UNIVERSITÀ DEGLI STUDI DI ROMA "TOR VERGATA"

FACOLTA' DI SCIENZE MATEMATICHE FISICHE E NATURALI

DOTTORATO DI RICERCA IN
MATEMATICA

CICLO DEL CORSO DI DOTTORATO
XIX

Titolo della tesi

BAYESIAN ALLOCATION USING THE SKEW NORMAL DISTRIBUTION

Nome e Cognome del dottorando

Francesco Simone Blasi

Coordinatore: Prof. S. Scarlatti

Contents

List of Figures

0

Introduction

Since the seminal work of Markowitz [25], portfolio theory has been improved in many directions. In order to adapt this theory to the great varieties of investment opportunities in modern large markets, many modifications of the mean-variance analysis have been developed.

One of the greatest limits of the theory of Markowitz consists in the assumption that all investors preferences can only be represented by the mean and the variance of returns. This assumption is coherent with the utility maximization only in two cases: either securities returns are assumed to be elliptically distributed or the investors utility function is quadratic (see for instance Ingersoll [18]).

Both hypothesis are subject to two traditional criticisms. As far as the elliptical distributional hypothesis is concerned, many papers show the importance to include higher moments of the portfolio return in the investment process (see e.g. [30] for the stock market). In [19], it is shown that, even though hedge funds indices are often very attractive in mean-variance terms, this is much less the case when skewness and kurtosis are taken into account. The restriction on the class of utility functions also leads to several inconsistences.

The non-elliptical modeling of financial returns has been the subject of many papers. In our opinion, the Skew-Normal distribution of Azzalini and Dalla Valle [5], represents one of the most attractive options. This distribution has been used by Meucci in [28] for investment problems.

This choice has two main advantages: first, in opposition to many modeling proposals, the skew-normal distribution has a coherent multivariate formulation; second, this

class maintains many useful properties that are typical of the normal distribution. Its characteristics are analyzed in details in [3], [4] and [9].
It is worth recalling that a random vector $\boldsymbol{Z} \in \mathbb{R}^n$ follows a skew-normal distribution with location parameter $\boldsymbol{\mu} \in \mathbb{R}^n$, $(n \times n)-$scale matrix Ω and shape parameter $\boldsymbol{\alpha} \in \mathbb{R}^n$ if its density has the following form:

$$f_{\boldsymbol{Z}}(\boldsymbol{x}) = 2\varphi_n(\boldsymbol{x}; \boldsymbol{\mu}, \Omega)\Phi(\boldsymbol{\alpha}^T \omega^{-1}(\boldsymbol{x} - \boldsymbol{\mu})) \tag{1}$$

where ω is the diagonal matrix $\omega = \text{diag}(\sqrt{\Omega_{11}}, \ldots, \sqrt{\Omega_{nn}})$, $\sqrt{\Omega_{ii}} \neq 0$ for $i = 1, \ldots, n$, $\varphi_n(\boldsymbol{x}; \boldsymbol{\mu}, \Omega)$ is the density of a $N_n(\boldsymbol{\mu}, \Omega)$-random vector and $\Phi(x)$ is the cumulative distribution function of a univariate standard normal. In this case we write $\boldsymbol{Z} \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$.
The main properties of the skew-normal distribution are analyzed in **Chapter** 1, in which the most basic facts are proved. Several modifications of this distribution have also been developed in literature. For instance, Liseo and Loperfido presented the so called "hierarchical" skew-normal (see [23]): this distribution has been conceived in a Bayesian framework, therefore showing the great adaptability of the skew-normal class to Bayesian inference. This class can also be obtained by conditioning a multivariate normal to one of its components, with a constraint on this component. The same generation method has been applied to other elliptical distributions, in particular to the t-Student. The classes of distributions generated in this way are called, respectively, skew-elliptical and skew-t. Sahu et al. in [34] introduced a slight modification of this method for linear regression models where errors are assumed skew-t. Many interesting applications of the skew-elliptical class are exposed in the survey of Genton [14] .
In this thesis, the skew-normal distribution is used to model securities returns. For the purpose of our research, it is important to focus on this modeling assumption (explicitly made in **Chapter** 4):
Given a vector of location parameters $\boldsymbol{\mu} \in \mathbb{R}^n$, a vector of parameters (related to the shape parameter $\boldsymbol{\alpha}$) $\boldsymbol{\delta} \in \mathbb{R}^n$, a diagonal matrix of standard deviations ω and a correlation matrix Ψ, we assume that the vector of returns of n risky securities is described by the following model:

$$\begin{aligned} \boldsymbol{R} = \boldsymbol{\mu} + (\omega\boldsymbol{\delta})|X| \quad &+ \quad \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z} \\ X \quad &\sim \quad N(0,1) \\ \boldsymbol{Z} \quad &\sim \quad N_n(0, \Psi) \end{aligned} \tag{2}$$

where X, $\boldsymbol{Z}$ are independent r.v.'s, $\Delta = \text{diag}(\delta_1, \ldots, \delta_n)$ and Id denotes the identity matrix (see Azzalini [3] on this representation). The link between (2) and (1) is given by the following result:

Proposition. *The random vector* $\boldsymbol{R}$ *given by (2) is skew normally distributed. More precisely* $\boldsymbol{R} \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$ *where*

$$\begin{aligned} \Omega &= \omega\overline{\Omega}\omega \\ \overline{\Omega} &= \boldsymbol{\delta}\boldsymbol{\delta}^T + (Id - \Delta^2)^{1/2}\Psi(Id - \Delta^2)^{1/2} \\ \boldsymbol{\alpha} &= \frac{\overline{\Omega}^{\,-1}\boldsymbol{\delta}}{(1 - \boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta})^{1/2}} \end{aligned}$$

Note that for $\boldsymbol{\delta} = \mathbf{0}$ we have $\Delta = 0$ (the null-matrix) and therefore from (2)

$$\boldsymbol{R} \sim N_n(\boldsymbol{\mu}, \omega\Psi\omega).$$

In the realistic example presented at the end of this thesis, we will show that for the hedge funds market, the hypothesis $\boldsymbol{\delta} \neq \mathbf{0}$ can be accepted (i.e. $\boldsymbol{\delta} = \mathbf{0}$ is rejected), validating the skew normal assumption.
We now come to the asset allocation problem which represents a considerable part of this work and has significant practical implications.
We assume that an investor selects at time t a portfolio of assets which is hold unchanged until time $\tau > t$. A portfolio is defined as a vector

$$\boldsymbol{w} \in \mathbb{R}^n \ \text{ such that} \sum w_i = 1.$$

By definition the portfolio return at time τ is a realization of the univariate random variable

$$R_{\boldsymbol{w}} = \boldsymbol{w}^T\boldsymbol{R}$$

Given two portfolios and considered their random returns $R_{\boldsymbol{w}_1}$ and $R_{\boldsymbol{w}_2}$, an investor is faced with the following problem:

Among the two portfolios $\mathbf{w}_1$ *and* $\mathbf{w}_2$*, which one should be preferred?*

If $R_{\boldsymbol{w}_2}(\cdot) \geq R_{\boldsymbol{w}_1}(\cdot)$ was true for each scenario then the choice would be obvious. Nonetheless portfolio returns usually do not satisfy the previous simple dominance relation. Indeed $R_{\boldsymbol{w}_1}$ and $R_{\boldsymbol{w}_2}$ may have intersecting probability densities on ample regions of the returns space.

The Stochastic Dominance theory (**SD**) mainly developed by Levy in [21], [22] and by Levy and Hanoch in [16], proposes a method to handle this problem. The final output of **SD** is a set of reasonable rules which should be considered as guidelines for the behavior of rational investors depending on their utility functions.

In **Chapter** 2, after having briefly recalled the principal aspects of **SD**, we develop this theory for skew-normal variables.

In short, the setup of **SD** is the following one: given two univariate random variables R_1 and R_2, which in this context are often called uncertain prospects, we write $R_1 \succeq R_2$ if R_1 is preferred to R_2. We denote by $\mathcal{U}_i$ the following sets of utility functions:

$$\mathcal{U}_1 = \{u \in \mathcal{C}^1(\mathbb{R}) \text{ with } u'(x) \geq 0\}$$

and

$$\mathcal{U}_2 = \{u \in \mathcal{C}^2(\mathbb{R}) \text{ with } u'(x) \geq 0, u''(x) \leq 0\},$$

then the core of **SD** relies in the following definition:

Definition.
(1)We say that R_1 stochastically dominates at first order R_2 , and write $R_1 \succeq_1 R_2$, if:

$$\mathbb{E}(u(R_1)) - \mathbb{E}(u(R_2)) \geq 0 \tag{3}$$

for every $u \in \mathcal{U}_1$.
(2)We say that R_1 stochastically dominates at second order R_2 , and write $R_1 \succeq_2 R_2$, if:

$$\mathbb{E}(u(R_1)) - \mathbb{E}(u(R_2)) \geq 0 \tag{4}$$

for every $u \in \mathcal{U}_2$.
*(whenever the inequalities (3) or (4) hold strictly for at least one u then we say that **SD** holds in strong sense).*

The use of **SD** criteria for portfolios ranking is motivated by this important classical result:

Proposition. *Given two uncertain prospects $R_1 \sim N(\mu_1, \sigma_1)$ and $R_2 \sim N(\mu_2, \sigma_2)$, suppose $\mu_1 \geq \mu_2$ and $\sigma_1 \leq \sigma_2$. Then $R_1 \succeq_2 R_2$.*

In the applications to portfolio selection and management the uncertain prospects are the returns of different portfolios. Due to the fact that the case of increasing and concave utility, describing risk-averse agents, is the most realistic one, the previous Proposition has a great worth: if two normal prospects have the same mean, the one with the smaller variance is preferable .
The following natural question arises: how can the previous Proposition be extended to skew-normal prospects whose distribution, compared to the normal law, contains a further parameter ?
Our results are summarized by the following Propositions proved in **Chapter** 2:

Proposition.
(i)Let $R_1 \sim SN(\mu, \sigma^2, \alpha_1)$ *and* $R_2 \sim SN(\mu, \sigma^2, \alpha_2)$ *be skew-normal r.v's. Suppose* $\alpha_1 \geq \alpha_2$ *then* $R_1 \succeq_1 R_2$.
(ii)Let $R_1 \sim SN(\mu, \sigma_1^2, \alpha)$ *and* $R_2 \sim SN(\mu, \sigma_2^2, \alpha)$ *be skew-normal r.v's with* $\alpha \leq 0$. *Suppose* $\sigma_1 \leq \sigma_2$ *then* $R_1 \succeq_2 R_2$.
(iii)Let $R_1 \sim SN(\mu_1, \sigma^2, \alpha)$ *and* $R_2 \sim SN(\mu_2, \sigma^2, \alpha)$ *be skew-normal r.v's. Suppose* $\mu_1 \geq \mu_2$ *then* $R_1 \succeq_1 R_2$.

Proposition. *Let* $R_1 \sim SN(\mu, \sigma_1^2, \alpha_1)$ *and* $R_2 \sim SN(\mu, \sigma_2^2, \alpha_2)$ *be skew-normal r.v's. Suppose* $\sigma_1 \leq \sigma_2$ *and* $\sigma_1\delta_1 = \sigma_2\delta_2$ *then* $R_1 \succeq_2 R_2$.

In **Chapter** 3 we outline the main aspects of the classical portfolio selection theory, due to Markowitz [25]. In this framework the preferences of the investor are simply codified by attitude towards mean and aversion towards variance. When market returns are normal so are portfolio returns and therefore the approach of Markowitz ranks portfolios in the same way as **SD** for risk-averse investors.
In the framework presented by Markowitz, the single investor fixes a level of expected portfolio return E and solves the following quadratic problem:

$$\begin{aligned} \text{Min}_{\boldsymbol{w}} \quad & \text{Var}(R_{\boldsymbol{w}}) \\ \text{with the constraints:} \quad & \mathbb{E}(R_{\boldsymbol{w}}) = E \\ & \mathbf{1}^T \boldsymbol{w} = 1 \end{aligned} \tag{5}$$

Equivalently the agent fixes a level of risk, i.e. of standard deviation of the portfolio returns, and tries to maximize the expected return of the portfolio. The space of means and variances of portfolio returns is called the mean-variance space (**MV**).

Note that no assumption is made on the distributional form of the random vector $\boldsymbol{R}$ describing returns (which drives the portfolio return $R_{\boldsymbol{w}}$), apart from having finite means and covariances.

Markowitz in his approach does not really take into consideration the Expected Utility (**EU**) maximization procedure, in the sense of Von-Neumann and Morgensten, rather he presents his theory as a self-referring one and looks forward to the practical implications. It is well known that, due to the absence of short-sale constraints[1], problem (5) has an explicit solution. The set of all solutions, spanned by varying the level E, is called the *minimum variance* set and its generic element is given by:

$$\boldsymbol{w}^* = \lambda_1 \frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} + \lambda_2 \frac{V^{-1}\boldsymbol{m}}{\mathbf{1}^T V^{-1}\boldsymbol{m}} \tag{6}$$

where $\lambda_i = \lambda_i(E)$, $i = 1, 2$, are such that $\lambda_1 + \lambda_2 = 1$, $\boldsymbol{m} = \mathbb{E}(\boldsymbol{R})$ and $V = \mathrm{Var}(\boldsymbol{R})$ (assumed to be not singular).

As it is evident, each portfolio given by (6) is a linear combination of only two portfolios:

$$\boldsymbol{w}_1 = \frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}}, \quad \boldsymbol{w}_2 = \frac{V^{-1}\boldsymbol{m}}{\mathbf{1}^T V^{-1}\boldsymbol{m}}.$$

Moreover it can be easily shown that any two distinct minimum variance portfolios will serve in place of $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$. This result is the first example of a mutual fund separation theorem : investors, accordingly to their selected level E of portfolio return, can form their *minimum variance* portfolio simply by buying (or selling) different amounts of $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$.

Let us now assume that the set of n basic market securities is enlarged by adding an extra riskless asset, which assures to the investor a fixed not random return R_f. We shall set $R_0 = R_f$. Portfolios are now vectors with $n+1$-components. It can be easily proven that in this case the solution to (5) is:

$$\boldsymbol{w}^* = \lambda_1 \boldsymbol{w}_f + \lambda_2 \boldsymbol{w}_t \tag{7}$$

with $\lambda_1 + \lambda_2 = 1$, and where:

$$\boldsymbol{w}_f = (1, 0, \ldots, 0), \qquad \boldsymbol{w}_t = \left(0, \frac{V^{-1}(\boldsymbol{m} - \mathbf{1}R_f)}{\mathbf{1}^T V^{-1}(\boldsymbol{m} - \mathbf{1}R_f)}\right)$$

This is again a linear combination of only two portfolios, the portfolio $\boldsymbol{w}_f$ and the "so called" tangency portfolio $\boldsymbol{w}_t$. Clearly, in this case, only the second one is risky. This

[1] In all this thesis the portfolio selection problem is studied admitting short-selling

implies that the risky component of the investor's portfolio is represented solely by the proportion of tangency portfolio that the investor holds.

Now suppose there are M rational agents investing in the market. Assume that they all are mean variance maximizers in the sense of (5). Furthermore assume they all have the same estimates of the means and of the covariances of the returns of the risky assets. These assumptions, together with some other ones fully listed in **Chapter** 3, lead to a classical result known as Capital Asset Pricing Model (CAPM), which gives the market securities expected returns and prices at equilibrium. The intuition behind the CAPM of Sharpe and Linter is based on the identification of the tangency portfolio with the market portfolio $\boldsymbol{w}_m$. Market portfolio is defined as the portfolio consisting of all securities in the market, where the weight of each security is given by its relative market value, that is, the aggregate market value of the security divided by the sum of the aggregate market values of all securities. The reasoning applied to the derivation of the CAPM can be summarized in this way: in equilibrium the aggregate demand of risky assets, which is solely represented by the tangency portfolio, is equal to the total supply of risky assets, i.e. the market portfolio. Being the tangency portfolio an efficient one so is the market portfolio at equilibrium.

The well known CAPM pricing equation, a consequence of the above identification and of the tangency condition, is then the following:

$$\boldsymbol{m}^e - \mathbf{1}R_f = \boldsymbol{\beta}_m(m_m - R_f) \tag{8}$$

where $m_m = \mathbb{E}(R_{\boldsymbol{w}_m})$, $\beta_{i,m} = \text{Cov}(R_i, R_{\boldsymbol{w}_m})/\text{Var}(R_{\boldsymbol{w}_m})$ and $\boldsymbol{m}^e$ is the vector of expected securities returns at equilibrium.

However if the investor's preferences are not fully captured by mean and variance of the portfolio return then the use of program (5) as decision rule for portfolio selection is questionable either from the point of view of **SD** or from that of **EU** (unless market returns are normal , or more in general elliptical). As a consequence the CAPM itself, in the form derived above, is affected by these limitations or inconsistencies.

To overcame these problems we consider in this thesis a more general approach due to Simaan [35]. More precisely, in **Chapter** 4 the framework considered by Simaan is explicitly worked out for the case of skew normal returns. Simaan's methodology provides an interesting extension of the mean variance analysis since it incorporates in an elegant way investors preferences on skewness.

We remark that our interest in this extension is mainly motivated by the need of obtaining an equilibrium result similar to (8), which is valid in case of skew-normal returns. This step is necessary in order to produce a suitable generalization of Black Litterman portfolio selection model, which we discuss in **Chapter** 5 and shortly present later on. In the Simaan framework each investor controls the choice of his portfolio through the control of the location parameter $\boldsymbol{w}^T\boldsymbol{\mu}$ and of the spherical and non-spherical components of the variance of the portfolio return $R_{\boldsymbol{w}}$, which, assuming (2) for the market returns distribution, is given by:

$$\mathrm{Var}(R_{\boldsymbol{w}}) = \boldsymbol{w}^T W \boldsymbol{w} + \left(1 - \frac{2}{\pi}\right) \boldsymbol{w}^T[(\omega\boldsymbol{\delta})(\omega\boldsymbol{\delta})^T]\boldsymbol{w}. \tag{9}$$

where $W = (Id - \Delta^2)^{1/2}\Psi(Id - \Delta^2)^{1/2}$. In (9) the first addendum represents the spherical component of the variance while the second one the non spherical component. With relation to portfolios, we shall call the three dimensional space (location,variance,non-spherical component of the variance) the location-variance-skewness space (**LVS**)[1] .The name "non spherical" component of the variance derives from the expression of the skewness of $R_{\boldsymbol{w}}$:

$$\mathrm{Skew}(R_{\boldsymbol{w}}) = \mathrm{Skew}(|X|) \cdot (\boldsymbol{w}^T(\omega\boldsymbol{\delta}))^3$$

which is non zero, unless $\boldsymbol{\delta} = \boldsymbol{0}$.
Maximization of expected utility turns out to be equivalent, for a risk-averse investor, to the following portfolio optimization program in the (**LVS**) space[2]:

$$\begin{aligned} \mathrm{Min}_{\boldsymbol{w}} \quad & \tfrac{1}{2}\boldsymbol{w}^T W \boldsymbol{w} \\ \text{with the constraints:} \quad & \boldsymbol{w}^T\boldsymbol{\mu} = L \\ & \boldsymbol{w}^T(\omega\boldsymbol{\delta}) = B \\ & \boldsymbol{1}^T\boldsymbol{w} = 1 \end{aligned} \tag{10}$$

for fixed $L, B \in \mathbb{R}$.
The attractive property of (5), which is to admit an explicit solution, remains true for

[1] In his paper ,([35]), Simaan represents portfolios in the mean-variance-skewness space (**MVS**). There is a one-to-one correspondence between points in the (**LVS**) space and points in the (**MVS**) space (see chapter 4). When $\boldsymbol{\delta} = \boldsymbol{0}$ both spaces coincide with the mean-variance space (**MV**). More important, when portfolio returns are skew-normal the expected utility of a risk-verse investor behaves, as function of the parameters, much in the same way over the above two spaces.

[2] We recall that here $\boldsymbol{\mu}$ identifies the locations of the returns $\boldsymbol{R}$, while in Simaan's paper the same symbol denotes the returns means. They coincide only for $\boldsymbol{\delta} = \boldsymbol{0}$.

the problem (10). We call the set of solution the *minimum spherical variance* set. This set includes the single portfolios with the smallest spherical part of the variance for given levels of location parameter and of non-spherical component of the variance. The generic solution can be expressed (without the riskless asset) as:

$$\boldsymbol{w}^* = \lambda_1 \frac{V^{-1}\boldsymbol{\mu}}{\mathbf{1}^T V^{-1}\boldsymbol{\mu}} + \lambda_2 \frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} + \lambda_3 \frac{V^{-1}(\omega\boldsymbol{\delta})}{\mathbf{1}^T V^{-1}(\omega\boldsymbol{\delta})} \tag{11}$$

where $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $V = \text{Var}(\boldsymbol{R})$.

Adding as before the riskless asset R_f to the vector of returns, the solution to (10) is given by:

$$\boldsymbol{w}^* = \lambda_1 \boldsymbol{w}_f + \lambda_2 \boldsymbol{w}_t + \lambda_3 \boldsymbol{w}_3 \tag{12}$$

where $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and where:

$$\boldsymbol{w}_t = \left(0, \frac{V^{-1}(\boldsymbol{\mu} - \mathbf{1}R_f)}{\mathbf{1}^T V^{-1}(\boldsymbol{\mu} - \mathbf{1}R_f)}\right), \quad \boldsymbol{w}_3 = \left(0, \frac{V^{-1}(\omega\boldsymbol{\delta})}{\mathbf{1}^T V^{-1}(\omega\boldsymbol{\delta})}\right)$$

It is evident both in (11) and (12) that the set of solution is now a linear combination of only three portfolios. The validity of this property leads, as for the mean variance analysis, to an equilibrium result. The pricing model we obtain is fully based on Simaan results in [35], and is different from the one of Kraus and Litzenberger [20] and of Adcock [1]. Kraus and Litzenberger CAPM is based on a three moment Taylor expansion of the utility function ignoring higher moments.

This modified CAPM, obtained under the assumption (2) of skew-normality of the random vector $\boldsymbol{R}$, relates the location parameters of the assets, μ_i, to the location parameters μ_m and μ_p of the portfolios $\boldsymbol{w}_m$ and $\boldsymbol{w}_p$ respectively. Here $\boldsymbol{w}_m$ is the market portfolio defined above and $\boldsymbol{w}_p$ is defined as a portfolio whose return is uncorrelated to the market portfolio return but which has its same skewness.

The CAPM equation we obtain is the following:

$$\boldsymbol{\mu}^e = \mathbf{1}R_f + \boldsymbol{\beta}_m[\mu_m - \mu_p] + (\boldsymbol{\gamma}_m - \boldsymbol{\beta}_m)[\mu_p - R_f] \tag{13}$$

where as before $\beta_{i,m} = \text{Cov}(R_i, R_{\boldsymbol{w}_m})/\text{Var}(R_{\boldsymbol{w}_m})$ and $\gamma_{i,m} = (\omega_i\delta_i)/B_m$ with $B_m = (\boldsymbol{w}_m^T(\omega\boldsymbol{\delta}))$

It can be proven that the *minimum spherical variance* set in the (**LVS**) space is an elliptical paraboloid.

The result (adapted from Simaan [35]), in absence of a risk-less asset, is stated below. Let us denote $B_i = (\omega\boldsymbol{\delta})^T \boldsymbol{a}_i$ and $L_i = \boldsymbol{\mu}^T \boldsymbol{a}_i$ where $\boldsymbol{a}_i$, $i = 1, 2, 3$ are the following portfolios:

$$\boldsymbol{a}_1 = \frac{V^{-1}\boldsymbol{\mu}}{\mathbf{1}^T V^{-1}\boldsymbol{\mu}}, \quad \boldsymbol{a}_2 = \frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} \text{ and } \boldsymbol{a}_3 = \frac{V^{-1}(\omega\boldsymbol{\delta})}{\mathbf{1}^T V^{-1}(\omega\boldsymbol{\delta})} \tag{14}$$

and denote by $B = (\omega\boldsymbol{\delta})^T \boldsymbol{w}$ and $L = \boldsymbol{\mu}^T \boldsymbol{w}$, then the following Proposition holds:

Proposition. *If there's no risk less asset and returns are skew normally distributed according to (2), the efficient set in the $(L, \mathcal{V}, B)-$space (the **LVS** space) is given by:*

$$\mathcal{V}^2 = \boldsymbol{w}^T V \boldsymbol{w} = \sigma_2^2 + \sigma_{h_3}^2 \left(\frac{B - B_2}{B_3 - B_2}\right)^2 + \sigma_{h_1}^2 c_1^2 \tag{15}$$

where $\sigma_2^2 = \boldsymbol{w}^T \boldsymbol{a}_2 \boldsymbol{w}$, $\sigma_{h_i}^2 = \boldsymbol{w}^T \boldsymbol{h}_i \boldsymbol{w}$,

$$c_1 = \frac{(L - E_2)/(E_3 - E_2) - (B - B_2)/(B_3 - B_2)}{(E_1 - E_2)/(E_3 - E_2) - (B_1 - B_2)/(B_3 - B_2)}$$

and portfolios $\boldsymbol{h}_i$, $i = 1, 3$, *are given by*

$$\boldsymbol{h}_1 = (\boldsymbol{a}_1 - \boldsymbol{a}_2) - \frac{B_1 - B_2}{B_3 - B_2}(\boldsymbol{a}_3 - \boldsymbol{a}_2) \;\; ; \;\; \boldsymbol{h}_3 = \boldsymbol{a}_3 - \boldsymbol{a}_2$$

The final parts of **Chapters** 3 and 4 are devoted to the presentation of the "so-called" *mutual funds separation results*. We have already presented in this introduction a particular instance of these type of theorems. More general results are available. They are directly linked to the derivation of CAPM in a general framework and for the sake of completeness are included in this thesis.

To give a rough description of general mutual funds separation results it is useful to introduce first the following definitions (Ross [32])

Definition. *We shall say that,*

(A) the distribution of $\boldsymbol{R}$ has the (strong) 2-funds separation property if there exist two portfolios $\boldsymbol{w}_1, \boldsymbol{w}_2$ such that for any portfolio $\boldsymbol{w}_b$ there is a portfolio $\boldsymbol{w}_a$ given by a linear combination of $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ for which it holds

$$\boldsymbol{w}_a^T \boldsymbol{R} \succeq_2 \boldsymbol{w}_b^T \boldsymbol{R}.$$

(B) the distribution of $\boldsymbol{R}$ has the (strong) 3-funds separation property if there exist three portfolios $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ such that for any portfolio $\boldsymbol{w}_b$ there is a portfolio $\boldsymbol{w}_a$ given by a linear combination of $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ for which it holds

$$\boldsymbol{w}_a^T \boldsymbol{R} \succeq_2 \boldsymbol{w}_b^T \boldsymbol{R}.$$

In a similar fashion k-funds separation ($k \geq 4$) can be defined and discussed. The expression for the efficient portfolios given by (6) and (11) strongly suggests that the normal and skew-normal distributions have respectively the (strong) 2-funds and 3-funds separation property.
Ross in [32] gives a set of properties to be satisfied by classes of market return distributions in order to have a k-funds separation property:

Theorem. *1 (2-funds separation with a risk-less asset) The distribution of* $\boldsymbol{R}$ *has the (strong) 2-funds separation property iff*
there exist a scalar r.v. Y*, a vector r.v.* $\boldsymbol{\epsilon}$*, a (deterministic) vector* $\boldsymbol{b}$*, and two portfolios* $\boldsymbol{\alpha}, \boldsymbol{\beta}$ *such that*
(a) each component of $\boldsymbol{R}$ *can be written as*

$$R_i = R_f + b_i Y + \epsilon_i, \quad for \;\; i = 1, \ldots, n+1$$

(b) $\mathbb{E}[\epsilon_i | Y] = 0 \;\; \forall i$
(c) $\sum \alpha_i \epsilon_i = 0 = \sum \beta_i \epsilon_i$.

Theorem. *2 (3-funds separation with a risk-less asset) The distribution of* $\boldsymbol{R}$ *has the (strong) 3-funds separation property iff:*
there exist two univariate r.v. Y *and* Q*, a random vector* $\boldsymbol{\epsilon}$ *, two (deterministic) vectors* $\boldsymbol{b}$*,* $\boldsymbol{c}$ *and three portfolios* $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ *such that:*
(a) each component of $\boldsymbol{R}$ *can be written as*

$$R_i = R_f + b_i Y + c_i Q + \epsilon_i \; for \; i = 1, \ldots, n+1$$

(b) $\mathbb{E}[\epsilon_i | Y, Q] = 0 \;\; \forall i$
(c) $\boldsymbol{\alpha}_i^T \boldsymbol{\epsilon} = 0$ *for* $i = 1, 2, 3$

In the thesis we show that normal and skew normal distributions satisfy the conditions of Theorems 1 and 2 respectively.

We now come to the exposition of the arguments discussed in **Chapter** 5.
The first step of the investment process, from the investor point of view, consists in the estimation of the parameters $\boldsymbol{\theta}$ contained in the distribution of the market returns $\boldsymbol{R}$. Within the normal assumption $\boldsymbol{\theta} = (\boldsymbol{m}, V)$, whereas in the skew-normal model $\boldsymbol{\theta} = (\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$. The true values of these parameters being unknown, there are two possible approaches to the estimation problem: the classical frequentist methodology or

the Bayesian point of view.

By taking the first viewpoint $\boldsymbol{\theta}$ is estimated directly using historical time series of returns: the result is a set of classical estimated parameters that we denote by $\hat{\boldsymbol{\theta}}$. These values can be used as imput data to solve the portfolio problem. For instance, if the investor agrees with the Markowitz's approach then the values $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{m}}, \hat{V})$ are used to solve (5), with $\mathrm{Var}(R_{\boldsymbol{w}}) = \boldsymbol{w}^T \hat{V} \boldsymbol{w}$ and $\mathbb{E}(R_{\boldsymbol{w}}) = \boldsymbol{w}^T \hat{\boldsymbol{m}}$.

There are two main drawbacks with this approach, both well known in the literature. The first is that the selected portfolios could be not so diversified neither so intuitive. Concerning the second, the solution weights turn out to be very sensitive to changes in the estimated parameter values: when the investor updates his estimations he faces the problem of a drastic portfolio change.

We take a Bayesian point of view which helps to smooth the great sensitivity of the allocation to variations in the input data.

In the Bayesian approach parameters are considered random variables distributed according to a prior distribution, which we denote by $f_\Theta(\boldsymbol{\theta})$. Then it is specified a model for the observations given the parameters, represented by the likelihood density $f_{\boldsymbol{R}|\Theta}(\boldsymbol{r}|\boldsymbol{\theta})$. From these two distributions it can be obtained the posterior density for parameters using the well known rule:

$$f^{po}_{\Theta|\boldsymbol{R}}(\boldsymbol{\theta}|\boldsymbol{r}) \propto f_\Theta(\boldsymbol{\theta}) f_{\boldsymbol{R}|\Theta}(\boldsymbol{r}|\boldsymbol{\theta}) \tag{16}$$

The expected utility of the portfolio return $R_{\boldsymbol{w}}$ is then defined conditionally on parameters:

$$\mathbb{E}(u(R_w|\Theta = \boldsymbol{\theta})) = \int u(r) f_{R_{\boldsymbol{w}}|\Theta}(r|\boldsymbol{\theta}) dr \tag{17}$$

The above quantity is hence averaged over parameters by using the posterior density $f^{po}_{\Theta|\boldsymbol{R}}$, that is

$$\int \mathbb{E}(u(R_w|\Theta)) f^{po}_{\Theta|\boldsymbol{R}_w}(\boldsymbol{\theta}|\boldsymbol{r}) d\boldsymbol{\theta}. \tag{18}$$

Then the investor can look for the optimal portfolio. However a more interesting approach is to introduce the so called predictive posterior density and take the average of the (conditional) expected utility with respect to it. Finding the optimum concludes the procedure.

However a great difficulty in handling the Bayesian approach is the fact that for many distributions the posterior density cannot be computed in closed form and numerical methods like MCMC (Monte Carlo Markov Chain) are needed.
A modified version of the classic Bayesian allocation is the Black and Litterman model (**BL**) appeared in [6], and discussed in detail by Black and Litterman in [7] and by He and Litterman in [17].
In the **BL** model security returns are assumed to be normally distributed. This model has the ability to blend together investors views and a prior on assets returns. Black and Litterman trust $\hat{V}$ (sample covariance) as good estimator of V but do not trust the sample mean $\hat{\boldsymbol{m}}$. A "modified" Bayesian approach to the estimation problem of $\boldsymbol{m}$ is considered. They assume the following model for returns $\boldsymbol{R}'$[1] :

$$\begin{aligned} \boldsymbol{R}'|\boldsymbol{M} = \boldsymbol{m} &\sim N(\boldsymbol{m}; \hat{V}) \quad (19) \\ \boldsymbol{M} &\sim N(\boldsymbol{\Pi}, \tau\hat{V}) \quad (20) \end{aligned}$$

where the first requirement sets up a normal model for observed returns (given their mean), while the second chooses the prior distribution on means inside the same family. τ is a scaling parameter, usually taken small. The vector $\boldsymbol{\Pi}$ is the vector of the so called "implied returns" . It is obtained by a slight modification of the CAPM pricing equation (8) and its derivation is discussed in the chapter.
Having given a model for observations and a prior on means we could implement the previous outlined Bayesian allocation scheme. However **BL** model aims to incorporate into the investment process a further layer of information. This is achieved by inserting random constraints on the prior representing investors opinions on the expected values of the returns:

$$\boldsymbol{v} - P\boldsymbol{M} \sim N(\boldsymbol{0}, \Omega_v)$$

where $\boldsymbol{v} \in \mathbb{R}^k$ is called the vector of the views, P is a $k \times n$ matrix, Ω_v is a $k \times k$ invertible diagonal matrix and $k \leq n$. It is useful introduce the r.v. $\boldsymbol{V}$ (do not confuse this vector with the covariance matrix V), and rewrite the constraint equations in "regression form":

$$\boldsymbol{V} = P\boldsymbol{M} + \boldsymbol{\epsilon}$$

[1] More precisely $\boldsymbol{R}' \equiv \boldsymbol{R} - R_f$ are the excess of returns w.r.t the risk-free rate.

with $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \Omega_v)$. We can look to the previous relation as a model for the views (given the means), that is

$$\boldsymbol{V}|(\boldsymbol{M} = \boldsymbol{m}) \sim N(P\boldsymbol{m}, \Omega_v)$$

Having a normal prior and a model for the views (again normal), we can easily obtain the posterior law of $\boldsymbol{M}$ given the views and then, integrating over $\boldsymbol{M}$, the posterior predictive distribution of $\boldsymbol{R}|\boldsymbol{V}$. Being this distribution normal the utility is maximized by solving problem (5) and using, in place of $\mathbb{E}(R_{\boldsymbol{w}})$ and $\text{Var}(R_{\boldsymbol{w}})$, the analogous moments of $R_{\boldsymbol{w}}|\boldsymbol{V}$ which are:

$$\begin{aligned} m_{BL}^{w} &= \boldsymbol{w}^T \boldsymbol{m}_{BL} \\ \Sigma_{BL}^{w} &= \boldsymbol{w}^T (\hat{V} + \Sigma_{BL}) \boldsymbol{w} \end{aligned}$$

where:

$$\boldsymbol{m}_{BL} = [(\tau \hat{V})^{-1} + P^T \Omega_v^{-1} P]^{-1} [(\tau \hat{V})^{-1} \boldsymbol{\Pi} + P^T \Omega_v^{-1} \boldsymbol{v}] \tag{21}$$

$$\Sigma_{BL} = [(\tau \hat{V})^{-1} + P^T \Omega_v^{-1} P]^{-1} \tag{22}$$

Meucci in [28] and [29] extends the **BL** model to non normal markets relying some new ideas. This is based on the previous two stages procedure, but the Bayesian inference is replaced by an opinion pooling approach. Then he uses copulas to obtain the market returns distribution. This approach has the great worth to be adaptable to many non-normal markets.
On the contrary our approach preserves the Bayesian framework. This requires to make the assumption that returns are skew normal, then it uses the good properties of the skew-normal distribution under Bayesian inference , see Liseo and Loperfido [23].
In this way an analytical result for the predictive posterior of returns is obtained. Moreover it can be used as "benchmark" for further non-normal distributional assumptions, such as the skew-t.
As already explained, the BL model relies on considering the expected values of returns random variables whose density can be combined with the views vector $\boldsymbol{V}$. In order to preserve this intuition in **Chapter** 6 we let the assumption (2) to be valid *conditionally* on locations $\boldsymbol{\mu}$. We denote by Θ_1 the random vector of location parameters and we impose on it a normal prior. As a result, the expected values, which depend on the

location parameters, turn out to be a random vector as well.
The market model which we assume is the following:

$$\boldsymbol{R}|(\Theta_1 = \boldsymbol{\mu}) \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha}) \quad (23)$$

$$\Theta_1 \sim N_n(\boldsymbol{\mu}^e, \tau\Omega) \quad (24)$$

where the vector $\boldsymbol{\mu}^e$ is given by the pricing equation (13), Ω is the scale matrix of $\boldsymbol{R}|\Theta_1$, $\boldsymbol{\alpha}$ its shape parameter and $\tau \in \mathbb{R}$ is a small scaling factor, as previously mentioned. Our interest in the three moment CAPM is mainly due to the need of centering the prior distribution (24) on an equilibrium vector of returns.
We prove that the model given by (23) and (24) has a skew-normal marginal for $\boldsymbol{R}$, validating our main assumption of skew normality for assets returns.
Denoting the r.v. of expected values by $\boldsymbol{M} = \Theta_1 + \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta})$, the previous model can also be written in the following way:

$$\boldsymbol{R}|(\boldsymbol{M} = \boldsymbol{m}) \sim SN_n(\boldsymbol{m} - \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta}), \Omega, \boldsymbol{\alpha}) \quad (25)$$

$$\boldsymbol{M} \sim N_n(\boldsymbol{m}^e, \tau\Omega) \quad (26)$$

where $\boldsymbol{m}^e = \boldsymbol{\mu}^e + \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta})$.
As in the **BL** classical model, the views are assumed to be normal :

$$\boldsymbol{V}|(\boldsymbol{M} = \boldsymbol{m}) \sim N_n(P\boldsymbol{m}, \Omega_v), \quad (27)$$

where, as before, P is a $k \times n$ matrix and Ω_v is a $k \times k$ invertible diagonal matrix and $k \leq n$.
The posterior distribution of $\boldsymbol{M}|\boldsymbol{V}$ is given by:

$$\boldsymbol{M}|(\boldsymbol{V} = \boldsymbol{v}) \sim N(\boldsymbol{m}_{BL}, \Sigma_{BL})$$

where:

$$\boldsymbol{m}_{BL} = [(\tau\Omega)^{-1} + P^T\Omega_v^{-1}P]^{-1}[(\tau\Omega)^{-1}\boldsymbol{m}^e + P^T\Omega_v^{-1}\boldsymbol{v}]$$

$$\Sigma_{BL} = [(\tau\Omega)^{-1} + P^T\Omega_v^{-1}P]^{-1}$$

The evaluation of the posterior predictive distribution of $\boldsymbol{R}|\boldsymbol{V}$ is now possible.
The result that we find combining $\boldsymbol{R}|\boldsymbol{M}$ and $\boldsymbol{M}|\boldsymbol{V}$ and integrating over $\boldsymbol{M}$ is the following:

$$\boldsymbol{R}|\boldsymbol{V} \sim SN_n(\boldsymbol{m}_{BL} - \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta}), \Omega + \Sigma_{BL}, \boldsymbol{\alpha}_{BL}) \quad (28)$$

where

$$\begin{aligned}
\Delta_{BL} &= (\Omega^{-1} + \Sigma_{BL}^{-1})^{-1} \\
\boldsymbol{\alpha}_{BL}^T &= \boldsymbol{\alpha}^T \omega^{-1} \Omega (\Omega + \Sigma_{BL})^{-1} (1 + \boldsymbol{\alpha}_\Delta^T \overline{\Delta}_{BL} \boldsymbol{\alpha}_\Delta)^{-1/2} \\
\boldsymbol{\alpha}_\Delta^T &= -\boldsymbol{\alpha}^T \omega^{-1} d_{BL}
\end{aligned}$$

and where d_{BL} is the diagonal matrix of standard deviations of Δ_{BL} and $\overline{\Delta}_{BL}$ its correlation matrix.

