Imperial College of Science, Technology and Medicine

Department of Mathematics

Active Fixed-Income Portfolio Management using the Black-Litterman Model

Alvaro Maggiar
CID: 00573448

Submitted in partial fulfillment of the requirements for the degree of
Master of Science in Mathematics and Finance

September 2009

Abstract

In this paper we study the application of the Black-Litterman model to an active fixed-income portfolio management. We present a rigorous derivation of the model in a general setting and then move on to apply it to active management. We derive the characteristics of the fixed-income portfolio we wish to manage, comprising government bonds, inflation-linked bonds and currencies and present results derived from the application of the Black-Litterman model. We also introduce and use some risk management tools to assess the benefits of using the Black-Litterman model. Finally, we discuss a way to calculate a sound covariance matrix to be used in the optimization process.

Contents

Introduction

Portfolio optimization is a vastly discussed topic. How to best allocate one's resources among different available assets is a crucial aspect in many investment banks and has become as much a science as an art. Markowitz [Mar52] published a seminal paper in 1952 which highlighted the importance of the diversification of risk. Since then, his mean-variance optimization problem has become a standard although highly criticized. One of the most successful models to improve a portfolio manager's investment methods is the Black-Litterman model presented in 1991 by Fischer Black and Robert Litterman [BL91] while working at Goldman Sachs.

In Black and Litterman's original paper, the asset universe to be optimized was constituted of equity. However, it can be extended to a wider class of assets and we will focus on a fixed-income portfolio. The type of assets we will deal with are government bonds, inflation-linked bonds and currencies. Moreover, we are interested in an active management problem, which is to maximize the return of a portfolio with respect to a benchmark.

The structure of the paper is the following:

The first chapter of this paper will introduce the Black-Litterman model and present a rigorous derivation thereof, along with an enlightening analogy with the Kalman filter. The assumptions and notations of the model will be used throughout the paper.

Since the goal of this paper is to apply the Black-Litterman model in an active management setting, we will present in chapter 2 the changes to be made in order to do so. We will also introduce useful risk analysis tools which will help us assess the benefits of the model.

In chapter 3 we will tackle the portfolio we are interested in, that is a fixed-income portfolio comprising government bonds, inflation-linked bonds and currencies. Once properly modeled, we will be able to apply the theory built over the first two chapters and will present and comment the results obtained.

Finally, we will discuss in chapter 4 one of the crucial points of the Black-Litterman model or any portfolio optimization problem for that matter: the covariance matrix. We will present a state-dependent model to try to more accurately capture the expected covariance matrix.

Chapter 1

The Black-Litterman Model

1.1 Introduction

1.1.1 Overview

In order to overcome the pitfalls inherent to the Mean-Variance optimization problem, Black and Litterman developed in 1991 a new framework in order to construct optimal portfolios ([BL91], [BL92]). Their model, known as the Black-Litterman model (BL hereafter), combines several advantages that have given it a large success among practitioners. Compared to the mean-variance optimized (MV) portfolio presented by Markowitz [Mar52], it provides much greater stability to changes in the inputs. Indeed, one of the main criticism received by the MV portfolio is its great sensitivity to small changes in the expected returns of the assets, as well as the extreme positions it holds in some assets (Michaud [Mic89]). This forces the use of severe constraints in the optimization process in order to prevent such problems.

Furthermore, the classical mean-variance optimization problem is very demanding for it requires the forecasting of absolute returns for all the available assets. The BL model on the other hand also enables investors to plug in their own views on the market into the optimizer; but these views need not be on all the assets, nor absolute. They may be absolute or relative views on the performance of a specific assets or much larger portfolios (e.g. a sector of industry portfolio versus another). This provides with a more flexible forecasting requirement which enables both top-down and bottom-up approaches. The optimizer then blends these views with an *equilibrium portfolio* (the notion of equilibrium portfolio will be discussed in the next section) to output a new optimized portfolio. Taking into account an equilibrium portfolio which serves as origin for the optimization is what gives its robustness to the model. It will counterbalance too strong or little confident opinions, and makes great sense in particular in the case of active management with respect to a benchmark for if the investor has no views on the market, they should hold the benchmark.

In summary, the Black-Litterman model is a portfolio optimization model which smoothly incorporates the expression of particular views on the market.

We can summarize the procedure in a few steps as is presented by Bevan and Winkelmann [BW98]:

Step	**Action**	**Purpose**
1	Calculate equilibrium returns	Set neutral reference point
2	Determine weightings for views	Dampen impact of aggressive views
3	Set target tracking error	Control risk relative to benchmark
4	Set target Market Exposure	Control directional effects
5	Determine optimal portfolio weights	Find allocations that maximize performance
6	Examine risk distribution	Determine whether risk is diversified

1.1.2 Universe

In the following we consider a discrete time portfolio optimization problem, that is, the investor at time t aims at constructing a portfolio which will yield the maximum *utility* at time $t+1$. Typically, this utility function is a quadratic function maximizing the expected returns of the portfolio with a penalty for risk, measured as the variance of the portfolio.

The portfolio can be constructed by investing in n different available assets. We further assume the existence of a “risk-free” rate r_f. When referring to the excess return of an asset, we mean its return over the risk free rate. Hence, if r_i is the return of the i-th asset, its excess return will be $(r_i - r_f)$.

1.1.3 Assumptions and Objective

Before we embark on calculations, we should write down the assumptions made by the model and state the objective we wish to achieve.

Assumptions

Information The model assumes that there exist two sources of information: public and private. All actors on the market share a same set of public information at time t, noted $\mathcal{I}_t$. This information is reflected by the market. In addition to this public information, an investor may have some private information or views on the market at time t, which we note $\mathcal{G}_t$. Therefore, the total information $\mathcal{F}_t$ held by an investor at time t is generated by the union of the public and the private information: $\mathcal{F}_t = \sigma\{\mathcal{I}_t, \mathcal{G}_t\}$. Of course, information is neither lost nor forgotten, so the sets of information are increasing and are thus filtrations.

An important assumption in the model is that public and private information, $\mathcal{I}_t$ and $\mathcal{G}_t$ respectively, are *independent*.

Since we consider a one-period optimization problem, we may drop the time subscript and we will use $\mathcal{I}$ for $\mathcal{I}_t$, $\mathcal{G}$ for $\mathcal{G}_t$ and $\mathcal{F}$ for $\mathcal{F}_t$.

Distribution of expected returns Let $\boldsymbol{r}$ be the vector of excess returns. We assume that the excess returns are normally distributed with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$, where $\boldsymbol{\mu}$ is a random variable and $\boldsymbol{\Sigma}$ is the covariance matrix of the assets' returns:

$$\boldsymbol{r} \sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \tag{1.1}$$

The fact that $\boldsymbol{\mu}$ is a random variable and not a constant accounts for the fact that the expected excess returns cannot be determined with certainty. *This is the variable we want to estimate.* For instance, in the absence of any views, that is, given only $\mathcal{I}$ the public information available today, the best estimation of $\boldsymbol{\mu}$ we can give is the expected excess returns predicted by the market equilibrium: $\boldsymbol{\pi}$. This value in turn is only an estimation and not a known value, which is why we model $\boldsymbol{\mu}$ by a random variable.

The assumption on $\boldsymbol{\mu}$ is the following: we assume $\boldsymbol{\mu}$ to be normally distributed with mean $\mathbb{E}\left[\boldsymbol{\mu}\right]$ and variance $\boldsymbol{\Sigma}_\mu$:

$$\boldsymbol{\mu} \sim \mathcal{N}\left(\mathbb{E}\left[\boldsymbol{\mu}\right], \boldsymbol{\Sigma}_\mu\right) \tag{1.2}$$

Going back to the case where we only have access to the public information $\mathcal{I}$, we can narrow down our estimation of $\boldsymbol{\mu}$. As we mentioned $\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]$ has mean $\boldsymbol{\pi}$ and its variance is set to be proportional to $\boldsymbol{\Sigma}$:

$$\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}] \sim \mathcal{N}(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma}) \tag{1.3}$$

τ can be understood as a scaler of the confidence in the estimation of the market expected excess returns. If we are certain of our estimation of the market expected excess returns, then $\tau = 0$ and $\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]$ becomes deterministic: $\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}] = \boldsymbol{\pi}$. Further discussion on the meaning and value of τ will be given later.

We further assume that $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_\mu$ are uncorrelated and we can thus rewrite the distribution of $\boldsymbol{r}$ as:

$$\boldsymbol{r} \sim \mathcal{N}\left(\mathbb{E}[\boldsymbol{\mu}], \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_\mu\right) \tag{1.4}$$

Objective

Let us write $\mathcal{F}$ the information available to the investor. $\mathcal{F}$ is generated by two sets of information: the set $\mathcal{I}$ containing the public information, and $\mathcal{G}$ the set of private information known to the investor. In a mathematical form, the σ-algebra $\mathcal{F}$ is generated by $\mathcal{I}$ and $\mathcal{G}$: $\mathcal{F} = \sigma\{\mathcal{I}, \mathcal{G}\}$. As mentioned earlier, it is assumed that public and private information are independent.

Given their additional private information $\mathcal{G}$, the excess returns expected by the investor, $\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]$, are most likely going to be different from $\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]$ predicted by the equilibrium portfolio. Using this additional information $\mathcal{G}$, the investor may express some opinions about the returns of several assets, and the question then becomes: how to blend the views of the investor with the equilibrium portfolio? Differently put: how do the investor's views update the equilibrium expected excess returns?

We should stress once again that $\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]$ and $\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]$ are *random variables* and that we are looking at their distribution.

1.2 Equilibrium Approach

1.2.1 Why Equilibrium?

As briefly mentioned in the introduction, a key notion in the BL model is the so-called *equilibrium portfolio*. This equilibrium portfolio plays a central role in the optimization process and serves as a “ center of gravity”. Indeed, in the BL model, a starting portfolio is needed, and the views expressed by the portfolio manager are going to modify the positions held in their original portfolio.

This reference portfolio has been chosen by Black and Litterman to be the market capitalization weighted portfolio, i.e. the portfolio obeying the Capital Asset Pricing Model (CAPM). The choice of this particular portfolio is explained by Litterman in the first chapter of his book, “Modern Investment Management: An Equilibrium Approach” [LG03]:

> There are many approaches to investing. Ours at Goldman Sachs is an equilibrium approach. In any dynamic system, equilibrium is an idealized point where forces are perfectly balanced. In economics, equilibrium refers to a state of the world where supply equals demand. But it should be obvious even to the most casual observer that equilibrium never really exists in actual financial markets. Investors, speculators, and traders are constantly buying and selling. Prices are constantly adjusting.

Thus, although markets are not assumed to be in equilibrium, Litterman argues that similarly to a dynamical system, there are underlying forces pushing towards this state, which makes it the best possible starting point.

The meaning of this portfolio is that if an investor has no view about the market, then this is the portfolio they should hold.

1.2.2 Equilibrium Returns

In this section we will derive the excess returns implied by the the equilibrium portfolio, under the CAPM. This is done using *reverse optimization* as presented by Sharpe [Sha74]. Contrarily to the usual optimization process, where given the expected excess returns and covariance matrix of the portfolio we deduce the optimal weights; we are here given the market capitalization weights and look for the vector of excess returns that would have provided with this result.

We are going to use the following notation:

$\boldsymbol{w}$	Vector of weights invested in each asset
$\boldsymbol{w}_m$	Vector of market capitalization weights
$\boldsymbol{r}$	Vector of excess returns of the assets
$\boldsymbol{\pi}$	Vector of excess equilibrium returns
$\boldsymbol{\Sigma}$	Covariance matrix of the assets
δ	Risk aversion parameter

We will try to stick to similar notations throughout the paper, using bold letters to identify matrix notations, with small letters indicating vectors and capital letters matrices.

Markowitz' optimization problem writes:

$$\arg\max_{\boldsymbol{w}} U_{\boldsymbol{r}}(\boldsymbol{w}) = \boldsymbol{w}^T\boldsymbol{r} - \frac{\delta}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}$$

U is a utility function and thus concave and has a single global maximum. We know that $\boldsymbol{w}_m$ solves the problem when $U = U_{\boldsymbol{\pi}}$. This observation implies:

$$\begin{aligned} \frac{\partial U_{\boldsymbol{\pi}}}{\partial \boldsymbol{w}}(\boldsymbol{w}_m) &= 0 = \boldsymbol{\pi} - \delta\boldsymbol{\Sigma}\boldsymbol{w}_m \\ \boldsymbol{\pi} &= \delta\boldsymbol{\Sigma}\boldsymbol{w}_m \end{aligned} \tag{1.5}$$

Conversely, we can retrieve the equilibrium weights if we are given the expected excess returns:

$$\boldsymbol{w}_m = \frac{1}{\delta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\pi}$$

Now, multiplying equation (1.5) by $\boldsymbol{w}_m^T$ will give a scalar relation, and noting that $\boldsymbol{w}_m^T\boldsymbol{\pi} = \mathbb{E}[r_m]$, the expected excess return of the market portfolio, and $\boldsymbol{w}_m^T\boldsymbol{\Sigma}\boldsymbol{w}_m = \sigma_m^2$, the variance of the market portfolio, we find:

$$\delta = \frac{\mathbb{E}[r_m]}{\sigma_m^2} = \frac{\mathrm{SR}_m}{\sigma_m}$$

where SR_m is the Sharpe ratio of the market portfolio.

1.3 Investor's Views

1.3.1 What are views?

An investor might have information or beliefs about the evolution of the market that might differ from what is provided by the equilibrium portfolio. One of the attractions of the BL model is the possibility to take into account these views. They may come in any combination of two criteria:

Absolute or Relative: Let us consider two traded assets A and B. An absolute view on asset A would be expressed as the absolute expected excess return of A, for instance "the excess return of A is going to be 3%". A relative view on the other hand involves several assets and how they are going to perform with respect to one another, regardless of their absolute performance. This could be "A is going to outperform B by 5%".

Asset specific or Global: Let us now consider four assets A, B, C and D. In the previous paragraph we have expressed views between single assets only. This could be the view of an analyst who expresses views about specific products they have evaluated. A strategist would for instance express views on a group of assets, which can be realized in the BL framework. This could be: "a portfolio consisting of A and B will outperform a portfolio consisting of C and D".

Furthermore, predictions on the market are rarely made with a hundred percent certainty and this has to be taken into account in the model. The more confidence in a prediction, the more it should affect the model, so when expressing a view, a degree of confidence must be attached to it. For example: "A will outperform B with a 25% certainty".

1.3.2 Expressing views in the model

Once we have stated a number of views, we would like to incorporate them in the model.

Expressing views is essentially giving information about $\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]$, the expected excess returns given investor's information. In BL, the views are expressed as linear combinations of the expected excess returns. We assume that these views are normally distributed. Mathematically, this is written as:

$$\begin{aligned} \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}] &= \boldsymbol{q}+\boldsymbol{\epsilon}, \quad \text{where } \epsilon \sim \mathcal{N}(0,\boldsymbol{\Omega}) \\ \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}] &\sim \mathcal{N}(\boldsymbol{q},\boldsymbol{\Omega}) \end{aligned} \tag{1.6}$$

where $\boldsymbol{P}$ is a K-row "pick" matrix, K being the number of expressed views, q a K-vector quantifying the expected excess return of each view and $\boldsymbol{\Omega}$ a covariance matrix quantifying the uncertainty thereof.

Example 1. *Let us consider a 3 asset market, where we can invest in assets A, B and C. We have two views:*

- *A will have an excess return of 4% with 15% confidence.*
- *A and C will outperform B by 2% with 20% confidence.*

We point out that the first view is absolute with respect to a single asset whereas the second one is relative to portfolios.
Here, applying the model we would write:

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & -1 & 0.5 \end{pmatrix} \text{ and } \boldsymbol{q} = \begin{pmatrix} 4\% \\ 2\% \end{pmatrix}.$$

We will discuss $\boldsymbol{\Omega}$ later.

In the above example, when dealing with the second view which consisted in the out-performance of a portfolio over another, we have weighted its components equally. Another option would have been to weight each component in the outperforming portfolio with its capitalization weight in that portfolio. Both choices can be found in the literature, for instance Satchell and Scowcroft [SS00] use the former and Idzorek [Idz04] uses the latter.

Another way of looking at (1.6) is to regard $\boldsymbol{q}$ not as a deterministic vector, but rather as a random variable. Using this interpretation, we rewrite (1.6) as:

$$\boldsymbol{q} = \boldsymbol{P}\boldsymbol{\mu} - \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon}$ is $\mathcal{G}$-measurable and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \boldsymbol{\Omega})$

Conditioned on a realization of $\boldsymbol{\mu}$, the conditional distribution of $\boldsymbol{q}$ is thus:

$$\mathbb{E}[\boldsymbol{q}|\boldsymbol{\mu}] \sim \mathcal{N}(\boldsymbol{P}\boldsymbol{\mu}, \boldsymbol{\Omega}) \tag{1.7}$$

If we further make the assumption that $\boldsymbol{\Sigma}_\mu$ and $\boldsymbol{\Omega}$ are independent, which means that the errors in the estimation of $\boldsymbol{\mu}$ and $\boldsymbol{q}$ are uncorrelated, we retrieve the unconditional distribution of $\boldsymbol{q}$:

$$\boldsymbol{q} \sim \mathcal{N}\left(\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}], \boldsymbol{P}\boldsymbol{\Sigma}_{\boldsymbol{\mu}}\boldsymbol{P}^T + \boldsymbol{\Omega}\right) \tag{1.8}$$

Note that the independence assumption was already made in the case where $\boldsymbol{\mu}$ is estimated thanks to the public information $\mathcal{I}$ for then $\boldsymbol{\Sigma}_\mu = \tau\boldsymbol{\Sigma}$ which is $\mathcal{I}$-measurable whilst $\boldsymbol{\Omega}$ is $\mathcal{G}$-measurable and $\mathcal{I}$ and $\mathcal{G}$ are independent.

In particular, we find the conditional distribution of $\boldsymbol{q}$ given the manager's private information:

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{q}|\mathcal{G}\right] &= \boldsymbol{P}\mathbb{E}\left[\boldsymbol{\mu}|\mathcal{G}\right] - \boldsymbol{\epsilon} \\ \mathbb{E}\left[\boldsymbol{q}|\mathcal{G}\right] &\sim \mathcal{N}\left(\boldsymbol{P}\mathbb{E}\left[\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]\right], \boldsymbol{\Omega}\right) \end{aligned}$$

One major problem that needs to be tackled is how to set $\boldsymbol{\Omega}$. This is, along with τ, one of the most discussed issues of the BL model. Let us not make any assumption for now as to what $\boldsymbol{\Omega}$ should be; and proceed with the construction of the model.

1.4 Combining Views

We have heretofore built the necessary bases to the solution of the problem we wish to solve. Let us recall that our problem is to retrieve the distribution of the expected excess returns given the views we have expressed. In other words, we would like to know how the views we have expressed thanks to our private information $\mathcal{G}$ "update" the excess returns expected by the market given only the public information $\mathcal{I}$. In a more mathematical notation, we are looking to derive an expression for the distribution function of the random variable:

$$\mathbb{E}\left[\mathbb{E}\left[\boldsymbol{\mu}|\mathcal{I}\right]|\boldsymbol{q}\right] \tag{1.9}$$

In their original article from 1992 [BL92, p.35], Black and Litterman suggest two methods for deriving the solution to the problem. The first is based on a sampling theory approach and the second on Bayesian theory:

> We can think of a view as representing a fixed number of observations drawn from the distribution of future returns. In this case we follow the "mixed estimation" strategy described in Theil (1971). Alternatively we can think of the view as directly reflecting a subjective distribution for the expected excess returns.

Of course, both methods lead to the same result, but having several means to reach the same solution provides with a better insight into a sometimes little intuitive problem.

The more common derivation in the literature is probably the Bayesian approach to the problem. Such derivations can be found in [Chr02], [SS00], [Meu08] or [Che09] among many. Theil's mixed estimation approach is not short of derivations either, we refer to [FKPF07], [BF03], [Wal09] or [TS09]. Finally, although explicitely mentioned in Black and Litterman's paper, it seems that the only derivation of the model using a strict sampling theory approach can be found in [Man06].

We will now present a Bayesian derivation and one based on a comparison of the problem with the Kalman filter. Indeed, although there is scarce mention of it in the literature, the BL problem can be interpreted as a particular filtering problem which nicely fits in the Kalman filter, as was suggested to me by Pr. Mark Davis. The only reference I have found linking the BL model to the Kalman filter is [Wei07][1].

1.4.1 Bayesian Theory Approach

We are going to use Bayes' rule to first derive the distribution function of $\mathbb{E}\left[\boldsymbol{\mu}|\boldsymbol{q}\right]$ and then use it in the case where we use $\mathbb{E}\left[\boldsymbol{\mu}|\mathcal{I}\right]$ for $\boldsymbol{\mu}$.
For any random variable X, we will write f_X its distribution function, and we recall that when applied to the distribution functions of two random variables X and Y, Bayes' rule reads:

$$f_{\mathbb{E}[Y|X]}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} \tag{1.10}$$

Substituting Y with $\boldsymbol{\mu}$ and X with $\boldsymbol{q}$, we get:

$$f_{\mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}]}(\boldsymbol{\mu}) = \frac{f_{\boldsymbol{q},\boldsymbol{\mu}}(\boldsymbol{q},\boldsymbol{\mu})}{f_{\boldsymbol{q}}(\boldsymbol{q})} \tag{1.11}$$

We first derive an expression for $f_{\boldsymbol{q},\boldsymbol{\mu}}$ using the fact that it can be rewritten as:

$$f_{\boldsymbol{q},\boldsymbol{\mu}}(\boldsymbol{q},\boldsymbol{\mu}) = f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) f_{\mathbb{E}[\boldsymbol{q}|\boldsymbol{\mu}]}(\boldsymbol{q})$$

and that according to (1.3):

$$f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) = \frac{|\boldsymbol{\Sigma}_\mu|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}(\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}])^T \left(\boldsymbol{\Sigma}_\mu\right)^{-1}\left(\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}]\right)\right)$$

and according to (1.7):

$$f_{\mathbb{E}[\boldsymbol{q}|\boldsymbol{\mu}]}(\boldsymbol{q}) = \frac{|\boldsymbol{\Omega}|^{-\frac{1}{2}}}{(2\pi)^{\frac{K}{2}}} \exp\left(-\frac{1}{2}(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu})\right)$$

[1] In an email exchange, Dr. Weinberger wrote me that he was not aware of any other reference either.

These observations yield:

$$f_{\boldsymbol{q},\boldsymbol{\mu}}(\boldsymbol{q},\boldsymbol{\mu}) \propto$$
$$\exp\left[-\frac{1}{2}\Big((\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}])^T(\boldsymbol{\Sigma}_\mu)^{-1}(\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}])+(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu})\Big)\right]$$

We have got rid of the constant for simplicity and will study it towards the end of the calculation.

Turning back to equation (1.11), we would like to express $f_{\boldsymbol{q},\boldsymbol{\mu}}$ as the product of a function of $\boldsymbol{q}$ only - which will be $f_{\boldsymbol{q}}$ the distribution function of $\boldsymbol{q}$ - and a function of $\boldsymbol{\mu}$ and $\boldsymbol{q}$ which will be the desired distribution function $f_{\mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}]}$.

Aiming at achieving this, we reshuffle the terms in the exponential above to make the decomposition apparent:

$$\begin{aligned}
& (\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}])^T(\boldsymbol{\Sigma}_\mu)^{-1}(\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}])+(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu}) \\
= \; & \boldsymbol{\mu}^T(\boldsymbol{\Sigma}_\mu)^{-1}\boldsymbol{\mu}-2\boldsymbol{\mu}^T(\boldsymbol{\Sigma}_\mu)^{-1}\mathbb{E}[\boldsymbol{\mu}]+\mathbb{E}[\boldsymbol{\mu}]^T(\boldsymbol{\Sigma}_\mu)^{-1}\mathbb{E}[\boldsymbol{\mu}] \\
& \qquad +\boldsymbol{q}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}-2\boldsymbol{\mu}^T\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}+\boldsymbol{\mu}^T\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\boldsymbol{\mu} \\
= \; & \boldsymbol{\mu}^T\Big((\boldsymbol{\Sigma}_\mu)^{-1}+\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\Big)\boldsymbol{\mu}-2\boldsymbol{\mu}^T\Big((\boldsymbol{\Sigma}_\mu)^{-1}\mathbb{E}[\boldsymbol{\mu}]+\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}\Big) \\
& \qquad +\mathbb{E}[\boldsymbol{\mu}]^T(\boldsymbol{\Sigma}_\mu)^{-1}\mathbb{E}[\boldsymbol{\mu}]+\boldsymbol{q}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}
\end{aligned} \tag{1.12}$$

We have used the fact that $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}_\mu$ are covariance matrices and are thus symmetric so $\boldsymbol{\Omega}^T=\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}_\mu^T=\boldsymbol{\Sigma}_\mu$. Similarly, we may use the fact that the transpose of a scalar is itself.

We write:

$$\begin{aligned}
\boldsymbol{K} &= (\boldsymbol{\Sigma}_\mu)^{-1}+\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P} \\
\boldsymbol{m} &= (\boldsymbol{\Sigma}_\mu)^{-1}\mathbb{E}[\boldsymbol{\mu}]+\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q} \\
a &= \mathbb{E}[\boldsymbol{\mu}]^T(\boldsymbol{\Sigma}_\mu)^{-1}\mathbb{E}[\boldsymbol{\mu}]+\boldsymbol{q}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}
\end{aligned}$$

Using these notations in (1.12) yields:

$$\begin{aligned}
& (\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}])^T(\boldsymbol{\Sigma}_\mu)^{-1}(\boldsymbol{\mu}-\mathbb{E}[\boldsymbol{\mu}])+(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\mu}) \\
= \; & \boldsymbol{\mu}^T\boldsymbol{K}\boldsymbol{\mu}-2\boldsymbol{m}^T\boldsymbol{\mu}+a \\
= \; & \boldsymbol{\mu}^T\boldsymbol{K}^T\boldsymbol{K}\boldsymbol{K}^{-1}\boldsymbol{\mu}-2\boldsymbol{m}^T\boldsymbol{K}\boldsymbol{K}^{-1}\boldsymbol{\mu}+a \\
= \; & (\boldsymbol{K}\boldsymbol{\mu}-\boldsymbol{m})^T\boldsymbol{K}^{-1}(\boldsymbol{K}\boldsymbol{\mu}-\boldsymbol{m})-\boldsymbol{m}^T\boldsymbol{K}^{-1}\boldsymbol{m}+a \\
= \; & \left(\boldsymbol{\mu}-\boldsymbol{K}^{-1}\boldsymbol{m}\right)^T\boldsymbol{K}\left(\boldsymbol{\mu}-\boldsymbol{K}^{-1}\boldsymbol{m}\right)-\boldsymbol{m}^T\boldsymbol{K}^{-1}\boldsymbol{m}+a
\end{aligned}$$

Plugging this result into the exponential gives:

$$\begin{aligned} f_{\boldsymbol{q},\boldsymbol{\mu}}(\boldsymbol{q},\boldsymbol{\mu}) &\propto \exp\left[-\frac{1}{2}\Big((\boldsymbol{K}\boldsymbol{\mu}-\boldsymbol{m})^T\boldsymbol{K}^{-1}(\boldsymbol{K}\boldsymbol{\mu}-\boldsymbol{m})-\boldsymbol{m}^T\boldsymbol{K}^{-1}\boldsymbol{m}+a\Big)\right] \\ &\propto \exp\left[-\frac{1}{2}\left(a-\boldsymbol{m}^T\boldsymbol{K}^{-1}\boldsymbol{m}\right)\right] \\ &\qquad \exp\left[-\frac{1}{2}\Big(\left(\boldsymbol{\mu}-\boldsymbol{K}^{-1}\boldsymbol{m}\right)^T\boldsymbol{K}\left(\boldsymbol{\mu}-\boldsymbol{K}^{-1}\boldsymbol{m}\right)\Big)\right] \end{aligned}$$

Provided correct transformation of the constant, we would like to identify the first part of the expression above with the distribution function of $\boldsymbol{q}$ which would allow us to state that the second part corresponds to the distribution function of $\mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}]$.

