i = 1$$

form a n-1-dimensional affine subspace of $span(x_1, \ldots, x_n)$.

If $\mathbf{x}$ is a vector of payoffs then $span(\mathbf{x}) = span(x_1, \ldots, x_n)$ where $x_1, \ldots, x_n$ are the components of $\mathbf{x}$. The definition of $span(\mathbf{x}, \mathbf{y})$ is similar. Note that $line(x) = span(x)$.

Two linear subspaces are orthogonal if each payoff of the first subspace is orthogonal to each payoff of the second subspace. Two affine subspaces are orthogonal if the linear subspaces that are obtained by translating them to the origin are orthogonal.

Two lines in a plane are either parallel or intersect. Two parallel lines can be considered as intersecting at infinity. Thus for each direction in the plane we have exactly one *point at infinity*. In projective geometry it is proved that the points at infinity of a plane form a line. In general, the points at infinity of a higher dimensional space form a hyperplane of that space. The concept of point at infinity may seem artificial in this context, but we will need it in the proof of the mean-variance spanning and intersection theorems of Huberman and Kandel. It allows to consider parallel lines as a special case of intersecting lines.

The following lemma, which we state without proof, will be useful.

Lemma 2.3 *Let l_1 and l_2 be two lines that do not intersect and are not parallel. Then there exist unique points x_1 on l_1 and x_2 on l_2 such that the distance between x_1 and x_2 is the shortest distance of all distances between a point of l_1 and a point of l_2. Moreover, $line(x_1, x_2)$ is orthogonal to both l_1 and l_2.*

The most important application of this lemma is when the first line is $line(1)$ and the second line is a minimum-variance line. x_2 is then the minimum-variance payoff of the minimum-variance line and $x_1 = E(x_2)$ (cfr. below).

2.4 Orthogonal projection

Let x be a payoff and X a subspace of the payoff space. Then we define $proj(x|X)$ as

$$proj(x|X) = x' \in X \text{ with } \|x - x'\| = \min_{y \in X} \|x - y\|$$

[2] The following definitions of *hypersphere* and *hyperboloid* are informal, as only an intuitive understanding is required.

that is, $proj(x|X)$ is the payoff of X with the shortest distance to x. It can be proved that $proj(x|X)$ exists and is unique. $proj(x|X)$ is called the *orthogonal projection* of x on X. In that case, $(x - proj(x|X)) \perp X$. [3]

If X and Y are subspaces then

$$proj(X|Y) = \{proj(x|Y)|x \in X\}.$$

It can be proved that $proj(X|Y)$ is a subspace of Y. For example, the projection of a line on a plane is again a line and the projection of a line on a point is a point.

Lemma 2.4 Transitivity of orthogonal projection. *Let x be a payoff, X a subspace of the payoff space and Y a subspace of X. Then*

$$proj(x|Y) = proj(proj(x|X)|Y). \tag{2.4}$$

Proof. For linear spaces a proof is given in [PDN90]. This proof can be generalized to affine spaces. The rationale behind the generalization is that, intuitively, orthogonal projection is a geometric operation and does not depend on the choice of the origin. □

2.5 Price functions and stochastic discount factors

We have seen that a price function is just a linear map $p(x)$ from the payoff space to the real numbers. When the price function is not degenerate, that is, when there exists a payoff x with $p(x) \neq 0$, then the payoffs with a given price form a hyperplane. In that case, there exists a unique payoff m orthogonal to the hyperplanes such that for any payoff x

$$p(x) = E(mx).$$

The payoff m is called the *stochastic discount factor* of the price function. There exists a one-one correspondence between stochastic discount factors and non-degenerate price functions. Any payoff different from the null payoff defines a stochastic discount factor and thus a non-degenerate price function.

Let X be a subspace not containing the origin. There may exist many hyperplanes containing X. Each of these hyperplanes defines a price function that assigns a price one to the payoffs of the hyperplane. To this price function corresponds a stochastic discount factor. It can be proved that all these discount factors form a subspace. This subspace reduces to a single payoff if and only if X is a hyperplane.

Stochastic discount factors play an important role in the recent finance literature and, more particularly, in Cochrane's manuscript. Mean-variance analysis however is independent of any pricing concept. Stochastic discount factors therefore play only a marginal role in this thesis.

[3] We could also have defined orthogonal projection using the orthogonality introduced earlier and then deduce that $proj(x|X)$ is the payoff of X with the shortest distance to x.

2.6 Geometric interpretion of moments

Lemma 2.5 *Let x be a payoff different from the null payoff. Then $E(x)$ is the projection of x on $line(1)$ and $x - E(x)$ is the projection of x on the hyperplane of payoffs with $E = 0$. The length of $x - E(x)$ is $\sigma(x)$. The payoffs with a given expectation lie in a hyperplane orthogonal to $line(1)$.*

Proof. The projection of x on $line(1)$ is $E(x1)1 = E(x)$. The projection of x on the hyperplane $E = 0$ is then $x - E(x)$. That the payoffs with a given expectation lie in a hyperplane orthogonal to $line(1)$ follows from the fact that they have the same orthogonal projection on $line(1)$. □

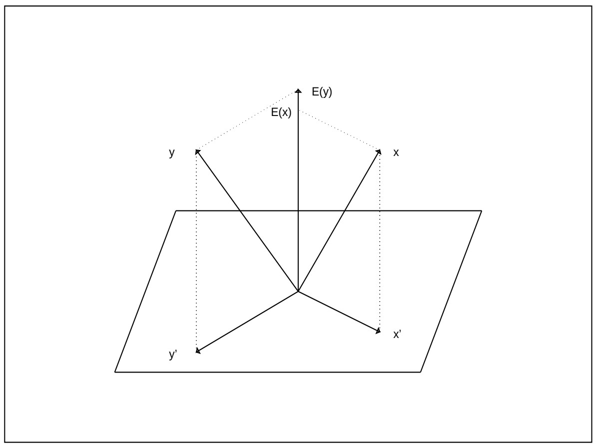

Figure 2.1: *The geometry of correlation*

Lemma 2.6 $\rho(x,y) = \cos(x',y')$ *where x' and y' are the orthogonal projections on the hyperplane $E = 0$.*

Proof. See figure 2.1. We have:

$$\begin{aligned} cov(x,y) &= E[(x - E(x))(y - E(y))] \\ &= E(x'y') \\ &= \|x'\| \, \|y'\| \cos(x',y') \end{aligned}$$

Substituting $\|x'\| = \sqrt{var(x')}$ and $\|y'\| = \sqrt{var(y')}$ gives the result. □

Corollary 2.7 *The payoffs y that have zero correlation with a given payoff x that is not a fixed payoff form a hyperplane containing $line(1)$ and orthogonal to $x - E(x)$.*

Corollary 2.8 $\rho(x,y) = \pm 1$ *if and only if x, y and 1 lie in a plane through the origin. $\rho = 1$ if x and y lie on the same side of the axis through the unit payoff. Otherwise $\rho = -1$.*

2.7 Decompositions

Let X and Y be linear subspaces[4] of a finite dimensional linear space V with inner product. Then $X+Y$ denotes the set $\{x+y|x \in X \text{ and } y \in Y\}$. $X+Y$ is called the *sum* of the two spaces and is again a linear subspace. Each vector of $X+Y$ can be written as a sum $x+y$ where $x \in X$ and $y \in Y$. If for each vector of $X+Y$ this decomposition is unique then the sum is called a *direct sum* and we write $X \oplus Y$. The following lemmas are well known and are given without proof.

Lemma 2.9 *Let U be a linear subspace of a linear space V. Then there exists a unique subspace $U^\perp$ such that $V = U \oplus U^\perp$. $U^\perp$ is called the* orthogonal complement *of U in V. We have $dim(V) = dim(U) + dim(U^\perp)$.*

Lemma 2.10 *Let U_1 and U_2 be linear subspaces of a linear space. Then*

$$dim(U_1 + U_2) = dim(U_1) + dim(U_2) - dim(U_1 \cap U_2). \tag{2.5}$$

These two lemmas are now used to prove the following decomposition.

Lemma 2.11 *Let U_1 and U_2 be linear subspaces of a linear space V with $U_1 \cap U_2 = \{0\}$. Then we have a direct sum decomposition*

$$V = U_1 \oplus U_2 \oplus (U_1^\perp \cap U_2^\perp) \tag{2.6}$$

Proof. The sum $U_1 + U_2 + (U_1^\perp \cap U_2^\perp)$ is a direct sum because the three linear subspaces have only the null vector in common. It then suffices to show that $dim(V) = dim(U_1 \oplus U_2 \oplus (U_1^\perp \cap U_2^\perp))$. Let $V' = U_1 \oplus U_2 \oplus (U_1^\perp \cap U_2^\perp)$. Because of the direct sum we have

$$\begin{aligned}
dim(V') &= dim(U_1) + dim(U_2) + dim(U_1^\perp \cap U_2^\perp) \\
&= dim(U_1) + dim(U_2) + dim(U_1^\perp) + dim(U_2^\perp) - dim(U_1^\perp + U_2^\perp) \\
&= dim(V) + dim(V) - dim(U_1^\perp + U_2^\perp) \\
&\geq dim(V)
\end{aligned}$$

From this it follows that $dim(V') = dim(V)$. □

Theorem 2.12 *Let y be a payoff and $\mathbf{x}$ be a vector of payoffs such that $1 \notin span(\mathbf{x})$. Then we have a unique decomposition*

$$y = a + \mathbf{x}'\mathbf{b} + e \tag{2.7}$$

where a is a fixed payoff and $\mathbf{b}$ a vector of scalars. We have

$$\mathbf{b} = cov(\mathbf{x}, \mathbf{x})^{-1} cov(y, \mathbf{x}) \tag{2.8}$$

and

$$cov(e, \mathbf{x}) = \mathbf{0} \tag{2.9}$$

[4] The definitions of sum and direct sum are valid for finite dimensional linear spaces in general, not only for payoff spaces.