After including the views in the investment process, the investor can complete the allocation process. Due to the fact that the posterior predictive distribution turns out to be skew-normal the expected utility can be maximized by the same procedure followed in chapter 4. In other terms the problem to be solved is program (10), which in this context takes the form :

$$\begin{aligned}
\text{Min}_w \qquad & \tfrac{1}{2} s_{BL}^2 & (29) \\
\text{with the constraints:} \qquad & \mu_{BL}^w = E \\
& \boldsymbol{w}^T (\gamma_{BL} \boldsymbol{\delta}_{BL}) = B \\
& \mathbf{1}^T \boldsymbol{w} = 1
\end{aligned}$$

where:

$$\mu_{BL}^w = \boldsymbol{w}^T (\boldsymbol{m}_{BL} - \sqrt{\frac{2}{\pi}} (\omega \boldsymbol{\delta})) \qquad (30)$$

$$s_{BL}^2 = \boldsymbol{w}^T (\Omega + \Sigma_{BL}) \boldsymbol{w} - (\boldsymbol{w}^T (\gamma_{BL} \boldsymbol{\delta}_{BL}))^2 \qquad (31)$$

with:

$$\boldsymbol{\delta}_{BL} = \frac{(\overline{\Omega + \Sigma}_{BL}) \boldsymbol{\alpha}_{BL}}{\sqrt{1 + \boldsymbol{\alpha}_{BL}^T (\overline{\Omega + \Sigma}_{BL}) \boldsymbol{\alpha}_{BL}}}$$

and γ_{BL} represents the diagonal matrix of standard deviations of $\Omega + \Sigma_{BL}$, and $\overline{\Omega + \Sigma}_{BL}$ its correlation matrix.

The last part of **Chapter** 6 contains the main numerical example of this thesis concerning a portfolio of 12 Hedge Fund Indexes (HFR Indexes), each one corresponding to a different Hedge Funds strategy. Our assumption is that the 12 Strategies are a good representation of the Hedge Funds Market. The example covers the entire investment

Likelihood ratio Test	**(null hypothesis: $\boldsymbol{\alpha} = \mathbf{0}$)**	
log-lik normal ($\boldsymbol{\alpha} = \mathbf{0}$)		-3967.99
log-lik skew-normal		-3923.89
lik-ratio		88
Prob		0

Table 1: Likelihood ratio test: the lik-ratio is the likelihood ratio test statistics and Prob the corresponding probability.

process in a skew normal market and its final output is a vector of weights that is generated conditionally on views.

The method used to validate the assumption of skew normality is a classical likelihood ratio test. The model with a restriction is the one with the vector $\boldsymbol{\alpha}$ of all zeros, which implies the normality of the restricted model. The values of the test, reported in the Table, have been compared with the values from a chi-squared distribution with 12 degrees of freedom.

The results are very promising: the skew-normal assumption seems to be much more appropriate for this very dynamic market.

Nomenclature

$\boldsymbol{R}$	Random vector of returns.
$\boldsymbol{w}$	Generic portfolio, i.e. $\sum_i w_i = 1$.
$R_{\boldsymbol{w}}$	Univariate random variable of portfolio returns, $R_{\boldsymbol{w}} = \boldsymbol{w}^T \boldsymbol{R}$.
u	Utility function.
n	The number of risky assets in the market.
$\mathbf{1}$	A vector whose elements are 1.
$\mathbf{0}$	A vector whose elements are 0.
$\boldsymbol{m}$	The expected value of the vector of returns.
V	The covariance matrix of the vector of returns.
$\boldsymbol{\mu}$	The location parameter for the vector of skew-normally distributed returns.
Ω	The scale matrix for the vector of skew-normally distributed returns.
$\boldsymbol{\alpha}$	The shape parameter for the vector of skew-normally distributed returns.
$\varphi_n(\cdot)$	The density of a multivariate standard normal.
$\Phi(\cdot)$	The cumulative distribution function of a univariate standard normal.
M	The number of investors in the market.

1

The family of Skewed Distributions

1.1 The univariate skew normal distribution

Lemma 1.1.1. *If f is a density symmetric with respect to 0, if G is a one-dimensional distribution function with density G' symmetric about 0, then*

$$\phi(z) = 2f(z)G(w(z))$$

is a density for every odd function $w(z)$

Proof. Denote by Y a random variable with density f, and by X a random variable with distribution function G, independent from Y. The first step in the proof consists in proving that $W = w(Y)$ has a distribution function symmetric about 0. Denote by A a Borel set of the real line and by $-A$ its mirror set obtained by reversing the sign of each element of A. The formula for the change of variables is

$$f_Y(y) = f_X(g^{-1}(y))|\frac{d}{dy}g^{-1}(y)| \tag{1.1}$$

for two random variables X and $Y = g(X)$, then, for $Z = -Y$ we have:

$$f_Y(t) = |-1| \cdot f_Z(-t) = f_Z(-t) \tag{1.2}$$

In addition, being f symmetric, we obtain

$$f_Y(t) = f_{-Y}(t). \tag{1.3}$$

The following equalities hold:

$$\mathbb{P}\{W \in -A\} \overset{(1)}{=} \mathbb{P}\{-W \in A\} \overset{(2)}{=} \mathbb{P}\{w(-Y) \in A\} \overset{(3)}{=} \mathbb{P}\{w(Y) \in A\} \tag{1.4}$$

Equality (1) is obtained by

$$\mathbb{P}\{W \in -A\} = \int_{-A} h_W(t)dt \; , \tag{1.5}$$

where $h_W(w)$ is the density of W and by

$$\begin{aligned} \mathbb{P}\{W \in -A\} &= \int_{-A} h_W(t)dt = -\int_A h_W(t)dt \\ &= -\int_A h_{-W}(-t)dt = \text{calling} s = -t \quad = -\int_A h_{-W}(s)(-ds) = \\ &= \int_A h_{-W}(s)ds = \mathbb{P}\{-W \in A\} \end{aligned} \tag{1.6}$$

where we used

$$h_W(t) = h_{-W}(-t). \tag{1.7}$$

As far as equality (2) is concerned, it holds

$$\mathbb{P}\{-W \in A\} = \mathbb{P}\{-w(Y) \in A\} = \mathbb{P}\{w(-Y) \in A\} \tag{1.8}$$

where we used $w(x)$ is odd.
Finally for the equality (3) we note that

$$\begin{aligned} \mathbb{P}\{w(-Y) \in A\} &= \int_A w(t)f_{-Y}(t)dt \\ &= \int_A w(t)f_Y(t)dt = \mathbb{P}\{w(Y) \in A\} \end{aligned} \tag{1.9}$$

As a result:

$$\mathbb{P}\{w(Y) \in -A\} = \mathbb{P}\{w(Y) \in A\} \tag{1.10}$$

that implies that $W = w(Y)$ has a distribution function symmetric about 0.
The second step in the proof consists in noting that the random variable $X - W$ has distribution function symmetric about 0:

$$\begin{aligned} &\mathbb{P}\{X - W < 0\} = \mathbb{P}\{X < W\} = \mathbb{P}\{X > -W\} \\ = \; &\mathbb{P}\{W > -X\} = \mathbb{P}\{W < X\} = \mathbb{P}\{X - W > 0\} \end{aligned} \tag{1.11}$$

Finally we have

$$\frac{1}{2} = \mathbb{P}\{X < W\} = \mathbb{E}\{\mathbb{P}\{X < w(Y)|Y = y\}\} = \int_{-\infty}^{\infty} f_Y(y)dy \int_{-\infty}^{w(y)} f_X(x)dx \quad (1.12)$$

and

$$\frac{1}{2} = \int_{-\infty}^{\infty} G\{w(y)\} f_Y(y)dy \quad (1.13)$$

which gives the desired result. □

Setting $f(x) = \varphi(x)$ and $G(x) = \Phi(x)$, the density and the distribution function of a standard normal r.v. respectively , and setting $w(x) = \alpha x$ with $\alpha \in \mathbb{R}$, we obtain the following density:

$$f(x) = 2\varphi(x)\Phi(\alpha x), \quad x \in \mathbb{R} \quad (1.14)$$

Definition 1.1.1. *A random variable Z having density (1.14) is called skew-normal (SN) with shape parameter α and denoted by*

$$Z \sim SN(\alpha).$$

If

$$Y = \mu + \sigma Z,$$

with $\mu, \sigma \in \mathbb{R}$ and $\sigma > 0$, then we write

$$Y \sim SN(\mu, \sigma^2, \alpha).$$

Its density is:

$$f_Y(y) = 2\varphi(\frac{y-\mu}{\sigma})\Phi(\alpha\frac{y-\mu}{\sigma}), \quad y \in \mathbb{R} \quad (1.15)$$

The following properties for (1.14) hold:
i) If $\alpha = 0$ then $Z \sim N(0, 1)$
ii) If $Z \sim SN(\alpha)$ then $-Z \sim SN(-\alpha)$
iii) As $\alpha \to \infty$ then (1.14) converges point-wise to the half normal density $2\varphi(z)$ for $z \geq 0$

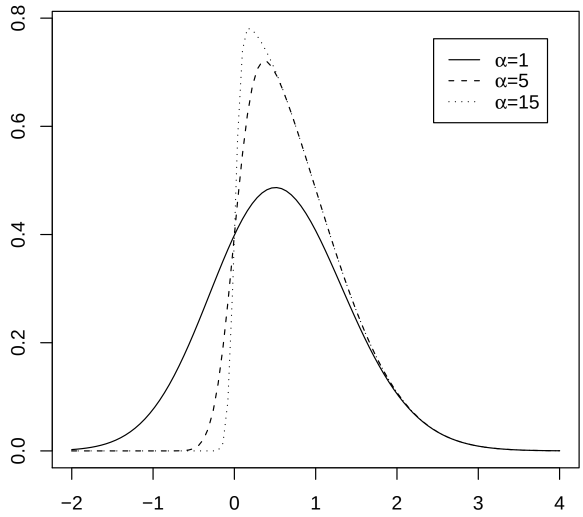


Figure 1.1: Density function $SN(\alpha)$ for three values of α

Proposition 1.1.1. *The moment generating function of $X \sim SN(\mu, \sigma^2, \alpha)$ is:*

$$M(t) = \mathbb{E}(e^{Xt}) = 2e^{\mu t + \sigma^2 t^2/2} \cdot \Phi(\delta\sigma t) \tag{1.16}$$

where

$$\delta = \frac{\alpha}{\sqrt{1+\alpha^2}} \in (-1, 1)$$

The previous Proposition is immediate given the following result:

Lemma 1.1.2. *If $U \sim N(0,1)$ and $a, b \in \mathbb{R}$ then*

$$\mathbb{E}(\Phi(a + bU)) = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right) \tag{1.17}$$

Proof. We have

$$\begin{aligned} \mathbb{E}(\Phi(a+bU)) &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{t+(a+bu)^2}{2}} dt\right] \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\ &= \int_{-\infty}^{0} \int_{-\infty}^{\infty} e^{\frac{-1}{2}[u^2(1+b^2)+2ub(a+t)+(a+t)^2]} du dt \end{aligned} \tag{1.18}$$

The following known result

$$\int_{-\infty}^{\infty} e^{-ax^2+bx}dx = \sqrt{\frac{\pi}{a}}e^{\frac{b^2}{4a}}$$

can be applied in order to obtain the desired equality

$$\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} \frac{b^2}{\sqrt{1+b^2}} e^{\frac{b^2(a+t)^2}{2(1+b^2)}} e^{\frac{(a+t)^2}{2}} dt = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1+b^2}} e^{-\frac{(a+t)^2}{2(1+b^2)}} dt$$

which gives the result. □

Proof of Proposition 1

We have for $X \sim SN(\mu, \sigma^2, \alpha)$ that

$$\begin{aligned}
& \int_{-\infty}^{\infty} 2\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}e^{tx}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left[\int_{-\infty}^{\frac{\alpha(x-\mu)}{\sigma}} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\right]dx \\
= & \int_{-\infty}^{\infty} 2\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}e^{-\frac{(x^2-2x(\mu+\sigma^2 t)+\mu^2}{2\sigma^2}} \left[\int_{-\infty}^{\frac{\alpha(x-\mu)}{\sigma}} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\right]dx \\
= & \int_{-\infty}^{\infty} 2\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}e^{-\frac{(x^2-2x(\mu+\sigma^2 t)+(\mu+\sigma^2 t)^2-2\mu\sigma^2 t-\sigma^4 t^2}{2\sigma^2}} \left[\int_{-\infty}^{\frac{\alpha(x-\mu)}{\sigma}} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\right]dx \\
= & \int_{-\infty}^{\infty} 2\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}e^{\mu t+\frac{\sigma^2t^2}{2}}e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} \left[\int_{-\infty}^{\frac{\alpha(x-\mu)}{\sigma}} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\right]dx. \qquad (1.19)
\end{aligned}$$

By the following change of variable

$$u = \frac{(x-(\mu+\sigma^2 t))}{\sigma}$$

we obtain

$$2e^{\mu t+\frac{\sigma^2t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}} \left[\int_{-\infty}^{\alpha\sigma t+\alpha u} \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt\right]du$$

and Lemma 1.1.2 can be applied to obtain the desired result:

$$M(t) = \mathbb{E}\{e^{Xt}\} = 2e^{\mu t+\sigma^2t^2/2}\cdot\Phi(\delta\sigma t) \qquad \square \qquad (1.20)$$

The cumulant generating function of $X \sim SN(\mu, \sigma^2, \alpha)$ is the following:

$$K(t) = \log M(t) = \mu t + \frac{\sigma^2 t^2}{2} + \log(2\Phi(\delta t\sigma))$$

Thus if $Z \sim SN(\alpha)$ we have:

$$K(t) = \frac{t^2}{2} + \log(2\Phi(\delta t))$$

and therefore

$$\mu_z := \mathbb{E}(Z) = \frac{d}{dt}K_z(t)|_{t=0} = (t + \frac{\Phi'(\delta t)}{\Phi(\delta t)}\delta)_{t=0} = \sqrt{\frac{2}{\pi}}\delta$$

and

$$\mathbb{E}(X) = \mu + \mu_z\sigma = \mu + \sqrt{\frac{2}{\pi}}(\sigma\delta)$$

The second and the third moment are given by:

$$\mathrm{Var}(X) = \frac{d^2}{dt^2}K(t)|_{t=0} = \sigma^2 - \frac{2}{\pi}(\delta\sigma)^2 = \sigma^2(1-\mu_z^2) \tag{1.21}$$

$$\mathrm{Skew}(X) = \frac{4}{(2\pi)^{3/2}}(4-\pi)(\sigma\delta)^3 \tag{1.22}$$

and furthermore

$$\gamma_1 = \frac{4-\pi}{2}\frac{\mu_z^3}{(1-\mu_z^2)^{3/2}}, \qquad \gamma_2 = 2(\pi-3)\frac{\mu_z^4}{(1-\mu_z^2)^2}$$

where γ_1, γ_2 denote the standardized third and fourth-order moments.

1.2 The multivariate skew normal distribution

The following Lemma can be easily proven , it is a simple generalization of Lemma 1.1.1 to the multivariate case.

Lemma 1.2.1. *If f_0 is a n-dimensional density function such that $f_0(\boldsymbol{x}) = f_0(-\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^n$, G is a one-dimensional differentiable distribution function such that G' is a density symmetric about 0, and w is a real valued function such that $w(-\boldsymbol{x}) = -w(-\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$, then*

$$f(\boldsymbol{z}) = 2f_0(\boldsymbol{z})G(w(\boldsymbol{z})) \tag{1.23}$$

is a density function on $\mathbb{R}^n$

Generalizing the univariate case, we set $f_0(\boldsymbol{x}) = \varphi_n(\boldsymbol{x}; 0, \Omega)$ where Ω is a positive definite matrix, $G(\boldsymbol{x}) = \Phi(\boldsymbol{x})$ and w a linear function. Then

$$f(\boldsymbol{x}) = 2\varphi_n(\boldsymbol{x}; 0, \Omega)\Phi(\boldsymbol{\alpha}^T\omega^{-1}\boldsymbol{x}) \tag{1.24}$$

is a multivariate density, where we denoted with ω the diagonal matrix of standard deviations of Ω. Allowing the presence of a location parameter $\boldsymbol{\mu}$ the density becomes:

$$f(\boldsymbol{x}) = 2\varphi_n(\boldsymbol{x}; \boldsymbol{\mu}, \Omega)\Phi(\boldsymbol{\alpha}^T\omega^{-1}(\boldsymbol{x}-\boldsymbol{\mu})) \tag{1.25}$$

Definition 1.2.1. *A n-dimensional random variable* $\boldsymbol{Z}$ *having density (1.25) is called multivariate skew-normal and denoted by:*

$$\boldsymbol{Z} \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha}).$$

Remark: The factor ω^{-1} in (1.25) is needed in order to keep the shape parameter unaltered when a location-scale transformation of the type $Y' = \boldsymbol{\xi} + BY$ is applied to Y, for some positive definite diagonal matrix B and location vector $\boldsymbol{\xi}$.

Proposition 1.2.1. *The moment generating function of a random variable distributed according to (1.25) is:*

$$M(\boldsymbol{t}) = 2exp(\boldsymbol{t}^T\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^T\Omega\boldsymbol{t})\Phi(\boldsymbol{t}^T(\omega\boldsymbol{\delta})) \tag{1.26}$$

where:

$$\boldsymbol{\delta} = \frac{\overline{\Omega}\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}^T\overline{\Omega}\boldsymbol{\alpha}}} \tag{1.27}$$

and $\overline{\Omega} = \omega^{-1}\Omega\omega^{-1}$

Proof. It is a simple extension of Proposition 1.1.1. □

The first two moments are obtained from $M(\boldsymbol{t})$:

$$\mathbb{E}(\boldsymbol{Z}) = \boldsymbol{\mu} + \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta}) \tag{1.28}$$

$$\text{Var}(\boldsymbol{Z}) = \Omega - \frac{2}{\pi}(\omega\boldsymbol{\delta})(\omega\boldsymbol{\delta})^T \tag{1.29}$$

and the multivariate index of skewness is:

$$\gamma_1 = \left(\frac{4-\pi}{2}\right)^2 \left(\frac{\frac{2}{\pi}(\omega\boldsymbol{\delta})^T\overline{\Omega}^{-1}(\omega\boldsymbol{\delta})}{1-\frac{2}{\pi}(\omega\boldsymbol{\delta})^T\overline{\Omega}^{-1}(\omega\boldsymbol{\delta})}\right)^3 \tag{1.30}$$

The skew-normal density can be "extended" to a more general form widely analyzed in literature (in particular see [9]). This extension is accomplished relaxing the condition on the normalization factor to be 1/2 and introducing a new parameter $\tau \in \mathbb{R}$.

Definition 1.2.2. *A n-dimensional random variable* $\boldsymbol{Z}$ *is distributed according to the "extended" skew normal distribution if its density is:*

$$f_{\boldsymbol{Z}}(\boldsymbol{x}) = \varphi(\boldsymbol{x}; \boldsymbol{\mu}, \Omega)\Phi(\alpha_0 + \boldsymbol{\alpha}^T\omega^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))/\Phi(\tau) \tag{1.31}$$

where $\alpha_0 = \tau(1+\boldsymbol{\alpha}^T\overline{\Omega}\boldsymbol{\alpha})^{1/2}$*. It is then denoted by*

$$\boldsymbol{Z} \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha}, \tau).$$

Remark: In the case $\tau = 0$ also $\alpha_0 = 0$ and (1.31) reduces to (1.25).
An important property of the class of distributions (1.31) is the closure under affine transformations. This property will be essential in the study of portfolios returns.

Proposition 1.2.2. *Given* $\boldsymbol{Z} \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha}, \tau)$, $\boldsymbol{\xi} \in \mathbb{R}^n$ *and* A *a* $(n \times d)$ *matrix then:*

$$\boldsymbol{\xi} + A\boldsymbol{Z} = Z_w \sim SN_d(\mu_w, \Omega_w, \alpha_w, \tau) \tag{1.32}$$

where:

$$\mu_w = \boldsymbol{\xi} + A\boldsymbol{\mu} \tag{1.33}$$
$$\Omega_w = A\Omega A^T \tag{1.34}$$
$$\alpha_w = \frac{\omega_w \Omega_w^{-1} H^T \boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}^T(\overline{\Omega} - H\Omega_w^{-1}H^T)\boldsymbol{\alpha}}} \tag{1.35}$$

where $H = \omega^{-1}\Omega A^T$ *and* ω_w *is the diagonal matrix of standard deviations of* Ω_w.

Proof. See [4]. □

Another important property satisfied by the family (1.31) is the closure under marginalization. If $\boldsymbol{Z} \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha}, \tau)$ is partitioned as follows:

$$\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \end{pmatrix} \;;\; \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \;;\; \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12} & \Omega_{22} \end{pmatrix} \;;\; \boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \alpha_2 \end{pmatrix}$$

then:

$$\boldsymbol{Z}_1 \sim SN_h(\boldsymbol{\mu}_1, \Omega_{11.}, \boldsymbol{\alpha}_{1(2)}, \tau) \tag{1.36}$$

where h is the dimension of $\boldsymbol{Z}_1$ and where:

$$\alpha_{1(2)} = \frac{\boldsymbol{\alpha}_1 + \overline{\Omega}_{11}^{-1}\overline{\Omega}_{12}\boldsymbol{\alpha}2}{\sqrt{1 + \boldsymbol{\alpha}_2^T \overline{\Omega}_{22.1}\boldsymbol{\alpha}_2}} \quad ; \quad \overline{\Omega}_{22.1} = \overline{\Omega}_{22} - \overline{\Omega}_{21}\overline{\Omega}_{11}^{-1}\overline{\Omega}_{12} \tag{1.37}$$

A detailed analysis of this property is presented in [4].

1.2.1 Bivariate skew normal

To better understand the properties of a skew normal distribution in this section we analyze the bivariate case. Consider the following covariances matrix:

$$\Omega = \begin{pmatrix} \omega_1^2 & \rho\omega_1\omega_2 \\ \rho\omega_1\omega_2 & \omega_2^2 \end{pmatrix} \tag{1.38}$$

where $|\rho| \leq 1$ and $\omega_i > 0$, and the shape parameter $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^T$ with $\alpha_i \in \mathbb{R}$.
The explicit expression of the bivariate skew normal density $\boldsymbol{Z} \sim SN_2(\boldsymbol{0}, \Omega, \boldsymbol{\alpha})$ is, from (1.25):

$$f_{\boldsymbol{Z}}(z_1, z_2) = \frac{2}{\pi\sqrt{\omega_1^2\omega_2^2(1-\rho^2)}} e^{-\frac{1}{2}\frac{\omega_1^2 z_1^2 - 2\rho z_1 z_2 \omega_1 \omega_2 + \omega_2^2 z_2^2}{(1-\rho^2)\omega_1^2\omega_2^2}} \Phi\left(\frac{\alpha_1}{\omega_1} z_1 + \frac{\alpha_2}{\omega_2} z_2\right) \tag{1.39}$$

From (1.36) the marginal distribution of Z_1 is still skew normal:

$$Z_1 \sim SN(0, \omega_1^2, \alpha_{1(2)})$$

where:

$$\alpha_{1(2)} = \frac{\alpha_1 + \rho\alpha_2}{\sqrt{1 + \alpha_2^2(1-\rho^2)}} \tag{1.40}$$

By (1.32) the random variable $Z_w = \boldsymbol{w}^T \boldsymbol{Z}$ with $\boldsymbol{w} = (w_1, w_2)^T$ and $\boldsymbol{Z} = (Z_1, Z_2) \sim SN_2(\boldsymbol{0}, \Omega, \boldsymbol{\alpha})$ is distributed according to $Z_w \sim SN(\mu_w, \omega_w^2, \alpha_w)$ where:

$$\begin{aligned} \mu_w &= \mu_1 w_1 + \mu_2 w_2 \\ \omega_w^2 &= w_1^2\omega_1^2 + 2\rho w_1 w_2 \omega_1 \omega_2 + w_2^2 \omega_2^2 \\ \alpha_w &= \frac{(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2\rho)\omega_1 w_1 + (\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_1\rho)\omega_2 w_2}{\sqrt{(1 + \boldsymbol{\alpha}_2\lambda\omega_1^2 w_1^2) + 2(\rho\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\lambda)\omega_1\omega_2 w_1 w_2 + (1 + \boldsymbol{\alpha}_1^2\lambda\omega_2^2 w_2^2)}} \end{aligned} \tag{1.41}$$

and where $\lambda = (1 - \rho^2)$.
Remark: As a consequence of (1.40) the distribution of Z_1 is not independent from Z_2 also in the case $\rho = 0$. To obtain the independence it is necessary the normality of Z_2, obtained setting $\alpha_2 = 0$.

1.3 Skew-t distribution

In this Section we present a further distribution generated by Lemma 1.2.1: the skew-t distribution. The main worth of this distribution is its ability to capture both the skewness and the thickness of the tails.
The expression of the density of a univariate t-distribution with n degree of freedom is the following:

$$t_X(x; \mu, \sigma, n) = \frac{\Gamma(\frac{n+1}{2})}{(\pi n)^{1/2}\Gamma(n/2)} \left(1 + \frac{(x-\mu)^2}{n\sigma}\right)^{-(n+1)/2}, \tag{1.42}$$

and its multivariate formulation is:

$$t_{\boldsymbol{X}}(\boldsymbol{x};\mu,\Sigma,n) = \frac{\Gamma(\frac{n+d}{2})}{(\pi n)^{d/2}\Gamma(n/2)}\left(1+\frac{(\boldsymbol{x}-\boldsymbol{m})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{m})}{n}\right)^{-(n+d)/2} \tag{1.43}$$

We denote by $T_1(x; n+d)$ the 1-dim t-cumulative distribution function with $n+d$ degrees of freedom. Being the t-density symmetric around the location parameter μ the assumptions of Lemma 1.2.1 are satisfied if one chooses: $f_0 = t_{\boldsymbol{Y}}$, $G(w) = T_1(w; n+d)$ and

$$w(\boldsymbol{y}) = \frac{\alpha^T\omega\boldsymbol{y}}{\boldsymbol{y}^T\Sigma^{-1}\boldsymbol{y}}.$$

In this case we obtain the following density:

$$f_{\boldsymbol{Y}} = t_{\boldsymbol{Y}}(\boldsymbol{y};\mu,\Sigma,n)\cdot T_1\left(\alpha^T\omega^{-1}(\boldsymbol{y}-\boldsymbol{m})\left(\frac{n+d}{Q_{\boldsymbol{y}}+n}\right)^{1/2}; n+d\right) \tag{1.44}$$

with $Q_{\boldsymbol{y}} = (\boldsymbol{y}-\boldsymbol{m})^T\Omega^{-1}(\boldsymbol{y}-\boldsymbol{m})$, for any definite positive matrix Σ, location vector $\mu \in \mathbb{R}^d$ and shape parameter $\boldsymbol{\alpha} \in \mathbb{R}^d$.

Definition 1.3.1. *A n-dimensional random variable $\boldsymbol{Z}$ having density (1.44) is called multivariate skew-t and we write $\boldsymbol{Z} \sim St_n(\boldsymbol{\mu},\Omega,\boldsymbol{\alpha},n)$.*

Remark: In the case $n \to \infty$ (1.44) becomes (1.25.)

We do not give here the expression of the moments of the distribution and its main properties. For a discussion of this distribution see Azzalini [3]. A modification of this distribution has been analyzed by Sahu *et. al.* in [34]. This new form turns out to be very useful in the regression models with skew-t errors. The main difference between the skew-t of Sahu and that one of Azzalini relies in the fact that Sahu uses the multivariate cumulative function of a t-Student as perturbation factor instead of the univariate one as in (1.44).

1.4 SUN

Several modifications of the original skew-normal density (1.31) have been developed in literature. Among these we mention the closed skew-normal (**CSN**) of Gonzalez-Farias *et al.* [15], the hierarchical skew-normal (**HSN**) of Liseo and Loperfido [23] and the unified skew normal (**SUN**) of Azzalini Arellano [2].

In this section we briefly recall the **SUN**.

Consider a $m+n$-normal variate $\boldsymbol{U} = (\boldsymbol{U}_1, \boldsymbol{U}_2)$:

$$\boldsymbol{U} \sim N_{m+n}(\mathbf{0}, \tilde{\Omega}) \;\; ; \;\; \tilde{\Omega} = \begin{pmatrix} \Gamma & \Delta^T \\ \Delta & \overline{\Omega} \end{pmatrix} \tag{1.45}$$

where $\tilde{\Omega}$ is a correlation matrix. Consider now the distribution of $\boldsymbol{Z} = (\boldsymbol{U}_2 | \boldsymbol{U}_1 + \boldsymbol{\gamma} > 0)$ for some $\boldsymbol{\gamma} \in \mathbb{R}^n$. The density function of $\boldsymbol{Z}$ is computed by the formula:

$$f_{\boldsymbol{Z}}(\boldsymbol{y}) = \frac{f_{\boldsymbol{U}_2}(\boldsymbol{y}) \mathbb{P}\{\boldsymbol{U}_1 > -\gamma | \boldsymbol{U}_2 = y\}}{\mathbb{P}\{\boldsymbol{U}_1 > c\}}.$$

After simple algebra one obtains the density of $\boldsymbol{Y} = \boldsymbol{\mu} + \omega \boldsymbol{Z} \in \mathbb{R}^n$:

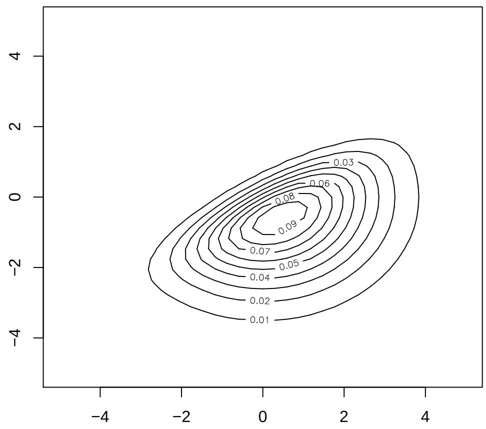
$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \varphi_n(\boldsymbol{y}; \boldsymbol{\mu}, \Omega) \frac{\Phi_m(\boldsymbol{\gamma} + \Delta^T \overline{\Omega}^{-1} \omega^{-1} (\boldsymbol{y} - \boldsymbol{\mu}); \Gamma - \Delta^T \overline{\Omega}^{-1} \Delta)}{\Phi_m(\boldsymbol{\gamma}; \Gamma)} \tag{1.46}$$

Consider the vector of standard deviations $\overline{\omega} = \omega \mathbf{1}$ then:

Definition 1.4.1. *A random variable $\boldsymbol{Y}$ with density given by (1.46) is called "unified" skew normal and is denoted by $\boldsymbol{Y} \sim SUN_{n,m}(\boldsymbol{\mu}, \boldsymbol{\gamma}, \overline{\omega}, \tilde{\Omega})$*

Remark: In the case $m = 1$, the density given by (1.46) collapses to the density of a multivariate skew-normal (1.25).

All the properties of a $\text{SUN}_{n,m}$ distribution are proved in the Appendices of Azzalini-Arellano [2]. In Appendix A.1 we apply these results in order to obtain the two different types of bivariate $\text{SUN}_{n,m}$, resulting from the following choice of parameters: $[n = 2, m = 1]$ and $[n = 2, m = 2]$.



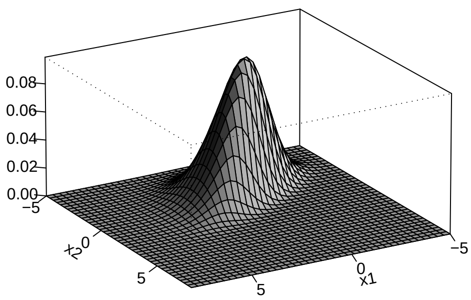


Figure 1.2: Contour plot and 3-d plot of a bivariate $SN_2(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$ with $\boldsymbol{\mu} = (0, 0)$, $\Omega = \text{diag}(3, 2.5)$ and $\boldsymbol{\alpha} = (2, -3)$

2

Stochastic Dominance for a skew-normal random variable

Assume an investor selects a portfolio $\boldsymbol{w}$ of risky assets. If $\boldsymbol{R}$ is the vector of assets returns we denote by

$$R_{\boldsymbol{w}} = w^T \boldsymbol{R}$$

the corresponding univariate random variable representing the portfolio return.
Given two portfolio returns $R_{\boldsymbol{w}_1}$ and $R_{\boldsymbol{w}_2}$ an investor is faced with the problem:

Among the two portfolios $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$, which one should be preferred?

If $R_{\boldsymbol{w}_2}(\cdot) \geq R_{\boldsymbol{w}_1}(\cdot)$ was true for each scenario then the choice would be obvious. Nonetheless portfolio returns usually do not satisfy the previous simple dominance relation. Indeed $R_{\boldsymbol{w}_1}$ and $R_{\boldsymbol{w}_2}$ may have intersecting probability densities on ample regions of the returns space.
The Stochastic Dominance Theory (**SD**), developed mainly by Levy in [21], [22] and by Levy and Hanoch in [16], represents an attractive method to solve this problem. This theory aims to find criterions to rank univariate r.v., which in this context are often called risky prospects, depending on the underlying utility functions.
Being the portfolios returns univariate r.v. they can be ranked using **SD**. In this way the set of all portfolios (the feasible set) breaks down into two sets: an efficient set and an inefficient set of portfolios, where a portfolio is efficient if, and only if, its return is not stochastically dominated by another. Hence **SD** represents the base of the Portfolio Selection Theory, developed in this thesis in Chapters 3 and 4 for normal and

skew-normal returns respectively.
In this Chapter after having briefly recalled the principal aspects of **SD** we analyze the case of normal and skew-normal prospects.