According to (1.8), $\boldsymbol{q}$ abides by the following distribution:

$$\boldsymbol{q} \sim \mathcal{N}\left(\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}], \boldsymbol{P}\boldsymbol{\Sigma}_{\boldsymbol{\mu}}\boldsymbol{P}^T+\boldsymbol{\Omega}\right)$$

Therefore, we would like to identify $(\boldsymbol{q}-\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])^T\left(\boldsymbol{P}\boldsymbol{\Sigma}_{\mu}\boldsymbol{P}^T+\boldsymbol{\Omega}\right)^{-1}(\boldsymbol{q}-\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])$ with $(a-\boldsymbol{m}^T\boldsymbol{K}^{-1}\boldsymbol{m})$, which would lead us to conclude that:

$$\mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}] \sim \mathcal{N}(\boldsymbol{K}^{-1}\boldsymbol{m}, \boldsymbol{K}^{-1})$$

The tedious transformation of $(a-\boldsymbol{m}^T\boldsymbol{K}^{-1}\boldsymbol{m})$ as well as that of the constant are relegated to appendix A.1.

Thanks to these calculations, we may assert that $\mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}]$ follows a normal distribution with mean $\boldsymbol{K}^{-1}\boldsymbol{m}$ and variance $\boldsymbol{K}^{-1}$, i.e.:

$$\begin{aligned} \mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}] &\sim \mathcal{N}(\boldsymbol{K}^{-1}\boldsymbol{m}, \boldsymbol{K}^{-1}) \qquad (1.13) \\ \mathbb{E}\left[\mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}]\right] &= \Big((\boldsymbol{\Sigma}_{\mu})^{-1}+\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\Big)^{-1}\left((\boldsymbol{\Sigma}_{\mu})^{-1}\mathbb{E}[\boldsymbol{\mu}]+\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}\right) \\ \mathrm{Var}\left[\mathbb{E}[\boldsymbol{\mu}|\boldsymbol{q}]\right] &= \Big((\boldsymbol{\Sigma}_{\mu})^{-1}+\boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\Big)^{-1} \end{aligned}$$

Finally, we deduce the desired result from these last expressions. Our initial wish was to derive the distribution of $\mathbb{E}\left[\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]|\boldsymbol{q}\right]$, the "updated" value of the mean of the equilibrium expected excess returns given the expressed views. This is a particular case of the above results, it suffices to change $\boldsymbol{\mu}$ into $\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]$, which also implies according to (1.3) that $\mathbb{E}[\boldsymbol{\mu}]=\boldsymbol{\pi}$ and $\boldsymbol{\Sigma}_{\mu}=\tau\boldsymbol{\Sigma}$. Substituting these values into (1.13) provides with the result:

$$\begin{aligned}
\mathbb{E}\left[\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]|\boldsymbol{q}\right] &\sim \mathcal{N}(\boldsymbol{\mu}_{BL}, \boldsymbol{\Sigma}^{\mu}_{BL}) \\
\boldsymbol{\mu}_{BL} &= \Big((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\Big)^{-1}\Big((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}\Big) \\
\boldsymbol{\Sigma}^{\mu}_{BL} &= \Big((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\Big)^{-1}
\end{aligned} \tag{1.14}$$

1.4.2 Kalman Filter Approach

In 1960, Rudolph Kalman published a seminal paper describing the solution to a discrete linear filtering problem [Kal60]. He tackled the problem of trying to estimate the unobservable state of a discrete time process thanks to the knowledge of a particular measurement of the system. The applications of the Kalman filter cover an extremely wide range of engineering areas. We refer to [WB06] for an introduction to the filter.

The Kalman Filter

In its common presentation, the Kalman filter addresses the problem of estimating the state $\boldsymbol{\mu}$ of a discrete-time process given only $\boldsymbol{q}$, a measurement of the system.

The discrete time linear system is given by:

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{A}_t\boldsymbol{\mu}_t + \boldsymbol{B}_t\boldsymbol{\pi}_t + \boldsymbol{\sigma}_t \tag{1.15}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\pi}$ are n-vectors and $\boldsymbol{\sigma} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\mu_t})$. $\boldsymbol{\sigma}$ represents the noise of the process.

The matrix $\boldsymbol{A}$ relates the state of the system at time $t+1$ to its state at time t, and $\boldsymbol{\pi}$ represents the inputs of the system. We should mention that these inputs were not included in the original presentation of the filter, but it is straightforward to include them.

The measurement is given by the linear relation:

$$\boldsymbol{q}_{t+1} = \boldsymbol{P}_{t+1}\boldsymbol{\mu}_{t+1} + \boldsymbol{\epsilon}_{t+1} \tag{1.16}$$

where $\boldsymbol{q}$ is a K-vector and $\boldsymbol{\epsilon}_{t+1} \sim \mathcal{N}(0, \boldsymbol{\Omega}_{t+1})$. $\boldsymbol{\epsilon}$ is the measurement noise.

It is assumed that the process and measurement noises $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$ are independent.

We write $\mathcal{H}_t$ the total information observed up to and including t: $\mathcal{H}_t = \sigma\{\boldsymbol{q}_0, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_t\}$. If we assume that the distribution function of $\mathbb{E}[\boldsymbol{\mu}_t|\mathcal{H}_t]$ is normal, and the previous assumptions hold, then the distribution function of $\mathbb{E}\left[\boldsymbol{\mu}_{t+1}|\mathcal{H}_{t+1}\right]$ will also be normal.

A graphical representation of this system can be seen below in Figure 1.1.

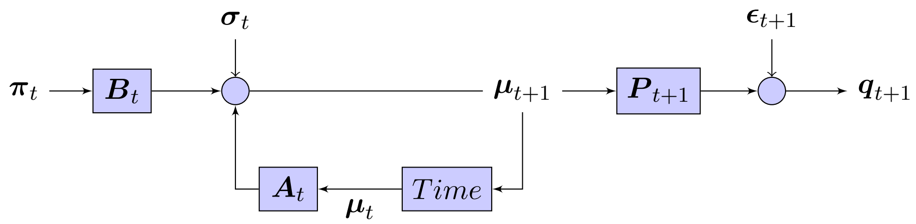


Figure 1.1: Kalman system

A derivation of the results can be found in [BH97], we will only present here the main ideas of the problem.

The algorithm of the Kalman filter comprises two main phases: predict and update the estimate of the next state of the system. We therefore distinguish two estimates, the prior estimate $\boldsymbol{e}_{t+1}^-$, and the posterior or updated estimate $\boldsymbol{e}_{t+1}^+$.

Predict The prior estimate $\boldsymbol{e}_{t+1}^-$ of $\boldsymbol{\mu}_{t+1}$ is a raw estimate using only the linear equation (1.15) relating successive states of the system. Assuming we know the initial state of the system $\boldsymbol{\mu}_0$, we can define a sequence of *unbiased* estimates of the state of the system at each time:

$$\begin{aligned} \boldsymbol{e}_{t+1}^- &= \boldsymbol{A}_t \boldsymbol{e}_t^- + \boldsymbol{B}_t \boldsymbol{\pi}_t \\ \boldsymbol{e}_0^- &= \boldsymbol{\mu}_0 \end{aligned} \tag{1.17}$$

From (1.17) we deduce the prior residual of the estimate $\boldsymbol{e}_{t+1}^-$:

$$\begin{aligned} \boldsymbol{r}_{t+1}^- &= \boldsymbol{e}_{t+1}^- - \boldsymbol{\mu}_{t+1} \\ &= \boldsymbol{A}_t \left(\boldsymbol{e}_t^- - \boldsymbol{\mu}_t\right) - \boldsymbol{\sigma}_t \\ &= \boldsymbol{A}_t \boldsymbol{r}_t^- - \boldsymbol{\sigma}_t \end{aligned}$$

Since $(\boldsymbol{e}_t^-)$ is a sequence of unbiased estimates, $\boldsymbol{r}^-$ has mean zero and thus its covariance matrix $(\boldsymbol{C}_t^-)$ propagates according to:

$$\begin{aligned} \boldsymbol{C}_{t+1}^- &= \mathbb{E}\left[(\boldsymbol{r}_{t+1}^-)(\boldsymbol{r}_{t+1}^-)^T\right] \\ &= \boldsymbol{A}_t\mathbb{E}\left[(\boldsymbol{r}_t^-)(\boldsymbol{r}_t^-)^T\right]\boldsymbol{A}_t^T + \mathbb{E}\left[\boldsymbol{\sigma}_t\boldsymbol{\sigma}_t^T\right] \\ &= \boldsymbol{A}_t\boldsymbol{C}_t^-\boldsymbol{A}_t^T + \boldsymbol{\Sigma}_{\mu_t} \end{aligned}$$

The predictive step hence provides with the following information:

$$\begin{aligned} \boldsymbol{e}_{t+1}^- &= \boldsymbol{A}_t\boldsymbol{e}_t^- + \boldsymbol{B}_t\boldsymbol{\pi}_t \\ \boldsymbol{r}_{t+1}^- &= \boldsymbol{A}_t\boldsymbol{r}_t^- + \boldsymbol{\sigma}_t \\ \boldsymbol{C}_{t+1}^- &= \boldsymbol{A}_t\boldsymbol{C}_t^-\boldsymbol{A}_t^T + \boldsymbol{\Sigma}_{\mu_t} \end{aligned}$$

Update In the predictive step, we have only used part of the information we have for we have not made use of the measurement information. The goal of the updating step is to improve that raw estimate thanks to the additional information on $\boldsymbol{\mu}$ brought by $\boldsymbol{q}$.

Similarly to the previous phase, we define a posterior estimate $\boldsymbol{e}_{t+1}^+$ of $\boldsymbol{\mu}_{t+1}$, which is the estimate found after incorporating the information contained in $\boldsymbol{q}_t$. We associate to this estimate its residual error $\boldsymbol{r}_{t+1}^+$ and the covariance thereof $\boldsymbol{C}_{t+1}^+$. The Kalman filter aims at minimizing the error in the posterior estimate, which becomes a minimum mean-square error estimator. It yields the following results:

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{\mu}_{t+1}|\mathcal{H}_{t+1}\right] &\sim \mathcal{N}(\boldsymbol{e}_{t+1}^+, \boldsymbol{C}_{t+1}^+) \\ \boldsymbol{e}_{t+1}^+ &= \boldsymbol{e}_{t+1}^- + \boldsymbol{K}_{t+1}\left(\boldsymbol{q}_{t+1} - \boldsymbol{P}_t\boldsymbol{\pi}_t\right) \\ \boldsymbol{r}_{t+1}^+ &= \left(\boldsymbol{I} - \boldsymbol{K}_{t+1}\boldsymbol{P}_{t+1}\right)\boldsymbol{r}_{t+1}^- - \boldsymbol{\epsilon}_{t+1} \\ \boldsymbol{C}_{t+1}^+ &= \left(\boldsymbol{I} - \boldsymbol{K}_{t+1}\boldsymbol{P}_{t+1}\right)\boldsymbol{C}_{t+1}^- \end{aligned}$$

where $\boldsymbol{K}_{t+1}$ is defined as the Kalman Filter gain and writes:

$$\boldsymbol{K}_{t+1} = \boldsymbol{C}_{t+1}^-\boldsymbol{P}_{t+1}\left(\boldsymbol{\Omega}_{t+1} + \boldsymbol{P}_{t+1}\boldsymbol{C}_{t+1}^-\boldsymbol{P}_{t+1}^T\right)^{-1} \tag{1.18}$$

Connection with the Black-Litterman model

We have presented the main results of the Kalman filter and we are now going to show that the Black Litterman model can be regarded as a one step Kalman filter.

Indeed, the problem the portfolio manager faces is to estimate at period t the excess returns of the assets, $\boldsymbol{\mu}$, over the next period, period $t+1$. At time t, we may give a raw estimate of $\boldsymbol{\mu}$ which is simply $\mathbb{E}\left[\boldsymbol{\mu}|\mathcal{I}\right] \sim \mathcal{N}(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma})$. This would correspond to the predictive step of the Kalman filter. Next, we may consider that the private information the manager holds, and which allows them to formulate their views, is simply a measurement of the state of the market at time $t+1$. Taking all these observations into consideration, we see that with the following equivalences, the BL model perfectly fits in the Kalman filter:

Kalman filter		Black Litterman
$\boldsymbol{\mu}_{t+1}$	$\longleftrightarrow$	$\boldsymbol{\mu}$
$\boldsymbol{A}_t$	$\longleftrightarrow$	$\boldsymbol{0}$
$\boldsymbol{B}_t$	$\longleftrightarrow$	$\boldsymbol{I}$
$\boldsymbol{\pi}_t$	$\longleftrightarrow$	$\boldsymbol{\pi}$
$\boldsymbol{\Sigma}_{\mu_t}$	$\longleftrightarrow$	$\tau\boldsymbol{\Sigma}$
$\boldsymbol{q}_{t+1}$	$\longleftrightarrow$	$\boldsymbol{q}$
$\boldsymbol{P}_{t+1}$	$\longleftrightarrow$	$\boldsymbol{P}$
$\boldsymbol{\epsilon}_{t+1}$	$\longleftrightarrow$	$\boldsymbol{\epsilon}$
$\boldsymbol{\Omega}_{t+1}$	$\longleftrightarrow$	$\boldsymbol{\Omega}$

With these substitutions, we then have:

$$\begin{aligned} \boldsymbol{\mu} &= \boldsymbol{\pi} + \boldsymbol{\sigma}, & \boldsymbol{\sigma} \sim \mathcal{N}(0, \tau\Sigma) \\ \boldsymbol{q} &= \boldsymbol{P}\boldsymbol{\mu} + \boldsymbol{\epsilon}, & \boldsymbol{\epsilon} \sim \mathcal{N}(0, \Omega) \end{aligned}$$

which leads to:

$$\begin{aligned} \boldsymbol{e}^- &= \boldsymbol{\pi} \\ \boldsymbol{C}^- &= \tau\boldsymbol{\Sigma} \end{aligned}$$

and:

$$\begin{aligned} \boldsymbol{e}^+ &= \boldsymbol{\pi} + \boldsymbol{K}\left(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\pi}\right) \\ \boldsymbol{C}^+ &= \left(\tau\boldsymbol{\Sigma} - \boldsymbol{K}\boldsymbol{P}\tau\boldsymbol{\Sigma}\right) \end{aligned}$$

where the Kalman filter gain $\boldsymbol{K}$ is now:

$$\boldsymbol{K} = \boldsymbol{\pi} + \tau\boldsymbol{\Sigma}\boldsymbol{P}^T\left(\boldsymbol{P}\tau\boldsymbol{\Sigma}\boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1}$$

So writing $\boldsymbol{\mu}_{BL}$ and $\boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}}$ the mean and variance of the posterior estimate, we find:

$$\begin{array}{rcl} \mathbb{E}\left[\boldsymbol{\mu}|\boldsymbol{q}\right] & \sim & \mathcal{N}(\boldsymbol{\mu}_{BL}, \boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}}) \\ \boldsymbol{\mu}_{BL} & = & \boldsymbol{\pi} + \tau\boldsymbol{\Sigma}\boldsymbol{P}^T\left(\boldsymbol{P}\tau\boldsymbol{\Sigma}\boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1}(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\pi}) \\ \boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}} & = & \tau\boldsymbol{\Sigma} - \tau\boldsymbol{\Sigma}\boldsymbol{P}^T\left(\boldsymbol{P}\tau\boldsymbol{\Sigma}\boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}\tau\boldsymbol{\Sigma} \end{array}$$

We will see in the next section that these results are equivalent to those found in (1.4.1).

1.5 The Black-Litterman Formula

Whichever the selected approach, we find the same final results.

First of all, we find the distribution of the optimized expected excess returns given the vector of views $\boldsymbol{q}$:

$$\mathbb{E}\left[\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]|\boldsymbol{q}\right] \sim \mathcal{N}\left(\boldsymbol{\mu}_{BL}, \boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}}\right) \tag{1.19}$$

where

$$\boldsymbol{\mu}_{BL} = \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right)^{-1}\left((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{q}\right) \tag{1.20}$$

$$\boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}} = \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right)^{-1} \tag{1.21}$$

Using formula (1.4), we retrieve the optimized expected excess returns knowing the views:

$$\mathbb{E}\left[\mathbb{E}[\boldsymbol{r}|\mathcal{I}]|\boldsymbol{q}\right] \sim \mathcal{N}\left(\boldsymbol{\mu}_{BL}, \boldsymbol{\Sigma}_{BL}\right) \tag{1.22}$$

where

$$\boldsymbol{\Sigma}_{BL} = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}} \tag{1.23}$$

We show in appendix (A.2) that equivalent ways to write $\boldsymbol{\mu}_{BL}$ and $\boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}}$ are:

$$\boldsymbol{\mu}_{BL} = \boldsymbol{\pi} + \boldsymbol{\Sigma}\boldsymbol{P}^T\left(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T + \frac{\boldsymbol{\Omega}}{\tau}\right)^{-1}(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\pi}) \tag{1.24}$$

$$\boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}} = \tau\boldsymbol{\Sigma} - \tau^2\boldsymbol{\Sigma}\boldsymbol{P}^T\left(\tau\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}\boldsymbol{\Sigma} \tag{1.25}$$

Via these transformations we have proved that the results from the Bayesian approach (1.4.1) and the Kalman filter approach (1.4.2) are indeed equivalent.

Equation (1.24) is interesting because written in this form, the role of $\boldsymbol{\pi}$ as "center of gravity" appears clearly. The new vector of expected excess returns is the sum of the equilibrium excess returns plus an additional term reflecting the views of the investor. How much weight is put on the views depends on $\boldsymbol{\Omega}$ and τ which are unfortunately not straightforward parameters; we discuss their meaning and how to set them in the next section. Nonetheless, all other things the same, we see that the greater τ or the smaller $\boldsymbol{\Omega}$, the more the views are going to be taken into account, and inversely the smaller τ or the bigger $\boldsymbol{\Omega}$, the less the views will matter and the more the weights are going to shift towards the equilibrium weights.

1.6 Discussions of tau and Omega

1.6.1 tau

There is little while very diverse information in the literature as to how to set τ. τ can be interpreted as a scalar weighting the confidence in the equilibrium expected excess returns estimation. Therefore, some authors including Black and Litterman [BL92] consider that since "*the uncertainty in the mean is much smaller than the uncertainty in the return itself*", it should set close to zero. In this case, τ generally takes values between 0.01 and 0.05.

Other authors such as Satchell and Scowcroft [SS00] state that τ can be set to 1 or not set at all in the case of He and Litterman [HL99]. This can find its explanation in the fact that in the BL formula (1.24), τ only appears as a factor of Ω and could thus be inserted into Ω, delegating the estimation of τ to an estimation of $\frac{\Omega}{\tau}$.

Finally, another approach by Bevan and Winkelman [BW98] in the case of active management is to set τ such that the information ratio (a performance measure) does not exceed 2.0. This leads in practice to values of τ between 0.5 and 0.7.

In any case, it must be remembered that setting τ very close to 0 will yield the equilibrium portfolio and conversely, a very high value of τ will shift all the weight towards the investor's views. A sound procedure would be to check for the sensitivity of the model with respect to τ.

1.6.2 Omega

Let us recall that according to (1.2), $\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}] \sim \mathcal{N}\left(\mathbb{E}[\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]], \boldsymbol{\Sigma}_{\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]}\right)$. Then, since $\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}] \sim \mathcal{N}(\boldsymbol{q}, \boldsymbol{\Omega})$ (cf. (1.6)), a direct calculation yields:

$$\boldsymbol{\Omega} = \boldsymbol{P}\boldsymbol{\Sigma}_{\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]}\boldsymbol{P}^T$$

However, $\boldsymbol{\Sigma}_{\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]}$ is the covariance matrix of the excess returns expected by the investor and is unknown, which prevents us from deriving an expression for $\boldsymbol{\Omega}$. It is therefore necessary to make further assumptions and guess an approximative structure for $\boldsymbol{\Omega}$.

Several approaches can be found in the literature:

- The first approach is to assume that the views are mutually independent. This assumption is very simplifying for in this case $\boldsymbol{\Omega}$ will become a diagonal matrix:

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \omega_k^2 \end{pmatrix}$$

  However, the terms on the diagonal ω_i still need to be evaluated. There is no definite answer as to how to handle this problem. A common solution is to interpret ω_i as an interval around the estimated value $\boldsymbol{q}_i$ of the corresponding view. We have assumed that the expected returns given the private information are normal, so using the normal distribution properties, we know that there is a 66% chance that the value will lie between the mean minus the standard deviation and the mean plus the standard deviation. Therefore, we can identify ω_i to play the role of the standard deviation of the view. For instance if we believe that asset A is going to have an excess return of between 100bp and 200bp with a 95% confidence, we can express that view as believing that A is going to have an excess return of 150bp with a standard deviation of 25bp for 95% of the normal distribution lies within ± 2 standard deviations.

- Another approach proposed by Meucci [Meu05, ch.9] is to make an assumption on $\boldsymbol{\Sigma}_{\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]}$ similar to the one made on $\boldsymbol{\Sigma}_{\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]}$. Recall that in (1.3) we assumed that $\boldsymbol{\Sigma}_{\mathbb{E}[\boldsymbol{\mu}|\mathcal{I}]}$ was proportional to $\boldsymbol{\Sigma}$. Meucci makes the same assumption on $\boldsymbol{\Sigma}_{\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}]}$ with a proportionality coefficient of $\left(\frac{1}{c}-1\right)$, where c takes values in $(0,1]$.He consequently writes:

$$\boldsymbol{\Omega} \equiv \left(\frac{1}{c}-1\right)\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T$$

The scalar c varies according to the absolute confidence in the investor's skills. $c \equiv 1$ neutralizes the volatility of the views and hence confers complete confidence to the investor's skills, whereas a c close to 0 produces an infinitely disperse distribution of the views, denying the investor any skill. An in between value of $c \equiv \frac{1}{2}$ corresponds to an equal confidence in the investor and in the market.

- He and Litterman [HL99] propose a method which can be perceived as a mix of the previous two. They suggest that $\boldsymbol{\Omega}$ should be diagonal, implying that the views are not correlated like in the first method; and with terms on the diagonal proportional to the prior distribution, like in Meucci's method. In fact, the terms on the diagonal of $\boldsymbol{\Omega}$ should be those on the diagonal of $\boldsymbol{P}\boldsymbol{\tau}\boldsymbol{\Sigma}\boldsymbol{P}^T$, so that:

$$\boldsymbol{\Omega} \equiv \mathrm{diag}(\boldsymbol{P}\boldsymbol{\tau}\boldsymbol{\Sigma}\boldsymbol{P}^T)$$

This choice for $\boldsymbol{\Omega}$ is rather convenient, for when substituted in the BL formula (1.24), it cancels out τ and leaves:

$$\boldsymbol{\mu}_{BL} = \boldsymbol{\pi} + \boldsymbol{\Sigma}\boldsymbol{P}^T\left(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1}(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\pi}) \tag{1.26}$$

Where $\boldsymbol{\Omega}$ is now $\mathrm{diag}(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T)$, thus making the model independent of τ.

- Other approaches can be found. For instance Idzorek [Idz04] provides with a mean for the investor to specify their confidence in the views as a percentage. Pezier [Pez07] introduces a Black-Litterman Singular model which derives from He and Litterman's approach.

1.7 Summary

Assumptions and inputs :

- There are two sources of information: public $\mathcal{I}$ and private $\mathcal{G}$.

- $\boldsymbol{w}_m$, vector of market capitalization weights.
- $\boldsymbol{\Sigma}$, covariance matrix of the assets.
- $\boldsymbol{r}$, vector of excess returns:

$$\begin{aligned} \boldsymbol{r} &\sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ \text{where } \boldsymbol{\mu} &\sim \mathcal{N}(\mathbb{E}[\boldsymbol{\mu}], \boldsymbol{\Sigma}_\mu) \end{aligned}$$

- knowledge on the distribution of $\boldsymbol{\mu}$:

Public: $\mathcal{I}$	Private: $\mathcal{G}$
$\mathbb{E}[\boldsymbol{\mu}\|\mathcal{I}] \sim \mathcal{N}(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma})$	$\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}\|\mathcal{G}] \sim \mathcal{N}(\boldsymbol{q}, \boldsymbol{\Omega})$

Procedure :

Equilibrium Returns Compute the vector of equilibrium returns $\boldsymbol{\pi}$:

$$\boldsymbol{\pi} = \delta\boldsymbol{\Sigma}\boldsymbol{w}_m$$

Express views

$$\begin{aligned} \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}] &= \boldsymbol{q} + \boldsymbol{\epsilon}, \quad \text{where } \epsilon \sim \mathcal{N}(0, \boldsymbol{\Omega}) \\ \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}|\mathcal{G}] &\sim \mathcal{N}(\boldsymbol{q}, \boldsymbol{\Omega}) \end{aligned}$$

Optimized Expected Returns Compute the optimized expected returns and its variance using the BL formulae:

$$\begin{aligned} \boldsymbol{\mu}_{BL} &= \boldsymbol{\pi} + \boldsymbol{\Sigma}\boldsymbol{P}^T \left(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T + \frac{\boldsymbol{\Omega}}{\tau}\right)^{-1} (\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\pi}) \\ \boldsymbol{\Sigma}_{BL} &= (1+\tau)\,\boldsymbol{\Sigma} - \tau^2\boldsymbol{\Sigma}\boldsymbol{P}^T \left(\tau\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1} \boldsymbol{P}\boldsymbol{\Sigma} \end{aligned}$$

Solve problem Solve the optimization problem using the optimized expected returns:

$$\begin{aligned} & \arg\max_{\boldsymbol{w}} \; \boldsymbol{w}^T\boldsymbol{\mu}_{BL} - \frac{\delta}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}_{BL}\boldsymbol{w} \qquad (1.27) \\ s.t. \quad & \boldsymbol{w}^T\boldsymbol{1} = 1 \\ & \text{other constraints} \end{aligned}$$

One important point that should be brought to attention regards the covariance matrix of the returns. In the framework developed thus far, the expected covariance matrix of the returns is a byproduct of the model. However, most author's and practitioners do not use the covariance matrix produced by the model and use their own prior covariance matrix instead.

The Black-Litterman model is therefore only used to derive the vector of expected returns, maintaining the two main issues in portfolio optimization separated: expected returns on the one side and covariance matrix on the other. The problem of calculating the covariance matrix of the returns is thus left as an entire other issue.

As a consequence, the optimization problem (1.27) should be rewritten as:

$$\begin{aligned} & \arg\max_{\boldsymbol{w}} \ \boldsymbol{w}^T \boldsymbol{\mu}_{BL} - \frac{\delta}{2} \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \\ s.t. \quad & \boldsymbol{w}^T \mathbf{1} = 1 \\ & \text{other constraints} \end{aligned}$$

where $\boldsymbol{\Sigma}$ is the prior covariance matrix.

