Diss. ETH No. 12397

Maximum Loss for Measurement of Market Risk

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Mathematics

presented by

GEROLD STUDER
Dipl. Math.–Ing. EPFL

born 9th March 1970
citizen of Brislach BL

accepted on the recommendation of
Prof. Dr. H.–J. Lüthi, examiner
Prof. Dr. P. Embrechts, co-examiner

1997

Inhaltsverzeichnis

Abbildungsverzeichnis

Acknowledgements

Many people have contributed to this thesis, from 'pure' academics to risk managers and members of supervisory authorities, and this thesis could only be realized with the financial support and the personal assistance of the members of the 'Forschungs–Interessengemeinschaft Risklab'. It would be impossible to list everyone who has been involved and I must therefore apologize to anyone who has not been mentioned below.
A very special vote of thanks to Prof. Hans–Jakob Lüthi, who has provided me with confirmation that Operations Research is much more than mathematical optimization and who has taught me how to pose questions. Moreover, I would like to thank Hans–Jakob Lüthi for the excellent cooperation: he has always taken time to discuss problems and he has been an unrenounceable guide through the confusions of risk management. I owe my thanks also to Prof. Paul Embrechts, whose many comments have led to significant improvements of this thesis. My colleague Lukas Finschi has been a tremendous support and I acknowledge his great implementation of the Levenberg–Marquardt algorithm.
At Swiss Bank Corporation, Michael Allen deserves a special acknowledgement since he was the person who introduced me with patience to the subject of risk measurement. I am particularly grateful to Marcel Rohner for his marvellous explanations of the principles — and many of the advanced topics — of risk management. Further, I am indebted to Thomas Kindler for the cordial teamwork which has led to the first implementation of Maximum Loss for practical operation. Thanks go to Peter Reiser for his dedicated help in proofreading the manuscript.
At Credit Suisse, my appreciation to Bernhard Hodler for allowing me insights into the daily business of his risk control group. There, I have benefited greatly from the collaboration with Christophe Rouvinez on delta–gamma techniques. I want also to express my gratitude to Michael Heintze and Lukas Gubler at Union Bank of Switzerland for the explanations of their risk management system. Last, but by no means least, I would like to thank Jürg Gabathuler from KPMG for providing me with lots of material and information about performance measurement.

Abstract

Effective risk management requires adequate risk measurement. A basic problem herein is the quantification of market risk: what is the overall effect on a portfolio's value if the market rates change? To answer this question, two fundamental difficulties have to be managed: first, market rates behave randomly and are correlated. Second, portfolio structures are high–dimensional and typically nonlinear.

The established risk measurement techniques can be divided into two categories. The purely stochastic approaches are based on the portfolio's profit and loss (P&L) distribution. The most popular method in this class is Value–at–Risk (VaR), which typically represents the 1 or 5 percent quantile of the P&L distribution.

The Maximum Loss (ML) methodology is a member of the second category, where risk is quantified by the value of some worst case scenario. Many of these worst case based measures examine a finite set of scenarios and do not take account of correlations (e.g. stress testing). The more elaborated methods determine the worst case scenario by solving constrained minimization problems, where the set of feasible scenarios is generally defined by the stochastic characteristics of the market rates.

Compared to other worst case techniques, the Maximum Loss methodology uses a very particular choice of feasible domains: the so–called trust regions cover a certain percentage of all future outcomes and their shape reflects the correlation structure of the market rates. Furthermore, ML uses a polynomial time algorithm to identify the global minimum, whereas other techniques employ rudimentary optimization procedures. The fact that the Maximum Loss computation is based on a fast and reliable algorithm allows to solve the optimization problem repeatedly, which leads to new insights into the risks of nonlinear portfolios.

This thesis is divided into five chapters: the first chapter derives the standard assumptions of risk management from the basics of financial calculus. Some of the established risk measurement techniques are presented in a unifying framework and their basic properties are discussed. The chapter ends with the introduction of the Maximum Loss concept.

Chapter 2 treats the modelling of risk factors in the ML approach. We discuss the transformation of market rates into more abstract risk factors, which are used

to construct a comprehensive and consistent risk model of moderate size.
The mathematical core of Maximum Loss is outlined in Chapter 3. Several optimization problems are formulated and solution strategies are discussed. Moreover, the relations of ML and Value–at–Risk are established and the importance of the worst case scenario is emphasized by sensitivity analyses.
Chapter 4 constitutes the engineering part of the thesis. A repetitive computation of worst case scenarios for growing trust regions leads to the ML path, which can be used for a risk characterization of portfolios. The concept of dynamic approximations is introduced to handle highly nonlinear functions. Finally, the key driving factors of risk are determined by solving restricted ML problems.
Risk measurement is the groundwork for an active management of risks at the firm level. The question of how to contrast return and risk leads to the problem of risk adjusted performance measurement, which is the subject of Chapter 5. Some of the most popular performance measures are presented and their relationships are investigated.

Zusammenfassung

Das Messen von Risiken ist Voraussetzung, um Risiken gezielt steuern zu können. Ein wichtiges Problem spielt dabei die Quantifizierung von Marktrisiken: Welchen Einfluss haben Schwankungen in den Marktsätzen auf den Wert eines Portefeuilles? Die Beantwortung dieser Frage erfordert, sich mit zwei grundlegenden Schwierigkeiten auseinander zu setzen: Zum einen verhalten sich Marktsätze zufällig und sind untereinander korreliert; zum anderen ist die Struktur von Portefeuilles hochdimensional und typischerweise nichtlinear.

Die bekannten Methoden zur Risikomessung können in zwei Gruppen eingeteilt werden. Die rein stochastischen Ansätze basieren auf der Wahrscheinlichkeitsverteilung des Gewinns/Verlusts. Die bekannteste Methode aus dieser Klasse ist der sogenannte Value–at–Risk (VaR), der typischerweise das 1 oder 5 Prozentquantil der Gewinn–/Verlustverteilung repräsentiert.

Der Maximum Loss (ML) gehört zur zweiten Kategorie der Risikomasse, welche Risiko als den Verlust eines Worst Case Szenarios quantifiziert. Viele der Worst Case basierten Methoden betrachten eine endliche Menge von Marktszenarien, ohne die Korrelationen zwischen den Marktsätzen zu berücksichtigen (z.B. Stress Testing). Die weiterentwickelten Verfahren lösen restringierte Minimierungsprobleme, wobei das Zulässigkeitsgebiet meist durch die stochastischen Eigenschaften der Marktsätze bestimmt wird.

Bei der Maximum Loss Methode werden, im Unterschied zu anderen Worst Case Techniken, die Zulässigkeitsgebiete ganz speziell gewählt: die sogenannten Trust Regions enthalten einen wohldefinierten Teil aller möglichen Realisationen und ihre Form widerspiegelt die Korrelationsstruktur der Marktsätze. Während die meisten anderen Techniken rudimentäre Optimierungsmethoden zur Bestimmung des Worst Case Szenarios benutzen, beruht das Optimierungsverfahren von Maximum Loss auf einem zeitpolynomialen Algorithmus. Durch den Einsatz eines schnellen, zuverlässigen Lösungsverfahrens wird es möglich, das Optimierungsproblem wiederholt zu berechnen und dadurch neue Einblicke in die Risiken nichtlinearer Portfolios zu gewinnen.

Diese Arbeit ist in fünf Kapitel gegliedert: Im ersten Kapitel werden, ausgehend von den Grundlagen der Finanzmathematik, die Standardmodelle des Riskmanagements entwickelt. Einige weitverbreitete Methoden zur Risikomessung werden unter einem einheitlichen Rahmen vorgestellt und ihre Eigenschaften eingehend

diskutiert. Am Ende des Kapitels wird das Grundkonzept von Maximum Loss eingeführt.
Im Kapitel 2 wird die Modellierung der Risikofaktoren im Maximum Loss Verfahren behandelt. Es wird gezeigt, wie man aus den Marktsätzen abstrakte Risikofaktoren gewinnt, welche es erlauben, ein umfassendes und konsistentes Modell von beschränktem Umfang zu entwickeln.
Der mathematische Kern von Maximum Loss wird in Kapitel 3 vorgestellt. Unterschiedliche Optimierungsprobleme werden formuliert und Lösungsstrategien werden diskutiert. Ausserdem wird auf den Zusammenhang zwischen Maximum Loss und Value–at–Risk eingegangen, und die Bedeutung des Worst Case Szenarios wird anhand von Sensitivitätsanalysen vertieft.
Kapitel 4 stellt den praktischen Einsatz von Maximum Loss in den Vordergrund. Durch wiederholte Berechnung des Worst Case Szenarios für grösser werdende Trust Regions lässt sich ein ML Pfad bestimmen, der dazu benutzt wird, Portfeuilles im Hinblick auf ihre Risiken und Chancen zu charakterisieren. Um hochgradig nichtlineare Funktionen zu behandeln, wird das Konzept der dynamischen Approximation eingeführt. Abschliessend diskutieren wir, wie sich die risikotreibenden Faktoren durch das Lösen von restringierten ML Problemen identifizieren lassen.
Risikomessung ist Grundlage für das Management von Risiken auf Unternehmensebene. Die Frage, wie Rendite und Risiko einander gegenüber zu stellen sind, führt zum Problem der risikoadjustierten Performancemessung. Diesem Thema wird in Kapitel 5 nachgegangen, wo einige der meistverbreiteten Performancemasse vorgestellt und ihre Gemeinsamkeiten und Zusammenhänge untersucht werden.

Chapter 1

Market Risk Measurement

1.1 Introduction

In a complex and changing world, stability in the sense of only minor volatility in the economy is not achievable: fast, global information leads to mature markets with increased competition. This promotes globalization and compels financial institutions to increase their performance. Thus, it becomes essential to understand the implications of market volatility.

Market risk, sometimes also called *price risk*, denotes the uncertainty engendered by market volatility, i.e. the possibility to lose money due to changes in market prices. Whereas quantitative techniques for the analysis of individual financial instruments are well established, further research on risk measurement for *entire portfolios* is still required.

The necessity to measure risks arose from the regulatory requirements of the supervisory authorities, who have to ensure that sufficient cushions of economic capital are held as a protection against any potential unexpected losses. Until the 1980's, the banking industry was heavily focused toward credit business. Increasing competitiveness resulted in a growing number of banking failures and in 1988, the Basle Committee on Banking Supervision ('Basle Committee'), which was set up under the auspices of the Bank for International Settlements in Basle and which will be referred to simply as the BIS, released the 'International Convergence of Capital Measurement and Capital Standards'. This document, which is also called the 'Basle Accord', established a *common* minimum framework for calculating the capital adequacy of banks with regard to credit risk.[1]

The augmenting importance of stock markets, the increasing volume of derivatives and decreasing profit margins forced banks to extend their trading activities and to take on new risks. The Basle Committee reacted to this development and issued in 1993 proposals for additional market risk requirements, which primarily explicated a standardized technique for measuring market risk to which we will

[1] See BIS (1988), pp. 21.

refer as the 'BIS Standardized Methodology'. This so–called 'building block' approach is a set of simple rules and formulas, which define the capital requirement for market risk as the sum of

- the interest rate risk of each currency
- the commodity risk of each commodity
- the equity risk of each market
- the foreign exchange risk of all currencies.

The proposal was heavily criticized from the banking industry as being too crude because it neither differentiates between the riskiness of different assets, nor takes into account the correlations between assets — which leads to an overestimation of the total risk. Nonetheless, this method became an element of the final 1996 'Amendment to the Capital Accord to Incorporate Market Risk'[2].

Increased competition and pressure from rating agencies and regulators produced a need for more sophisticated risk measurement techniques, leading to the lowering of the capital requirements. Also, the demand for bank internal risk measurement intensified: the availability of a common yardstick, which allows to compare the risk of different portfolios and business units, is an indispensable basis for an operative management of risks. Furthermore, the ability to link expected earnings and risk is a prerequisite for a risk adjusted performance measurement and, consequently, a conditio sine qua non for an effective capital allocation and the strategic management of the firm. The dialog between the Basle Committee and the financial industry resulted in the fact that the *Value–at–Risk*[3] (VaR) approach became an accepted risk measure in the 1996 amendment.[4]

Since the beginning of the 1990s, the trading volume of derivative instruments has drastically augmented and given rise to highly nonlinear portfolios[5]. The consequence is that simple answers to the question 'what is the overall risk of a portfolio?' exist less than ever before. This is reflected in the 1996 amendment, where a *combined* use of several risk measurement techniques is prescribed.[6]

The Maximum Loss methodology which is described in this thesis has been developed to provide new insights into the risks of nonlinear portfolios. It can be understood as a systematic way to identify the 'black holes' of a portfolio under normal market conditions, without ignoring the correlations between the market rates. In its basic form, Maximum Loss enables to actively manage risks

[2]See BIS (1996a) and BIS (1996a).

[3]See Section 1.3.1.

[4]See BIS (1995).

[5]A nonlinear portfolio is a portfolio whose value does not linearly depend on the market rates.

[6]The Basle Committee specifies that VaR models have to be accompanied by stress tests, cf. BIS (1996b), pp. 46.

since it clarifies the sources of risk, provides the ability to find the key drivers of a portfolio's profit and loss, helps to determine risk reducing transactions, and provides estimates of the incremental risk of single trades. More sophisticated applications of the Maximum Loss concept examine the dynamic development of risks and therefore enable advanced risk characterizations of portfolios.

1.2 Definitions and Basic Assumptions

This section introduces the terminology used in risk measurement. In particular, the fundamental notions of *risk factors* and *profit and loss functions* are explained in Section 1.2.1, and the hypothesis of normally distributed risk factors is justified in Section 1.2.2.

1.2.1 Notations and Definitions

When we measure the market risk of a portfolio, we are interested in the *joint* effect of *changes* of the market rates (e.g. commodity prices, foreign exchange rates, equity indices, interest rates, implied volatilities) on the portfolio value. In a mathematical model, these market rates can be represented by stochastic processes $(X_t^i)_{t\geq 0}, i = 1, \dots, M$, which are defined on the probability space $(\Omega, \mathcal{A}, \mathrm{P})$.[7]
We will restrict the analysis to the time interval $[0, T]$, where $t = 0$ is the current time and T stands for the time horizon which is considered relevant for the risk analysis. T is also called *holding period* and it represents the maximum time required to liquidate or hedge the portfolio. In other words, it is assumed that the whole portfolio is exposed to changes in market rates during a period of length T.
The term 'risk' can be understood as *the possibility that the real development deviates negatively from the planned development*;[8] positive deviations are usually called 'chances'. Consequently, we will compare the values X_T^i of the stochastic processes $i = 1, \dots, M$ with the expected outcomes at the end of the holding period:

Definition 1.1 *$\psi_i = X_T^i$ is the random variable which represents the outcome of market rate $i = 1, \dots, M$ at time T. The expected value at time T is denoted by $\bar{\psi}_i = \mathrm{E}(\psi_i)$. The difference $\omega_i = \psi_i - \bar{\psi}_i$ is called risk factor.*[9]

Thus, the risk factors $\omega = (\omega_1, \dots, \omega_M)$ describe the deviations of the market rates from their expected value at time T. It follows from Definition 1.1 that

[7] See Appendix A for an exhibition of the fundamentals of stochastic calculus.
[8] See Moser and Quast (1994), p. 665.
[9] The term 'risk factor' means 'source of risk' and is not related to statistical multiple factor models; cf. Beckers (1996), pp. 175.

$\omega : \Omega \to \mathbb{R}^M$ is a random vector with mean 0 (and not, as in usual notation, $\omega \in \Omega$). The distribution of ω corresponds to the centered marginal distribution of the stochastic processes $(X_t^i)_{t \geq 0}, i = 1, \ldots, M$, at time T (the end of the holding period).
Although $\omega : \Omega \to \mathbb{R}^M$ is by Definition 1.1 a random vector, the symbol ω is also used — by abuse of notation — for realizations of risk factors.[10] A particular realization is also called a *scenario* and the *set of realizable scenarios* is defined as $\mathcal{O} = \{\omega \in \mathbb{R}^M \mid \exists X_T : \omega = X_T - \mathrm{E}(X_T)\}$. For subsets $\mathcal{U} \subseteq \mathbb{R}^M$, we denote by $\mathrm{P}(\mathcal{U}) := \mathrm{P}(\{\omega \mid \omega \in \mathcal{U}\})$ the probability which is induced by the probability space $(\Omega, \mathcal{A}, \mathrm{P})$. For technical reasons, we require that:

Assumption 1.2 *The set of realizable scenarios $\mathcal{O}$ is a connected subset of $\mathbb{R}^M$ with $\mathrm{P}(\mathcal{O}) = 1$ and it contains the scenario $\omega = 0$.*

Moreover, we assume that the values of all market rates at time $t = 0$ are known; they are denoted by $\psi_i^0 = X_0^i, i = 1, \ldots, M$.
In practice, pricing models are used for assessing the effect of future market rates $\psi = (\psi_1, \ldots, \psi_M)$ on the portfolio value. We postulate to have a function $u(\psi)$ at disposal, which gives the *value of the portfolio* at time T for every possible outcome of ψ. Since the risk factors ω represent the deviations of the market rates ψ from their expected values $\bar{\psi}$ (i.e. $\psi = \bar{\psi} + \omega$), we define a profit and loss function $v(\omega)$, which measures the change in portfolio value caused by the deviations in the market rates:

Definition 1.3 *Let $u(\psi)$ be the portfolio value function. The profit and loss (P&L) function $v(\omega)$ is defined as*

$$v(\omega) = u(\bar{\psi} + \omega) - u(\bar{\psi}).$$

Hence, the P&L function $v(\omega)$ can be seen as a deterministic function of the random vector ω; positive values of $v(\omega)$ represent profits and negative values represent losses.