2.1 First and Second order Stochastic Dominance

We denote by $\mathcal{U}_i$ the following sets of functions:
$\mathcal{U}_1 = \{u \in \mathcal{C}^1(\mathbb{R}) \text{ with } u'(x) \geq 0\}$
$\mathcal{U}_2 = \{u \in \mathcal{C}^2(\mathbb{R}) \text{ with } u'(x) \geq 0, u''(x) \leq 0\}$

Definition 2.1.1. *Consider two random variables X_1 and X_2, we say that X_1 stochastically dominates at* first *order X_2 , and we write $X_1 \succeq_1 X_2$, if:*

$$\mathbb{E}(u(X_1)) - \mathbb{E}(u(X_2)) \geq 0$$

for every $u \in \mathcal{U}_1$.
*(whenever the inequality holds strictly for at least one u, then we say that **SD** holds in strong sense).*

Lemma 2.1.1. *Consider two random variables X_1 and X_2 which have $F_1(x)$ and $F_2(x)$ as their cumulative distributions functions. Then if $u \in \mathcal{U}_1$:*

$$\mathbb{E}(u(X_1)) - \mathbb{E}(u(X_2)) = \int_{-\infty}^{\infty} (F_2(x) - F_1(x))d(u(x))$$

Proof. In [16] pag. 336. □

Theorem 2.1.1. *Under the hypothesis of Lemma 2.1.1, then:*

$$X_1 \succeq_1 X_2 \quad \Leftrightarrow \quad F_1(x) \leq F_2(x) \; \forall x$$

Proof. In [16] pag. 337. □

This criterion is simply interpretable: the probability to take a value smaller than x for X_1 is not larger than the same probability for X_2.

Definition 2.1.2. *Consider two random variables X_1 and X_2, we say that X_1 stochastically dominates at* second *order X_2 , and we write $X_1 \succeq_2 X_2$, if:*

$$\mathbb{E}(u(X_1)) - \mathbb{E}(u(X_2)) \geq 0$$

for every $u \in \mathcal{U}_2$.
(whenever the inequality holds strictly for at least one u*, then we say that* ***SD*** *holds in strong sense).*

Remark 1: We have:

$$X_1 \succeq_1 X_2 \Rightarrow X_1 \succeq_2 X_2$$

Remark 2: If $X_1 \succeq_i Y$ and $Y \succeq_i X_2$ then $X_1 \succeq_i X_2$, for $i = 1, 2$; this is obvious considering that for any $u \in \mathcal{U}_i$:

$$\mathbb{E}(u(X_1)) - \mathbb{E}(u(X_2)) = \mathbb{E}(u(X_1)) - \mathbb{E}(u(Y)) + \mathbb{E}(u(Y)) - \mathbb{E}(u(X_2)) \geq 0$$

Theorem 2.1.2. *Consider two random variables* X_1 *and* X_2 *which have* $F_1(x)$ *and* $F_2(x)$ *as their cumulative distributions functions. Then if* $u \in \mathcal{U}_2$*:*

$$X_1 \succeq_2 X_2 \qquad \Leftrightarrow \qquad \int_{-\infty}^{x} (F_2(t) - F_1(t))dt \geq 0 \; \forall x \tag{2.1}$$

Proof. In [16] pag. 338. □

Given two functions g_1, g_2 defined on $\mathbb{R}$. We say that g_1 intersects only once from *below* g_2 if $g_1 < g_2$ to the left of the intersection point.

Lemma 2.1.2. *Consider two random variables* X_1 *and* X_2 *which have* $F_1(x)$ *and* $F_2(x)$ *as their cumulative distributions functions. If* $F_1(x)$ *intersects only once from below* $F_2(x)$ *in* x_0 *, then :*

$$\mathbb{E}(X_1) - \mathbb{E}(X_2) \geq 0 \iff \int_{-\infty}^{x} (F_2(t) - F_1(t))dt \geq 0 \quad \forall x.$$

Proof. ($\Rightarrow$): From Lemma 2.1.1:

$$\mathbb{E}(X_1) - \mathbb{E}(X_2) = \int_{-\infty}^{\infty} (F_2(x) - F_1(x))dx \tag{2.2}$$

If $x \leq x_0$ then

$$\int_{-\infty}^{x} (F_2(t) - F_1(t))dt > 0$$

If $x > x_0$:

$$\int_{-\infty}^{x} (F_2(t) - F_1(t))dt \geq \int_{x}^{\infty} (F_1(t) - F_2(t))dt > 0$$

($\Leftarrow$): For $x \to \infty$:

$$\int_{-\infty}^{\infty} (F_2(t) - F_1(t))dt \geq 0 \Rightarrow \mathbb{E}(X_1) - \mathbb{E}(X_2) \geq 0$$

□

In the next Theorem we give sufficient conditions for second order stochastic dominance:

Theorem 2.1.3. *Consider two random variables X_1 and X_2 which have $F_1(x)$ and $F_2(x)$ as their cumulative distributions functions. If $F_1(x)$ intersects only once from below $F_2(x)$ in x_0 then:*

$$\mathbb{E}(X_1) - \mathbb{E}(X_2) \geq 0 \iff X_1 \succeq_2 X_2$$

Proof. : Obvious from Theorem 2.1.2 and Lemma 2.1.2. □

Normal random variables

Proposition 2.1.1. *Consider two normal random variables $X_1 \sim N(\mu_1, \sigma)$ and $X_2 \sim N(\mu_2, \sigma)$. Suppose $\mu_1 \geq \mu_2$, then $X_1 \succeq_1 X_2$.*

Proof. If $\mu_1 = \mu_2$ then $X_1 \equiv X_2$. If $\mu_1 > \mu_2$ the result it's immediate considering that $F_{X_1}(x)$ and $F_{X_2}(x)$ have no intersection points and $F_{X_1}(x) < F_{X_2}(x)$ for any x and then applying Theorem 2.1.1. □

Proposition 2.1.2. *Consider two normal random variables $X_1 \sim N(\mu_1, \sigma_1)$ and $X_2 \sim N(\mu_2, \sigma_2)$. Suppose $\mu_1 \geq \mu_2$ and $\sigma_1 \leq \sigma_2$, then $X_1 \succeq_2 X_2$.*

The proof is based on the following Lemma and on Theorem 2.1.3.

Lemma 2.1.3. *Consider two normal random normal variables $X_1 \sim N(\mu_1, \sigma_1)$ and $X_2 \sim N(\mu_2, \sigma_2)$ which have $F_1(x)$ and $F_2(x)$ as their cumulative distributions functions. Then $F_1(x)$ intersects only once from below $F_2(x)$ if and only if $\sigma_1 < \sigma_2$*

Proof. The cumulative distribution function of $Z \sim N(\mu, \sigma)$ can be written as:

$$F_Z(x; \mu, \sigma) = \int_{-\infty}^{x} \frac{1}{\sigma} \varphi(\frac{y - \mu}{\sigma}) dy = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \varphi(t) dt = \Phi(\frac{x - \mu}{\sigma}) \tag{2.3}$$

The intersection points of $F_1(x)$ and $F_2(x)$ are obtained by:

$$\Phi(\frac{x - \mu_1}{\sigma_1}) = \Phi(\frac{x - \mu_2}{\sigma_2}) \tag{2.4}$$

assuming $\sigma_1 \neq \sigma_2$ the previous expression implies the existence of a unique intersection point in $x^* = (\mu_1 \sigma_2 - \mu_2 \sigma_1)/(\sigma_1 - \sigma_2)$. Furthermore it holds by De l'Hopital:

$$\lim_{x \to -\infty} \frac{F_1(x)}{F_2(x)} = \lim_{x \to -\infty} \text{Exp}[-\frac{(x - \mu_1)^2}{2\sigma_1^2} + \frac{(x - \mu_2)^2}{2\sigma_2^2}] = \lim_{x \to -\infty} \text{Exp}[-\frac{x^2}{2}(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2})]$$

if $\sigma_1 < \sigma_2$ ($\sigma_1 > \sigma_2$) then the above limit is 0 (∞). This ends the proof. □

Proof. (*of Proposition 2.1.2*) In the case $\sigma_1 = \sigma_2$:

$$\mu_1 \geq \mu_2 \Rightarrow X_1 \succeq_1 X_2 \Rightarrow X_1 \succeq_2 X_2 \tag{2.5}$$

where the first row is implied by Proposition 2.1.1.
The previous Lemma implies that, if $\sigma_1 < \sigma_2$, then F_1 intersects only once from below F_2. In addition, considering the condition $\mu_1 \geq \mu_2$ and the Theorem 2.1.3 we obtain the result. $\square$

Skew-Normal random variables

Theorem 2.1.4. : *(i) Let $X_1 \sim SN(\mu, \sigma^2, \alpha_1)$ and $X_2 \sim SN(\mu, \sigma^2, \alpha_2)$ be skew-normal r.v's. Suppose $\alpha_1 \geq \alpha_2$ then $X_1 \succeq_1 X_2$.*

(ii) Let $X_1 \sim SN(\mu, \sigma_1^2, \alpha)$ and $X_2 \sim SN(\mu, \sigma_2^2, \alpha)$ be skew-normal r.v's with $\alpha \leq 0$. Suppose $\sigma_1 \leq \sigma_2$ then $X_1 \succeq_2 X_2$.

(iii) Let $X_1 \sim SN(\mu_1, \sigma^2, \alpha)$ and $X_2 \sim SN(\mu_2, \sigma^2, \alpha)$ be skew-normal r.v's. Suppose $\mu_1 \geq \mu_2$ then $X_1 \succeq_1 X_2$.

Proof:

(i): Let $X \sim SN(\mu, \sigma^2, \alpha)$ and denote by $F_{\mu,\sigma,\alpha}(x)$ the corresponding distribution function. For each fixed (μ, σ) and arbitrary x we consider the function $\alpha \to h(\alpha) := F_{\mu,\sigma,\alpha}(x)$. We have:

$$h'(\alpha) = \frac{2}{\sigma} \int_{-\infty}^{x} \varphi(\frac{y-\mu}{\sigma}) \frac{\partial}{\partial \alpha} \Phi(\alpha \frac{y-\mu}{\sigma}) dy = 2 \int_{-\infty}^{x} \frac{y-\mu}{\sigma} \varphi(\frac{y-\mu}{\sigma}) \varphi(\alpha \frac{y-\mu}{\sigma}) d(\frac{y-\mu}{\sigma})$$

$$= 2 \int_{-\infty}^{\frac{x-\mu}{\sigma}} t\varphi(t\sqrt{1+\alpha^2}) dt = \frac{2}{1+\alpha^2} \int_{-\infty}^{\frac{\sqrt{1+\alpha^2}}{\sigma}(x-\mu)} t\varphi(t) dt = -\frac{2}{1+\alpha^2} \varphi(\frac{\sqrt{1+\alpha^2}}{\sigma}(x-\mu))$$

Therefore $h(\alpha)$ is decreasing, $F_{\mu,\sigma,\alpha_1}(x) \leq F_{\mu,\sigma,\alpha_2}(x)$ for all x, and the result follows by Theorem 2.1.1.

(ii): For each fixed (μ, α) and arbitrary z we consider the function $\sigma \to l(\sigma) := \int_{-\infty}^{z} F_{\mu,\sigma,\alpha}(x) dx$. We have:

$$l'(\sigma) = \int_{-\infty}^{z} \frac{\partial}{\partial \sigma} [\int_{-\infty}^{x} \frac{2}{\sigma} \varphi(\frac{y-\mu}{\sigma}) \Phi(\alpha \frac{y-\mu}{\sigma}) dy] dx = 2 \int_{-\infty}^{z} \frac{\partial}{\partial \sigma} [\int_{-\infty}^{\frac{x-\mu}{\sigma}} \varphi(t) \Phi(\alpha t) dt] dx$$

$$= -2 \int_{-\infty}^{z} \frac{x-\mu}{\sigma^2} \varphi(\frac{x-\mu}{\sigma}) \Phi(\alpha \frac{x-\mu}{\sigma}) dx = -2 \int_{-\infty}^{\frac{z-\mu}{\sigma}} t\varphi(t) \Phi(\alpha t) dt$$

and by integrating by parts the last integral

$$l'(\sigma) = 2\varphi(\frac{z-\mu}{\sigma})\Phi(\alpha\frac{z-\mu}{\sigma}) - \frac{2\alpha}{\sqrt{1+\alpha^2}}\Phi(\frac{\sqrt{1+\alpha^2}}{\sigma}(z-\mu))$$

which is non negative for all $\alpha \leq 0$. Henceforth $l(\sigma)$ is increasing, $\int_{-\infty}^{z} F_{\mu,\sigma_1,\alpha}(x)dx \leq \int_{-\infty}^{z} F_{\mu,\sigma_2,\alpha}(x)dx$ for all z, and we get the result by Theorem 2.1.2.

(iii): For each fixed (σ, α) and arbitrary x we consider the function $\mu \to k(\mu) := F_{\mu,\sigma,\alpha}(x)$. We have:

$$k'(\mu) = \int_{-\infty}^{x} \frac{2}{\sigma}\frac{\partial}{\partial\mu}[\varphi(\frac{y-\mu}{\sigma})\Phi(\alpha\frac{y-\mu}{\sigma})]dy$$

$$= \frac{2}{\sigma}\int_{-\infty}^{x}(\frac{y-\mu}{\sigma})\varphi(\frac{y-\mu}{\sigma})\Phi(\alpha\frac{y-\mu}{\sigma})d(\frac{y-\mu}{\sigma}) - \frac{2\alpha}{\sigma}\int_{-\infty}^{x}\varphi(\frac{y-\mu}{\sigma})\varphi(\alpha\frac{y-\mu}{\sigma})d(\frac{y-\mu}{\sigma})$$

$$= -\frac{2}{\sigma}\int_{-\infty}^{\frac{x-\mu}{\sigma}}\varphi'(t)\Phi(\alpha t)dt - \frac{2\alpha}{\sigma\sqrt{1+\alpha^2}}\int_{-\infty}^{\frac{x-\mu}{\sigma}}\varphi(t)dt = -\frac{2}{\sigma}\varphi(\frac{x-\mu}{\sigma})\Phi(\alpha\frac{x-\mu}{\sigma})$$

where we have integrated by parts the first integral.

Therefore $k(\mu)$ is decreasing and $F_{\mu_1,\sigma,\alpha}(x) \geq F_{\mu_2,\sigma,\alpha}(x)$ for all x, the result follows by Theorem 2.1.1.□

Corollary 2.1.5. :

(i)Let $X_1 \sim SN(\mu_1, \sigma^2, \alpha_1)$ and $X_2 \sim SN(\mu_2, \sigma^2, \alpha_2)$ be skew-normal r.v's. Suppose $\mu_1 \geq \mu_2$ and $\alpha_1 \geq \alpha_2$, then $X_1 \succeq_1 X_2$.

(ii)Let $X_1 \sim SN(\mu_1, \sigma_1^2, \alpha_1)$ and $X_2 \sim SN(\mu_2, \sigma_2^2, \alpha_2)$ be skew-normal r.v's with $\alpha_1, \alpha_2 \leq 0$. Suppose $\mu_1 \geq \mu_2$, $\sigma_1 \leq \sigma_2$ and $\alpha_1 \geq \alpha_2$, then $X_1 \succeq_2 X_2$.

Proof. (i)Let $Y \sim SN(\mu_1, \sigma, \alpha_2)$ then:

$$X_1 \succeq_1 Y \succeq_1 X_2.$$

(ii) Let $Y \sim SN(\mu_1, \sigma_1^2, \alpha_2)$ and $Z \sim SN(\mu_1, \sigma_2^2, \alpha_2)$ then, by Theorem 2.1.4,

$$X_1 \succeq_2 Y \succeq_2 Z \succeq_2 X_2.$$

□

Given a skew normal random variable $X \sim SN(\mu, \sigma^2, \alpha)$ we shall call:

$$B^2 := (\sigma\delta)^2 : \text{ the } \textit{non-spherical} \text{ component of the variance: } V_{\text{ns}} \quad (2.6)$$

$$s^2 := \sigma^2 - (\sigma\delta)^2 : \text{ the } \textit{spherical} \text{ component of the variance: } V_{\text{S}} \quad (2.7)$$

where we recall the expression of δ :

$$\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}.$$

The expression of $\mathrm{Var}(X)$ in terms of (σ, B), reported below, justifies this terminology (adapted from Simaan [35]).

In terms of $B = B(\sigma, \alpha)$ the density of $X \sim SN(\mu, \sigma^2, \alpha)$ becomes:

$$f_X(x) = \frac{2}{\sigma}\varphi(\frac{x-\mu}{\sigma})\Phi(\frac{B}{\sqrt{\sigma^2 - B^2}}\frac{(x-\mu)}{\sigma}) \tag{2.8}$$

and the first three moments are:

$$\mathbb{E}(X) = \mu + \sqrt{\frac{2}{\pi}}B \tag{2.9}$$

$$\mathrm{Var}(X) = \sigma^2 - \frac{2}{\pi}B^2 = [\sigma^2 - B^2] + [(1-\frac{2}{\pi})B^2] = V_{\mathrm{S}} + (1-\frac{2}{\pi})V_{\mathrm{ns}} \tag{2.10}$$

$$\mathrm{Skew}(X) = \frac{4}{(2\pi)^{3/2}}(4-\pi)B^3 \tag{2.11}$$

From the last relation we see that for $B = 0$ the skewness of X is zero and viceversa. This justifies the name "*non-spherical component*" for the part of the variance proportional to B^2. The following second order stochastic dominance result will be used later in Section 4.3.1.

Proposition 2.1.6. *Let $X_1 \sim SN(\mu, \sigma_1^2, \alpha_1)$ and $X_2 \sim SN(\mu, \sigma_2^2, \alpha_2)$ be skew-normal r.v's. Suppose $\sigma_1 \leq \sigma_2$ and $B_1 = \sigma_1\delta_1 = \sigma_2\delta_2 = B_2$ then $X_1 \succeq_2 X_2$.*

Proof. Let $X \sim SN(\mu, \sigma^2, \alpha)$ in terms of (μ, σ^2, B) its cumulative distribution function is:

$$\begin{aligned} F_{\mu,\sigma,B}(x) &= \frac{2}{\sigma}\int_{-\infty}^{x}\varphi(\frac{z-\mu}{\sigma})\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(z-\mu)}{\sigma})dz \\ &= \int_{-\infty}^{\frac{x-\mu}{\sigma}} 2\varphi(t)\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}t)dt \end{aligned}$$

For each fixed (μ, B) and arbitrary y we consider the function

$$\sigma \to m(\sigma) := \int_{-\infty}^{y} F_{\mu,\sigma,B}(x)dx.$$

We have:

$$\begin{aligned} m'(\sigma) &= -2\int_{-\infty}^{y}\frac{x-\mu}{\sigma^2}\varphi(\frac{x-\mu}{\sigma})\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(x-\mu)}{\sigma})dx \\ &+ 2\int_{-\infty}^{y}\left(\int_{-\infty}^{\frac{x-\mu}{\sigma}}\varphi(t)\varphi(\frac{B}{\sqrt{\sigma^2-B^2}}t)\frac{-Bt\sigma}{(\sigma^2-B^2)^{3/2}}dt\right)dx \end{aligned}$$

Denote by:

$$A = -2\int_{-\infty}^{y} \frac{x-\mu}{\sigma^2}\varphi(\frac{x-\mu}{\sigma})\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(x-\mu)}{\sigma})dx$$

then:

$$m'(\sigma) = A + 2\int_{-\infty}^{y}\left(\int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{-Bt\sigma}{\pi^{1/2}(\sigma^2-B^2)^{3/2}}\varphi(\frac{\sigma t}{\sqrt{\sigma^2-B^2}})dt\right)dx$$

Integrating the gaussian part we obtain:

$$m'(\sigma) = A + 2\int_{-\infty}^{y} \frac{B}{\sigma(\pi(\sigma^2-B^2))^{1/2}}\varphi(\frac{x-\mu}{\sqrt{\sigma^2-B^2}})dx$$

Notice that A can be written as:

$$A = 2\int_{-\infty}^{y}(\partial_x\varphi(\frac{x-\mu}{\sigma}))\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(x-\mu)}{\sigma})dx$$

therefore integrating by parts we get:

$$\begin{aligned} A &= 2\left[\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(x-\mu)}{\sigma})\varphi(\frac{x-\mu}{\sigma}))\right]_{-\infty}^{y} && (2.12)\\ &- 2\int_{-\infty}^{y}\varphi(\frac{x-\mu}{\sigma})\varphi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(x-\mu)}{\sigma})\frac{B}{\sigma((\sigma^2-B^2))^{1/2}}dx \\ &= 2\left[\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(x-\mu)}{\sigma})\varphi(\frac{x-\mu}{\sigma}))\right]_{-\infty}^{y} && (2.13)\\ &- 2\int_{-\infty}^{y}\frac{B}{\sigma(\pi(\sigma^2-B^2))^{1/2}}\varphi(\frac{(x-\mu)}{\sqrt{\sigma^2-B^2}})dx \end{aligned}$$

Inserting this result into the expression of $m'(\sigma)$ we obtain:

$$m'(\sigma) = 2\Phi(\frac{B}{\sqrt{\sigma^2-B^2}}\frac{(y-\mu)}{\sigma})\varphi(\frac{y-\mu}{\sigma}))$$

Hence $\int_{-\infty}^{y} F_{\mu,\sigma,B}(x)dx$ is increasing in σ and for fixed (μ, B):

$$\int_{-\infty}^{y} F_{\mu,\sigma_1,B}(x)dx \leq \int_{-\infty}^{y} F_{\mu,\sigma_2,B}(x)dx \qquad (2.14)$$

for any y. The result follows by Theorem 2.1.2. □

Corollary 2.1.1. *Let $X_1 \sim SN(\mu_1, \sigma_1^2, \alpha_1)$ and $X_2 \sim SN(\mu_2, \sigma_2^2, \alpha_2)$ be skew-normal r.v's. Suppose $\mu_1 \geq \mu_2$, $\sigma_1 \leq \sigma_2$ and $B_1 = \sigma_1\delta_1 = \sigma_2\delta_2 = B_2$ then $X_1 \succeq_2 X_2$.*

Proof. Let $Y \sim SN(\mu_1, \sigma_2^2, \alpha_2)$. We have $X_1 \succeq_2 Y$ and $Y \succeq_1 X_2$. Hence:

$$X_1 \succeq_2 Y \succeq_2 X_2.$$

□

3

Mean Variance Analysis and CAPM

3.1 The mean variance framework

Consider a market of n risky assets; we denote by $\boldsymbol{R} \in \mathbb{R}^n$ the random vector of assets returns and by

$$R_{\boldsymbol{w}} = \boldsymbol{w}^T \boldsymbol{R}$$

the portfolio $\boldsymbol{w} \in \mathbb{R}^n$ representing the univariate r.v. of portfolio return.
The Mean Variance analysis is based on the following assumptions:
i) the preferences of all investors about portfolios are based on an expected utility function depending only on two parameters: the mean and the variance of the portfolio returns
ii) all investors prefer high portfolio means and small portfolio variances.

Denoting by $f_{R_{\boldsymbol{w}}}(\cdot)$ the density of $R_{\boldsymbol{w}}$, the condition i) is expressed by the following formula:

$$\mathbb{E}(u(R_{\boldsymbol{w}})) = \int u(r) f_{R_{\boldsymbol{w}}}(r) dr \equiv \lambda(\mathbb{E}(R_{\boldsymbol{w}}), \mathrm{Var}(R_{\boldsymbol{w}})), \tag{3.1}$$

that is the investor expected utility depends only on the first two moments of the portfolio returns distribution.
In addition, to assure the preference towards high means and small variances, the function λ has to be increasing in means and decreasing in variances. Therefore denoting

by $m_{\boldsymbol{w}} = \mathbb{E}(R_{\boldsymbol{w}})$ and by $v_{\boldsymbol{w}}^2 = \mathrm{Var}(R_{\boldsymbol{w}})$, the condition ii) is expressed by:

$$\frac{\partial}{\partial m_{\boldsymbol{w}}}\lambda(m_{\boldsymbol{w}}, v_{\boldsymbol{w}}^2) > 0 \;\; ; \;\; \frac{\partial}{\partial v_{\boldsymbol{w}}^2}\lambda(m_{\boldsymbol{w}}, v_{\boldsymbol{w}}^2) < 0 \tag{3.2}$$

Hence, in this framework, given two portfolios $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ and denoting by $R_{\boldsymbol{w}_1}$ and $R_{\boldsymbol{w}_2}$ their portfolio returns, the condition

$$\mathbb{E}(u(R_{\boldsymbol{w}_1})) \geq \mathbb{E}(u(R_{\boldsymbol{w}_2})), \tag{3.3}$$

is satisfied if the following inequalities

$$\mathbb{E}(R_{\boldsymbol{w}_1}) \geq \mathbb{E}(R_{\boldsymbol{w}_2}) \quad \text{and} \quad \mathrm{Var}(R_{\boldsymbol{w}_1}) \leq \mathrm{Var}(R_{\boldsymbol{w}_2}). \tag{3.4}$$

hold. Note that the conditions (3.4) are sufficient but not necessary for (3.3). Indeed mean variance analysis does not provide any advices for the cases:
1) $\mathbb{E}(R_{\boldsymbol{w}_1}) > \mathbb{E}(R_{\boldsymbol{w}_2})$ and $\mathrm{Var}(R_{\boldsymbol{w}_1}) > \mathrm{Var}(R_{\boldsymbol{w}_2})$
2) $\mathbb{E}(R_{\boldsymbol{w}_1}) < \mathbb{E}(R_{\boldsymbol{w}_2})$ and $\mathrm{Var}(R_{\boldsymbol{w}_1}) < \mathrm{Var}(R_{\boldsymbol{w}_2})$,
In these cases the investor has to choose between $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ using other rules. Notice that no explicit assumption is made on the distribution of returns $\boldsymbol{R}$ which drives the portfolio return $R_{\boldsymbol{w}}$. However when the utility function is quadratic or when $\boldsymbol{R}$ is elliptically distributed the requirement implicitly made in (3.1) holds true. This will be shown later on.

3.2 Portfolio selection

Definition 3.2.1. *The mean-variance* efficient set *is the set of all portfolios with smallest variance for a given level of mean and greatest mean for a given level of variance.*

The *efficient set* is a subset of the *minimum-variance* set. This is the set of portfolios of minimum variance for each level of mean. They are compared in Figure 3.2. In the next two subsections we analyze the fundamental property of the *minimum variance* set: portfolios in this set can all be obtained as linear combination of only two distinct minimum variance portfolios (called the *separating funds*).

Separating funds without risk-less asset

In this section we analyze the nature of the *minimum variance* set and we find the two *separating funds* in absence of a risk-less asset. Let $\boldsymbol{R} \in \mathbb{R}^n$ be the random vector of assets returns (all assets are assumed to be risky). Denote by

$$\boldsymbol{m} = \mathbb{E}(\boldsymbol{R}) \quad \text{and} \quad V = \text{Var}(\boldsymbol{R})$$

For each portfolio $\boldsymbol{w}$, we have:

$$m_{\boldsymbol{w}} = \mathbb{E}(R_{\boldsymbol{w}}) = \boldsymbol{w}^T\boldsymbol{m} \tag{3.5}$$

$$v_w^2 = \text{Var}(R_{\boldsymbol{w}}) = \boldsymbol{w}^T V \boldsymbol{w} \tag{3.6}$$

with $R_{\boldsymbol{w}} = \boldsymbol{w}^T\boldsymbol{R}$.

Consider a fixed level of poertfolio returns $E \in \mathbb{R}$ and the following associated optimization problem:

$$\begin{aligned} \text{Min}_w \quad & \tfrac{1}{2}\boldsymbol{w}^T V \boldsymbol{w} \\ \text{with the constraints:} \quad & \boldsymbol{w}^T\boldsymbol{m} = E \\ & \boldsymbol{1}^T\boldsymbol{w} = 1 \end{aligned} \tag{3.7}$$

Short sales are permitted: no constraints of the kind $w_i \geq 0$ have been imposed on the components of $\boldsymbol{w}$.

Each portfolio which solves this problem is said to belong to the *minimum variance* set. Varying the values of $E \in \mathbb{R}$ we span the *minimum variance* set. To obtain a portfolio lying in the mean-variance *efficient set* an investor selects among portfolios in the *minimum variance* set which have equal variance the portfolio with the highest value of $E \in \mathbb{R}$.

Proposition 3.2.1. *It holds the following: the* minimum variance *set is spanned by the two portfolios (called* separating funds*):*

$$\boldsymbol{a}_1 = \frac{V^{-1}\boldsymbol{m}}{\boldsymbol{1}^T V^{-1}\boldsymbol{m}}, \quad \boldsymbol{a}_2 = \frac{V^{-1}\boldsymbol{1}}{\boldsymbol{1}^T V^{-1}\boldsymbol{1}} \tag{3.8}$$

Proof. Starting from the problem (3.7) we can write the following Lagrangian:

$$L = \frac{1}{2}\boldsymbol{w}^T V \boldsymbol{w} - \lambda_1(\boldsymbol{1}^T\boldsymbol{w} - 1) - \lambda_2(\boldsymbol{w}^T\boldsymbol{m} - E) \tag{3.9}$$

and derive the first order conditions:

$$\frac{\partial}{\partial \boldsymbol{w}} L = V\boldsymbol{w} - \lambda_1 \mathbf{1} - \lambda_2 \boldsymbol{m} = 0 \tag{3.10}$$

so the efficient set satisfies:

$$\boldsymbol{w} = \lambda_1 V^{-1}\mathbf{1} + \lambda_2 V^{-1}\boldsymbol{\mu} \tag{3.11}$$

Multiplying (3.11) once for $\mathbf{1}^T$ and once for $\boldsymbol{m}^T$ and using constraints we obtain respectively:

$$1 = \lambda_1 A + \lambda_2 B \quad ; \quad E = \lambda_1 B + \lambda_2 C \tag{3.12}$$

where $A = \mathbf{1}^T V^{-1}\mathbf{1}$, $B = \boldsymbol{m}^T V^{-1}\mathbf{1}$ and $C = \boldsymbol{m}^T V^{-1}\boldsymbol{m}$. Solving (3.12) for λ_1, λ_2 we obtain:

$$\lambda_1 = \frac{C - EB}{AC - B^2} \quad ; \quad \lambda_2 = \frac{EA - B}{AC - B^2} \tag{3.13}$$

where $AC - B^2 > 0$. Inserting expressions of λ_1, λ_2 into (3.11):

$$\boldsymbol{w} = \frac{C - EB}{AC - B^2} V^{-1}\mathbf{1} + \frac{EA - B}{AC - B^2} V^{-1}\boldsymbol{m} \tag{3.14}$$

which leads to:

$$\boldsymbol{w}^* = c_1 \frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} + c_2 \frac{V^{-1}\boldsymbol{m}}{\boldsymbol{m}^T V^{-1}\mathbf{1}} \tag{3.15}$$

where $c_1 + c_2 = 1$. This is the desired result. □

We call the two dimensional space generated by $(m_{\boldsymbol{w}}, v_{\boldsymbol{w}}^2)$ the mean variance space (**MV**). The equation of the *minimum variance* set in the **MV** space is obtained multiplying equation (3.14) for $\boldsymbol{w}^T V$ (where now $E = m_{\boldsymbol{w}}$):

$$v_w^2 = \boldsymbol{w}^T V \boldsymbol{w} = \frac{C - m_{\boldsymbol{w}} B}{AC - B^2} + \frac{m_{\boldsymbol{w}} A - B}{AC - B^2} m_{\boldsymbol{w}} = \frac{m_{\boldsymbol{w}}^2 A - 2B m_{\boldsymbol{w}} + C}{AC - B^2} \tag{3.16}$$

The previous equation describes a parabola. If we plot the mean on the vertical axes we obtain, in the mean-standard deviation space, an hyperbola of equation:

$$m_{\boldsymbol{w}} = \frac{B}{A} \pm \frac{\sqrt{A\Delta v_w^2 - \Delta}}{A} \tag{3.17}$$

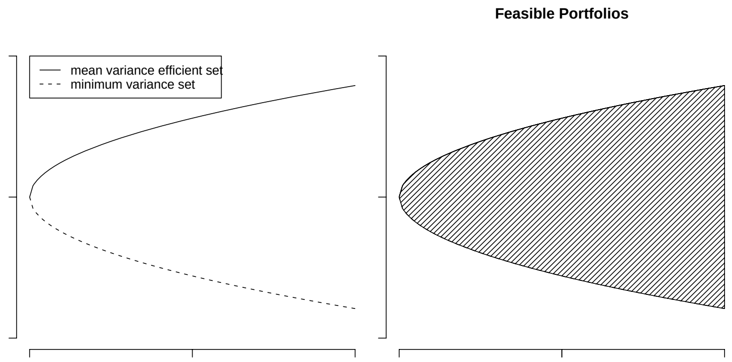


Figure 3.1: Two graphs in the space $(m_{\boldsymbol{w}}, v_{\boldsymbol{w}})$, where the mean is represented in the vertical axes. On the left: the *minimum variance* set and the *efficient* subset. On the right: the set of feasible portfolios.

where $\Delta = AC - B^2$. The graph of this function is illustrated in Figure 3.2. The slopes of the asymptotes of the hyperbola are obtained by the partial derivative of $m_{\boldsymbol{w}}$ with respect to v_w^2:

$$\begin{aligned}\lim_{m_{\boldsymbol{w}}\to\pm\infty} \frac{\partial}{\partial v_w^2} m_{\boldsymbol{w}} = \lim_{m_{\boldsymbol{w}}\to\pm\infty} v_w^2 \frac{AC - B^2}{Am_{\boldsymbol{w}} - B} &= \lim_{m_{\boldsymbol{w}}\to\pm\infty} \frac{\sqrt{AC - B^2}}{Am_{\boldsymbol{w}} - B} \sqrt{m_{\boldsymbol{w}}^2 A - 2Bm_{\boldsymbol{w}} + C} \\ &= \pm\sqrt{\frac{AC - B^2}{A}} \qquad (3.18)\end{aligned}$$

Separating funds with a risk-less asset

We now enlarge the assets set adding an asset whose return R_f is not random. In this case we have

$$(R_0, \boldsymbol{R}) \in \mathbb{R}^{n+1},$$

, with the first component given by $R_0 = R_f$. Furthermore, continuing to denote by V the covariance matrix of the risky returns $\boldsymbol{R}$, we set:

$$\Sigma = \mathrm{Var}((R_0, \boldsymbol{R})) = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & V \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}, \qquad (3.19)$$

hence the first row and the first column of Σ are made of all zeros. Each portfolio $\boldsymbol{w}$ is a vector in $\mathbb{R}^{n+1}$: we denote it by $\boldsymbol{w} = (w_0, \tilde{w})^T$, where $\tilde{w} \in \mathbb{R}^n$.

The optimization problem (3.7), which defines a minimum variance portfolio, now takes the following form:

$$\text{Min} \qquad \tfrac{1}{2}\boldsymbol{w}^T\Sigma\boldsymbol{w} \tag{3.20}$$

$$\text{with the constraints:} \quad w_0 R_f + \tilde{w}^T\boldsymbol{m} = E \tag{3.21}$$

$$w_0 + \mathbf{1}^T\tilde{w} = 1 \tag{3.22}$$

Proposition 3.2.2. *Under the previous assumptions it holds the following: the* minimum variance *set is spanned by the following two portfolios (or* separating funds*):*

$$\boldsymbol{w}_f = (1, 0, \ldots, 0)^T \quad \boldsymbol{w}_t = \left(0, \frac{V^{-1}(\boldsymbol{m} - \mathbf{1}R_f)}{\mathbf{1}^T V^{-1}(\boldsymbol{m} - \mathbf{1}R_f)}\right)^T \tag{3.23}$$

Proof. The Lagrangian of the problem assumes the form:

$$\mathfrak{L} = \frac{1}{2}\boldsymbol{w}^T\Sigma\boldsymbol{w} - \lambda(w_0 + \mathbf{1}^T\tilde{w} - 1) - \gamma(w_0 R_f + \tilde{w}^T\boldsymbol{m} - E) \tag{3.24}$$

and we can derive the first order conditions:

$$\frac{\partial}{\partial\tilde{w}}\mathfrak{L} = V\tilde{w} - \lambda\mathbf{1} - \gamma\boldsymbol{m} = 0 \tag{3.25}$$

$$\frac{\partial}{\partial w_0}\mathfrak{L} = -\lambda - \gamma R_f = 0 \tag{3.26}$$

Inserting the expression $\lambda = -\gamma R_f$ in the first of the two equations above and extracting $\tilde{w}$ we obtain:

$$\tilde{w} = \gamma V^{-1}(\boldsymbol{m} - \mathbf{1}R_f) \;\; ; \;\; w_0 = 1 - \mathbf{1}^T\tilde{w} \tag{3.27}$$

Using both the constraints γ can be obtained, its value is:

$$\gamma = \frac{E - R_f}{(\boldsymbol{m} - \mathbf{1}R_f)^T V^{-1}(\boldsymbol{m} - \mathbf{1}R_f)} = \frac{E - R_f}{C - 2R_f B + R_f^2 A} \tag{3.28}$$

Using this result we can write the generic solution to (3.20) as:

$$\boldsymbol{w}^* = \boldsymbol{w}_f w_0 + \frac{(E - R_f)\mathbf{1}^T V^{-1}(\boldsymbol{m} - \mathbf{1}R_f)}{(\boldsymbol{m} - \mathbf{1}R_f)^T V^{-1}(\boldsymbol{m} - \mathbf{1}R_f)}\boldsymbol{w}_t \tag{3.29}$$

where $w_0 = 1 - 1^T\tilde{w}$. □

As in the previous Section we can now obtain the equation of the *minimum variance* set in the **MV** space. Multiplying (3.27) for $\tilde{w}^T V$, denoting $m_{\boldsymbol{w}} = w_0 R_f + m_{\tilde{w}}$, $m_{\tilde{w}} = \tilde{w}^T\boldsymbol{m}$ and considering that $E - R_f = m_{\boldsymbol{w}} - R_f = m_{\tilde{w}} - R_f\mathbf{1}^T\tilde{w}$, we get:

$$v_{\boldsymbol{w}}^2 = \boldsymbol{w}^T\Sigma\boldsymbol{w} = \tilde{w}^T V\tilde{w} = \frac{(E - R_f)^2}{C - 2R_f B + R_f^2 A} \tag{3.30}$$

In the M-V space this is the equation of two rays with common intercept in R_f :

$$m_{\boldsymbol{w}} = R_f \pm v_{\boldsymbol{w}} \sqrt{C - 2R_f B + R_f^2 A} \tag{3.31}$$

Again the minimum variance set is spanned by only two funds. The portfolio $\boldsymbol{w}_t$ is called the *tangency portfolio*: it is the only minimum variance portfolio that belongs to the hyperbola representing portfolios investing only in risky assets. The mean and variance of this portfolio are:

$$m_t = \boldsymbol{w}_t^T \boldsymbol{m} = \frac{C - BR_f}{B - AR_f} \tag{3.32}$$

$$v_t^2 = \boldsymbol{w}_t^T V \boldsymbol{w}_t = \frac{C - 2R_f B + R_f^2 A}{(B - AR_f)^2} = \frac{m_t - R_f}{B - AR_f} \tag{3.33}$$

The tangent line drawn from the risk-less asset through the tangency portfolio in the **MV** space is called *security market line*.