Chapter 2

BL and Active Management

2.1 Introduction

The BL framework used so far focused on maximizing the classical mean/variance problem, i.e. maximizing the Sharpe Ratio of the portfolio. Many investors do not manage their portfolio with that objective in mind, but rather have a benchmark target they wish to outperform. In this context, it does not make sense to use the equilibrium portfolio as a starting point any more, for it would be much more relevant to have the benchmark portfolio take this role instead. This is legitimized when we think at the case where we have no particular information other than the public information. In this case, the best we can do is simply match the performance of the benchmark, which is achieved by holding the benchmark portfolio.

This section partly uses from Roll's 1992 paper [Rol92], which underlined the discrepancy between minimizing the volatility of tracking error and maximizing the mean/variance efficiency of a portfolio; and from S. da Silva, Lee and Pornrojnangkool [SdSLP09] and Herold [Her03] who have looked at BL under an active management perspective.

2.2 Formulation of the new problem

We will use the following notations:

$\boldsymbol{w}_b$	vector of benchmark weights
$\boldsymbol{w}_a$	vector of active weights
$\boldsymbol{w}$	vector of portfolio weights ($= \boldsymbol{w}_a + \boldsymbol{w}_b$)
$\boldsymbol{w}_{GMV}$	vector of the Global Minimum Variance portfolio weights
$\boldsymbol{\mu}$	vector of expected excess returns
μ_{GMV}	expected excess returns of the GMV portfolio ($= \boldsymbol{\mu}^T \boldsymbol{w}_{GMV}$)
λ	active risk aversion parameter
δ	absolute risk aversion parameter

The notations above mention the Global Minimum Variance Portfolio (GMV). It is the efficient portfolio which has minimum variance, i.e.:

$$\left.\begin{array}{l} \boldsymbol{w}_{GMV} = \arg\min_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \\ s.t. \quad \boldsymbol{w}^T \mathbf{1} = 1 \end{array}\right\} \boldsymbol{w}_{GMV} = \frac{\mathbf{1}^T \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}}$$

In the previous section, the goal was to express the expected excess returns of assets in order to use them in the following problem:

$$\begin{aligned} & \arg\max_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{\mu} - \frac{\delta}{2} \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} \\ s.t. \quad & \boldsymbol{w}^T \mathbf{1} = 1 \\ & \text{other constraints} \end{aligned} \tag{2.1}$$

In the active management setting in the presence of a benchmark, we no longer wish to solve this problem but rather to maximize the active returns with a penalty on the square of active risk:

$$\begin{aligned} & \arg\max_{\boldsymbol{w}} \boldsymbol{w}_a^T \boldsymbol{\mu} - \frac{\lambda}{2} \boldsymbol{w}_a^T \boldsymbol{\Sigma} \boldsymbol{w}_a \\ s.t. \quad & \text{constraints} \end{aligned} \tag{2.2}$$

We can see a first difference between the two problems. The former problem aims at maximizing the expected excess returns (over the risk-free rate) of the portfolio with a penalty for its volatility, which comes down to maximizing its Sharpe Ratio; while the latter aims at maximizing the expected excess returns over the benchmark with a penalty for the tracking error variance, which comes down to maximizing the Information Ratio. This difference in the objectives has some consequences, as [SdSLP09] show.

2.3 Particular case

In order to get a sense of what can differ in this new problem setting, let us solve (2.2) in the particular case where we have the constraint $\boldsymbol{w}_a^T \mathbf{1} = 0$. This is a self-financing condition, it corresponds to the case where the investor simply readjusts their portfolio positions, without investing nor withdrawing any value.

The Lagrangian of the problem is:

$$L(\boldsymbol{w}_a, \nu) = \boldsymbol{w}_a^T \boldsymbol{\mu} - \frac{\lambda}{2} \boldsymbol{w}_a^T \boldsymbol{\Sigma} \boldsymbol{w}_a - \nu \boldsymbol{w}_a$$

The first derivative of the Langrangian with respect to $\boldsymbol{w}_a$ must be set to 0:

$$\frac{\partial L}{\partial \boldsymbol{w}_a}(\boldsymbol{w}_a, \nu) = \boldsymbol{\mu} - \lambda \boldsymbol{\Sigma} \boldsymbol{w}_a - \nu \mathbf{1}$$

$$\boldsymbol{w}_a = \frac{1}{\lambda} \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\mu} - \nu \mathbf{1}\right) \tag{2.3}$$

The budget constraint yields:

$$\frac{1}{\lambda} \left(\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \mathbf{1} - \nu \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}\right) = 0$$

$$\nu = \frac{\mathbf{1}^T \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}} \boldsymbol{\mu} = \boldsymbol{w}_{GMV}^T \boldsymbol{\mu} = \mu_{GMV}$$

Substituting in (2.3) finally gives:

$$\boldsymbol{w}_a = \frac{1}{\lambda} \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{I} - \mathbf{1} \boldsymbol{w}_{GMV}^T\right) \boldsymbol{\mu} \tag{2.4}$$

This solution is intriguing. It makes pairwise comparisons between the expected excess returns and the GMV portfolio expected excess return, and if they are different (and they are very likely to be) it will induce non-null active positions. This is true even if no view has been expressed, when we would have expected nothing to happen. This unfortunate consequence is tackled in the next section.

2.4 Change of referential

We recall the formula (1.24) for the expected excess returns obtained in the first section:

$$\boldsymbol{\mu}_{BL} = \boldsymbol{\pi} + \tau \boldsymbol{\Sigma} \boldsymbol{P}^T \left(\tau \boldsymbol{P} \boldsymbol{\Sigma} \boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1} \left(\boldsymbol{q} - \boldsymbol{P} \boldsymbol{\pi}\right)$$

A naive approach to implementing BL in an active management setting would be to use directly this formula in (2.4). However, this would not be the right way to proceed.

As we have seen in the previous section, and as S. da Silva, Lee and Pornrojnangkool [SdSLP09] show, doing this would imply unintended trades, even in the case where one does not have any views, if the equilibrium portfolio is not chosen to be the GMV portfolio. They claim that the root of this inconsistency lies in the mismatch between the objectives in the optimization problems (2.1) and (2.2) which are a to maximize the Sharpe Ratio in the former and the Information Ratio in the latter.

When justifying the choice of the equilibrium portfolio as a reference for the optimization (c.f. section 1.2.1), Litterman made an analogy with a dynamic system. I believe this analogy can be carried on in the active management setting. Here, we are assumed to hold a benchmark. This benchmark moves, its expected returns are continuously changing and we are studying the returns of some movements relative to this moving referential. This is similar to studying dynamics in a non Galilean referential, a change of referential is necessary. Indeed, in the previous setting, we were looking at expected excess returns *over the risk-free rate* whereas we are now interested in the expected excess returns *over the benchmark*. That means that the origin of the previous study was the risk-free rate and it must now be set to the benchmark; and in this new referential, the equilibrium vector must be null. The meaning of this is that when we do not have any view on the market, and thus we do not hold any active position, our expected excess return *over the benchmark* will be the equilibrium vector, which is zero. But in absolute, the expected excess returns *over the risk-free rate* will be the same as the benchmark.

The consequence of this is that when applying BL, we must set $\boldsymbol{\pi} = \mathbf{0}$ in the Black-Litterman formula:

$$\boldsymbol{\mu}_{BL}^{a} = \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^{T}\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right)^{-1}\left(\boldsymbol{P}^{T}\boldsymbol{\Omega}^{-1}\boldsymbol{q}\right)$$

This is how Winkelmann [LG03, ch.13], Herold [Her03] and Rachev, Hsu, Bagasheva and Fabozzi [RHBF08, ch.8] treat the problem. Indeed, we may directly consider the alphas instead of the excess returns over the benchmark, which is equivalent since we are starting from a portfolio equal to the benchmark and which thus has a beta equal to one. We can therefore rewrite the above formula as:

$$\boldsymbol{\alpha}_{BL} = \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^{T}\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right)^{-1}\left(\boldsymbol{P}^{T}\boldsymbol{\Omega}^{-1}\boldsymbol{q}\right) \tag{2.5}$$

$$\boldsymbol{\alpha}_{BL} = \boldsymbol{\Sigma}\boldsymbol{P}^{T}\left(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^{t} + \frac{\boldsymbol{\Omega}}{\tau}\right)^{-1}\boldsymbol{q} \tag{2.6}$$

We stress out that the vector $\boldsymbol{q}$ now contains predictions on the *alphas*.

2.5 Implementation

2.5.1 General Case

Once predictions have been made and we have a new vector of expected *alphas* according to (2.6), we consider the corresponding active management problem:

$$\begin{aligned} & \arg\max_{\boldsymbol{w}_a} \boldsymbol{w}_a^T \boldsymbol{\alpha}_{BL} - \frac{\lambda}{2} \boldsymbol{w}_a^T \boldsymbol{\Sigma} \boldsymbol{w}_a \\ s.t. \quad & \text{constraints} \end{aligned} \tag{2.7}$$

We may imagine different sort of constraints. For instance we could bound the total active weights by absolute bounds, or by bounds proportional to the benchmark portfolio weight; we could also consider bounding some or all of the particular weights on some assets. Those are some among many possibilities.

We also point out that the above problem could be substituted by a maximization of the expected alpha subject to a risk budget constraint, i.e. a constraint on the tracking error.

The views and information of a portfolio manager may come in two distinct ways, quantitative and qualitative. When expressed quantitatively, the views enable a direct application of the above formulae. Another case is that of qualitative approach, in which a manager only expresses preferences about some assets with respect to one or several others. A typical example is the expression of a ranking of the assets in the portfolio, ranging from a strong sell opinion to a strong buy opinion, with possibly a stratification of the assets along different groups. The BL model is flexible enough to enable such a case.

2.5.2 Qualitative views

In this section we will investigate how the BL model can treat the case where we only express qualitative views.

A qualitative view can be stated as going long one portfolio and going short another: going long the bearish portfolio and going short the bullish one. The resulting long/short portfolio induces a tracking error that can be located on the diagonal of the matrix $\boldsymbol{P\Sigma P}^T$. The manager then expresses their degree of confidence in that view. This information is captured by the Information Coefficient (IC), which is a number between 0 and 1 and that can be understood as the probability of that particular view to happen.

When referring to a manager, the information coefficient gives an information as to how often their predictions have been correct and thus gives an indication of their skill. A high IC denotes a high confidence and inversely for a small IC.

Grinold and Kahn [GK00] present their forecasting rule of thumb for active management:

$$\mathrm{IR} \approx \mathrm{IC}.\sqrt{\mathrm{BR}}$$

where:

IR	Information Ratio
IC	Information Coefficient
BR	Strategy's Breadth

The Breadth is defined as the number of independent forecasts of exceptional return made per year for a strategy ([GK00, ch.6]).

We recall that the Information Ratio is defined as the ratio of active return to active risk, i.e. alpha divided by the tracking error:

$$\begin{aligned} \mathrm{IR} &= \frac{\alpha}{\sigma_a} \\ \text{where: } \alpha &= \boldsymbol{\alpha}^T \boldsymbol{w}_a \\ \sigma_a &= \sqrt{\boldsymbol{w}_a^T \boldsymbol{\Sigma} \boldsymbol{w}_a} \end{aligned}$$

It follows that we can express the investor's expected alphas, $\boldsymbol{q}$, as the product of their IR times the tracking error:

$$\begin{aligned} \boldsymbol{q}_i &= \boldsymbol{\sigma}_{ai}.\mathrm{IC}_i.\sqrt{\mathrm{BR}}_i \\ \text{where } \boldsymbol{\sigma}_{ai} &= \sqrt{\mathrm{diag}(\boldsymbol{P\Sigma P}^T)_i} \end{aligned} \tag{2.8}$$

$\boldsymbol{\sigma}_{ai}$ is the tracking error of the i-th view, IC_i and BR_i its respective Information Coefficient and Breadth.

Equation (2.8) shows that in order to state a view on a strategy, it suffices to provide a corresponding information coefficient. As we mentioned, the tracking error $\boldsymbol{\sigma}_{ai}$ is located on the diagonal of the matrix $\boldsymbol{P\Sigma P}^T$, and the breadth is already known.

It remains to set the values of τ and $\boldsymbol{\Omega}$. In the previous section, we discussed several options regarding how to do this, and a convenient way in the present setting is to use He and Litterman's method and set $\boldsymbol{\Omega}$ to be a diagonal matrix with diagonal terms equal to $\mathrm{diag}(\boldsymbol{P\tau\Sigma P}^T)$, thus cancelling out τ.

The role of $\boldsymbol{\Omega}$ is to curb the risk. Indeed, the terms on its diagonal are the volatility of the tracking error of each view. Thus, the higher the volatility of one view and hence the riskier, the more it is going to be dampened by $\boldsymbol{\Omega}$. This is going to ensure that the risks undertaken by each view will be spread more evenly amongst them, and that most of the risk is not borne by only one or only a few.

We may consider the case where $\boldsymbol{\Omega}$ is null; this is the case of perfect foresight, or complete certainty in each of the views. In this case, it suffices to assume an Information Ratio for each asset and compute the expected returns that fit them perfectly. Here, equation (2.6) would simply become $\boldsymbol{P\alpha} = \boldsymbol{q}$. This is approach is treated in the literature and also used in practice, it can be found for instance in Lee and Lam (Credit Suisse) [LL01], or in the Optimal Risk Budgeting with Skill (ORBS) approach of Dynkin, Gould, Hyman, Konstantinovsky and Phelps (Lehman Brothers) [DGH$^+$06, ch.24]. We will present some examples in order to compare both methodologies and evaluate the importance of forecast uncertainty.

We summarize the procedure for qualitative forecasts:

$$\begin{aligned} \mathbb{E}[\boldsymbol{\alpha}|\mathcal{I}] &\sim \mathcal{N}(0, \tau\boldsymbol{\Sigma}) \\ \boldsymbol{P}\mathbb{E}[\boldsymbol{\alpha}|\mathcal{G}] &\sim \mathcal{N}(\boldsymbol{q}, \boldsymbol{\Omega}) \\ \text{where} \quad \boldsymbol{q}_i &= \boldsymbol{\sigma}_{ai}.\mathrm{IC}_i.\sqrt{\mathrm{BR}_i} \\ \boldsymbol{\Omega} &\equiv \mathrm{diag}(\boldsymbol{P\Sigma P}^T) \end{aligned}$$

Given this information, BL yields the following expression for the expected alphas:

$$\boldsymbol{\alpha}_{BL} = \boldsymbol{\Sigma P}^T \left(\boldsymbol{P\Sigma P}^T + \boldsymbol{\Omega}\right)^{-1} \boldsymbol{q}$$

(note that we have here cancelled out τ and that $\boldsymbol{\Omega} \equiv \mathrm{diag}(\boldsymbol{P\Sigma P}^T)$)

With this alphas in hand, we may then turn to the optimization problem:

$$\begin{aligned} &\arg\max_{\boldsymbol{w}_a} \boldsymbol{w}_a^T \boldsymbol{\alpha}_{BL} - \frac{\lambda}{2}\boldsymbol{w}_a^T \boldsymbol{\Sigma w}_a \\ s.t. \quad &\text{constraints} \end{aligned}$$

2.6 Risk Analysis

An important part of the optimization process lies in the ability to assess and correct the risks undertaken as a consequence of the new positions. This is achieved in two steps: first by analysing the coherence of the expressed views used as inputs in the problem, and second by analysing the importance of each view in the total undertaken risk (active risk).

2.6.1 Views Consistency

One of the priorities in the optimization process is to analyse the coherence and consistency of the views. Views may be expressed by many different experts and in many different ways, so it can be sometimes hard to embrace all of them at one single glance and be able to detect suspicious ones. There are two ways in which views can be assessed: individually and globally. Indeed, we must check whether the views actually reflect our opinion and that we have not unintentionally expressed contradictory views. That is done by making pairwise analysis of the views thanks to their covariance matrix. However, after this first step is successfully taken and we have checked individual views, it does not mean that the overall view they amount to is acceptable. Each individual view could be perfectly possible but their combination might be very unlikely because too far from the consensus expected behaviour of the market.

Individual assessment :
The covariance of the assets are given by the matrix $\mathbf{\Sigma}$. Since we express the views as a linear combinations of these assets via the "pick" matrix $\boldsymbol{P}$, the variance-covariance matrix of the views is given by $\boldsymbol{P\Sigma P}^T$. As we mentioned, the diagonal terms of $\boldsymbol{P\Sigma P}^T$ are the volatilities of the tracking error for each view, and the non-diagonal terms are the covariance between the views. It is then easy to deduce the correlation matrix of the views given their variance-covariance matrix.

Analysing the coherence of the views is achieved by looking at their correlation matrix. A high positive correlation indicates that the views are consistent with historical data of the market. It means that we have expressed the same "ideas" as to how the market is going to behave. If most of the views are positively correlated, then it means that there is little room for diversification for their implied alphas should move in the same way, all up or all down.

On the other hand, strongly negatively correlated views indicate that we are somehow betting against market historical data, and that portfolios which used to move accordingly will now drift apart, which accounts for the manager going long one portfolio and short the other.

Finally, little correlated views present diversification opportunities since movements in one of the views portfolio are unlikely to affect the other.

Global assessment :
The second step in assessing the views is to check for their overall consistency. Fusai and Meucci [FM03] introduced a way to achieve this using the Mahalanobis distance. We adapt their method to the case of active management. The principle is to measure how far from the consensus expected returns we have landed. In the case of active management, the consensus expected returns are the alphas of the benchmark portfolio, which are null. The Mahalanobis distance $M(\boldsymbol{q})$ of the optimized alphas provided by the BL model $\boldsymbol{\alpha}_{BL}(\boldsymbol{q})$ is given by:

$$M(\boldsymbol{q})^2 = \boldsymbol{\alpha}_{BL}(\boldsymbol{q})^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha}_{BL}(\boldsymbol{q}) \tag{2.9}$$

According to Mahalanobis' law, $M(\boldsymbol{q})^2$ approximately follows a chi-square distribution with N degrees of freedom where N is the number of assets. So we can assign a probability to the vector or forecasts $\boldsymbol{q}$ such that:

$$\mathbb{P}(\boldsymbol{q}) = 1 - F_{\chi^2}\left(M(\boldsymbol{q})^2, N\right) \tag{2.10}$$

where $F_{\chi^2}(\ , N)$ is the cumulative probability of the chi-square distribution with N degrees of freedom.

The manager sets a probability threshold under which they reject the current vector of forecasts $\boldsymbol{q}$ for being too extreme. For instance, they could decide that $\boldsymbol{q}$ must be revised if $\mathbb{P}(\boldsymbol{q})$ falls under, say, 0.95.

In the case where the threshold is not overcome, the manager must readjust their forecasts. The next question is which forecast to modify, and to what extent? The answer to this question is given by analyzing the sensibility of the probability $\mathbb{P}(\boldsymbol{q})$ to $\boldsymbol{q}$. Computing the derivative of $\mathbb{P}(\boldsymbol{q})$ with respect to $\boldsymbol{q}$ will yield the more influential forecasts on $\mathbb{P}(\boldsymbol{q})$. Applying the chain rule, we get:

$$\begin{aligned} \frac{\partial \mathbb{P}(\boldsymbol{q})}{\partial \boldsymbol{q}} &= \frac{\partial \mathbb{P}(\boldsymbol{q})}{\partial M(\boldsymbol{q})^2} \frac{\partial M(\boldsymbol{q})^2}{\partial \boldsymbol{\alpha}_{BL}(\boldsymbol{q})} \frac{\partial \boldsymbol{\alpha}_{BL}(\boldsymbol{q})}{\partial \boldsymbol{q}} \\ &= -2 f_{\chi^2}\left(M(\boldsymbol{q})^2, N\right) \left(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T + \boldsymbol{\Omega}\right)^{-1} \boldsymbol{P}\boldsymbol{\alpha}_{BL}(\boldsymbol{q}) \end{aligned} \tag{2.11}$$

where $f_{\chi^2}(\ ,N)$ is the probability density function of the chi-square distribution with N degrees of freedom.

Once this derivative has been computed, the manager needs to find the entry with the highest absolute value, and hence the one with the most influence on $\mathbb{P}(\boldsymbol{q})$. If that entry is positive (resp. negative), then the forecast for that view must be decreased (resp. increased).

The procedure must then be repeated to check whether the tuning of the forecasts has been sufficient. Repeating this operation as many times as necessary will guarantee to reach the required threshold.

This procedure provides with a means to assess *post-optimization* returns. However, we know that the BL model can be seen as a method minimizing the distance between the manager's expected returns and the market expected expected returns. Now, consider a negligent manager who expresses equally strong opinions but in opposite directions to linked assets. We might think for instance of the US dollar and the HK dollar whose correlation is almost 1 since the HK dollar is indexed on the US dollar. Expressing opposite opinions on these two currencies would be careless, but our manager accidentally does so, thinking that the US dollar will provide an alpha of $-\alpha_v$ and the HK dollar an alpha of $+\alpha_v$. In minimizing the distance between these views and the benchmark expected alphas - which are null -, the BL model will somehow annihilate them and yield alphas close to zero for both of them. If we then applied the above method, the optimized alphas might satisfy the threshold condition, and we would not necessarily realize how poor a prediction we have made.

Therefore, it might be a sound procedure to apply the same methodology to the vector of raw forecasts $\boldsymbol{q}$ in order to assess the overall consistency of the prior forecasts.

In this case we would calculate:

$$M'(\boldsymbol{q})^2 = \boldsymbol{q}^T \left(\boldsymbol{P\Sigma P}^T\right)^{-1} \boldsymbol{q} \tag{2.12}$$

and the corresponding probability:

$$\mathbb{P}'(\boldsymbol{q}) = 1 - F_{\chi^2}\left(M'(\boldsymbol{q})^2, K\right) \tag{2.13}$$

where K is the number of expressed views.

2.6.2 Risk Contributions

Once the views have been checked and fed to the optimizer, the manager will be provided with the active positions $\boldsymbol{w}_a$ recommended by the optimizer. This new positions will induce some risk, but some in more extent than the others. Intuitively and logically, the manager will be willing to take more risk in the assets and portfolios upon which they have a stronger view. However, they will probably not want to take more risk in a portfolio if they do not have any particular preference for it. Therefore, if the manager were to have an equally strong feeling about all of their view portfolios, then they would want the risk to be approximately evenly spread among them.

The tools used to assess the risks inherent to active positions on the assets - $\boldsymbol{w}_a$ - and on the view portfolios - $\boldsymbol{w}_{v_a}$ - are the Marginal Contribution to Tracking Error (MCTE) and Percentage Contribution to Tracking Error (PCTE) also called Absolute Marginal Contribution to Tracking Error (ACTE) and Relative Marginal Contribution to Tracking Error (RCTE) respectively. We derive the formulae for the assets, it is then straightforward to deduce them for the views.

The MCTE of an asset is the amount by which the tracking error will increase when the asset's weight is increased by a small amount. In other words, the MCTE of an asset is the sensibility of the tracking error to its active weight, hence the derivative of the tracking error with respect to its active weight. The vector of MCTEs is thus:

$$\begin{aligned} \mathbf{MCTE} &= \frac{\partial \sigma_a}{\partial \boldsymbol{w}_a} && (2.14) \\ &= \frac{\partial \sqrt{\boldsymbol{w}_a^T \boldsymbol{\Sigma} \boldsymbol{w}_a}}{\partial \boldsymbol{w}_a} \\ \mathbf{MCTE} &= \frac{\boldsymbol{\Sigma} \boldsymbol{w}_a}{\sigma_a} && (2.15) \end{aligned}$$

We recall that $\boldsymbol{w}_a^T \boldsymbol{\Sigma} \boldsymbol{w}_a = \sigma_a^2$, so multiplying (2.15) by $\boldsymbol{w}_a^T$ yields:

$$\begin{aligned} \boldsymbol{w}_a^T \mathbf{MCTE} &= \frac{\boldsymbol{w}_a^T \boldsymbol{\Sigma} \boldsymbol{w}_a}{\sigma_a} = \sigma_a \\ \text{i.e.} \quad \sum_i \boldsymbol{w}_{a_i} \mathbf{MCTE}_i &= \sigma_a && (2.16) \end{aligned}$$

where we use the subscript i to denote the values corresponding to the i-th asset.

We see in (2.16) that the tracking error can be written as a weighted sum of the assets' MCTEs where the weights are the respective assets' active weights. We rewrite (2.16) as:

$$\sum_i \frac{\boldsymbol{w}_{a_i}\mathbf{MCTE}_i}{\sigma_a} = 1 \tag{2.17}$$

Under this formulation, the relative contribution of each asset to the total tracking error appears clearly as $\frac{\boldsymbol{w}_{a_i}\mathbf{MCTE}_i}{\sigma_a}$ which we define to be the PCTE of the i-th asset:

$$\mathbf{PCTE}_i = \frac{\boldsymbol{w}_{a_i}\mathbf{MCTE}_i}{\sigma_a} \tag{2.18}$$

We may give an other interpretation to the PCTEs rewriting them as:

$$\mathbf{PCTE}_i = \frac{\partial\sigma_a/\sigma_a}{\partial\boldsymbol{w}_{a_i}/\boldsymbol{w}_{a_i}} \tag{2.19}$$

The PCTEs can thus be seen as the relative change in tracking error given a relative change in active weights.

We have derived the expressions of the MCTEs and PCTEs relative to the *assets*, but a manager would find more relevant to compute these values relative to the *views* for these are the inputs they used in the optimization process. Knowing how much a view affects their risk is of utmost importance.

We write $\boldsymbol{w}_{v_a}$ the views' active weights and use the subscript j to denote the values relative to the j-th view. We recall that the tracking error can also be computed as:

$$\sigma_a = (\boldsymbol{w}_a^T\boldsymbol{\Sigma}\boldsymbol{w}_a)^{\frac{1}{2}} = (\boldsymbol{w}_{v_a}^T\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T\boldsymbol{w}_{v_a})^{\frac{1}{2}} \tag{2.20}$$

Therefore, it is straightforward to write:

$$\mathbf{MCTE}_v = \frac{\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T\boldsymbol{w}_{v_a}}{\sigma_a} \tag{2.21}$$

Similarly, we find:

$$\mathbf{PCTE}_{v_j} = \frac{\boldsymbol{w}_{v_{a_j}}\mathbf{MCTE}_{v_j}}{\sigma_a} \tag{2.22}$$

2.7 Worked Examples

In this section we will present some examples to illustrate how the BL model applies in active management with qualitative forecasts. We will use the diagnostic tools introduced earlier to evaluate how risk is handled, in particular with respect to the case of perfect foresight.

We consider an investor who may invest in 6 different countries: Australia, Canada, France, Germany, Japan and the United Kingdom. They use a benchmark whose weights in each country are shown in Table 2.1 [1].

Country	Weight (in %)
Australia	7.14
Canada	8.63
France	12.80
Germany	9.52
Japan	29.46
UK	32.44

Table 2.1: Benchmark Weights

View 1	AU vs (FR & JP)	bullish
View 2	CA vs GE	bearish
View 3	FR vs JP	bullish
View 4	CA vs FR	bearish

Table 2.2: Investor Views

The investor has several views expressed in Table 2.2. The corresponding 4×5 view matrix $\boldsymbol{P}$ in the BL model is the following:

$$\begin{array}{cc} & \begin{array}{cccccc} \text{AU} & \text{CA} & \text{FR} & \text{GE} & \text{JP} & \text{UK} \end{array} \\ \boldsymbol{P} = & \begin{pmatrix} 1 & 0 & -0.3028 & 0 & -0.6972 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

The rows correspond to the views and the columns to the weights of each asset in that view. Note that the weights of each portfolio sum up to 0 and that in the first view, where more than two assets are involved, we have used benchmark-weighted weights.