Remark 1.4 *Definitions 1.1 and 1.3 imply that*

$$v(0) = u(\bar{\psi} + 0) - u(\bar{\psi}) = 0, \tag{1.1}$$

which states that there are neither unexpected profits nor losses if the market rates behave as expected.

[10] The meaning of the symbol ω will be clear from the context: when we discuss stochastic risk measures (e.g. Value–at–Risk, Section 1.3.1), ω has to be interpreted as a random vector. On the other hand, in worst case based risk measures (e.g. Maximum Loss, Section 1.5 and Chapters 2–4), ω stands mostly for scenarios.

Remark 1.4 motivates the definition of meaningful P&L functions:

Definition 1.5 *The set of admissible P&L functions is* $\mathcal{V} = \{v : \mathcal{O} \to \mathbb{R} \mid v(0) = 0\}$.

To avoid technical problems, we assume that

Assumption 1.6 *Every P&L function* $v : \mathcal{O} \to \mathbb{R}$ *is a continuous function which is zero on all open subsets* $\mathcal{U} \subset \mathcal{O}$ *with measure* $\mathrm{P}(\mathcal{U}) = 0$.[11]

In the sequel we will mainly work with linear and quadratic portfolio structures. Thus, we distinguish:

Definition 1.7 *The subset of linear portfolios is*

$$\mathcal{V}_l = \{v_l \in \mathcal{V} \mid v_l(\omega) = a^T\omega, a \in \mathbb{R}^M\},$$

and the subset of quadratic portfolios is

$$\mathcal{V}_q = \{v_q \in \mathcal{V} \mid v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega, G = G^T \in \mathbb{R}^{M\times M}, g \in R^M\}.$$

Obviously, linear and quadratic P&L functions can always be obtained from the local price sensitivities δ and Γ:

Definition 1.8 $\delta \in \mathbb{R}^M$ *is the first order price sensitivity defined by* $\delta_i = \frac{\partial u(\psi)}{\partial \psi_i} \mid_{\bar{\psi}}$ *for* $i = 1, \ldots, M$, *and* $\Gamma \in \mathbb{R}^{M\times M}$ *is the symmetric matrix of second order sensitivities* $(\Gamma)_{i,j} = \frac{\partial^2 u(\psi)}{\partial \psi_i \partial \psi_j} \mid_{\bar{\psi}}$.[12]

Finally, a *risk measure* ρ is a function which assigns to a portfolio $v(\omega)$ a real number $\rho(v)$ to be interpreted as its risk; the possibility to lose money is expressed by negative values of ρ. We allow risk measures to be defined only for specific subclasses of portfolios, e.g. linear or quadratic portfolios. Thus, $\rho : \mathcal{V}_\rho \to \mathbb{R}$, where $\mathcal{V}_\rho$ is a subset of $\mathcal{V}$.

1.2.2 Risk Factor Distribution

For risk measurement purposes, it is usually claimed that the risk factors are normally distributed. This section presents the mathematical backgrounds of this choice.

[11] The second condition is trivially satisfied if the risk factors are multinormally distributed, cf. Assumption 1.12.

[12] For many financial instruments, the local sensitivities can directly be obtained from valuation models. The formulas for the Black–Scholes model, for example, can be found in Cox and Rubinstein (1985), pp. 210.

Let $(X_t)_{t\geq 0}$ represent the process of a market rate, e.g. a stock price. According to the standard assumption of finance,[13] $(X_t)_{t\geq 0}$ typically follows a geometric Brownian motion:

$$dX_t = \mu_X X_t dt + \sigma_X X_t dB_t, \tag{1.2}$$

where μ_X is the drift rate, σ_X the volatility and $(B_t)_{t\geq 0}$ a standard Brownian motion.[14] If we set $X_0 = x_0$, this expression can be rewritten as (and indeed is defined via) $X_t = x_0 + \int_0^t \mu_X X_s ds + \int_0^t \sigma_X X_s dB_s$. Then, the *formal*[15] application of Itô's lemma[16] to $g(X_t) = \log(X_t)$ leads to

$$\log(X_t) = \log(X_0) + \int_0^t \frac{1}{X_s} dX_s - \frac{1}{2}\int_0^t \frac{\sigma_X^2 X_s^2}{X_s^2} ds.$$

Using equation (1.2), it follows that

$$X_t = x_0 \exp\left(\left(\mu_X - \frac{\sigma_X^2}{2}\right)t + \sigma_X B_t\right). \tag{1.3}$$

It is easy to verify that (1.3) is a solution to the problem (1.2), and it can be proved indeed that this solution is unique.[17] Hence, the logreturns $\log(\frac{X_t}{X_0})$ are normally distributed with mean $(\mu_X - \frac{\sigma_X^2}{2})t$ and variance $\sigma_X^2 t$. It should be noted that many empirical studies[18] have shown that the hypothesis of normally distributed logreturns is only a rough approximation of reality.[19] Nonetheless, this assumption is widely used in practice, mainly because of the analytical tractability of the normal distribution.

In our usual notation, where ψ_i denotes the value of the stochastic process at time T (the end of the holding period) and ψ_i^0 is the value at time 0, equation (1.3) can be restated as $\psi_i = \psi_i^0 \exp((\mu_i - \frac{\sigma_i^2}{2})T + \sigma_i B_T^i)$, where μ_i is the drift, σ_i the volatility and $(B_t^i)_{t\geq 0}$ the standard Brownian motion for market rate $i, i = 1, \ldots, M$. It follows that

$$\log\left(\frac{\psi_i}{\psi_i^0}\right) = \left(\mu_i - \frac{\sigma_i^2}{2}\right)T + \sigma_i B_T^i. \tag{1.4}$$

[13] See Hull (1993), pp. 197.
[14] See Appendix A.
[15] In fact, the function $\log(X_t)$ is not in C^2.
[16] See Appendix A.
[17] See Lamberton and Lapeyre (1993), pp. 56.
[18] See Taylor (1986), p. 44.
[19] Indeed, analyses of time series show that the distributions of logreturns are slightly skewed (i.e. asymmetric) and leptokurtic (i.e. fat tailed).

A first order Taylor expansion around ψ_i^0 results in

$$\begin{aligned}\log\left(\frac{\psi_i}{\psi_i^0}\right) &= \log(\psi_i^0 + (\psi_i - \psi_i^0)) - \log(\psi_i^0) \\ &\approx \left[\log(\psi_i^0) + \frac{\psi_i - \psi_i^0}{\psi_i^0}\right] - \log(\psi_i^0) \\ &\approx \frac{\psi_i - \psi_i^0}{\psi_i^0},\end{aligned}$$

which implies together with (1.4) that $\frac{\psi_i - \psi_i^0}{\psi_i^0} \approx \left(\mu_i - \frac{\sigma_i^2}{2}\right)T + \sigma_i B_T^i$. If we take this approximation as an exact equality, we get normally distributed risk factors:

Theorem 1.9 *If* $\frac{\psi_i - \psi_i^0}{\psi_i^0} = \left(\mu_i - \frac{\sigma_i^2}{2}\right)T + \sigma_i B_T^i$, *then* $\omega_i \sim \mathcal{N}(0, (\psi_i^0\sigma_i)^2 T)$.

Proof: Since $(B_t^i)_{t\geq 0}$ is a standard Brownian motion it follows that $B_T^i \sim \mathcal{N}(0,T)$, which implies

$$\frac{\psi_i - \psi_i^0}{\psi_i^0} \sim \mathcal{N}\left((\mu_i - \frac{\sigma_i^2}{2})T, \sigma_i^2 T\right). \tag{1.5}$$

Thus, we get $\psi_i \sim \mathcal{N}\left(\psi_i^0[1 + (\mu_i - \frac{\sigma_i^2}{2})T], (\psi_i^0\sigma_i)^2 T\right)$ because ψ_i^0 is a constant. Then, the definition of the risk factors $\omega_i = \psi_i - \bar{\psi}_i$, where $\bar{\psi}_i = \mathrm{E}(\psi_i)$, proves the theorem. ■

Sometimes, the behavior of interest rates is modelled directly by Brownian motion with drift.[20] In this case, it is easy to show that $\omega_i \sim \mathcal{N}(0, \sigma_i^2 T)$; in the present context, this would be equivalent to setting $\psi_i^0 \equiv 1$.
Finally, it remains to introduce the correlations between the risk factors:

Lemma 1.10 *If* $\mathrm{Corr}(\frac{\psi_i - \psi_i^0}{\psi_i^0}, \frac{\psi_j - \psi_j^0}{\psi_j^0}) = \rho_{i,j}$, *then* $\mathrm{Corr}(\omega_i, \omega_j) = \rho_{i,j}$.

Proof: The basic properties of the correlation assert that

$$\mathrm{Corr}\left(\frac{\psi_i - \psi_i^0}{\psi_i^0}, \frac{\psi_j - \psi_j^0}{\psi_j^0}\right) = \mathrm{Corr}(\psi_i - \bar{\psi}_i, \psi_j - \bar{\psi}_j) = \mathrm{Corr}(\omega_i, \omega_j).$$

■

[20] This means, that the process $(X_t)_{t\geq 0}$ satisfies the differential equation $dX_t = \mu_X dt + \sigma_X dB_t$; see Chen (1996), p. 3.

Corollary 1.11 *If* $\mathrm{Var}(\frac{\psi_i-\psi_i^0}{\psi_i^0}) = \sigma_i^2 T$ *for* $i = 1, \ldots, M$ *and* $\mathrm{Corr}(\frac{\psi_i-\psi_i^0}{\psi_i^0}, \frac{\psi_j-\psi_j^0}{\psi_j^0}) = \rho_{i,j}$*, then* $\omega \sim \mathcal{N}(0, \Sigma)$*, where the elements of the covariance matrix are given by* $(\Sigma)_{i,j} = \rho_{i,j}\psi_i^0\psi_j^0\sigma_i\sigma_j T$*.*

Proof: This is a direct consequence of Theorem 1.9 and Lemma 1.10. ∎

Hence, the model of normally distributed returns (1.5) justifies the following assumption:

Assumption 1.12 *The risk factors* ω *are multinormally distributed with mean 0 and positive definite covariance matrix* Σ*.*

It should be noted that this assumption does not imply that the return *processes* $(\frac{X_{t+T}-X_t}{X_t})_{t\geq 0}$ are Gaussian; it only states that the conditional distribution of $\frac{X_T-X_0}{X_0}$ at time $t = 0$ is normal.[21]

1.3 Established Risk Measurement Techniques

This section presents some popular risk measurement techniques under the unifying framework established in Section 1.2.1.

Remark 1.13 *The factor push methodology*[22]*, stress testing*[23]*, as well as the scanning charge calculation*[24] *of the SPAN*[25] *framework are special cases of the more general Maximum Loss concept, which will be introduced in Section 1.5.*

1.3.1 Value–at–Risk

The currently most popular methodology for measuring market risk is Value–at–Risk (VaR).[26] It is defined as

> the level of P&L, which is exceeded
> - with probability α,
> - over some time interval T.

Turned the other way round, VaR reflects below which level of P&L a fraction of $(1-\alpha)$ of all outcomes will be. It is clear that this risk measure heavily depends on

[21] Hence, the covariance matrix Σ (for a fixed holding period T) may depend on time (i.e. today's Σ can be different from tomorrow's) and might be estimated from various forecast methodologies, e.g. moving averages or GARCH models (cf. Alexander (1996), pp. 234).

[22] See Wilson (1996), pp. 224.

[23] See Dembo et al. (1995).

[24] See SPAN (1997).

[25] Standard Portfolio Analysis of Risk; cf. SPAN (1993).

[26] See Beckström and Campbell (1995).

the choice of parameters, which is a controversial subject in practice. The length of the holding period T should reflect the time needed to liquidate or hedge the portfolio and may be influenced by the type of products and the markets on which they are traded, the size of the owned positions relative to the issue size, the credit rating of the issuer, the turnover of the instruments, etc. The confidence level α is usually chosen between 95% and 99%, the latter being the choice of the Basle Committee[27]. Furthermore, it should not be neglected that every risk assessment comprises predictive elements and therefore heavily depends on our beliefs and models (e.g. the forecast of variances and correlations)[28]. Since ω is a random vector, we define:

Definition 1.14 *For a continuous, strictly increasing P&L distribution function, Value–at–Risk for a confidence level α is that level of loss* $\mathrm{VaR}(\alpha)$ *for which* $\mathrm{P}(\{\omega \in \mathcal{O} \mid v(\omega) \leq \mathrm{VaR}(\alpha)\}) = 1 - \alpha$.[29]

If we look at $v(\omega)$ as a random variable (and not as a deterministic function of a random vector), then VaR corresponds to the $(1-\alpha)$–*quantile* of the P&L distribution. This interpretation also agrees with the way in which VaR is calculated: First, the *distribution of the P&L* is constructed from the (deterministic) P&L function $v(\omega)$ and the distribution of the risk factors. Then, the $(1-\alpha)$–quantile of this one–dimensional distribution gives the Value–at–Risk (cf. Figure 1.1).

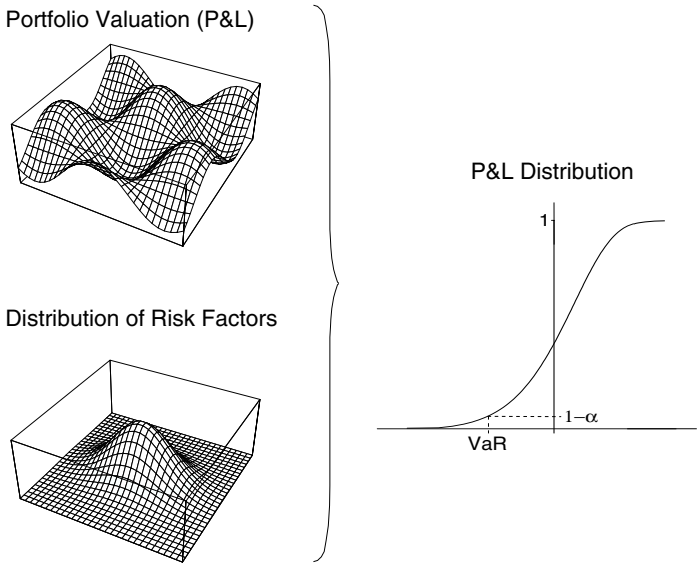

Figure 1.1: Interpretation of VaR

[27] See BIS (1996a), p. 2.

[28] See Alexander (1996), pp. 233.

[29] For general P&L distribution functions, VaR might be defined as $\mathrm{VaR}(\alpha) = \inf\{y \in \mathbb{R} \mid \mathrm{P}(\{\omega \in \mathcal{O} \mid v(\omega) \leq y\}) \geq 1 - \alpha\}$; this generalized definition keeps all results about VaR valid.