Expected returns of assets

In this Section we characterize the expected returns of the assets in the market.
Given any portfolio $\boldsymbol{w}_p$ and its return $R_{\boldsymbol{w}_p} = \boldsymbol{w}_p^T \boldsymbol{R}$ we define the *beta* of the portfolio to be the vector

$$\boldsymbol{\beta}_p = \boldsymbol{v}_p / v_p^2$$

where $\boldsymbol{v}_p = V \boldsymbol{w}_p$ is the vector of covariances of $\boldsymbol{R}$ with $R_{\boldsymbol{w}_p}$ and where $v_p^2 = \mathrm{Var}(R_{\boldsymbol{w}_p})$.

We begin to analyze expected returns when a risk-less asset is available.
For the tangency portfolio we have:

$$\boldsymbol{v}_t = V \boldsymbol{w}_t = \frac{(\boldsymbol{m} - \mathbf{1} R_f)}{\mathbf{1}^T V^{-1} (\boldsymbol{m} - \mathbf{1} R_f)} = \frac{(\boldsymbol{m} - \mathbf{1} R_f)}{B - AR_f}, \tag{3.34}$$

so, by using (3.33), we obtain the fundamental relation:

$$\boldsymbol{m} - \mathbf{1} R_f = \boldsymbol{v}_t \frac{m_t - R_f}{v_t^2} = \boldsymbol{\beta}_t (m_t - R_f) \tag{3.35}$$

Therefore any asset uncorrelated with the tangency portfolio, has an expected return equal to R_f
Let us consider a market where all assets are risky. As mentioned above, in this case the minimum variance set is an hyperbola spanned by the two portfolios $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.

The covariance vector $\boldsymbol{v}_a$ of the generic portfolio $\boldsymbol{w}_a$ lying in the *minimum variance* set,

$$\boldsymbol{w}_a = (1-\lambda)\boldsymbol{a}_1 + \lambda\boldsymbol{a}_2,$$

is given by:

$$\boldsymbol{v}_a = V\boldsymbol{w}_a = \frac{1-\lambda}{A}\mathbf{1} + \frac{\lambda}{B}\boldsymbol{m} \tag{3.36}$$

where we have used (3.15). Let $\boldsymbol{w}_p$ a further portfolio of the *minimum variance* set, multiplying (3.36) once for $\boldsymbol{w}_a$ and once for $\boldsymbol{w}_p$:

$$v_a^2 = \frac{1-\lambda}{A} + \frac{m_a\lambda}{B}, \qquad \boldsymbol{v}_{pa} = \frac{1-\lambda}{A} + \frac{m_p\lambda}{B} \tag{3.37}$$

Substituting into (3.36) the above equalities we obtain:

$$\boldsymbol{m} = \frac{m_p v_a^2 - m_a v_{ap}}{v_a^2 - v_{ap}}\mathbf{1} + \frac{m_a - m_p}{m_a^2 - v_{ap}}\boldsymbol{v}_a \tag{3.38}$$

If in addition the portfolio $\boldsymbol{w}_a$ is uncorrelated to portfolio $\boldsymbol{w}_p$, namely $\sigma_{ap} = 0$, then

$$\boldsymbol{m} = m_p\mathbf{1} + \frac{m_a - m_p}{v_a^2}\boldsymbol{v}_a = m_p\mathbf{1} + (m_a - m_p)\boldsymbol{\beta}_a. \tag{3.39}$$

3.3 Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) is maybe the most widely used model in finance. The CAPM equation is a pricing equation relating the expected return of each asset to the expected return of the market portfolio.
The CAPM result relies on the following assumptions:

Assumption 3.3.1. *All the M investors are utility maximizers and they all maximize expected utility through mean variance analysis, that is among all portfolios of equal mean returns they prefer the portfolio with the smallest variance of returns. All have the same time horizon and the same beliefs about the values of the parameters $(\boldsymbol{m}, V)$, representing mean and covariance of securities returns.*
N is the total number of Firms and each Firm contributes in the market with n_i securities with $\sum n_i = n$.
The risk-less asset can be bought or sold in unlimited amounts.

Under these assumptions all investors hold a minimum variance portfolio. Since all portfolios in the *minimum variance* set are combination of only two portfolios the j-th investor's portfolio can be written as:

$$\boldsymbol{w}^{(j)} = \gamma_j \boldsymbol{w}_f + (1-\gamma_j)\boldsymbol{w}_t.$$

Then the aggregate portfolio $\boldsymbol{w}^* = \sum_{j=1}^{M} \boldsymbol{w}^{(j)}$ representing the investments of all the M investors will be again a linear combination of only the two portfolios written above. Furthermore the risky component of the aggregate portfolio will consist solely of the *tangency portfolio*. Next we have the following

Definition 3.3.1. *The market portfolio is the portfolio representing the total supply of assets in the market.*

By consequence if p_i is the price of the security of the i-th Firm, and

$$C_{tot} = \sum_{i=1}^{N} n_i p_i,$$

then the market portfolio is given by:

$$\boldsymbol{w}_m = \left(\frac{n_1 p_1}{C_{tot}}, \ldots, \frac{n_N p_N}{C_{tot}}\right) \tag{3.40}$$

In equilibrium the aggregate demand of risky assets, represented solely by the *tangency portfolio*, is equal to the total supply, i.e. the market portfolio. Therefore:

$$\boldsymbol{w}_m = \boldsymbol{w}_t$$

The CAPM equation is then easily obtained inserting this result into (3.35):

$$\boldsymbol{m}^e - \mathbf{1}R_f = \boldsymbol{\beta}_m(m_m - R_f) \tag{3.41}$$

where $m_m = \boldsymbol{w}_m^T \boldsymbol{R}$.
In absence of a risk-less asset a similar way of reasoning together with the use of (3.39), leads to the following version of the CAPM equation

$$\boldsymbol{m}^e = m_z \mathbf{1} + (m_m - m_z)\boldsymbol{\beta}_m \tag{3.42}$$

due to F.Black. In (3.42) $m_z = \boldsymbol{w}_z^T \boldsymbol{R}$ where $\boldsymbol{w}_z$ is a zero beta portfolio whose returns are uncorrelated with the market.

3.4 Compatibility with expected utility maximization

As argued in Section 3.2 mean variance analysis is fully compatible with utility maximization only in two cases: when the utility is quadratic or when the returns are elliptically distributed. In this Section we prove this statement.
First of all consider the case of n risky assets whose returns are described by $\boldsymbol{R}$ which is normal $N_n(\boldsymbol{m}, V)$. We know from the standard properties of the normal distribution that each affine transformation of $\boldsymbol{R}$ gives rise to another normal r.v. Specifically the random variable $R_{\boldsymbol{w}} = \boldsymbol{w}^T\boldsymbol{R}$ is distributed as

$$R_{\boldsymbol{w}} \sim N(\boldsymbol{w}^T\boldsymbol{m}, \boldsymbol{w}^T V\boldsymbol{w})$$

Hence expected utility is a function only of the mean and of the variance of portfolio return:

$$\mathbb{E}(u(R_{\boldsymbol{w}})) = \int \frac{1}{(\boldsymbol{w}^T V\boldsymbol{w})^{1/2}} u(r)\varphi(\frac{r-\boldsymbol{w}^T\boldsymbol{m}}{(\boldsymbol{w}^T V\boldsymbol{w})^{1/2}})dr \equiv \lambda(\boldsymbol{w}^T\boldsymbol{m}, \boldsymbol{w}^T V\boldsymbol{w}) \tag{3.43}$$

In addition Proposition 2.1.2, which states the **SD** properties for normal prospects, implies that , given two portfolios $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$, if:

$$\mathbb{E}(R_{\boldsymbol{w}_1}) \geq \mathbb{E}(R_{\boldsymbol{w}_2}) \quad \text{and} \quad \mathrm{Var}(R_{\boldsymbol{w}_1}) \leq \mathrm{Var}(R_{\boldsymbol{w}_2}) \tag{3.44}$$

then:

$$\mathbb{E}(u(R_{\boldsymbol{w}_1})) \geq \mathbb{E}(u(R_{\boldsymbol{w}_2})) \tag{3.45}$$

for all $u \in \mathcal{U}_2$. Therefore we recover the same conditions stated in (3.2).
These properties assure that in the normal assumption mean variance analysis and utility maximization are compatible.
The normal is not the only distribution that guarantees compatibility between utility maximization and mean variance analysis. The larger class of elliptical distributions preserves the same fundamental property.

Elliptical distributions

It can be shown that a class of distributions will give rise to an utility maximization procedure compatible with the mean variance analysis whenever: i)the class is closed under affine transformations; ii) all moments can be written as functions only of the first

two moments;iii) a portfolio with given variance and expected return is stochastically dominated (at second order) by another one with same expected return but smaller variance.
From results of Appendix A we already know that the class of elliptical distributions satisfies conditions i) and ii). We now prove the validity of condition iii).
From (A.1) a 1-dim elliptical random variable Y has a density of the form:

$$f_Y(y;\mu,\sigma^2) = \frac{1}{\sigma} f(\frac{y-\mu}{\sigma}) \tag{3.46}$$

Proposition 3.4.1. *Given two elliptical random normal variables, with the same distribution, $X_1 \sim Ell(\mu_1, \sigma_1)$ and $X_2 \sim Ell(\mu_2, \sigma_2)$ which have $F_1(x)$ and $F_2(x)$ as their cumulative distributions functions, then $F_1(x)$ intersects only once from below $F_2(x)$ if and only if $\sigma_1 < \sigma_2$*

Proof. By (3.46) the cumulative distribution function of $Z \sim \text{Ell}(\mu, \sigma)$ can be written in the form:

$$F_Z(x;\mu,\sigma) = \int_{-\infty}^{x} \frac{1}{\sigma} f_Z(z;\mu,\sigma^2)dz = \int_{-\infty}^{\frac{x-\mu}{\sigma}} f_Z(z)dz \tag{3.47}$$

Performing the partial derivative with respect to σ pointwise in x and with μ fixed we obtain:

$$\frac{\partial}{\partial\sigma} F_Z(x;\mu,\sigma) = -\frac{x-\mu}{\sigma^2} f_Z(\frac{x-\mu}{\sigma}) \tag{3.48}$$

This implies $\partial_\sigma F_Z(x;\mu,\sigma) < 0$ for each $x > \mu$ and $\partial_\sigma F_Z(x;\mu,\sigma) > 0$ for each $x < \mu$. That is the intersection point is unique in $x = \mu$ and for $x < \mu$ $F_Z(x;\mu,\sigma_1) < F_Z(x;\mu,\sigma_2)$ if $\sigma_1 < \sigma_2$. □

Corollary 3.4.1. *Given two elliptical random normal variables, with the same distribution, $X_1 \sim Ell(\mu_1, \sigma_1)$ and $X_2 \sim Ell(\mu_2, \sigma_2)$, if $\mu_1 \geq \mu_2$ $\sigma_1 \leq \sigma_2$ then $X_1 \succeq_2 X_2$*

Proof. The result is obvious considering the previous result and the Theorem 2.1.2. □

This Corollary implies the validity of iii). As a consequence the elliptical class is consistent with utility maximization in case of concave utility functions.

Quadratic Utility

Here we drastically restrict the class of utility functions assuming the utility is quadratic:

$$u(x) = x - \frac{1}{2}bx^2 \qquad b > 0 \tag{3.49}$$

For this specific form the expected utility of any random variable (or uncertain prospect) X can be explicitly computed:

$$\begin{aligned} \mathbb{E}(u(X)) = \int u(x)f_X(x)dx &= \int (x - \frac{1}{2}bx^2)f_X(x)dx \\ &= \mathbb{E}(X) - \frac{1}{2}b(\mathbb{E}(X)^2 + \mathrm{Var}(X)) \end{aligned} \tag{3.50}$$

Clearly it holds:

$$\frac{\partial}{\partial \mathrm{Var}(X)}\mathbb{E}(u(X)) < 0$$

and, if we restrict our considerations to the region of increasing utility, which is $x < 1/b$, then we also have:

$$\frac{\partial}{\partial \mathbb{E}(X)}\mathbb{E}(u(X)) > 0.$$

3.5 Ross's Separation Theorems

In this Section we briefly recall the Theory developed by Ross on k-funds separability, presented in [32]. The main references are Ross [32] and Ingersoll [18].
All the proofs of this Section are reported in Appendix 2.

Proposition 3.5.1. *(i)* $Y \succeq_1 X$ *iff* $X \sim Y + Z$ *for some r.v.* Z *such that* $Z \leq 0$
(ii) $Y \succeq_2 X$ *iff* $X \sim Y + Z + \epsilon$ *for some r.v.'s* Z, ϵ *such that* $Z \leq 0$ *and* $\mathbb{E}(\epsilon | Y + Z) = 0$.

Proof. : Given in Appendix 2 □

Definition 3.5.1. *: We shall say that,*
(A) the distribution of $\boldsymbol{R}$ *has the (strong) 1-fund separation property if there exists a portfolio* $\boldsymbol{w}_a$ *such that for any portfolio* $\boldsymbol{w}_b$ *it holds*

$$\boldsymbol{w}_a^T \boldsymbol{R} \succeq_2 \boldsymbol{w}_b^T \boldsymbol{R}$$

(B) the distribution of $\boldsymbol{R}$ has the (strong) 2-funds separation property if there exist two portfolios $\boldsymbol{w}_1, \boldsymbol{w}_2$ such that for any portfolio $\boldsymbol{w}_b$ there is a portfolio $\boldsymbol{w}_a$ given by a linear combination of $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ for which it holds

$$\boldsymbol{w}_a^T \boldsymbol{R} \succeq_2 \boldsymbol{w}_b^T \boldsymbol{R}.$$

(C) the distribution of $\boldsymbol{R}$ has the (strong) 3-funds separation property if there exist three portfolios $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ such that for any portfolio $\boldsymbol{w}_b$ there is a portfolio $\boldsymbol{w}_a$ given by a linear combination of $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ for which it holds

$$\boldsymbol{w}_a^T \boldsymbol{R} \succeq_2 \boldsymbol{w}_b^T \boldsymbol{R}.$$

Remark: In a similar fashion k-funds ($k \geq 4$) can be defined and discussed. We give details of the three funds separation in Section 4.6.

Ross in [32] characterizes the families of distributions of $\boldsymbol{R}$ which have the k-fund separation property. Here we discuss such results for the case of 1-fund SP and 2-fund SP.

Theorem 3.5.1. *The distribution of $\boldsymbol{R}$ has the (strong) 1-fund separation property iff there exist a scalar r.v. Y, a vector r.v. $\boldsymbol{\epsilon}$ and a portfolio $\boldsymbol{\alpha}$ such that*
(a) each component of $\boldsymbol{R}$ can be written as $R_i = Y + \epsilon_i$, for $i = 1, \ldots, n$
(b) it holds $\mathbb{E}(\epsilon_i | Y) = 0$
(c) the portfolio $\boldsymbol{\alpha}$ is orthogonal to the vector $\boldsymbol{\epsilon}$ (i.e. $\boldsymbol{\alpha}^T \boldsymbol{\epsilon} = 0$).

Remark: For obvious reasons Y is called the "common (risky) factor" and the noise ϵ_i the asset-specific "residual risk".

Proof. : Given in Appendix 2 □

Theorem 3.5.2. *(2-funds separation without risk-less asset) The distribution of $\boldsymbol{R}$ has the (strong) 2-fund separation property iff*
there exist two scalar r.v.'s Y, Z, a vector r.v. $\boldsymbol{\epsilon}$, a (deterministic) vector $\boldsymbol{b}$, and two portfolios $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that
(a) each component of $\boldsymbol{R}$ can be written as $R_i = Y + b_i Z + \epsilon_i$, for $i = 1, \ldots, n$
(b) $\mathbb{E}[\epsilon_i | Y + b_i Z] = 0 \ \ \forall i$
(c) $\sum \alpha_i \epsilon_i = 0 = \sum \beta_i \epsilon_i$

Proof. : Given in [18], pag 156. □

Theorem 3.5.3. *(2-funds separation with a risk-less asset) The distribution of* $\boldsymbol{R}$ *has the (strong) 2-fund separation property iff*
there exist a scalar r.v.Y*, a vector r.v.* $\boldsymbol{\epsilon}$*, a (deterministic) vector* $\boldsymbol{b}$*, and two portfolios* $\boldsymbol{\alpha}, \boldsymbol{\beta}$ *such that*
(a) each component of $\boldsymbol{R}$ *can be written as* $R_i = R_f + b_i Y + \epsilon_i$*, for* $i = 1, \ldots, n+1$
(b) $\mathbb{E}[\epsilon_i|Y] = 0 \forall i$
(c) $\sum \alpha_i \epsilon_i = 0 = \sum \beta_i \epsilon_i$

Proof. : Given in [18], pag 152. □

Remark 1:Theorem 3.5.2 provides for two conditions to be checked in a market model for risky assets:
1) the return of each asset can be written in the form specified by (a)
2) there are two distinct portfolios orthogonal to the vector $\boldsymbol{\epsilon}$ of the residual risks, namely $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.
Whenever these conditions are verified by the distributional model for the asset returns, any portfolio generated by the linear combination of the portfolios $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ will have a return which always stochastically dominates (at the second order) the returns of any other portfolio. The viceversa holds true as well.

Proposition 3.5.2. *The normal distribution has the 2-funds separating property.*

Proof. When risky assets are assumed to be normal distributed the mean variance analysis is consistent with utility maximization. Then for normal returns Proposition 3.2.1 holds true.
Consequently each portfolio is dominated by a portfolio lying in the *minimum variance* set. In addition, the expression (3.15) implies that the generic portfolio of the *minimum variance set* is a linear combination of only two portfolios. □

The previous result is equivalent to the following one:

Proposition 3.5.3. *If the vector of returns is normally distributed then it satisfies conditions of Theorem 3.5.3*

Proof. : Given in Appendix 2. □

4

Investing in non-normal Markets

4.1 The Simaan model

If the returns of assets are assumed to be non-elliptical or the utility function to be non quadratic, the mean variance analysis loses its validity. As a consequence the CAPM equation itself needs to be reconsidered.

In relation with these problems Simaan in [35] presents a novel framework. In his work, among various results, he is able to obtain a consistent equation of CAPM-type relaxing the standard assumption of normality (or ellipticity) on the distribution of assets returns. It is worth noting that in Simaan's approach the expected utility , rather than being a function of only the mean and the variance, is a function of three parameters: the mean, the variance and the skewness of the portfolio return. That is to say:

$$\mathbb{E}(u(R_{\boldsymbol{w}})) = \lambda(\mathbb{E}(R_{\boldsymbol{w}}), \mathrm{Var}(R_{\boldsymbol{w}}), \mathrm{Skew}(R_{\boldsymbol{w}})) \tag{4.1}$$

Furthermore, in his model, the distribution of the assets returns turns out to have the 3-funds separation property.

The peculiarity of the model is that, contrary to the mean variance analysis where the only risk measure was the variance, it admits two measures of risk. This is a direct consequence of the specific form of the variance of the portfolio return which splits into two components respectively called the *spherical* and the *non-spherical* components. The returns are no more symmetric, moreover their skewness is a function only of the *non-spherical* part of the variance.

In the Simaan model an investor controls his portfolio choice through three parameters: the mean and the two components of the portfolio variance. Minimizing the spherical variance of the portfolio return for given mean and non-spherical variance is the correct optimization procedure which a risk-averse agent must implement in order to select the most efficient portfolio.
In the following Section we recall briefly the main aspects of the Simaan model.

4.1.1 The Simaan Market Model

Assumption 4.1.1. *Given two deterministic vectors $\boldsymbol{\mu}, \boldsymbol{b} \in \mathbb{R}^n$, the assets returns are distributed according to:*

$$R_i = \mu_i + b_i\, Y + \epsilon_i \tag{4.2}$$

$\imath = 1, \ldots, n$, *where the joint distribution of $\boldsymbol{\epsilon}$ is elliptical* [1] *conditionally on Y, with law:*

$$\boldsymbol{\epsilon}|Y \sim Ell_n(0, W, \psi), \tag{4.3}$$

and where Y follows a univariate non-elliptical distribution.

In vector notation (4.2) can be written:

$$\boldsymbol{R} = \boldsymbol{\mu} + \boldsymbol{b}\, Y + \boldsymbol{\epsilon} \tag{4.4}$$

The following straightforward results, see Simaan [35] , illustrate some aspects of $\boldsymbol{R}$:

Theorem 4.1.2. *The characteristic function of any random vector $\boldsymbol{R}$ distributed according to formula (4.2) is given by:*

$$\Phi_{\boldsymbol{R}}(\boldsymbol{t}) = exp(i\boldsymbol{t}^T \boldsymbol{\mu})\Phi_{\boldsymbol{\epsilon}}(\boldsymbol{t})\Phi_Y(\boldsymbol{t}^T \boldsymbol{b}) \tag{4.5}$$

where $\Phi_{\boldsymbol{\epsilon}}(\boldsymbol{t})$ and $\Phi_Y(\boldsymbol{t}^T\boldsymbol{b})$ are the characteristic functions of the vector $\boldsymbol{\epsilon}$ and of the random variable Y respectively.

[1] See Appendix 1 for elliptical distributions. We remark that Simaan in [35] (Definition1. on pg.569) calls spherical what here and in Appendix 1 is called an elliptical distribution. Usually the statistical literature assigns the name "spherical" only to the class $Ell_n(0, Id, \psi)$.

Theorem 4.1.3. *If $\mathbb{E}(Y) = 0$ and $Var(\boldsymbol{\epsilon})$ is finite then the first three joint moments of $\boldsymbol{R}$ are given as*

$$\mathbb{E}(\boldsymbol{R}) = \boldsymbol{\mu} \tag{4.6}$$

$$Var(\boldsymbol{R}) = k\,W + \boldsymbol{b}\boldsymbol{b}^T\,\sigma_Y^2 = V_S + \sigma_Y^2\,V_{ns} \tag{4.7}$$

$$Skew_{ijk}(\boldsymbol{R}) = \mathbb{E}[(R_i - \mu_i)(R_j - \mu_j)(R_k - \mu_k)] = b_i b_j b_k \mathbb{E}(Y^3) \tag{4.8}$$

where k is a positive constant and $\sigma_Y^2 = Var(Y)$.

This shows that the skewness of $\boldsymbol{R}$ is a function of $\boldsymbol{b}$ and, in general, of the distribution parameters of Y. We also observe in (4.7) the splitting of the variance into two components that Simaan calls respectively the *spherical* and *non-spherical* components[1] (only the latter depends on the vector $\boldsymbol{b}$).

We now come to the properties of the portfolio return $R_{\boldsymbol{w}} = \boldsymbol{w}^T\boldsymbol{R}$:

Theorem 4.1.4. *The characteristic function of the r.v. $R_{\boldsymbol{w}}$ depends only on $\boldsymbol{w}^T\boldsymbol{\mu}$, $\boldsymbol{w}^T\boldsymbol{b}$ and $\boldsymbol{w}^T W\boldsymbol{w}$*

Proof. By (4.5) we have:

$$\begin{aligned}\Phi_{\boldsymbol{w}^T\boldsymbol{R}}(t) &= \exp(it(\boldsymbol{w}^T\boldsymbol{\mu})))\Phi_{\boldsymbol{\epsilon}}(t\boldsymbol{w})\Phi_Y(t\boldsymbol{w}^T\boldsymbol{b}) \\ &= \exp(it\boldsymbol{w}^T\boldsymbol{\mu})\psi(t^2\boldsymbol{w}^T W\boldsymbol{w})\Phi(t\boldsymbol{w}^T\boldsymbol{b})) \\ &= h(t, \boldsymbol{w}^T\boldsymbol{\mu}, \boldsymbol{w}^T\boldsymbol{b}, \boldsymbol{w}^T W\boldsymbol{w}) \end{aligned} \tag{4.9}$$

□

Corollary 4.1.1. *The expected utility of any portfolio $\boldsymbol{w}$ is determined by its mean, its variance and its skewness*

4.2 The skew normal case

The main purpose of this Section is to show that a skew normal distributed random vector $\boldsymbol{R}$, with parameters taking a certain form, can be recognized to satisfy the Simaan's assumptions [35] .

[1] We have chosen to preserve here the original Simaan's names

The following Proposition provides for a new representation (called the representation for convolution) of a multivariate skew-normal r.v. See Azzalini [3] for more details.

Consider two vectors $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T \in \mathbb{R}^n$ and $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)^T$ where $\delta_i \in (-1, 1)$, and two $(n \times n)$ matrices:

$$\Delta = \begin{pmatrix} \delta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \delta_n \end{pmatrix}, \quad \omega = \begin{pmatrix} \omega_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \omega_n \end{pmatrix}$$

where $\omega_i > 0 \; \forall i$.

Then it holds:

Proposition 4.2.1. *Given two random independent variables $X \sim N(0,1)$ and $\boldsymbol{Z} \sim N_n(0, \Psi)$, the random vector:*

$$\boldsymbol{R} = \boldsymbol{\mu} + (\omega\boldsymbol{\delta})|X| + \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z} \tag{4.10}$$

is skew normally distributed with $\boldsymbol{R} \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$ where $\Omega = \omega\overline{\Omega}\omega$,

$$\overline{\Omega} = \boldsymbol{\delta}\boldsymbol{\delta}^T + (Id - \Delta^2)^{1/2}\Psi(Id - \Delta^2)^{1/2} \tag{4.11}$$

and where $\boldsymbol{\alpha} = \frac{\overline{\Omega}^{-1}\boldsymbol{\delta}}{(1-\boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta})^{1/2}}$.

Remark: The matrix $\overline{\Omega}$ is a correlation matrix: calling with ψ_{ij} the elements of Ψ, we have ($\psi_{ii} = 1$ because Ψ is a correlation matrix):

$$\overline{\Omega} = \begin{pmatrix} \delta_1^2 & \dots & \delta_1\delta_n \\ \vdots & \ddots & \vdots \\ \delta_1\delta_n & \dots & \delta_n^2 \end{pmatrix} + \begin{pmatrix} 1-\delta_1^2 & \dots & \psi_{1n}\sqrt{1-\delta_1^2}\sqrt{1-\delta_n^2} \\ \vdots & \ddots & \vdots \\ \psi_{1n}\sqrt{1-\delta_1^2}\sqrt{1-\delta_n^2} & \dots & 1-\delta_n^2 \end{pmatrix}$$

and so:

$$\overline{\Omega} = \begin{pmatrix} 1 & \dots & \delta_1\delta_n + \psi_{1n}\sqrt{1-\delta_1^2}\sqrt{1-\delta_n^2} \\ \vdots & \ddots & \vdots \\ \delta_1\delta_n + \psi_{1n}\sqrt{1-\delta_1^2}\sqrt{1-\delta_n^2} & \dots & 1 \end{pmatrix}$$

with $-1 \leq \delta_i\delta_j + \psi_{ij}\sqrt{1-\delta_i^2}\sqrt{1-\delta_j^2} \leq 1$ being $|\psi_{ij}| \leq 1$ and $|\delta_i| < 1$. [1]

[1] Claim: If $a, b, c \in [-1, 1]$ then $-1 \leq ab + c\sqrt{1-a^2}\sqrt{1-b^2} \leq 1$. Proof: it holds $ab + c\sqrt{1-a^2}\sqrt{1-b^2} \leq ab + \sqrt{1-a^2}\sqrt{1-b^2} = y$ and $y^2 \leq 1$ because $a^2 + b^2 \geq 2ab\sqrt{1-a^2}\sqrt{1-b^2}$. Then $ab + c\sqrt{1-a^2}\sqrt{1-b^2} \leq 1$. The proof that $ab + c\sqrt{1-a^2}\sqrt{1-b^2} \geq -1$ is analogue.

Proof. We have to show that the random vector $\boldsymbol{R}$ has (1.25) as density . We denote by $\tilde{\boldsymbol{R}}$ the random vector such that:

$$\boldsymbol{R} = \boldsymbol{\mu} + \omega\tilde{\boldsymbol{R}}$$

or equivalently:

$$\tilde{\boldsymbol{R}} = \boldsymbol{\delta}|X| + (Id - \Delta^2)^{1/2}\boldsymbol{Z} \tag{4.12}$$

If we prove that $\tilde{\boldsymbol{R}} \sim SN_n(0, \overline{\Omega}, \boldsymbol{\alpha})$ the result is then easily obtained applying the standard properties of a multivariate skew-normal r.v.
To obtain the density of $\tilde{\boldsymbol{R}}$ first we consider the distribution of $\tilde{\boldsymbol{R}}$ conditionally on X:

$$\tilde{\boldsymbol{R}}|X = x \quad \sim \quad N(\boldsymbol{\delta}x, W) \tag{4.13}$$

where

$$W = (Id - \Delta^2)^{1/2}\Psi(Id - \Delta^2)^{1/2}. \tag{4.14}$$

Then

$$\begin{aligned} f_{\tilde{\boldsymbol{R}}}(\boldsymbol{z}) &= \int f_{\tilde{\boldsymbol{R}}|X}(\boldsymbol{r}, x) f_{|X|}(x)dx \\ &= \int_0^{\infty} 2\varphi(x; 0, 1)\ \varphi(\boldsymbol{r} - \boldsymbol{\delta}x; 0, W)dx \end{aligned} \tag{4.15}$$

Some algebraic manipulation is needed in order to prove that:

$$(\boldsymbol{r} - \boldsymbol{\delta}x)^T W^{-1}(\boldsymbol{r} - \boldsymbol{\delta}x) + x^2 = \boldsymbol{r}^T(w + \boldsymbol{\delta}\boldsymbol{\delta}^T)^{-1}\boldsymbol{r} + \frac{(x - \boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{r})^2}{1 - \boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta}} \tag{4.16}$$

Using the previous equality, formula (4.15) becomes:

$$\begin{aligned} f_{\tilde{\boldsymbol{R}}}(\boldsymbol{r}) &= 2\varphi(\boldsymbol{r}; 0, W + \boldsymbol{\delta}\boldsymbol{\delta}^T)\int_0^{\infty} \varphi(x; \boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{r}, 1 - \boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta})dx \\ &= 2\varphi(\boldsymbol{r}; 0, \overline{\Omega})\int_{-\frac{\boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{r}}{\sqrt{1-\boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta}}}}^{\infty} \varphi(x; 0, 1)dx \\ &= 2\varphi(\boldsymbol{r}; 0, \overline{\Omega})\int_{-\infty}^{\frac{\boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{r}}{\sqrt{1-\boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta}}}} \varphi(x; 0, 1)dx \\ &= 2\varphi(\boldsymbol{r}; 0, \overline{\Omega})\Phi\left(\frac{\boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{r}}{\sqrt{1 - \boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta}}}\right) \end{aligned} \tag{4.17}$$

Finally adding the location parameter $\boldsymbol{\mu}$ and the scale parameter ω we obtain the desired result:

$$f_{\boldsymbol{R}}(\boldsymbol{r}) = 2\varphi(\boldsymbol{r};\boldsymbol{\mu},\Omega)\Phi\left(\frac{\boldsymbol{\delta}^T\overline{\Omega}^{-1}\omega^{-1}(\boldsymbol{r}\boldsymbol{\mu})}{\sqrt{1-\boldsymbol{\delta}^T\overline{\Omega}^{-1}\boldsymbol{\delta}}}\right) \tag{4.18}$$

$$= 2\varphi(\boldsymbol{r};\boldsymbol{\mu},\Omega)\Phi\left(\boldsymbol{\alpha}^T\omega^{-1}(\boldsymbol{r}-\boldsymbol{\mu})\right) \tag{4.19}$$

□

To summarize, the model (4.10) given by:

$$\boldsymbol{R} = \boldsymbol{\mu} + (\omega\boldsymbol{\delta})|X| + \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z} \tag{4.20}$$

can be obtained by (4.2):

$$\boldsymbol{R} = \boldsymbol{\mu} + \boldsymbol{b}\,Y + \boldsymbol{\epsilon} \tag{4.21}$$

with the following choices:

$$\begin{aligned} \boldsymbol{\mu} &\Rightarrow \boldsymbol{\mu} \\ \boldsymbol{b} &\Rightarrow \omega\boldsymbol{\delta} \\ Y &\Rightarrow |X| \\ \boldsymbol{\epsilon} &\Rightarrow \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z}. \end{aligned} \tag{4.22}$$

The moments of (4.20) can be obtained by formulas (1.28), (1.29) but it is useful to derive them directly from (4.101). Being $\mathbb{E}(|X|) = \sqrt{\frac{2}{\pi}}$, the mean value is

$$\mathbb{E}[\boldsymbol{R}] = \boldsymbol{\mu} + \sqrt{\frac{2}{\pi}}\omega\boldsymbol{\delta}. \tag{4.23}$$

and the covariance matrix is

$$V = \text{Var}(\boldsymbol{R}) = \text{Cov}((\omega\boldsymbol{\delta})|X| + \omega(Id-\Delta^2)^{1/2}\boldsymbol{Z}, \omega\boldsymbol{\delta}|X| + \omega(Id-\Delta^2)^{1/2}\boldsymbol{Z})$$

$$= \text{Var}(\omega\boldsymbol{\delta}|X|) + \text{Var}(\omega(Id-\Delta^2)^{1/2}\boldsymbol{Z}))$$

$$= \text{Var}(|X|)[(\omega\boldsymbol{\delta})(\omega\boldsymbol{\delta})^T] + W$$

with W as in (4.14).

Following Simaan, we call

- $V_{\mathbf{S}} := W$ the *spherical* component of the variance
- $V_{\mathbf{ns}} := [(\omega\boldsymbol{\delta})(\omega\boldsymbol{\delta})^T]$ the *non spherical* component of the variance,

then

$$V = V_{\mathbf{S}} + \sigma^2_{|X|} V_{\mathbf{ns}},$$

We notice that the covariance V can also be written in the form

$$V = \Omega - \frac{2}{\pi}[(\omega\boldsymbol{\delta})(\omega\boldsymbol{\delta})^T] \tag{4.24}$$

Indeed this follows easily from $\Omega = \omega\overline{\Omega}\omega$, (4.11), (4.14) and the fact that $\sigma^2_{|X|} = 1 - \frac{2}{\pi}$. Furthermore, the generic element (i, j, k) of $Skew(\boldsymbol{R})$,the skewness of the returns , takes the form:

$$\text{Skew}_{ijk}(\boldsymbol{R}) = (\omega_i\delta_i)(\omega_j\delta_j)(\omega_k\delta_k)Skew(|X|)$$

To conclude the section we remark that the representation (4.10) can also be rewritten in the form

$$\begin{aligned} \boldsymbol{R} &= (\boldsymbol{\mu} + \sqrt{\frac{2}{\pi}}\omega\boldsymbol{\delta}) + (\omega\boldsymbol{\delta})(|X| - \sqrt{\frac{2}{\pi}}) + \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z} && (4.25) \\ &\equiv \boldsymbol{\mu}' + \boldsymbol{b}\, Y' + \boldsymbol{\epsilon} && (4.26) \end{aligned}$$

with $\boldsymbol{\mu}' = \mathbb{E}[\boldsymbol{R}]$. This decomposition verifies the hypothesis of Theorem 4.1.3. However, in this thesis we have chosen to continue to work with the representation (4.21)rather than (4.25)[1].

4.3 Portfolio selection

In this section we discuss the expected utility of an investor who holds a portfolio in a market of risky assets. The main underlying assumption is that assets returns $\boldsymbol{R}$ follow a (multivariate) skew-normal distribution according to (4.10).

At time t, the investor is faced with the decision to choose the best portfolio between all feasible portfolios. We assume that when the investment is made, he invests 1 euro. Then the portfolio is hold unchanged until time $\tau > t$.