It might be interesting to look at the correlation matrix of the views to apply the analysis we mentioned earlier. The views' correlation matrix is presented in Table 2.3. We first observe that no two views are negatively correlated, meaning that the investor is not betting against historical movements and correlations. Secondly, it is noticeable that views 2 and 4 are

[1]The origin of the data and the covariance matrix can be found in appendix B.1

strongly correlated. Finally, view 1 is virtually uncorrelated to views 2 and 4 and little correlated to view 3, which allows for risk diversification; we therefore expect a stronger active allocation to that view than to the others which present stronger correlations.

	View 1	View 2	View 3	View 4
View 1	1	0.03602	0.2895	0.0509
View 2		1	0.2069	0.9150
View 3			1	0.2694
View 4				1

Table 2.3: Views' Correlations

2.7.1 First Case: Equal Preferences

The first case we consider is one where the investor has equal confidence in each one of their views and therefore gives them a same level of Information Coefficient which we set to 0.2. In order to compute the investor's expected returns, we further assume a Breadth of 1. As discussed earlier, we choose to set $\mathbf{\Omega} \equiv \text{diag}\boldsymbol{P\Sigma P}^T$ for convenience.

They wish to optimize active risk allocation subject to an active risk constraint on the tracking error of, for example, 100bp. The results obtained after optimization are showed in Table 2.4. We see that the risk is very well spread across the different views, with almost equal contributions. As predicted, view 1 received a greater allocation and bears a little more risk than the others for it is less correlated to them.

		$\mathbf{\Omega} \equiv \text{diag}\boldsymbol{P\Sigma P}^T$		$\mathbf{\Omega} \equiv 0$	
View	IC	$\boldsymbol{w}_{v_a}$	$\mathbf{PCTE}_v$	$\boldsymbol{w}_{v_a}$	$\mathbf{PCTE}_v$
1	0.2	2.00%	26.76%	2.04%	32.43%
2	0.2	1.35%	24.51%	1.45%	24.49%
3	0.2	1.14%	25.85%	1.32%	35.24%
4	0.2	1.24%	22.88%	0.47%	7.84%

Table 2.4: Optimized Portfolio (IC= 0.2 for all views)

Let us now investigate the model where the confidence in the views is not accounted for. In this case, $\mathbf{\Omega}$ is set to 0 and the expected alphas are now calculated as:

$$\boldsymbol{\alpha}_{\boldsymbol{\Omega}=0} = \boldsymbol{\Sigma}\boldsymbol{P}^T\left(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T\right)^{-1}\boldsymbol{q}$$

The corresponding results are presented in Table 2.4. They present some important differences with respect to the previous case. Views 1 and 3 carry almost 70% of the total risk while the fourth view only accounts for under 8%. Overall, the positions are sharper than when the confidence in the views was taken into account. This underlines the importance of using the confidence in the views in order to obtain a more balanced and less risky portfolio.

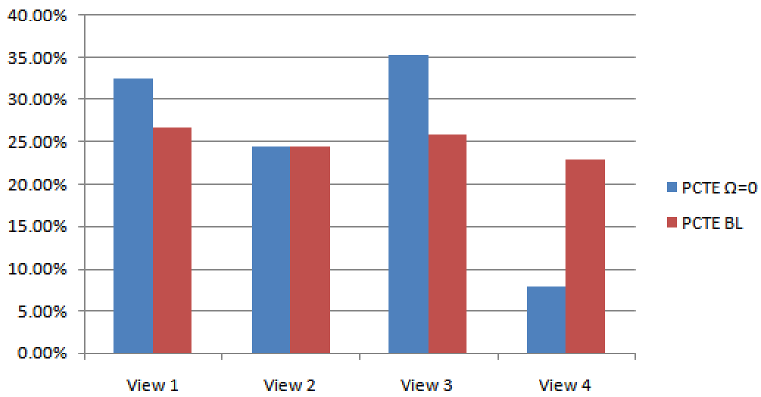


Figure 2.1: Risk distribution across views

2.7.2 Second Case: One Expressed Preference

In the first case, we worked with equal ICs for all the views. We are now going to look at the consequences of giving more preference to one view, which translates into an increase in its Information Coefficient.

Assume that the investor has a preference for the first view. They think it is going to provide with a better performance than the other and are hence willing to allocate more risk to it. We may for instance consider that they will give an IC of 0.3 to that view and maintain the other ICs at 0.2.

We then move on to the optimization, keeping the 100bp constraint on the tracking error. The results for both the $\boldsymbol{\Omega} \equiv \text{diag}\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^T$ case and the $\boldsymbol{\Omega} \equiv 0$ case are showed in Table 2.5.

The conclusions drawn from these results are quite similar to those of the previous example. Here, given the expressed preference for the first view, more risk was allocated to it. In the case where confidence is taken into account, the increase in that view seems adequate and the remaining risk

View	IC	$\Omega \equiv \text{diag}\boldsymbol{P\Sigma P}^T$		$\Omega \equiv 0$	
		$\boldsymbol{w}_{v_a}$	$\mathbf{PCTE}_v$	$\boldsymbol{w}_{v_a}$	$\mathbf{PCTE}_v$
1	0.3	2.77%	46.41%	3.26%	60.79%
2	0.2	1.16%	18.15%	1.49%	20.69%
3	0.2	0.89%	18.35%	0.68%	12.77%
4	0.2	1.07%	17.09%	0.42%	5.75%

Table 2.5: Optimized Portfolio with a Preference

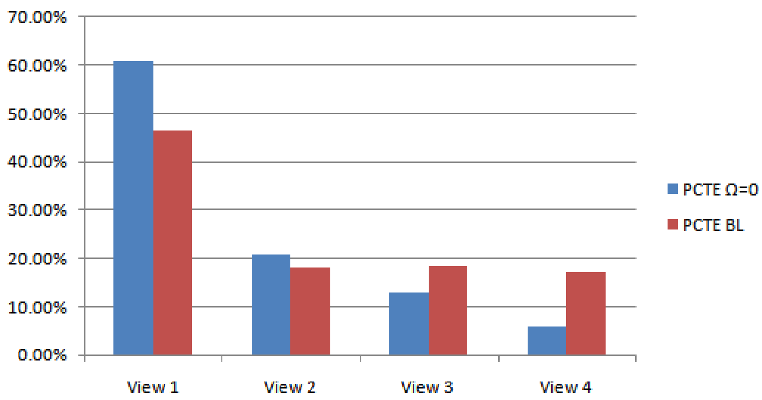


Figure 2.2: Risk distribution across views

is evenly spread across the three other views. On the contrary, in the case of perfect foresight ($\boldsymbol{\Omega} \equiv \mathbf{0}$), a disproportionately large amount (60%) of the risk was allocated to the first view alone, and the remaining risk is not evenly spread.

2.7.3 Third Case: Ranking

A possible case scenario is one where the investor ranks the assets along a ranking. This ranking may take different forms, we could for instance imagine a simple case where the investor places the assets into three view categories: bearish, neutral and bullish. From this simple case we could elaborate into "very bearish, bearish, neutral, bullish or very bullish" and so forth, making the categories more and more precise. Whichever the choice, since each view is associated to a single asset, the pick matrix $\boldsymbol{P}$ is going to contain as many rows as there are assets, and only one term per row, either 1 or -1 depending on the view (1 for a bullish view and -1 for a bearish view).

Let us go back to our investor. Given their information, they have managed to give either a bearish or bullish opinion to each asset. The assets in the "bearish basket" are Australia, Canada and Japan, and the assets in the "bullish basket" are France, Germany and the UK. Unfortunately they could not be more precise than that, so the best they can do is assign an equal IC of 0.15 to each asset.

The pick matrix in this case is:

$$
\boldsymbol{P} = \begin{array}{c} \begin{array}{cccccc} \text{AU} & \text{CA} & \text{FR} & \text{GE} & \text{JP} & \text{UK} \end{array} \\ \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}
$$

The results with a tracking error constraint of 100bp are showed in Table 2.6. We see here again that using a non null $\boldsymbol{\Omega}$ helps dealing with the risk. There are less extreme positions in the assets and the risk is better diversified.

			$\boldsymbol{\Omega} \equiv \text{diag}\boldsymbol{P\Sigma P}^T$		$\boldsymbol{\Omega} \equiv 0$	
Country	View	IC	$\boldsymbol{w}_a$	**PCTE**	$\boldsymbol{w}_a$	**PCTE**
Australia	bearish	0.15	-1.75%	16.28%	-2.33%	23.89%
Canada	bearish	0.15	-2.33%	10.15%	-3.05%	28.95%
Japan	bearish	0.15	-1.36%	18.20%	-0.11%	0.92%
France	bullish	0.15	1.57%	18.53%	1.02%	9.19%
Germany	bullish	0.15	1.55%	18.62%	1.23%	10.89%
UK	bullish	0.15	1.72%	18.22%	2.96%	26.16%

Table 2.6: Optimized Portfolio for a simple ranking

We now assume the investor has received some more information and is able to produce a more detailed ranking of the assets. They have ranked the assets in each category and allocated the ICs accordingly. The modification in the results are showed in Table 2.7.

The results are not as easy to interpret in the case where $\boldsymbol{\Omega} \equiv 0$ for some of the PCTEs are negative; but we can still say a few things. Here the different assets have been given different ICs depending on how much the investor thinks they are going to over- or under-perform. We therefore expect to allocate more risk to those assets with a higher IC and less to

Country	View	IC	$\boldsymbol{\Omega} \equiv \mathrm{diag}\boldsymbol{P\Sigma P}^T$ $\boldsymbol{w}_a$	PCTE	$\boldsymbol{\Omega} \equiv 0$ $\boldsymbol{w}_a$	PCTE
Australia	bearish	0.15	-2.76%	32.83%	-3.08%	33.92%
Canada	bearish	0.1	-2.27%	10.24%	-2.35%	15.93%
Japan	bearish	0.05	-0.21%	2.17%	1.14%	-3.35%
Germany	bullish	0.05	0.32%	3.23%	-2.07%	-6.59%
France	bullish	0.1	1.65%	17.54%	1.05%	6.86%
UK	bullish	0.15	3.16%	33.99%	5.61%	53.23%

Table 2.7: Optimized Portfolio for a detailed ranking

the ones with low IC. Furthermore, these allocation should present some homogeneity in their distribution, that is to say if two assets have been given a similar IC, we expect them to bear a similar portion of risk. All these requests are very well handled in the case where $\boldsymbol{\Omega}$ is taken into account as we can see on Figure 2.3; although this cannot be said to be entirely true in the case where $\boldsymbol{\Omega} \equiv 0$, which shows much larger active positions.

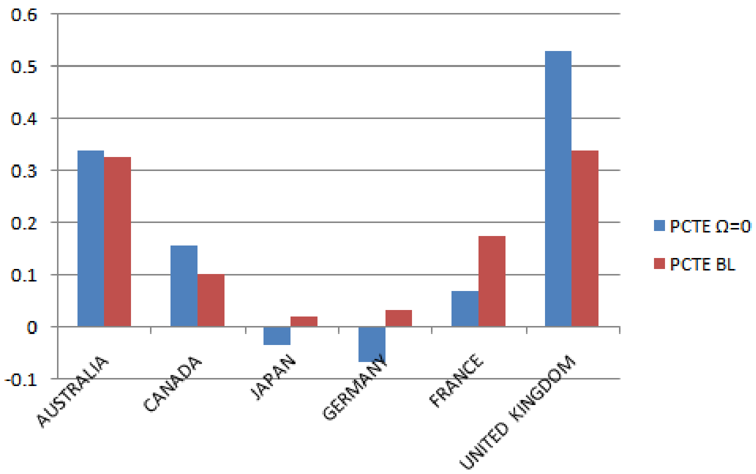


Figure 2.3: Risk distribution across views in a ranking case

2.8 Testing Views' Independence

In the theoretical construction of the Black-Litterman model, we have assumed a diagonal structure for the covariance matrix of the views $\boldsymbol{\Omega}$, hence implying that they were independent.

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \omega_k^2 \end{pmatrix}$$

It would be a sound exercise to verify on a numerical example whether this assumption is tolerable or whether it must abandoned.

Let us test this on an over-simplified example wherein all the views share a common correlation ρ. This would be a model where all the views are affected by a single factor:

$$\begin{aligned} \boldsymbol{q} &= \boldsymbol{P}\mathbb{E}[\boldsymbol{\alpha}|\mathcal{G}] + \boldsymbol{\omega} \\ \boldsymbol{\omega} &= \begin{pmatrix} \boldsymbol{\omega}_1 \\ \vdots \\ \boldsymbol{\omega}_K \end{pmatrix} \\ \boldsymbol{\omega}_i &= \omega_i \left(\sqrt{1-\rho^2} Z_i + \rho Z \right) \end{aligned}$$

where $Z, Z_1, \dots, Z_K$ are independent standard normal random variables and ω_i is the variance of the corresponding view, i.e. $\mathrm{diag}(\boldsymbol{P\Sigma P}^T)_i$, to be consistent with the previous model.

We will test this model on the data used in section 2.7 by fixing a set of views and analyzing how much the correlation factor ρ affects the values of the expected alphas and active weights.

Let us set the following views:

Country	View	IC
Australia	bullish	0.15
Canada	bullish	0.15
Japan	bullish	0.15
Germany	bearish	0.15
France	bearish	0.15
UK	bearish	0.15

Table 2.8: Views used for testing the influence of the views'correlation

We then solve the problem for values of ρ ranging from 0 - independent views - to 1 - fully correlated views. For each value of ρ we find a vector of expected alphas $\boldsymbol{\alpha}_\rho$ and active weights $\boldsymbol{w}_\rho$, and in order to measure how distant this new vectors are from the initial ones obtained with independent views, we calculate their Mahalanobis distance with respect to the original vectors:

$$\begin{aligned} d_\alpha &= \sqrt{(\boldsymbol{\alpha}_\rho - \boldsymbol{\alpha}_0)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\alpha}_\rho - \boldsymbol{\alpha}_0)} \\ d_w &= \sqrt{(\boldsymbol{w}_\rho - \boldsymbol{w}_0)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{w}_\rho - \boldsymbol{w}_0)} \end{aligned}$$

2.8.1 Influence on expected alphas

Country	0	0.2	0.4	0.6	0.8	1
Australia	1.61%	1.63%	1.73%	1.93%	2.32%	3.21%
Canada	0.76%	0.79%	0.92%	1.20%	1.75%	2.97%
France	-2.05%	-2.47%	-2.95%	-3.55%	-4.40%	-6.00%
Germany	-2.08%	-2.49%	-2.96%	-3.54%	-4.33%	-5.32%
Japan	2.32%	2.43%	2.59%	2.80%	3.02%	3.11%
UK	-1.84%	-2.22%	-2.66%	-3.21%	-3.97%	-5.31%
d_α	0	0.02365	0.05465	0.09833	0.16663	0.30335

Table 2.9: Alphas $\boldsymbol{\alpha}_\rho$ in each country and Mahalanobis distance d_α

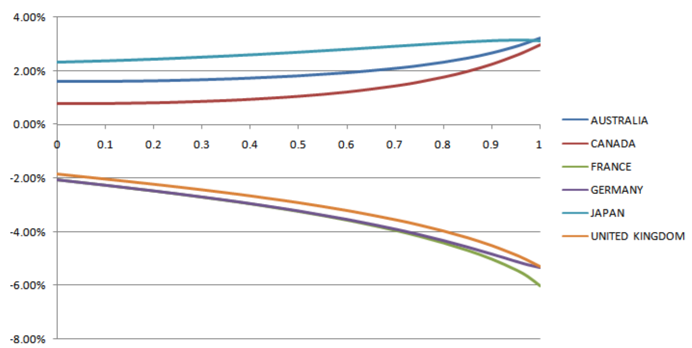


Figure 2.4: $\boldsymbol{\alpha}_\rho$ against ρ

2.8.2 Influence on active weights

	ρ					
Country	0	0.2	0.4	0.6	0.8	1
Australia	1.76%	1.76%	1.77%	1.81%	1.96%	2.19%
Canada	2.34%	2.37%	2.41%	2.52%	2.73%	2.88%
France	-1.59%	-1.61%	-1.63%	-1.68%	-1.90%	-3.24%
Germany	-1.54%	-1.56%	-1.57%	-1.57%	-1.38%	-0.16%
Japan	1.36%	1.28%	1.19%	1.01%	0.63%	0.13%
UK	-1.72%	-1.75%	-1.79%	-1.84%	-1.98%	-1.87%
d_w	0	0.00491	0.01115	0.02399	0.06769	0.33144

Table 2.10: Active weights $\boldsymbol{w}_\rho$ in each country and d_w against ρ

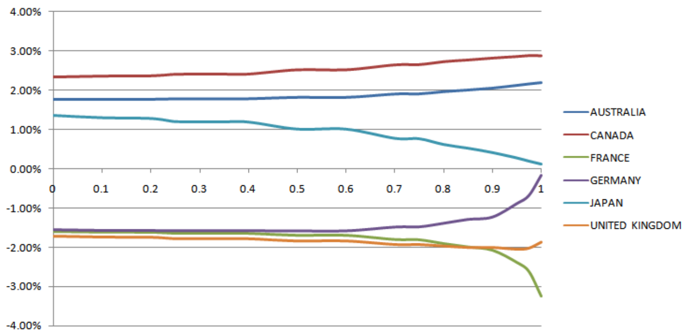


Figure 2.5: $\boldsymbol{w}_\rho$ against ρ

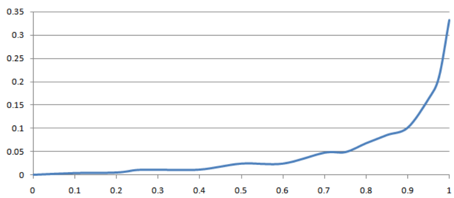


Figure 2.6: d_w against ρ

2.8.3 Analysis

We observe on the results for the expected alphas that the effect of an increasing correlation is to sharpen the results, positive alphas are more positive and negative alphas are more negative. The results seem to depart the initial vector of expected alphas quite quickly.

However, it is interesting to notice that the effects on the active weights are not that severe. Indeed, for relatively small values of ρ ($\rho \leq 0.4$), we can consider the independence of the views' errors as a relatively good and convenient approximation.

We should point out that this conclusion is only based on a particular case and that it might not apply to any application of the model. It seems nonetheless that I have found similar results on all the examples I have tried.

2.9 Influence of tau

Similarly to the study realized in the previous section, we may want to investigate the impact of "canceling out" τ by using He and Litterman's estimation of $\boldsymbol{\Omega}$. Using the same views as in Table 2.8, we plot the resulting active weights against τ.

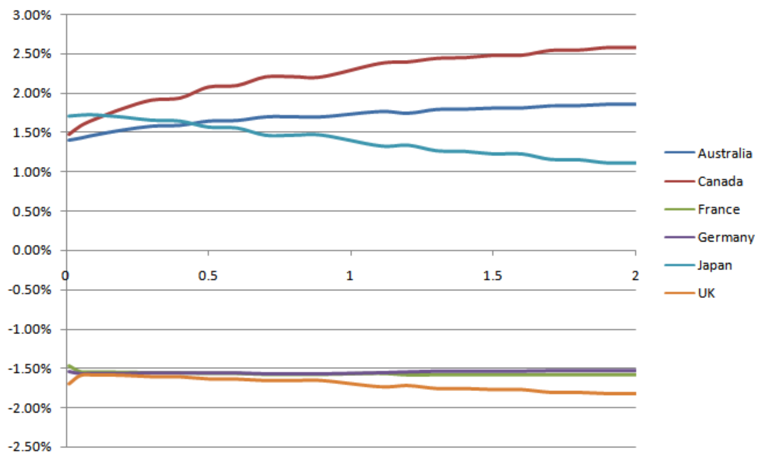


Figure 2.7: w against τ

These results show that the influence of τ is not insignificant, even for small changes around the value we have used so far ($\tau = 1$). There might thus be some work to be done in order to fine-tune this value.

Chapter 3

Fixed-Income Portfolio Optimization

3.1 Introduction

We have introduced in the previous chapters the Black-Litterman model and its application in an active management context. The purpose of this chapter is to apply the framework and tools we have presented thus far to a Fixed-Income Portfolio which may consist of various instruments such as bonds, futures, currency forwards or swaps to name but a few.

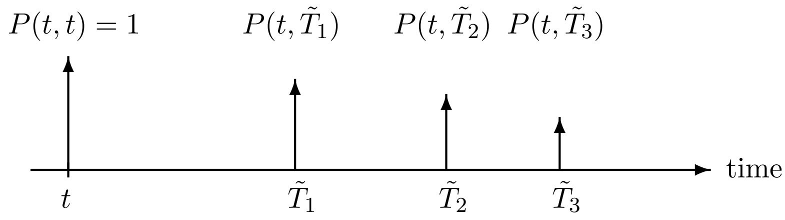
In this new context, we will consider a portfolio manager whose universe is constituted of bonds from different countries or monetary zones and their currencies, as well as inflation-linked bonds. They manage their portfolios with respect to a given benchmark and aim at maximizing their expected alpha with a tracking error constraint.

The active positions they might take in any asset are motivated by views or opinions on how the currencies, yield curves and inflation are going to behave in the future. Furthermore, this manager might handle portfolios for different clients with different base currencies. When switching from one base currency to another, the risks - measured by the assets' variances and covariances - might change, thus modifying the positions to be taken, which we will have to investigate.

We will proceed as follows: in the first sections we will tackle the modeling of the portfolio and its returns, raising some practical issues as we come across them; and then apply it to some concrete examples.

3.2 Yield curve

Let $P(t,T)$ be the discount factor observed at time t for maturity T. In the following t shall refer to the present time and T to a future time, often the maturity of a bond. In order to keep notations as simple as possible, we will write $\tilde{T} := t + T$.



Given $P(t,T)$, we may define the associated zero-coupon continuously compounded rate $z(t,T)$:

$$\begin{aligned} P(t,\tilde{T}) &= e^{-z(t,\tilde{T})T} \\ \Rightarrow z(t,\tilde{T}) &= -\frac{1}{T}\ln P(t,\tilde{T}). \end{aligned}$$

By Principal Component Analysis (PCA), z is broken down into a linear combination of n factors $(z_i(T))_{1\leq i\leq n}$ with respective loadings $(\kappa_i(t))_{1\leq i\leq n}$:

$$z(t,\tilde{T}) \approx \sum_{i=1}^{n} \kappa_i(t) z_i(T). \tag{3.1}$$

These loadings are time dependent as is underlined by the notation. They are continuously re-evaluated in order to best fit the yield curve.

Let us stress that the functions P and z are functions of two variables: present time t and maturity $\tilde{T}$; but that the functions $(\kappa_i)_{1\leq i\leq n}$ and $(z_i)_{1\leq i\leq n}$ are univariate; $(\kappa_i)_{1\leq i\leq n}$ being functions solely of the present time t, and $(z_i)_{1\leq i\leq n}$ being functions only of the maturity date $\tilde{T}$.

Furthermore, we may choose n specific terms $(\tilde{T}_i)_{1\leq i\leq n}$ and express the loadings $(\kappa_i(t))_{1\leq i\leq n}$ as a linear combination of the values $(z(t,\tilde{T}_i))_{1\leq i\leq n}$ of z for these n particular terms. The interest in doing this is that the effect of the change in the yield curve around one of these terms can then easily be assessed. A common choice for these term dates would be treasury bills, notes and bonds' maturity dates.

Letting $(\tilde{T}_i)_{1\leq i\leq n}$ be the chosen n different terms, for each term $\tilde{T}_i$, we have:

$$z(t,\tilde{T}_i) = \sum_{j=1}^{n} \kappa_j(t) z_j(T_i).$$

In matrix form, this can be written as:

$$\boldsymbol{z}(t,T_1,\ldots,T_n) = \boldsymbol{A}(T_1,\ldots,T_n)\boldsymbol{\kappa}(t) \tag{3.2}$$

Where:

$$\boldsymbol{z}(t,T_1,\ldots,T_n) = \begin{pmatrix} z(t,\tilde{T}_1) \\ \vdots \\ z(t,\tilde{T}_n) \end{pmatrix}, \ \boldsymbol{A}(T_1,\ldots,T_n) = \begin{pmatrix} z_1(T_1) & \ldots & z_n(T_1) \\ \vdots & & \vdots \\ z_1(T_n) & \ldots & z_n(T_n) \end{pmatrix}$$

and $\boldsymbol{\kappa}(t) = \begin{pmatrix} \kappa_1(t) \\ \vdots \\ \kappa_n(t) \end{pmatrix}$.

Inverting the above matrix equation (3.2), the values of the loadings $(\kappa_i(t))_{1\leq i\leq n}$ can then be expressed as linear combinations of the $\left(z(t,\tilde{T}_i)\right)_{1\leq i\leq n}$:

$$\boldsymbol{\kappa}(t) = \boldsymbol{A}^{-1}(T_1,\ldots,T_n)\boldsymbol{z}(t,T_1,\ldots,T_n),$$

letting $\boldsymbol{A}^{-1}(T_1,\ldots,T_n) = (\alpha_{ij}(T_1,\ldots,T_n))_{1\leq i,j\leq n}$.
Consequently:

$$\kappa_i(t) \quad = \quad \sum_{j=1}^{n} \alpha_{ij}(T_1,\ldots,T_n) z(t,\tilde{T}_j) \tag{3.3}$$

Substituting (3.3) into (3.1) we may thus write any term rate as a linear combination of the n specific terms rates:

$$z(t,\tilde{T}) = \sum_{i=1}^{n} \beta_i(T,T_1,\ldots,T_n) z(t,\tilde{T}_i). \tag{3.4}$$

3.3 Sensitivity to yield curve changes

We want to investigate the consequences of changes in the yield curve on the value of a *bond* portfolio in one particular monetary zone. We look at the change in the value of the portfolio between times t and $t+dt$, where dt is a positive time. We assume a self-financing condition between t and $t+dt$, that is, no money is inserted in or taken out of the portfolio and the changes in the value of the portfolio are solely due to fluctuations of the portfolio's assets.

For any function of time $g(t)$, we will write $\Delta g(t) := g(t + dt) - g(t)$.

The value of a bond portfolio is a function of the yield curve and consequently a function of z.

- Writing $V(z(t,.))$ the value of the portfolio at time t, we may then write:

$$V(z(t,.)) = V(\kappa_1(t), \ldots, \kappa_n(t)).$$

  Changes in the yield curve are represented by changes in the values of the loadings. Therefore we may investigate the sensitivity of the portfolio to changes in the yield curve thanks to a first order Taylor approximation:

$$\Delta V(\kappa_1(t), \ldots, \kappa_n(t)) \approx \sum_{i=1}^{n} \frac{\partial V}{\partial \kappa_i} \Delta \kappa_i(t)$$

- However, the above expression is not fully satisfactory. Indeed, we would like to express the variation of the portfolio not as depending on the changes in the loadings $(\kappa_i(t))_{1 \leq i \leq n}$, but rather depending on the log-returns of the yield curve at the chosen terms. We aim at expressing the change in the value of the portfolio in terms of the $(\ln z(t, T_i))_{1 \leq i \leq n}$ which we know is possible thanks to equation (3.4). We would like an expression of the form:

$$\Delta V\left(z(t,.)\right) \approx \sum_{i=1}^{n} c_i(t) \Delta \ln z(t, \tilde{T}_i)$$

According to (3.4), the yield curve can be expressed as a function of a selected number of term rates, so in particular:

$$V\left(z(t,.)\right) = V\left(\ln z(t, \tilde{T}_1), \ldots, \ln z(t, \tilde{T}_n)\right)$$

which leads to the following first order Taylor expansion:

$$\Delta V\left(\ln z(t, \tilde{T}_1), \ldots, \ln z(t, \tilde{T}_n)\right) \approx \sum_{i=1}^{n} \frac{\partial V}{\partial \ln z(t, \tilde{T}_i)} \Delta \ln z(t, \tilde{T}_i) \tag{3.5}$$

Unfortunately, this expression cannot be readily used for we do not have immediate access to the partial derivatives with respect to the logarithms of $z(t, \tilde{T}_i)$. We must thus transform these expressions further, using that:

$$\frac{\partial V}{\partial \ln z(t, \tilde{T}_i)} = \frac{\partial V}{\partial z(t, \tilde{T}_i)} z(t, \tilde{T}_i) \tag{3.6}$$

In the above equality, the term $\frac{\partial V}{\partial z(t,\tilde{T}_i)}$ represents the exposure of the portfolio to the yield curve at $\tilde{T}_i$. It is the opposite of the dollar duration $d_{\$,i}(t)$ at time t of the portfolio with respect to the corresponding rate.