VaR of Linear Portfolios

The Value–at–Risk of linear portfolios can be calculated analytically:

Lemma 1.15 *Let $v_l(\omega) = a^T\omega$ be a linear portfolio with normally distributed risk factors $\omega \sim \mathcal{N}(0, \Sigma)$. Then $v_l(\omega) \sim \mathcal{N}(0, a^T\Sigma a)$.*

Proof: First, note that any linear combination of normally distributed variables is again normally distributed. The result follows immediately since $\mathrm{E}(a^T\omega) = a^T\mathrm{E}(\omega)$ and $\mathrm{Cov}(a^T\omega) = a^T\mathrm{Cov}(\omega)a$. ■

Theorem 1.16 *The Value–at–Risk of a linear portfolio $v_l(\omega) = a^T\omega$ is equal to $\mathrm{VaR}(\alpha) = -z_\alpha\sqrt{a^T\Sigma a}$, where z_α is the α–quantile of the standard normal distribution.*

Proof: Let $Y \sim \mathcal{N}(0, \sigma^2)$, then $\frac{Y}{\sigma} \sim \mathcal{N}(0, 1)$ and $(-z_\alpha\sigma)$ is the $(1-\alpha)$–quantile of Y. ■

Example 1.17 *The portfolio of Figure 1.2 will be used in the sequel to illustrate various aspects of risk measurement methodologies. It is an equity portfolio with seven risk factors $\omega_1, \ldots, \omega_7$ which represent each an equity index. The dots display the P&L if each of the risk factors moves up or down 1, respectively 2 standard deviations separately. The solid lines show the quadratic functions $v_q^{(i)}(\omega_i) = \frac{1}{2}G_{i,i}\omega_i^2 + g_i\omega_i, i = 1, \ldots, 7$, which best fit to these points (in a least squares sense). Apparently, the P&L function of this portfolio can reasonably be approximated by a quadratic function $v_q(\omega)$ (which is, however, different from the δ–Γ approximation of Definition 1.8, which only considers local information). Figure 1.3 displays the P&L distribution of the quadratic approximation $v_q(\omega) = \sum_{i=1}^7 v_q^{(i)}(\omega_i)$ as well as for the linear approximation $v_l(\omega) = \sum_{i=1}^7 g_i\omega_i$.*[30] *The fact that the portfolio is almost perfectly delta hedged results in a Delta–Normal VaR of almost 0, whereas the $(1-\alpha)$–quantile of the quadratic model is nearly -1000.*

This example shows that nonlinearity can have tremendous effects on a portfolio's risk. Hence, it is essential to dispose of risk measurement techniques which consider nonlinear portfolio structures.

VaR of Quadratic Portfolios

For nonlinear portfolios, there exists no analytic expression of VaR. However, it is possible to calculate the Value–at–Risk of a quadratic portfolio $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$ numerically.

[30] The separability of the global P&L functions $v(\omega) = \sum_{i=1}^7 v^{(i)}(\omega_i)$ is justified because there are no instruments in the portfolio which depend one more than one risk factor.

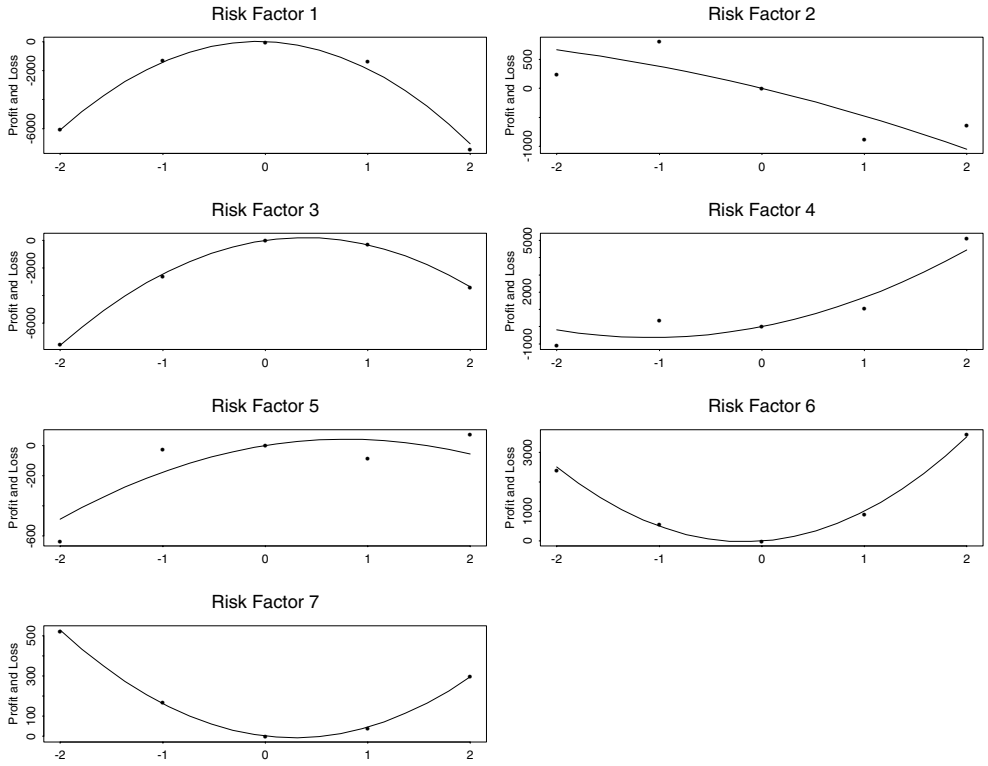

Figure 1.2: Risk profiles of test portfolio

Definition 1.18 *Let $\Sigma = U^T U$ be the Cholesky decomposition of the covariance matrix and let O be the orthonormal matrix of eigenvectors of UGU^T.*[31] *The transformed system is defined as $\tilde{v}_q(\tilde{\omega}) = \frac{1}{2}\tilde{\omega}^T\tilde{G}\tilde{\omega} + \tilde{g}^T\tilde{\omega}$, where $\tilde{g} = OUg$, $\tilde{G} = OUGU^TO^T$ and $\tilde{\omega} = OU^{-T}\omega$.*

Lemma 1.19 *The matrix $\tilde{G}$ is diagonal.*

Proof: Since O is the matrix of the eigenvectors of UGU^T, it satisfies $UGU^To_i = \lambda_i o_i$, where o_i is the ith eigenvector of UGU^T (i.e. the ith vector of O) and λ_i the corresponding eigenvalue. The fact that O is orthonormal implies that $\tilde{G}$ is the diagonal matrix made up of the eigenvalues of UGU^T. ■

Lemma 1.20 *The transformed risk factors $\tilde{\omega}_i, i = 1, \ldots, M$, are independent standard normally distributed variables.*

Proof: $\tilde{\omega}$ is a linear combination of normally distributed variables with $\mathrm{E}(\tilde{\omega}) = 0$ and $\mathrm{Cov}(\tilde{\omega}) = \mathrm{E}(\tilde{\omega}\tilde{\omega}^T) = OU^{-T}\Sigma U^{-1}O^T = \mathrm{I}$. ■

[31] There exist M orthonormal eigenvectors of UGU^T since UGU^T is symmetric.

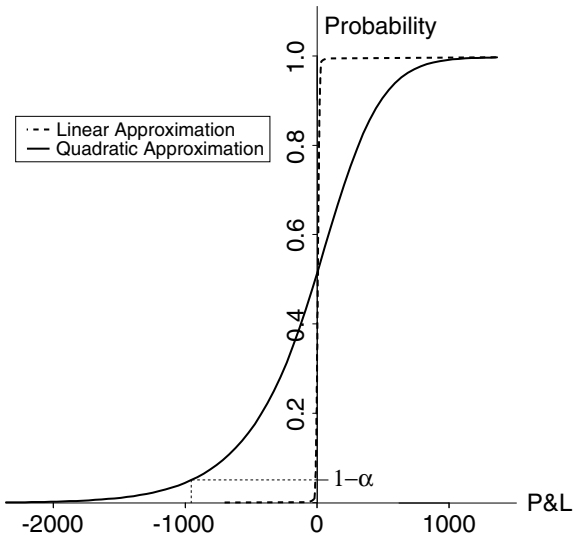

Figure 1.3: P&L distribution of test portfolio

Lemma 1.21 *The functions $v_q(\omega)$ and $\tilde{v}_q(\tilde{\omega})$ are identical: $v_q(\omega) = \tilde{v}_q(\tilde{\omega})$ for $\tilde{\omega} = OU^{-T}\omega$.*

Proof: This is an immediate consequence of $\tilde{g}^T\tilde{\omega} = g^T U^T O^T O U^{-T}\omega = g^T\omega$ and $\tilde{\omega}^T\tilde{G}\tilde{\omega} = \omega^T U^{-1}O^T(OUGU^TO^T)OU^{-T}\omega = \omega^T G\omega$. ∎

Thus, the original risk factors ω have become independent $\mathcal{N}(0,1)$ variables and the new quadratic P&L function $\tilde{v}_q(\tilde{\omega})$ has no mixed terms. Hence, the P&L function $\tilde{v}_q(\tilde{\omega})$ can be separated into a sum of purely quadratic and a sum of purely linear terms:[32]

Definition 1.22 *Let $I = \{i \in \{1,\dots,M\} \mid (\tilde{G})_{i,i} \neq 0\}$, $J = \{i \in \{1,\dots,M\} \mid (\tilde{G})_{i,i} = 0\}$. Define $\breve{\omega}_i = \tilde{\omega}_i + \frac{\tilde{g}_i}{(\tilde{G})_{i,i}}, i \in I$, and $\breve{\omega}_i = \tilde{\omega}_i, i \in J$. The separated P&L function $\breve{v}_q(\breve{\omega})$ is defined as*

$$\breve{v}_q(\breve{\omega}) = \sum_{i\in I}\frac{1}{2}(\tilde{G})_{i,i}\breve{\omega}_i^2 + \sum_{i\in J}\tilde{g}_i\breve{\omega}_i - \sum_{i\in I}\frac{\tilde{g}_i^2}{2(\tilde{G})_{i,i}}.$$

Theorem 1.23 *The random variables $\breve{\omega}_i, i = 1,\dots,M$, are independent and the functions $\breve{v}_q(\breve{\omega})$ and $\tilde{v}_q(\tilde{\omega})$ are identical.*

Proof: The mutual independence of $\breve{\omega}_i, i = 1,\dots,M$, is due to the fact that each $\breve{\omega}_i$ corresponds to a translated variable $\tilde{\omega}_i$, which are themselves mutually independent.

[32] See Schaefer (1995).

On the other hand, $\frac{1}{2}(\tilde{G})_{i,i}\check{\omega}_i^2 - \frac{\tilde{g}_i^2}{2(\tilde{G})_{i,i}} = \frac{1}{2}(\tilde{G})_{i,i}\tilde{\omega}_i^2 + \tilde{g}_i\tilde{\omega}_i$ for $i \in I$, and $\tilde{g}_i\check{\omega}_i = \tilde{g}_i\tilde{\omega}_i$ for $i \in J$, which proves that $\check{v}_q(\check{\omega})$ and $\tilde{v}_q(\tilde{\omega})$ are identical. ∎

As a result, $\check{v}_q(\check{\omega})$ is a linear combination of M independent random variables: for $i \in J$, the separated risk factors $\check{\omega}_i$ are standard normally distributed, whereas for $i \in I$, the quantities $\check{\omega}_i^2$ are noncentral χ^2 variables with one degree of freedom and noncentrality parameter $(\frac{\tilde{g}_i}{(\tilde{G})_{i,i}})^2$.
To calculate VaR, Rouvinez (1997) uses the fact that the moment generating function of a linear combination of independent variables is the product of the moment generating functions of the individual variables. Then, the inversion theorem[33] allows to calculate numerically the probability of a given fractile z, i.e. $\mathrm{P}(\check{v}_q(\check{\omega}) \leq z)$; a detailed exhibition is given in Appendix B. VaR, which is the $(1-\alpha)$–quantile of the P&L distribution, is then determined by using a bisectionning technique, which converges to that value of z for which $\mathrm{P}(\check{v}_q(\check{\omega}) \leq z) = 1 - \alpha$.

VaR by Monte Carlo Simulation

Monte Carlo simulation is used to determine the P&L distribution of more general, not quadratic, P&L functions, where (in practice) it is still assumed that the risk factors are normally distributed.[34] The basic idea is to draw a random set of representative scenarios, to value the portfolio for each scenario, and to take the distribution of the resulting P&Ls as a representative of the true P&L distribution. More concretely, the algorithm of the Monte Carlo simulation looks as follows:

1. Draw a set of n independent scenarios $\mathcal{S} = \{\omega^{(1)}, \ldots, \omega^{(n)}\}$, where each $\omega^{(i)} \sim \mathcal{N}(0, \Sigma)$.[35]
2. Value the portfolio for each scenario: $\mathcal{P} = \{v(\omega^{(1)}), \ldots, v(\omega^{(n)})\}$.
3. Sort $\mathcal{P}$ in increasing order: $v^{(1)} \leq \ldots \leq v^{(n)}$.
4. VaR is estimated by some interpolation of $v^{(\lfloor(1-\alpha)n\rfloor)}$ and $v^{(\lceil(1-\alpha)n\rceil)}$.

Pritsker (1996) discusses computational aspects of Monte Carlo simulation as well as the construction of confidence intervals for such estimates.[36]

[33] See Davies (1973), pp. 415.

[34] See Jorion (1997), pp. 241.

[35] To generate a random variable $\omega \sim \mathcal{N}(0, \Sigma)$, first draw M independent components $y_i \sim \mathcal{N}(0,1), i = 1, \ldots, M$. Let $\Sigma = U^T U$ be the Cholesky decomposition of the covariance matrix, then $\omega = U^T y$ has the desired distribution since $\mathrm{Cov}(\omega) = U^T \mathrm{Cov}(y) U = \Sigma$.

[36] The confidence bounds are based on the order statistics of n iid random variables and are constructed from a *single* set $\mathcal{P} = \{v(\omega^{(1)}), \ldots, v(\omega^{(n)})\}$ of random draws; multiple runs of the simulation procedure are not required.

VaR by Historical Simulation

Historical simulation is commonly used to avoid making assumptions about the distribution of risk factors. It differs from the previously discussed Monte Carlo simulation in step 1, where the scenario set $\mathcal{S}$ is constructed: instead of using scenarios that come from a random distribution, $\mathcal{S}$ is a set of historical scenarios, e.g. the changes in the market rates which have been observed during the last n holding periods.

1.3.2 BIS Standardized Methodology

The Basle Committee on Banking Supervision of the Bank for International Settlements (BIS) proposed in 1993 a standardized measurement methodology for establishing the overall market risk of financial institutions.[37]
The aim of this section is to present the so–called 'building–block' approach in the framework which has been established in Section 1.2.1. Readers who are not familiar to this methodology will get an impression of the complexity of the calculation procedures and understand the shortcomings of this technique. The mathematical formulation, which is elaborated in the following, will be used in Section 1.4.2 where it is shown that the BIS Standardized Methodology is weakly coherent.
The BIS approach distinguishes between general market risk and specific risk: general market risk is the risk caused by changes in rates related to a whole market, whereas specific risks are due to price movements which are purely related to the individual issue/issuer. More precisely, the methodology uses individual risk factors for:[38]

- the interest rate risk of each currency
- the equity risk of each market
- the foreign exchange risk of *all exchange rates together*
- the commodity risk of each commodity
- the gamma risk of all options with identical underlying
- the vega risk[39] of all options with identical underlying
- the specific risk of each interest rate related issue
- the specific risk of each individual equity.

[37] See BIS (1996a), p. 7.
[38] See BIS (1996b), pp. 9.
[39] Vega risk is the risk due to changes in the volatility of the underlying; it is also called *volatility risk*, cf. Remark 2.1.

Although this model leads to a considerable number M of risk factors, the simple structure of the function $v(\omega) = \sum_{i=1}^{M} v_i(\omega_i)$ compensates for this drawback. The separated functions $v_i(\omega_i)$ are defined by the market values of the instruments:

Definition 1.24 *Let J be the set of all instruments of the portfolio. Denote by m_j the market value*[40] *of instrument $j, j \in J$. Define $H_i = \{j \in J \mid$ value of instrument j depends on risk factor $\omega_i\}$. The net long position in risk factor ω_i is $p_i^{(l)} = \sum_{j \in H_i} m_j^+$, the net short position is $p_i^{(s)} = \sum_{j \in H_i} |m_j^-|$, where $m_j^+ = \max\{0, m_j\}$ and $m_j^- = \min\{0, m_j\}$. Then,*

- $p_i^{(N)} = |p_i^{(l)} - p_i^{(s)}|$ *is the net position of risk factor ω_i*
- $p_i^{(G)} = |p_i^{(l)} + p_i^{(s)}|$ *is the gross position of risk factor ω_i*
- $p_i^{(H)} = \max\{p_i^{(l)}, p_i^{(s)}\}$ *is the higher net position of risk factor ω_i*
- $p_i^{(M)} = \min\{p_i^{(l)}, p_i^{(s)}\}$ *is the matched position of risk factor ω_i*
- $p_i^{(\gamma)} = |\min\{0, \frac{1}{2}\sum_{j \in H_i} \Gamma_j (\theta_i m_j)^2\}|$ *is the gamma–impact of all options on the underlying i. The symbol Γ_j denotes the second order price sensitivity*[41] *of instrument j with respect to the underlying i, and θ_i is a weighting coefficient for the underlying i.*
- $p_i^{(\upsilon)} = \frac{1}{4}\sigma_i |\sum_{j \in H_i} \upsilon_j|$ *is the vega–impact of all options with underlying i; σ_i is the implied volatility of the underlying.*[42]

In our usual notation, the individual P&L functions become

$$v_i(\omega_i) = (p_i^{(l)} - p_i^{(s)})\omega_i. \tag{1.6}$$

For each such function, a risk figure $\rho_i^{(\cdot)}(v_i)$ is then defined on the basis of the positions $p_i^{(\cdot)}$:

Definition 1.25 *Depending on the risk category to which risk factor ω_i belongs, the individual risk component $\rho_i(v_i)$ is defined as:*

- $\rho_i^{(1)}(v_i) = -(p_i^{(N)} + \kappa_i p_i^{(M)}), 0 \leq \kappa_i \leq 1$ *for interest rate risk, commodity risk and general equity risk*
- $\rho_i^{(2)}(v_i) = -p_i^{(G)}$ *for specific equity risk*
- $\rho_i^{(3)}(v_i) = -p_i^{(H)}$ *for foreign exchange risk*[43]

[40] For options, the market value of the underlying is δ–weighted, i.e. multiplied by the δ of the instrument.