Proof. We apply lemma 2.11 with $U_1 = span(1)$, $U_2 = span(\mathbf{x})$ and $U_1^\perp$ and $U_2^\perp$ the orthogonal complements. This gives decomposition 2.7. From $e \in U_1^\perp \cap U_2^\perp$ it follows that $e \perp 1$ or $E(e) = 0$ and $e \perp span(\mathbf{x})$ or $E(e\mathbf{x}) = \mathbf{0}$. This implies $cov(e, \mathbf{x}) = \mathbf{0}$. From equation 2.7 we then have

$$cov(y, \mathbf{x}) = cov(\mathbf{x}, \mathbf{x})\mathbf{b}$$

or

$$\mathbf{b} = cov(\mathbf{x}, \mathbf{x})^{-1} cov(y, \mathbf{x}).$$

□

The decomposition of theorem 2.12 looks like a regression equation. The difference with a classical regression equation is that the components of the payoffs are indexed by the states of the sample space whereas in a classical regression equation the components are indexed by observations. For example, the payoff y can be considered as a vector of values $y(s)$ where s is a state. In classical regression, the y would be a vector y_i where i indexes the i-th observation. Theorem 2.12 then essentially states that we can regress any payoff on any set of payoffs.

In the finance literature, regression arguments are often used in proofs. These arguments are heuristics rather than exact proofs. Examples are Huberman and Kandel's proof of mean-variance spanning and intersection ([HK87]) and various arguments in Cochrane's manuscript ([Coc97]). The decomposition in theorem 2.12 seems better suited for these kinds of arguments.

As an example of how the decomposition can be used to formalize the regression arguments, we now deduce the Hansen-Jagannathan inequality.

Lemma 2.13 *Let y be a payoff and $\mathbf{x}$ a vector of payoffs. Then*

$$\sigma^2(y) \geq cov(y, \mathbf{x})' cov(\mathbf{x}, \mathbf{x})^{-1} cov(y, \mathbf{x}). \tag{2.10}$$

Proof. We apply lemma 2.11 to y and $\mathbf{x} - E(\mathbf{x})$. This gives

$$y = E(y) + (\mathbf{x} - E(\mathbf{x}))'\mathbf{b} + e.$$

where the three payoffs on the right hand side are mutually orthogonal. Taking the variances of left and right hand side we obtain

$$\sigma^2(y) = \sigma^2(\mathbf{b}'(\mathbf{x} - E(\mathbf{x}))) + \sigma^2(e).$$

Using equation 2.8 and the fact that $\sigma^2(\mathbf{b}'\mathbf{z}) = \mathbf{b}' cov(\mathbf{z}, \mathbf{z})\mathbf{b}$ for any vector $\mathbf{z}$ of payoffs we obtain

$$\sigma^2(y) = cov(y, \mathbf{x})' cov(\mathbf{x}, \mathbf{x})^{-1} cov(y, \mathbf{x}) + \sigma^2(e)$$

from which them lemma results. □

Inequality 2.10 is the Hansen-Jagannathan inequality that defines the Hansen-Jagannathan bounds ([HJ91]). In that case, y is a stochastic discount factor that prices the assets $\mathbf{x}$. Cochrane ([Coc97]) gives two proofs of the inequality, a quick regression derivation taken from Hansen and Jagannathan and a geometric argument. He prefers the geometric argument because he thinks it can be more easily adapted for infinite dimensional spaces. However, he fails to see that the regression he uses is in fact a decomposition and thus a hundred percent geometrical. The regression argument captures the essence whereas his geometrical argument only indirectly arrives at the result.

Chapter 3

Mean-variance analysis

3.1 Introduction

In this chapter we discuss mean-variance analysis in payoff space. We deduce a number of basic results and show how they relate to classical mean-variance analysis in the E-σ plane.

3.2 The E-σ plane

Each payoff x has a mean $E(x)$ and a standard deviation $\sigma(x)$. $E(x)$ and $\sigma(x)$ can be considered as the coordinates of a point in a plane. This plane is called the *E-σ plane*. $\sigma(x)$ is the coordinate on the horizontal axis and $E(x)$ the coordinate on the vertical axis. The relation between E-σ plane and payoff space is made clear in figure 3.1. $E(x)$ is the

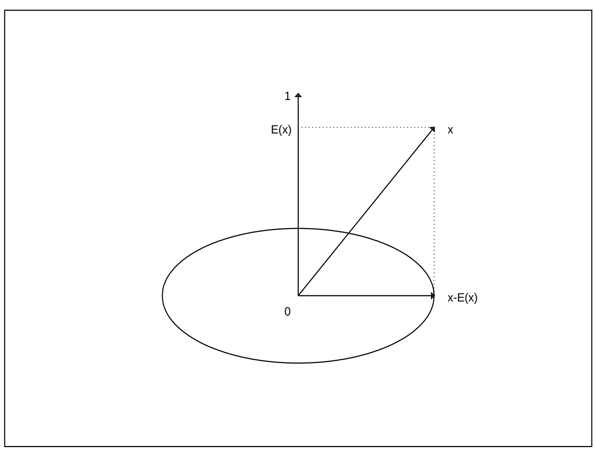

Figure 3.1: *From payoff space to E-σ plane*

orthogonal projection on $line(1)$ and $x - E(x)$ the orthogonal projection on the hyperplane

$E = 0$. The length of $x - E(x)$ is

$$\|x - E(x)\| = \sqrt{E[(x - E(x))^2]} = \sigma(x).$$

The E-σ plane can be obtained by rotating a plane through $line(1)$ and registering for each payoff where it passes through the plane. Alternatively, we can fix the plane and rotate the payoff space around $line(1)$.

It should be noticed that every payoff will appear in the E-σ plane, not only the returns. Moreover, payoffs that lie in a plane through the rotation axis $line(1)$ will have the same relative position in the E-σ plane as in the payoff space. The payoffs with a given E and σ form a hypersphere of the hyperplane through $E(x)$ and orthogonal to $line(1)$. All these payoffs are mapped to the same point in the E-σ plane. That means that each point in the E-σ plane that is not on the vertical axis corresponds to an infinity of payoffs in the payoff space. The points on the vertical axis are in one-to-one correspondence with the fixed payoffs.

A line in payoff space will be mapped to a hyperbola in the E-σ plane. This follows from the fact that in the payoff space the line will describe a hyperboloid which intersects a plane through the axis in a hyperbola[1]. The payoff on the line with shortest distance to $line(1)$ is the payoff with minimum variance, the *minimum-variance payoff* of the line.

The asymptotes of the hyperbola can be found as follows. Let x be the minimum-variance payoff on the line. Through the payoff $E(x)$ we draw a line parallel to the given line. This line will be mapped to a degenerate hyperbola consisting of two lines that intersect on the vertical axis. These lines cut the hyperbola at infinity and are thus parallel to the asymptotes[2]. Moreover they go through the origin of the hyperbola, the point $E(x)$. So they coincide with the asymptotes.

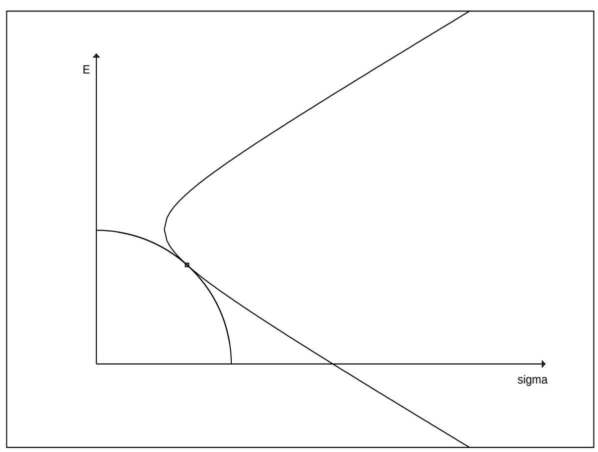

Figure 3.2: *Minimum distance payoff*

Lemma 3.1 *Let l be a line in payoff space that is mapped to a hyperbola in the E-σ plane. The payoff on l with the shortest distance to the origin is mapped in the E-σ plane to the point where the hyperbola is tangent to a concentric circle. See figure 3.2.*

[1] This is not really evident, but nevertheless intuitively acceptable.
[2] The asymptotes are tangent to the hyperbola at the points at infinity of the hyperbola.