At time τ the value of the portfolio is therefore given by the realization of the univariate r.v.

$$R_{\boldsymbol{w}} = \boldsymbol{w}^T\boldsymbol{R} \sim SN(\mu_{\boldsymbol{w}}, \sigma^2_{\boldsymbol{w}}, \alpha_{\boldsymbol{w}}), \tag{4.27}$$

[1]There, differently from Simaan's paper, the symbol $\boldsymbol{\mu}$ denotes the location of the returns distribution

representing the portfolio return.

For market parameters $(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$ as in (4.2.1), by using (1.32)[1], the portfolio parameters in (4.27)are given by

$$\mu_{\boldsymbol{w}} = \boldsymbol{w}^T \boldsymbol{\mu} \tag{4.28}$$

$$\sigma^2_{\boldsymbol{w}} = \boldsymbol{w}^T \Omega \boldsymbol{w} \tag{4.29}$$

$$\alpha_{\boldsymbol{w}} = \frac{\sigma_{\boldsymbol{w}} \sigma_{\boldsymbol{w}}^{-2} H^T \boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}^T (\overline{\Omega} - H \sigma_{\boldsymbol{w}}^{-2} H^T) \boldsymbol{\alpha}}}, \tag{4.30}$$

(notice we have set the new notation $\sigma^2_{\boldsymbol{w}}$ for $\Omega_{\boldsymbol{w}}$).

Clearly the portfolio mean is

$$\mathbb{E}(R_{\boldsymbol{w}}) = \mu_{\boldsymbol{w}} + \sqrt{\frac{2}{\pi}} (\sigma_{\boldsymbol{w}} \delta_{\boldsymbol{w}}) \tag{4.31}$$

where $\delta_{\boldsymbol{w}} = \frac{\alpha_{\boldsymbol{w}}}{\sqrt{1+\alpha^2_{\boldsymbol{w}}}}$. At the same time the portfolio variance is

$$\mathrm{Var}(R_{\boldsymbol{w}}) = \sigma^2_{\boldsymbol{w}} - \frac{2}{\pi} \delta^2_{\boldsymbol{w}} \sigma^2_{\boldsymbol{w}} \tag{4.32}$$

By using (4.24)we reach an alternative way of writing it

$$\mathrm{Var}(R_{\boldsymbol{w}}) = \boldsymbol{w}^T V \boldsymbol{w} = \boldsymbol{w}^T \Omega \boldsymbol{w} - \frac{2}{\pi} \boldsymbol{w}^T [(\omega \boldsymbol{\delta})(\omega \boldsymbol{\delta})^T] \boldsymbol{w} = \sigma^2_{\boldsymbol{w}} - \frac{2}{\pi} (\boldsymbol{w}^T (\omega \boldsymbol{\delta}))^2 \tag{4.33}$$

By comparing the two formulas we deduce

$$\delta_{\boldsymbol{w}} \sigma_{\boldsymbol{w}} = \boldsymbol{w}^T (\omega \boldsymbol{\delta}) \tag{4.34}$$

Lastly, the variance can also be expressed by means of the matrix W, which has been previously introduced. In this case

$$\mathrm{Var}(R_{\boldsymbol{w}}) = \boldsymbol{w}^T W \boldsymbol{w} + (1 - \frac{2}{\pi})(\boldsymbol{w}^T (\omega \boldsymbol{\delta}))^2. \tag{4.35}$$

By defining the three symbols

$$v^2_{\boldsymbol{w}} := \mathrm{Var}(R_{\boldsymbol{w}}), \qquad s^2_{\boldsymbol{w}} := \boldsymbol{w}^T W \boldsymbol{w}, \qquad b_{\boldsymbol{w}} := \boldsymbol{w}^T (\omega \boldsymbol{\delta})$$

[1] In this section $\boldsymbol{w}$ replaces A appearing in (1.32)

one can immediately verify that the following relationships hold

$$v_{\boldsymbol{w}}^2 = s_{\boldsymbol{w}}^2 + (1 - \frac{2}{\pi})b_{\boldsymbol{w}}^2 \tag{4.36}$$

$$s_{\boldsymbol{w}}^2 = \sigma_{\boldsymbol{w}}^2 - b_{\boldsymbol{w}}^2 \tag{4.37}$$

$$\sigma_{\boldsymbol{w}}^2 = v_{\boldsymbol{w}}^2 + \frac{2}{\pi}b_{\boldsymbol{w}}^2 \tag{4.38}$$

$$b_{\boldsymbol{w}} = \sigma_{\boldsymbol{w}}\delta_{\boldsymbol{w}} \tag{4.39}$$

For obvious reasons it is natural to call $s_{\boldsymbol{w}}^2$ the *spherical* component of the portfolio variance and $b_{\boldsymbol{w}}^2$ the *non spherical* component.

Coming back to the investment process, we recall that the investor's utility is a function of the total wealth and therefore the investment decision is based on maximizing the following expected utility:

$$\mathbb{E}(u(R_{\boldsymbol{w}})) = \int u(r)f_{R_{\boldsymbol{w}}}(r)dr. \tag{4.40}$$

We have:

Proposition 4.3.1. *The density of $R_{\boldsymbol{w}}$ is a function only of $\mu_{\boldsymbol{w}}, \sigma_{\boldsymbol{w}}^2, b_{\boldsymbol{w}}$.*

Proof. Immediate from previous considerations. □

By consequence, defining

$$\Lambda := \mathbb{E}(u(R_{\boldsymbol{w}})),$$

the previous result states that

$$\Lambda = \lambda(\mu_{\boldsymbol{w}}, \sigma_{\boldsymbol{w}}^2, b_{\boldsymbol{w}}) \tag{4.41}$$

4.3.1 Portfolio Selection with increasing and concave utilities functions

In this section we consider an utility function $u(\cdot) \in \mathcal{U}_2$. The following result holds:

Theorem 4.3.1. *Suppose $\mu_{\boldsymbol{w}} = L, b_{\boldsymbol{w}} = B$, then the expected utility $\mathbb{E}(u(R_{\boldsymbol{w}}))$ is a non-increasing function of the spherical part of the variance $s_{\boldsymbol{w}}^2$.*

Proof. By Proposition 2.1.6 we know that expected utility is a non increasing function of $\sigma_{\boldsymbol{w}}^2$. Furthermore we have:

$$\sigma_{\boldsymbol{w}}^2 = b_{\boldsymbol{w}}^2 + s_{\boldsymbol{w}}^2 \tag{4.42}$$

This expression implies that expected utility, for fixed $b_{\boldsymbol{w}}$, is non increasing in $s_{\boldsymbol{w}}$. □

The following Corollary states that a risk-averse investor who aims to maximize the expected utility of his terminal wealth can reframe the problem in terms of an equivalent one, based on a quadratic program.

Corollary 4.3.1. *Let $u(\cdot) \in \mathcal{U}_2$. Then there exists a pair (L, B) for which the following quadratic problem*

$$\begin{aligned} Min_{\boldsymbol{w}} \quad & s^2_{\boldsymbol{w}} \\ \textit{with the constraints:} \quad & \mu_{\boldsymbol{w}} = L \\ & b_{\boldsymbol{w}} = B \\ & \mathbf{1}^T \boldsymbol{w} = 1 \end{aligned} \tag{4.43}$$

is solved by the same portfolio that maximizes the expected utility.

Proof. [1] Denote by $\bar{\boldsymbol{w}}$ the portfolio which gives the highest expected utility, that is

$$\mathbb{E}(u(R_{\bar{\boldsymbol{w}}})) \geq \mathbb{E}(u(R_{\boldsymbol{w}}))$$

for any other portfolio $\boldsymbol{w}$. Define $B = \bar{\boldsymbol{w}}^T(\omega\boldsymbol{\delta})$ and $L = \bar{\boldsymbol{w}}^T\boldsymbol{\mu}$.; for fixed $b_{\boldsymbol{w}}$ the expected utility is non increasing in $s^2_{\boldsymbol{w}}$, then for $b_{\boldsymbol{w}} = B$ and $\mu_{\boldsymbol{w}} = L$

$$s_{\bar{\boldsymbol{w}}} \leq s_{\boldsymbol{w}}$$

therefore $\bar{\boldsymbol{w}}$ solves problem (4.43). □

Remark:

i) Considering that:

$$\mathbb{E}(R_{\boldsymbol{w}}) = \mu_{\boldsymbol{w}} + \sqrt{\frac{2}{\pi}} b_{\boldsymbol{w}}$$

this problem implies that by fixing the pair (L, B) the expected portfolio return is fixed to the value $\mathbb{E}(R_{\boldsymbol{w}}) = L + \sqrt{\frac{2}{\pi}}B \equiv E$. Viceversa, as in Simaan's paper, one can take the point of view of fixing $\mathbb{E}(R_{\boldsymbol{w}}) = E$ and $b_{\boldsymbol{w}} = B$. This last choice looks clearly more natural from a mean-variance point of view, however in the present framework it is tantamount to fixing $L = E - \sqrt{\frac{2}{\pi}}B$ as location parameter for the portfolio return distribution. In other terms, an investor who explicitly wishes to consider the pair (E, B) then he is implicitly considering the pair (L, B), the other way round being true

[1] adapted from Simaan [35]

as well.
ii) Minimizing $s_{\boldsymbol{w}}$ with fixed $b_{\boldsymbol{w}}$ means maximizing $\alpha_{\boldsymbol{w}}$ because

$$\delta_{\boldsymbol{w}} = \frac{b_{\boldsymbol{w}}}{\sqrt{b_{\boldsymbol{w}}^2 + s_{\boldsymbol{w}}^2}}$$

and the function $x/\sqrt{1-x^2}$ is increasing for any $x \in (-1, 1)$.
iii) The quadratic shape of the objective function typical of the mean-variance analysis is preserved by this approach.

We call the space generated by the location/variance/*non spherical* component of the variance of portfolio return, $(\mu_{\boldsymbol{w}}, v_{\boldsymbol{w}}^2, b_{\boldsymbol{w}})$, the (**LVS**) space. There is a one-to-one correspondence between points in the (**LVS**) space and points in the (**MVS**) space (mean/variance/*non spherical* component of the variance) discussed by Simaan in [35] : concepts expressed in one space can be easily translated in the other space.
We call *minimum spherical variance* set, the set of all the solutions (portfolios) to the problem (4.43) obtained by varying the pair values (L, B). Below we prove that the *minimum spherical variance* set is represented in the (**LVS**) space by points of an elliptical paraboloid, with the location parameter plotted on the vertical axes.
Furthermore we say that a portfolio belongs to the *efficient* set, or it is an efficient portfolio, if it is a portfolio of minimum spherical variance with the highest location among all minimum spherical variance portfolios having the same spherical variance and skewness [1]. In the market under consideration, where short sales allowed, the *efficient* set corresponds to the upper surface of the paraboloid. Indeed, by Theorem 2.1.4 (iii), we know that for fixed $(\sigma_{\boldsymbol{w}}^2, \alpha_{\boldsymbol{w}})$ the expected utility is increasing in location; however, the same is remains true by fixing $(v_{\boldsymbol{w}}^2, b_{\boldsymbol{w}})$ because of the relationships existing among the two pairs (see the previous section). The following Proposition states that the *minimum spherical variance* set is spanned by three portfolios:

Proposition 4.3.2. *Assume the asset returns* $\boldsymbol{R}$ *are distributed as in Prop.4.2.1. The* minimum spherical variance *set is spanned by the following three funds:*

$$\boldsymbol{a}_1 = \frac{V^{-1}\boldsymbol{\mu}}{\mathbf{1}^T V^{-1}\boldsymbol{\mu}}, \quad \boldsymbol{a}_2 = \frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} \quad \textit{and} \quad \boldsymbol{a}_3 = \frac{V^{-1}(\omega\boldsymbol{\delta})}{\mathbf{1}^T V^{-1}(\omega\boldsymbol{\delta})} \tag{4.44}$$

[1] Simaan, who works in the (**MVS**) space, makes in [35] the following comment "With restriction on short sales and non negativity of constraints is quite difficult to provide a characterization of the efficient set of portfolios...". This is not the case when short sales are allowed, as in this thesis.

Proof. [1] Set $\boldsymbol{b} = (\omega\boldsymbol{\delta})$, the objective function to be minimized is

$$s_{\boldsymbol{w}}^2 = v_{\boldsymbol{w}}^2 - \sigma_{|X|}^2 b_{\boldsymbol{w}}^2 \tag{4.45}$$

The Lagrangian is obtained by the constraints of the problem (4.43):

$$\mathfrak{L} = \frac{1}{2}v_{\boldsymbol{w}}^2 - \frac{1}{2}b_{\boldsymbol{w}}^2\sigma_{|X|}^2 + \delta_1(L - \mu_{\boldsymbol{w}}) + \delta_2(1 - \boldsymbol{w}^T\mathbf{1}) + \delta_3(B - b_{\boldsymbol{w}}) \tag{4.46}$$

The first order conditions are easily derived:

$$\begin{aligned} \frac{\partial \mathfrak{L}}{\partial \boldsymbol{w}} &= V\boldsymbol{w} - \boldsymbol{b}\sigma_{|X|}^2 - \delta_1\boldsymbol{\mu} - \delta_2\mathbf{1} - \delta_3\boldsymbol{b} = 0 \\ &= V\boldsymbol{w} - \delta_1\boldsymbol{\mu} - \delta_2\mathbf{1} - (\delta_3 + \sigma_{|X|}^2)\boldsymbol{b} = 0 \end{aligned} \tag{4.47}$$

Taking the linear combinations of the last equation respectively by $\mathbf{1}^T V^{-1}$, $\boldsymbol{\mu}^T V^{-1}$ and $\boldsymbol{b}^T V^{-1}$ and calling

$$A = \mathbf{1}^T V^{-1}\mathbf{1}, \;\; C = \mathbf{1}^T V^{-1}\boldsymbol{\mu}, \;\; D = \mathbf{1}^T V^{-1}\boldsymbol{b}, \tag{4.48}$$

$$F = \boldsymbol{\mu}^T V^{-1}\boldsymbol{b}, \;\; G = \boldsymbol{\mu}^T V^{-1}\boldsymbol{\mu}, \;\; H = \boldsymbol{b}^T V^{-1}\boldsymbol{b} \tag{4.49}$$

we obtain the following linear system:

$$\begin{aligned} 1 - D\sigma_{|X|}^2 &= \delta_1 C + \delta_2 A + \delta_3 D \\ L - F\sigma_{|X|}^2 &= \delta_1 G + \delta_2 C + \delta_3 F \\ B - H\sigma_{|X|}^2 &= \delta_1 F + \delta_2 D + \delta_3 H \end{aligned}$$

Solving for δ_1, δ_2 and δ_3 and inserting the result into (4.47) we obtain

$$\begin{aligned} \boldsymbol{w}^* &= \delta_1 C\frac{V^{-1}\boldsymbol{\mu}}{\mathbf{1}^T V^{-1}\boldsymbol{\mu}} + \delta_2 A\frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} + (\delta_3 + \sigma_{|X|}^2)D\frac{V^{-1}\boldsymbol{b}}{\mathbf{1}^T V^{-1}\boldsymbol{b}} \\ &= \lambda_1\frac{V^{-1}\boldsymbol{\mu}}{\mathbf{1}^T V^{-1}\boldsymbol{\mu}} + \lambda_2\frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} + \lambda_3\frac{V^{-1}\boldsymbol{b}}{\mathbf{1}^T V^{-1}\boldsymbol{b}} \end{aligned} \tag{4.50}$$

where $\lambda_1 + \lambda_2 + \lambda_3 = 1$. This ends the Proof. □

Remark: Portfolio $\boldsymbol{a}_2$ is the global minimum variance portfolio. It plays a role analogous to that of the risk-less asset for this risky market.

[1] adapted from Simaan [35]

Portfolio $\boldsymbol{a}_3$ is the portfolio whose return maximizes the correlation with $|X|$. To show this let us write the correlation between a generic portfolio return $\boldsymbol{w}^T R$ and $|X|$: Given $\boldsymbol{R} = \boldsymbol{\mu} + \omega\boldsymbol{\delta}|X| + \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z}$ the correlation is:

$$\text{Cor}(\boldsymbol{w}^T\boldsymbol{R}, |X|) = \frac{\sigma^2_{|X|}(\boldsymbol{w}^T\boldsymbol{b})^2 + \text{Cov}(\boldsymbol{w}^T(\omega(Id - \Delta^2)^{1/2}\boldsymbol{Z}), |X|)}{\boldsymbol{w}^T V\boldsymbol{w}} = \frac{\sigma^2_{|X|}(\boldsymbol{w}^T\boldsymbol{b})^2}{\boldsymbol{w}^T V\boldsymbol{w}} \tag{4.51}$$

where in the last step we used $\mathbb{E}(\boldsymbol{w}^T(\omega(Id-\Delta^2)^{1/2}\boldsymbol{Z})||X|) = 0$ that implies $\text{Cov}(\boldsymbol{w}^T(\omega(Id-\Delta^2)^{1/2}\boldsymbol{Z}), |X|) = 0$. Solving for $\boldsymbol{w}$ the first order condition we obtain:

$$\frac{d}{d\boldsymbol{w}}\text{Cor}(\boldsymbol{w}^T\boldsymbol{R}, |X|) = 0 \quad \Rightarrow \quad \boldsymbol{w} = \frac{\boldsymbol{w}^T V\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{b}}V^{-1}\boldsymbol{b} \tag{4.52}$$

Since $\text{Cor}(\boldsymbol{w}^T\boldsymbol{R}, |X|)$ is independent of the scale of $\boldsymbol{w}$ it is clear that $\boldsymbol{a}_3$ is the portfolio with the maximum correlation with the factor $|X|$.

We now enlarge the list of the market primary assets by adding a risk-less asset which offers a fixed return R_f.
At the same time we modify the vector of the asset returns (4.10) in the following way:

$$\tilde{\boldsymbol{R}} = \begin{pmatrix} R_f \\ \boldsymbol{R} \end{pmatrix} = \begin{pmatrix} R_f \\ \boldsymbol{\mu} \end{pmatrix} + \begin{pmatrix} 0 \\ (\omega\boldsymbol{\delta}) \end{pmatrix}|X| + \begin{pmatrix} 0 \\ \tilde{\boldsymbol{Z}} \end{pmatrix} \tag{4.53}$$

where $\tilde{\boldsymbol{Z}} = \omega(Id - \Delta^2)^{1/2}Z$. Notice that, so doing, we interpret the value $\boldsymbol{R}_f$ as a location. Indeed, for this asset location and mean do coincide.

It is immediate to prove that now the portfolio problem is equivalent to the following quadratic problem:

$$\min \ s^2_{\boldsymbol{w}} \tag{4.54}$$

with the constraints:

$$\boldsymbol{w}^T\boldsymbol{\mu} + w_0 R_f = L,$$

$$b_{\boldsymbol{w}} = B,$$

$$\boldsymbol{w}^T\boldsymbol{1} + w_0 = 1$$

and we have

Proposition 4.3.3. *Assume that assets returns are described by (4.53). Then the efficient set is spanned by the risk free asset portfolio* $\boldsymbol{w}_f = (1, 0, \ldots, 0)^T$,

$$\boldsymbol{w}_3 = (0, \boldsymbol{a}_3)^T$$

and:

$$\boldsymbol{w}_t = (0, \boldsymbol{a})^T$$

where

$$\boldsymbol{a} = \frac{V^{-1}(\boldsymbol{\mu} - R_f \mathbf{1})}{\mathbf{1}^T V^{-1}(\boldsymbol{\mu} - R_f \mathbf{1})} \tag{4.55}$$

Proof. [1] Writing down the Lagrangian and the first order conditions for the portfolio problem (4.54), after some easy computations it is obtained:

$$\boldsymbol{w}^* = \lambda_1 \frac{V^{-1}(\boldsymbol{\mu} - R_f \mathbf{1})}{\mathbf{1}^T V^{-1}(\boldsymbol{\mu} - R_f \mathbf{1})} + \lambda_2 \frac{V^{-1}\boldsymbol{b}}{\mathbf{1}^T V^{-1}\boldsymbol{b}} \tag{4.56}$$

and $w_0 = 1 - \lambda_1 - \lambda_2$. □

4.4 Location Variance Skewness efficient frontier without a risk-less asset

The following result characterizes the geometry of the *efficient* set in the $(\mu_{\boldsymbol{w}}, v^2_{\boldsymbol{w}}, b_{\boldsymbol{w}})-$space, the (**LVS**) space, in absence of riskless asset.

Denote by $B_i = (\omega\boldsymbol{\delta})^T \boldsymbol{a}_i$ and $L_i = \boldsymbol{\mu}^T \boldsymbol{a}_i$ where $\boldsymbol{a}_i$ are the portfolios given by (4.44).

We have the following parameters correspondence with the notation used in (4.48), (4.49):

$$E_1 = \frac{G}{C}\ ;\ E_2 = \frac{C}{A}\ ;\ E_3 = \frac{F}{D}\ ;\ B_1 = \frac{F}{C}\ ;\ B_2 = \frac{D}{A}\ ;\ B_3 = \frac{H}{D}$$

Proposition 4.4.1. *The* minimum spherical variance *set in the* ***LVS*** *space is given by:*

$$v^2_{\boldsymbol{w}} = v^2_2 + v^2_{h_3}\left(\frac{b_{\boldsymbol{w}} - B_2}{B_3 - B_2}\right)^2 + v^2_{h_1} c_1^2 \tag{4.57}$$

where $v_2^2 = \boldsymbol{a}_2^T V \boldsymbol{a}_2$, $v^2_{h_i} = \boldsymbol{h}_i^T V \boldsymbol{h}_i$,

$$c_1 = \frac{(\mu_{\boldsymbol{w}} - E_2)/(E_3 - E_2) - (b_{\boldsymbol{w}} - B_2)/(B_3 - B_2)}{(E_1 - E_2)/(E_3 - E_2) - (B_1 - B_2)/(B_3 - B_2)}$$

[1] adapted from Simaan [35]

and portfolios $\boldsymbol{h}_i$, $i = 1, 3$, *are given by*

$$\boldsymbol{h}_1 = (\boldsymbol{a}_1 - \boldsymbol{a}_2) - \frac{B_1 - B_2}{B_3 - B_2}(\boldsymbol{a}_3 - \boldsymbol{a}_2) \;\; ; \;\; \boldsymbol{h}_3 = \boldsymbol{a}_3 - \boldsymbol{a}_2$$

Proof. [1] By Corollary (4.3.2) any efficient portfolio is a combination of the $\boldsymbol{a}_i$ portfolios; it is therefore of the form:

$$\boldsymbol{w} = c_1\boldsymbol{a}_1 + (1 - c_1 - c_3)\boldsymbol{a}_2 + c_3\boldsymbol{a}_3 = \boldsymbol{a}_2 + c_1(\boldsymbol{a}_1 - \boldsymbol{a}_2) + c_3(\boldsymbol{a}_3 - \boldsymbol{a}_2)$$

Multiplying for $(\omega\boldsymbol{\delta})^T$ we obtain:

$$c_3 = \frac{b_{\boldsymbol{w}} - B_2}{B_3 - B_2} - c_1\frac{B_1 - B_2}{B_3 - B_2}$$

and inserting c_3 in the expression for $\boldsymbol{w}$:

$$\boldsymbol{w} = \boldsymbol{a}_2 + \boldsymbol{h}_1 c_1 + \boldsymbol{h}_3\frac{b_{\boldsymbol{w}} - B_2}{B_3 - B_2} \tag{4.58}$$

Multiplying now for $\boldsymbol{\mu}^T$ and solving for c_1:

$$c_1 = \frac{(\mu_{\boldsymbol{w}} - E_2)/(E_3 - E_2) - (b_{\boldsymbol{w}} - B_2)/(B_3 - B_2)}{(E_1 - E_2)/(E_3 - E_2) - (B_1 - B_2)/(B_3 - B_2)}$$

To prove (4.57) is now sufficient to show that:

$$\boldsymbol{a}_2^T V \boldsymbol{h}_3 = \boldsymbol{a}_2^T V \boldsymbol{h}_1 = \boldsymbol{h}_3^T V \boldsymbol{h}_1 = 0 \tag{4.59}$$

Since $\boldsymbol{a}_2 \propto \mathbf{1}^T V^{-1}$ it is immediate that being $\boldsymbol{a}_i$ portfolios it holds:

$$\mathbf{1}^T V^{-1} V(\boldsymbol{a}_3 - \boldsymbol{a}_2) = \mathbf{1}^T V^{-1} V(\boldsymbol{a}_1 - \boldsymbol{a}_2) = 0$$

The third equality in (4.59) is valid because:

$$\boldsymbol{a}_3^T V \boldsymbol{a}_1 - \frac{b_{\boldsymbol{w}} - B_1}{B_3 - B_1}\boldsymbol{a}_3^T V(\boldsymbol{a}_3 - \boldsymbol{a}_2)$$
$$\propto \left[(\omega\boldsymbol{\delta})^T(\boldsymbol{a}_1 - \boldsymbol{a}_2) - \frac{B_1 - B_2}{B_3 - B_2}(\omega\boldsymbol{\delta})^T(\boldsymbol{a}_3 - \boldsymbol{a}_2)\right] = 0$$

□

Remark: If the location parameter is represented on the vertical axes then equation (4.57) describes an elliptical paraboloid in the **LVS** space. The *efficient* set is the upper surface of this paraboloid.

[1] adapted from Simaan [35]

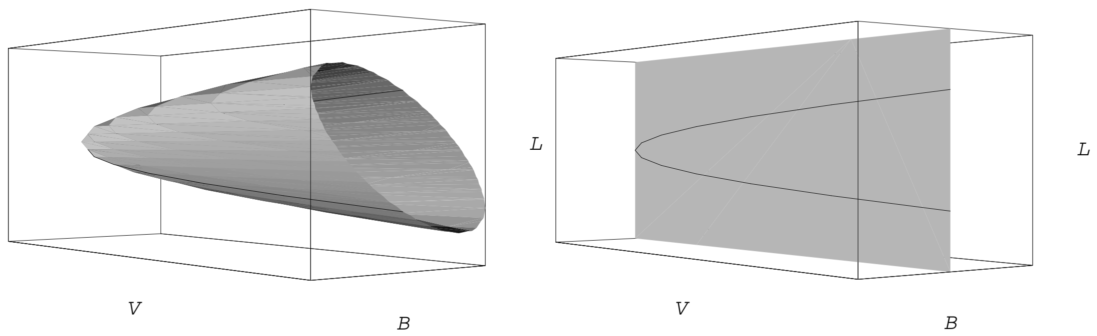


Figure 4.1: Plot of the *minimum spherical variance* set in the $(L, \mathcal{V}, \mathrm{B})$ space and intersection curve with the generic plane $B = k$

4.5 CAPM in three moment space

In this Section, following Simaan [35], we prove an *exact* equilibrium result which takes into account skewness and extends the classical CAPM.
The result we obtain is different from that one of Kraus and Litzenberger [20] and the more recent one of Adcock [1]. Their CAPM is based on a Taylor expansion of the utility function, on the contrary we do not need such an expansion: we completely base our analysis on the three parameters driving the expected utility values.
To derive the pricing model we keep all the assumptions listed for the classical CAPM (see the corresponding section) and add the further requirement that a Pareto optimal market equilibrium exists. This assumption is needed to guarantee the efficiency of the market portfolio, defined as in Section 3.3(see Ingersoll in [18] pag. 194-195 for a proof of this fact).
Under this set of conditions we are able to derive a CAPM equation when assets returns are assumed to be skew normally distributed.

4.5.1 Three moment CAPM

We recall here equality (4.47):

$$V\boldsymbol{w} - \delta_1\boldsymbol{\mu} - \delta_2\mathbf{1} - (\delta_3 + \sigma^2_{|X|})(\omega\boldsymbol{\delta}) = 0 \tag{4.60}$$

where $B = \boldsymbol{w}^T(\omega\boldsymbol{\delta})$ and the three separating funds are:

$$\boldsymbol{w}^* = \lambda_1\frac{V^{-1}\boldsymbol{\mu}}{\mathbf{1}^T V^{-1}\boldsymbol{\mu}} + \lambda_2\frac{V^{-1}\mathbf{1}}{\mathbf{1}^T V^{-1}\mathbf{1}} + \lambda_3\frac{V^{-1}(\omega\boldsymbol{\delta})}{\mathbf{1}^T V^{-1}(\omega\boldsymbol{\delta})} \tag{4.61}$$

Let us call c^k the amount of money invested by k-th investor in the market and by

$$\tilde{\lambda}_1^k = \lambda_1 \frac{c^k}{\mathbf{1}^T V^{-1} \boldsymbol{\mu}} \quad ; \quad \tilde{\lambda}_2^k = \lambda_2 \frac{c^k}{\mathbf{1}^T V^{-1} \mathbf{1}} \quad ; \quad \tilde{\lambda}_3^k = \lambda_3 \frac{c^k}{\mathbf{1}^T V^{-1} (\omega \boldsymbol{\delta})}$$

Denote by $\boldsymbol{w}^k$ the minimum spherical variance portfolio which is held by the k-th investor , and with $\tilde{w}_i^k$ the money invested by the k-th investor in the i-th Firm, namely $\tilde{w}_i^k = w_i^k c^k$. Considering the equality (4.61) for the k-th investor and multiplying it for V we obtain:

$$\sum_{j=1}^{M} V_{ij} \tilde{w}_j^k \quad = \quad \tilde{\lambda}_1^k \tilde{\mu}_i + \tilde{\lambda}_2^k + \tilde{\lambda}_3^k (\omega \boldsymbol{\delta})_i \tag{4.62}$$

valid for $i = 1, \ldots, n$ and $k = 1, \ldots, M$. The market portfolio, defined in Section 3.3, is defined by the vector of weights $\boldsymbol{w}_m$.

The i-th component of this vector is the relative weight of the the i-th Firm in the market and satisfies

$$\sum_{k=1}^{M} \tilde{w}_j^k = (\boldsymbol{w}_m)_j C_{tot} = n_j p_j \quad \text{for } \ j = 1, \ldots, M \tag{4.63}$$

where $C_{tot} = \sum_{k=1}^{M} c^k$ is the total amount of money in the Market and p_i is the price of the security of the i-th Firm.

Equality (4.63) states that the aggregate demand by all investors for the i-th stock must be equal to the value of the i-th Firm in the market which is true in equilibrium, that is when the supply is equal to the demand.

Summing (4.62) over all investors and using (4.63) we obtain:

$$C_{tot} \sum_{j=1}^{M} V_{ij} (\boldsymbol{w}_m)_j \quad = \quad \sum_{k=1}^{M} \tilde{\lambda}_1^k \tilde{\mu}_i + \sum_{k=1}^{M} \tilde{\lambda}_2^k + \sum_{k=1}^{M} \tilde{\lambda}_3^k (\omega \boldsymbol{\delta})_i \tag{4.64}$$

valid for $i = 1, \ldots, n$. The covariance of the i-th asset with the market portfolio is given by:

$$\text{Cov}(R_i, \boldsymbol{R}^T \boldsymbol{w}_m) = \sum_{j=1}^{n} V_{ij} (\boldsymbol{w}_m)_j =: v_{i,m} \tag{4.65}$$

Denoting with $\bar{\lambda}_i = \sum_{k=1}^{M} \tilde{\lambda}_i^k / C_{tot}$ we can rewrite (4.64) in the following form:

$$v_{i,m} \quad = \quad \bar{\lambda}_1 \tilde{\mu}_i + \bar{\lambda}_2 + \bar{\lambda}_3 (\omega \boldsymbol{\delta})_i \tag{4.66}$$

Let $\mu_m = \boldsymbol{\mu}^T \boldsymbol{w}_m$ and $B_m = (\omega\boldsymbol{\delta})^T \boldsymbol{w}_m$ then:

$$\sigma_m^2 = \bar{\lambda}_1 \mu_m + \bar{\lambda}_2 + \bar{\lambda}_3 B_m \tag{4.67}$$

Furthermore given a portfolio $\boldsymbol{w}_a$ we denote by

$$\beta_a := \sum_{i=1}^{n} w_a^i v_{i,m} / \sigma_m^2$$

and with

$$\gamma_a := \boldsymbol{w}_a^T (\omega\boldsymbol{\delta})_i / B_m.$$

The following two portfolios exist:

1)a portfolio $\boldsymbol{w}_0$ such that $\beta_0 = 0$ and $\gamma_0 = 0$

2)a portfolio $\boldsymbol{w}_p$ such that $\beta_p = 0$ and $\gamma_p = 1$

Since $1 = \gamma_p := \boldsymbol{w}_p^T (\omega\boldsymbol{\delta})_i / B_m$ then $(\omega\boldsymbol{\delta})^T \boldsymbol{w}_p = (\omega\boldsymbol{\delta})^T \boldsymbol{w}_m$. Let $\boldsymbol{\mu}^T \boldsymbol{w}_0 = \mu_0$ and $\boldsymbol{\mu}^T \boldsymbol{w}_p = \mu_p$, multiplying (4.65) for $(\boldsymbol{w}_0)_i$ and for $(\boldsymbol{w}_p)_i$ respectively and adding over i we obtain equalities:

$$\bar{\lambda}_1 E_0 + \bar{\lambda}_2 + \bar{\lambda}_2 = 0 \tag{4.68}$$

$$\bar{\lambda}_1 E_p + \bar{\lambda}_2 + \bar{\lambda}_2 + \bar{\lambda}_3 B_m = 0 \tag{4.69}$$

Solving for $\bar{\lambda}_i$ equations (4.67),(4.68) and (4.69) and inserting the results into (4.65) we obtain the following pricing model:

$$\mu_i^e = \mu_0 + \beta_i [\mu_m - E_p] + \gamma_i [\mu_p - \mu_0] \tag{4.70}$$

where β_i and γ_i are the β and γ of the portfolio having all zeros except than 1 in the $i-$th component. If the presence of a risk-less asset is taken into account then R_f replaces μ_0 in the previous equation, and we obtain the following result:

$$\mu_i^e = R_f + \beta_i [\mu_m - R_f] + (\gamma_i - \beta_i)[\mu_p - R_f] \tag{4.71}$$

4.6 Three funds Ross Separation Theorem

Definition 4.6.1. *The distribution of $\boldsymbol{R}$ has the (strong) 3-funds separation property if there exist three portfolios $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ such that for any portfolio $\boldsymbol{\beta}$ there is a portfolio $\boldsymbol{\alpha}$ given by a linear combination of $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_3$ for which it holds*

$$\boldsymbol{\alpha}^T \boldsymbol{R} \succeq_2 \boldsymbol{\beta}^T \boldsymbol{R}.$$

The separating properties of the *minimum spherical variance* set, previously showed, imply that the 3-funds separation obtains for the skew-normal distribution. Nonetheless it is also interesting to prove directly that the skew-normal has the 3-funds separating properties of Ross.