[41] See Definition 1.8.

[42] υ_j is the price sensitivity of instrument j with respect to changes in the volatility σ_i. υ_j has positive value for long positions and negative value for short positions.

[43] Note that there is only one risk factor which accounts for all exchange rates together.

- $\rho_i^{(4)}(v_i) = -p_i^{(\gamma)}$ *for gamma risk.*
- $\rho_i^{(5)}(v_i) = -p_i^{(\upsilon)}$ *for vega risk.*

The components of type $\rho_i^{(1)}(v_i)$ are utilized for risk factors which have a term structure. They are determined by the following procedure:

1. All instruments are slotted into a maturity ladder[44] and their market value is multiplied with a time–band dependent weight, which reflects the price sensitivity with respect to a one basis point move.
2. The long and short positions *within* each time–band are matched and a fraction of these matched amounts reflect the *basis* risk.
3. The remaining net long/short positions of each time–band are matched against each other (*across* the time–bands) and the matched amounts account for the *shape* risk of the yield curve.
4. Finally, the net value $p_i^{(N)}$ is identical to the unmatched part of the portfolio and mirrors the yield curve *shift* risk.

Example 1.26 *Assume that we are short a call option on a 3 month interest rate future maturing in 2 months. For simplicity, we suppose that the market value of the underlying is* $m = 1000$ *and the price sensitivities are* $\delta = 0.3, \Gamma = 0.05$ *and* $\upsilon = 5$.
For determining the interest rate risk, the future has to be mapped into the maturity ladder (step 1.): a future can be decomposed into a combination of a long and a short position of different maturities. The 2 month time–band has a weight of 0.002 *and the 5 month time–band of* 0.004*; therefore we get (by* δ*–weighting the market value of the option's underlying):*

- *2 month risk–weighted position:* $0.002\delta m = 0.6$
- *5 month risk–weighted position:* $-0.004\delta m = -1.2$

Since there is only one position in each time–band, step 2. is dropped. Step 3. relates to the matching across the time–bands and leads to $p^{(M)} = \min\{0.6, 1.2\} = 0.6$*; the weighting coefficient for matching across time–bands of 0–12 months is* 0.4*. The net position of step 4. is* $p^{(N)} = |0.6 - 1.2| = 0.6$*, which results in an overall interest rate risk of* $\rho^{(1)} = -(p^{(N)} + 0.4p^{(M)}) = -0.84$.
The gamma–impact of the option is equal to $|\min\{0, -\frac{1}{2}\Gamma(0.002m)^2\}| = 0.1$*, which leads to an overall gamma risk of* $\rho^{(4)} = -0.1$.
Finally, the vega risk is $-\frac{1}{4}0.3|\upsilon| = -0.375$*, where it has been assumed that the implied volatility is equal to* 0.3.

[44] The relevant time–band depends on the residual term to maturity.

Once all the elements $\rho_i^{(\cdot)}(v_i)$ have been determined, the total risk of the portfolio $\rho(v)$ is obtained by adding the weighted[45] risk components

$$\rho(v) = \sum_{i=1}^{M} w_i \rho_i^{(\cdot)}(v_i), \quad w_i > 0, i = 1, \dots, M. \tag{1.7}$$

Remark 1.27 *The BIS Standardized Methodology does not take account of individual correlations among the risk factors nor of diversification effects. In fact, the addition of absolute values corresponds to assuming that all correlations are equal to* ± 1.

Example 1.28 *Suppose that the two commodities A and B are perfectly correlated in the sense that the price of B is l times the price of A (i.e. each price change of 1 unit for commodity A goes along with a price change of l units in B). We construct the following portfolio:*

- *take a long position of l physical units of commodity A*
- *take a short position of 1 physical unit of commodity B.*

This portfolio is perfectly hedged since each increase of x units in the price of A results in a profit of lx units for commodity A and an equal loss for B. However, the BIS rules lead to a total risk of $-2wl\psi_A$, *where* ψ_A *is the current price of 1 physical unit of A and w is the risk weight of the two assets.*

1.4 Coherency

What are the characteristics which a practical risk measure should have? Whereas the risk measurement techniques of Section 1.3 were first defined and only later their properties were investigated, Artzner et al. (1996) went the other way round and defined a set of axioms which are to be satisfied by every 'reasonable' measure of risk.

1.4.1 Definition of Coherent Risk Measures

The axiomatic approach of Artzner et al. (1996) prescribes what mathematical properties a meaningful risk measure, more precisely *capital requirement measure* $r(v)$, should have:[46]

[45] The fact that the risk weights w_i are not well differentiated for different assets i was heavily criticized by practitioners since it encourages taking positions in risky assets.

[46] A capital requirement measure $r(v)$ defines how much capital has to be held for portfolio v as a cushion against potential future losses.

Definition 1.29 *Let $\mathcal{G} = \{v \mid v : \mathcal{O} \to \mathbb{R}\}$ be the set of all functions on the set of realizable scenarios $\mathcal{O}$. A capital requirement measure $r(v)$ is called coherent if it satisfies the following properties:*

(P0) $r : \mathcal{G} \to \mathbb{R}^+$

(P1) $r(v) = r(v^-), \quad \forall v \in \mathcal{G}$

(P2) $r(\lambda v) = \lambda r(v), \quad \forall v \in \mathcal{G}, \lambda \in \mathbb{R}^+$

(P3) $r(v_1 + v_2) \leq r(v_1) + r(v_2), \quad \forall v_1, v_2 \in \mathcal{G}$

(P4) $r(v) \leq -\inf_{\omega \in \mathcal{O}} v(\omega), \quad \forall v \in \mathcal{G}$

(P5) $r(v - c) = r(v) + c, \quad \forall v \leq 0, c \in \mathbb{R}^+$

The condition (P0) states that the capital requirement should be a positive amount. Property (P1) says that the capital requirement may only depend on the losses and (P2) implies that the losses can be scaled: the capital for λ identical portfolios should be λ times the capital for one portfolio.[47]
The so–called *subadditivity* property (P3) is the most interesting one: it says that the capital requirement of two combined portfolios should not be greater than the sum of the individual requirements, hence, global risk management is encouraged. On the other hand, splitting a portfolio in two smaller parts should not reduce the capital requirement: there should be no possibilities of manipulating the regulatory standards by creation of subsidiaries. Subadditivity means also that risk measurement can safely be decentralized and that risk control can safely be delegated.

Example 1.30 *Imagine a risk control system based on additive limits: the total capital l_T is split up on two portfolios $l_T = l_1 + l_2$. If the capital requirement is not subadditive, it can happen that each portfolio is within its limits*

$$r(v_1) \leq l_1, \quad r(v_2) \leq l_2,$$

but that the capital limits of the entire institution are exceeded: $r(v_1 + v_2) \geq l_T$.

(P4) states that the capital requirements of a portfolio should not be higher than its maximal possible loss. Finally, (P5) says that if a portfolio, which has only negative outcomes, is affected by a supplementary deterministic loss of c units, this should increase the capital requirement by an amount of c.
At this point, it becomes important to distinguish the notions of 'risk' and 'capital requirement': A risk measure ρ provides a means for comparing the risk of different portfolios — its absolute value (unit of measurement) is unimportant.

[47] The assumption (P2) is only reasonable for liquid markets.

Contrarily, the *capital requirement* $r(v)$ says how much capital must be held as a protection against potential future losses.[48] Thus, the capital requirement of a portfolio is measured in some monetary unit and defined as

$$r(v) = -[p(\bar{\psi}) + h(\rho(v))], \tag{1.8}$$

where $p(\bar{\psi})$ is the profit or loss which incurs over the interval $[0, T]$ if the expected outcome $\bar{\psi}$ is attained at time T. The function $h : \mathbb{R} \to \mathbb{R}$ transforms the risk figure $\rho(v)$ into a capital charge.[49]
To define what properties a 'good' *risk measure* $\rho(v)$ should have, the notion of *coherent capital requirement* needs to be adapted. Since the main purpose of a risk measure $\rho(v)$ is to *enable a comparison* of several portfolios, a limitation of $\rho(v)$ to some absolute value, as stated in (P4), is unimportant. Moreover, (P5) can be seen as a condition which relates $\rho(v)$ and $p(\bar{\psi})$, but it makes no statement about $\rho(v)$ itself. Therefore, we will define a 'reasonable' measure of risk as one which satisfies the properties (P0)–(P3) of the original framework:

Definition 1.31 *A risk measure $\rho : \mathcal{V}_\rho \to \mathbb{R}$ is called weakly coherent if it satisfies conditions (P0') to (P3'):*

(P0') $\rho : \mathcal{V}_\rho \to \mathbb{R}^-$

(P1') $\rho(v) = \rho(v^-), \quad \forall v \in \mathcal{V}_\rho$

(P2') $\rho(\lambda v) = \lambda\rho(v), \quad \forall v \in \mathcal{V}_\rho, \lambda \in \mathbb{R}^+$

(P3') $\rho(v_1 + v_2) \geq \rho(v_1) + \rho(v_2), \quad \forall v_1, v_2 \in \mathcal{V}_\rho$

The conditions (P1'), (P2') and (P3') imply several characteristics:

Lemma 1.32 *If ρ is weakly coherent, then $\rho(0) = 0$.*

Proof: This is an immediate consequence of (P2'). ∎

Thus, an empty portfolio has now risk. More generally, a portfolio which produces only gains has no risk:

Corollary 1.33 *If ρ is weakly coherent and $v \geq 0$, then $\rho(v) = 0$.*

Proof: Trivial, since $\rho(v) = \rho(v^-) = \rho(0)$ by (P1'). ∎

Finally, a weakly coherent risk measure allows to compare the risk of portfolios since it defines a partial ordering:

[48] See Duffie and Pan (1997), p. 4.
[49] The BIS *multiplication factor* of 3 is an example of such a transformation, cf. BIS (1996a), pp. 3.

Theorem 1.34 *If ρ is weakly coherent, then $\rho(v_1) \leq \rho(v_2)$ for all $v_1 \leq v_2$.*

Proof: Let $v_2 = v_1 + (v_2 - v_1)$. The superadditivity property (P3') implies that

$$\rho(v_2) \geq \rho(v_1) + \rho(v_2 - v_1).$$

But $\rho(v_2 - v_1) = 0$ by Corollary 1.33, hence $\rho(v_2) \geq \rho(v_1)$. ■

1.4.2 Discussion of Risk Measurement Techniques

Value–at–Risk

Remark 1.35 *In general, VaR is not weakly coherent since it does not satisfy the superadditivity property (P3').*

Example 1.36 *Consider two portfolios A and B, the first with P&L function $v_A(\omega) = -\omega_1^2$, having no positions in risk factor ω_2, and the second $v_B(\omega) = -\omega_2^2$. If the two risk factors are independent and standard normally distributed, the P&L functions of A and B are both χ^2 distributed with one degree of freedom and therefore*

$$\mathrm{VaR}_{\mathrm{A}}(75\%) = \mathrm{VaR}_{\mathrm{B}}(75\%) = -1.32.$$

On the other hand, the composite portfolio $(A \cup B)$ is χ^2 distributed with two degrees of freedom: $\mathrm{VaR}_{\mathrm{A\cup B}}(75\%) = -2.77$, which shows that the superadditivity condition $\mathrm{VaR}_{\mathrm{A\cup B}}(75\%) \geq \mathrm{VaR}_{\mathrm{A}}(75\%) + \mathrm{VaR}_{\mathrm{B}}(75\%)$ is violated.

However, there is one important exception among the quantile based risk measures:

Theorem 1.37 *Delta–Normal VaR for confidence levels $\alpha \geq 50\%$ is weakly coherent.*

Proof: It is easy to verify that Delta–Normal $\mathrm{VaR}(\alpha) = -z_\alpha \sqrt{a^T \Sigma a}$ satisfies conditions (P0') to (P2') if $\alpha \geq 50\%$, since in this case $z_\alpha \geq 0$.
The superadditivity property (P3') is also satisfied: let $v_{l,A}(\omega) = a^T\omega$ and $v_{l,B}(\omega) = b^T\omega$ be two linear portfolios and take the Cholesky decomposition of the covariance matrix $\Sigma = U^T U$. If we define $\tilde{a} = Ua$ and $\tilde{b} = Ub$, then

$$\begin{aligned}
(a+b)^T\Sigma(a+b) &= \tilde{a}^T\tilde{a} + \tilde{b}^T\tilde{b} + 2\tilde{a}^T\tilde{b} \\
&\leq \tilde{a}^T\tilde{a} + \tilde{b}^T\tilde{b} + 2\sqrt{\tilde{a}^T\tilde{a}}\sqrt{\tilde{b}^T\tilde{b}} \\
&\leq (\sqrt{\tilde{a}^T\tilde{a}} + \sqrt{\tilde{b}^T\tilde{b}})^2 \\
&\leq (\sqrt{a^T\Sigma a} + \sqrt{b^T\Sigma b})^2,
\end{aligned}$$

where the inequality is due to the Schwarz inequality. Taking the square root and multiplying by $(-z_\alpha)$, which is negative for $\alpha \geq 50\%$, proves the superadditivity. ■

BIS Standardized Methodology

Under the BIS Standardized Methodology[50], the total risk of a portfolio is simply a weighted sum of individual risk components. Each component is a function of the long and short positions in one risk factor; either their sum, maximum, net value or matched value.[51]

Theorem 1.38 *The BIS Standardized Methodology is a weakly coherent risk measure.*

Proof: By formula (1.7), $\rho(v) = \sum_{i=1}^{M} w_i \rho_i^{(\cdot)}(v_i)$ is the weighted sum of individual risk components, where $w_i > 0, i = 1, \ldots, M$. Thus, it is sufficient to prove that each component of type $\rho^{(1)}, \ldots, \rho^{(4)}$ is weakly coherent. Property (P0') is obviously satisfied because each $\rho^{(\cdot)}$ is of the form $-|x|$. (P1') and (P2') are fulfilled because each individual P&L function $v_i(\omega_i)$ is linear by formula (1.6). The superadditivity property is not obvious, since the matching procedure does not satisfy (P3').[52] However, matched positions $p^{(M)}$ always appear together with netted positions $p^{(N)}$ and we will show that their combination is superadditive. To shorten the notation, we fix one risk factor for all four cases. $p_A^{(l)}$ and $p_A^{(s)}$ denote the net long and net short positions in portfolio A, $p_B^{(l)}$ and $p_B^{(s)}$ those in portfolio B:

- Component $\rho^{(1)}$:
  Since $p_A^{(N)} + p_A^{(M)} = \max\{p_A^{(l)}, p_A^{(s)}\}$, it follows that

$$\begin{aligned} p_{A\cup B}^{(N)} + p_{A\cup B}^{(M)} &= \max\{p_{A\cup B}^{(l)}, p_{A\cup B}^{(s)}\} \\ &= \max\{p_A^{(l)} + p_B^{(l)}, p_A^{(s)} + p_B^{(s)}\} \\ &\leq \max\{p_A^{(l)}, p_A^{(s)}\} + \max\{p_B^{(l)}, p_B^{(s)}\} \\ &\leq (p_A^{(N)} + p_A^{(M)}) + (p_B^{(N)} + p_B^{(M)}). \end{aligned} \tag{1.9}$$

[50] See Section 1.3.2.