We write:

$$\boldsymbol{d}_{\$}(t) = -\begin{pmatrix} \frac{\partial V}{\partial z(t,\tilde{T}_1)} \\ \vdots \\ \frac{\partial V}{\partial z(t,\tilde{T}_n)} \end{pmatrix} \tag{3.7}$$

Inserting (3.7) into (3.6) gives:

$$\frac{\partial V}{\partial \ln z(t,\tilde{T}_i)} = -\boldsymbol{d}_{\$_i}(t) z(t,\tilde{T}_i) \tag{3.8}$$

In turn, inserting (3.8) into equation (3.5) finally yields:

$$\Delta V = -\sum_{i=1}^{n} \boldsymbol{d}_{\$_i}(t) z(t,\tilde{T}_i) \Delta \ln z(t,\tilde{T}_i) \tag{3.9}$$

Equation (3.9) gives a first order approximation of the variations of the value of a bond portfolio in terms of the rates' log-returns of a chosen number of terms. It is then straightforward to compute the return of the portfolio:

$$\begin{aligned} R &= \frac{\Delta V}{V(t)} = -\sum_{i=1}^{n} \frac{\boldsymbol{d}_{\$_i}(t)}{V(t)} z(t,\tilde{T}_i) \Delta \ln z(t,\tilde{T}_i) \\ R &= -\sum_{i=1}^{n} \boldsymbol{d}_i(t) z(t,\tilde{T}_i) \Delta \ln z(t,\tilde{T}_i) \qquad (3.10) \end{aligned}$$

where $\boldsymbol{d}_i(t) = \frac{\boldsymbol{d}_{\$_i}(t)}{V(t)} = -\frac{1}{V} \frac{\partial V}{\partial z(t,\tilde{T}_i)}$ are the durations in each term.

3.4 Going global

In the previous section we considered a bond portfolio in one country and in its domestic currency. This was a very simplistic example, in most cases a bond investor will hold a portfolio of bonds from different countries and consequently might also invest in currencies, or simply hedge its currency exposure. Therefore, the accessible universe of the portfolio manager has extended from one country bonds to bonds from several countries and their currencies.

In order to take these new opportunities into account, let us consider K countries or monetary zones in which the portfolio manager may invest. We will use the subscript j to refer to the j-th country ($1 \leq j \leq K$).

Let $V_j(t)$ be the value of the portion of the portfolio invested in the j-th country. $V_j(t)$ is the sum of the amount invested in country j's bonds $V_{z_j}(t)$ and the amount invested in country j's currency c_j, $V_{c_j}(t)$.

$$V_j(t) = V_{z_j}(t) + V_{c_j}(t) \tag{3.11}$$

The previous section applies to any country's yield curve, so for each country in the investor's universe we have that the change in the value of the fraction of the portfolio invested in that country, expressed in its *domestic currency* c_j, is (according to (3.9)):

$$\Delta V_{zj}^{c_j} = -\sum_{i_j=1}^{n_j} \boldsymbol{d}^j_{\$_{i_j}}(t) z_j(t, \tilde{T}_{i_j}) \Delta \ln z_j(t, \tilde{T}_{i_j}) \tag{3.12}$$

Remark that we have here used the notations i_j and n_j to take into account the possibility that the number of terms and their values might not be a constant across the different countries. For instance, we might have $T_{i_{j_1}} = \{1m, 3m, 6m, 1y, 2y, 5y, 10y, 20y\}$ in country j_1 and $T_{i_{j_2}} = \{3m, 6m, 1y, 3y, 7y, 10y, 30y\}$ in country j_2.

3.4.1 Conversion into the base currency

We would like to sum up the contributions of all countries to the entire portfolio to form the expression of the variation thereof. In order for the sums to make sense, all the values must be expressed in the same unit, that is, we must choose a base currency in which all the values will be converted.

Let b be the chosen base currency and c_j be the currency in country j. We consider $f_t^{c_{j1}/c_{j2}}$ the exchange rate at time t to convert one unit of c_{j2} into c_{j1}. For instance, $f_t^{\$/£}$ is the value of one pound in dollars at time t. It is also obvious that for any j and any t $f_t^{c_j/c_j} = 1$, and that in order to satisfy no-arbitrage conditions the following cyclic condition is satisfied at all times:

$$\begin{aligned} f_t^{c_{j1}/c_{j2}} f_t^{c_{j2}/c_{j3}} &= f_t^{c_{j1}/c_{j3}} \\ i.e. \qquad \ln f_t^{c_{j1}/c_{j2}} + \ln f_t^{c_{j2}/c_{j3}} &= \ln f_t^{c_{j1}/c_{j3}}. \end{aligned} \tag{3.13}$$

For any country j $(1 \leq j \leq K)$, there are two factors that contribute to the value of the portfolio: the amount invested in country j's bonds and the amount invested in its currency c_j. We must therefore investigate how each factor affects the value of the portfolio.

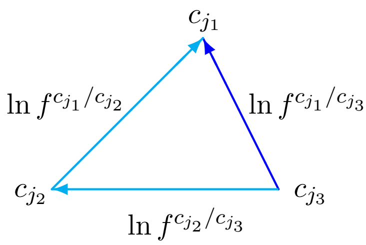


Figure 3.1: FX cyclic relation

Bonds Equation (3.12) expresses the change in the value of the amount invested in country j's bonds in its domestic currency c_j. Converting this into the base currency is straightforward, it suffices to multiply it by the exchange rate:

$$\begin{aligned}\Delta V_{z_j}^b &= -\sum_{i_j=1}^{n_j} f_t^{b/c_j} \boldsymbol{d}_{\$_{i_j}}^j(t) z_j(t,\tilde{T}_{i_j}) \Delta \ln z_j(t,\tilde{T}_{i_j}) \\ &= -\sum_{i_j=1}^{n_j} \boldsymbol{d}_{\$_{i_j}}^b(t) z_j(t,\tilde{T}_{i_j}) \Delta \ln z_j(t,\tilde{T}_{i_j}) \qquad (3.14)\end{aligned}$$

where $\boldsymbol{d}_{\$_{i_j}}^b(t) = \boldsymbol{d}_{\$_{i_j}}^j(t) f_t^{b/c_j}$ are now dollar durations with respect to the base currency.

Currencies At time t, the portfolio manager invests $V_{c_j}^b(t)$ units of b in currency c_j, thus acquiring $V_{c_j}^{c_j} = f_t^{c_j/b} V_{cj}^b(t)$ units of currency c_j. Here $V_{cj}^{c_j}$ is written as independent of time because the number of units of currency c_j is independent of time, and so the value of the amount invested in currency c_j in units of c_j is independent of time.

The value at time $t + dt$ of the amount invested in currency c_j at time t is then in units of the base currency:

$$V_{c_j}^b(t + dt) = f_{t+dt}^{b/c_j} V_{cj}^{c_j}$$

The change in value between t and $t + dt$ is given as before by a first order Taylor approximation:

$$\begin{aligned}
\Delta V_{c_j}^b &= \frac{\partial V_{c_j}^b}{\partial \ln f_t^{b/c_j}}(t)\Delta \ln f_t^{b/c_j} \\
&= \frac{\partial V_{c_j}^b}{\partial f_t^{b/c_j}}(t) f_t^{b/c_j} \Delta \ln f_t^{b/c_j} \\
&= \frac{\partial (f_t^{b/c_j} V_{cj}^{c_j})}{\partial f_t^{b/c_j}} f_t^{b/c_j} \Delta \ln f_t^{b/c_j} \\
&= V_{c_j}^{c_j} f_t^{b/c_j} \Delta \ln f_t^{b/c_j} \\
\Delta V_{c_j}^b &= V_{cj}^b(t)\Delta \ln f_t^{b/c_j} \qquad (3.15)
\end{aligned}$$

3.4.2 Adding up the contributions

We have derived the expression for the change in value of the amount invested in each country j ($j = 1, \ldots, K$) of the investor's universe, expressed in the base currency:

$$\begin{aligned}
\Delta V_j^b &= \Delta V_{cj}^b(t) + \Delta V_{zj}^b(t) \\
&= V_{cj}^b(t)\Delta \ln f_t^{b/c_j} - \sum_{i_j=1}^{n_j} \boldsymbol{d}_{\$_{i_j}}^b(t) z_j(t, \tilde{T}_{i_j})\Delta \ln z_j(t, \tilde{T}_{i_j}) \qquad (3.16)
\end{aligned}$$

Since all these contributions are expressed in units of the base currency, we may add them up to find the change in the value of the total portfolio:

$$\Delta V^b = \sum_{j=1}^{K} \Delta V_j^b \qquad (3.17)$$

3.4.3 Returns

Once we have obtained the expression for the change in the value of the portfolio, it is easy to retrieve the returns of this portfolio in the base currency. Here again we will separate returns due to bond investment and to currency investment.

The return perceived in the base currency for the currency investments is:

$$\begin{aligned}
R_c^b &= \frac{1}{V(t)} \sum_{j=1}^{K} \Delta V_{c_j}^b \\
&= \sum_{j=1}^{K} w_j(t) \Delta \ln f_t^{b/c_j} \qquad (3.18)
\end{aligned}$$

where $w_j(t)$ is the fraction of wealth invested in currency c_j at time t: $w_j(t) = \frac{V^b_{c_j}(t)}{V^b(t)}$.

Similarly, we find the return perceived in the base currency for the bonds investments:

$$R^b_z = -\sum_{j=1}^{K}\sum_{i_j=1}^{n_j} \boldsymbol{d}_{i_j}(t) z_j(t,\tilde{T}_{i_j}) \Delta \ln z_j(t,\tilde{T}_{i_j}) \tag{3.19}$$

where $(\boldsymbol{d}_{i_j}(t))_{\substack{1\le j\le K\\ 1\le i_j\le n_j}}$ are durations.

The total return perceived in the base currency is then:

$$R^b = R^b_c + R^b_z \tag{3.20}$$

3.4.4 Practicalities

In order to apply the previous relations in practice, a few transformations must often be made.

Calculate the equivalent zero-coupon continuously compounded rates In all our calculations, we have used $z(t,\tilde{T})$, the zero-coupon bond continuously compounded rate. However, in practice, quoted bonds like treasury bonds or gilt are not zero-coupons and are not continuously compounded but annualized. For instance US treasury bonds have coupon payments twice a year every six months. In order to apply our results, we must convert all rates into equivalent zero-coupon continuously compounded rates.

Consider at time t the discount factor for maturity at $\tilde{T}$. We considered the corresponding zero-coupon continuously compounded rate $z(t,\tilde{T})$:

$$P(t,\tilde{T}) = e^{-z(t,\tilde{T})T}$$

Using its quoted rate r and frequency h, its value is given by:

$$P(t,\tilde{T}) = \left(1+\frac{r}{h}\right)^{-Th}$$

Combining these two expressions yields:

$$z(t,\tilde{T}) = h\ln\left(1+\frac{r}{h}\right) \tag{3.21}$$

Converting to the base currency Practitioners often do not store exchange rates between all possible K currencies, for it would require storing $\frac{K(K-1)}{2}$ time series of data when in fact, thanks to the cyclic relation (3.13), it is only necessary to store $(K-1)$ time series, that is, all the exchange rates with respect to a given currency. (The K-th relation is just the exchange rate between the reference currency and itself, which is 1 at all times)

The reference currency for storing exchange rates might not be the same as the base currency for the portfolio. Let us assume that the reference currency is the US dollar. The exchange rate between currencies c_{j_1} and c_{j_2} can be deduced from to the cyclic relation:

$$f_t^{c_{j1}/c_{j}2} = f_t^{c_{j1}/\$} f_t^{\$/c_{j}2}$$

3.4.5 Covariance matrix and change of numeraire

Motivations

This section finds its motivations on the same grounds as the previous paragraph. For portfolio optimization processes and many other daily operations, it is necessary to calculate the covariance matrix of the assets' log-returns in the manager's universe. This task requires time and is numerically expensive since it requires extracting the time series of each of the assets and calculating pairwise covariances. Furthermore, the portfolio manager might handle accounts in different base currencies, so for each account, they would have to calculate that covariance matrix in the corresponding base currencies.

A sound question is then to ask whether it is possible to avoid too many tedious calculations, and given the covariance matrix in one specific reference currency, recover the covariance matrix in any other base currency via a simple transformation. This would enable the manager to calculate the covariance matrix only once every day and spare some computational time.

We will first derive a method in the case where the only assets are foreign currencies which will be easily extended in order to include foreign bonds.

Starting Point

As we mentioned, the reference currency used to quote exchange rates is often the US Dollar. We will assume this in our case for illustration purposes. Under this assumption, the exchange rates are quoted as the amount of USD that can be purchased with 1 unit of foreign currency. Thus, if we consider K countries in which we may invest (including the USA for simplicity), we are given $(K-1)$ time series: $\left(f_t^{\$/c_j}\right)_{1\leq j\leq K-1}$.

Our starting point is the covariance matrix of the log-returns of these exchange rates:

$$\begin{array}{c} \\ \Delta\ln f^{\$/c_1} \\ \Delta\ln f^{\$/c_2} \\ \vdots \\ \Delta\ln f^{\$/c_{K-1}} \end{array}
\begin{array}{c} \begin{array}{cccc} \Delta\ln f^{\$/c_1} & \Delta\ln f^{\$/c_2} & \dots & \Delta\ln f^{\$/c_{K-1}} \end{array} \\
\begin{pmatrix} \boldsymbol{C}_{1,1} & \boldsymbol{C}_{1,2} & \dots & \boldsymbol{C}_{1,K-1} \\ \boldsymbol{C}_{2,1} & \boldsymbol{C}_{2,2} & \dots & \boldsymbol{C}_{2,K-1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{C}_{K-1,1} & \dots & \dots & \boldsymbol{C}_{K-1,K-1} \end{pmatrix} \end{array} = \boldsymbol{C}$$

Where we have written $\boldsymbol{C}_{i,j} = \mathrm{cov}(f^{\$/c_i}, f^{\$/c_j})$, $(i,j = 1,\dots,K-1)$.

Of course, we have that $\boldsymbol{C}_{i,i} = \mathrm{cov}(f^{\$/c_i}, f^{\$/c_i}) = \mathrm{var}(f^{\$/c_i})$ and that $\boldsymbol{C}_{i,j} = \boldsymbol{C}_{j,i}$.

Algebraic structure

Let us consider E the $\mathbb{R}$-vector space span by the log-returns of all the possible exchange rates:

$$E = \mathrm{vect}\left\{\Delta\ln f^{c_i/c_j}, \quad i,j = 1,\dots,K\right\}$$

We know that given a base currency b, any vector $\Delta\ln f^{c_i/c_j}$ can be written as a linear combination of the $K-1$ vectors $\left(\Delta\ln f^{b/c_j}\right)_{1\leq j\leq K-1}$:

$$\begin{aligned} \Delta\ln(f^{c_i/c_j}) &= \Delta\ln(f^{c_i/b} f^{b/c_j}) \\ &= \Delta\ln(f^{b/c_j}) - \Delta\ln(f^{b/c_i}) \end{aligned}$$

This shows that $\dim E \leq K-1$. In fact it is easy to see that $\dim E = K-1$. So any free family of $K-1$ log-returns of exchange rates forms a base of E. There are some bases worth being noticed, these are the bases constructed from the log-returns of all the exchange rates relative to one particular currency. For instance we associate to any currency b the base of E, $\mathcal{B}^b$, given by:

$$\mathcal{B}^b = \left\{\Delta\ln(f^{b/c_i}), \quad i = 1,\dots,K \text{ and } c_i \neq b\right\} \tag{3.22}$$

Now, given any two vectors u and v belonging to E, we may consider the form ϕ that returns their covariance:

$$\begin{array}{rccl} \phi: & E\times E & \longrightarrow & \mathbb{R} \\ & u,v & \longmapsto & \phi(u,v) = \mathrm{cov}(u,v) \end{array}$$

Furthermore, using the properties of the covariance, we have that:

- $\forall (u,v) \in E^2, \quad \phi(u,v) = \phi(v,u)$
- $\forall (u,v,w) \in E^3, \quad \phi(u+w,v) = \phi(u,v) + \phi(w,v)$
- $\forall (u,v) \in E^2, \forall \lambda \in \mathbb{R}, \quad \phi(\lambda u, v) = \lambda \phi(u,v)$
- $\forall u \in E, \phi(u,u) = \mathrm{var}(u) \geq 0$

These properties show that ϕ is in fact a bilinear symmetric positive form on E - the associated quadratic form being the variance. Given a base of E, $\mathcal{B} = \{e_i, \quad i = 1, \ldots, K-1\}$, we may look at the matrix representation of ϕ:

$$\boldsymbol{V}^{\mathcal{B}} = (\phi(e_i, e_j))_{\substack{1 \leq i \leq K-1 \\ 1 \leq j \leq K-1}} .$$

Moreover, given another base of E, $\mathcal{B}'$, and the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$, P, the matrix representation of ϕ in $\mathcal{B}'$ is:

$$\boldsymbol{V}^{\mathcal{B}'} = \boldsymbol{P}^t \boldsymbol{V}^{\mathcal{B}} \boldsymbol{P} \tag{3.23}$$

When $\mathcal{B} = \mathcal{B}^b$, we will write $\boldsymbol{V}^{\mathcal{B}^b} = \boldsymbol{V}^b$ for simplicity.

Solution to the problem

We may now return to our problem and realize that in fact our original matrix $\boldsymbol{C}$ turns out to be $\boldsymbol{V}^{\$}$, and that our problem is to transform $\boldsymbol{V}^{\$}$ into $\boldsymbol{V}^b$, where b is our desired base currency. According to equation (3.23) this problem comes down to finding the transition matrix from $\mathcal{B}^{\$}$ to $\mathcal{B}^b$.

In order to find the transition matrix from $\mathcal{B}^{\$}$ to $\mathcal{B}^b$, all we need to do is write down the vectors of $\mathcal{B}^b$ as linear combinations of the vectors of $\mathcal{B}^{\$}$, for a transition matrix contains the coordinates of the new base in the old base.

The old base is:

$$\mathcal{B}^{\$} = \left\{ \Delta \ln f^{\$/c_1}, \Delta \ln f^{\$/c_2}, \ldots, \Delta \ln f^{\$/b}, \ldots, \Delta \ln f^{\$/c_{K-1}} \right\}$$

whereas the new base is:

$$\mathcal{B}^b = \left\{ \Delta \ln f^{b/c_1}, \Delta \ln f^{b/c_2}, \ldots, \Delta \ln f^{b/\$}, \ldots, \Delta \ln f^{b/c_{K-1}} \right\}$$

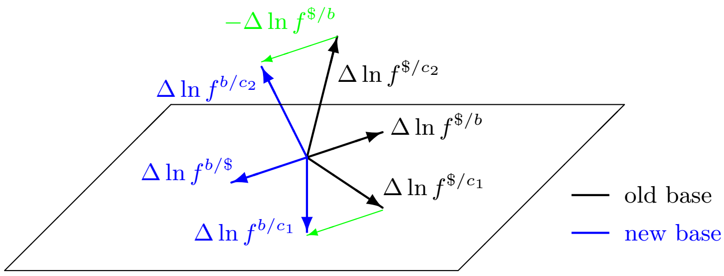


Figure 3.2: Covariance matrix' change of base

For any currency c_i, $i = 1, \ldots, K-1$, we have:

$$\begin{aligned} \Delta \ln f^{b/c_i} &= \Delta \ln(f^{b/\$} f^{\$/c_i}) \\ &= \Delta \ln f^{\$/c_i} - \Delta \ln f^{\$/b} \end{aligned}$$

We easily deduce the transition matrix from these relations:

$$\begin{array}{cc} & \begin{array}{cccccc} \Delta \ln f^{b/c_1} & \Delta \ln f^{b/c_2} & \ldots & \Delta \ln f^{b/\$} & \ldots & \Delta \ln f^{b/c_{K-1}} \end{array} \\ \begin{array}{c} \Delta \ln f^{\$/c_1} \\ \Delta \ln f^{\$/c_2} \\ \vdots \\ \Delta \ln f^{\$/b} \\ \vdots \\ \Delta \ln f^{\$/c_{K-1}} \end{array} & \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ -1 & -1 & -1 & -1 & -1 & -1 \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix} \end{array} = \boldsymbol{P}$$

Therefore, the covariance matrix of the log-returns of the exchange rates in the base currency b is simply:

$$\begin{aligned} \boldsymbol{V}^b &= \boldsymbol{P}^T \boldsymbol{V}^\$ \boldsymbol{P} \\ &= \boldsymbol{P}^T \boldsymbol{C} \boldsymbol{P} \end{aligned}$$

Extension

The result we have derived above is only valid in the case where the assets only consist of foreign currencies, but our framework also includes foreign bonds. The formulae we have found in sections (3.4.2) and (3.4.3) show that we need to calculate covariances between the log-returns of the foreign exchange rates and of the bond yield rates. Thankfully, the latter do not depend on foreign exchange rates so incorporating them is quite simple.

The original covariance matrix we deal with has the form:

$$\boldsymbol{\Sigma}^{\$} = \begin{pmatrix} \boldsymbol{C} & \boldsymbol{F} \\ \boldsymbol{F}^T & \boldsymbol{G} \end{pmatrix}$$

where $\boldsymbol{C}$ is as above the covariance matrix of the log-returns of the foreign exchange rates and $\boldsymbol{G}$ is the covariance matrix of the log-returns of the bonds' yield rates.

Here we consider the $\mathbb{R}$-vector space E' which is span by E and the log-returns of the bond yield rates. Since the log-returns of the bonds' yield rates are independent of the exchange rates, the transition matrix $\boldsymbol{Q}$ from the "dollar base" of E' to the new base currency base is just:

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{P} & 0 \\ 0 & \boldsymbol{I} \end{pmatrix}$$

where $\boldsymbol{P}$ is as above and $\boldsymbol{I}$ is the identity matrix.

The covariance matrix in the base currency b is then:

$$\boldsymbol{\Sigma}^b = \boldsymbol{Q}^T \boldsymbol{\Sigma}^{\$} \boldsymbol{Q} \tag{3.24}$$

3.5 Inflation-linked Bonds

We now throw in our manager's universe inflation-linked bonds. Inflation-linked bonds bear great similarities with conventional bonds in that they pay interest at fixed intervals and return the principal at maturity. However, unlike those bonds, the principal is adjusted each time on the general price inflation.

If we write N the principal, α the interest rate payed by the bond, $(\tilde{T}_1, \ldots, \tilde{T}_n)$ the fixed term dates, $i(t, \tilde{T}_j)$ the continuously compounded inflation rate between t and $\tilde{T}_j$; the present value of such a bond is:

$$V = \sum_{j=1}^{n} \alpha \left(N e^{i(t,\tilde{T}_j)T_j} \right) e^{-z(t,\tilde{T}_j)T_j} + \left(N e^{i(t,\tilde{T}_n)T_n} \right) e^{-z(t,\tilde{T}_n)T_n}$$

Let us now consider a simple zero-coupon inflation-linked bond maturing at $\tilde{T}$. Its present value is simply:

$$V = e^{i(t,\tilde{T})T} e^{-z(t,\tilde{T})T}$$

Assuming the nominal rate $z(t,\tilde{T})$ remains constant, we find the duration of this bond with respect to the inflation rate:

$$\begin{aligned} -\frac{1}{V}\frac{\partial V}{\partial i} &= -i(t,\tilde{T})T \\ d &= -i(t,\tilde{T})T \leq 0 \end{aligned}$$

We see here that while most conventional bonds have a positive duration (with respect to nominal interest rates), inflation-linked bonds on the contrary have a negative duration (with respect to inflation rates). This is understandable for when nominal interest rates go up, the value of a conventional bond goes down; but inversely, when the value of inflation rates go up (while maintaining a constant nominal rate), the value of an inflation-linked bond goes up as well.

Applying the exact same reasoning to the inflation curves as the one applied to the yield curves, we find that the return in the base currency b from investing in L countries' inflation curves is:

$$R_i^b = +\sum_{j=1}^{L}\sum_{l_j=1}^{m_j} \boldsymbol{d}_{l_j}(t) i_j(t,\tilde{T}_{l_j})\Delta \ln i_j(t,\tilde{T}_{l_j}) \tag{3.25}$$

Note that we have stressed the difference with the return from a conventional bond portfolio by explicitly writing the "+" sign.

3.6 Optimization Objective

Let us consider the problem of optimizing a portfolio formed of currencies, conventional bonds and inflation-linked bonds over a one period interval. We consider a manager who aims at maximizing the return of their portfolio with respect to a benchmark, over the period spanning from t to $t+1$. We drop the time index in our notations for clarity and simplicity since they can be inferred.

Optimization function Using the previous notations, the return of the portfolio in the base currency is:

$$R^b = R_c^b + R_z^b + R_i^b$$

where R_c^b is the return from investing in currencies, R_z^b the return from investing in the yield curves and R_i^b the return from investing in the inflation curves.

According to equations (3.18), (3.19) and (3.25), the manager must maximize the following function:

$$\begin{aligned} R^b = \sum_{i=1}^{K} \boldsymbol{w}_j^a \Delta \ln f^{b/c_j} \quad &- \sum_{\substack{1\le j\le K\\ 1\le i_j\le n_j}} \boldsymbol{d}_{i_j}^a z_j(\tilde{T}_{i_j}) \Delta \ln z_j(\tilde{T}_{i_j}) \\ &+ \sum_{\substack{1\le j\le L\\ 1\le l_j\le m_j}} \boldsymbol{d}_{l_j}^a i_j(\tilde{T}_{l_j}) \Delta \ln i_j(\tilde{T}_{l_j}) \end{aligned} \tag{3.26}$$

where $\boldsymbol{d}^a$ and $\boldsymbol{w}^a$ are active durations and active weights respectively.

The problem faced by the manager can therefore be written as:

$$\begin{aligned} &\arg\max_{\boldsymbol{w}^a,\boldsymbol{d}^a} R^b(\boldsymbol{w}^a, \boldsymbol{d}^a) \\ s.t. \quad &\text{constraints} \end{aligned} \tag{3.27}$$

Constraints One constraint of particular interest in an active management setting will be a tracking error constraint defined as:

$$(\boldsymbol{w}^a \boldsymbol{d}^a)\, \boldsymbol{K}^{\$} \begin{pmatrix} \boldsymbol{w}^a \\ \boldsymbol{d}^a \end{pmatrix} \le \alpha^2$$

where we have "stacked" the vectors of active weights and active durations. $\boldsymbol{K}^{\$}$is the covariance matrix as defined in (3.4.5) and α the imposed tracking error upper bound.

Another constraint we may consider is bounding the active duration of the portfolio, so that:

$$l \le \mathbf{1}^T \boldsymbol{d}^a \le u$$

where l and u are lower and upper bounds respectively and $\mathbf{1}$ is a vector of 1s.