Proof. The payoff x on l with the shortest distance to the origin is the payoff that minimizes $E(x^2)$. As $E(x^2) = E(x)^2 + \sigma^2(x)$, this means that x minimizes the distance to the origin in the E-σ plane. It then follows that the circle through the point $(E(x), \sigma(x))$ must be tangent to the hyperbola. An intuitive, though less rigorous proof is as follows. The payoff on l with the shortest distance to the origin is $0' = proj(0, l)$. We have $0' \perp l$. Consider the hypersphere through $0'$ and centered at the origin. The line l is tangent to this hypersphere. When we rotate l around $line(1)$, l generates a hyperboloid that is tangent to the hypersphere. If we cut the hyperboloid and the hypersphere with a 2-dimensional plane containing $line(1)$ we obtain a circle and a hyperbola that are tangent. The point of tangency is the payoff $0'$. □

3.3 Zero-correlated payoffs on a line

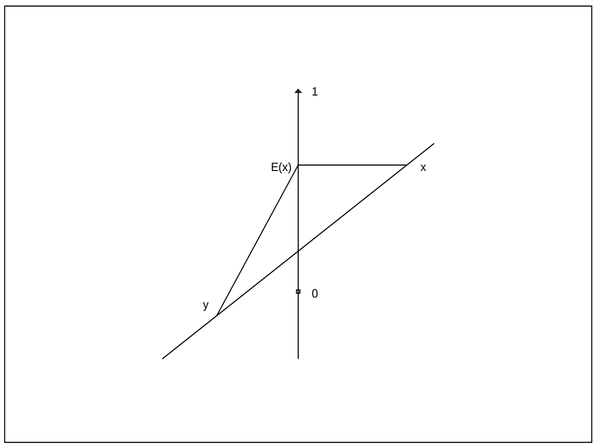

Figure 3.3: *Zero correlated payoffs*

Lemma 3.2 *Let l be a line that does not intersect $line(1)$ and is not parallel with it*[3]*. For each payoff x on l there exists exactly one payoff y that is zero-correlated with x. The zero-correlated payoff of the minimum-variance payoff of l is the payoff at infinity of l. Moreover, x and y lie on different sides of the minimum-variance payoff.*

Proof. This follows from corrollary 2.7. The hyperplane through the origin and orthogonal to $x - E(x)$ cuts line l in exactly one point y. $x - E(x)$ orthogonal to the hyperplane is equivalent to $line(0, x - E(x))$ orthogonal to the hyperplane. As $line(E(x), x)$ is parallel to $line(0, x - E(x))$ this implies that $line(E(x), x)$ is orthogonal to the hyperplane and thus orthogonal to every line in the hyperplane. In particular, $line(E(x), x)$ is orthogonal to $line(E(x), y)$, not only in payoff space but also in the plane through the payoffs x, y and $E(x)$. It is now clear from figure 3.3 that for each x there exists exactly one zero-correlated y and that y is the payoff at infinity when x is the minimum-distance payoff. It is also clear from the figure that x and y lie at different sides of the minimum-distance payoff. □

[3]The special cases of intersection and parallellism can be handled in a similar way. But they are not particularly interesting.

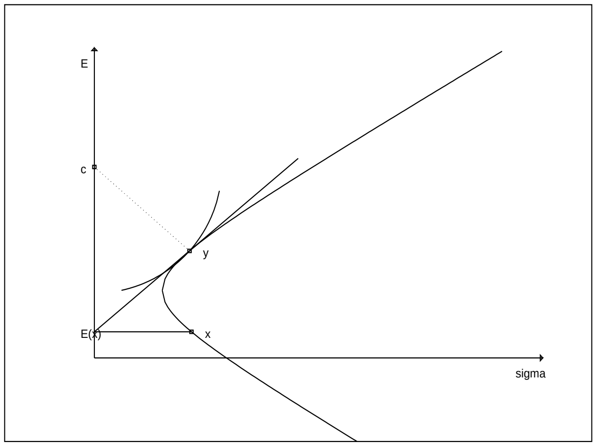

Figure 3.4: *Zero-correlated payoffs*

Figure 3.4 shows how the configuration in figure 3.3 is mapped to the E-σ plane: $line(x, y)$ is mapped to a hyperbola, $line(E(x), x)$ is mapped to a horizontal line and $line(E(x), y)$ is mapped to a line that is tangent to the hyperbola in y. This last point can be proved using an argument similar to the intuitive proof of lemma 3.1. Consider the fixed payoff c such that $line(c, y) \perp line(E(x), y)$. As $line(E(x), x)$ is orthogonal to $line(1)$ and to $line(E(x), y)$ it is also orthogonal to $line(c, y)$. But then $line(c, y)$ is orthogonal to the plane containing x, y and $E(x)$ and thus orthogonal to l. The hypersphere through y and centered at c is tangent to l and also to the hyperboloid generated by l. But, as $line(E(x), y) \perp line(c, y)$, $line(E(x), y)$ is tangent to the hypersphere and thus to the circle in the E-σ plane. Consequently $line(E(x), y)$ will also be tangent to the hyperbola.

3.4 The minimum-variance line

Definition 3.3 *The* minimum-variance set *for a subspace X is the set*

$$\{x \in X | \sigma(x) = \min_{y \in X} \sigma(y) \text{ and } E(x) = E(y)\} \tag{3.1}$$

That is, the set of payoffs that minimize the variance for a given expected value. A payoff in X with minimum variance is called a minimum-variance payoff *of X.*

Theorem 3.4 *The minimum-variance set of a subspace X is the orthogonal projection of $line(1)$ on X. When $line(1) \not\perp X$* [4]*, then this projection is a line, the* minimum-variance line . *The minimum-variance line is orthogonal to the hyperplanes of X of constant E. The minimum-variance line of X is noted $mvl(X)$. When $mvl(X)$ is not parallel with $line(1)$ there exists a unique minimum-variance payoff.*

Proof. Let x be a payoff in the minimum-variance set of X. Let X_E be the intersection of X and the hyperplane of payoffs with expected value equal to $E(x)$. Let c be any fixed

[4] In the following we will suppose that this condition is always satisfied unless it is stated otherwise

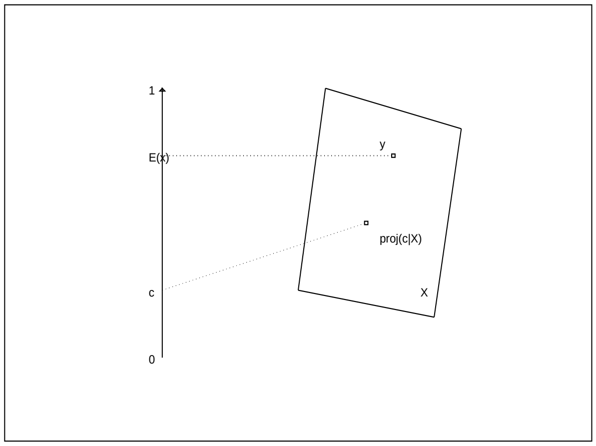

Figure 3.5: *The minimum-variance set is a line*

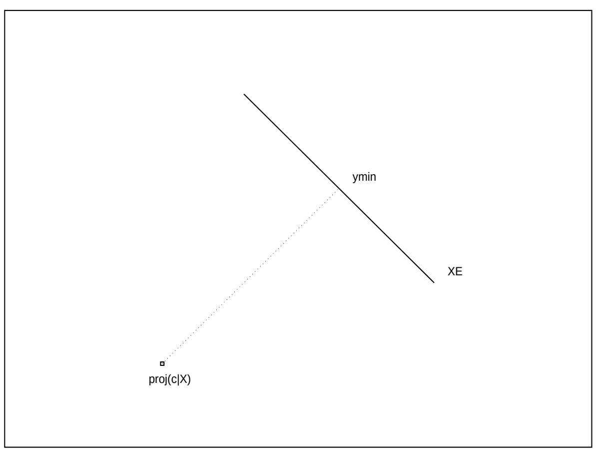

Figure 3.6: *The minimum-variance set is a line*

payoff. Then

$$\begin{aligned}
\sigma(x) &= \min_{y \in X_E} \sigma(y) \\
&= \min_{y \in X_E} dist(y, E(x)) \\
&= \min_{y \in X_E} dist(y, c) \\
&= \min_{y \in X_E} dist(y, proj(c|X))
\end{aligned}$$

where $dist(x, y)$ is the distance between two payoffs. These equations follow from Pythagoras' theorem ($line(E(x), c) \perp line(E(x), y)$ and $line(c, proj(c, X) \perp line(y, proj(c, X)))$ and from the fact that $dist(E(x), c)$ and $dist(c, proj(c|X))$ remain constant when y varies over X_E (see figure 3.5). The y_{min} satisfying the condition is thus $proj(proj(c|X)|X_E)$. If we let c vary over the fixed payoffs, $proj(c|X)$ describes a line in X that is the orthogonal projection of $line(1)$ and this line must go through y_{min}. Thus $mvl(X) = proj(line(1)|X)$ and $mvl(X) \perp X_E$. This is illustrated in figure 3.6. The standard deviation of a payoff is equal to the distance of the payoff to $line(1)$. Thus, when $mvl(X)$ is not parallel with $line(1)$, the minimum-variance payoff must be unique (lemma 3.1). □

The minimum-variance line of the payoff space itself consists of the payoffs with variance equal to zero, that is, $line(1)$. The minimum-variance line of a subspace is thus the orthogonal projection of the minimum-variance line of the payoff space. This result is intuitively satisfying and is similar to what we have in regression theory. More generally, the minimum-variance line of a subspace is the orthogonal projection of the minimum-variance line of any enclosing space. This follows from the transitivity of orthogonal projection (lemma 2.4).