[51] See Definitions 1.24 and 1.25.

[52] To see this, take a portfolio A which holds only long positions $p^{(l)}$ of an asset and a portfolio B which holds only short positions $p^{(s)}$. Then,

$$p_{A\cup B}^{(M)} = \min\{p^{(l)}, p^{(s)}\} \geq \min\{p^{(l)}, 0\} + \min\{0, p^{(s)}\} = p_A^{(M)} + p_B^{(M)},$$

since $p^{(l)}, p^{(s)} \geq 0$ by Definition 1.24. Thus, we get for the risk, which is a negative amount, $-p_{A\cup B}^{(M)} \leq -(p_A^{(M)} + p_B^{(M)})$, which contradicts superadditivity.

Hence, $\kappa(p^{(N)}_{A\cup B} + p^{(M)}_{A\cup B}) \leq \kappa(p^{(N)}_A + p^{(M)}_A + p^{(N)}_B + p^{(M)}_B)$ for $\kappa \geq 0$. The definition of $\rho^{(1)}$ leads to

$$\begin{aligned}\rho^{(1)}_{A\cup B}(v_{A\cup B}) &= -(1-\kappa)p^{(N)}_{A\cup B} - \kappa(p^{(N)}_{A\cup B} + p^{(M)}_{A\cup B}) \\ &\geq -(1-\kappa)(p^{(N)}_A + p^{(N)}_B) - \kappa(p^{(N)}_A + p^{(M)}_A + p^{(N)}_B + p^{(M)}_B) \\ &\geq \rho^{(1)}_A(v_A) + \rho^{(1)}_B(v_B)\end{aligned}$$

as long as $0 \leq \kappa \leq 1$.

- Component $\rho^{(2)}$:
  Using the definition of the gross value, we get

$$\begin{aligned}\rho^{(2)}_A(v_A) + \rho^{(2)}_B(v_B) &= -(p^{(l)}_A + p^{(s)}_A + p^{(l)}_B + p^{(s)}_B) \\ &= -(p^{(l)}_{A\cup B} + p^{(s)}_{A\cup B}) \\ &= \rho^{(2)}_{A\cup B}(v_{A\cup B}).\end{aligned}$$

- Component $\rho^{(3)}$:
  The proof for the higher net has already been given in equation (1.9), since $\rho^{(3)}(v) = \max\{p^{(l)}, p^{(s)}\} = p^{(N)} + p^{(M)}$.

- Component $\rho^{(4)}$:
  Set $p^{(l)}_A = p^{(l)}_B = 0$, $p^{(s)}_A = |\min\{0, \frac{1}{2}\sum_{j\in H_i\cap A}\Gamma_j m_j^2\}|$, $p^{(s)}_B = |\min\{0, \frac{1}{2}\sum_{j\in H_i\cap B}\Gamma_j m_j^2\}|$ and compare to the case of $\rho^{(2)}$.

- Component $\rho^{(5)}$:
  Set $p^{(l)}_A = p^{(l)}_B = 0$, $p^{(s)}_A = \frac{1}{4}\sigma_i|\sum_{j\in H_i\cap A} v_j|$, $p^{(s)}_B = \frac{1}{4}\sigma_i|\sum_{j\in H_i\cap B} v_j|$ and compare to $\rho^{(2)}$, noting that the volatility σ_i is nonnegative.

■

1.5 Maximum Loss

This section introduces the fundamentals of the Maximum Loss (ML) risk measure. Maximum Loss is principally a methodology to determine the worst case scenario under normal market conditions, without ignoring the correlations among the risk factors. It is defined as

The maximum loss that can occur

- over some holding period T
- if the risk factors are restricted to a given set $\mathcal{T} \subseteq \mathcal{O}$,

where $\mathcal{T}$ is a closed set which contains the scenario $\omega = 0$ and has a confidence level $\mathrm{P}(\mathcal{T}) = \alpha$. This definition of ML has some similarities to the definition of VaR in Section 1.3.1. However, there is one important difference: Whereas for calculating VaR the *distribution* of P&L has to be known (and whose determination is the crucial point in all VaR methods), ML is directly defined on the set of realizable scenarios $\mathcal{O}$. The mathematical definition of Maximum Loss is

Definition 1.39 *Let $v(\omega)$ be the P&L function. Maximum Loss is the solution of the minimization problem*

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad v(\omega) \\ & \text{s.t.} \quad \omega \in \mathcal{T}, \end{aligned}$$

where $\mathcal{T}$ is a given closed subset of $\mathcal{O}$ which must satisfy the conditions $0 \in \mathcal{T}$ and $\mathrm{P}(\mathcal{T}) = \alpha$.

For a given portfolio, VaR depends on two parameters: the holding period T and the confidence level α. ML, however, has a supplementary degree of freedom, called 'trust region' $\mathcal{T}$: any closed subset of $\mathcal{O}$ which contains the scenario $\omega = 0$ and which has a probability α is a valid trust region. ML represents the maximal loss over such a trust region (cf. Figure 1.4).
The most important difference to the previously discussed risk measurement techniques is that the solution process of ML not only produces a risk figure, but that it also determines the worst case scenario ω^*. This scenario tells which combination of risk factors is most dangerous for the portfolio and it is a basic element for a proactive risk management. For example, the knowledge of ω^* allows to determine risk reducing transactions or to estimate the incremental risk of a single trade; a detailed discussion follows in Chapter 3.

1.5.1 Coherency of Maximum Loss

Theorem 1.40 *Maximum Loss is a weakly coherent risk measure.*

Proof: The conditions (P0') and (P1') are satisfied because the scenario $\omega = 0$ is, by construction, an element of the trust region $\mathcal{T}$ and $v(0) = 0$. Basic properties of the minimum operator assert (P2') and (P3') for a fixed set $\mathcal{T}$. ■

Hence, Maximum Loss has some theoretical advantages over the previously discussed methods: ML is a risk measure for nonlinear portfolios, it is weakly coherent and respects (by the choice of the trust region $\mathcal{T}$) correlations. In contrast,

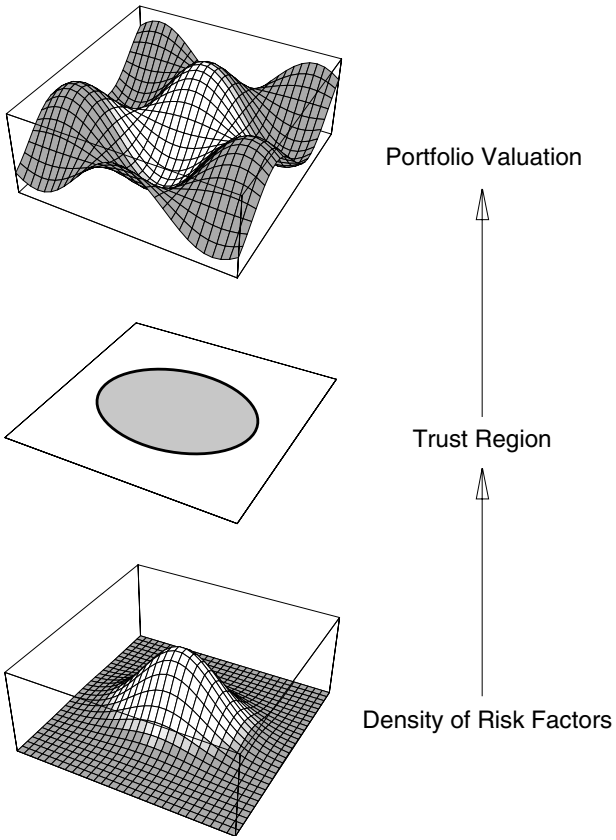

Figure 1.4: Modelling Maximum Loss

VaR is not weakly coherent for nonlinear portfolios and the BIS Standardized Methodology does not take correlations between risk factors into account.

Remark 1.41 *All risk measurement techniques which are based on a worst case analysis of a discrete set of scenarios, such as stress testing, the factor push methodology or the scanning charge of the SPAN framework*[53]*, are weakly coherent.*

[53]In the SPAN methodology, however, the inter–commodity spread credits are not weakly coherent, which implies that SPAN's overall risk measure is *not* weakly coherent.

Chapter 2

Risk Factor Models

2.1 Introduction

This section shows how to transform the financial realm into a mathematical model which is adapted to the Maximum Loss methodology. More precisely, it will be explained of what type the relations between market rates and risk factors should be and how these risk factors can be modelled. The key idea is to choose a set of risk factors which allows to value the portfolio in a straightforward manner. Contrarily to most other risk measurement techniques, the ML methodology does not only determine a risk figure, but it also offers the capability of identifying worst case *scenarios*. Consequently, the set of the realizable scenarios $\mathcal{O}$ can be seen as a set in the *space of the risk factors*, and the following conditions should be satisfied:

- the set of risk factors $\{\omega_1, \ldots, \omega_M\}$ should be rich enough[1] to reflect all relevant risks[2]
- every scenario $\omega \in \mathcal{O}$ should give all the information which is required to value the portfolio, i.e. to determine the P&L $v(\omega)$
- the risk factors should be interpretable, i.e. they should reflect real market rates and not some abstract quantity
- every scenario $\omega \in \mathcal{O}$ should represent a consistent set of market rates (no–arbitrage conditions should be satisfied).

Whereas RiskMetrics requires a preliminary step where cash flows are mapped into a maturity ladder[3], there is no corresponding operation in the ML technique.

[1]For example, market rates which have a term structure need to be modelled as a finite number of risk factors, or different equities are often aggregated into a single variable.

[2]In practice, the number M of risk factors is limited by the availability of reliable data for estimating the covariance matrix Σ and by the restricted resources (space and time) for numerical computations.

[3]See RiskMetrics (1996), pp. 107.

All that needs to be done is to value the portfolio for a prescribed set of scenarios $\mathcal{S} = \{\omega^{(1)}, \ldots, \omega^{(n)}\}$. The resulting profit and losses $\mathcal{P} = \{\xi^{(1)}, \ldots, \xi^{(n)}\}$ are then used to construct an approximating P&L function $v(\omega)$; details of this procedure are discussed in Section 4.2.
Since one of the goals of Maximum Loss is to identify the most dangerous developments of *a market*, we will limit the risk factor model to general market risks; *specific risks* which are related to individual issues or issuers will not be part of the risk factor model.[4]
In the following, we discuss risk factor models which satisfy all of the above conditions; the sections are organized by the nature of risks:

- currency risk
- equity risk
- interest rate risk
- commodity risk.

Remark 2.1 *The value of contingent claims depends not only on the underlying ψ_i, but also on the volatility[5] of ψ_i. For ease of notation, we will denote the volatility of the market rate ψ_i by $\psi_{\sigma,i}$. It should be noted that $\psi_{\sigma,i}$ is not directly related to the variance $(\Sigma)_{i,i}$ of risk factor ω_i because $\psi_{\sigma,i}$ and $(\Sigma)_{i,i}$ refer to different time horizons.*
Since the volatility $\psi_{\sigma,i}$ can vary during the time interval $[0, T]$, we need a risk factor $\omega_{\sigma,i}$ which reflects the deviation of the real outcome $\psi_{\sigma,i}$ from the expected value $\bar{\psi}_{\sigma,i}$ at time T. Consequently, the covariance matrix Σ must also hold the covariances of the various volatility risk factors.[6]

2.2 Currency Risk

If positions in foreign currencies are held, two types of risk can occur:

- FX rate risk: the risk due to changes in exchange rates
- FX volatility risk: the risk related to changes *in the volatilities* of the exchange rates.

[4]However, once the most adverse movements of the total market have been identified, it is possible to assess the Maximum Loss of specific risks in a second model, which uses risk factors which are related to individual issues/issuers.

[5]When speaking about valuation, the technical term 'volatility' usually refers to the annualized standard deviation.

[6]Procedures for estimating the variance of the volatility $\psi_{\sigma,i}$ are described in Cox and Rubinstein (1985), pp. 255 and pp. 278.

2.2.1 FX Rate Risk

When working with currencies, consistency plays a particularly important role:

Definition 2.2 *Let H denote the reporting currency*[7]*, and let A and B be two foreign currencies. The value of the foreign exchange rate $\frac{H}{A}$ at time T is denoted by $\psi_{\frac{H}{A}}$ and the corresponding risk factor is $\omega_{\frac{H}{A}}$. Similarly, $\omega_{\frac{H}{B}}$ is the risk factor of the rate $\frac{H}{B}$ with value $\psi_{\frac{H}{B}}$.*

The triangular relationship (no–arbitrage condition) between the foreign exchange rates $\frac{H}{A}$ and $\frac{H}{B}$ determines the value of the *cross rate* $\psi_{\frac{A}{B}}$ uniquely:

Lemma 2.3 *The value of the cross rate $\frac{A}{B}$ is*

$$\psi_{\frac{A}{B}} = \frac{\omega_{\frac{H}{B}} + \bar{\psi}_{\frac{H}{B}}}{\omega_{\frac{H}{A}} + \bar{\psi}_{\frac{H}{A}}},$$

where, according to Definition 1.1, $\bar{\psi}_i$ denotes the expected outcome of risk factor i at the end of the holding period.

Proof: The triangular relationship implies that $\psi_{\frac{A}{B}} = \frac{\psi_{\frac{H}{B}}}{\psi_{\frac{H}{A}}} = \frac{\omega_{\frac{H}{B}} + \bar{\psi}_{\frac{H}{B}}}{\omega_{\frac{H}{A}} + \bar{\psi}_{\frac{H}{A}}}$, since $\omega_i = \psi_i - \bar{\psi}_i$ by Definition 1.1. ■

Hence, the risk factors of the *foreign exchange rates* determine uniquely the values of all cross rates. Moreover, the *distributions* of the foreign exchange rates determine the distributions of the cross rates:

2.2.2 FX Volatility Risk

In Section 1.2.2 it was shown that the standard assumption of finance leads to $\frac{X_t - X_s}{X_s} \approx (\mu_X - \frac{\sigma_X^2}{2})(t-s) + \sigma_X(B_t - B_s), 0 \leq s < t$, where $(X_t)_{t \geq 0}$ is an Itô–process[8] satisfying $dX_t = \mu_X X_t dt + \sigma_X X_t dB_t$. If we take this approximation as an equality for infinitesimal time intervals dt, we get with a slight abuse of notation

$$\frac{dX_t}{X_t} \sim \mathcal{N}\left(\left(\mu_X - \frac{\sigma_X^2}{2}\right)dt, \sigma_X^2 dt\right).$$

If $\left(X_t^{\frac{H}{A}}\right)_{t \geq 0}$ denotes the process of the foreign exchange rate $\frac{H}{A}$ and $\left(X_t^{\frac{H}{B}}\right)_{t \geq 0}$ of $\frac{H}{B}$, it follows that

[7] The reporting currency is the currency used for measuring the P&L.

[8] See Appendix A.