Inputs The inputs in the optimization problem are the manager's views on the movements of the exchange rates, yield and inflation curves. These views will be translated into expected alphas for each investment strategy through the Black-Litterman model developed in the first chapters.

Results As shown in (3.27), the outputs of the optimization procedure are active weights in the foreign currencies and active durations in the interest rate and inflation-linked bonds.

3.7 Views and Strategies

In order to actively manage their portfolio, the manager must state views on currencies, yield curves and inflation curves.

Currencies The views on the currencies are rather simple, they are expressed in a Bullish/Bearish form. We allow more precise opinions such as "very bearish" or "lightly bullish". Depending on the strength of the opinion, a higher or lower Information Coefficient will be given to it.

Yield Curves We have modeled the yield curves as a function of a few selected term rates. The manager's views must thus be expressed as linear combinations of these term rates. We select three particular terms' rates which we will use in order to state the views: 2 years, 10 years and 30 years. Using these terms, we will allow three different strategies as detailed below.

Duration The manager may express an opinion on a duration strategy. This corresponds to expressing an opinion on the *level* of the yield curve. For instance, if the manager believes the level of the yield curve is going to rise, then they will want to *short* duration, and hence express a bearish opinion on the duration.

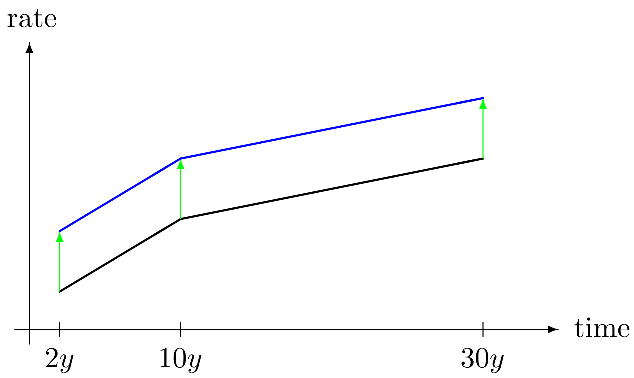


Figure 3.3: Bearish view on duration

2s-10s A view on the 2s-10s is a view on the slope of the yield curve between the 2 year and 10 year terms. It is expressed in a Steepening/Flattening form. If the manager expects the slope of the yield curve to steepen between the 2 year and 10 year terms, then the strategy is to buy the 2 year bond and sell the 10 year bond. Consequently, duration in the 2 year bond will increase whereas duration in the 10 year bond will decrease, which should lead to an overall increase of the duration of the portfolio.

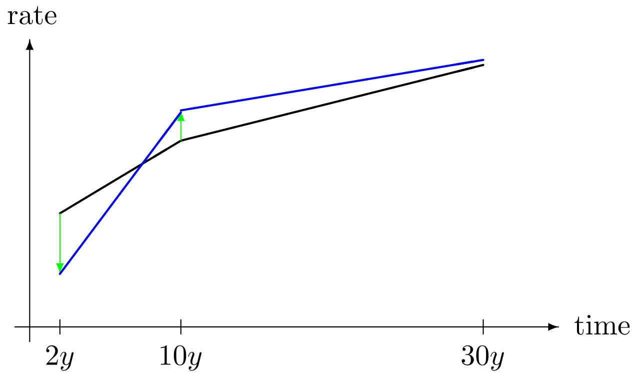


Figure 3.4: Steepening view on the 2s-10s

10s-30s Similarly to the 2s-10s, a view on the 10s-30s is a view on the slope of the yield curve between the 10 year and 30 year terms. It is expressed in a Steepening/Flattening form. If the manager expects the slope of the yield curve to flatten between the 10 year and 30 year terms, then the strategy is to buy the 30 year bond and sell the 10 year bond.

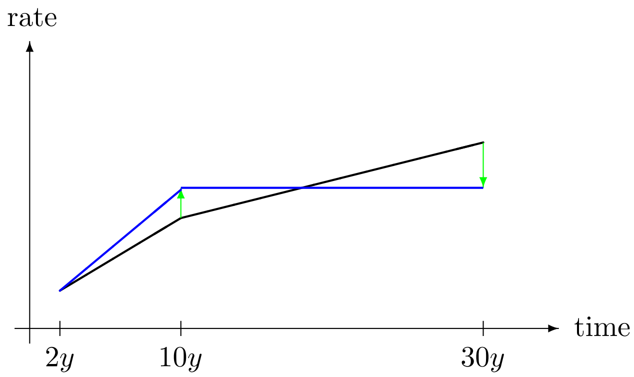


Figure 3.5: Flattening view on the 10s-30s

Inflation We only consider the 10 year inflation-linked bond which we assume to account for changes in the level of the inflation curve. The manager expresses their views in a Higher/Lower form. For instance if they believe that the inflation is going to rise, then they will want to *long* the inflation-linked bond, and hence increase duration in the bond.

3.8 Worked Examples

3.8.1 Universe and Methodology

We are going to apply the established framework to a fixed-income portfolio constituted of currencies, bonds and inflation-linked bonds from the following countries:

Monetary Zone	Currency	2y Bond	10y Bond	30y Bond	Inflation
Australia	AUD	×	×	×	
Canada	CAD	×	×	×	
Switzerland	CHF	×	×	×	
Denmark	DKK	×	×	×	
Eurozone	EUR	×	×	×	×
UK	GBP	×	×	×	×
Hong Kong	HKD	×	×	×	
Japan	JPY	×	×	×	
Norway	NOK	×	×	×	
New Zealand	NZD	×	×	×	
Sweden	SEK	×	×	×	
USA	USD	×	×	×	×

Table 3.1: Benchmark Universe

The portfolio manager expresses their views on the different strategies outlined in the previous section. These views are independent of any base currency and stated *with respect to the benchmark*, so that we are able to use them for any portfolio. The procedure is then to calculate the covariance matrix of the log-returns of the managed assets and apply the active framework developed in Chapter 2 and summarized in section 2.5.2.

3.8.2 Views

We consider the more simple way of expressing views which only allows two states: bearish or bullish, steepening or flattening and higher or lower. The manager has the following views:

Monetary Zone	Duration	2s10s	10s30s	Inflation	Currency
Australia	Bearish				Bullish
Canada	Bearish		Steepening		
Switzerland			Steepening		
Denmark					
Eurozone	Bullish	Flattening	Flattening	Higher	Bullish
UK	Bearish		Steepening	Higher	Bearish
Hong Kong		Flattening			
Japan					Bearish
Norway		Steepening			Bullish
New Zealand					
Sweden	Bullish		Flattening		
USA	Bearish			Lower	

Table 3.2: Manager's views

When no view is stated, it means that the manager has a neutral view on that strategy.

3.8.3 Results

Parameters We assume a breadth of 1 and since all views are equally strong, they are all given an Information Coefficient of 0.2. We also use a historical covariance matrix.

Constraint We impose a risk budget constraint by setting a limit *ex-ante* tracking error not to be overcome. In the example case we set it to 100 bps. We also impose a *self-financing* constraint, which means the changes made to the cash positions in the different currencies must sum up to 0.

Results We present the results in a more practical way. We give the active weights to be taken in the currencies, the durations in the 2s10s, 10s30s and inflation strategies, and the total duration in each monetary zone (excluding inflation). We recall that the active weights are dimensionless whereas the durations have the dimension of time (expressed in years).

The results[1] for a USD-based investor are presented in Table 3.3

[1]calculated on 11 September 2009

Monetary Zone	Total Duration	2s10s	10s30s	Inflation	Currency
Australia	-0.180				1.49%
Canada	-0.496		0.139		
Switzerland	0.053		0.053		
Denmark					
Eurozone	0.200	-0.667	-0.136	0.326	2.60%
UK	-0.623		0.080	0.285	-3.56%
Hong Kong	-0.589	-0.589			
Japan					-1.94%
Norway	0.141	0.141			1.41%
New Zealand					
Sweden	0.580		-0.128		
USA	-0.358			-0.301	

Table 3.3: Results USD-based

We also present the Percentage Contribution to Tracking Error of each strategy in Figure 3.6, in the USD-based scenario. We see that the risk is fairly evenly distributed, which is what we expected since we gave the same Information Coefficient to all strategies.

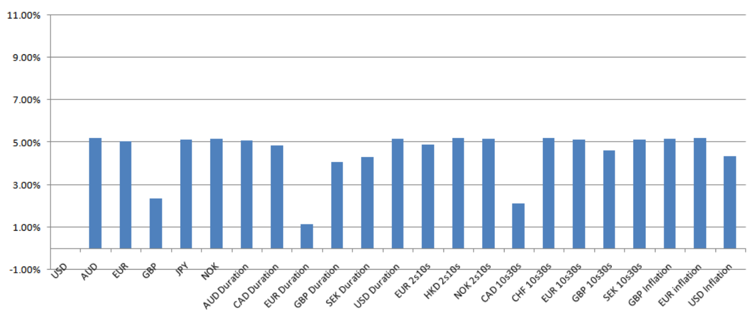


Figure 3.6: PCTEs of the results (USD-based)

We may investigate a case where we want to invest more in currencies than in bonds and hence take more risk in the former. A simple way to do this is to yield a greater IC to currency strategies. In Figure 3.7 we present the results obtained giving an IC of 0.3 to currency strategies and leaving 0.2 to the others. In that case we do observe a greater risk in currencies while still well spread.

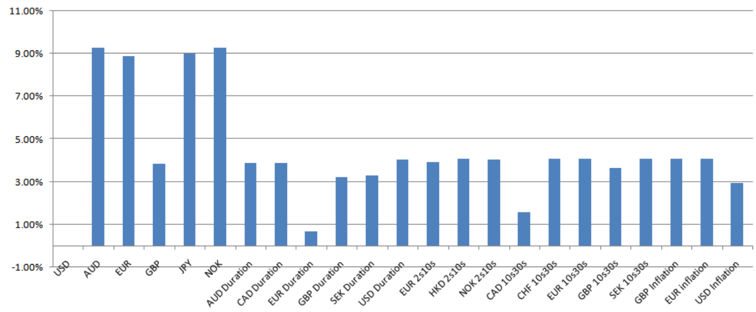


Figure 3.7: PCTEs of the results with different ICs (USD-based)

3.8.4 Refining the views

Methodology

The results in the previous section were derived using a simple expression of the views where they all had the same strength. This is obviously very simplistic, a manager's views will probably not be equally strong, they will prefer some assets more than others and will thus be willing to take more risk in the former.

View Strength	Scaler	IC
Strong	1.5	0.3
Moderate	1	0.2
Light	0.5	0.1
Neutral	0	0

Table 3.4: IC per View

To that end, we allow the manager to express their views along a more detailed ladder such as: very bearish, moderately bearish, lightly bearish, neutral, lightly bullish, moderately bullish and very bullish. Depending on the strength of their opinion, we will yield a more or less important IC to the corresponding strategy.

A simple way to achieve this is to use a scaling scheme: the moderate views will serve as neutral points, with a scaling factor of 1, the strong views will have a scaling factor of 1.5, the light views a scaling factor of 0.5 and finally the neutral views will have a scaling factor of 0. Consistently with the previous section, we yield an Information Coefficient of 0.2 to the moderate views. Consequently, the strong views will have an IC of 0.3, the light views an IC of 0.1 and the neutral views an IC of 0.

Results

The manager then refines the views expressed in Table 3.2 according to Table 3.5.

Monetary Zone	Duration	2s10s	10s30s	Inflation	Currency
Australia	Very Bearish				Lightly Bullish
Canada	Lightly Bearish		Moderately Steepening		
Switzerland			Moderately Steepening		
Denmark					
Eurozone	Very Bullish	Very Flattening	Lightly Flattening	Moderately Higher	Very Bullish
UK	Very Bearish		Very Steepening	Moderately Higher	Very Bearish
Hong Kong		Moderately Flattening			
Japan					Moderately Bearish
Norway		Very Steepening			Moderately Bullish
New Zealand					
Sweden	Moderately Bullish		Lightly Flattening		
USA	Lightly Bearish			Much Lower	

Table 3.5: Manager's refined views

We have now used all the possible strength for the views. We could continue on refining the views and allowing more and more precise specifications of their strength but since the views are expressed only in a qualitative way and given the relatively small number of possible strategies, we believe this is sufficient.

We present the results in Figure 3.8 along with those obtained in section 3.8.3 to see how they have been affected. We observe that the positions have moved according to our refining of the views, stronger positions where we have expressed stronger views and lighter positions where we have expressed lighter views.

We can check on Figure 3.9 which presents the PCTEs of each strategy in this new scenario that the level of risk undertaken in each strategy is indeed proportional to the strength of the corresponding views. Furthermore, we see that strategies with the same level of strength contribute to the overall risk at a similar amount, which ensures a good repartition of risk. This well-balanced distribution of risk is one of the strengths of the Black-Litterman

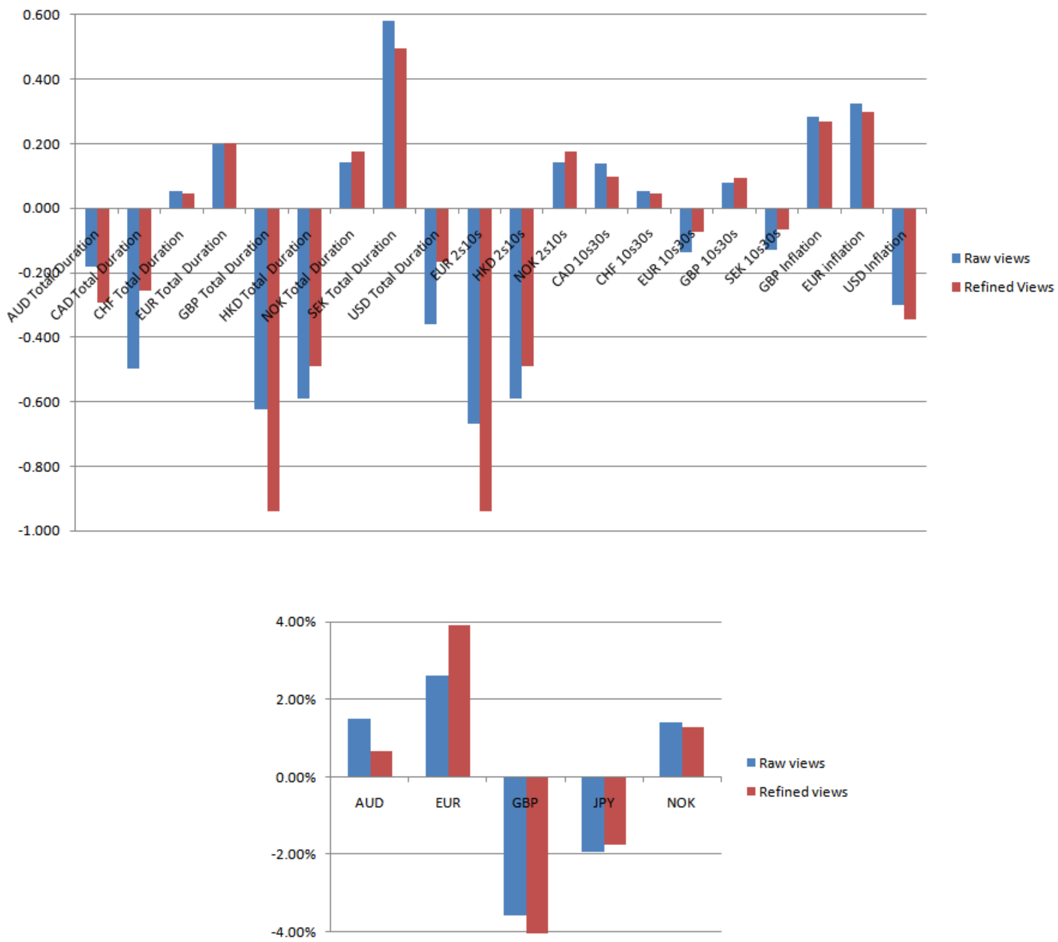


Figure 3.8: Refined views' active positions

model.

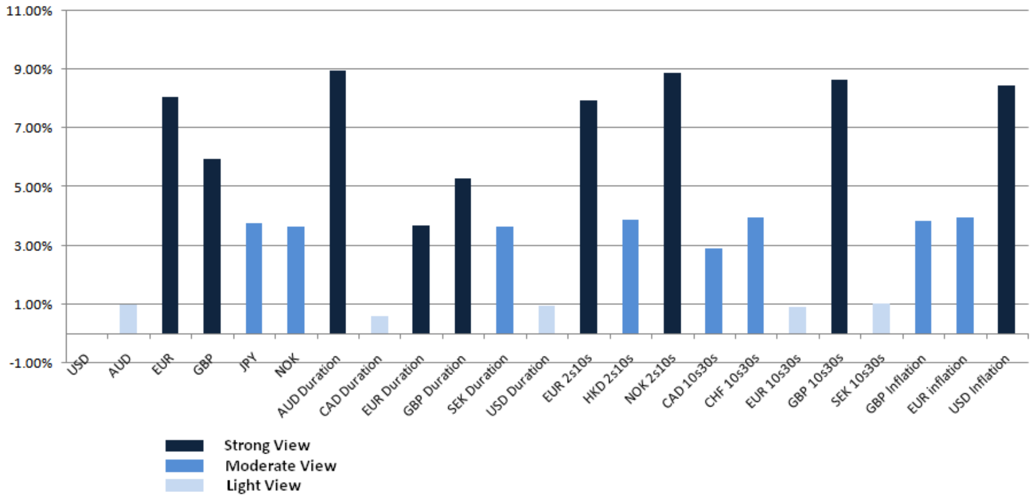


Figure 3.9: Refined views' PCTEs (USD-based)

However, a closer look at Figure 3.9 reveals that some of the strategies seem to have experienced a "downgrade" in the strength of the corresponding views, namely the EUR duration and GBP duration strategies. The answer to this observation lies in the use of the analysis of the views as presented in section 2.6.1. Indeed, looking at the correlation of the EUR and GBP duration strategies, we see that they exhibit a strong correlation of 0.75. In spite of their high correlation we have stated contradictory views, very bullish in EUR duration and very bearish in GBP duration. The Black-Litterman takes into account this inconsistency and tries to correct it for we know that the Black-Litterman model can be interpreted as a distance minimizer. As a consequence the opportunities perceived in each strategy (their alphas), are going to be dampened to reflect this inconsistency, which explains our observation.

3.8.5 Independence of the allocations relatively to the base currency

All the results we have presented in the previous section were results calculated for a USD-based portfolio. Naturally, we may ask ourselves whether the allocations we have derived are going to be modified if the base currency is changed.

In this section we will show that *under a self-financing constraint*, the optimal active allocations of the portfolio are independent of the chosen base-currency, which will enable us to derive the results in a convenient currency and avoid some troublesome transformations.

Universe We recall that the universe we deal with is constituted of currencies and bonds. We thus consider:

- n currencies + USD
- m bonds (rates $(z_j)_{1\leq j\leq m}$)

Covariance matrix We keep similar notations to section 3.4.5 with a few additional changes. We consider the covariance matrix of the log-returns of the FX rates relative to the USD and the log-returns of the bond's rates:

$$\boldsymbol{\Sigma}^{\$} = \begin{pmatrix} \boldsymbol{C} & \boldsymbol{F} \\ \boldsymbol{F}^T & \boldsymbol{G} \end{pmatrix}$$

where

- $\boldsymbol{C}$ is an $(n+1)\times(n+1)$ matrix (covariance matrix of the log-returns of the FX rates):

$$\begin{array}{c} \\ \$/\$ \\ \$/c_1 \\ \vdots \\ \$/c_n \end{array} \begin{array}{c} \begin{array}{cccc} \$/\$ & \$/c_1 & \dots & \$/c_n \end{array} \\ \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & \sigma_{11} & \dots & \sigma_{1n} \\ 0 & \vdots & \vdots & \vdots \\ 0 & \sigma_{1n} & \dots & \sigma_{nn} \end{pmatrix} \end{array}$$

- $\boldsymbol{G}$ is an $m\times m$ matrix (covariance matrix of the log-returns of the bond rates):

$$\begin{array}{c} \\ z_1 \\ z_2 \\ \vdots \\ z_m \end{array} \begin{array}{c} \begin{array}{cccc} z_1 & z_2 & \cdots & z_m \end{array} \\ \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{12} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{1n} & g_{2n} & \cdots & g_{nn} \end{pmatrix} \end{array}$$

- $\boldsymbol{F}$ is an $(n+1) \times m$ matrix (covariance matrix of the log-returns of the FX rates vs log-returns of the bonds rates):

$$\begin{array}{cc} & \begin{array}{cccc} z_1 & z_2 & \dots & z_m \end{array} \\ \begin{array}{c} \$/\$ \\ \$/c_1 \\ \vdots \\ \$/c_n \end{array} & \begin{pmatrix} 0 & 0 & \dots & 0 \\ f_{11} & f_{12} & \dots & f_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{1n} & f_{2n} & \dots & f_{nn} \end{pmatrix} \end{array}$$

Transition matrix We find the transition matrix from the dollar-based universe to the c_i-based universe in the same fashion as in section 3.4.5:

$$Q_{\$ \to c_i} = \begin{pmatrix} \boldsymbol{P}_i & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_m \end{pmatrix}$$

where

$$\boldsymbol{P}_i = \begin{array}{cc} & \begin{array}{cccccccc} \$ & & & & c_i & & & \end{array} \\ \begin{array}{c} \$ \\ \\ \\ \\ c_i \\ \\ \\ \\ \end{array} & \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & \dots & 1 & 0 & 0 & \dots & 0 \\ -1 & -1 & \dots & -1 & 0 & -1 & \dots & -1 \\ 0 & & \dots & 0 & 0 & 1 & & 0 \\ \vdots & & & & \vdots & & \ddots & \vdots \\ 0 & & \dots & & 0 & \dots & \dots & 1 \end{pmatrix} \end{array}$$

and $\boldsymbol{I}_m$ is the $m \times m$ identity matrix.

Expected alphas Let $\boldsymbol{\alpha}^{\$}$ be the vector of expected alphas dollar-based. When switching to a c_i-based universe, the expected alphas resulting from investments in the bonds remain unchanged, whereas the alphas resulting from investments in the currencies experience a translation of $-\alpha_i$, where $-\alpha_i$ is the expected alpha of currency c_i for a USD-based investor.

$$\boldsymbol{\alpha}^{\$} = \begin{array}{c} \$ \\ c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \\ z_1 \\ \vdots \\ z_m \end{array} \begin{pmatrix} 0 \\ \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \\ \alpha_{z_1} \\ \vdots \\ \alpha_{z_m} \end{pmatrix}$$

c_i-based, the expected alphas become:

$$\boldsymbol{\alpha}^{c_i} = \begin{array}{c} \$ \\ c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \\ z_1 \\ \vdots \\ z_m \end{array} \begin{pmatrix} -\alpha i \\ \alpha_1 - \alpha i \\ \vdots \\ 0 \\ \vdots \\ \alpha_n - \alpha i \\ \alpha_{z_1} \\ \vdots \\ \alpha_{z_m} \end{pmatrix} = \boldsymbol{\alpha}^{\$} - \begin{pmatrix} \alpha_i \\ \alpha_i \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Optimization problem Our objective is to maximize the expected alpha of the portfolio with a tracking error constraint and a self-financing constraint. The optimization problem dollar-based writes:

$$\begin{aligned} & \arg\max_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{\alpha}^{\$} \\ s.t. \quad & \boldsymbol{w}^T \boldsymbol{\Sigma}^{\$} \boldsymbol{w} \leq \kappa^2 \\ & \sum_{i=0}^{n} w_i = 0 \end{aligned}$$

c_i-based, it becomes:

$$\begin{aligned} & \arg\max_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{\alpha}^{c_i} \\ s.t. \quad & \boldsymbol{w}^T \boldsymbol{\Sigma}^{c_i} \boldsymbol{w} \leq \kappa^2 \\ & \sum_{i=0}^{n} w_i = 0 \end{aligned}$$

Equivalence of the problems We now show that both problems are in fact equivalent. In order to achieve that, we prove that the function to maximize as well as the constraints are equivalent.

- $$\begin{aligned} \boldsymbol{w}^T \boldsymbol{\alpha}^{c_i} &= \boldsymbol{w}^T \boldsymbol{\alpha}^{\$} - \alpha_i \sum_{i=0}^{n} w_i \\ &= \boldsymbol{w}^T \boldsymbol{\alpha}^{\$} \end{aligned}$$

  So the functions to be maximized are the same in both cases.

- Furthermore:

  $$\begin{aligned} \boldsymbol{w}^T \boldsymbol{\Sigma}^{c_i} \boldsymbol{w} &= \boldsymbol{w}^T (\boldsymbol{Q}^T \boldsymbol{\Sigma}^{\$} \boldsymbol{Q}) \boldsymbol{w} \\ &= (\boldsymbol{Q}\boldsymbol{w})^T \boldsymbol{\Sigma}^{\$} (\boldsymbol{Q}\boldsymbol{w}) \end{aligned}$$

  and

  $$\boldsymbol{Q}\boldsymbol{w} = \begin{pmatrix} 0 \\ w_1 \\ \vdots \\ w_{i-1} \\ -\sum_{\substack{j=0 \\ j\neq i}}^{n} w_j \\ w_{i+1} \\ \vdots \\ w_{n+m} \end{pmatrix} = \boldsymbol{w} - \begin{pmatrix} w_0 \\ 0 \\ \\ \vdots \\ \\ \vdots \\ \\ 0 \end{pmatrix} := \boldsymbol{w} - \boldsymbol{w}_0$$

  where we have used that $-\sum_{\substack{j=0 \\ j\neq i}}^{n} w_j = w_i$ by the self-financing condition. Since $\boldsymbol{\Sigma}^{\$} \boldsymbol{w}_0 = \boldsymbol{0}_{n+1}$, we find that

  $$\boldsymbol{w}^T \boldsymbol{\Sigma}^{c_i} \boldsymbol{w} = \boldsymbol{w}^T \boldsymbol{\Sigma}^{\$} \boldsymbol{w} - 2\boldsymbol{w}^T \boldsymbol{\Sigma}^{\$} \boldsymbol{w}_0 + \boldsymbol{w}_0^T \boldsymbol{\Sigma}^{\$} \boldsymbol{w}_0 = \boldsymbol{w}^T \boldsymbol{\Sigma}^{\$} \boldsymbol{w}$$

  So the constraint conditions are equivalent as well.

Conclusion We have showed in this section that *under a self-financing constraint*, the optimal allocations for a given vector of expected alphas in a reference currency are independent of the portfolio base currency. This enables to optimize the portfolio allocations in any convenient base currency (in particular USD).

3.8.6 Changes in the risk contributions

In the previous section we have showed that under a self-financing constraint, the optimal allocations are independent of the base currency. However, the risk contributions of each strategy may change, in particular the risk undertaken in any base currency is null. The purpose of this section is to investigate the redistribution of the PCTEs when switching from one currency to another.