It should be noted that a minimum-variance line is defined for any subspace that is not orthogonal to $line(1)$ and not only for subspaces of returns. Cochrane ([Coc97]) for example develops the theory for subspaces of returns. When deriving the minimum-variance frontier of the subspace of stochastic discount factors, he is obliged to repeat (with different notation) the whole argument, even though he recognizes that the argument is exactly the same. Roll ([Rol80]), still working in mean-variance space, faced the same problem when discussing orthogonal portfolios. These orthogonal portfolios form a subspace and thus have an associated minimum-variance line. He recognized that the arguments were similar but did not have the tools (payoff spaces) to express this similarity.

3.5 From minimum-variance line to CAPM

Lemma 3.5 *Let x_1 and x_2 be any payoffs with $var(x_2) \neq 0$ and $cov(x_1, x_2) = 0$. Let x a payoff on $line(x_1, x_2)$. Then*

$$x = x_1 + \frac{cov(x, x_2)}{var(x_2)}(x_2 - x_1) \tag{3.2}$$

Proof. Any payoff x on $line(x_1, x_2)$ can be written as

$$x = \lambda x_1 + \mu x_2 \text{ with } \lambda + \mu = 1$$

or

$$x = x_1 + \mu(x_2 - x_1) \tag{3.3}$$

From this we obtain

$$cov(x, x_2) = \mu var(x_2)$$

Solving for μ and substituting in equation 3.3 gives the desired result. □

Lemma 3.6 *Let X be a subspace and x and y payoffs in X with x on $mvl(X)$. Then*

$$cov(x, y) = cov(x, proj(y|mvl(X))). \tag{3.4}$$

That is, the covariance of x and y is equal to the covariance of x and the orthogonal projection of y on $mvl(X)$ (the efficient payoff corresponding to y).

Proof. Let $y' = proj(y|mvl(X))$ and $n = y - y'$. Then $E(n) = E(y) - E(y') = 0$ and $n \perp mvl(X)$ because the hyperplanes of X of constant E are orthogonal to $mvl(X)$. Let $0' = proj(0|X)$ and $x_2 = x - 0'$. Then $n \perp 0'$ because $0' \perp X$. If $x = 0'$ then $n \perp x$ or $E(xn) = 0$. If $x \neq 0'$ then $n \perp x_2$ because $n \perp mvl(X)$ and thus $n \perp x$ or $E(xn) = 0$. It follows that $cov(x, n) = E(xn) - E(x)E(n) = 0$. Consequently,

$$\begin{aligned} cov(x, y) &= cov(x, y' + n) \\ &= cov(x, y') + cov(x, n) \\ &= cov(x, y') \end{aligned}$$

□

Theorem 3.7 *Let x_1 and x_2 be two zero-correlated payoffs on $mvl(X)$. For any payoff x of X we have:*

$$E(x) = E(x_1) + \frac{cov(x, x_2)}{var(x_2)}(E(x_2) - E(x_1)) \tag{3.5}$$

Proof. We apply lemma 3.5 to $x' = proj(x|mvl(X))$. That gives:

$$x' = x_1 + \frac{cov(x', x_2)}{var(x_2)}(x_2 - x_1) \tag{3.6}$$

Taking expectations we obtain

$$E(x') = E(x_1) + \frac{cov(x', x_2)}{var(x_2)}(E(x_2) - E(x_1)) \tag{3.7}$$

Now, $mvl(X)$ is orthogonal to the hyperplanes of X of constant expectation. That means that $E(x) = E(x')$. Moreover, $cov(x, x_2) = cov(x', x_2)$ according to lemma 3.6. After substitution in the preceding equation we obtain equation 3.5. □

From equation 3.5 we deduce

$$\frac{E(x) - E(x_1)}{\sigma(x)} = \rho(x, x_2)\frac{E(x_2) - E(x_1)}{\sigma(x_2)}$$

or

$$\rho(x, x_2) = \frac{\alpha}{\alpha_2}$$

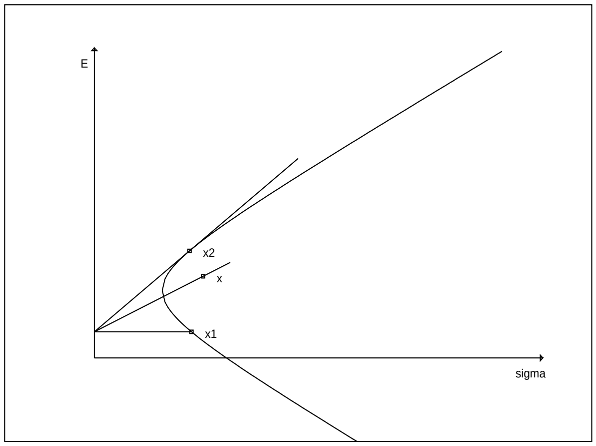

Figure 3.7: *Zero correlated payoffs*

where

$$\alpha = \frac{E(x) - E(x_1)}{\sigma(x)} \text{ and } \alpha_2 = \frac{E(x_2) - E(x_1)}{\sigma(x_2)}$$

This is illustrated in figure 3.7. Equation 3.5 is the familiar security market line equation from the Capital Asset Pricing Model.

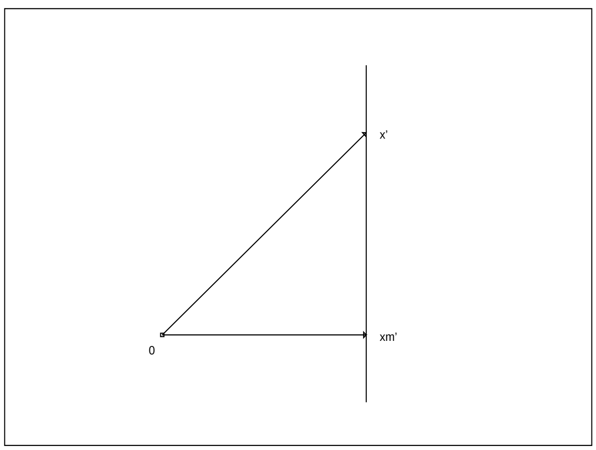

Figure 3.8: *Covariance with the minimum-variance payoff*

Lemma 3.8 *Let l be a line not through the origin and x_m be the minimum variance payoff of l, that is, the payoff with shortest distance to $line(1)$. Then*

$$cov(x, x_m) = \sigma^2(x_m) \tag{3.8}$$

for any x on l.

Proof. From the proof of 2.6 we have

$$cov(x, x_m) = \|x'\| \, \|x'_m\| \cos(x', x'_m) \tag{3.9}$$

where x' and x'_m are the orthogonal projections of x and x_m on the hyperplane $E = 0$. The line l is projected to a line l' that is orthogonal to x'_m. This is shown in figure 3.8. From this figure it is clear that $\|x\| \cos(x', x'_m) = \|x'_m\|$ and thus $cov(x, x_m) = cov(x_m, x_m) = var(x_m)$. □

Corollary 3.9 *Let X be a subspace not parallel and not orthogonal to $line(1)$. Let x be any payoff of X and x_m the minimum-variance payoff of X. Then*

$$cov(x, x_m) = \sigma^2(x_m) \tag{3.10}$$

Proof. This follows immediately from lemmas 3.6 and 3.8. □

3.6 Classical mean-variance analysis

In this section we start with a vector $\mathbf{x}$ of linearly independent payoffs and develop mean-variance analysis in terms of the mean and variance of $\mathbf{x}$. We then obtain most of the formulas and results of classical mean-variance analysis.

Let $y \in span(\mathbf{x})$ then, applying lemma 2.12, we have

$$y = \mathbf{x}'\mathbf{b} \tag{3.11}$$

$$\text{with } \mathbf{b} = cov(\mathbf{x}, \mathbf{x})^{-1} cov(y, \mathbf{x}). \tag{3.12}$$

Let us denote $\Sigma = cov(\mathbf{x}, \mathbf{x})$. Let X be the affine subspace containing the payoffs $\mathbf{x}$. X is a hyperplane of $span(\mathbf{x})$.

Lemma 3.10 *Let*

$$v_1 = \mathbf{x}'\Sigma^{-1}E(\mathbf{x}) \tag{3.13}$$

$$v_2 = \mathbf{x}'\Sigma^{-1}\mathbf{1}_K. \tag{3.14}$$

Then $span(v_1, v_2)$ contains $mvl(X)$.

Proof. We show that there exist linearly independent payoffs x_1 and x_2 that are linear combinations of v_1 and v_2, such that $span(x_1, x_2)$ contains $mvl(X)$. Define

$$x_1 = proj(0|X) \tag{3.15}$$

$$x_2 = proj(1|span(\mathbf{x})). \tag{3.16}$$

Then $x_1 \in mvl(X)$ because $0 \in line(1)$ and $mvl(X) = proj(line(1)|X))$ and $\lambda x_2 \in mvl(X)$. This last statement follows from the fact that we can obtain $mvl(X)$ by projecting $line(1)$ on $span(\mathbf{x})$ and then further projecting the result on X. Now $proj(line(1)|span(\mathbf{x})) = line(x_2)$ and $proj(line(x_2)|X) = mvl(X)$. As the dimensions of X and $span(\mathbf{x})$ differ by only one, $line(x_2)$ and $mvl(X)$ must intersect (possibly at infinity). This proves that $span(x_1, x_2)$ contains $mvl(X)$. We now prove that both x_1 and x_2 are linear combinations of v_1 and v_2.