Theorem 2.4 *If* $\frac{dX_t^{\frac{H}{A}}}{X_t^{\frac{H}{A}}} \sim \mathcal{N}(\mu_{\frac{H}{A}}dt, \sigma^2_{\frac{H}{A}}dt)$ *and* $\frac{dX_t^{\frac{H}{B}}}{X_t^{\frac{H}{B}}} \sim \mathcal{N}(\mu_{\frac{H}{B}}dt, \sigma^2_{\frac{H}{B}}dt)$ *with* $\mathrm{Corr}\left(\frac{dX_t^{\frac{H}{A}}}{X_t^{\frac{H}{A}}}, \frac{dX_t^{\frac{H}{B}}}{X_t^{\frac{H}{B}}}\right) = \rho_{\frac{H}{A},\frac{H}{B}}$, *then* $\frac{dX_t^{\frac{A}{B}}}{X_t^{\frac{A}{B}}} \sim \mathcal{N}(\mu_{\frac{A}{B}}dt, \sigma^2_{\frac{A}{B}}dt)$, *where*

$$\sigma^2_{\frac{A}{B}} = \sigma^2_{\frac{H}{A}} + \sigma^2_{\frac{H}{B}} - 2\rho_{\frac{H}{A},\frac{H}{B}}\sigma_{\frac{H}{A}}\sigma_{\frac{H}{B}}.$$

Proof: Since $X_t^{\frac{A}{B}} = \frac{X_t^{\frac{H}{B}}}{X_t^{\frac{H}{A}}}$, it follows that $dX_t^{\frac{A}{B}} = \frac{1}{X_t^{\frac{H}{A}}}dX_t^{\frac{H}{B}} - \frac{X_t^{\frac{H}{B}}}{\left(X_t^{\frac{H}{A}}\right)^2}dX_t^{\frac{H}{A}}$, and consequently $\frac{dX_t^{\frac{A}{B}}}{X_t^{\frac{A}{B}}} = \frac{dX_t^{\frac{H}{B}}}{X_t^{\frac{H}{B}}} - \frac{dX_t^{\frac{H}{A}}}{X_t^{\frac{H}{A}}}$, which shows that $\frac{dX_t^{\frac{A}{B}}}{X_t^{\frac{A}{B}}}$ is the difference of two normally distributed variables with $\mu_{\frac{A}{B}} = \mu_{\frac{H}{B}} - \mu_{\frac{H}{A}}$ and $\sigma^2_{\frac{A}{B}} = \sigma^2_{\frac{H}{A}} + \sigma^2_{\frac{H}{B}} - 2\rho_{\frac{H}{A},\frac{H}{B}}\sigma_{\frac{H}{A}}\sigma_{\frac{H}{B}}$. ■

This relation allows to calculate the volatility of the cross rate from the volatility risk factors of the foreign exchange rates:

Corollary 2.5 *Let* $\omega_{\sigma,\frac{H}{A}}$ *be the volatility risk factor of the foreign exchange rate* $\frac{H}{A}$ *and* $\omega_{\sigma,\frac{H}{B}}$ *the volatility risk factor of* $\frac{H}{B}$. *If* $\mathrm{Corr}\left(\frac{dX_t^{\frac{H}{A}}}{X_t^{\frac{H}{A}}}, \frac{dX_t^{\frac{H}{B}}}{X_t^{\frac{H}{B}}}\right) = \rho_{\frac{H}{A},\frac{H}{B}}$, *then the volatility of the cross rate* $\frac{A}{B}$ *is*

$$\psi_{\sigma,\frac{A}{B}} = \sqrt{(\omega_{\sigma,\frac{H}{A}} + \bar{\psi}_{\sigma,\frac{H}{A}})^2 + (\omega_{\sigma,\frac{H}{B}} + \bar{\psi}_{\sigma,\frac{H}{B}})^2 - 2\rho_{\frac{H}{A},\frac{H}{B}}(\omega_{\sigma,\frac{H}{A}} + \bar{\psi}_{\sigma,\frac{H}{A}})(\omega_{\sigma,\frac{H}{B}} + \bar{\psi}_{\sigma,\frac{H}{B}})}.$$

Proof: From Theorem 2.4 we conclude that

$$\psi^2_{\sigma,\frac{A}{B}} = \psi^2_{\sigma,\frac{H}{A}} + \psi^2_{\sigma,\frac{H}{B}} - 2\rho_{\frac{H}{A},\frac{H}{B}}\psi_{\sigma,\frac{H}{A}}\psi_{\sigma,\frac{H}{B}}.$$

Then, Definition 1.1 of risk factors implies that $\psi_{\sigma,\frac{H}{A}} = \omega_{\sigma,\frac{H}{A}} + \bar{\psi}_{\sigma,\frac{H}{A}}$, which proves the corollary. ■

Consequently, all the quantities which concern the cross rates are uniquely determined by the foreign exchange rates:

Remark 2.6 *The risk factors which are used to model currency risk are*

- *the changes in the values of the foreign exchange rates (i.e.* $\omega_{\frac{H}{A}}, \omega_{\frac{H}{B}}$*)*
- *the changes in the volatilities of the foreign exchange rates (i.e.* $\omega_{\sigma,\frac{H}{A}}, \omega_{\sigma,\frac{H}{B}}$*).*

Hence, when a scenario ω is evaluated in order to determine $v(\omega)$, the implied values of the cross rates and their volatilities have to be considered in the valuation of the portfolio.

Example 2.7 *Assume that we have a portfolio depending on two foreign currencies A, B and that H denotes the reporting currency. The expected outcomes of the foreign exchange rates at the end of the holding period are*

$$\bar{\psi}_{\frac{H}{A}} = 1.20, \quad \bar{\psi}_{\frac{H}{B}} = 0.80$$

and the expected volatilities are

$$\bar{\psi}_{\sigma,\frac{H}{A}} = 0.15, \quad \bar{\psi}_{\sigma,\frac{H}{B}} = 0.15.$$

Furthermore, we suppose that the correlation between the two rates is $\rho_{\frac{H}{A},\frac{H}{B}} = 0.25$. If the portfolio is valued for the scenario

$$\omega = (\omega_{\frac{H}{A}}, \omega_{\frac{H}{B}}, \omega_{\sigma,\frac{H}{A}}, \omega_{\sigma,\frac{H}{B}}) = (0.05, -0.05, 0.05, -0.05),$$

then not only the instruments depending on the foreign exchange rates $\psi_{\frac{H}{A}}, \psi_{\frac{H}{B}}$ and the volatilities $\psi_{\sigma,\frac{H}{A}}, \psi_{\sigma,\frac{H}{B}}$ have to be considered, but also all positions which depend on the cross rate $\frac{A}{B}$. In fact, Lemma 2.3 implies that

$$\psi_{\frac{A}{B}} = \frac{-0.05 + 0.80}{0.05 + 1.20} = 0.60,$$

and

$$\begin{aligned}\psi_{\sigma,\frac{A}{B}} &= \sqrt{(0.05+0.15)^2 + (-0.05+0.15)^2 - 2 \cdot 0.25(0.05+0.15)(-0.05+0.15)} \\ &= 0.20.\end{aligned}$$

by Corollary 2.5.

2.3 Equity Risk

In the equity markets, we distinguish two main sources of risk:

- changes of equity prices
- changes in the volatilities of equity prices.

2.3.1 Equity Price Risk

This section shows how to model the general market risk of equity positions (i.e. the risk which is due to overall movements in one market). The state of a market will be represented by a market index I. The relationship between the individual assets and this index is derived from the capital asset pricing model (CAPM)[9]:

[9] See Copeland and Weston (1995), pp. 195.

Definition 2.8 *The return of asset j during the holding period of length T is $r_j = \frac{\psi_j - \psi_j^0}{\psi_j^0}$, where ψ_j^0 is the value of the asset at the beginning of the holding period and ψ_j its value at the end of the period. Similarly, $r_I = \frac{\psi_I - \psi_I^0}{\psi_I^0}$ is the return of the market index I.*

Assumption 2.9 *According to Section 1.2.2, it is assumed that that the returns are normally distributed: $r_j \sim \mathcal{N}(\mu_j, \sigma_j^2)$ and $r_I \sim \mathcal{N}(\mu_I, \sigma_I^2)$ with covariance* $\mathrm{Cov}(r_j, r_I) = \sigma_{j,I}$.

The capital asset pricing model pretends a linear relationship between the return of an individual asset r_j and the return of the market index r_I:

$$r_j = a_j + b_j r_I + \varepsilon_j, \tag{2.1}$$

where a_j and b_j are constants and ε_j is a random variable with mean 0 and $\mathrm{Cov}(r_I, \varepsilon_j) = 0$. Thus, ε_j plays the role of a random perturbation which is related to asset j and independent of the market index. Of course, the relationship between r_j and r_I is strongest when the perturbation ε_j has a minimal effect:

Theorem 2.10 *If $r_j = a_j + b_j r_I + \varepsilon_j$, then ε_j is normally distributed and it has minimal variance if $b_j = \frac{\sigma_{j,I}}{\sigma_I^2}$.*

Proof: Since r_j and r_I are normally distributed, $\varepsilon_j = r_j - a_j - b_j r_I$ is also normally distributed, i.e. $\varepsilon_j \sim \mathcal{N}(0, \sigma_{\varepsilon_j}^2)$. Formula (2.1) implies that that

$$\sigma_{\varepsilon_j}^2 = \sigma_j^2 + b_j^2 \sigma_I^2 - 2 b_j \sigma_{j,I}.$$

It is easy to verify that this expression is minimized if $b_j = \frac{\sigma_{j,I}}{\sigma_I^2}$; in this case, $\sigma_{\varepsilon_j}^2 = \sigma_j^2 - \frac{\sigma_{j,I}^2}{\sigma_I^2}$ is the minimal variance.[10] ∎

Corollary 2.11 *If $r_j = a_j + b_j r_I + \varepsilon_j$ and ε_j has minimal variance, then $a_j = \mu_j - \frac{\sigma_{j,I}}{\sigma_I^2}\mu_I$.*

Proof: In Theorem 2.10, b_j and $\sigma_{\varepsilon_j}^2$ have been determined uniquely. Taking the expectation of formula (2.1), it follows that $a_j = \mu_j - \frac{\sigma_{j,I}}{\sigma_I^2}\mu_I$. ∎

Hence, we have determined all parameters of an optimal model of the form (2.1). It remains to make the link with the risk factors, i.e. we have to express the equity price ψ_j as a function of risk factor ω_I:

[10] The quantity $b_j^2 \sigma_I^2$ is usually called *systematic risk*, whereas $\sigma_{\varepsilon_j}^2$ is called *unsystematic risk*.

Theorem 2.12 *If* $r_j = a_j + b_j r_I + \varepsilon_j$ *with* $\varepsilon_j \sim \mathcal{N}(0, \sigma^2_{\varepsilon_j})$ *and* $\mathrm{Cov}(r_I, \varepsilon_j) = 0$*, then there exist constants* α_j, β_j *and a random variable* $\eta_j \sim \mathcal{N}(0, \sigma^2_{\eta_j})$ *with* $Cov(\omega_I, \eta_j) = 0$ *such that*

$$\psi_j = \alpha_j + \beta_j \omega_I + \eta_j.$$

Proof: Rewriting formula (2.1) leads to $\frac{\psi_j - \psi_j^0}{\psi_j^0} = a_j + b_j \frac{\psi_I - \psi_I^0}{\psi_I^0} + \varepsilon_j$ and hence

$$\psi_j = (1 + a_j)\psi_j^0 + b_j \frac{\psi_j^0}{\psi_I^0}(\psi_I - \psi_I^0) + \psi_j^0 \varepsilon_j.$$

Using Definition 1.1 of risk factors $\omega_I = \psi_I - \bar{\psi}_I$, it follows that

$$\begin{aligned} \alpha_j &= (1 + a_j)\psi_j^0 + b_j \frac{\psi_j^0}{\psi_I^0}(\bar{\psi}_I - \psi_I^0) \\ \beta_j &= b_j \frac{\psi_j^0}{\psi_I^0} \\ \eta_j &= \psi_j^0 \varepsilon_j. \end{aligned}$$

■

Since we know that $\mathrm{E}(\eta_j) = 0$, we have an explicit relation between changes in the market index and the price of asset j. Similarly, there exists a relationship between the volatility $\psi_{\sigma,j}$ of asset j and the volatility $\psi_{\sigma,I}$ of the market index:

2.3.2 Equity Volatility Risk

Theorem 2.13 *Let* $\hat{r}_j \sim \mathcal{N}(\hat{\mu}_j, \hat{\sigma}_j^2)$ *denote the annualized return of asset* j *and* $\hat{r}_I \sim \mathcal{N}(\hat{\mu}_I, \hat{\sigma}_I^2)$ *the annualized return of the market index* I*. If* $\hat{r}_j = \hat{a}_j + \hat{b}_j \hat{r}_I + \hat{\varepsilon}_j$ *with* $\mathrm{Cov}(\hat{r}_I, \hat{\varepsilon}_j) = 0$*, then*

$$\psi_{\sigma,j} = \sqrt{\hat{b}_j^2(\omega_{\sigma,I} + \bar{\psi}_{\sigma,I})^2 + \hat{\sigma}^2_{\hat{\varepsilon}_j}},$$

where $\bar{\psi}_{\sigma,I}$ *is the expected value of the market index volatility and* $\omega_{\sigma,I}$ *is the corresponding risk factor.*

Proof: Taking the variance on both sides of $\hat{r}_j = \hat{a}_j + \hat{b}_j \hat{r}_I + \hat{\varepsilon}_j$ we get $\psi^2_{\sigma,j} = \hat{b}_j^2 \psi^2_{\sigma,I} + \hat{\sigma}^2_{\hat{\varepsilon}_j}$. Then, the definition of the risk factors: $\omega_{\sigma,I} = \psi_{\sigma,I} - \bar{\psi}_{\sigma,I}$ leads immediately to the result. ■

The relationships between individual assets and the market index have been established in Theorems 2.12 and 2.13. Thus, we have been able to model the general equity risk of a market with only two risk factors: ω_I for the market index

and $\omega_{\sigma,I}$ for its volatility. It should be noticed that it is easily possible to refine this model by adding supplementary risk factors for specific market sectors.

Example 2.14 *Over a ten–day holding period, the return r_j of stock j is a random variable with $r_j \sim \mathcal{N}(0, 0.02)$ and the market index I has a return $r_I \sim \mathcal{N}(0, 0.01)$ with $\mathrm{Cov}(r_j, r_I) = 0.005$.*
According to Theorem 2.10 and Corollary 2.11, the optimal parameters of the model $r_j = a_j + b_j r_I + \varepsilon_j$ are

$$a_j = 0, \quad b_j = \frac{0.005}{0.01} = 0.5.$$

Assume that the stock price at the beginning of the holding period is $\psi_j^0 = 250$, the market index is $\psi_I^0 = 100$ and its expected value at the end of the period is $\bar{\psi}_I = 110$. Then, Theorem 2.12 implies that the value of the stock ψ_j depends linearly on the risk factor of the market index: $\psi_j = \alpha_j + \beta_j \omega_I$, where

$$\alpha_j = 250 + 0.5\frac{250}{100}(110 - 100) = 262.5, \quad \beta_j = 0.5\frac{250}{100} = 1.25.$$

Hence, the scenario $\omega_I = -2$ implies a stock price of $\psi_j = 262.5 + 1.25(-2) = 260$.
What concerns volatilities, we will assume that the annualized returns can be modelled as $\hat{r}_j = \hat{a}_j + \hat{b}_j \hat{r}_I + \hat{\varepsilon}_j$. By Theorem 2.10, the perturbation $\varepsilon_j \sim \mathcal{N}(0, \sigma_{\varepsilon_j}^2)$ has a variance of

$$\sigma_{\varepsilon_j}^2 = 0.02 - \frac{0.005^2}{0.01} = 0.0175.$$

Under the hypothesis of geometric Brownian motion, the 'square root of time rule' leads to $\hat{\sigma}_{\varepsilon_j}^2 = 0.0175(\frac{250}{10}) = 0.4375$ (where it is assumed that a year counts 250 trading days).
Hence, if the expected volatility of the index is $\bar{\psi}_{\sigma,I} = 0.55$ at the end of the ten–day holding period and if $\omega_{\sigma,I} = -0.05$, then Theorem 2.13 implies that

$$\psi_{\sigma,j} = \sqrt{0.5^2(-0.05 + 0.55)^2 + 0.4375} = \sqrt{0.5}.$$

2.4 Interest Rate Risk

The problem in dealing with interest rate risk is that a single interest rate of a market does not exist. Indeed, the yield of a particular instrument may depend on its maturity, possible call features, coupon payments, the credit rating of the issuer, etc.
In the sequel, *spot rate* refers to the current market yield of a default–free zero–coupon instrument and the spot rate curve is the graphical depiction of the

relationship between spot rates and their maturities. This curve[11], together with its volatility, can be used to price any default–free instrument.[12]
The price of non default–free instruments additionally depends on the credit rating of their issuers.[13] The lower their rating, the higher the risk premium that has to be paid. This is reflected by higher *spreads*, i.e. by increasing differences between the spot rate and zero–coupon rates of rating classes of lower quality (cf. Figure 2.1).[14]

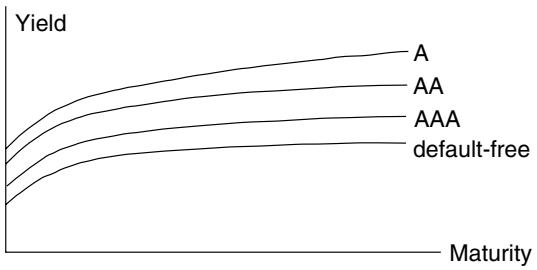

Figure 2.1: Term structures of zero–coupon yields

In connection with interest rate related instruments we thus emphasize on the following types of general interest rate risk:

- yield curve risk, being the risk due to changes of the spot rate curve, i.e. its level and its shape
- credit spread risk, the risk due to changes in the zero–coupon rate of a particular rating class, *relative* to the corresponding spot rate
- yield curve volatility risk, i.e. the risk due to changes in the volatilities of the spot rates.

2.4.1 Yield Curve Risk

The representation of the spot rate curve as a finite set of rates of differing maturities, as used in RiskMetrics[15] or in the BIS proposal[16], is not appropriate

[11] Alternatively, the *forward rate curve* or the *swap rate curve* might be used instead of the spot rate curve since they are all equivalent in the sense that each curve can be constructed from each other.