Let us assume that we have solved the optimization problem for a USD-based investor. According to section 2.6.2, we may derive the MCTEs and PCTEs using equations 2.15 and 2.18:

$$\begin{aligned} \mathbf{MCTE}^{\$} &= \frac{\boldsymbol{\Sigma}^{\$}\boldsymbol{w}_a}{\sigma_a} \\ \mathbf{PCTE}_j^{\$} &= \frac{\boldsymbol{w}_{a_j}\mathbf{MCTE}_j^{\$}}{\sigma_a} \end{aligned}$$

Since the optimal allocations are independent of the base currency, the optimal allocations for a currency c_i-based investor are $\boldsymbol{w}_a$ as well. The corresponding MCTEs and PCTEs are:

$$\begin{aligned} \mathbf{MCTE}^{c_i} &= \frac{\boldsymbol{\Sigma}^{c_i}\boldsymbol{w}_a}{\sigma_a} \\ \mathbf{PCTE}_j^{c_i} &= \frac{\boldsymbol{w}_{a_j}\mathbf{MCTE}_j^{c_i}}{\sigma_a} \end{aligned}$$

But $\boldsymbol{\Sigma}^{c_i}\boldsymbol{w}_a = \boldsymbol{Q}^T\boldsymbol{\Sigma}^{\$}\boldsymbol{Q}\boldsymbol{w}_a$, and we have seen in the previous section that $\boldsymbol{\Sigma}^{\$}\boldsymbol{Q}\boldsymbol{w}_a = \boldsymbol{\Sigma}^{\$}\boldsymbol{w}_a$. Consequently:

$$\begin{aligned} \mathbf{MCTE}^{c_i} &= \frac{\boldsymbol{Q}^T\boldsymbol{\Sigma}^{\$}\boldsymbol{w}_a}{\sigma_a} \\ &= \boldsymbol{Q}^T\mathbf{MCTE}^{\$} \end{aligned}$$

Next,

$$
\mathbf{MCTE}^{c_i} = \boldsymbol{Q}^T \mathbf{MCTE}^{\$} = \begin{array}{c} \$ \\ c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \\ z_1 \\ \vdots \\ z_m \end{array} \begin{pmatrix} -\mathbf{MCTE}^{\$}_i \\ \mathbf{MCTE}^{\$}_1 - \mathbf{MCTE}^{\$}_i \\ \vdots \\ 0 \\ \vdots \\ \mathbf{MCTE}^{\$}_n - \mathbf{MCTE}^{\$}_i \\ \mathbf{MCTE}^{\$}_{z_1} \\ \vdots \\ \mathbf{MCTE}^{\$}_{z_m} \end{pmatrix}
$$

We then deduce the PCTEs:

$$
\mathbf{PCTE}^{c_i} = \frac{\boldsymbol{w}_{a_j} \mathbf{MCTE}^{c_i}_j}{\sigma_a} = \begin{array}{c} \$ \\ c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \\ z_1 \\ \vdots \\ z_m \end{array} \begin{pmatrix} -\frac{\boldsymbol{w}_{a_0} \mathbf{MCTE}^{\$}_i}{\sigma_a} \\ \mathbf{PCTE}^{\$}_1 - \frac{\boldsymbol{w}_{a_1} \mathbf{MCTE}^{\$}_i}{\sigma_a} \\ \vdots \\ 0 \\ \vdots \\ \mathbf{PCTE}^{\$}_n - -\frac{\boldsymbol{w}_{a_n} \mathbf{MCTE}^{\$}_i}{\sigma_a} \\ \mathbf{PCTE}^{\$}_{z_1} \\ \vdots \\ \mathbf{PCTE}^{\$}_{z_m} \end{pmatrix} \tag{3.28}
$$

Conclusion We see in equation 3.28 that the PCTEs of the bond investments remain unchanged when switching the base currency, whereas there is a redistribution of the risk within the currencies.

This was to be expected: if we have a neutral view on the USD and are lightly bearish GBP and lightly bullish EUR, from a GBP perspective we become moderately bullish EUR, and the perceived undertaken risk should hence be greater.

Example We use the same set of views as in section 2.7. Figure 3.10 recalls the PCTEs we found for a USD-based investor.

Figure 3.11 on the other hand shows how these values change when we choose the Japanese Yen as our base currency.

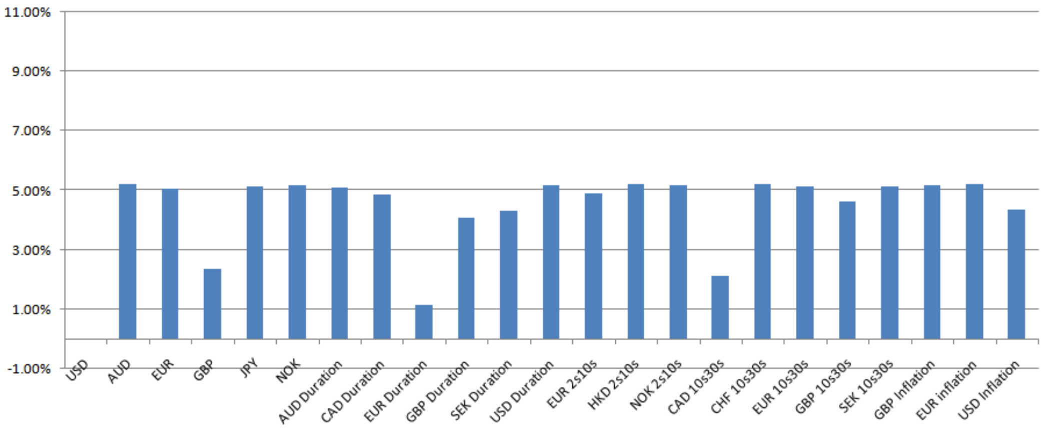


Figure 3.10: PCTEs (USD-based)

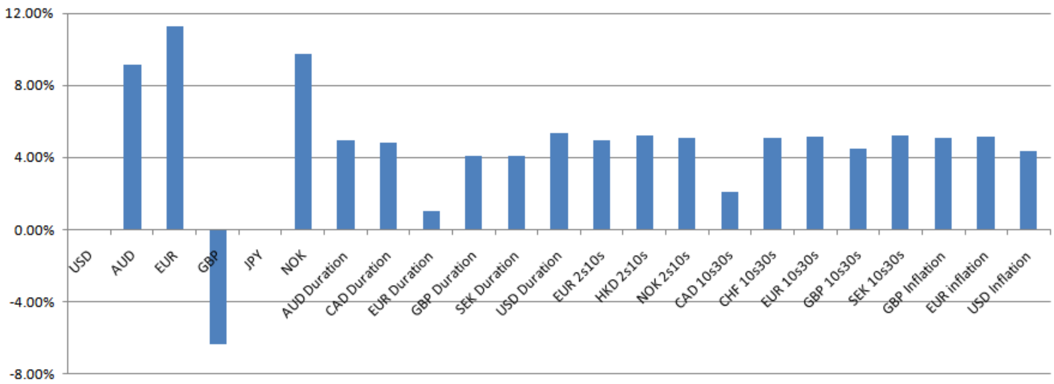


Figure 3.11: PCTEs (Yen-based)

We see that the risks undertaken in EUR, AUD and NOK are perceived as more important for a Yen-based investor. This comes from the fact that we expressed a moderately bearish view on the Yen and a moderately bullish view on the EUR, AUD and NOK; so from a Yen-based point of view, these views are actually strong. Similarly, the view on the GBP is equal to that on the Yen, which enables for diversification. Those two currencies are used to invest in the bullish currencies; stating that the GBP will have a similar performance to the Yen allows to diversify the risk in the short positions, which explains the negative PCTE of the GBP.

3.8.7 Discussion of the results

Now that we have constructed an optimization framework that produces seemingly good results, we may want to conduct an *ex-post* analysis thereof and check their consistency.

At first glance, it seems that although the size of the currency positions are sensible, the duration positions are higher - in absolute terms - than what would be expected.

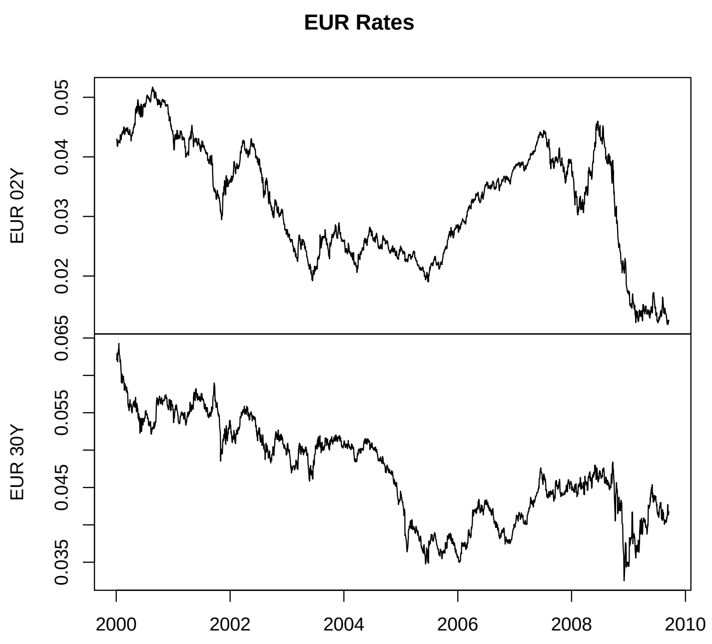


Figure 3.12: 2 year and 30 year EUR rates

We believe an explanation for this phenomenon lies in the fact that due to the historically low interest rates we deal with, the log-normality assumption of the interest rates can no longer be trusted. This highlights one of the limitations of the Black-Litterman model which is the necessity of making assumptions on the distribution of the returns.

Since we make a log-normal assumption on the returns of the interest rates, we must divide by the current interest rates to retrieve the duration positions (see formula 3.19), and given their current extremely low values, the resulting durations shoot up.

In order to circumvent this undesired problem, we suggest to transform the lower rates which are the ones most affected during crises. Indeed, during difficult times, shorter term rates tend to be lowered while long-term rates remain more steady (Figure 3.12). We therefore want to pass the rates through a “filter” which would pull up lower rates and leave unchanged higher rates.

This method is almost purely empirical and the way we choose to conduct it is to apply a function that would resemble an exponential in the lower rates and which would then continuously and smoothly reconnect with the $y(x) = x$ function. We choose the value of this function to be 2% at 0 and the connection to be made at 3.5%. The function is thus of the form:

$$y(x) = \begin{cases} Ae^{Bx} + C & \text{if } 0 \leq x < 3.5\% \\ x & \text{if } x \geq 3.5\% \end{cases} \tag{3.29}$$

where A, B and C are constants calculated in appendix C.1. This transforming function is plotted in Figure 3.13.

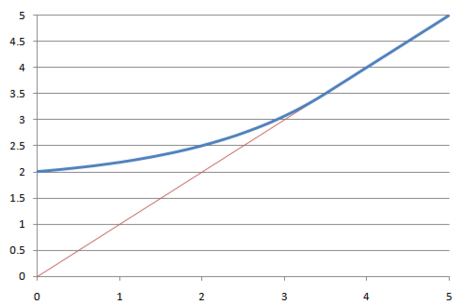


Figure 3.13: Transforming function

Equipped with new values for the current rates obtained passing the rates through the transforming functions, we retrieve new active duration positions - the currency positions remain unchanged. The new positions are represented in Figure 3.14 along with the former positions from our main example (Table 3.2).

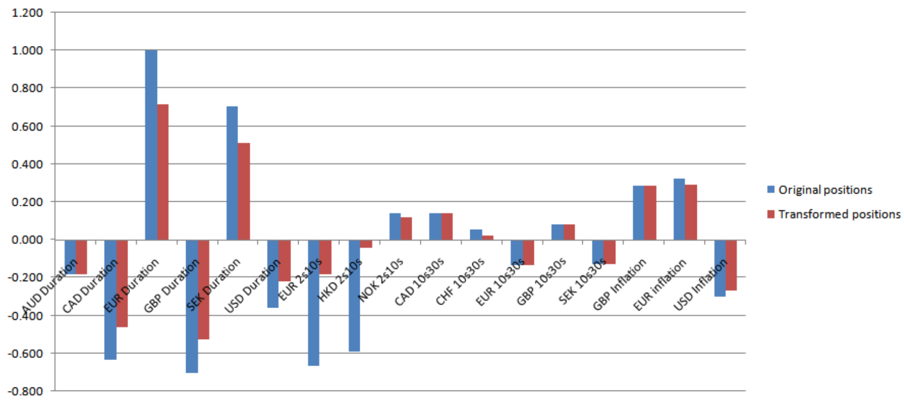


Figure 3.14: Duration positions after modification of the rates

We observe that using the transforming function achieves its objective, that is reducing the active duration positions. Given its empirical nature, the use of this method must therefore be conducted after careful analysis of the original outputs. Some important fine-tuning might be necessary in order for this method to be fully satisfactory.

Chapter 4

Covariance Matrix: a state dependent model

4.1 Introduction

As we have already stressed it, the covariance matrix holds a key role in the portfolio optimization process, and in the Black-Litterman model in particular. In the traditional mean-variance problem, risk is associated to the volatility of the assets which is captured, along with their correlations, by the covariance matrix of their log-returns. Estimating it is thus of utmost importance. Typically, it will be computed using data covering a wide range of time in order to provide with consistent data. This procedure gives an equal weight to all the sample observations of the returns, which might be satisfactory in a long term horizon or in times of great stability, but in times of turbulence or financial crisis it might not be representative of the market behaviour anymore and lead to wrongful investments. Consequently, many authors have discussed the appropriate ways to calculate the covariance matrix, we cite Litterman and Winkelmann in [LW98] in particular who discuss several ways of calculating the covariance matrix, presenting a range of possibilities.

4.2 State Model

In an attempt to remedy the problem, Chow, Jacquier, Kritzman and Lowry [CJKL99] and in a subsequent paper Kritzman, Lowry and Van Royen [KLVR01] suggested a new method to estimate the covariance matrix. It is based on a *state* model of the market, which is characterized as either in a normal or perturbed state. Using this assumption as a starting point, the estimation of a covariance matrix $\mathbf{\Sigma}_Q$ during quiet or unperturbed times and a covariance matrix $\mathbf{\Sigma}_P$ during troubled or perturbed times yields an expression for the actual covariance matrix at time t:

$$\mathbf{\Sigma}_t = \gamma_t \mathbf{\Sigma}_P + (1 - \gamma_t)\mathbf{\Sigma}_Q \tag{4.1}$$

where γ_t is at time t the probability of falling in the perturbed state.

In order to apply this method, we must therefore specify the unknown parameters of the model: $\mathbf{\Sigma}_P$, $\mathbf{\Sigma}_Q$ and the probability of falling into the perturbed state γ_t. Given all these parameters, we will be able to calculate $\mathbf{\Sigma}_t$.

The problem then breaks down into two subordinate problems. The first is to find a way to estimate and separate the two states in order to calculate $\mathbf{\Sigma}_P$ and $\mathbf{\Sigma}_Q$, and the second is to evaluate γ_t, the probability of falling into the perturbed state.

We fill focus on the first problem and present a method to separate states and some results as to how they affect the active allocations of the portfolio.

4.3 Characterizing states

4.3.1 Principle

We assumed the existence of a quiet and a perturbed state but we have not given a proper sense to what these terms refer to. While it is sometimes obvious from market observations whether we find ourselves in one or the other state, we need to establish a way to link the data to a corresponding state. We need to ask ourselves what characterizes and differentiates market states.

Perturbed times are characterized by great changes in returns of the assets as well as in their variance and covariances, which is the very reason why they make portfolio optimization problems more complicated since we cannot rely on previous covariance matrices.

We would like to find a value that would incorporate precisely these changes and would quantify how far away from quiet times we find ourselves at a particular time. Such a variable can be found in the Mahalanobis distance.

We consider over the discrete time horizon $t_0, t_1, \ldots, t_n$ a time dependent $n \times 1$ vector $\boldsymbol{y}_t$ containing the log-returns at time t of n different assets or indices. Let $\bar{\boldsymbol{y}}$ be the mean of $\boldsymbol{y}_t$ over this period and $\boldsymbol{\Sigma}$ be its covariance matrix. The Mahalanobis distance d_t at time t associated to $\boldsymbol{y}_t$ is defined as:

$$d_t^2 = (\boldsymbol{y}_t - \bar{\boldsymbol{y}})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}_t - \bar{\boldsymbol{y}}) \tag{4.2}$$

For multivariate variables, the square of the Mahalanobis distance approximately follows a chi-square distribution with n degrees of freedoms. Since we assume log-normality of the returns, $\boldsymbol{y}_t$ is indeed multivariate. Therefore, we may assign a probability to each value d_t thanks to the cumulative probability function of the chi-square function:

$$\mathbb{P}(d_t) = 1 - F_{\chi^2}(d_t^2, n) \tag{4.3}$$

where $F_{\chi^2}(, n)$ is the cumulative probability function of the chi-square function with n degrees of freedom. Thus we may fix a probability threshold and find the corresponding distance d_0. Then all the dates with $d_t > d_0$ will be classified as perturbed market dates and the dates such that $d_t \leq d_0$ will be classified as quiet market dates. The frontier drawn by all the possible vectors with a Mahalanobis distance equal to d_0 will represent the limit between both states.

Let us consider an over-simplified example where $\boldsymbol{y}_t^{ex}$ is a 2×1 vector of normal random variables with covariance matrix $\boldsymbol{\Sigma}^{ex}$ and mean $\bar{\boldsymbol{y}}^{ex}$.

$$\boldsymbol{\Sigma}^{ex} = \begin{pmatrix} \sigma_1 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2 \end{pmatrix} \qquad \bar{\boldsymbol{y}}^{ex} = \begin{pmatrix} \bar{\boldsymbol{y}}_1 \\ \bar{\boldsymbol{y}}_2 \end{pmatrix}$$

We simulate 50 realizations of $\boldsymbol{y}^{ex}$ and plot them along with the frontier corresponding to a probability of $\mathbb{P}(d) = 0.8$ in Figure 4.1[1].

The points outside the ellipse representing the probability threshold are considered to be outliers and would correspond to the perturbed state whereas the points inside the ellipse would correspond to the quiet state.

In this simple example, the frontier is drawn with an ellipse, with more variables and hence more dimensions it would be an ellipsoid.

All this previous discussion has equipped us with a means of characterizing quiet and perturbed states with respect to a certain threshold. The question is now to know where to set this threshold.

[1] $\sigma_1 = 1, \sigma_2 = 5, \sigma_{1,2} = 1.2 \qquad \bar{\boldsymbol{y}}_1 = 2, \bar{\boldsymbol{y}}_2 = 3$

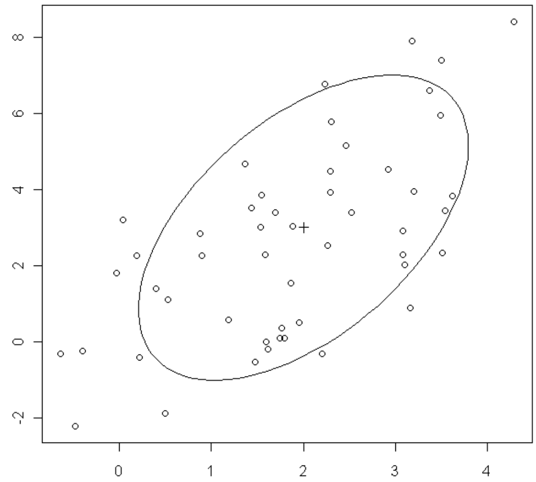


Figure 4.1: $\mathbb{P}(d) = 0.8$ frontier

4.3.2 Methodology

We here present a methodology to apply the tool we found in the Mahalanobis distance in order to characterize the state of the market. This methodology is as much based on theoretical results as on empirical ones.

The financial market is driven by a lot of factors. In order to characterize its state, we must thus use a set of assets and indices which best describes the overall activity of the market. We select a handful of time series which we believe will reflect market drivers and provide with good indicators. These indices are chosen to be:

FX Volatility Index We design an index in order to reflect the bigger movements of the FX market. It is a weighted index of volatility indices of the most important FX crosses: AUD/USD, CHF/USD, EUR/USD, GBP/USD, JPY/USD.

VIX The CBOE Volatility Index (VIX) is a key measure of market expectations of near-term volatility conveyed by S&P 500 stock index option prices. Since its introduction in 1993, VIX has been considered by many to be the world's premier barometer of investor sentiment and market volatility.[2]

EMBI+ The JP Morgan Emerging Markets Bond Index Plus (EMBI+)

[2]Source: `http://www.cboe.com/micro/vix/introduction.aspx`

tracks returns for actively traded external debt instruments in emerging market[3].

CORPB Spread of BBB-rated American industries.

USSP10 US 10 year semi-annual Swap Spread. The swap spread reflects the risk premium that is involved in a swap transaction instead of holding risk-free government bonds.

MOVE The Merrill Option Volatility Expectations Index (MOVE) reflects a market estimate of future Treasury bond yield volatility.

Once we have pulled out the data series for those indices, we compute their log-returns $\boldsymbol{y}_t$ and their mean vector $\bar{\boldsymbol{y}}$ and covariance matrix $\boldsymbol{\Sigma}$ in order to compute the Mahalanobis distances between $\boldsymbol{y}_t$ and $\bar{\boldsymbol{y}}$ according to relation 4.2:

$$d_t^2 = (\boldsymbol{y}_t - \bar{\boldsymbol{y}})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}_t - \bar{\boldsymbol{y}})$$

We have applied this methodology to time series of the indices spanning over almost 11 years, from September 1998 to August 2009. We plot d_t in Figure 4.2 and signal out the values overcoming a certain threshold. We observe that the signal is extremely noisy. This was to be expected since even in quiet periods, some out-of-the-ordinary trading days may occasionally occur and inversely, quiet days may occur in perturbed times.

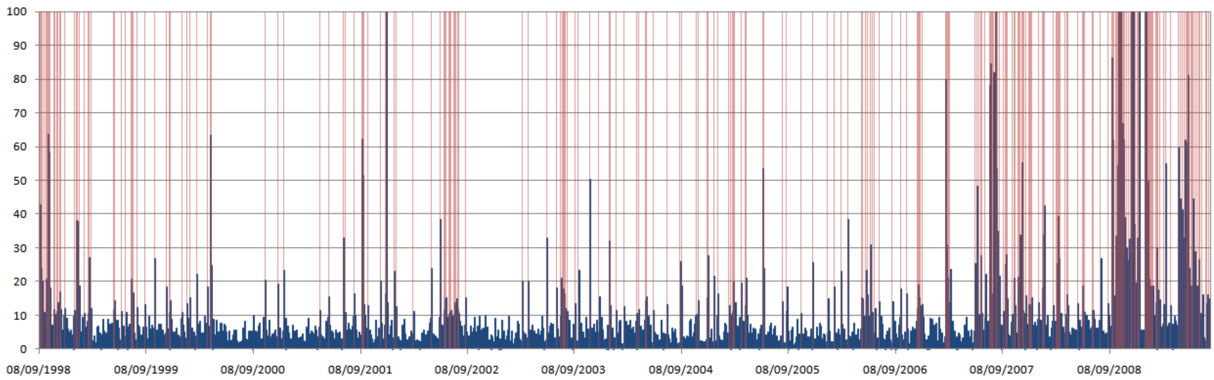


Figure 4.2: Raw Mahalanobis distance

In order to overcome this problem, we filter out the noise by smoothing the data through a 30 day moving averaging. The result obtained after this

[3]Source: `http://www.jpmorgan.com/pages/jpmorgan/investbk/solutions/research/EMBI`

operation is a great improvement over the first attempt and we present it in Figure 4.3.

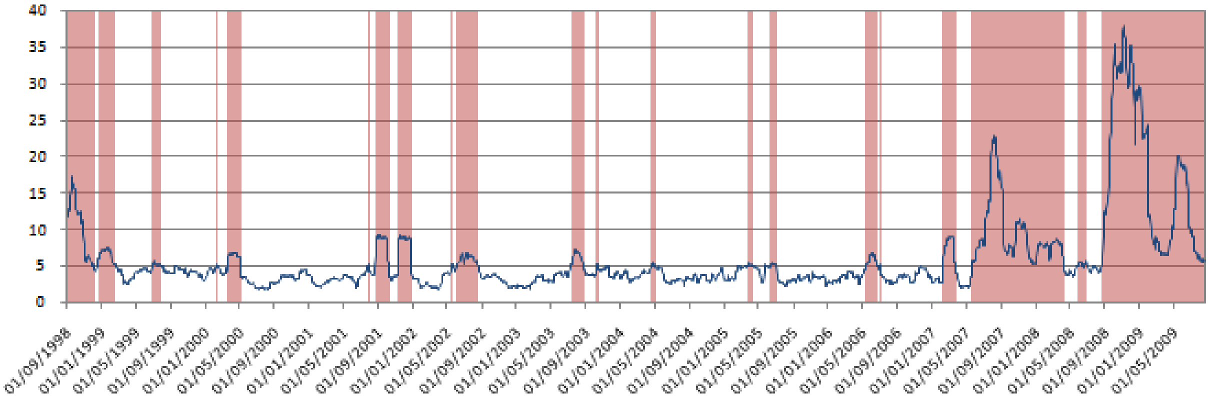


Figure 4.3: Smoothed Mahalanobis distance

We observe in this new figure that the studied time interval is well split in ranges of time of either high or low values of d_t, with very little residual noise. This arises from the moving averaging which yields to any date t some of the properties of its neighbour dates. Thus, if a date t has a low value of d_t but many of its surrounding dates exhibit crisis values of their Mahalanobis distance, the former date will probably be recast as a crisis date. The inverse also holds.

We must though investigate whether in spite of the apparently good results we have obtained, these are indeed meaningful. To achieve that, we try to link the perturbed states brought out by the analysis to known financial crises or periods of high activity; and conversely check whether most important crises are picked up.

We present in Figure 4.4 a brief identification of potentially perturbed states to known events in the financial markets. It does seem like we are able to link financial crises and major events to the periods we have brought out by our analysis since we are able to identify the bigger incidents which affected the market during the last decade: Russian default and Long-Term Capital Management (LTCM) collapse in 1998, the burst of the internet and new technologies bubble in 2000, 9/11 and the accounting scandals of Enron and WorldCom among others in 2001 and 2002 or the more recent subprime and credit-crunch crises to name a few.

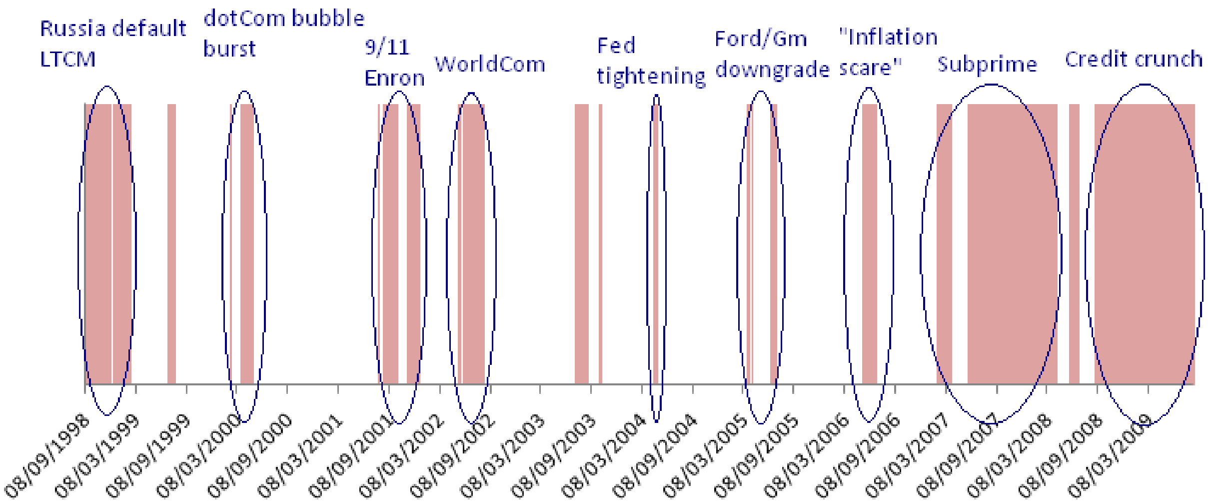


Figure 4.4: Identification of crisis times

4.4 Examples

We are now able to separate perturbed periods from quiet periods and we may therefore compute the corresponding covariance matrices $\mathbf{\Sigma}_P$ and $\mathbf{\Sigma}_Q$. Given γ_t, the probability of being in a perturbed state during the next period, we retrieve a new covariance matrix $\mathbf{\Sigma}_t$ to be used in the optimization process. In this section, we will present some results to investigate how this method affects the results found in section 3.8.3.