We start with x_1. As both x_1 and the payoffs of $\mathbf{x}$ are in X, the orthogonallity $x_1 \perp X$ is equivalent to

$$E[x_1(\mathbf{x} - x_1\mathbf{1}_K)] = 0.$$

From this equation we deduce

$$E(x_1\mathbf{x}) = E(x_1^2)\mathbf{1}_K$$

and thus

$$\begin{aligned} cov(x_1, \mathbf{x}) &= E(x_1\mathbf{x}) - E(x_1)E(\mathbf{x}) \\ &= E(x_1^2)\mathbf{1}_K - E(x_1)E(\mathbf{x}) \end{aligned}$$

Substituting in 3.12 and then in 3.11, with x_1 playing the role of y, shows that x_1 is a linear combination of v_1 and v_2. We now prove the same for x_2. As x_2 is the orthogonal projection of 1 on $span(\mathbf{x})$ we have

$$E[(1 - x_2)\mathbf{x}] = 0.$$

This gives $E(\mathbf{x}) = E(x_2\mathbf{x})$ and thus

$$\begin{aligned} cov(x_2, \mathbf{x}) &= E(x_2\mathbf{x}) - E(x_2)E(\mathbf{x}) \\ &= E(\mathbf{x}) - E(x_2)E(\mathbf{x}) \\ &= (1 - E(x_2))E(\mathbf{x}) \end{aligned}$$

Substituting in 3.12 and then in 3.11 shows that x_2 is scalar multiple of v_1. This concludes the proof. □

Let A, B and C be defined by the following equations [5] :

$$A = E(\mathbf{x})'\Sigma^{-1}E(\mathbf{x}) \quad (3.17)$$

$$B = E(\mathbf{x})'\Sigma^{-1}\mathbf{1}_K \quad (3.18)$$

$$C = \mathbf{1}_K'\Sigma^{-1}\mathbf{1}_K \quad (3.19)$$

where $\mathbf{1}_K$ is a vector of K ones. The first and second order moments of v_1 and v_2 can be easily expressed in terms of A, B and C.

Lemma 3.11 *The first and second order moments of v_1 and v_2 are given by*

$$E(v_1) = A \quad (3.20)$$

$$\sigma^2(v_1) = A \quad (3.21)$$

$$E(v_2) = B \quad (3.22)$$

$$\sigma^2(v_2) = C \quad (3.23)$$

$$cov(v_1, v_2) = B \quad (3.24)$$

Proof. The expected values of v_1 and v_2 can be immediately deduced from their definitions. We show the calculation of $\sigma^2(v_1)$. The calculations for $\sigma^2(v_2)$ and $cov(v_1, v_2)$ are similar.

[5] These constants are defined in various ways in the literature. We follow [Rol77] and [Coc97].

We have

$$\begin{aligned}
\sigma^2(v_1) &= E[(v_1 - E(v_1))^2] \\
&= E[E(\mathbf{x})'\Sigma^{-1}(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))'\Sigma^{-1}E(\mathbf{x})] \\
&= E(\mathbf{x})'\Sigma^{-1}E[(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))']\Sigma^{-1}E(\mathbf{x}) \\
&= E(\mathbf{x})'\Sigma^{-1}\Sigma\Sigma^{-1}E(\mathbf{x})] \\
&= E(\mathbf{x})'\Sigma^{-1}E(\mathbf{x})] \\
&= A
\end{aligned}$$

□

The payoffs v_1 and v_2 can be used to parametrize the minimum-variance line.

Theorem 3.12 *The payoff $y = \mathbf{x}'\mathbf{b}$ lies on $mvl(X)$ if and only if there exist scalars λ and μ for which*

$$\mathbf{b} = \Sigma^{-1}(\lambda E(\mathbf{x}) + \mu \mathbf{1}_K).$$

where the coefficients satisfy

$$\lambda A + \mu B = E(y) \tag{3.25}$$
$$\lambda B + \mu C = 1. \tag{3.26}$$

Proof. Suppose that y lies on $mvl(X)$. Then $y \in span(v_1, v_2)$ and thus

$$\begin{aligned}
y &= \lambda v_1 + \mu v_2 \\
&= \mathbf{x}'\Sigma^{-1}\lambda E(\mathbf{x}) + \mathbf{x}'\Sigma^{-1}\mu \mathbf{1}_K \\
&= \mathbf{x}'\Sigma^{-1}(\lambda E(\mathbf{x}) + \mu \mathbf{1}_K)
\end{aligned}$$

We now prove that the coefficients must satisfy 3.25 and 3.26. To prove 3.25 we calculate $E(y)$:

$$\begin{aligned}
E(y) &= E[\mathbf{x}'\Sigma^{-1}(\lambda E(\mathbf{x}) + \mu \mathbf{1}_K)] \\
&= \lambda E(\mathbf{x})'\Sigma^{-1}E(\mathbf{x}) + \mu E(\mathbf{x})'\Sigma^{-1}\mathbf{1}_K \\
&= \lambda A + \mu B
\end{aligned}$$

To prove 3.26 we note that y lies X in and thus the coefficients $\mathbf{b}$ must sum to one. This gives:

$$\begin{aligned}
\mathbf{1}_K'\mathbf{b} = 1 &\Leftrightarrow \mathbf{1}_K'\Sigma^{-1}(\lambda E(\mathbf{x}) + \mu \mathbf{1}_K) = 1 \\
&\Leftrightarrow \lambda \mathbf{1}_K'\Sigma^{-1}E(\mathbf{x}) + \mu \mathbf{1}_K'\Sigma^{-1}\mathbf{1}_K = 1 \\
&\Leftrightarrow \lambda B + \mu C = 1
\end{aligned}$$

Conversely, if y satisfies the conditions, then y lies in $span(v_1, v_2)$. Moreover, $\mathbf{1}_K'\mathbf{b} = 1$ thus y lies in X. As $span(v_1, v_2) \cap X = mvl(X)$, it follows that y lies on $mvl(X)$. □

The equations 3.25 and 3.26 can be used to calculate the coefficients λ and μ. We have:

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} E(y) \\ 1 \end{pmatrix} \tag{3.27}$$

From which we obtain

$$\begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \frac{1}{AC - B^2} \begin{pmatrix} C & -B \\ -B & A \end{pmatrix} \begin{pmatrix} E(y) \\ 1 \end{pmatrix} \tag{3.28}$$

Lemma 3.13 *Let*

$$\begin{aligned} y_1 &= \lambda_1 v_1 + \mu_1 v_2 \\ y_2 &= \lambda_2 v_1 + \mu_2 v_2 \end{aligned}$$

where v_1 and v_2 are defined in lemma 3.10. Then

$$\begin{aligned} cov(y_1, y_2) &= \lambda_2 E(y_1) + \mu_2 && (3.29) \\ &= \lambda_1 E(y_2) + \mu_1 && (3.30) \end{aligned}$$

Proof.

$$\begin{aligned} cov(y_1, y_2) &= \lambda_1 \lambda_2 var(v_1) + \lambda_1 \mu_2 cov(v_1, v_2) + \mu_1 \lambda_2 cov(v_2, v_1) + \mu_1 \mu_2 var(v_2) && (3.31) \\ &= \lambda_1 \lambda_2 A + \lambda_1 \mu_2 B + \mu_1 \lambda_2 B + \mu_1 \mu_2 C && (3.32) \\ &= \lambda_1 (\lambda_2 A + \mu_2 B) + \mu_1 (\lambda_2 B + \mu_2 C) && (3.33) \\ &= \lambda_1 E(y_2) + \mu_1 && (3.34) \end{aligned}$$

The last equality follows from equations 3.25 and 3.26. This proves the first equality. The second equality can be proved in the same way. □

Corollary 3.14 *Let*

$$y = \lambda v_1 + \mu v_2 \tag{3.35}$$

then

$$var(y) = \lambda E(y) + \mu \tag{3.36}$$

Lemma 3.15 *The mean and variance of the minimum-variance payoff of X are given by*

$$\begin{aligned} E(y) &= \frac{B}{C} && (3.37) \\ var(y) &= \frac{1}{C} && (3.38) \end{aligned}$$

Proof. From the previous lemmas we have the following equations:

$$\begin{aligned} \lambda A + \mu B &= E(y) \\ \lambda B + \mu C &= 1 \\ var(y) &= \lambda E(y) + \mu \end{aligned}$$

From these equations we deduce

$$\begin{aligned} var(y) &= \lambda E(y) + \mu \\ &= \lambda^2 A + \lambda\mu B + \mu \\ &= \lambda^2 A + \lambda B\left(\frac{1-\lambda B}{C}\right) + \left(\frac{1-\lambda B}{C}\right) \\ &= \lambda^2 A + \frac{\lambda B}{C} - \frac{\lambda^2 B^2}{C} + \frac{1}{C} - \frac{\lambda B}{C} \\ &= \lambda^2\left(\frac{AC - B^2}{C}\right) + \frac{1}{C} \end{aligned}$$