[12] There exist several methods to construct the spot rate curve from observed bond prices, cf. Vasicek and Fong (1982), pp. 339.

[13] For example, Standard & Poor designate the highest credit quality of corporate securities by AAA, followed by AA, A, BBB, BB, B and CCC.

[14] See Fabozzi (1993), pp. 209.

[15] See RiskMetrics (1996), pp. 107.

[16] See BIS (1996b), pp. 9.

in the Maximum Loss approach. With such a model, the choice of an interest rate scenario might define a curve whose shape is not in accord with empirically observed rates.[17]
Instead of modelling the rates of specific maturities, it is preferable to model changes of the *whole* curve. This can be achieved by using a generic basis, as the one shown in Figure 2.2.

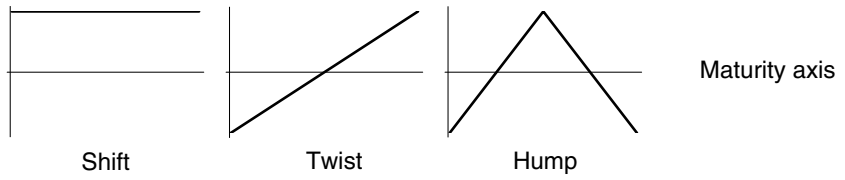

Figure 2.2: Elementary changes in the spot rate curve

Example 2.15 *Figure 2.2 shows a basis which is made up of three different modifications of the spot rate curve:*

- *parallel shifts*
- *twists, i.e. a flattening or steepening of the yield curve*
- *humps, i.e. a change in the curvature of the spot rate curve.*

This way, we have three risk factors $(\omega_1, \omega_2, \omega_3)$*, each one representing a modification of the whole curve. Hence, an interest rate scenario consists of a linear combination of the three basic changes. It defines a new spot rate curve, for which the portfolio can be valued to determine the corresponding P&L* $v(\omega)$*.*

According to Assumption 1.12, each of the yield curve risk factors ω_i is normally distributed with mean zero. If $\omega_i = 0$, this means that there are no deviations from the expected outcome and therefore the corresponding curve coincides with the abscissa.

Remark 2.16 *This framework represents a yield curve model with given initial term structure. As in the HJM approach*[18]*, the drifts have to be chosen appropriately to guarantee an arbitrage–free model, i.e. the expected outcome of the the yield curve at time* T *has to match with the distribution of the yield curve risk factors.*

[17] For example, a scenario where the 3 year and 5 year rates increase, but the 4 year rate decreases, is economically not very meaningful.

[18] See Heath et al. (1992), pp. 77.

Compared to a time bucketing approach, where each time–band has its own risk factor, this model has the advantage that the number of risk factors is small and no mapping of the instruments into a sequence of cash flows[19] is required.

Example 2.17 *Let ω_1 be the risk factor of the parallel shifts, ω_2 of the twists and ω_3 of the humps, as it is shown in Figure 2.2. Then, the scenario $\omega = (\frac{3}{4}, \frac{1}{2}, \frac{1}{4})$ reflects an increase in the medium– and long–term rates, whereas the short–term rates remain almost unaffected, as Figure 2.3 shows.*

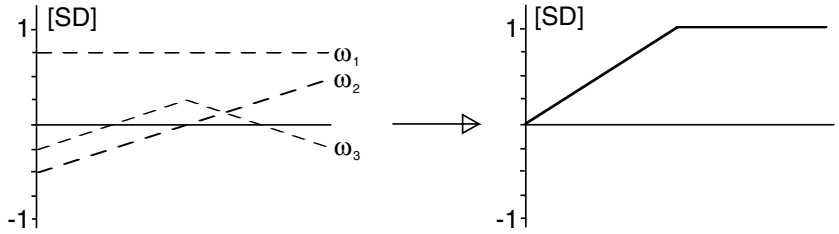

Figure 2.3: Example of interest rate scenario

Rather than choosing an arbitrary basis of yield curve changes (as in Example 2.15), it may be advantageous to use statistical techniques to determine an 'optimal' basis. For example, principal components analysis (PCA)[20] allows to determine an uncorrelated set of curves which minimize the unexplained variability, and may therefore be appropriate for risk measurement purposes. In fact, empirical studies have shown that the first three principal components typically account for 95–99 percent of the total variance.[21]

2.4.2 Credit Spread Risk

To capture the changes in the spreads, we will use risk factors which represent moves of the whole *spread curve*. As before, either an arbitrary basis of elementary modifications might be chosen, or a basis constructed with statistical techniques.

Example 2.18 *A very rough model would have one risk factor for the spread changes of each credit class, e.g. $(\omega_A, \omega_B, \omega_C)$ for the classes A, B and C. Rating specific coefficients $\kappa \geq 0$ would allow to model differing spread proportions for each individual credit rating by taking, for example, changes equal to $\kappa_{AAA}\omega_A, \kappa_{AA}\omega_A$ and ω_A for the credit qualities AAA, AA and A.*

[19] In the RiskMetrics approach and the BIS Standardized Methodology, the instruments need to be mapped in a preliminary step into a sequence of cash flows at prescribed vertices, cf. BIS (1996b), pp. 11 and RiskMetrics (1996), pp. 108.

[20] See Jackson (1991).

[21] See Litterman and Scheinkman (1991), pp. 54.

2.4.3 Yield Curve Volatility Risk

As is the case for the interest rates themselves, also their volatilities have a term structure. Obviously, changes in this term structure can be captured by the same models that have been used in Section 2.4.1 for modelling yield curve risk.

Example 2.19 *If it is assumed that the changes in the volatility of all credit ratings are identical, it is sufficient to have one generic basis of volatility curve modifications (cf. Example 2.15), which is used for all credit classes simultaneously.*

To summarize, risk factors that represent modifications of the entire term structure lead to compact models, which reflect all relevant interest rate risks in a consistent manner.

2.5 Commodity Risk

By commodities we understand physical products, such as agricultural products, precious metals or minerals, which can be traded on a secondary market. These products are typically traded as future transactions (forwards or futures). Consequently, there exists a term structure of commodity prices and the following types of risk have to be distinguished:

- commodity price risk: the risk due to movements in commodity spot prices
- commodity volatility risk: the risk owing to changes in the annualized volatility of a commodity
- commodity forward gap risk: the risk that forward and futures prices[22] may change *for reasons other than changes of the interest rates*.

2.5.1 Commodity Price Risk and Volatility Risk

To model the commodity price risk of commodity i, we use one risk factor ω_i which reflects the changes in the price ψ_i. The changes in the volatility $\psi_{\sigma,i}$ of commodity i can be represented by a risk factor $\omega_{\sigma,i}$.

2.5.2 Commodity Forward Gap Risk

Let us assume that the price F_t of a future with maturity t is of the form

$$F_t = S_0 \exp\left[(r_t + c_t)t\right],$$

[22] Cox et al. (1981) show indeed that forward prices and futures prices are equal provided that interest rates are constant.

where S_0 is the spot price of the commodity, r_t is the risk–free interest rate and c_t the *commodity specific yield*. If the forward price is expressed in this way, it becomes obvious that the changes of the forward price which are neither due to changes of the spot price nor to changes of the interest rate r_t are caused by changes in the commodity specific yield c_t.
Hence, it is natural to use a model which is analogous to the yield curve risk model of Section 2.4.1: modifications of the commodity specific *yield curve* are expressed as a linear combination of elementary curve changes.

Example 2.20 *In a simplistic model, we could have for each commodity a single risk factor which reflects a parallel shift in the commodity specific yield curve. More elaborated models would have a richer basis, such as the one presented in Example 2.15.*
Assume that ω_4 is the risk factor of the parallel shifts in the commodity specific yield curve, ω_5 of the twists and ω_6 of the humps. Then, the commodity specific yield scenario $(\omega_4, \omega_5, \omega_6) = (-\frac{1}{4}, -\frac{1}{4}, 0)$ leads, together with the yield curve scenario $(\omega_1, \omega_2, \omega_3) = (\frac{3}{4}, \frac{1}{2}, \frac{1}{4})$ of Example 2.17, to an overall change in $(r_t + c_t)$ which is shown on the right of Figure 2.4.

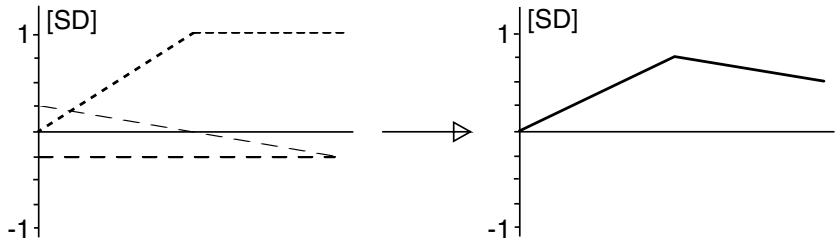

Figure 2.4: Overall effect of combined IR and CO scenarios

Chapter 3

Fundamentals of Maximum Loss

3.1 General Relationship to VaR

Although VaR and ML have apparently different formal definitions[1] and the Maximum Loss approach has, with its trust region $\mathcal{T}$, a supplementary degree of freedom, there exists a strong relationship between the two methods:

Theorem 3.1 *Let $\mathcal{T}$ be an arbitrary feasible trust region (i.e. a closed subset of $\mathcal{O}$ which contains the scenario $\omega = 0$) with $\mathrm{P}(\mathcal{T}) = \alpha$. Then, for every continuous P&L function $v(\omega)$, the relation $\mathrm{ML}(\alpha) \leq \mathrm{VaR}(\alpha)$ holds.*

Proof: Denote by $\mathcal{U}$ the set of active risk factors for VaR

$$\mathcal{U} = \{\omega \in \mathcal{O} \mid v(\omega) \leq \mathrm{VaR}(\alpha)\},$$

which is a closed set since $v(\omega)$ is a continuous function. Two cases have to be distinguished:
If $\mathcal{U} \cap \mathcal{T} \neq \emptyset$, simply choose any scenario $\omega^* \in \mathcal{U} \cap \mathcal{T}$, then obviously $\mathrm{ML}(\alpha) \leq v(\omega^*) \leq \mathrm{VaR}(\alpha)$.
On the other hand, if $\mathcal{U} \cap \mathcal{T} = \emptyset$, then define $\mathcal{W} = (\mathcal{T} \cup \mathcal{U})^C$. Note that

1. $\mathcal{W}$ is an open, nonempty set since $\mathcal{T}$ and $\mathcal{U}$ are both closed
2. $\mathrm{P}(\mathcal{T}) + \mathrm{P}(\mathcal{U}) = \alpha + (1 - \alpha) = 1$ implies that $\mathrm{P}(\mathcal{W}) = 0$, hence $v(\omega) = 0, \forall \omega \in \mathcal{W}$ by Assumption 1.6.

Since the set of realizable scenarios $\mathcal{O} \subseteq \mathbb{R}^M$ is connected by Assumption 1.2 and $\mathcal{U}$ is closed, there exists a sequence $\{\omega^{(1)}, \omega^{(2)}, \dots\} \subset \mathcal{W}$ such that $\lim_{i\to\infty} \omega^{(i)} = \omega^* \in \mathcal{U}$ (in the Euclidean metric). The continuity of the P&L function v implies that $v(\omega^*) = \lim_{i\to\infty} v(\omega^{(i)}) = 0$ and consequently $\mathrm{VaR}(\alpha) \geq 0$. The fact that any feasible trust region $\mathcal{T}$ holds the scenario $\omega = 0$ and $v(0) = 0$

[1] See Definitions 1.14 and 1.39.

then leads to $\mathrm{ML}(\alpha) \leq 0 \leq \mathrm{VaR}(\alpha)$. ∎

Remark 3.2 *It follows from the previous proof that*

$$\max_{\substack{\mathcal{U}\subseteq\mathcal{O}:\mathrm{P}(\mathcal{U})=\alpha \\ \mathcal{U}\text{ closed}}} \min_{\omega\in\mathcal{U}} v(\omega) = \mathrm{VaR}(\alpha).$$

In other words, the maximum ML over all generalized trust regions (in the sense that the scenario $\omega = 0$ needs not to be an element of $\mathcal{U}$, contrarily to Definition 1.39) is equal to the Value–at–Risk for the same confidence level. More precisely,

$$\operatorname*{argmax}_{\substack{\mathcal{U}\subseteq\mathcal{O}:\mathrm{P}(\mathcal{U})=\alpha \\ \mathcal{U}\text{ closed}}} \min_{\omega\in\mathcal{U}} v(\omega) = \{\omega \mid v(\omega) \geq \mathrm{VaR}(\alpha)\},$$

which means that the generalized trust region which produces to the highest ML is equal to the closure of the complement of the set of active risk factors $\mathcal{U}^* = \{\omega \mid v(\omega) \leq \mathrm{VaR}(\alpha)\}$ for VaR.

Thus, Maximum Loss is a risk measure which is always more conservative than VaR, independent of the risk factor distribution and the trust region which have been chosen.

Example 3.3 *Take a portfolio consisting of one linear instrument $v_l(\omega) = -\omega$, whose underlying risk factor $\omega \sim \mathcal{N}(0,1)$. For a confidence level $\alpha = 95\%$, VaR is -1.64. The Maximum Loss however depends heavily on the choice of the trust region $\mathcal{T}$, as Figure 3.1 shows.*

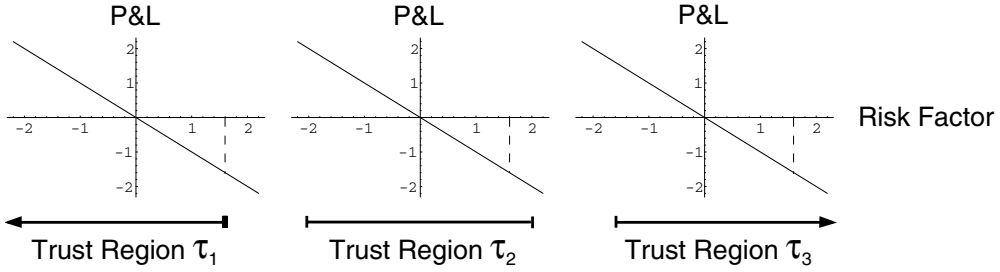

Figure 3.1: Different choices of trust regions

- *For trust region $\mathcal{T}_1 =]-\infty, 1.64]$:* $\mathrm{ML}(\alpha) = -1.64$.
- *For trust region $\mathcal{T}_2 = [-1.96, 1.96]$:* $\mathrm{ML}(\alpha) = -1.96$.
- *For trust region $\mathcal{T}_3 = [-1.64, \infty[$:* $\mathrm{ML}(\alpha) = -\infty$.

In any of these cases, the relation $\mathrm{ML}(\alpha) \leq \mathrm{VaR}(\alpha)$ holds.

Theorem 3.1 has important consequences for risk management: if the portfolio is restructured for the worst case scenario ω^* such that the new P&L $v(\omega^*)$ becomes greater than VaR — without reducing the P&L for any of the other scenarios — then we can be sure that the VaR of the new portfolio will be closer to 0. Consequently, the ML approach can be used to determine VaR reducing transactions.