In order to do this, we keep the same views on the market as in Table 3.2 and the same parameters (i.e. an IC of 0.2 for all strategies and a breadth of 1).

We then solve the optimization problem for different values of γ_t and present the results in Figure 4.5. We observe that in general, the active positions exhibit a larger sensitivity to γ_t for low values of γ_t, between 0 and 0.3/0.4. For larger values of γ_t, there is a certain stability in the active positions. It is also interesting to see that the positions found with the whole historical matrix as in section 3.8.3, which are roughly found using $\gamma_t \approx 0.32$ (the ratio of the number of crisis dates over the total of observed dates), are approximately half-way between the most extreme positions corresponding to $\gamma_t = 0$ and $\gamma_t = 1$.

Finally, we may find the results somewhat intuitive. Indeed, we see that as γ_t grows, so the more weight we give to the perturbed state, the smaller the active positions are. This is what we would expect since in times of crisis the assets tend to be much more volatile and hence have a larger variance. Because the tracking error constraint remains unchanged, in order to stay beneath that limit, we must restrain the positions, which is what we observe.

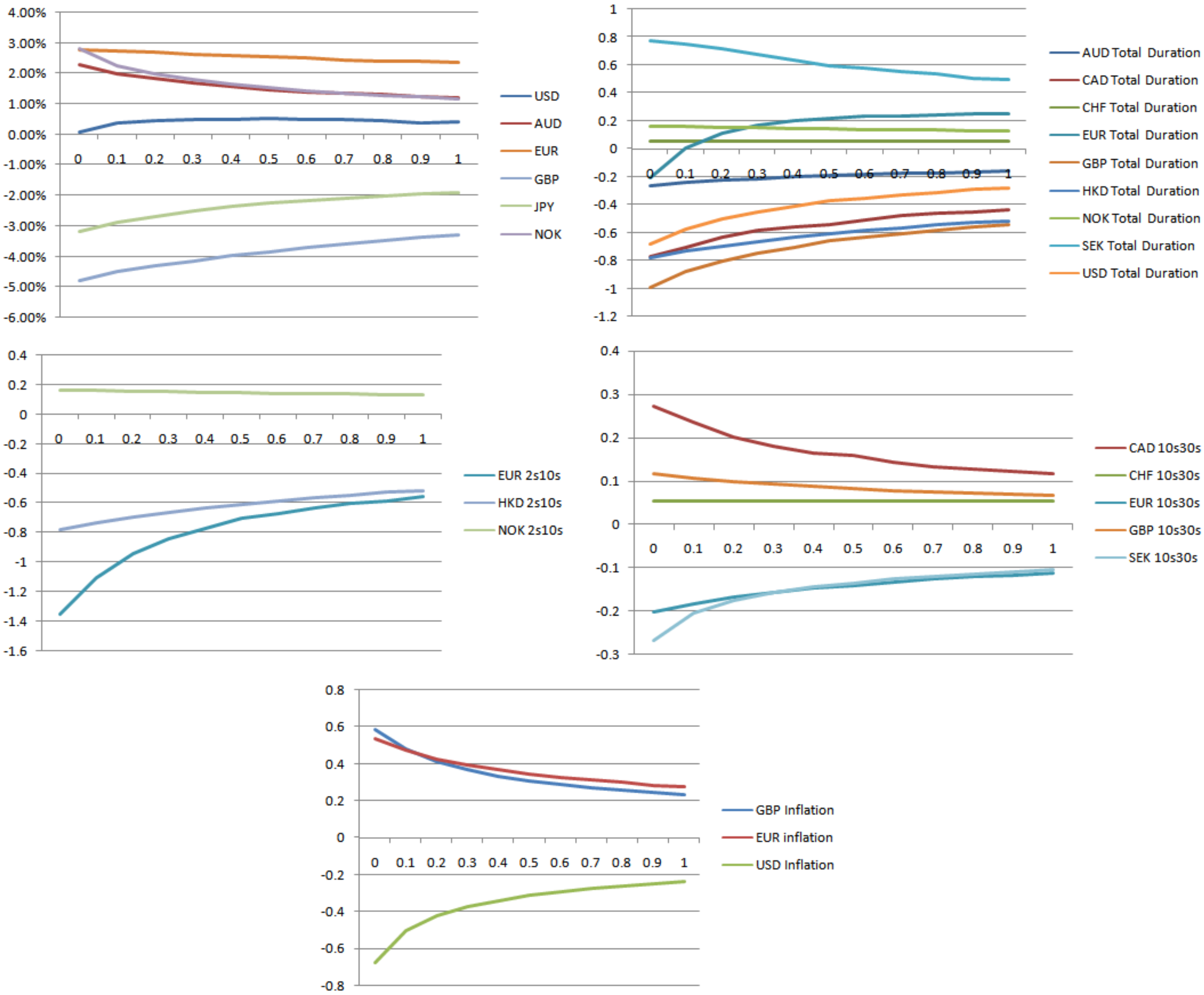


Figure 4.5: Influence of γ_t on the active positions

Conclusion

We have studied how to actively manage a multi-currency fixed-income portfolio comprising bonds, inflation-linked bonds and currencies using the Black-Litterman model. The Black-Litterman model represents a great tool in allocating the portfolio's weights and is a great improvement over the classical mean-variance optimization process developed by Markowitz more than 50 years ago. In particular it does a great job allocating risk according to the manager's confidence in their views and offers a greater flexibility in the way in which they can be expressed.

However, in spite of its many qualities and advantages over other allocating methods, the Black-Litterman model still suffers from some constraints. The assumptions on the distribution of the returns, the parameters τ and $\mathbf{\Omega}$ and the imposed linearity of the expression of the views are all factors which hinder its optimal use.

A possible way to improve some of these constraints would be to explore further the analogy with the Kalman filter that we presented in chapter 1. Using the advances in filtering and solutions to non-linear filtering problems (Extended Kalman Filter, Unscented Kalman filter and Particle Filtering), we might find ways to use the Black-Litterman model without making any assumption regarding the distribution of the returns and relax the linear condition on the views to allow for non-linear views. This would for instance enable us to overcome the problem faced in section 3.8.7 where we were handicapped by the failure of the log-normal assumption of the interest rates.

Finally, in our discussion of the covariance matrix, we focused on how to distinguish perturbed states from quiet states but it still remains to study how to determine the probability of falling into a perturbed state to make the method complete. An idea might be to look for autoregressive properties of the Mahalanobis distance, or using Hidden Markov Chains to predict what this value might be.

References

[BF03] Daniel Blamont and Nick Firoozye. Quantitative Strategy, TOPIC 1: Bayesian Asset Allocation: Black-Litterman. *Deutsche Bank Fixed Income Weekly*, pages 99–103, December 2003.

[BH97] Robert G. Brown and Patrick Y.C. Hwang. *Introduction to random Signals and Applied Kalman Filtering.* John Wiley & Sons, Hoboken, New Jersey, third edition, 1997.

[BL91] Fischer Black and Robert Litterman. Asset Allocation: Combining Investor Views with Market Equilibrium. *The Journal of Fixed Income*, 1(2):7 –, September 1991.

[BL92] Fischer Black and Robert Litterman. Global Portfolio Optimization. *Financial Analysts Journal*, 48(5):28–43, September/October 1992.

[BN07] Maria D. Braga and Francesco P. Natale. TEV Sensitivity to Views in Black-Litterman Model. *SSRN eLibrary*, 2007.

[BW98] Andrew Bevan and Kurt Winkelmann. Using the Black-Litterman Global Asset Allocation Model: Three Years of Practical Experience. *Goldman Sachs Fixed Income Research Paper*, 1998.

[Che09] Wing Cheung. The Black-Litterman Model Explained. *SSRN eLibrary*, February 2009.

[Chr02] George A. Christodoulakis. Bayesian Optimal Portfolio Selection: the Black-Litterman Approach. Notes for Quantitative Asset Pricing, MSc Mathematical Trading and Finance, Sir John Cass Business School, November 2002.

[CJKL99] George Chow, Eric Jacquier, Mark Kritzman, and Kenneth Lowry. Optimal Portfolios in Good Times and Bad. *Financial Analysts Journal*, 55(3):65–73, May/June 1999.

[DGH+06] Lev Dynkin, Anthony Gould, Jay Hyman, Vadim Konstantinovsky, and Bruce Phelps. *Quantitative Management of Bond Portfolios*. Princeton University Press, Princeton, New Jersey, 2006.

[Dro01] Wolfgang Drobetz. How to avoid the pitfalls in portfolio optimization? Putting the Black-Litterman approach at work. *Financial Markets and Portfolio Management*, 15(1):59–75, March 2001.

[FKPF07] Frank J. Fabozzi, Peter N. Kolm, Dessislava A. Pachamanova, and Sergio M. Focardi. *Robust Portfolio Optimization and Management*. Wiley Finance. John Wiley & Sons, Hoboken, New Jersey, 2007.

[FM03] Gianluca Fusai and Attilio Meucci. Assessing Views. *Risk*, 16(3):S18–S21, March 2003.

[GK00] Richard C. Grinold and Ronald N. Kahn. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. McGraw-Hill, New York, NY, second edition, 2000.

[Her03] Ulf Herold. Portfolio Construction with Qualitative Forecasts. *Journal of Portfolio Management*, 30(1):61–72, 2003.

[HL99] Guangliang He and Robert Litterman. The Intuition Behind Black-Litterman Model Portfolios. *Goldman Sachs Asset Management Working Paper*, 1999.

[Idz04] Thomas M. Idzorek. A Step-by-Step Guide to the Black Litterman Model: Incorporating User-Specified Confidence Levels. Working paper series, Zephyr Associates, 2004.

[Kal60] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[KLVR01] Mark Kritzman, Kenneth Lowry, and Anne-Sophie Van Royen. Risk, Regimes, and Overconfidence. *Journal of Derivatives*, 8(3):32–42, Spring 2001.

[LG03] Robert Litterman and Quantitative Resources Group. *Modern Investment Management: An Equilibrium Approach*. John Wiley & Sons, Hoboken, New Jersey, 2003.

[LL01] Wai Lee and Daniel Y. Lam. Implementing optimal risk Budgeting. *The Journal of Portfolio Management*, 28(1), Fall 2001.

[LW98] Robert Litterman and Kurt Winkelmann. Estimating Covariance Matrices. *Goldman Sachs Risk Management Series*, January 1998.

[Man06] Charlotta Mankert. The Black-Litterman Model - mathematical and behavioral finance approaches towards its use in practice. Master's thesis, Royal Institute of Technology, Stockholm, Sweden, June 2006.

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

[Meu05] Attilio Meucci. *Risk and Asset Allocation*. Springer Finance. Springer Verlag, New York, New York, 2005.

[Meu06] Attilio Meucci. Beyond Black-Litterman in Practice: A Five-Step Recipe to Input Views on Non-Normal Markets. *SSRN eLibrary*, 2006.

[Meu08] Attilio Meucci. The Black-Litterman Approach: Original Model and Extensions. *SSRN eLibrary*, 2008.

[Mic89] Richard O. Michaud. The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, 45(1):31–42, January/February 1989.

[Pez07] Jacques Pezier. Global Portfolio Optimization Revisited: a Least Discrimination Alternative to Black-Litterman. ICMA Center Discussion Papers in Finance, ICMA Center, University of Reading, July 2007.

[RHBF08] Svetlozar T. Rachev, John S. J. Hsu, Biliana S. Bagasheva, and Frank J. Fabozzi. *Bayesian Methods in Finance*. John Wiley & Sons, Hoboken, New Jersey, 2008.

[Rol92] Richard Roll. A Mean-Variance Analysis of Tracking Error. *The Journal of Portfolio Management*, 18(4):13–22, Summer 1992.

[Sch07] Bernd Scherer. *Portfolio Construction and Risk Budgeting*. Risk Books, third edition, 2007.

[SdSLP09] Alexandre Schutel da Silva, Wai Lee, and Bobby Pornrojnangkool. The Black-Litterman Model for Active Management. *Journal of Portfolio Management*, 35(2):61–70, Winter 2009.

[Sha74] William Sharpe. Imputing Expected Returns From Portfolio Composition. *Journal of Financial and Quantitative Analysis*, pages 463–472, June 1974.

[SS00] Stephen Satchell and Alan Scowcroft. A Demystification of the Black-Litterman Model: Managing Quantitative and Traditional Portfolio Construction. *Journal of Asset Management*, 1(2):138–150, September 2000.

[TS09] Mateo Eduardo Trujillo Segura. Construcción y Gestión de Portafolios con el Modelo Black-Litterman: Una Aplicación a los Fondos de Pensiones Obligatorias en Colombia. Master's thesis, Universidad de los Andes, Bogota, Columbia, January 2009.

[Wal09] Jay Walters. The Black-Litterman Model in Detail. *SSRN eLibrary*, 2009.

[WB06] Greg Welch and Gary Bishop. An Introduction to the Kalman Filter. Working paper series, Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, July 2006.

[Wei07] Edward D. Weinberger. The Black Litterman Model and the Kalman Filter. Working paper series, Department of Finance and Risk Engineering, Polytechnic University and Aloha Enterprises, Inc., New York, NY, October 2007. Available at `http://www.alohaenterprises.net/BlowhardCentral/quantfinance/BLandKF.htm`.

Appendix A

Appendix to Chapter 1

A.1 Transformations in the proof of the Bayesian approach

In section (1.4.1), we derive the Black-Litterman formula using a Bayesian approach. There are two tedious tasks that need to be carried out:

1. Prove that $(a - \boldsymbol{m}^t \boldsymbol{K}^{-1} \boldsymbol{m})$ is in fact equal to $(\boldsymbol{q} - \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])^t \left(\boldsymbol{P}\boldsymbol{\Sigma}_\mu \boldsymbol{P}^t + \boldsymbol{\Omega}\right)^{-1} (\boldsymbol{q} - \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])$, where:

$$
\begin{aligned}
\boldsymbol{K} &= (\boldsymbol{\Sigma}_{\boldsymbol{\mu}})^{-1} + \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{P} \\
\boldsymbol{m} &= (\boldsymbol{\Sigma}_{\boldsymbol{\mu}})^{-1} \mathbb{E}[\boldsymbol{\mu}] + \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q} \\
a &= \mathbb{E}[\boldsymbol{\mu}]^t (\boldsymbol{\Sigma}_{\boldsymbol{\mu}})^{-1} \mathbb{E}[\boldsymbol{\mu}] + \boldsymbol{q}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q}
\end{aligned}
$$

2. Prove that the constant we got rid of along the way does indeed transform as we wish, i.e.:

$$
\begin{aligned}
\frac{|\boldsymbol{\Sigma}_\mu|^{\frac{-1}{2}} |\boldsymbol{\Omega}|^{-\frac{1}{2}}}{(2\pi)^{\frac{n+K}{2}}} &= \frac{|\boldsymbol{P}\boldsymbol{\Sigma}_\mu \boldsymbol{P}^t + \Omega|^{-\frac{1}{2}} |\boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{P} + \boldsymbol{\Sigma}_\mu^{-1}|^{\frac{1}{2}}}{(2\pi)^{\frac{n+K}{2}}} \\
|\boldsymbol{\Omega} + \boldsymbol{P}\boldsymbol{\Sigma}_\mu \boldsymbol{P}^t| &= |\boldsymbol{\Sigma}_\mu| |\boldsymbol{\Omega}| |\boldsymbol{\Sigma}_\mu^{-1} + \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{P}|
\end{aligned}
$$

Let us prove these two points:

Proof. 1.

$$
\begin{aligned}
& \boldsymbol{m}^t \boldsymbol{K}^{-1} \boldsymbol{m} \\
= & \left(\boldsymbol{\Sigma}_\mu^{-1} \mathbb{E}[\boldsymbol{\mu}] + \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q}\right)^t \boldsymbol{K}^{-1} \left(\boldsymbol{\Sigma}_\mu^{-1} \mathbb{E}[\boldsymbol{\mu}] + \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q}\right) \\
= & \mathbb{E}[\boldsymbol{\mu}]^t \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{K}^{-1} \boldsymbol{\Sigma}_\mu^{-1} \mathbb{E}[\boldsymbol{\mu}] + \boldsymbol{q}^t \boldsymbol{\Omega}^{-1} \boldsymbol{P} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q} \\
& \qquad + 2\mathbb{E}[\boldsymbol{\mu}]^t \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q}
\end{aligned}
$$

This yields:

$$\begin{aligned}
& a - \boldsymbol{m}^t \boldsymbol{K}^{-1} \boldsymbol{m} \\
= \quad & \mathbb{E}[\boldsymbol{\mu}]^t \left(\boldsymbol{\Sigma}_\mu^{-1} - \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{K}^{-1} \boldsymbol{\Sigma}_\mu^{-1} \right) \mathbb{E}[\boldsymbol{\mu}] + \boldsymbol{q}^t \left(\boldsymbol{\Omega}^{-1} \boldsymbol{P} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \right) \boldsymbol{q} \\
& \qquad -2\mathbb{E}[\boldsymbol{\mu}]^t \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q}
\end{aligned} \tag{A.1}$$

We will now rearrange the terms using the Sherman-Morrison-Woodburry formula:

Theorem A.1 (Sherman-Morrison-Woodburry formula)**.**

$$(\boldsymbol{A} + \boldsymbol{U}\boldsymbol{C}\boldsymbol{V})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{U} \left(\boldsymbol{C}^{-1} + \boldsymbol{V}\boldsymbol{A}^{-1}\boldsymbol{U} \right)^{-1} \boldsymbol{V}\boldsymbol{A}^{-1}$$

where $\boldsymbol{A}$, $\boldsymbol{U}$, $\boldsymbol{C}$ and $\boldsymbol{V}$ all denote matrices of the correct size and invertible when required in the formula.

We also write:

$$\boldsymbol{G} = \boldsymbol{\Omega} + \boldsymbol{P}\boldsymbol{\Sigma}_\mu \boldsymbol{P}^t$$

Applying the Sherman-Morrison-Woodburry formula, we find:

$$\begin{aligned}
\boldsymbol{K}^{-1} &= \boldsymbol{\Sigma}_\mu - \boldsymbol{\Sigma}_\mu \boldsymbol{P}^t \boldsymbol{G}^{-1} \boldsymbol{P} \boldsymbol{\Sigma}_\mu \\
\boldsymbol{G}^{-1} &= \boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1} \boldsymbol{P} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1}
\end{aligned}$$

Substituting these two equalities into (A.1):

$$\begin{aligned}
& a - \boldsymbol{m}^t \boldsymbol{K}^{-1} \boldsymbol{m} \\
= \quad & (\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])^t \boldsymbol{G}^{-1} (\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}]) + \boldsymbol{q}^t \boldsymbol{G}^{-1} \boldsymbol{q} - 2\mathbb{E}[\boldsymbol{\mu}]^t \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q}
\end{aligned}$$

This last equality shows that the final step in the proof is establishing that $\mathbb{E}[\boldsymbol{\mu}]^t \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \boldsymbol{q} = (\boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])^t \boldsymbol{G}^{-1} \boldsymbol{q}$, for then we will be able to conclude that indeed:

$$(a - \boldsymbol{m}^t \boldsymbol{K}^{-1} \boldsymbol{m}) = (\boldsymbol{q} - \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])^t \boldsymbol{G}^{-1} (\boldsymbol{q} - \boldsymbol{P}\mathbb{E}[\boldsymbol{\mu}])$$

We have:

$$\begin{aligned}
& \boldsymbol{\Sigma}_\mu^{-1} \boldsymbol{K}^{-1} \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \\
= \quad & \left(\boldsymbol{I} - \boldsymbol{P}^t \boldsymbol{G}^{-1} \boldsymbol{P} \boldsymbol{\Sigma}_\mu \right) \boldsymbol{P}^t \boldsymbol{\Omega}^{-1} \\
= \quad & \boldsymbol{P}^t \left(\boldsymbol{I} - \boldsymbol{G}^{-1} \boldsymbol{P} \boldsymbol{\Sigma}_\mu \boldsymbol{P}^t \right) \boldsymbol{\Omega}^{-1}
\end{aligned} \tag{A.2}$$

But the Sherman-Morrison-Woodburry formula applied to $\boldsymbol{G}^{-1}$ when identifying $\boldsymbol{UC}$ to the identity matrix gives:

$$\begin{aligned}\boldsymbol{G}^{-1} &= \boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}\left(\boldsymbol{I} + \boldsymbol{\Sigma}_\mu \boldsymbol{P}\boldsymbol{\Sigma}_\mu^{-1}\boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\right)^{-1}\boldsymbol{P}\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\\ &= \boldsymbol{\Omega}^{-1} - \left(\boldsymbol{\Omega} + \boldsymbol{\Sigma}_\mu \boldsymbol{P}\boldsymbol{\Sigma}_\mu^{-1}\boldsymbol{P}^t\right)^{-1}\boldsymbol{P}\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\\ &= \left(\boldsymbol{I} - \boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t\right)\boldsymbol{\Omega}^{-1}\end{aligned}$$

which inserted into (A.2) yields:

$$\boldsymbol{\Sigma}_\mu^{-1}\boldsymbol{K}^{-1}\boldsymbol{P}^t\boldsymbol{\Omega}^{-1} = \boldsymbol{P}^t\boldsymbol{G}^{-1}$$

This concludes the proof.

2.

$$\begin{aligned}|\boldsymbol{\Sigma}_\mu||\boldsymbol{\Omega}||\boldsymbol{\Sigma}_\mu^{-1} + \boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\boldsymbol{P}| &= |\boldsymbol{\Omega}||\boldsymbol{\Sigma}_\mu\left(\boldsymbol{\Sigma}_\mu^{-1} + \boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right)|\\ &= |\boldsymbol{\Omega}||\boldsymbol{I} + (\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t)(\boldsymbol{\Omega}^{-1}\boldsymbol{P})|\end{aligned}$$

Sylvester's determinant theorem states that:

Theorem A.2 (Sylvester's determinant theorem)**.** *Let $\boldsymbol{A}$, $\boldsymbol{B}$ be matrices of size $p \times n$ and $n \times p$ respectively, then*

$$|\boldsymbol{I}_p + \boldsymbol{AB}| = |\boldsymbol{I}_n + \boldsymbol{BA}|$$

where $\boldsymbol{I}_a$ is the identity matrix of order a.

So $|\boldsymbol{I}_n + (\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t)(\boldsymbol{\Omega}^{-1}\boldsymbol{P})| = |\boldsymbol{I}_K + (\boldsymbol{\Omega}^{-1}\boldsymbol{P})(\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t)|$ and:

$$\begin{aligned}|\boldsymbol{\Sigma}_\mu||\boldsymbol{\Omega}||\boldsymbol{\Sigma}_\mu^{-1} + \boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\boldsymbol{P}| &= |\boldsymbol{\Omega}\left(\boldsymbol{I}_K + (\boldsymbol{\Omega}^{-1}\boldsymbol{P})(\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t)\right)|\\ &= |\boldsymbol{\Omega} + \boldsymbol{P}\boldsymbol{\Sigma}_\mu\boldsymbol{P}^t|\end{aligned}$$

which is the desired result.

□

A.2 Rearranging of the BL formulae

We make use of the Sherman-Morrison-Woodburry formula (A.1) to reshuffle the BL formulae. We can write (1.20) as:

$$
\begin{aligned}
\boldsymbol{\mu}_{BL} &= \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right)^{-1}\left((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\boldsymbol{q}\right) \\
&= \left((\tau\boldsymbol{\Sigma}) - (\tau\boldsymbol{\Sigma})\boldsymbol{P}^t\left(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}(\tau\boldsymbol{\Sigma})\right) \\
&\qquad \left((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\boldsymbol{q}\right) \\
&= \boldsymbol{\pi} + (\tau\boldsymbol{\Sigma})\boldsymbol{P}^t\left(\boldsymbol{\Omega}^{-1} - \left(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\right)\boldsymbol{q} \\
&\qquad -(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t\left(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}\boldsymbol{\pi}
\end{aligned}
$$

it follows from the Sherman-Morrison-Woodburry formula that:

$$
\left(\boldsymbol{\Omega}^{-1} - \left(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\right) = \left(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}^t + \boldsymbol{\Omega}\right)^{-1}
$$

which inserted in the previous equality yields:

$$
\boldsymbol{\mu}_{BL} = \boldsymbol{\pi} + \boldsymbol{\Sigma}\boldsymbol{P}^t\left(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^t + \frac{\boldsymbol{\Omega}}{\tau}\right)^{-1}(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\pi})
$$

Similarly, applying the Sherman-Morrison-Woodburry formula to (1.23) gives:

$$
\begin{aligned}
\boldsymbol{\Sigma}_{BL}^{\boldsymbol{\mu}} &= \boldsymbol{\Sigma} + \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^t\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right)^{-1} \\
&= (1+\tau)\,\boldsymbol{\Sigma} - \tau^2\boldsymbol{\Sigma}\boldsymbol{P}^t\left(\tau\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^t + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}\boldsymbol{\Sigma}
\end{aligned}
$$

Appendix B

Appendix to Chapter 2

B.1 Worked Example Data

The data for the worked example in section 2.7 consists of 4 years of daily values for each of the six countries' (Australia, Canada, France, Germany, Japan and the United Kingdom) MSCI Barra Standard Core Index. These values correspond to the index values (in USD) from the 27th June 2005 to the 26th June 2009 and was downloaded from MSCI Barra website[1].

The daily returns of each country and their variance/covariance matrix were calculated in order to be used in the example. The correlation matrix and annualized variance follow:

Country	AU	CA	FR	GE	JP	UK
AU	1	0.4654	0.6217	0.5889	0.6037	0.6273
CA		1	0.6673	0.6451	0.2231	0.6780
FR			1	0.9377	0.2772	0.9331
GE				1	0.2570	0.8738
JP					1	0.2704
UK						1
Variance	0.1117	0.0958	0.0866	0.0839	0.0714	0.0835

Table B.1: Data correlation matrix and annualized variance

[1] http://www.mscibarra.com/products/indices/stdindex/performance.html

Appendix C

Appendix to Chapter 3

C.1 Constants of the transforming function

We want to find the constants A, B, C of the function $y(x) = Ae^{Bx} + C$ subject to the conditions:

$$\begin{aligned} y(0) &= 2\% \\ y(3.5\%) &= 3.5\% \quad \text{(continuity condition)} \\ y'(3.5\%) &= 1 \quad \text{(smoothness condition)} \end{aligned}$$

i.e.:

$$A + C = 2\% \tag{C.1}$$

$$Ae^{3.5\%B} + C = 3.5\% \quad \text{continuity condition} \tag{C.2}$$

$$ABe^{3.5\%B} = 1 \quad \text{smoothness condition} \tag{C.3}$$

From C.3 we get:

$$A = \frac{1}{B}e^{-3.5\%B} \tag{C.4}$$

Using C.2 we then retrieve:

$$C = 2 - \frac{1}{B}e^{-3.5\%B} \tag{C.5}$$

And inserting C.4 into C.1 we find the relation:

$$1 - e^{-3.5\%B} = 1.5\%B \tag{C.6}$$

Numerically we find:

$$B \approx 57.9 \tag{C.7}$$

We then derive A and C using this value in C.4 and C.5.

Index