We consider $E(y)$ and μ as functions of λ. Then we can find the minimum-variance payoff by putting the derivative of $var(y)$ with respect to λ equal to zero.

$$\frac{d\ var(y)}{d\lambda} = 2\lambda\left(\frac{AC - B^2}{C}\right) = 0$$

It follows that λ must equal zero. We then find $var(y) = \mu = \frac{1}{C}$ and $E(y) = \mu B = \frac{B}{C}$. □

We are now ready to give an intuitive interpretation to v_1 and v_2. Theorem 3.12 shows that any payoff y on $mvl(X)$ can be parametrized as

$$y = \lambda v_1 + \mu v_2 \text{ with } \lambda B + \mu C = 1$$

It follows that y can also be parametrized as

$$y = \lambda' v_1' + \mu' v_2' \text{ with } \lambda' + \mu' = 1$$

where

$$\begin{aligned} \lambda' &= \lambda B \\ \mu' &= \mu C \\ v_1' &= v_1/B \\ v_2' &= v_2/C \end{aligned}$$

and v_1' and v_2' lie on $mvl(X)$. In that theorem it was also proved that v_1 is a scalar multiple of $proj(1|span(\mathbf{x}))$. Consequently v_1' is the intersection of $mvl(span(\mathbf{x}))$ and $mvl(X)$. From lemma 3.15 it follows that v_2' is the minimum-variance payoff of X.

The previous lemmas can also be used to calculate the zero-correlated payoff of a given payoff on the minimum-variance line.

Lemma 3.16 *If y_1 and y_2 are on $mvl(X)$ and $cov(y_1, y_2) = 0$ then*

$$E(y_2) = \frac{A - BE(y_1)}{B - CE(y_1)}. \tag{3.39}$$

Proof. From $\lambda_1 E(y_2) + \mu_1 = 0$ we deduce $E(y_2) = -\mu_1/\lambda_1$. Substituting λ_1 and μ_1 by the values obtained from solving equation 3.28 gives the desired result. □

3.7 Mean-variance spanning and intersection

Let $\mathbf{y}$ be a vector of N payoffs and $\mathbf{x}$ be a vector of K payoffs and assume that the payoffs $\mathbf{y}$ and $\mathbf{x}$ are linearly independent. Then theorem 2.12 implies a unique decomposition

$$\mathbf{y} = \mathbf{a} + \mathbf{B}\mathbf{x} + \mathbf{e} \tag{3.40}$$

where $\mathbf{a}$ is a vector of fixed payoffs, $\mathbf{e}$ a vector of payoffs with expected value equal to zero and $\mathbf{B}$ a matrix of scalars. Let p be a price function and Y and X be the return spaces of $span(\mathbf{y})$ and $span(\mathbf{x})$. Let $X + Y$ be the return space of $span(\mathbf{x}, \mathbf{y})$.

Theorem 3.17 Mean-variance intersection. *$mvl(X+Y)$ intersects $mvl(X)$ if and only if there exists a constant w_0 such that*

$$\mathbf{a} = w_0(\mathbf{1}_N - \mathbf{B}\mathbf{1}_K). \tag{3.41}$$

Moreover, w_0 is the expected value of any return of $X + Y$ that is zero-correlated with the intersection point.

Proof.[6] If $mvl(X+Y)$ intersects $mvl(X)$ then $mvl(X+Y)$ intersects $span(\mathbf{x})$. Inversely, if $mvl(X+Y)$ intersects $span(\mathbf{x})$ in a payoff s then $proj(s|X) = s$ and thus s lies on $mvl(X)$. Thus $mvl(X+Y)$ intersects $mvl(X)$ if and only if $mvl(X+Y)$ intersects $span(\mathbf{x})$. We now use the parametrization of theorem 3.12 to translate this condition in terms of $\mathbf{a}$. Let $\mathbf{z}$ denote the vector of payoffs $\mathbf{x}$ and $\mathbf{y}$. Let

$$\begin{aligned} v_1 &= \mathbf{z}'\Sigma_{\mathbf{zz}}^{-1}E(\mathbf{z}) \\ v_2 &= \mathbf{z}'\Sigma_{\mathbf{zz}}^{-1}\mathbf{1}_{K+N} \end{aligned}$$

where $\Sigma_{\mathbf{zz}} = cov(\mathbf{z}, \mathbf{z})$. We have

$$\Sigma_{\mathbf{zz}} = \begin{pmatrix} \Sigma_{\mathbf{xx}} & \Sigma_{\mathbf{xy}} \\ \Sigma_{\mathbf{yx}} & \Sigma_{\mathbf{yy}} \end{pmatrix}$$

Any payoff y on $mvl(X+Y)$ can then be written as

$$\begin{aligned} y &= \lambda v_1 + \mu v_2 \\ &= \mathbf{z}'\Sigma_{\mathbf{zz}}^{-1}(\lambda E(\mathbf{z}) + \mu\mathbf{1}_{K+N}) \end{aligned}$$

For the coefficients $\mathbf{b}_{\mathbf{z}}$ we must have

$$\mathbf{b}_{\mathbf{z}} = \Sigma_{\mathbf{zz}}^{-1}(\lambda E(\mathbf{z}) + \mu\mathbf{1}_{K+N})$$

or

$$\Sigma_{\mathbf{zz}}\mathbf{b}_{\mathbf{z}} = \lambda E(\mathbf{z}) + \mu\mathbf{1}_{K+N}$$

Translating in terms of $\mathbf{x}$ and $\mathbf{y}$ gives

$$\begin{aligned} \Sigma_{\mathbf{xx}}\mathbf{b}_{\mathbf{x}} + \Sigma_{\mathbf{xy}}\mathbf{b}_{\mathbf{y}} &= \lambda E(\mathbf{x}) + \mu\mathbf{1}_K \\ \Sigma_{\mathbf{yx}}\mathbf{b}_{\mathbf{x}} + \Sigma_{\mathbf{yy}}\mathbf{b}_{\mathbf{y}} &= \lambda E(\mathbf{y}) + \mu\mathbf{1}_N \end{aligned}$$

[6] The following proof is adapted from [dR97]. The proof given by de Roon is essentially the proof given by Huberman and Kandel, but it is more complete and less obscure.

The payoff y lies in $span(\mathbf{x})$ if and only if $\mathbf{b_y} = \mathbf{0}$. The equations then reduce to

$$\begin{aligned} \Sigma_{\mathbf{xx}}\mathbf{b_x} &= \lambda E(\mathbf{x}) + \mu\mathbf{1}_K \\ \Sigma_{\mathbf{yx}}\mathbf{b_x} &= \lambda E(\mathbf{y}) + \mu\mathbf{1}_N \end{aligned}$$

Solving for $\mathbf{b_x}$ in the first equation and substituting in the second equation gives

$$\begin{aligned} \Sigma_{\mathbf{xy}}\Sigma_{\mathbf{xx}}^{-1}(\lambda E(\mathbf{x}) + \mu\mathbf{1}_K) &= \lambda E(\mathbf{y}) + \mu\mathbf{1}_N \\ \mathbf{B}(\lambda E(\mathbf{x}) + \mu\mathbf{1}_K) &= \lambda E(\mathbf{y}) + \mu\mathbf{1}_N \end{aligned}$$

where we have used $\mathbf{B} = \Sigma_{\mathbf{xy}}\Sigma_{\mathbf{xx}}^{-1}$. Rearanging left and right hand side, we obtain

$$\begin{aligned} \lambda(E(\mathbf{y}) - \mathbf{B}E(\mathbf{x})) &= \mu(\mathbf{B1}_K - \mathbf{1}_N) \\ \lambda\mathbf{a} &= \mu(\mathbf{B1}_K - \mathbf{1}_N) \\ \mathbf{a} &= \frac{\mu}{\lambda}(\mathbf{B1}_K - \mathbf{1}_N) \\ \mathbf{a} &= -\frac{\mu}{\lambda}(\mathbf{1}_N - \mathbf{B1}_K) \end{aligned}$$

where we have used $\mathbf{a} = E(\mathbf{y} - \mathbf{B})$. From lemma 3.16 we know that $-\frac{\mu}{\lambda}$ is the expected value of the payoff that is zero-correlated with y. □

Theorem 3.18 Mean-variance spanning. *$mvl(X+Y)$ coincides with $mvl(X)$ if and only if $\mathbf{a} = 0$ and $\mathbf{1}_N = \mathbf{B1}_K$.*

Proof. This follows immediately from the intersection theorem. Equation 3.41 must be satisfied for all payoffs on $mvl(X+Y)$, that is, for all corresponding w_0. This is possible if and only if $\mathbf{a} = 0$ and $\mathbf{1}_N = \mathbf{B1}_K$. □

The spanning and intersection theorems were used by Huberman and Kandel to construct statistical tests for spanning and intersection. Such tests allows to verify whether it is worth to add a new asset to an existing portfolio by verifying whether the introduction of the asset shifts the minimum-variance frontier of the portfolio and thus extends the investment opportunities.