Ross in [32] gives the following sufficient condition for three funds separability with risk-less asset:

Theorem 4.6.1. *(3-funds separation with a risk-less asset) The distribution of $\boldsymbol{R}$ has the (strong) 3-fund separation property iff:*
there exist two univariate r.v. Y and Q, a multivariate r.v. $\tilde{\boldsymbol{\epsilon}}$ (the residuals), two (deterministic) vectors $\boldsymbol{b}$, $\boldsymbol{c}$ and three portfolios $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ such that:
(a) each component of $\boldsymbol{R}$ can be written as

$$R_i = R_f + b_i Y + c_i Q + \tilde{\epsilon}_i \text{ for } i = 1, \ldots, n+1$$

(b) $\mathbb{E}[\tilde{\epsilon}_i | Y, Q] = 0 \;\; \forall i$
(c) $\boldsymbol{\alpha}_i^T \tilde{\boldsymbol{\epsilon}} = 0$ *for* $i = 1, 2, 3$

Proposition 4.6.1. *If $\boldsymbol{R}$ is skew-normal then it satisfies conditions of Theorem 4.6.1*

Proof. The expression of the return vector $\tilde{\boldsymbol{R}} \in \boldsymbol{R}^{n+1}$ in presence of the risk-less asset is given by formula (4.53) that can be rewritten in such a way $\mathbb{E}(Y) = 0$ (i.e. $Y = |X| - \sqrt{\frac{2}{\pi}}$):

$$\tilde{\boldsymbol{R}} = \begin{pmatrix} R_f \\ \boldsymbol{R} \end{pmatrix} = \begin{pmatrix} R_f \\ \tilde{\boldsymbol{\mu}} \end{pmatrix} + \begin{pmatrix} 0 \\ (\omega\boldsymbol{\delta}) \end{pmatrix} Y + \begin{pmatrix} 0 \\ \boldsymbol{Z} \end{pmatrix} \tag{4.72}$$

where $\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu} + \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta})$, $\boldsymbol{Z} \sim N_n(\boldsymbol{0}, W)$, and $W = \omega(Id - \Delta^2)^{1/2}\Psi(Id - \Delta^2)^{1/2}\omega$. The variable $\tilde{\boldsymbol{\mu}} + \boldsymbol{Z}$ being normal satisfies a two fund separability Theorem, and can be decomposed in the following way (see Appendix 2):

$$\tilde{\boldsymbol{\mu}} + \boldsymbol{Z} = \boldsymbol{R}_f + \boldsymbol{\xi} Q + \boldsymbol{\epsilon}$$

where $\boldsymbol{\xi} = \boldsymbol{\mu} - \boldsymbol{R}_f$ and:

$$\begin{aligned} \boldsymbol{R}_f &= R_f \boldsymbol{1} \\ Q &= (\boldsymbol{\xi}^T V^{-1} \boldsymbol{\xi})^{-1} \boldsymbol{\xi}^T V^{-1} (\boldsymbol{Z} - \boldsymbol{R}_f) \\ \epsilon_i &= Z_i - \xi_i Q \end{aligned}$$

Furthermore the portfolio:

$$\boldsymbol{\alpha}_1 = \frac{V^{-1}\boldsymbol{\xi}}{\mathbf{1}^T V^{-1}\boldsymbol{\xi}}$$

is orthogonal to the residuals vector $\boldsymbol{\epsilon}$, i.e. $\boldsymbol{\alpha}_1^T\boldsymbol{\epsilon} = 0$. Inserting this expression of $\tilde{\boldsymbol{\mu}} + \boldsymbol{Z}$ expression into (4.72) we obtain:

$$\tilde{\boldsymbol{R}} = \begin{pmatrix} R_f \mathbf{1} \\ \boldsymbol{R} \end{pmatrix} = \begin{pmatrix} R_f \\ \boldsymbol{R}_f \end{pmatrix} + \begin{pmatrix} 0 \\ (\omega\boldsymbol{\delta}) \end{pmatrix} Y + \begin{pmatrix} 0 \\ \boldsymbol{\xi} \end{pmatrix} Q + \begin{pmatrix} 0 \\ \boldsymbol{\epsilon} \end{pmatrix} \tag{4.73}$$

Let $\boldsymbol{\alpha}_i$ for $i = 1, 2, 3$ the three following portfolios:

$$\tilde{\boldsymbol{\alpha}}_1 = (1, 0, \ldots, 0)^T \quad ; \quad \tilde{\boldsymbol{\alpha}}_2 = (0, \boldsymbol{\alpha}_1)^T \quad ; \quad \tilde{\boldsymbol{\alpha}}_3 = (0, \boldsymbol{\alpha}_3)^T \tag{4.74}$$

where

$$\boldsymbol{\alpha}_3 = \frac{V^{-1}(\omega\boldsymbol{\delta})}{\mathbf{1}^T V^{-1}(\omega\boldsymbol{\delta})}$$

and let:

$$\boldsymbol{b} = (0, (\omega\boldsymbol{\delta}))^T \quad ; \quad \boldsymbol{c} = (0, \boldsymbol{\xi})^T. \tag{4.75}$$

In addiction we define the vector $\tilde{\boldsymbol{\epsilon}}$ in the following way:

$$\tilde{\epsilon}_0 = 0 \quad ; \quad \tilde{\epsilon}_i = \epsilon_i \; i = 1, \ldots, n-2 \tag{4.76}$$

$$\tilde{\epsilon}_{n-1} = \epsilon_{n-1} + B \quad ; \quad \tilde{\epsilon}_n = \epsilon_n + A \tag{4.77}$$

where:

$$A = \frac{(\boldsymbol{\xi}^T V^{-1})_{n-1}}{(\boldsymbol{\xi}^T V^{-1})_n} \frac{(\omega\boldsymbol{\delta})^T V^{-1}\boldsymbol{\epsilon}}{((\omega\boldsymbol{\delta})^T V^{-1})_{n-1} - \frac{(\boldsymbol{\xi}^T V^{-1})_{n-1}}{(\boldsymbol{\xi}^T V^{-1})_n}} \tag{4.78}$$

$$B = -\frac{(\omega\boldsymbol{\delta})^T V^{-1}\boldsymbol{\epsilon}}{((\omega\boldsymbol{\delta})^T V^{-1})_{n-1} - \frac{(\boldsymbol{\xi}^T V^{-1})_{n-1}}{(\boldsymbol{\xi}^T V^{-1})_n}} \tag{4.79}$$

the reason of this definition will be clear below. The previous definition completes the set $(Y, Q, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \tilde{\boldsymbol{\epsilon}})$ so that:

$$\tilde{\boldsymbol{R}} = \boldsymbol{R}_f + \boldsymbol{b}Y + \boldsymbol{c}Q + \tilde{\boldsymbol{\epsilon}} \tag{4.80}$$

Now we have to prove (b) and (c).
A little bit of algebra is needed to show that:

$$\begin{aligned} \tilde{\boldsymbol{\alpha}}_3^T\tilde{\boldsymbol{\epsilon}} &= \frac{(\omega\boldsymbol{\delta})^T V^{-1}}{\mathbf{1}^T V^{-1}(\omega\boldsymbol{\delta})}\tilde{\boldsymbol{\epsilon}} = 0 \\ \tilde{\boldsymbol{\alpha}}_2^T\tilde{\boldsymbol{\epsilon}} &= \frac{\boldsymbol{\xi}^T V^{-1}}{\mathbf{1}^T V^{-1}\boldsymbol{\xi}}\tilde{\boldsymbol{\epsilon}} = 0 \end{aligned} \tag{4.81}$$

and furthermore is obvious:

$$\tilde{\boldsymbol{\alpha}}_1^T \tilde{\boldsymbol{\epsilon}} = 0 \tag{4.82}$$

so (c) is proved.
As far as condition (b) is concerned, it is enough to show that $\mathbb{E}(\epsilon_i|Y, Q) = 0$. The two fund separability result (4.73) implies $\mathbb{E}(\epsilon_i|Q) = 0$. Furthermore we have $\mathbb{E}(\epsilon_i|Z) = \mathbb{E}(\epsilon_i) = 0$ because ϵ_i and Z are independent random variables for any i. □

5

Black Litterman model

5.1 Bayesian allocation

Two well known problems of the mean-variance portfolio optimization are the high sensitivity to the input parameters and the lack of diversification in the optimal portfolio weights.

In this Section we face the problem of reducing the sensitivity of optimal portfolios to changes in the values of the inputs using a Bayesian allocation approach. In particular we will show that, through the Bayesian approach, it is possible to modify the utility maximization process in order to select a set of more stable optimal portfolios.

As in the previous sections utility is a function of the investor terminal wealth which is represented by the final period return of the portfolio, assuming a unitary initial investment. Once again, given a portfolio $\boldsymbol{w}$, the portfolio return will be given by

$$R_{\boldsymbol{w}} = \boldsymbol{w}^T \boldsymbol{R},$$

where $\boldsymbol{R}$ represents the vector of assets returns. We know that the classical approach to the expected utility maximization, described in Chapter 3 and 4 for normal and skew-normal market returns respectively, prescribes that each risk-averse investor solves the problem:

$$\text{Max}_{\boldsymbol{w}\in C}\mathbb{E}(u(R_{\boldsymbol{w}})) = \text{Max}_{\boldsymbol{w}\in C}\int u(r) f_{R_{\boldsymbol{w}}}(r)dr \tag{5.1}$$

where $f_{R_w}(r)$ is the density of $R_{\boldsymbol{w}}$, $u \in \mathcal{U}_2$ is the investor's utility function and C is a set of constraints.

In general the assets returns distribution depends on a vector of parameters $\boldsymbol{\theta}$ and therefore the portfolio return density depends on the same vector as well. In this case $f_{R_w}(r|\boldsymbol{\theta})$ is a more appropriate notation for the return and the associated expected utility is better written in the form

$$\mathbb{E}(u(R_{\boldsymbol{w}})|\boldsymbol{\theta}) = \int u(r) f_{R_w}(r|\boldsymbol{\theta}) dr \tag{5.2}$$

To give a couple of examples consider first a market in which returns are assumed to be normal distributed . Therefore we have:

$$\boldsymbol{\theta} = (\boldsymbol{m}, V)$$

and the investor's decision depends on solving

$$\text{Max}_{\boldsymbol{w}\in C}\mathbb{E}(u(R_{\boldsymbol{w}}|)\boldsymbol{\theta}) = \text{Max}_{\boldsymbol{w}\in C}\lambda(\boldsymbol{w}^T\boldsymbol{m}, \boldsymbol{w}^T V\boldsymbol{w}). \tag{5.3}$$

In this special case the maximum can be obtained by solving the quadratic program (3.7).
Similarly, for a market with skew normal returns, we have

$$\boldsymbol{\theta} = (\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$$

and the problem is

$$\text{Max}_{\boldsymbol{w}\in C}\mathbb{E}(u(R_{\boldsymbol{w}})|\boldsymbol{\theta}) = \text{Max}_{\boldsymbol{w}\in C}\lambda(\boldsymbol{w}^T\boldsymbol{\mu}, \boldsymbol{w}^T\Omega\boldsymbol{w}, \boldsymbol{w}^T(\omega\boldsymbol{\delta})) \tag{5.4}$$

The solution gives the optimal portfolio.
In both cases small changes in $\boldsymbol{\theta}$ can lead to an optimal allocation completely different from the initial one, a behaviour that many investors and portfolio managers greatly dislike.
In the Bayesian allocation approach the vector $\boldsymbol{\theta}$ is considered to be the realization of a random variable Θ which is modeled by means of a prior distribution $f^{pr}_{\Theta}(\boldsymbol{\theta})$.
The density of the posterior distribution of the parameters given the observed vector of returns $\boldsymbol{R} = \boldsymbol{r}$ is then obtained by the Bayes rule:

$$f^{po}_{\Theta|\boldsymbol{R}}(\boldsymbol{\theta}|\boldsymbol{r}) \propto f^{pr}_{\Theta}(\boldsymbol{\theta}) f_{\boldsymbol{R}|\Theta}(\boldsymbol{r}|\boldsymbol{\theta}) \tag{5.5}$$

In the same way, given observed portfolio returns $R_{\boldsymbol{w}} = r'$ the posterior density of the parameters is

$$f^{po}_{\Theta|R_{\boldsymbol{w}}}(\boldsymbol{\theta}|r') \propto f^{pr}_{\Theta}(\boldsymbol{\theta}) f_{R_{\boldsymbol{w}}|\Theta}(r'|\boldsymbol{\theta}) \tag{5.6}$$

where $f_{R_{\boldsymbol{w}}|\Theta}(r'|\boldsymbol{\theta}) \equiv f_{R_{\boldsymbol{w}}}(r'|\boldsymbol{\theta})$.

The key idea of the Bayesian allocation is then the following: to limit the sensitivity of the allocation process to inputs changes we can average over all values $\boldsymbol{\theta}$ by using $f^{po}_{\Theta|R_{\boldsymbol{w}}}$.

In this framework the utility maximization problem, given the observation $R_{\boldsymbol{w}} = r'$, becomes:

$$\text{Max}_{\boldsymbol{w}\in C} \int \mathbb{E}(u(R_{\boldsymbol{w}})|\boldsymbol{\theta}) f^{po}_{\Theta|R_{\boldsymbol{w}}}(\boldsymbol{\theta}|r') d\boldsymbol{\theta} \tag{5.7}$$

By inserting (5.2) into (5.7) and interchanging the order of the double integration it is easily seen that the previous maximization problem can be rewritten in following form

$$\text{Max}_{\boldsymbol{w}\in C} \int u(r) f^{pre}(r|r') dr \tag{5.8}$$

where the density $f^{pre}(r|r')$ is given by

$$f^{pre}(r|r') = \int f_{R_{\boldsymbol{w}}|\Theta}(r|\boldsymbol{\theta}) f^{po}_{\Theta|R_{\boldsymbol{w}}}(\boldsymbol{\theta}|r') d\boldsymbol{\theta} \tag{5.9}$$

and takes the name of predictive posterior density. This is the Bayesian allocation approach to the utility maximization. In this context the utility is maximized using the predictive posterior distribution instead of the density of the portfolio returns.

5.2 Black Litterman model

Black and Litterman in [6] proposed an allocation method based on ideas similar to those of the Bayesian approach but introducing some novelty. The resulting model is known as the Black Litterman model (**BL**).

The main difference between the two approaches is the following one: rather than evaluating the posterior distribution of Θ given the returns $\boldsymbol{R}$, as in the Bayesian allocation (see (5.5)), **BL** evaluates the posterior of Θ on a new vector, $\boldsymbol{V}$, which represents the views of the investors, or of market experts, on the mean values of (future) returns. In other words in this model the investors views and the prior on the parameters are directly blended in order to obtain a posterior distribution on the parameters. So doing Black and Litterman are able to include into the investment process the investors beliefs about future returns of assets or of classes of assets, a fact which helps to obtain more diversified portfolios and bounds the sensitivity of the optimal weights to the inputs.

5.2.1 The Black Litterman framework

Black and Litterman trust the estimate of the covariance of the returns by the historical covariance matrix but do not trust the historical mean as good estimator for the mean of the returns. As a consequence, looking to the estimation problem, they are only "half-bayesian" and put a prior distribution just on the vector of expected returns. More precisely, they consider the vector

$$\boldsymbol{R}' = \boldsymbol{R} - \mathbf{1}R_f \tag{5.10}$$

whose components measure the excess of return of the risky assets over the risk-free asset, and make the following assumption

Assumption 5.2.1. *The vector of returns $\boldsymbol{R}'$, given its mean, is distributed as:*

$$\boldsymbol{R}'|\Theta = \boldsymbol{\theta} \sim N_n(\boldsymbol{\theta}; \hat{V}), \tag{5.11}$$

moreover

$$\Theta \sim N_n(\boldsymbol{\Pi}, \tau\hat{V}) \tag{5.12}$$

where $\hat{V}$ and τ are known parameters ($\hat{V}$ represents the estimated covariance matrix) and $\boldsymbol{\Pi}$ is the vector of implied (excess) returns.

The definition of the vector $\boldsymbol{\Pi}$ which appears in (5.12) is important and goes as follows. Since returns are assumed to be normally distributed all market investors, who are believed to maximize their utility functions, select their portfolios by mean-variance theory. As shown in chapter 3, in presence of a risk-less asset the risky part of the portfolio will be given by (see (3.27)):

$$\tilde{w} = \gamma\hat{V}^{-1}(\boldsymbol{m} - \mathbf{1}R_f) = \gamma\hat{V}^{-1}\mathbb{E}(\boldsymbol{R}'). \tag{5.13}$$

Following an idea of Sharpe, Black and Litterman then look to the equation (5.13) in a reversed way, that is: they assign to $\tilde{w}$ the role of input variable in the equation and consider $\mathbb{E}(\boldsymbol{R}')$ as the value to be determined, a procedure known as "reversed optimization". So doing, by inserting different numerical values $\tilde{w}$ in (5.13) one can get different estimates for $\mathbb{E}(\boldsymbol{R}')$. **BL** estimate $\mathbb{E}(\boldsymbol{R}')$ by using as input $\tilde{w}$ the weights determined by the relative capitalization of each risky asset traded in the concrete market where the investment process takes place. Their choice correspond to the assumption that the market is at equilibrium so that the tangency portfolio is given by the market portfolio. The result of the estimate obtained in this way is $\mathbf{\Pi}$, the vector of implied (excess) returns. Denoting by $\tilde{w}_m$ the previous input weights (which now, differently from chapter 3, sum up to 1), the formula defining $\mathbf{\Pi}$ becomes

$$\tilde{w}_m = \gamma \hat{V}^{-1} \mathbf{\Pi} \tag{5.14}$$

or explicitly

$$\mathbf{\Pi} = \frac{1}{\gamma} \hat{V} \tilde{w}_m \tag{5.15}$$

In order to get $\mathbf{\Pi}$ numerically it remains to asses the value of γ appearing in (5.15). Recalling equations (3.28) and (3.30) we see that $\delta \equiv \frac{1}{\gamma}$[1] can be written in Sharpe's ratio form

$$\delta = \frac{m_{\tilde{w}_m} - R_f}{v^2_{\tilde{w}_m}}. \tag{5.16}$$

R_f being known, to evaluate δ it is sufficient to plug-in (5.16)the estimated values for mean and standard deviation of the market portfolio return. So we are back to the CAPM point of view: at equilibrium the expected return of the i-th asset deviates from the risk-free return by a term proportional to the market risk premium $\tilde{w}_m - R_f$, the proportionality coefficient $\beta_i = v^{-2}_{\tilde{w}_m}(V\tilde{w}_m)_i$ being asset-specific.
Further properties of the vector $\mathbf{\Pi}$ are analyzed by He and Litterman in [17].
By standard computations the model given by (5.11) and (5.12) implies a marginal distribution for $\boldsymbol{R}'$ of normal type:

$$\boldsymbol{R}' \sim N_n(\mathbf{\Pi}; (1+\tau)\hat{V}) \tag{5.17}$$

$$\tag{5.18}$$

The parameter τ is usually chosen in the range $0.02 < \tau < 0.05$, corresponding to a prior belief on expected returns having a small spread around the values $\mathbf{\Pi}$.

[1]The coefficient δ takes the name of parameter of risk aversion and measures the aggregate propencity of the market investors to invest in risky securities

5.2.2 Incorporating views

The model for the returns involved in formulas (5.11) and (5.12) could be used directly for implementing the Bayesian allocation outlined in Section 5.1. However, **BL** aims to incorporate investor views into the investment process. This is accomplished by putting random constraints on the vector $\boldsymbol{m}$ using the views vector $\boldsymbol{V} \in \mathbb{R}^k$, $k \leq n$. The constraints are of the form:

$$P\boldsymbol{M} - \boldsymbol{V} \sim N(0, \Omega_v) \tag{5.19}$$

P is a $(k \times n)$ matrix and Ω_v is $(k \times k)$ invertible diagonal matrix expressing confidence in the views.

The number k represents the total number of views, and it can be different from the total number of assets involved in the constraints. Each view may contain as many assets as the investor wants. This information is codified in the matrix P, called the "pick" matrix. The rows of this matrix represent the views and the columns represent the assets. See the example in the next Section for a realistic form of the matrix P.

The constraints (5.19) can also be written in the following regression form

$$\begin{aligned} \boldsymbol{V} &= P\boldsymbol{M} + \varepsilon \\ \varepsilon &\sim N(0, \Omega_v) \end{aligned} \tag{5.20}$$

from which we get

$$\boldsymbol{V}|\boldsymbol{M} = \boldsymbol{m} \sim N(P\boldsymbol{m}, \Omega_v) \tag{5.21}$$

This last relation can be interpreted as a model for the "observed" views, given expected values of assets. The prior distribution specified by (5.11) and the previous distribution for the views given the expected returns can be combined in order to obtain the following posterior distribution:

$$\boldsymbol{M}|\boldsymbol{V} = \boldsymbol{v} \sim N(\boldsymbol{m}_{BL}, \Sigma_{BL}) \tag{5.22}$$

where:

$$\boldsymbol{m}_{BL} = [(\tau V)^{-1} + P^T \Omega_v^{-1} P]^{-1}[(\tau V)^{-1}\boldsymbol{\Pi} + P^T \Omega_v^{-1}\boldsymbol{v}] \tag{5.23}$$

$$\Sigma_{BL} = [(\tau V)^{-1} + P^T \Omega_v^{-1} P]^{-1} \tag{5.24}$$

These last formulas are classical and well-known in the bayesian approach to parameters estimation of multivariate normal models. We refer also to Meucci [27] for a proof and further discussions. However it is important to notice that for $P = 0$ (no views!) they reproduces the original "equilibrium estimate" for the mean values, that is $\boldsymbol{m}_{BL} = \boldsymbol{\Pi}$. The vector $\boldsymbol{m}_{BL}$ represents the mean of the posterior distribution of expected returns given the views. It includes both information coming from the market and the investor views.

Given the posterior distribution of expected returns it is possible to evaluate the predictive distribution of returns given the views.

5.2.3 Predictive distribution

The evaluation of the predictive distribution is easy because both the posterior (5.22) and the likelihood (5.11) are normally distributed. The predictive posterior, defined by formula (5.8), in this case takes the form:

$$f^{pre}_{\boldsymbol{R}|\boldsymbol{V}}(\boldsymbol{r}|\boldsymbol{v}) = \int f_{\boldsymbol{R}|\boldsymbol{M}}(\boldsymbol{r}|\boldsymbol{m}) f^{po}_{\boldsymbol{M}|\boldsymbol{V}}(\boldsymbol{m}|\boldsymbol{v}) d\boldsymbol{m} \tag{5.25}$$

The distribution of $\boldsymbol{R}|\boldsymbol{V}$ is obtained by applying the standard technique to handle the product of two normal densities and the result is:

$$\boldsymbol{R}|\boldsymbol{V} \sim N(\boldsymbol{m}_{BL}, V + \Sigma_{BL}) \tag{5.26}$$

5.2.4 The modified Portfolio Problem

The predictive posterior density (5.26) substitutes the density of returns in the bayesian approach to utility maximization, see formula (5.8). For any portfolio $\boldsymbol{w} \in \mathbb{R}^n$ the random variable $R_{\boldsymbol{w}}|\boldsymbol{V} = \boldsymbol{w}^T \boldsymbol{R}|\boldsymbol{V}$, describing the portfolio return given the views, is distributed in the following way:

$$R_{\boldsymbol{w}}|\boldsymbol{V} \sim N(m^w_{BL}, \Sigma^w_{BL}) \tag{5.27}$$

where

$$\begin{aligned} m^w_{BL} &= \boldsymbol{w}^T \boldsymbol{m}_{BL} \ \in \mathbb{R} \\ \Sigma^w_{BL} &= \boldsymbol{w}^T (V + \Sigma_{BL}) \boldsymbol{w} \ \in \mathbb{R} \end{aligned} \tag{5.28}$$

Assume now $u \in \mathcal{U}_2$. The expectation of the utility with respect to the predictive posterior is given by:

$$\mathbb{E}(u(R_{\boldsymbol{w}})|\boldsymbol{V}) = \int u(r) f^{pre}_{R_{\boldsymbol{w}}|\boldsymbol{V}}(r)dr = \int u(r)\varphi(\frac{r - m^w_{BL}}{\Sigma^w_{BL}})dr \tag{5.29}$$

The normality of the predictive distribution implies the validity of mean-variance analysis, see the remark in Section 5.1. Therefore the expected utility can be maximized by solving the portfolio problem (3.7).

Fixing $E \in \mathbb{R}$ the related quadratic program is then the following one:

$$\text{Min} \quad \Sigma^w_{BL} \tag{5.30}$$

$$\text{with the constraints:} \quad m^w_{BL} = E \tag{5.31}$$

$$\mathbf{1}^T \boldsymbol{w} = 1. \tag{5.32}$$

6

Black Litterman model in non-normal markets

In [28] and in [29], Meucci extends the **BL** approach to markets in which assets returns are assumed to be non-normal. Rather than using a Bayesian inference approach, Meucci is able to include views into the investment process by a two steps computation based on opinion pooling and copulas. This approach has the attractive property to be adaptable to several non-normal distributions, including the skew elliptical class. Nonetheless his approach does not provide for a closed analytical form of the distribution of returns given the views.
The theory developed in this Chapter shows that, with the assumption of skew-normality for the returns of assets, it is possible to extend **BL** theory to non-normal markets preserving its Bayesian nature. In particular, we will show that, under this assumption, it is possible to obtain a closed form for the predictive distribution of returns given the views. In addition, thanks to the fact that this density is skew-normal, it is also possible to complete the allocation process by using a bayesian utility maximization procedure.

6.1 Market model

We have seen in the previous chapter that **BL** framework is characterized by its ability to blend historical market information with subjective investors views. In this approach, the manager rather than expressing views directly on possible realizations of the vector

of returns, expresses views on the realizations of its expected value, which is considered a random variable. As underlined by Meucci in [28], this could create some difficulties due to the fact that in many distributions there is no clear relationship between the expected returns and the parameters that characterize the distribution itself. In the case of a market modeled by normal returns, the parameters on which the manager expresses his views are the means, because of the equivalence between location parameters and means. Therefore expressing views on the location parameters in case of normality is quite intuitive. However, this correspondence may not hold for other distributions. This problem occurs in general for the class of skew-elliptical distributions, due to the fact that the expected return is a complex combination of all the parameters contained in the density function.
In order to solve this problem, we will first define a model that refers to a non normal market. We will work under the following assumption:

Assumption 6.1.1. *The vector of returns $\boldsymbol{R}$ is described by the following model:*

$$\boldsymbol{R}|(\Theta_1 = \boldsymbol{\mu}) \sim SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha}) \quad (6.1)$$
$$\Theta_1 \sim N_n(\boldsymbol{\mu}^e, \tau\Omega) \quad (6.2)$$

where Ω, $\boldsymbol{\alpha}$ and τ are known parameters and $\boldsymbol{\mu}^e$ is given by the equilibrium result (4.71).

Remark: The reader should notice that in this chapter we make our assumptions directly on the returns vector $\boldsymbol{R}$ rather than on the excess of returns $\boldsymbol{R}'$. We do this for our own convenience in handling some formulas, the two formulations being equivalent. It is worth mentioning that Θ_1 represents the uncertainty of the location parameter. This vector coincides with the vector of expected values of returns only in the case $\boldsymbol{\alpha} = \boldsymbol{0}$.
The parameter τ, as in **BL**, is a small parameter (we already mentioned that its value is commonly fixed in the interval $(0.02, 0.05)$). It assures that the location parameters are much less volatile than the returns. Furthermore we recall the expression of $\boldsymbol{\mu}^e$ (see (4.71)):

$$\mu_i^e = R_f + \beta_i[\mu_m - R_f] + (\gamma_i - \beta_i)[\mu_p - R_f]$$

where μ_m and μ_p are the location parameters respectively of the market portfolio $\boldsymbol{w}_m$ and of the portfolio $\boldsymbol{w}_p$ having $\beta_p = 0$ and $\gamma_p = 1$ (see sec. 5 in chapter 4) and where

$\beta_i = v_{i,m}/\sigma_m^2$ and $\gamma_i = (\omega\boldsymbol{\delta})_i/B_m$.
The vector of the expected values of $\boldsymbol{R}|\Theta_1$, denoted by $\boldsymbol{M}$, is given by :

$$\boldsymbol{M} := \mathbb{E}(\boldsymbol{R}|\Theta_1) = \Theta_1 + \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta})$$

where ω and δ are linked to Ω and $\boldsymbol{\alpha}$ by the usual relations.
Rather than expressing the returns conditionally to the location parameters, as in (6.1), we express them conditionally to the vector of expected values $\boldsymbol{M}$:

$$\begin{aligned} \boldsymbol{R}|(\boldsymbol{M} = \boldsymbol{m}) &\sim SN_n(\boldsymbol{m} - \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta}), \Omega, \boldsymbol{\alpha}) &(6.3)\\ \boldsymbol{M} &\sim N_n(\boldsymbol{m}^e, \tau\Omega) &(6.4)\end{aligned}$$

where $\boldsymbol{m}^e = \boldsymbol{\mu}^e + \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta})$.
For the sake of completeness we give two other equivalent expressions of the market model. From the Proposition 4.2.1, the previous model can be written as:

$$\begin{aligned} \boldsymbol{R}|(\Theta_1 = \boldsymbol{\mu}) &= \boldsymbol{\mu} + (\omega\boldsymbol{\delta})|X| + \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z} &(6.5)\\ \Theta_1 &\sim N_n(\boldsymbol{\mu}^e, \tau\Omega)\\ X &\sim N(0,1)\\ \boldsymbol{Z} &\sim N_n(0, \Psi)\end{aligned}$$

where $\Theta_1, X, \boldsymbol{Z}$ are independent and we refer the reader to Section 4.2 for the expression of the other parameters.
Finally using the expected returns vector $\boldsymbol{M}$, the model can be further modified in the following way:

$$\begin{aligned} \boldsymbol{R}|(\boldsymbol{M} = \boldsymbol{m}) &= \boldsymbol{m} + (\omega\boldsymbol{\delta})\tilde{X} + \omega(Id - \Delta^2)^{1/2}\boldsymbol{Z} &(6.6)\\ \Theta_1 &\sim N_n(\boldsymbol{\mu}^e, \tau\Omega)\\ \tilde{X} &= |X| - \sqrt{\frac{2}{\pi}}\\ X &\sim N(0,1)\\ Z &\sim N_n(0, \Psi)\end{aligned}$$

6.1.1 Marginal distribution of $\boldsymbol{R}$

Lemma 6.1.1. *Given the following Bayesian model:*

$$\boldsymbol{R}|\boldsymbol{M}=\boldsymbol{m} \sim SN_n(\boldsymbol{m}-\boldsymbol{\chi},\Omega,\boldsymbol{\alpha}_0) \tag{6.7}$$
$$\boldsymbol{M} \sim N_n(\boldsymbol{m}_0,\Omega_0) \tag{6.8}$$

where $\boldsymbol{\chi}\in\mathbb{R}^n$ is a deterministic vector, then the marginal distribution of $\boldsymbol{R}$ is skew-normal.
More specifically, if the parameters $(\boldsymbol{\chi},\Omega,\Omega_0,\boldsymbol{\alpha}_0,\boldsymbol{m}_0)$ are those ones specified in (6.3) and (6.4), then the marginal distribution of $\boldsymbol{R}$ is:

$$\boldsymbol{R}\sim SN_n(\boldsymbol{m}^e-\sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta}),(1+\tau)\Omega,\frac{\frac{\boldsymbol{\alpha}}{\sqrt{1+\tau}}}{\sqrt{1+\frac{\tau}{1+\tau}\boldsymbol{\alpha}^T\overline{\Omega}\boldsymbol{\alpha}}}) \tag{6.9}$$

Proof. Calling $\tilde{\boldsymbol{r}}=\boldsymbol{r}+\boldsymbol{\chi}$ the marginal distribution of $\boldsymbol{R}$ is given by:

$$\begin{aligned} f_{\boldsymbol{R}}(\tilde{\boldsymbol{r}}) &= \int f_{\boldsymbol{R}|\boldsymbol{M}}(\tilde{\boldsymbol{r}},\boldsymbol{m})\cdot f_{\boldsymbol{M}}(\boldsymbol{m})d\boldsymbol{m} \\ &= \int \frac{2}{\sqrt{|\Omega|}}\varphi(\tilde{\boldsymbol{r}};\boldsymbol{m},\Omega)\cdot\frac{1}{\sqrt{|\Omega_0|}}\varphi(\boldsymbol{m};\boldsymbol{m}_0,\Omega_0)\Phi(\boldsymbol{\alpha}_0^T\omega^{-1}(\tilde{\boldsymbol{r}}-\boldsymbol{m}))d\boldsymbol{m} \end{aligned} \tag{6.10}$$

Using the standard decomposition for the product of two gaussians we obtain:

$$\begin{aligned} & \frac{2}{\sqrt{|\Omega|}}\varphi(\tilde{\boldsymbol{r}};\boldsymbol{m},\Omega)\cdot\frac{1}{\sqrt{|\Omega_0|}}\varphi(\boldsymbol{m};\boldsymbol{m}_0,\Omega_0) \\ = & \frac{2}{\sqrt{|\Omega||\Omega_0|}}\varphi(\tilde{\boldsymbol{r}};\boldsymbol{m}_0,\Omega+\Omega_0)\cdot\varphi(\boldsymbol{m};\boldsymbol{z}(\tilde{\boldsymbol{r}},\boldsymbol{m}_0)),\Delta) \end{aligned} \tag{6.11}$$

where

$$\begin{aligned} z(\tilde{\boldsymbol{r}},\boldsymbol{m}_0) &= \Delta(\Omega^{-1}\tilde{\boldsymbol{r}}+\Omega_0^{-1}\boldsymbol{m}_0) \\ \Delta &= (\Omega^{-1}+\Omega_0^{-1})^{-1} \end{aligned}$$

In this way formula (6.10) becomes:

$$\begin{aligned} f_{\boldsymbol{R}}(\tilde{\boldsymbol{r}}) &= \int\frac{2}{\sqrt{|\Omega||\Omega_0|}}\varphi(\tilde{\boldsymbol{r}};\boldsymbol{m}_0,\Omega+\Omega_0)\cdot\varphi(\boldsymbol{m};\boldsymbol{z}(\tilde{\boldsymbol{r}},\boldsymbol{m}_0),\Delta)\Phi(\boldsymbol{\alpha}_0^T\omega^{-1}(\tilde{\boldsymbol{r}}-\boldsymbol{m}))d\boldsymbol{m} \\ &= \int\frac{2}{\sqrt{|\Omega||\Omega_0|}}\varphi(\tilde{\boldsymbol{r}};\boldsymbol{m}_0,\Omega+\Omega_0)\cdot\varphi(\boldsymbol{m};\boldsymbol{z}(\tilde{\boldsymbol{r}},\boldsymbol{m}_0),\Delta) \\ &\quad \Phi(\boldsymbol{\alpha}_0^T\omega^{-1}(\tilde{\boldsymbol{r}}-\boldsymbol{z}(\tilde{\boldsymbol{r}},\boldsymbol{m}_0))-\boldsymbol{\alpha}_0^T\omega^{-1}(\boldsymbol{m}-\boldsymbol{z}(\tilde{\boldsymbol{r}},\boldsymbol{m}_0)))d\boldsymbol{m} \\ &= \int\frac{2}{\sqrt{|\Omega||\Omega_0|}}\varphi(\tilde{\boldsymbol{r}};\boldsymbol{m}_0,\Omega+\Omega_0)\cdot\varphi(\boldsymbol{m};\boldsymbol{z}(\tilde{\boldsymbol{r}},\boldsymbol{m}_0),\Delta) \\ &\quad \Phi(\rho_0+\boldsymbol{\alpha}_1^T\delta^{-1}(\boldsymbol{m}-\boldsymbol{z}(\tilde{\boldsymbol{r}},\boldsymbol{m}_0)))d\boldsymbol{m} \end{aligned} \tag{6.12}$$

where δ is the diagonal matrix of standard deviations of Δ and

$$\begin{aligned}
\boldsymbol{\alpha}_1^T &= -\boldsymbol{\alpha}_0^T \omega^{-1} \delta \\
\rho_0 &= \rho \sqrt{1 + \boldsymbol{\alpha}_1^T \overline{\Delta} \boldsymbol{\alpha}_1} \\
\rho &= \boldsymbol{\alpha}_0^T \omega^{-1} (1 + \boldsymbol{\alpha}_1^T \overline{\Delta} \boldsymbol{\alpha}_1)^{-1/2} (\tilde{\boldsymbol{r}} - \boldsymbol{z}(\tilde{\boldsymbol{r}}, \boldsymbol{m}_0)) \\
&= \boldsymbol{\alpha}_0^T \omega^{-1} \Omega (\Omega + \Omega_0)^{-1} (1 + \boldsymbol{\alpha}_1^T \overline{\Delta} \boldsymbol{\alpha}_1)^{-1/2} (\tilde{\boldsymbol{r}} - \boldsymbol{m}_0) \qquad (6.13)
\end{aligned}$$

Finally formula (6.12) can be rearranged in the following form:

$$\begin{aligned}
f_{\boldsymbol{R}}(\tilde{\boldsymbol{r}}) &= \int \frac{2}{\sqrt{|\Omega + \Omega_0|}} \varphi(\tilde{\boldsymbol{r}}; \boldsymbol{m}_0, \Omega + \Omega_0) \Phi(\rho) \cdot \frac{1}{\Phi(\rho)} \varphi(\boldsymbol{m}; \boldsymbol{z}(\tilde{\boldsymbol{r}}, \boldsymbol{m}_0), \Delta) \\
&\quad \Phi(\rho_0 + \boldsymbol{\alpha}_1^T \delta^{-1} (\boldsymbol{m} - \boldsymbol{z}(\tilde{\boldsymbol{r}}, \boldsymbol{m}_0))) d\boldsymbol{m} \\
&= \frac{2}{\sqrt{|\Omega + \Omega_0|}} \varphi(\tilde{\boldsymbol{r}}; \boldsymbol{m}_0, \Omega + \Omega_0) \\
&\quad \Phi(\boldsymbol{\alpha}_0^T \omega^{-1} \Omega (\Omega + \Omega_0)^{-1} (\tilde{\boldsymbol{r}} - \boldsymbol{m}_0) (1 + \boldsymbol{\alpha}_1^T \overline{\Delta} \boldsymbol{\alpha}_1)^{-1/2}) \\
&\cdot \int \frac{1}{\Phi(\rho)} \varphi(\boldsymbol{m}; \boldsymbol{z}(\tilde{\boldsymbol{r}}, \boldsymbol{m}_0), \Delta) \Phi(\rho_0 + \boldsymbol{\alpha}_1^T \delta^{-1} (\boldsymbol{m} - \boldsymbol{z}(\tilde{\boldsymbol{r}}, \boldsymbol{m}_0))) d\boldsymbol{m} \qquad (6.14)
\end{aligned}$$

The expression inside the integral is the density of a multivariate skew normal random variable of the extended form (see 1.31) and its integral is one. Hence it holds:

$$f_{\boldsymbol{R}}(\tilde{\boldsymbol{r}}) = \frac{2}{\sqrt{|\Omega + \Omega_0|}} \varphi(\tilde{\boldsymbol{r}}; \boldsymbol{m}_0, \Omega + \Omega_0) \Phi(\boldsymbol{\alpha}_0^T \omega^{-1} \Omega (\Omega + \Omega_0)^{-1} (\tilde{\boldsymbol{r}} - \boldsymbol{m}_0) (1 + \boldsymbol{\alpha}_1^T \overline{\Delta} \boldsymbol{\alpha}_1)^{-1/2})$$

Using an opportune shape parameter $\boldsymbol{\alpha}_2$ it is possible to rewrite the previous expression in the following form:

$$f_{\boldsymbol{R}}(\tilde{\boldsymbol{r}}) = \frac{2}{\sqrt{|\Omega + \Omega_0|}} \varphi(\tilde{\boldsymbol{r}}; \boldsymbol{m}_0, \Omega + \Omega_0) \Phi(\boldsymbol{\alpha}_2^T \gamma^{-1} (\tilde{\boldsymbol{r}} - \boldsymbol{m}_0)) \qquad (6.15)$$

where γ is the diagonal matrix of standard deviations of $\Omega + \Omega_0$.

Assuming that the values of the parameters in (6.7) and (6.8) are those ones specified in (6.3) and (6.4), we get the following expression for the marginal:

$$\boldsymbol{R} \sim SN_n(\boldsymbol{m}^e - \sqrt{\frac{2}{\pi}}(\omega \boldsymbol{\delta}), (1+\tau)\Omega, \frac{\frac{\boldsymbol{\alpha}}{\sqrt{1+\tau}}}{\sqrt{1 + \frac{\tau}{1+\tau} \boldsymbol{\alpha}^T \overline{\Omega} \boldsymbol{\alpha}}}). \qquad (6.16)$$

□

6.2 Incorporating views

We assume that all investors have the same beliefs about the values of the parameters $(\boldsymbol{\mu}^e, \Omega, \boldsymbol{\alpha}, \tau)$. They represent the information coming from the market.
The second step of **BL** relies on modeling the investors views. These add an investor or market-expert subjective input into the investment process.
The information about investors' opinions on the expected values of assets is carried by the q-vector of views $\boldsymbol{V}$ and by the $(q \times n)$ "pick" matrix P, where q represents the number of views.
Even though expected returns are non-normal there is no reason to modify the conditional distribution of the investor's views given the expected returns.

Assumption 6.2.1. *The random vector of views is given by:*

$$\boldsymbol{V}|(\boldsymbol{M} = \boldsymbol{m}) \sim N(P\boldsymbol{m}, \Omega_v) \tag{6.17}$$

where Ω_v is a $(q \times q)$ matrix , with $(\Omega_v)_{ij} = 0$ if $i \neq j$, which measures the confidence on the market-expert opinions, and P is a $(q \times n)$ pick-matrix.

In the following Section we evaluate the posterior distribution of expected values given the views following the same procedure described in Chapter 5

6.2.1 Posterior distribution of $\boldsymbol{M}|\boldsymbol{V}$

To find the posterior distribution $f^{po}_{\boldsymbol{M}|\boldsymbol{V}}(\boldsymbol{m}|\boldsymbol{v})$, the prior distribution of the expected values, given by formula (6.4), is combined with the distribution of the views, specified by formula (6.17).
Let us write the basic Bayes rule:

$$f_{\boldsymbol{V}}(\boldsymbol{v}) f^{po}_{\boldsymbol{M}|\boldsymbol{V}}(\boldsymbol{m}, \boldsymbol{v}) = f_{\boldsymbol{M}}(\boldsymbol{m}) f_{\boldsymbol{V}|\boldsymbol{M}}(\boldsymbol{v}|\boldsymbol{m}) \tag{6.18}$$

Substituting the corresponding densities one finds:

$$f_{\boldsymbol{V}}(\boldsymbol{v}) f^{po}_{\boldsymbol{M}|\boldsymbol{V}}(\boldsymbol{m}|\boldsymbol{v}) = \left(\frac{2}{\sqrt{|(\tau\Omega)|}}\right) \varphi(\boldsymbol{m}; \boldsymbol{m}^e, (\tau\Omega)) \left(\frac{1}{\sqrt{|\Omega_v|}}\right) \varphi(\boldsymbol{v}; P\boldsymbol{m}, \Omega_v)$$

Using the standard result for the product of two gaussians densities, the previous expression becomes:

$$
\begin{aligned}
f_{\boldsymbol{V}}(\boldsymbol{v}) f^{po}_{\boldsymbol{M}|\boldsymbol{V}}(\boldsymbol{m}|\boldsymbol{v}) &= \left(\frac{2}{\sqrt{|(\tau\Omega)+\Omega_v|}}\right) \varphi(\boldsymbol{v}; \boldsymbol{m}^e, (\tau\Omega)+\Omega_v) \cdot \\
&\cdot \left(\frac{1}{\sqrt{|\Sigma_{BL}|}}\right) \varphi(\boldsymbol{m}; \boldsymbol{m}_{BL}, \Sigma_{BL}) \qquad (6.19)
\end{aligned}
$$

where the location parameter and the covariances matrix in the second line of the previous expression are the same as those ones specified in Section 5.2.3, namely:

$$
\boldsymbol{m}_{BL} = [(\tau\Omega)^{-1} + \Omega'_v]^{-1}[(\tau\Omega)^{-1}\boldsymbol{m}^e + P^T\Omega_v^{-1}\boldsymbol{v}] \qquad (6.20)
$$

$$
\Sigma_{BL} = [(\tau\Omega)^{-1} + \Omega'_v]^{-1} \qquad (6.21)
$$

$$
\Omega'_v = P^T\Omega_v^{-1}P \qquad (6.22)
$$

Therefore the posterior distribution of returns given the views is the same as the one obtained in the classical **BL**:

$$
f^{po}_{\boldsymbol{M}|\boldsymbol{V}}(\boldsymbol{m}|\boldsymbol{v}) = \left(\frac{1}{\sqrt{|\Sigma_{BL}|}}\right) \varphi_n(\boldsymbol{m}; \boldsymbol{m}_{BL}, \Sigma_{BL}) \qquad (6.23)
$$

6.2.2 Posterior predictive distribution of $\boldsymbol{R}|\boldsymbol{V}$

The predictive posterior density $f^{pr}_{\boldsymbol{R}|\boldsymbol{V}}(\boldsymbol{r}|\boldsymbol{v})$ is given by the following equality:

$$
f^{pr}_{\boldsymbol{R}|\boldsymbol{V}}(\boldsymbol{r}|\boldsymbol{v}) = \int f_{\boldsymbol{R}|\boldsymbol{M}}(\boldsymbol{r}|\boldsymbol{m}) \cdot f^{po}_{\boldsymbol{M}|\boldsymbol{V}}(\boldsymbol{m}|\boldsymbol{v}) d\boldsymbol{m} \qquad (6.24)
$$

Introducing (6.3) and (6.23) in the previous equation and using the scaled variable $\tilde{\boldsymbol{r}}$ we get:

$$
\begin{aligned}
f^{pr}_{\boldsymbol{R}|\boldsymbol{V}}(\tilde{\boldsymbol{r}}|\boldsymbol{v}) &= \int \frac{2}{\sqrt{|\Omega|}} \varphi_n(\tilde{\boldsymbol{r}}; \boldsymbol{m}, \Omega) \cdot \left(\frac{1}{\sqrt{|\Sigma_{BL}|}}\right) \varphi_n(\boldsymbol{m}; \boldsymbol{m}_{BL}, \Sigma_{BL}) \\
&\quad \Phi(\boldsymbol{\alpha}^T\omega^{-1}(\tilde{\boldsymbol{r}} - \boldsymbol{m})) d\boldsymbol{m} \qquad (6.25)
\end{aligned}
$$

Using the standard rule for the product of two gaussians we obtain:

$$
\begin{aligned}
f^{pr}_{\boldsymbol{R}|\boldsymbol{V}}(\tilde{\boldsymbol{r}}|\boldsymbol{v}) &= \int \frac{2}{\sqrt{|\Omega+\Sigma_{BL}|}} \varphi_n(\tilde{\boldsymbol{r}}; \boldsymbol{m}_{BL}, \Omega+\Sigma_{BL}) \\
&\cdot \left(\frac{1}{\sqrt{|\Delta_{BL}|}}\right) \varphi_n(\boldsymbol{m}; \boldsymbol{z}_{BL}(\tilde{\boldsymbol{r}}, \boldsymbol{m}_{BL}), \Delta_{BL}) \Phi(\xi + \boldsymbol{\gamma}^T\sigma_{BL}^{-1}(\boldsymbol{m} - \boldsymbol{m}_{BL})) d\boldsymbol{m}
\end{aligned}
$$

where:

$$\begin{aligned} \boldsymbol{z}_{BL}(\tilde{\boldsymbol{r}}, \boldsymbol{m}_{BL}) &= [\Omega^{-1} + \Sigma_{BL}^{-1}]^{-1}[\Omega^{-1}\tilde{\boldsymbol{r}} + \Sigma_{BL}^{-1}\boldsymbol{m}_{BL}] \\ \Delta_{BL} &= (\Omega^{-1} + \Sigma_{BL}^{-1})^{-1} \end{aligned}$$

With the same algebra of Section 6.1.1 it can be proven that the posterior predictive density of $\boldsymbol{R}|\boldsymbol{V}$ is given by the following:

$$f^{pr}_{\boldsymbol{R}|\boldsymbol{V}}(\tilde{\boldsymbol{r}}|\boldsymbol{v}) = \frac{2}{\sqrt{|\Omega + \Sigma_{BL}|}}\varphi_n(\tilde{\boldsymbol{r}}; \boldsymbol{m}_{BL}, \Omega + \Sigma_{BL})\Phi(\boldsymbol{\alpha}_{BL}^T\gamma_{BL}^{-1}(\tilde{\boldsymbol{r}} - \boldsymbol{m}_{BL})) \tag{6.26}$$

where γ_{BL} is the diagonal matrix of the standard deviations of $\Omega + \Sigma_{BL}$. The shape parameter $\boldsymbol{\alpha}_{BL}$ is given by:

$$\boldsymbol{\alpha}_{BL}^T = \boldsymbol{\alpha}^T\omega^{-1}\Omega(\Omega + \Sigma_{BL})^{-1}(1 + \boldsymbol{\alpha}_{\Delta}^T\overline{\Delta}_{BL}\boldsymbol{\alpha}_{\Delta})^{-1/2} \tag{6.27}$$

where

$$\boldsymbol{\alpha}_{\Delta}^T = -\boldsymbol{\alpha}^T\omega^{-1}d_{BL},$$

d_{BL} is the diagonal matrix of standard deviations of Δ_{BL} and $\overline{\Delta}_{BL}$ its correlation matrix.

Therefore the random vector $\boldsymbol{R}|\boldsymbol{V}$ is distributed according to:

$$\boldsymbol{R}|\boldsymbol{V} \sim SN_n(\boldsymbol{m}_{BL} - \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta}), \Omega + \Sigma_{BL}, \boldsymbol{\alpha}_{BL}) \tag{6.28}$$

and we have the following expressions for the parameters:

$$\begin{aligned} \Omega'_v &= P^T\Omega_v P \\ \boldsymbol{m}_{BL} &= [(\tau\Omega)^{-1} + \Omega'_v]^{-1}[(\tau\Omega)^{-1}\boldsymbol{m}^e + P^T\Omega^{-1}\boldsymbol{v}] \\ \Sigma_{BL} &= [(\tau\Omega)^{-1} + \Omega'_v]^{-1} \end{aligned}$$

From (6.28) the expected value of $\boldsymbol{R}|\boldsymbol{V}$ can be easily evaluated:

$$\mathbb{E}(\boldsymbol{R}|\boldsymbol{V}) = \boldsymbol{m}_{BL} + \sqrt{\frac{2}{\pi}}((\gamma_{BL}\boldsymbol{\delta}_{BL}) - (\omega\boldsymbol{\delta})) \tag{6.29}$$

where $\boldsymbol{\delta}_{BL}$ is given by:

$$\boldsymbol{\delta}_{BL} = \frac{(\overline{\Omega + \Sigma}_{BL})\boldsymbol{\alpha}_{BL}}{\sqrt{1 + \boldsymbol{\alpha}_{BL}^T(\overline{\Omega + \Sigma}_{BL})\boldsymbol{\alpha}_{BL}}} \tag{6.30}$$

$(\overline{\Omega + \Sigma}_{BL})$ being the correlation matrix of $\Omega + \Sigma_{BL}$.

6.3 The Portfolio Problem revisited

The evaluation of the predictive posterior distribution, described in the previous Section, completes the Bayesian "part" of the investment process. At this stage the investor has all the inputs he needs for the implementation of the portfolio problem.

For any portfolio $\boldsymbol{w} \in \mathbb{R}^n$ we consider again the random variable $R_{\boldsymbol{w}}|\boldsymbol{V} := \boldsymbol{w}^T \boldsymbol{R}|\boldsymbol{V}$. It describes the portfolio return given the views. The distribution of $R_{\boldsymbol{w}}|\boldsymbol{V}$ is obtained by an affine transformation of $\boldsymbol{R}|\boldsymbol{V}$, whose expression is given by formula (6.28). Using standard rules we therefore obtain:

$$R_{\boldsymbol{w}}|\boldsymbol{V} \sim SN(\mu_{BL}^w, \Omega_{BL}^w, \alpha_{BL}^w) \tag{6.31}$$

where the parameters $(\mu_{BL}^w, \Omega_{BL}^w, \alpha_{BL}^w) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, depending on $\boldsymbol{w}$, are given by :

$$\begin{aligned} \mu_{BL}^w &= \boldsymbol{w}^T(\boldsymbol{m}_{BL} - \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta})) \\ \Omega_{BL}^w &= \boldsymbol{w}^T(\Omega + \Sigma_{BL})\boldsymbol{w} \\ \alpha_{BL}^w &= \frac{\delta_w^{BL}}{\sqrt{1-(\delta_w^{BL})^2}} \end{aligned}$$

with:

$$\delta_w^{BL} = \frac{\boldsymbol{w}^T(\gamma_{BL}\boldsymbol{\delta}_{BL})}{\sqrt{\Omega_{BL}^w}}$$

The portfolio problem that each risk-averse investor, with utility function $u \in \mathfrak{U}_2$, faces in a market of skew-normal assets returns is described in Section 4.3.1. The expression of the expected utility given the views is the following:

$$\begin{aligned} \mathbb{E}(u(R_{\boldsymbol{w}}|\boldsymbol{V}=\boldsymbol{v})) &= \int u(r) f^{pre}_{R_{\boldsymbol{w}}|\boldsymbol{V}}(r|\boldsymbol{v})dr \\ &= \int u(r)\frac{2}{\sqrt{\Omega_{BL}^w}}\varphi(r;\mu_{BL}^w,\Omega_{BL}^w)\Phi(\alpha_{BL}^w(\Omega_{BL}^w)^{-1/2}(r-\alpha_{BL}^w))dr \end{aligned}$$

Denoting by

$$s_{BL}^2 := \Omega_{BL}^w - (\boldsymbol{w}^T(\gamma_{BL}\boldsymbol{\delta}_{BL}))^2$$

the *spherical* component of the variance, the portfolio problem resulting is simply obtained adapting the problem (4.43). Fixed the values of the location and of the *non* *spherical* component of the variance respectively to $L, B \in \mathbb{R}$, then each *minimum spherical variance* portfolio is a solution of the following problem:

$$\begin{aligned} \text{Min}_w \qquad & s^2_{BL} \\ \text{with the constraints:} \qquad & \mu^w_{BL} = L \\ & \boldsymbol{w}^T(\gamma_{BL}\boldsymbol{\delta}_{BL}) = B \\ & \mathbf{1}^T\boldsymbol{w} = 1. \end{aligned} \tag{6.32}$$

6.4 The Hedge Funds Market

In this Section we apply the main theoretical results obtained in this Thesis to a Portfolio composed of 12 Hedge Funds Strategies Indexes. The data set used in this study is made of 12 HFR Indexes Strategies monthly returns. Although those indexes are not investable they represent an ideal framework to analyze skewness in returns. We assume the 12 Hedge Funds Strategies cover the entire Hedge Funds Market (excluding Funds of Hedge Funds).

The historical monthly series of the 12 indexes goes from 01/1990 to 10/2007 and the total number of observation is 214. They are: Convertible Arbitrage, Distressed Securities, Emerging Markets (Total), Equity Hedge, Equity Market Neutral, Equity Non-Hedge, Event-Driven, Fixed Income (Total), Macro, Market Timing, Merger Arbitrage and Relative Value Arbitrage.

The risk-free rate taken is the 1-month Euribor.

We assume that investors invest 1 Euro at time t in their portfolio and that portfolios are hold unchanged until time $\tau > t$, with $\tau - t = 1$ month. As a result, the returns are assumed to be monthly returns.

6.4.1 Mean Variance Analysis

In the next section we prove that the normality assumption is much less attractive than the skew normal one. This is proven both statistically by the likelihood ratio test and graphically by the QQ-plots of some strategies.

Nonetheless the mean variance analysis continues to be interesting due to the fact that the results obtained in this framework can used as benchmark for more complex cases.

We assume here that the joint distribution of the 12 strategies is normal:

$$\boldsymbol{R} \sim N_{12}(\hat{\boldsymbol{m}}, \hat{V}) \tag{6.33}$$

where $R_1, \ldots, R_{12}$ represent respectively the returns of the 12 strategies in the order in which they have been mentioned above. This assumption validates the mean variance analysis.

The estimated parameters $(\hat{\boldsymbol{m}}, \hat{V})$ have been obtained by the maximum likelihood method. The values of the estimated $\hat{\boldsymbol{m}}$ are reported in Table 6.1 whereas we do not report the values of $\hat{V}$. Given a portfolio $\boldsymbol{w}$ the portfolio return $R_w = \boldsymbol{w}^T\boldsymbol{R}$ is normally distributed:

$$\boldsymbol{R}_w \sim N(\boldsymbol{w}^T\hat{\boldsymbol{m}}, \boldsymbol{w}^T\hat{V}\boldsymbol{w}) \tag{6.34}$$

Suppose an investor sets the monthly target return to $1,2\%$, then its portfolio problem is by (3.7):

$$\begin{aligned} \text{Min} \quad & \tfrac{1}{2}\boldsymbol{w}^T\hat{V}\boldsymbol{w} \\ \text{with the constraints:} \quad & \mathbf{1}^T\boldsymbol{w} = 1 \\ & \boldsymbol{w}^T\hat{\boldsymbol{m}} = 1,2\% \end{aligned} \tag{6.35}$$

The optimal portfolio $\boldsymbol{w}^*$ resulting from this quadratic programming problem together with optimal portfolio return and standard deviation are reported in Table 6.2.
It results:

$$R_{\boldsymbol{w}^*} \sim N(1.2, 1.45) \tag{6.36}$$

In Figure 6.1 is represented the efficient frontier in the mean variance space.

Suppose now the investor owns personal views relating the Hedge Funds Strategies. The **BL** model described in chapter 5, permits to include this views in the investment process. The **BL** model is implementable imposing uncertainty on the values of expected returns by a prior distribution.
The market model we assume is (as in 5.11):

$$\boldsymbol{R}|\boldsymbol{M} = \boldsymbol{m} \quad \sim \quad N_{12}(\boldsymbol{m}; \hat{V}) \tag{6.37}$$

$$\boldsymbol{M} \quad \sim \quad N_{12}(\boldsymbol{\mu}^e, \tau\hat{V}) \tag{6.38}$$

where $\boldsymbol{m}^e$ are the CAPM equilibrium values given by (3.41):

$$\boldsymbol{m}^e - \mathbf{1}R_f = \boldsymbol{\beta}_m(m_m - R_f) \tag{6.39}$$

and $\tau = 0,03$. Even though in the original **BL** model the prior (5.11) is centered on the implied returns we prefer here to use the CAPM equilibrium values, the two quantities being related as shown in chapter 5. This choice will allow us to compare the mean values derived from the CAPM with those derived from the three moments CAPM.
The market capitalization weights (the market portfolio) $\boldsymbol{w}_m$ are reported in Table 6.1

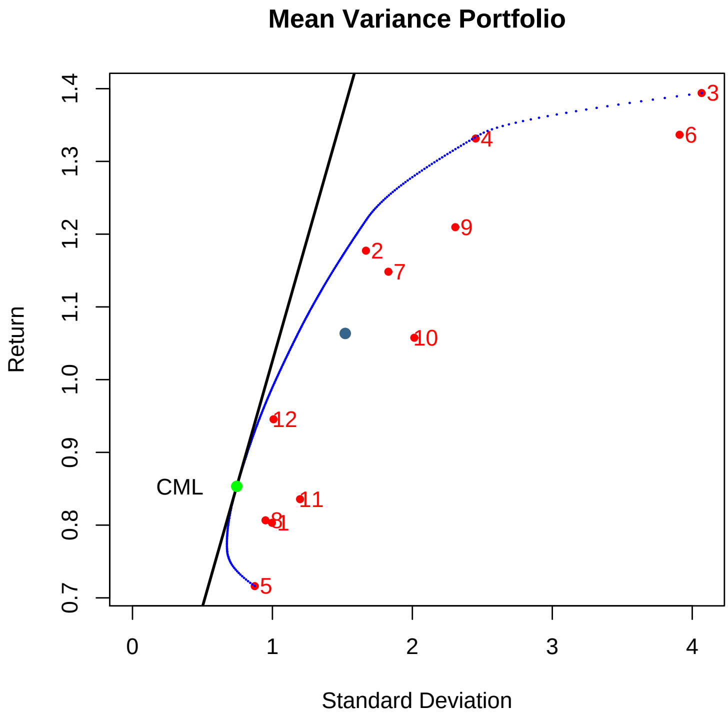


Figure 6.1: Mean variance *efficient set* for the portfolio of 12 strategies in the normal assumption. The tangency portfolio $(0,85\%, 0,74\%)$ and the equally weighted portfolio $(1,06\%, 2,03\%)$ are also represented.

	1	2	3	4	5	6	7	8	9	10	11	12
$\boldsymbol{w}_m$	0.04	0.03	0.06	0.40	0.06	0.08	0.08	0.10	0.03	0.04	0.04	0.04
$\boldsymbol{\beta}_m$	0.28	0.63	1.70	1.28	0.17	1.98	0.86	0.34	0.84	0.80	0.38	0.34
$\hat{\boldsymbol{\mu}}$	0,80	1,17	1,39	1,33	0,71	1,33	1,14	0,80	1,20	1,05	0,83	0,94
$\boldsymbol{m}^e$	0.51	0.82	1.79	1.41	0.40	2.05	1.02	0.56	1.01	0.97	0.59	0.56
$\boldsymbol{m}_{BL}$	0.47	0.73	1.94	1.37	0.38	1.99	0.95	0.54	1.02	0.98	0.55	0.52

Table 6.1: $\boldsymbol{w}_m$ is the vector of weight of the market portfolio, $\boldsymbol{\beta}_m$ and $\boldsymbol{m}^e$ are the coefficients needed for the two moment CAPM. $\hat{\boldsymbol{m}}$ is the vector of historical means. $\boldsymbol{m}_{BL}$ is the **BL** mean vector.

together with the values of the vector of "betas" $\boldsymbol{\beta}_m$.
It gives

$$R_m \sim N(1.15, 3.42).$$

and the results relating the equilibrium values $\boldsymbol{m}^e$ are reported in Table 6.1.

Suppose now the investor owns the following personal views on the expected monthly returns:
i) The Distressed Security Strategy will almost surely perform -2% the following month
ii) It's highly probable that the Emerging Markets Strategy will over-perform the Event Driven Strategy by $+4\%$
iii) It's probable that the Fixed Income Strategy will perform the same as the average between the three Equity Strategies (Equity Hedge, Equity Market Neutral and Equity non-Hedge)
By the **BL** model it is possible to codify those views in a vector of views $\boldsymbol{v}$ (the realization of the r.v. V), an opportune "pick" $(3 \times 12)-$matrix P, a confidence (3×3) diagonal matrix Ω_v. By (5.21) the views are distributed according to:

$$\boldsymbol{V}|\boldsymbol{M} = \boldsymbol{m} \sim N_3(P\boldsymbol{m}, \Omega_v) \tag{6.40}$$

By i) ii) and iii):

$$\boldsymbol{v} = \begin{pmatrix} \text{-2\%} & +4\% & 0\% \end{pmatrix} \tag{6.41}$$

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.42}$$

$$\Omega_v = \begin{pmatrix} 1{,}5\% & 0 & 0 \\ 0 & 2{,}5\% & 0 \\ 0 & 0 & 4\% \end{pmatrix} \tag{6.43}$$

The views can also be written in the following more intuitive form.

$$\begin{aligned} m_2 &= v_1 + \epsilon_1 \\ m_3 - m_5 &= v_2 + \epsilon_2 \\ \frac{m_4 + m_5 + m_6}{3} - m_8 &= v_3 + \epsilon_3 \end{aligned} \tag{6.44}$$

where $\boldsymbol{\epsilon} \sim N_3(0, \Omega_v)$. Using the prior for the means (6.38) and the distribution of the views (6.41) we can obtain the posterior of $\boldsymbol{M}|\boldsymbol{V}$ (see formula (5.22)), which is normal with the following mean and variance:

$$\boldsymbol{m}_{BL} = [(\tau\hat{V})^{-1} + P^T\Omega_v^{-1}P]^{-1}[(\tau\hat{V})^{-1}\boldsymbol{m}^e + P^T\Omega_v^{-1}\boldsymbol{v}] \tag{6.45}$$

$$\Sigma_{BL} = [(\tau\hat{V})^{-1} + P^T\Omega_v^{-1}P]^{-1} \tag{6.46}$$

In Table 6.1 are reported values of $\boldsymbol{m}_{BL}$ whereas values of Σ_{BL} are not reported here to save space. Setting the target return again to $1,2\%$ the modified portfolio problem is:

$$\begin{aligned} \text{Min} \quad & \boldsymbol{w}^T(\hat{V} + \Sigma_{BL})\boldsymbol{w} \\ \text{with the constraints:} \quad & \mathbf{1}^T\boldsymbol{w} = 1 \\ & w^T\boldsymbol{m}_{BL} = 1,2\% \end{aligned} \tag{6.47}$$

brings to the optimal weights showed in Table 6.2. It results:

$$R_{\boldsymbol{w}^*_{BL}} \sim N(1.2, 3.76) \tag{6.48}$$

6.4.2 The skew-normal assumption

The portfolio of the 12 strategies is assumed now to be joint skew normal:

$$\boldsymbol{R} \sim SN_{12}(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha}) \tag{6.49}$$

	1	2	3	4	5	6	7	8	9	10	11	12
$\boldsymbol{w}^*$	-0.26	0.42	-0.07	0.54	-0.25	-0.49	0.15	0.05	0.10	0.29	0.06	0.47
$\boldsymbol{w}^*_{BL}$	0.04	0.04	0.06	0.42	0.05	0.09	0.09	0.09	0.04	0.03	0.03	0.03

Table 6.2: Optimal portfolio w^*_{BL} obtained by the portfolio problem (6.47) compared with the optimal weights w^* obtained by (6.35)

where $R_1, \ldots, R_{12}$ represent respectively the 12 strategies in the order in which they have been mentioned at the beginning of this Section.
The results of the ML estimation of parameters $(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$ (in this Section we don't use a different symbol for the estimated parameters) are reported in Tables 6.4 and 6.5. The assumption (6.49) has been tested by a standard likelihood ratio test. The model with a restriction is that one with the vector $\boldsymbol{\alpha}$ of all zeros , implying the normality of the restricted model. The values of the test have been compared with the values from a chi-squared distribution with 12 degrees of freedom. The results given in Table 6.3 prove that the skew-normal assumption is much more appropriate.

Likelihood ratio Test	**(null hypothesis: $\boldsymbol{\alpha} = \mathbf{0}$)**
log-lik normal ($\boldsymbol{\alpha} = \mathbf{0}$)	-3967.99
log-lik skew-normal	-3923.89
lik-ratio	88
Prob	0

Table 6.3: Likelihood ratio test: lik-ratio is the likelihood ratio test statistics and Prob the corresponding probability.

	1	2	3	4	5	6	7	8	9	10	11	12
$\boldsymbol{\mu}$	1.13	2.01	3.68	2.34	0.84	3.34	2.51	1.17	1.86	1.40	1.98	1.31
$\boldsymbol{\alpha}$	0.42	0.66	-1.63	1.25	-0.22	-1.12	-1.72	0.01	0.14	1.41	-4.49	0.57

Table 6.4: Values of $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$

In Figure 6.2 are represented the marginal densities of two Strategies with the corresponding histograms. It can be noticed, also from the graphical point of view, that it would be inappropriate a normal assumption for these two skewed Strategies.

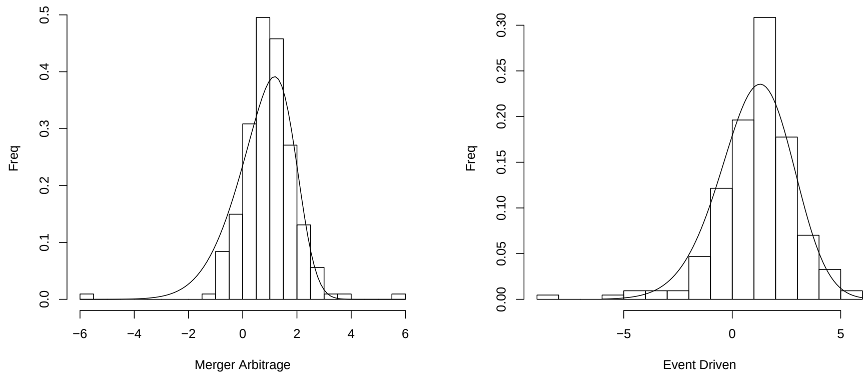


Figure 6.2: Histograms of Merger Arbitrage and Event Driven Strategies and the two marginal densities. The curve on the left is the density of a $SN(1.98, 1.65, -2.81)$, on the right of a $SN(2.51, 2.27, -1.40)$

This is also evident from the QQ-plots represented in Figure 6.3.

For comparison purposes we solve three location variance skewness problems all having fixed expected value:

$$\mathbb{E}(R_{\boldsymbol{w}}) = 1.2$$

in order to do this we choose the following three couples of location/non-spherical component of the variance: $(L, B) = (4, -3.5); (0, 1.5); (-4, 6.5)$.
The location variance skewness problem (4.43) is:

$$\begin{aligned} \text{Min}_{\boldsymbol{w}} \qquad & s^2_{\boldsymbol{w}} \\ \text{with the constraints:} \qquad & \boldsymbol{\mu}_{\boldsymbol{w}} = L \\ & \boldsymbol{w}^T(\omega\boldsymbol{\delta}) = B \\ & \mathbf{1}^T\boldsymbol{w} = 1 \end{aligned} \tag{6.50}$$

The values of the the three optimal portfolios $\boldsymbol{w}_i^*$, $i = 1, 2, 3$ solutions of the previous problem are reported in Table 6.6.

Suppose now the investor has the same views as in section 6.4.1. The skew-normal

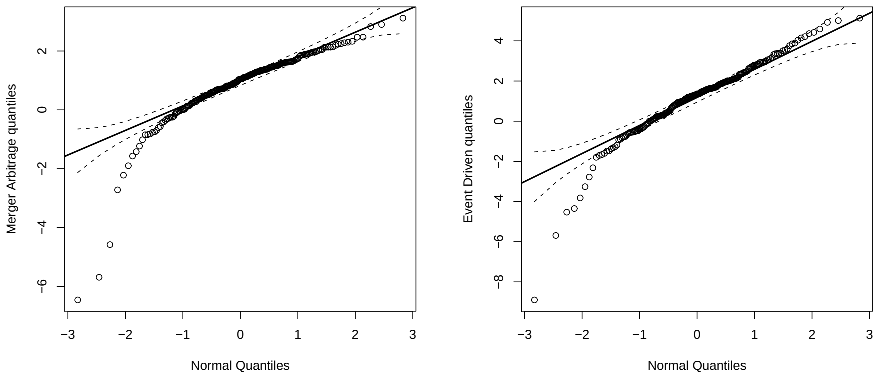


Figure 6.3: QQ-plots for the Merger Arbitrage and the Event Driven Startegies

	1	2	3	4	5	6	7	8	9	10	11	12
1	1.10	1.19	2.33	1.44	0.23	2.31	1.48	0.58	1.15	0.75	0.93	0.73
2	1.19	3.46	6.17	3.24	0.40	5.75	3.54	1.38	2.34	1.56	1.98	1.44
3	2.33	6.17	21.71	8.77	0.63	15.83	8.22	3.14	7.08	5.23	4.72	2.82
4	1.44	3.24	8.77	7.01	0.93	10.59	4.84	1.64	4.06	3.84	2.63	1.71
5	0.23	0.40	0.63	0.93	0.78	0.97	0.55	0.24	0.64	0.46	0.40	0.30
6	2.31	5.75	15.83	10.59	0.97	19.24	8.56	2.98	6.38	6.52	4.78	2.74
7	1.48	3.54	8.22	4.84	0.55	8.56	5.19	1.63	3.24	2.51	3.16	1.68
8	0.58	1.38	3.14	1.64	0.24	2.98	1.63	1.03	1.44	0.83	0.83	0.69
9	1.15	2.34	7.08	4.06	0.64	6.38	3.24	1.44	5.72	2.67	1.63	1.18
10	0.75	1.56	5.23	3.84	0.46	6.52	2.51	0.83	2.67	4.15	1.22	0.79
11	0.93	1.98	4.72	2.63	0.40	4.78	3.16	0.83	1.63	1.22	2.74	0.99
12	0.73	1.44	2.82	1.71	0.30	2.74	1.68	0.69	1.18	0.79	0.99	1.15

Table 6.5: Matrix Ω

	1	2	3	4	5	6	7	8	9	10	11	12
$\boldsymbol{w}_1^*$	−0.74	0.35	0.14	0.42	−0.49	−0.34	0.79	−0.52	0.14	−0.23	1.38	0.09
$\boldsymbol{w}_2^*$	−0.01	0.49	−0.20	0.66	−0.15	−0.60	−0.21	0.35	0.09	0.61	−0.74	0.72
$\boldsymbol{w}_3^*$	−2.43	2.32	−0.10	2.87	−2.65	−2.10	1.93	−2.19	0.75	0.49	0.55	1.57

Table 6.6: Three optimal portfolios w_i^* obtained by the portfolio problem (6.50)

market model is given by formula (6.3) which is recalled here:

$$\boldsymbol{R}|(M = \boldsymbol{m}) \sim SN_n(\boldsymbol{m} - \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta}), \Omega, \boldsymbol{\alpha}) \tag{6.51}$$

$$M \sim N_n(\boldsymbol{m}^e, \tau\Omega) \tag{6.52}$$

where $\boldsymbol{m}^e = \boldsymbol{\mu}^e + \sqrt{\frac{2}{\pi}}(\omega\boldsymbol{\delta})$.

The equilibrium vector is given by (its values are reported in Table 6.7):

$$\boldsymbol{\mu}_i^e = R_f + \beta_i[\mu_m - R_f] + (\gamma_i - \beta_i)[\mu_p - R_f] \tag{6.53}$$

The univariate skew-normal distribution of the market is:

$$R_m \sim SN(2.13, 2.09, -0.83)$$

hence $\mathbb{E}(R_m) = 1.39$, $\text{Var}(R_m) = 1.55$ and $B_m = \boldsymbol{w}_m^T(\omega\boldsymbol{\delta}) = -0.91$.

In Table 6.7 are reported the values of the covariances $v_{i,m}$, of the vector $\boldsymbol{\beta}_m$ of components $(\boldsymbol{\beta}_m)_i = v_{i,m}/\sigma_m^2$ and of the vector of components $\gamma_i = (\omega\boldsymbol{\delta})_i/(\omega\boldsymbol{\delta})_m$.

The portfolio $\boldsymbol{w}_p$ is defined as a portfolio uncorrelated to the market portfolio $v_{p,m} = 0$ but having its skewness $\gamma_p = 1$. It has been obtained solving the following linear programming problem:

$$\text{Min}_w \quad v_{1,m}w_1 + \ldots + v_{12,m}w_{12} \tag{6.54}$$

$$\text{with the constraints:} \quad v_{1,m}w_1 + \ldots + v_{12,m}w_{12} \geq 0$$

$$(\omega\boldsymbol{\delta})_1 w_1 + \ldots + (\omega\boldsymbol{\delta})_{12} w_{12} = B_m$$

$$w_1 + \ldots + w_{12} = 1 \tag{6.55}$$

The results relating the values of $\boldsymbol{\mu}^e$ and $\boldsymbol{m}^e$ are in Table 6.7.