Chapter 4

Minimum-variance estimators

4.1 Introduction

In this chapter we show how mean-variance analysis in payoff space can be used to study minimum-variance estimators. We give a proof of Barankin's theorem and of two special cases of it, the Chapman-Robbins-Kiefer inequality and the Cramér-Rao inequality.

4.2 Random samples

Let X be a random variable with a probability distribution that depends on a parameter θ. *Statistical estimation* is concerned with estimating the true value of θ from a sample $\{X_1, X_2, \dots, X_n\}$.

A *random sample* from a distribution F is defined as a set $\{X_1, X_2, \dots, X_n\}$ of random variables that are independently and identically distributed with distribution F. This definition is somewhat counterintuitive. Taking a sample for a random variable X defined on a sample space Ω means, repeating the experiment and for each outcome recalculating the value of X. That is, the random variable is the same for each trial. A sample then is a random vector defined on the product sample space Ω^n and taking values in $\mathbf{R}^n$. The components of this vector are random variables that are also defined on Ω^n. One can then easily prove that they are independently and identically distributed with the same distribution as X. The point is that these random variables are defined on the product sample space Ω^n and not on Ω.

In the payoff space model there is a clear separation between the function that defines the random variable and the probability measure that gives the random variable its random character. We can only represent random variables that are defined on the same sample space underlying the payoff space. As we have not defined product payoff spaces, we can only discuss estimation for samples of size one.

4.3 Estimation

It is generally, although most of the time implicitely, assumed that the probability distribution $f(x, \theta)$ of the random variable X only depends on θ through the probability measure and that the function that defines the random variable does not depend on θ. Instead of a family of probability distributions $f(x, \theta)$, we then have a family of probability payoffs π_θ with $\theta \in \Theta$, where Θ is a set of parameters. Each probability payoff π_θ defines an expectation

$$E_\theta(x) = \sum_{s \in \Omega} \pi_\theta(s) x(s) \tag{4.1}$$

and a corresponding inner product. Let θ_0 be the true value of the parameter. We define $E(x) = E_{\theta_0}(x)$.

Lemma 4.1 *We have the following equality:*

$$E_\theta(x) = E(x \frac{\pi_\theta}{\pi_{\theta_0}}) \tag{4.2}$$

For $x = 1$ we have

$$E(\frac{\pi_\theta}{\pi_{\theta_0}}) = 1 \tag{4.3}$$

Proof.

$$\begin{aligned} E_\theta(x) &= \sum_{s \in \Omega} \pi_\theta(s) x(s) \\ &= \sum_{s \in \Omega} \pi_{\theta_0}(s) \frac{\pi_\theta(s)}{\pi_{\theta_0}(s)} x(s) \\ &= E_{\theta_0}(x \frac{\pi_\theta}{\pi_{\theta_0}}) \\ &= E(x \frac{\pi_\theta}{\pi_{\theta_0}}) \end{aligned}$$

□

Definition 4.2 *An* estimator *is just a payoff. The estimator x is an* unbiased estimator *of θ if*

$$E_\theta(x) = \theta \text{ for all } \theta \in \Theta \tag{4.4}$$

Lemma 4.3 *The unbiased estimators form a subspace M orthogonal (with respect to θ_0) to the linear subspace P generated by the payoffs $\frac{\pi_\theta}{\pi_{\theta_0}}$. The subspace P contains line(1).*

Proof. Let x_1 and x_2 be unbiased estimators and $x = \lambda_1 x_1 + \lambda_2 x_2$ with $\lambda_1 + \lambda_2 = 1$. From the linearity of the operator E_θ it follows that

$$\begin{aligned} E_\theta(x) &= \lambda_1 E_\theta(x_1) + \lambda_2 E_\theta(x_2) \\ &= \lambda_1 \theta + \lambda_2 \theta \\ &= (\lambda_1 + \lambda_2)\theta \\ &= \theta \end{aligned}$$

From this it follows that the unbiased estimators form an affine subspace. To verify the orthogonality we translate the affine subspace of unbiased estimators to the origin and verify the orthogonality of the resulting linear spaces. We take an unbiased estimator x_0. We then have

$$\begin{aligned} E[(x - x_0)\frac{\pi_\theta}{\pi_{\theta_0}}] &= E(x\frac{\pi_\theta}{\pi_{\theta_0}}) - E(x_0\frac{\pi_\theta}{\pi_{\theta_0}}) \\ &= \theta - \theta \\ &= 0. \end{aligned}$$

That means that the linear subspaces are orthogonal and thus also the corresponding affine subspaces. As P contains $1 = \frac{\pi_{\theta_0}}{\pi_{\theta_0}}$ it also contains the subspace generated by 1, that is, $line(1)$. □

4.4 Minimum-variance estimators

Let M be the subspace of unbiased estimators and P be the linear subspace generated by the payoffs $\frac{\pi_\theta}{\pi_{\theta_0}}$.

Lemma 4.4 *The subspace P contains exactly one unbiased estimator. This estimator is the minimum-variance unbiased estimator for θ_0.*

Proof. The unbiased estimators are characterized by

$$E(x\frac{\pi_\theta}{\pi_{\theta_0}}) = \theta \text{ for all } \theta \in \Theta. \tag{4.5}$$

Among the $\frac{\pi_\theta}{\pi_{\theta_0}}$ we can select a basis. Equation 4.5 then defines coordinates for x by projecting orthogonally on the coordinate axes[1]. There is exactly one payoff in P whose coordinate on the θ-axis is θ. For this payoff we have $E(x(\pi_\theta/\pi_{\theta_0})) = E_\theta(x) = \theta$. This payoff is thus an unbiased estimator. We now prove that it is the minimum-variance estimator of M. The minimum-variance set of M is equal to $proj(line(1)|M)$ where $line(1)$ is contained in P. But P and M are orthogonal subspaces whose intersection $M \cap P = \{x\}$ consists of a single payoff. Therefore, $proj(P|M) = \{x\}$ and thus also $proj(line(1)|M) = \{x\}$. The minimum-variance set of M consists of a single payoff x which is the minimum-variance unbiased estimator. □

It becomes clear now why there may be a lower bound on the variance of a set of estimators. For a given θ the payoffs satisfying $E_\theta(x) = \theta$ form a hyperplane. When we vary θ, we generate a family of hyperplanes whose intersection is the subspace of unbiased estimators. This subspace contains a minimum-variance set (which in this case is not a line, but a single payoff) and thus a minimum-variance payoff. The variance of the minimum-variance payoff is a lower bound on the variances of the payoffs of the subspace. In general, the subspace does not cut $line(1)$ and therefore the lower bound will in general be different from zero.

[1]Coordinates are normally obtained by projecting parallel on the coordinate axes. When the axes are orthogonal both coordinate systems give the same result.

Lemma 4.5 *The variance of the minimum-variance unbiased estimator is given by*

$$\sigma^2(x) = (\theta - \theta_0)'\mathbf{E}^{-1}(\theta - \theta_0) \tag{4.6}$$

where θ is a vector[2] *of θ_i such that the $\frac{\pi_{\theta_i}}{\pi_{\theta_0}}$ form a basis of P, θ_0 is a constant vector $\theta_0\mathbf{1}$ and $\mathbf{E}$ is the matrix*

$$E_{i,j} = E(\frac{\pi_{\theta_i}}{\pi_{\theta_0}}\frac{\pi_{\theta_j}}{\pi_{\theta_0}}).$$

.

Proof. Let $e_i = \frac{\pi_{\theta_i}}{\pi_{\theta_0}}$, $\mathbf{e}$ the vector of e_i and λ a vector of coefficients. Then $\mathbf{E} = E(\mathbf{e}\mathbf{e}')$. Let $x = \lambda'\mathbf{e}$ be any payoff in P. Then we have

$$E(x\mathbf{e}) = E(\mathbf{e}\mathbf{e}')\lambda$$

and

$$x = \mathbf{e}'\mathbf{E}^{-1}E(x\mathbf{e}).$$

If x is moreover an unbiased estimator then we have $E(x) = \theta_0$ and $E(x\mathbf{e}) = \theta$ and thus

$$\begin{aligned} x - E(x) &= \mathbf{e}'\mathbf{E}^{-1}E[(x - E(x))\mathbf{e}] \\ &= \mathbf{e}'\mathbf{E}^{-1}E[(x - \theta_0)\mathbf{e}] \\ &= \mathbf{e}'\mathbf{E}^{-1}(\theta - \theta_0) \end{aligned}$$

Consequently:

$$\begin{aligned} \sigma^2(x) &= E[(x - E(x))^2] \\ &= E[(\mathbf{e}'\mathbf{E}^{-1}(\theta - \theta_0))^2] \\ &= (\theta - \theta_0)'\mathbf{E}^{-1}E(\mathbf{e}\mathbf{e}')\mathbf{E}^{-1}(\theta - \theta_0) \\ &= (\theta - \theta_0)'\mathbf{E}^{-1}(\theta - \theta_0) \end{aligned}$$

□

The following lemma gives another characterization of the variance of the minimum-variance estimator.

Lemma 4.6 *The variance of the minimum-variance unbiased estimator is given by*

$$\sigma^2(x) = \sup_{a_i} \frac{(\sum_{i=1}^n a_i(\theta_i - \theta_0))^2}{E\left[\left(\sum_{i=1}^n a_i \frac{\pi_{\theta_i}}{\pi_{\theta_0}}\right)^2\right]} \tag{4.7}$$

where the supremum is taken over all a_i.

[2] Latex does not allow to put lower case Greek characters in bold. It should be clear from the context when θ denotes a vector of parameters and when it denotes a single parameter.

Proof. For any two payoffs x and y we have the Cauchy-Schwartz inequality

$$E^2(xy) \leq E(x^2)E(y^2)$$

This inequality becomes an equality for $x = y$. Therefore we have

$$E(x^2) = \sup_y \frac{E^2(xy)}{E(y^2)}$$

where the supremum is over all y. Let x be the minimum-variance payoff. Then x is in P. Any other payoff y in P can be written as

$$y = \sum_{i=1}^{n} a_i \frac{\pi_{\theta_i}}{\pi_{\theta_0}}$$

We then have

$$\begin{aligned}\sigma^2(x) &= E[(x - E(x))^2] \\ &= \sup_y \frac{E^2[(x - E(x))y]}{E(y^2)}\end{aligned}$$

But

$$\begin{aligned}E[(x - E(x))y] &= E\left[(x - Ex)\left(\sum_{i=1}^{n} a_i \frac{\pi_{\theta_i}}{\pi_{\theta_0}}\right)\right] \\ &= \sum_{i=1}^{n} a_i \left[E\left(x\frac{\pi_{\theta_i}}{\pi_{\theta_0}}\right) - E(x)E\left(\frac{\pi_{\theta_i}}{\pi_{\theta_0}}\right)\right] \\ &= \sum_{i=1}^{n} a_i(\theta_i - \theta_0)\end{aligned}$$

Substituting y and $E[(x - E(x))y]$ in the above formule for $\sigma^2(x)$ gives the desired result. □

The three lemmas together constitute Barankin's theorem ([Itô93] page 1488) for finite sample spaces.

Theorem 4.7 Barankin. *Let M be the set of all unbiased estimators of a parameter θ. Assume that M is not empty. Then there exists an estimator x in M that minimizes the variance at θ_0 within M. Actually $\{x\} = M \cap P$ where P is the linear space generated by $\{\frac{\pi_\theta}{\pi_{\theta_0}} | \theta \in \Theta\}$. The minimum variance is given by*

$$V(x) = \inf_{y \in M} V(y) \tag{4.8}$$

$$= \sup_{a_i} \frac{\left(\sum_{i=1}^{n} a_i(\theta_i - \theta_0)\right)^2}{E\left[\left(\sum_{i=1}^{n} a_i \frac{\pi_{\theta_i}}{\pi_{\theta_0}}\right)^2\right]} \tag{4.9}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} (\theta_i - \theta_0)(\theta_j - \theta_0)\lambda^{ij} \tag{4.10}$$

where the supremum is taken over all a_i and where λ^{ij} is the (i, j)-component of the inverse of the $n \times n$ matrix $(\lambda_{ij}) = E_{\theta_0}(\frac{\pi_{\theta_i}}{\pi_{\theta_0}} \frac{\pi_{\theta_j}}{\pi_{\theta_0}})$ and the $\frac{\pi_{\theta_i}}{\pi_{\theta_0}}$ form a basis of S.

Proof. Equation 4.9 follows from lemma 4.6 and equation 4.10 follows from lemma 4.5. □

Two special cases of Barankin's theorem are the Chapman-Robbins-Kiefer inequality and the Cramér-Rao inequality ([Itô93] ibidem).

Theorem 4.8 Chapman-Robbins-Kiefer inequality. *Let x be the minimum-variance unbiased estimator. Then*

$$\sigma^2(x) \geq \sup_{\theta\in\Theta} \frac{(\theta-\theta_0)^2}{E\left[\left(\frac{\pi_\theta}{\pi_{\theta_0}} - \frac{\pi_{\theta_0}}{\pi_{\theta_0}}\right)^2\right]} \tag{4.11}$$

Proof. This is a direct consequence of lemma 4.6 with $a_1 = -1$, $a_2 = 1$ and $a_i = 0$ for $i > 2$. As lemma 4.6 is valid for any choice of basis θ_i, we can choose a basis such that $\theta_1 = \theta_0$ and $\theta_2 = \theta$. We then obtain

$$\sigma^2(x) \geq \frac{(\theta-\theta_0)^2}{E\left[\left(\frac{\pi_\theta}{\pi_{\theta_0}} - \frac{\pi_{\theta_0}}{\pi_{\theta_0}}\right)^2\right]} \tag{4.12}$$

This equality is valid for any choice of θ. So we may replace the right hand side by its supremum. □

Theorem 4.9 Cramér-Rao inequality. *Let x be the minimum-variance unbiased estimator. Then, under certain regularity conditions, we have the inequality*

$$\sigma^2(x) \geq \frac{1}{E\left[\left(\frac{\partial log \pi_\theta}{\partial\theta}|_{\theta=\theta_0}\right)^2\right]} \tag{4.13}$$

Proof. We will not give a completely rigorous proof but limit ourselves to the intuitive argument. We apply the Chapman-Robbins-Kiefer inequality to $\theta = \theta_0 + h$, divide numerator and denominator by h^2 and take the limit for h to zero:

$$\begin{aligned}
\sigma^2(x) &\geq \sup_h \frac{(\frac{\theta_0+h-\theta_0}{h})^2}{E\left[\left(\frac{\frac{\pi_{\theta_0+h}-\pi_{\theta_0}}{h}}{\pi_{\theta_0}}\right)^2\right]} \\
&\geq \lim_{h\to 0} \frac{1}{E\left[\left(\frac{\frac{\pi_{\theta_0+h}-\pi_{\theta_0}}{h}}{\pi_{\theta_0}}\right)^2\right]} \\
&= \frac{1}{E\left[\left(\frac{\frac{\partial\pi_\theta}{\partial\theta}|_{\theta=\theta_0}}{\pi_{\theta_0}}\right)^2\right]} \\
&= \frac{1}{E\left[\left(\frac{\partial log \pi_\theta}{\partial\theta}|_{\theta=\theta_0}\right)^2\right]}
\end{aligned}$$

□

Chapter 5

Conclusion

This thesis contains a systematic study of payoff spaces and their application to mean-variance analysis and minimum- variance estimators. This study leads to the following conclusions.

First, payoff spaces are a natural framework for the study of mean-variance analysis. We have shown that the minimum-variance frontier of a subspace of a payoff space is a line, the minimum-variance line, and that each minimum-variance line is the orthogonal projection of the minimum-variance line of the total payoff space. Working with lines and, in general, linear and affine subspaces is simpler and more intuitive than working with hyperbola as in classical mean-variance analysis. We have shown how many classical results can be formulated and proved in a geometric and intuitive way.

Second, mean-variance analysis in payoff space is a natural tool for the study of minimum-variance estimators. We have given a geometric-algebraic proof of Barankin's theorem. Even though the proof covers only the special case of finite sample spaces and samples of size one, it is not unlikely that the proof contains the essential elements of the general case. From the proof it also is clear that Barankin's theorem is essentially a geometric-algebraic theorem. It is not only more general than its derivative, the Cramér-Rao theorem, but also more natural in that it is purely geometric-algebraic and does not require any concepts from analysis.

Third, payoff spaces are an excellent pedagogical tool for the study of statistics in general. We think that certain parts of probability theory and statistics might benefit from a geometric-algebraic treatment within the payoff space framework. Limiting ourselves to finite sample spaces may considerably simplify a problem. Once the problem has been clarified and solved in the finite dimensional payoff space framework, we may generalize the results obtained to the general case of an infinite dimensional vector space of random variables defined on a general sample space. For many problems this generalization will only involve technicalities that are independent of the problem itself.

Bibliography

[Coc97] John H. Cochrane. Asset pricing. Unpublished manuscript. Graduate School of Business. University of Chicago, April 1997.

[dR97] Frans de Roon. *Essays on Testing for Spanning and on Modeling Futures Risk Premia.* PhD thesis, Tilburg University, September 1997.

[HJ91] Lars Peter Hansen and Ravi Jagannathan. Implications of security market data for models of dynamic economics. *Journal of Political Economy*, 99(2):225–262, 1991.

[HK87] Ger Huberman and Shmuel Kandel. Mean-variance spanning. *Journal of Finance*, 42(4):873–888, 1987.

[Itô93] Kiyosi Itô, editor. *Encyclopedic Dictionary of Mathematics.* The MIT Press, Cambridge, Massachusetts and London, England, 1993.

[Lin65] John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 67(1):13–37, 1965.

[Mar52] Harry Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[PDN90] Jacques Pontier, Anne-Béatrice Dufour, and Myriam Normand. *Le Modèle Euclidien en Analyse des Données.* Editions de l'Université de Bruxelles, Bruxelles, 1990.

[Rol77] Richard Roll. A critique of the asset pricing theory's tests. Part I: On past and potential testability of the theory. *Journal of Financial Economics*, 4:129–176, 1977.

[Rol80] Richard Roll. Orthogonal portfolios. *Journal of Financial and Quantitative Analysis*, 15(5):1005–1023, 1980.

[Sha64] William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

Index