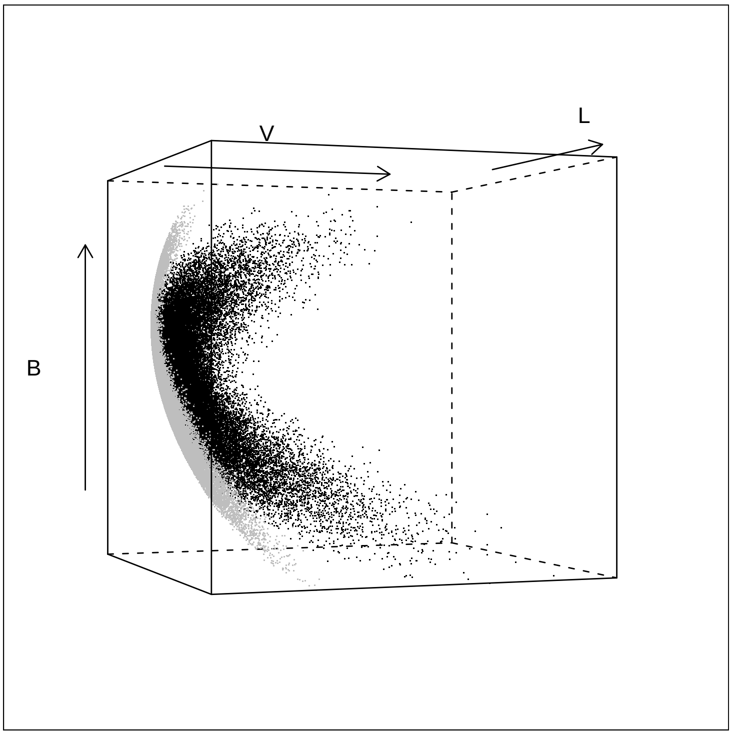
To obtain the r.v $\boldsymbol{M}|\boldsymbol{V}$ we use the prior on the expected return and the distribution of the views, the theoretical result is the same as in the normal case. The values of the vectors $\boldsymbol{m}_{BL}$ and $\boldsymbol{\alpha}_{BL}$ are reported in Table 6.7. While the values of Σ_{BL} are not reported to save space. Finally it's possible to evaluate the predictive posterior $\boldsymbol{R}|\boldsymbol{V}$. The resulting location variance skewness problem is:

$$\text{Min}_w \quad s_{BL}^2 \tag{6.56}$$

$$\text{with the constraints:} \quad \mu_{BL}^w = L$$

$$\boldsymbol{w}^T(\gamma_{BL}\boldsymbol{\delta}_{BL}) = B$$

$$\boldsymbol{1}^T\boldsymbol{w} = 1$$



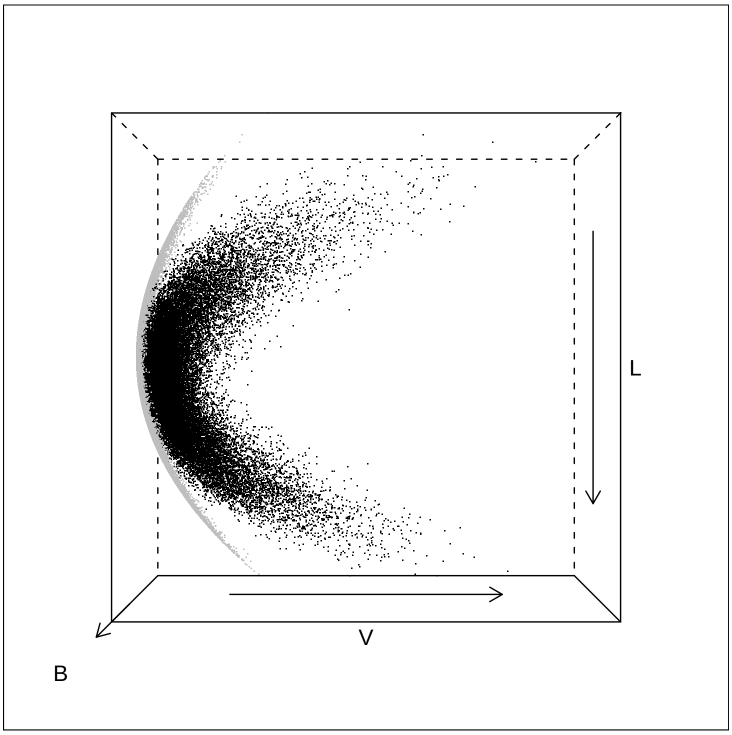


Figure 6.4: Plots of 30000 feasible portfolios (in grey) and of 30000 portfolios lying in the *minimum spherical variance* set (in black) in the $(L, V = v_{\boldsymbol{w}}^2, B)$-space for the portfolio of the 12 HF Strategies

	1	2	3	4	5	6	7	8	9	10	11	12
$v_{i,m}$	0.92	2.02	5.41	4.21	0.55	6.44	2.69	1.11	2.76	2.68	1.08	1.09
$\boldsymbol{\beta}_m$	0.59	1.30	3.49	2.71	0.35	4.15	1.73	0.72	1.78	1.73	0.70	0.70
γ	0.49	1.24	3.40	1.50	0.18	2.98	2.03	0.54	0.96	0.51	1.70	0.54
$\boldsymbol{\mu}$	1.13	2.01	3.68	2.34	0.84	3.34	2.51	1.17	1.86	1.40	1.98	1.31
$\boldsymbol{\mu}^e$	1.25	2.63	6.71	4.01	0.73	6.76	3.83	1.40	2.69	2.15	2.67	1.40
$\boldsymbol{m}^e$	0.89	1.72	4.22	2.92	0.59	4.58	2.35	1.01	1.99	1.78	1.43	1.00
$\boldsymbol{m}_{BL}$	0.82	1.52	4.06	2.73	0.56	4.27	2.14	0.94	1.89	1.71	1.30	0.92
$\boldsymbol{\alpha}_{BL}$	0.29	0.25	−0.25	0.34	−0.18	−0.18	−0.53	0.01	0.04	0.49	−1.91	0.38
$\boldsymbol{\delta}_{BL}$	−0.26	−0.41	−0.44	−0.34	−0.19	−0.42	−0.62	−0.30	−0.21	−0.09	−0.82	−0.28

Table 6.7:

	1	2	3	4	5	6	7	8	9	10	11	12
$\boldsymbol{w}_p$	0.00	0.00	−0.22	0.00	0.21	0.00	0.00	0.00	0.00	0.00	1.01	0.00

Table 6.8: Portfolio $\boldsymbol{w}_p$

We solve this problem for the following three couples of location/non-spherical part of the variance (the same as above): $(L, B) = (4, -3.5); (0, 1.5); (-4, 6.5)$.
Results are reported in Table 6.9.

	1	2	3	4	5	6	7	8	9	10	11	12
$\boldsymbol{w}^*_{BL,1}$	−0.66	−0.27	0.04	−0.46	0.53	0.03	0.48	0.01	−0.06	−0.54	2.74	−0.85
$\boldsymbol{w}^*_{BL,2}$	0.61	0.07	0.04	0.44	−0.04	−0.06	−0.41	0.56	0.01	0.49	−1.45	0.74
$\boldsymbol{w}^*_{BL,3}$	−0.61	−0.31	0.20	−0.10	0.27	0.19	0.61	−0.22	−0.00	−0.59	2.34	−0.78

Table 6.9: Three optimal portfolios $w^*_{BL,i}$ obtained by the portfolio problem (6.56)

Appendix A

Mathematical results

A.1 Bivariate SUN

In this section the fundamental properties of the bivariate **SUN** are analyzed, the main reference is [2].
The general expression of the density of a **SUN**, given by (1.46), leads to two different bivariate random variables: $\mathrm{SUN}_{2,1}$ and $\mathrm{SUN}_{2,2}$.
Starting from the former and considering the general formula of the density (1.46), we have that the r.v.

$$\boldsymbol{Y} \sim \mathrm{SUN}_{2,1}(\xi, 0, \bar{\omega}, \Omega^*)$$

has the following density :

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2\varphi_2(\boldsymbol{y} - \boldsymbol{\xi}; \Omega)\Phi(\Delta^T \bar{\Omega}^{-1} \omega^{-1} (\boldsymbol{y} - \boldsymbol{\xi}); 1 - \Delta^T \bar{\Omega}^{-1} \Delta)$$

where Ω, Δ are obtained by the opportune partition of the (3×3) correlation matrix:

$$\Omega^* = \begin{pmatrix} 1 & \Delta^T \\ \Delta & \bar{\Omega} \end{pmatrix}$$

and where:

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \quad ; \quad \Delta = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \quad ; \quad \Omega = \omega \bar{\Omega} \omega$$

with

$$\omega = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \quad ; \quad \bar{\Omega} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

with $\sigma_i > 0$ and $|\rho| \leq 1$. Expliciting the previous expression of the density we have:

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}) &= 2\varphi_2(\boldsymbol{y}-\boldsymbol{\xi};\Omega) \\ & \Phi(\frac{\delta_1-\rho\delta_2}{\sigma_1(1-\rho^2)\sqrt{1-\delta_1^2-\delta_2^2+2\rho\delta_1\delta_2}}(y_1-\xi_1) \\ &+ \frac{\delta_2-\rho\delta_1}{\sigma_2(1-\rho^2)\sqrt{1-\delta_1^2-\delta_2^2+2\rho\delta_1\delta_2}}(y_2-\xi_2)) \end{aligned}$$

which corresponds to the density of of a r.v distributed according to $SN_2(\boldsymbol{\xi},\Omega,\boldsymbol{\alpha})$. This can be easily proven introducing the parametrization given by

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = (1-\delta^T\bar{\Omega}^{-1}\boldsymbol{\delta})^{-1/2}\bar{\Omega}^{-1}\boldsymbol{\delta}$$

and inserting this expression in the cumulative function. The result is:

$$f_{\boldsymbol{Y}}(y) = 2\varphi_2(\boldsymbol{y}-\boldsymbol{\xi};\Omega)\Phi(\alpha^T\omega^{-1}(\boldsymbol{y}-\boldsymbol{\xi}))$$

A linear transformation applied to the vector $\boldsymbol{Y}$ leads obviously to the same results (1.41) obtained for the bivariate skew normal. It is however helpful to rewrite the same results in the parametrization given by δ. If we take

$$\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

then $Z = \boldsymbol{w}^T\boldsymbol{Y}$ is a univariate random variable with the following density:

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2\varphi(\boldsymbol{z}-\tilde{\boldsymbol{\xi}};\tilde{\Omega})\Phi(\frac{w_1\delta_1-\delta_2 w_2}{w_1\sigma_1+w_2\sigma_2}\Theta(\boldsymbol{z}-\tilde{\boldsymbol{\xi}}); 1-(\delta_1^2+\delta_2^2)\Theta)$$

where:

$$\begin{aligned} \tilde{\boldsymbol{\xi}} &= w\xi_1 + (1-w)\xi_2 \\ \tilde{\Omega} &= w^2\sigma_{11} + 2w(1-w)\sigma_{12} + (1-w)^2\sigma_{22} \\ \Theta &= \frac{w^2+(1-w)^2}{w^2+(1-w)^2+2\rho w(1-w)} \end{aligned}$$

Let's now consider a r.v distributed according to $\boldsymbol{Y} \sim \mathrm{SUN}_{2,2}(\xi, 0, \bar{\omega}, \Omega^*)$, its density has the following expression:

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2\varphi_2(\boldsymbol{y}-\boldsymbol{\xi};\Omega)\Phi_2(\Delta^T\bar{\Omega}^{-1}\omega^{-1}(\boldsymbol{y}-\boldsymbol{\xi}); I-\Delta^T\bar{\Omega}^{-1}\Delta)$$

where the parameters Ω, Δ are obtained by the opportune partition of:

$$\Omega^* = \begin{pmatrix} I & \Delta^T \\ \Delta & \bar{\Omega} \end{pmatrix}$$

and where:

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \quad ; \quad \Delta = \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \quad ; \quad \Omega = \omega\bar{\Omega}\omega$$

with

$$\omega = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \quad ; \quad \bar{\Omega} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Differently from the previous case, now Ω^* is a 4×4 matrix (in the $\text{SUN}_{2,1}$ it was 3×3). Its density can therefore be written as:

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}) &= 2\varphi_2(\boldsymbol{y} - \boldsymbol{\xi}; \Omega) \\ & \Phi_2\left(\frac{1}{1-\rho^2}\begin{pmatrix} \frac{\delta_1}{\sigma_1}(y_1-\xi_1) - \frac{\delta_1\rho}{\sigma_2}(y_2-\xi_2) \\ \frac{\delta_2}{\sigma_2}(y_2-\xi_2) - \frac{\delta_2\rho}{\sigma_1}(y_1-\xi_1) \end{pmatrix}; \begin{pmatrix} 1-\delta_1^2 & -\rho\delta_1\delta_2 \\ -\rho\delta_1\delta 2 & 1-\delta_2^2 \end{pmatrix}\right) \end{aligned}$$

It is worth mentioning that the main differences between the two bivariate **SUN** is that if $\rho = 0$ only $\text{SUN}_{2,2}$ factorizes in two independent components. In order to get the same property for $\text{SUN}_{2,1}$, is required the normality of one of the two components. Therefore a complete set of assumptions for the independence is in this case either $[\rho = 0, \delta_1 = 0]$ or $[\rho = 0, \delta_2 = 0]$.
Given a non singular matrix B of dimension $k \times k$ acting on $Y \sim \text{SUN}_{2,2}(\xi, 0, \bar{\omega}, \Omega^*)$, the random variable

$$B^T\boldsymbol{Y} \sim \text{SUN}_{2,2}(\tilde{\boldsymbol{\xi}}, 0, \tilde{\omega}, \tilde{\Omega}^*)$$

is obtained by the opportune partition of

$$\tilde{\Omega}^* = \begin{pmatrix} I & \tilde{\Delta}^T \\ \tilde{\Delta} & \tilde{\bar{\Omega}} \end{pmatrix}$$

with

$$\begin{aligned}
\tilde{\boldsymbol{\xi}} &= \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = B^T \xi \\
\tilde{\omega} &= \begin{pmatrix} \tilde{\omega}_1 & 0 \\ 0 & \tilde{\omega}_2 \end{pmatrix} = B^T \omega B \\
\tilde{\bar{\Omega}} &= \begin{pmatrix} \tilde{\bar{\Omega}}_{11} & \tilde{\bar{\Omega}}_1 \\ \tilde{\bar{\Omega}}_{12} & \tilde{\bar{\Omega}}_{22} \end{pmatrix} = B^{-1}\bar{\Omega}(B^T)^{-1} \\
\tilde{\Omega} &= \begin{pmatrix} \tilde{\Omega}_{11} & \tilde{\Omega}_1 \\ \tilde{\Omega}_{12} & \tilde{\Omega}_{22} \end{pmatrix} = B^T \Omega B = \tilde{\omega}\tilde{\bar{\Omega}}\tilde{\omega} \\
\tilde{\Delta} &= \begin{pmatrix} \tilde{\Delta}_{11} & \tilde{\Delta}_{12} \\ \tilde{\Delta}_{21} & \tilde{\Delta}_{22} \end{pmatrix} = B^{-1}\Delta
\end{aligned}$$

If we take

$$B = \begin{pmatrix} w_1 & -w_2 \\ w_2 & w_1 \end{pmatrix} \quad ; \quad |B| = w_1^2 + w_2^2 \neq 0 \; \forall \, w_1, w_2 \in \mathbb{R}$$

we obtain a bivariate random variable $\boldsymbol{Z}$ given by

$$\boldsymbol{Z} = B^T \boldsymbol{Y} = \begin{pmatrix} w_1 & w_2 \\ -w_2 & w_1 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} w_1 Y_1 + w_2 Y_2 \\ w_1 Y_2 - w_2 Y_1 \end{pmatrix} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

The marginal distribution of Z_1 is:

$$Z_1 \sim \mathrm{SUN}_{1,2}(\tilde{\xi}_1, 0, \tilde{\omega}_1, \tilde{\Omega}_1^*)$$

obtained by an opportune partition of

$$\tilde{\Omega}_1^* = \begin{pmatrix} I & \tilde{\Delta}_1^T \\ \tilde{\Delta}_1 & \tilde{\bar{\Omega}}_{11} \end{pmatrix} \quad ; \quad \tilde{\Delta}_1 = \begin{pmatrix} \tilde{\Delta}_{11} & \tilde{\Delta}_{12} \end{pmatrix}$$

and the explicit expression of the density is:

$$\begin{aligned}
f_{Z_1}(z) &= 2\varphi_2(z - \tilde{\xi}_1; \tilde{\Omega}_{11}) \\
& \quad \Phi_2\left(\begin{pmatrix} w_1\delta_1 \\ w_2\delta_2 \end{pmatrix} \Theta \frac{z - \tilde{\xi}_1}{w_1\sigma_1 + w_2\sigma_2}; I - \Theta \begin{pmatrix} w_1^2\delta_1^2 & w_1 w_2 \delta_1 \delta_2 \\ w_1 w_2 \delta_1 \delta 2 & (1-w)^2 \delta_2^2 \end{pmatrix} \right)
\end{aligned}$$

with:

$$\begin{aligned}
\tilde{\Omega}_{11} &= w_1^2 \sigma_{11} + 2 w_1 w_2 \sigma_{12} + w_2^2 \sigma_{22} \\
\Theta &= \frac{w_1^2 + w_2^2}{w_1^2 + w_2^2 + 2\rho w_1 w_2}
\end{aligned}$$

A.2 Elliptical Distributions

A d-dimensional random variable $\boldsymbol{Y}$ is elliptical if its density function is constant on ellipsoids, that is to say:

$$f_{\boldsymbol{Y}}(\boldsymbol{y};\boldsymbol{m},\Omega) = \frac{c_d}{|\Omega|^{1/2}} f_d\{(\boldsymbol{y}-\boldsymbol{m})^T\Omega^{-1}(\boldsymbol{y}-\boldsymbol{m})\} \quad \boldsymbol{y}\in\mathbb{R}^d \tag{A.1}$$

where $\boldsymbol{m}\in\mathbb{R}^d$, Ω is a covariance matrix, f_d is a suitable function from $\mathbb{R}^+$ to $\mathbb{R}^+$ called the density generator and c_d is a normalizing factor; we then write $Y \sim \mathrm{Ell}_d(\boldsymbol{m},\Omega,f_d)$.
The condition

$$\int_0^\infty y^{d/2-1} f_d(y)dy \ < \infty \tag{A.2}$$

guarantees that $f_d(y)$ is a density generator. In addition the normalizing constant c_d can be explicitly determined:

$$c_d = \frac{\Gamma(n/2)}{(2\pi)^{n/2}}\left[\int_0^\infty y^{d/2-1} f_d(y)dy\right]^{-1} \tag{A.3}$$

If $\boldsymbol{Y}$ is elliptically distributed then its characteristic function has the following form

$$\varphi_{\boldsymbol{Y}}(\boldsymbol{t}) = e^{i\boldsymbol{t}^T\boldsymbol{m}}\psi\left(\boldsymbol{t}^T\Omega\boldsymbol{t}\right). \tag{A.4}$$

The function $\psi(t)$ is called the characteristic generator. To denote elliptical laws we shall also use the notation $\boldsymbol{Y} \sim \mathrm{Ell}_d(\boldsymbol{m},\Omega,\psi)$.
The density of a normal variate is obtained by taking

$$\tilde{f}(x) = e^{-\frac{x}{2}} \ , \ c_d = (2\pi)^{-d/2}$$

The multivariate Pearson type VII distribution is obtained by

$$\tilde{f}(x) = (1+x/\nu)^{-M} \ , \ c_d = \frac{\Gamma(M)}{(\pi\nu)^{d/2}\Gamma(M-d/2)}$$

where $\nu > 0, M > d/2$ and which gives the multivariate t density in the case $M = (d+\nu)/2$.
The multivariate Pearson type II density is obtained by

$$\tilde{f}(x) = (1-x)^{-\nu} \ , \ c_d = \frac{\Gamma(d/2+\nu+1)}{(\pi)^{d/2}\Gamma(\nu+1)}$$

where $0 \le x \le 1, \nu > -1$.

Theorem A.2.1. *Suppose* $\boldsymbol{X} \in \mathbb{R}^n$ *is elliptically distributed and* A *a* $(k \times n)$ *matrix. Then the random vector* $\boldsymbol{Y} = A \cdot \boldsymbol{X}$ *is elliptically distributed.*

Proof. This can be easily proven showing that the the characteristic function of $\boldsymbol{Y}$ has the same form of the characteristic function of any subvector of $\boldsymbol{X}$ with the same number o components. □

Appendix B

Stochastic Dominance and Separation's Theorems

Proposition B.0.1. *(i)* $Y \succeq_1 X$ *iff* $X \sim Y + Z$ *for some r.v.* Z *such that* $Z \leq 0$
(ii) $Y \succeq_2 X$ *iff* $X \sim Y + Z + \epsilon$ *for some r.v.'s* Z, ϵ *such that* $Z \leq 0$ *and* $\mathbb{E}(\epsilon | Y + Z) = 0$.

Proof. :(i) we just prove the "if" part. From $X \sim Y + Z$ follows $\mathbb{E}(u(X)) = \mathbb{E}(u(Y + Z)) \leq \mathbb{E}(u(Y))$, because $Z \leq 0$ and u is increasing.
(ii) again we just prove the "if" part. Set $S = Y + Z$, we have

$$\mathbb{E}(u(X)) = \mathbb{E}(u(S + \epsilon)) = \mathbb{E}(\mathbb{E}(u(S + \epsilon)|S)) \leq \mathbb{E}(u(\mathbb{E}(S + \epsilon)|S)))$$

$$= \mathbb{E}(u(\mathbb{E}(S|S) + \mathbb{E}(\epsilon|S))) = \mathbb{E}(u(S)) \leq \mathbb{E}(u(Y)).$$

We have used in the first line the conditional Jensen inequality (which applies to concave functions) and in the second line all the remaining hypothesis (moreover we have used everywhere the basic properties of conditional expectation). □

Consider now a market of n risky assets (numbered from 1 to n). Denote by

$$\boldsymbol{R} = (R_1, \ldots, R_n)$$

the random vector of assets returns (at a future fixed time) and by $\boldsymbol{R}^T$ the corresponding column vector. Set $\mathbf{1} = (1, \ldots, 1)$, then the return of a portfolio $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$, is simply the scalar r.v.

$$\boldsymbol{\beta}^T \boldsymbol{R} = \sum_{i=1}^{n} \beta_i R_i$$

Theorem B.0.2. *The distribution of $\boldsymbol{R}$ has the (strong) 1-fund separation property iff there exist a scalar r.v. Y, a vector r.v. $\boldsymbol{\epsilon}$ and a portfolio $\boldsymbol{\alpha}$ such that*
(a) each component of $\boldsymbol{R}$ can be written as $R_i = Y + \epsilon_i$, for $i = 1, \ldots, n$
(b)it holds $\mathbb{E}(\epsilon_i|Y) = 0$
(c) the portfolio $\boldsymbol{\alpha}$ is orthogonal to the vector $\boldsymbol{\epsilon}$ (i.e. $\boldsymbol{\alpha}^T\boldsymbol{\epsilon} = 0$).

Remark: For obvious reasons Y is called the "common (risky) factor" and the noise ϵ_i the asset-specific "residual risk".

Proof. We prove the sufficiency. Using (a), and the existence of $\boldsymbol{\alpha}$ we have $\boldsymbol{\alpha}^T\boldsymbol{R} = \boldsymbol{\alpha}^T\mathbf{1}Y + \boldsymbol{\alpha}^T\boldsymbol{\epsilon}$. Therefore, by (c), $\boldsymbol{\alpha}^T\boldsymbol{R} = \boldsymbol{\alpha}^T\mathbf{1}Y = Y$. Let $\boldsymbol{\beta}$ be any portfolio and set $\boldsymbol{\eta} = \boldsymbol{\beta} - \boldsymbol{\alpha}$, then clearly $\mathbf{1}^T\boldsymbol{\eta} = 0$. Hence

$$\mathbb{E}(\boldsymbol{\beta}^T\boldsymbol{R}|\boldsymbol{\alpha}^T\boldsymbol{R}) = \mathbb{E}(\boldsymbol{\beta}^T\boldsymbol{R}|Y) \tag{B.1}$$

$$= \mathbb{E}(\boldsymbol{\alpha}^T\boldsymbol{R} + \boldsymbol{\eta}^T\boldsymbol{R}|Y) = \mathbb{E}(Y + \boldsymbol{\eta}^T\boldsymbol{\epsilon}|Y) = \mathbb{E}(Y|Y) + \sum_i \mathbb{E}(\epsilon_i|Y) = Y \tag{B.2}$$

where we have used (b). Summarizing we have $\boldsymbol{\beta}^T\boldsymbol{R} = \boldsymbol{\alpha}^T\boldsymbol{R} + \boldsymbol{\eta}\boldsymbol{R}$. Now considering PropositionB.0.1(ii) we can take $Z = 0$, moreover we have

$$\mathbb{E}(\boldsymbol{\eta}^T\boldsymbol{R}|\boldsymbol{\alpha}^T\boldsymbol{R}) = \mathbb{E}(\boldsymbol{\beta}^T\boldsymbol{R} - \boldsymbol{\alpha}^T\boldsymbol{R}|\boldsymbol{\alpha}^T\boldsymbol{R}) = \mathbb{E}(\boldsymbol{\beta}^T\boldsymbol{R}|\boldsymbol{\alpha}^T\boldsymbol{R}) - \boldsymbol{\alpha}^T\boldsymbol{R} = Y - \boldsymbol{\alpha}^T\boldsymbol{R} = 0$$

therefore all the conditions of (ii) are verified so that $\boldsymbol{\alpha}^T\boldsymbol{R} \succeq_2 \boldsymbol{\beta}^T\boldsymbol{R}$. The necessity argument requires some more efforts. □

Proposition B.0.2. *If the vector of returns is normally distributed then it satisfies conditions of Theorem 3.5.3*

Proof. Consider a market of $n+1$ assets with modified vector of returns

$$\boldsymbol{R} = (R_0, R_1, \ldots, R_n)$$

where $R_0 = R_f$ and assume it has finite mean and variance, i.e. $\mathbb{E}(\boldsymbol{R}) = E \in \mathbb{R}^{(n+1)}$ and $\mathrm{Var}(\boldsymbol{R}) = \Sigma \in \mathbb{R}^{(n+1)\times(n+1)}$.

Since $cov(R_0, R_i) = 0$ for all $i = 0, \ldots, n$ the first row and the first column of the covariance matrix Σ are made of zeros. The remaining square matrix V of dimension n is assumed to be not degenerate, i.e. $detV \neq 0$, and the vector $\tilde{\boldsymbol{R}} = (R_1, \ldots, R_n)$ such that

$$\tilde{\boldsymbol{R}} \sim N_n(\tilde{\boldsymbol{\mu}}, V)$$

where $\tilde{\boldsymbol{\mu}} = \mathbb{E}(\tilde{\boldsymbol{R}})$. We also set $\tilde{\boldsymbol{R}}_f = (R_f, \dots, R_f) \in \mathbb{R}^n$, $\boldsymbol{\xi} = \tilde{\boldsymbol{\mu}} - \tilde{\boldsymbol{R}}_f$. We shall prove that $\boldsymbol{R}$ verifies all the hypothesis of Theorem 3, that is we shall exhibit $(Y, Z, \boldsymbol{b}, \boldsymbol{\epsilon}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ such that (a) holds true for $\boldsymbol{R}$ and all the other requirements are also satisfied.
First, for all $i = 0, \dots, n$, we choose

$$Y = R_f \qquad \text{and} \qquad Z = \tilde{\boldsymbol{\nu}}^T(\tilde{\boldsymbol{R}} - \tilde{\boldsymbol{R}}_f)$$

with the vector $\tilde{\boldsymbol{\nu}}$ to be specified below.

Then for $i = 0$ we choose $b_0 = 0 = \epsilon_0$ and therefore we have $R_0 = Y + 0Z + 0 = R_f$ which is obviously true. Next, for $i = 1, \dots, n$, we wish to show that

$$R_i = R_f + b_i Z + \epsilon_i$$

holds for suitable choices of $(b_i, \tilde{\boldsymbol{\nu}}, \epsilon_i)$ and $\mathbb{E}(\epsilon_i|Z) = 0$.
Indeed define,

$$\tilde{\boldsymbol{\nu}} = (\boldsymbol{\xi}^T V^{-1} \boldsymbol{\xi})^{-1} V^{-1} \boldsymbol{\xi}$$

so that $\tilde{\boldsymbol{\nu}}^T \boldsymbol{\xi} = \boldsymbol{\xi}^T \tilde{\boldsymbol{\nu}} = 1$).

Hence $Z \sim N(1, \sigma_Z^2)$, with $\sigma_Z^2 = \tilde{\boldsymbol{\nu}}^T V \tilde{\boldsymbol{\nu}}$.
We then choose

$$b_i = \xi_i,$$

$(i = 1, \dots, n)$, and , for $i = 1, \dots, n$, we have

$$R_i = R_f + \xi_i Z + (R_i - E_0 - \xi_i Z) =: E_0 + \xi_i Z + \epsilon_i$$

with $\epsilon_i \sim N(0, \sigma_{\epsilon_i}^2)$ $(\mathbb{E}(\epsilon_i) = \xi_i - \xi_i 1 = 0)$, being a difference of two normal r.v.'s. We now show that $\mathbb{E}(\epsilon_i|E_0 + \xi Z) = \mathbb{E}(\epsilon_i|Z) = 0$. To this aim we prove ϵ_i is independent from Z, and therefore $\mathbb{E}(\epsilon_i|Z) = \mathbb{E}(\epsilon_i) = 0$. Since Z, ϵ_i are both gaussian r.v.'s their independence is equivalent to the condition $cov(\epsilon_i, Z) = 0$. Indeed, for any $\boldsymbol{\gamma} = (\gamma_0, \tilde{\gamma}) \in \mathbb{R}^{(n+1)}$, we have

$$\begin{aligned} cov(\boldsymbol{\gamma}^T \boldsymbol{\epsilon}, Z) &= cov(\tilde{\gamma}^T [\tilde{\boldsymbol{R}} - \tilde{\boldsymbol{R}}_f - \boldsymbol{\xi} Z], Z) \\ &= cov(\tilde{\gamma}^T [\tilde{\boldsymbol{R}} - \tilde{\boldsymbol{R}}_f], Z) - cov((\tilde{\gamma}^T \boldsymbol{\xi}) Z, Z) \\ &= (\tilde{\gamma}^T \Sigma \tilde{\boldsymbol{\nu}}) - cov((\tilde{\gamma}^T \boldsymbol{\xi}) \tilde{\boldsymbol{\nu}}^T \boldsymbol{R}, \tilde{\boldsymbol{\nu}}^T \boldsymbol{R}) \end{aligned}$$

$$= (\tilde{\gamma}^T \Sigma \tilde{\boldsymbol{\nu}}) - cov(\tilde{\gamma}^T (\tilde{\boldsymbol{\nu}}^T \boldsymbol{\xi}) \boldsymbol{R}, \boldsymbol{\nu}^T \boldsymbol{R}) = (\tilde{\gamma}^T \Sigma \tilde{\boldsymbol{\nu}}) - cov(\tilde{\gamma} \boldsymbol{R}, \tilde{\boldsymbol{\nu}}^T \boldsymbol{R}) = 0$$

from which $cov(\epsilon_i, Z) = 0$ follows as particular case. We now set $\boldsymbol{\nu} = (0, \tilde{\boldsymbol{\nu}})$ and consider the portfolios (the riskless portfolio and the *tangency portfolio*)

$$\boldsymbol{w}_t = \frac{\boldsymbol{\nu}}{\boldsymbol{1}^T \boldsymbol{\nu}} \quad , \quad \boldsymbol{w}_f = (1, 0, \dots, 0).$$

Since $\epsilon_0 = 0$ we have $\boldsymbol{w}_f^T \boldsymbol{\epsilon} = 0$, moreover

$$\boldsymbol{w}_t^T \boldsymbol{\epsilon} = \tilde{\boldsymbol{\nu}}^T [\tilde{\boldsymbol{R}} - \tilde{\boldsymbol{R}}_f - \boldsymbol{\xi} Z] = \tilde{\boldsymbol{\nu}}^T (\tilde{\boldsymbol{R}} - \tilde{\boldsymbol{R}}_f) - (\tilde{\boldsymbol{\nu}} \boldsymbol{\xi}) Z = Z - Z = 0$$

which shows that these two portfolios are both orthogonal to ϵ and therefore can be chosen to span the whole set of dominating portfolios. This ends the Proof $\square$

Bibliography

[1] Adcock C.J., (2004), "Capital Asset Pricing for UK Stocks Under the Multivariate Skew-Normal Distribution", in *Skew Elliptical Distributions and Their Applications: A Journey Beyond Normality*, M. Genton, Chapman and Hall, Boca Raton, pp191-204. 9, 68

[2] Arellano R. B. and A. Azzalini, (2006), "On the Unification of Families of Skew-normal Distributions", *Scandinavian Journal of Statistics 33:3*, 1999, 561574 28, 29, 103

[3] Azzalini A., (2005), "The skew-normal distribution and related multivariate families (with discussion)", *Scandinavian Journal of Statistics 32*, 2005, 159-188 (CR: 189-200) 2, 3, 28, 56

[4] Azzalini A. and A. Capitanio, (1999), "Statistical applications of the multivariate skew normal distributions", *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61*, 1999, Issue 3, 579-602 2, 26

[5] Azzalini A. and A. Dalla Valle, (1996), "The multivariate skew normal distribution", *Biometrika 83*, 1996, 715-726 1

[6] Black F. and R. Litterman, (1990), "Asset Allocation: combining investor views with market equilibrium", *Goldman Sachs Fixed Income Research*, 1990 13, 76

[7] Black F. and R. Litterman, (1992), "Global Portfolio Optimization", *Financial Analysts Journal 48:5* , September October 1992, 28 13

[8] Branco M. D. and D. K. Dey, (2001), "A general class of multivariate skew-elliptical distributions", *J. Multivariate Analysis 79*, 2001, 99-113

[9] Capitanio A., Azzalini A. and E. Stanghellini, (1996), "Graphical models for skew-normal variates", *Scandinavian Journal of Statistics 30*, 1996, 120-144 2, 25

[10] Chib S. , (1995), "Marginal likelihood from the Gibbs output", *Journal of the American Statistical Association 90*, 1995, 1313-1321

[11] Chib S. and I. Jeliazkov , (2001), "Marginal likelihood from the Metropolis Hastings output", *Journal of the American Statistical Association 96*, 2001, 270-281

[12] Fusai G. A. Meucci, (2003), "Assessing Views", *Risk*, March 2003, 18-21

[13] Gamba A. and Rossi F., (1998), "A Three-moment Based Portfolio Selection Model", *Rivista di Matematica per le Scienza Economiche e Sociali 20*, 1998, 25-48

[14] Genton M. G. (ed.), (2004), *Skew-Elliptical distributions and their applications: a journey beyond normality*, Chapman & Hall/CRC, Boca Raton 2

[15] G. Gonzlez-Faras, J. Domnguez-Molina and A. Gupta, (2003)), *The multivariate closed skew normal distribution*, Technical Report No.03-12, Department of Mathematics and Statistics, Bowling Green State University, 2003 28

[16] Hanoch G. and H. Levy, (1969), "The Efficiency Analysis of Choices Involving Risk", *Review of Economic Studies XXXVI*, July 1969, 335-346 4, 31, 32, 33

[17] He G. and R. Litterman, (1999), "The intuition beyond Black-Litterman Model Portfolios", *Investment Management Research*, Goldman Sachs & Company, December 1999 13, 78

[18] Ingersoll J., (1987), *Theory of Financial Decision Making*, Rowman & Littlefield, 1987 1, 50, 52, 68

[19] Kat H.M. and C. Brooks, (2002), "The statistical Properties of hedge Fund Index Returns and Their Implications for Investors", *Journal of Alternative Investments*, Winter 2002, 25-41 1

[20] Kraus A. and Litzenberger, (1976), "Skewness Preference and the Valuation of Risk Assets", *Journal of Finance 31*, September 1976, 1085-1099 9, 68

[21] Levy H and M. Sarnat, (1972), *Investemt and Portfolio Analysis*, 1972, New York; Wiley 4, 31

[22] Levy H, (1973), "Stochastic Dominance among Log-Normal Prospects", *International Economic Review 14:3*, October 1973, 601-614 4, 31

[23] Liseo B. and N. Loperfido, (2003), "A Bayesian interpretation of the multivariate skew-normal distribution", *Statistics and Probability Letters 61:4*, February 2003, 395-401(7) 2, 14, 28

[24] Liseo B. and N. Loperfido, (2006), "Default Bayesian analysis of the skew-normal distribution", *Journal of Statistical Planning and Inference 136:2*, 2006, 373-389.

[25] Markowitz H., (1952), "Portfolio Selection", *Journal of Finance 7*, 1952, 77-91 1, 5

[26] Markowitz H., (1991), "Foundations of Portfolio Theory", *Journal of Finance 46*, June 1991, 469-477

[27] Meucci A., (2005), *Risk and Asset Allocation*, 2005, Springer 80

[28] Meucci A., (2006), "Beyond Black-Litterman: views on non-normal markets", *Risk*, February 2006, 87-92 1, 14, 82, 83

[29] Meucci A., (2006), "Beyond Black-Litterman in practice", *Risk*, September 2006, 114-119 14, 82

[30] Mills T. C., (1995), "Modelling skewness and kurtosis in the London Stock Exchange FT-SE index return distributions", *The Statistician 44:3*, 1995, 323-332 1

[31] Rockafellar R. and S. Uryasev, (2000), "Optimization of Conditional Value at Risk", *Journal of Risk 2*, 2000, 21-41

[32] Ross S., (1978), "Mutual Fund Separation in Financial Theory - The Separating Distributions", *Journal of Economic Theory 17*, 1978 10, 11, 50, 51, 71

[33] Rubinstein, (1973), "The Foundamental Theorem of Parameter-Preference Security Valuation", *Journal of Financial and Quantitative Analysis 8*, January 1973, 61-69

[34] Sahu S. K., Dey D. K. and M.D. Branco, (2003), "A new class of multivariate skew distributions with applications to Bayesian regression models", *The Canadian Journal of Statistics 31:2*, 2003, 129-150 2, 28

[35] Simaan Y., (1993), "Portfolio Selection and Asset Pricing Three Parameter Framework", *Management Science 5*, May 1993, 578-587 7, 8, 9, 10, 37, 53, 54, 55, 62, 63, 64, 66, 67, 68

[36] Tutuncu R. and G. Cornueejols, (2006), *Optimization Methods in Finance*, Cambridge University Press, December 2006

[37] Yarii M.E, (1965), *Convexity in the Theory of Choice under Risk*, The Quarterly Journal of Economics, May 1965, 79-2, pag. 278-290