3.2 Trust Regions

In Section 1.5, it has been said that any closed subset $\mathcal{T} \subseteq \mathcal{O}$ with $0 \in \mathcal{T}$ and probability $\mathrm{P}(\mathcal{T}) = \alpha$ is a feasible trust region. For multinormally distributed risk factors $\omega \sim \mathcal{N}(0, \Sigma)$, there exists a particularly appealing choice of $\mathcal{T}$:

Lemma 3.4 *For multinormally distributed risk factors $\omega \sim \mathcal{N}(0, \Sigma)$, the quantity $\omega^T \Sigma^{-1} \omega$ is χ^2 distributed with M degrees of freedom.*

Proof: Let U be the Cholesky decomposition of $\Sigma = U^T U$, then

$$\mathrm{Cov}(U^{-T}\omega) = \mathrm{E}[(U^{-T}\omega)(U^{-T}\omega)^T] = U^{-T}\mathrm{Cov}(\omega)U^{-1} = \mathrm{I}.$$

Hence, $(U^{-T}\omega) \sim \mathcal{N}(0, \mathrm{I})$ and $\omega^T\Sigma^{-1}\omega = (U^{-T}\omega)^T(U^{-T}\omega)$ is the sum of M squared, independent standard normal variables. ■

Theorem 3.5 *If $\omega \sim \mathcal{N}(0, \Sigma)$, the choice $\mathcal{T} = \{\omega \in \mathbb{R}^M \mid \omega^T\Sigma^{-1}\omega \leq c_\alpha\}$ is a feasible trust region; c_α is the α–quantile of a χ^2 distribution with M degrees of freedom.*

Proof: $\mathcal{T}$ is a closed set which holds the scenario $\omega = 0$ and has probability $\mathrm{P}(\mathcal{T}) = \alpha$. ■

Geometrically, this choice defines an M–dimensional ellipsoid which is centered at the origin. Besides being symmetrical, this trust region has another important property:

Corollary 3.6 *If $\omega \sim \mathcal{N}(0, \Sigma)$, then $\mathcal{T} = \{\omega \in \mathbb{R}^M \mid \omega^T\Sigma^{-1}\omega \leq c_\alpha\}$ is the trust region of minimal volume.*

Proof: Since the risk factors are normally distributed, their density is $f(y) = \vartheta \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right)$, where $\vartheta = \frac{1}{(2\pi)^{M/2}\sqrt{\det\Sigma}} > 0$. The trust region $\mathcal{T}$ is the set

$$\begin{aligned}\{y \in \mathbb{R}^M \mid y^T\Sigma^{-1}y \leq c_\alpha\} &= \left\{y \in \mathbb{R}^M \mid \exp\left(-\frac{1}{2}y^T\Sigma^{-1}y\right) \geq \exp\left(-\frac{1}{2}c_\alpha\right)\right\} \\ &= \left\{y \in \mathbb{R}^M \mid f(y) \geq \vartheta \exp\left(-\frac{1}{2}c_\alpha\right)\right\},\end{aligned}$$

which shows that $\mathcal{T}$ is the set of risk factors with highest density (i.e. $f(y)$ is greater than a constant). Since the density is a nonnegative function, we conclude that for every $\mathcal{U} \subset \mathbb{R}^M$ with $\mathrm{Vol}(\mathcal{U}) = \mathrm{Vol}(\mathcal{T})$ the relation $\mathrm{P}(\mathcal{U}) \leq \mathrm{P}(\mathcal{T}) = \alpha$ holds. Thus, $\mathcal{T}$ is the trust region of minimal volume. ■

These very special characteristics justify to fix the choice of the trust region once and for all:

Assumption 3.7 *In the sequel, we will always take the trust region* $\mathcal{T} = \{\omega \in \mathbb{R}^M \mid \omega^T \Sigma^{-1} \omega \leq c_\alpha\}$, *where* c_α *is the* α*–quantile of a* χ^2 *distribution with* M *degrees of freedom.*

3.2.1 Relative Coordinates

So far, the changes in market rates have always been measured in absolute, i.e. physical units. In some cases, it might be advantageous to express the market moves in units of standard deviations:

Definition 3.8 *Let* $\Sigma = \mathrm{Cov}(\omega)$ *be the covariance matrix of the risk factors. The matrix of standard deviations* V *is the diagonal matrix defined by*

$$V = \mathrm{diag}(\sqrt{(\Sigma)_{1,1}}, \ldots, \sqrt{(\Sigma)_{\mathrm{M,M}}}).$$

Lemma 3.9 *If* $\mathrm{Cov}(\omega) = \Sigma$, *then* $R = V^{-1}\Sigma V^{-1}$ *is the correlation matrix of the risk factors.*

Proof: By definition, the correlation between ω_i and ω_j is

$$\mathrm{Corr}(\omega_i, \omega_j) = \frac{(\Sigma)_{i,j}}{\sqrt{(\Sigma)_{i,j}}\sqrt{(\Sigma)_{i,j}}} = (R)_{i,j},$$

for $1 \leq i, j \leq M$. ■

Definition 3.10 *The standardized risk factors* $\underline{\omega}$ *are defined as* $\underline{\omega} = V^{-1}\omega$.

This way, $\underline{\omega}_i$ expresses market rate changes which are measured in units of the standard deviation of risk factor ω_i. We will show that ML can be calculated in either of the two systems — but, in general, computations in relative coordinates provide more flexibility:

Lemma 3.11 *If* $\omega \sim \mathcal{N}(0, \Sigma)$, *the sets* $\mathcal{U}_1 = \{\omega \in \mathbb{R}^M \mid \omega^T \Sigma^{-1} \omega \leq c_\alpha\}$ *and* $\mathcal{U}_2 = \{\underline{\omega} \in \mathbb{R}^M \mid \underline{\omega}^T R^{-1} \underline{\omega} \leq c_\alpha\}$ *are identical.*

Proof: For each ω we have $\omega^T\Sigma^{-1}\omega = \omega^T V^{-1}R^{-1}V^{-1}\omega = \underline{\omega}^T R^{-1}\underline{\omega}$. ∎

To calculate ML for the standardized risk factors, the objective function has to be adapted:

Definition 3.12 *Take a quadratic portfolio $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$. The standardized P&L function is $\underline{v}_q(\underline{\omega}) = \frac{1}{2}\underline{\omega}^T\underline{G}\underline{\omega} + \underline{g}^T\underline{\omega}$, where $\underline{G} = VGV$ and $\underline{g} = Vg$.*

Theorem 3.13 *ω^* is solution to the ML problem $\mathrm{ML}(\alpha) = \min_{\omega^T\Sigma^{-1}\omega\le c_\alpha} v_q(\omega)$ if and only if $\underline{\omega}^* = V^{-1}\omega^*$ is solution to $\underline{\mathrm{ML}}(\alpha) = \min_{\underline{\omega}^T R^{-1}\underline{\omega}\le c_\alpha} \underline{v}_q(\underline{\omega})$.*

Proof: Lemma 3.11 shows that $\underline{\omega}^*$ is feasible for the problem $\underline{\mathrm{ML}}(\alpha)$ if and only if ω^* is feasible for $\mathrm{ML}(\alpha)$. Moreover, for every $\omega \in \mathbb{R}^M$ the relations $\omega^T G\omega = \omega^T V^{-1}VGVV^{-1}\omega = \underline{\omega}^T\underline{G}\underline{\omega}$ and $g^T\omega = g^T VV^{-1}\omega = \underline{g}^T\underline{\omega}$ hold. ∎

Hence, there is no difference whether the original problem $\mathrm{ML}(\alpha)$ or the standardized problem $\underline{\mathrm{ML}}(\alpha)$ is solved. However, the use of standardized risk factors has the advantage that not all instruments of the portfolio need to have the same holding period T:

Example 3.14 *Suppose that the standard deviation σ_T for a T day holding period follows the square root of time rule of formula (1.5), i.e. $\sigma_T = \sigma\sqrt{T}$. Take two instruments i and j depending on the same risk factor ω_k. If instrument i has a holding period of T_i days and instrument j of T_j days, then an upward move of $\underline{\omega}_k = 2$ standard deviations corresponds to an increase of $\omega_{k(i)} = 2\sigma\sqrt{T_i}$ units for instrument i and of $\omega_{k(j)} = 2\sigma\sqrt{T_j}$ units for instrument j.*

3.3 Linear Portfolios

As for Value–at–Risk, it is possible to obtain an analytic expression of ML for linear portfolios:

Theorem 3.15 *The Maximum Loss of a linear portfolio $v_l(\omega) = a^T\omega$ is $\mathrm{ML}(\alpha) = -\sqrt{c_\alpha}\sqrt{a^T\Sigma a}$ and the worst case scenario is $\omega^* = -\frac{\sqrt{c_\alpha}}{\sqrt{a^T\Sigma a}}\Sigma a$.*

Proof: We have to solve the minimization problem

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad v_l(\omega) \\ & \text{s.t.} \quad h(\omega) \le 0, \end{aligned} \tag{3.1}$$

where $h(\omega) = \omega^T\Sigma^{-1}\omega - c_\alpha$. It is easy to verify that $\omega^* = -\frac{\sqrt{c_\alpha}}{\sqrt{a^T\Sigma a}}\Sigma a$ and $\mu = \frac{\sqrt{a^T\Sigma a}}{2\sqrt{c_\alpha}}$ satisfy the Kuhn–Tucker conditions

$$\begin{aligned} \nabla v_l(\omega^*) &= -\mu \nabla h(\omega^*) \\ \mu h(\omega^*) &= 0 \\ h(\omega^*) &\leq 0 \\ \mu &\geq 0. \end{aligned}$$

Since $v_l(\omega)$ is linear and $h(\omega)$ is convex, this implies that ω^* is the global optimum of problem (3.1). ■

The expression of $\mathrm{ML}(\alpha) = -\sqrt{c_\alpha}\sqrt{a^T \Sigma a}$ is very similar to Delta–Normal Value–at–Risk $\mathrm{VaR}(\alpha) = -z_\alpha \sqrt{a^T \Sigma a}$. The only difference lies in the scaling factor $\sqrt{c_\alpha}$ vs. z_α. This means that ML and VaR have a constant relation, which is independent of the portfolio. However, the ratio depends on the confidence level α and on the number M of risk factors, since c_α is the α–quantile of a χ^2 distribution with M degrees of freedom:

	$M = 2$	$M = 5$	$M = 10$	$M = 50$
$\alpha = 90.0\%$	1.67	2.37	3.12	6.20
$\alpha = 95.0\%$	1.49	2.02	2.60	5.00
$\alpha = 97.5\%$	1.39	1.83	2.31	4.31
$\alpha = 99.0\%$	1.30	1.67	2.07	3.75

Table 3.1: Ratios $r_{\alpha,M}$ of ML/VaR

The fact that the ratios $r_{\alpha,M}$ of ML/VaR depend on the number M of risk factors is due to Definition 1.39 of a trust region $\mathcal{T}$ as a subset of $\mathbb{R}^M$ with $\mathrm{P}(\mathcal{T}) = \alpha$: the more dimensions the scenario space has, the larger $\mathcal{T}$ has to be chosen in order to cover a probability of α.[2]

At this point, it should be noted that the computation of the worst case scenario $\omega^* = -\frac{\sqrt{c_\alpha}}{\sqrt{a^T \Sigma a}} \Sigma a$ requires no other information(except the constant c_α) than the information used in the Delta–Normal VaR computation; the identification of the worst outcome is for free.

[2]In practical applications, this property might be unwanted. In fact, the dependence on the number of risk factors implies that the ML of a model which contains 'empty' risk factors is different from the ML of the same portfolio where the superfluous risk factors have been removed. A way to overcome this problem is to use the coefficients $r_{\alpha,M}$ of Table 3.1 as standardization factors for $\mathrm{ML}^*(\alpha) = \frac{\mathrm{ML}(\alpha)}{r_{\alpha,M}}$. This choice implies in particular that $\mathrm{ML}^*(\alpha) = \mathrm{VaR}(\alpha)$ for linear portfolios.

3.4 Quadratic Portfolios

Whereas VaR can only be computed efficiently for linear portfolios,[3] it is possible to calculate ML for quadratic portfolios $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$ with a fast algorithm, the so–called Levenberg–Marquardt algorithm. In the sequel we show how the algorithm can be used to calculate ML and explain the principles of the optimization procedure.
The Levenberg–Marquardt algorithm allows to calculate numerically the *global* minimum of a quadratic function in a ball. Since the trust region $\mathcal{T}$ has ellipsoidal form, it has first to be transformed into spherical form.

Definition 3.16 *Let $\omega \sim \mathcal{N}(0, \Sigma)$ and denote the Cholesky decomposition of the covariance matrix by $\Sigma = U^T U$. The transformed risk factors are defined as $\hat{\omega} = U^{-T}\omega$.*

Lemma 3.17 *The two sets $\mathcal{U}_1 = \{\omega \in \mathbb{R}^M \mid \omega^T\Sigma^{-1}\omega \leq c_\alpha\}$ and $\mathcal{U}_2 = \{\hat{\omega} \in \mathbb{R}^M \mid \hat{\omega}^T\hat{\omega} \leq c_\alpha\}$ are identical.*

Proof: For every $\omega \in \mathbb{R}^M$ we have $\omega^T\Sigma^{-1}\omega = \omega^T U^{-1}U^{-T}\omega = \hat{\omega}^T\hat{\omega}$. ∎

Hence, the trust region can be transformed by a linear function into a ball. It remains to adapt the objective function:

Definition 3.18 *Let $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$ be a quadratic portfolio. The transformed P&L function is $\hat{v}_q(\hat{\omega}) = \frac{1}{2}\hat{\omega}^T\hat{G}\hat{\omega} + \hat{g}^T\hat{\omega}$, where $\hat{G} = UGU^T$ and $\hat{g} = Ug$.*

Theorem 3.19 *The scenario $\omega^* = U^T\hat{\omega}^*$ is the global solution to the problem $\min_{\omega^T\Sigma^{-1}\omega \leq c_\alpha} v_q(\omega)$ if and only if $\hat{\omega}^*$ is the global solution to $\min_{\omega^T\hat{\omega} \leq c_\alpha} \hat{v}_q(\hat{\omega})$.*

Proof: Lemma 3.17 proves that ω^* is feasible if and only if $\hat{\omega}^*$ is feasible. Furthermore, the two functions are identical since $\hat{\omega}^T\hat{G}\hat{\omega} = \omega^T U^{-1}UGU^T U^{-T}\omega = \omega^T G\omega$ and $\hat{g}^T\hat{\omega} = g^T U^T U^{-T}\omega = g^T\omega$, for every $\omega \in \mathbb{R}^M$. ∎

Consequently, the original quadratic ML problem has been restated as the minimization of a quadratic function $\hat{v}_q(\hat{\omega})$ in a ball:

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad \frac{1}{2}\hat{\omega}^T\hat{G}\hat{\omega} + \hat{g}^T\hat{\omega} \\ & \text{s.t.} \quad \hat{\omega}^T\hat{\omega} \leq c_\alpha. \end{aligned} \tag{3.2}$$

[3] See Section 1.3.1.

3.4.1 The Levenberg–Marquardt Algorithm

The Levenberg–Marquardt algorithm will be the main ingredient of all numerical Maximum Loss applications of Chapter 4 because it allows to determine the global solution to problem (3.2) in an efficient way. The key idea behind this algorithm is to search for a one–dimensional variable ν instead of the M–dimensional scenario ω:

Theorem 3.20 *$\hat{\omega}^*$ is a global solution to (3.2) if and only if there exists $\nu^* \in \mathbb{R}$ such that*

$$(\hat{G} + \nu^* \mathrm{I}) \text{ is positive semidefinite} \tag{3.3}$$

and the following conditions hold:

$$\begin{aligned} (\hat{G} + \nu^* \mathrm{I})\hat{\omega}^* &= -\hat{g} && (3.4)\\ \nu[c_\alpha - (\hat{\omega}^*)^T \hat{\omega}^*] &= 0 && (3.5)\\ \nu^* &\geq 0. && (3.6) \end{aligned}$$

Moreover, if ν^* exists, then it is unique, and if $(\hat{G} + \nu^* \mathrm{I})$ is positive definite, then $\hat{\omega}^*$ is unique.

Proof: The complete proof is given in Fletcher (1987), pp.101. ∎

In the following, it is explained how Theorem 3.20 can be used to solve problem (3.2) numerically.
Suppose for a moment that we knew an orthonormal basis consisting of the eigenvectors of $\hat{G}$, i.e. $\mathcal{B} = \{\hat{e}_1, \ldots, \hat{e}_M\}$ with corresponding eigenvalues $\lambda_1 \leq \ldots \leq \lambda_M$.[4] Then, we could express the vectors $\hat{g}$ and $\hat{\omega}$ in this basis: $\hat{g} = \sum_{i=1}^{M} \alpha_i \hat{e}_i$ and $\hat{\omega} = \sum_{i=1}^{M} \beta_i \hat{e}_i$, where the coefficients β_i are the values we would like to determine. Relation (3.4) implies that $\beta_i = -\frac{\alpha_i}{\lambda_i + \nu^*}$ for $i = 1, \ldots, M$. If we consider $\hat{\omega}$ as a function of ν we get

$$\|\hat{\omega}(\nu)\|^2 = \sum_{i=1}^{M} \left(\frac{\alpha_i}{\lambda_i + \nu} \right)^2. \tag{3.7}$$

For $\nu \geq -\lambda_1$, the function $\|\hat{\omega}(\nu)\|^2$ is decreasing with $\lim_{\nu \to \infty} \|\hat{\omega}(\nu)\|^2 = 0$. Conditions (3.3) and (3.6) imply that $\nu \geq \max(-\lambda_1, 0)$. This phenomenon is represented in Figure 3.2.
Two cases could arise:

[4] An orthonormal basis $\mathcal{B}$ exists because the matrix $\hat{G} \in \mathbb{R}^{M \times M}$ is symmetric.

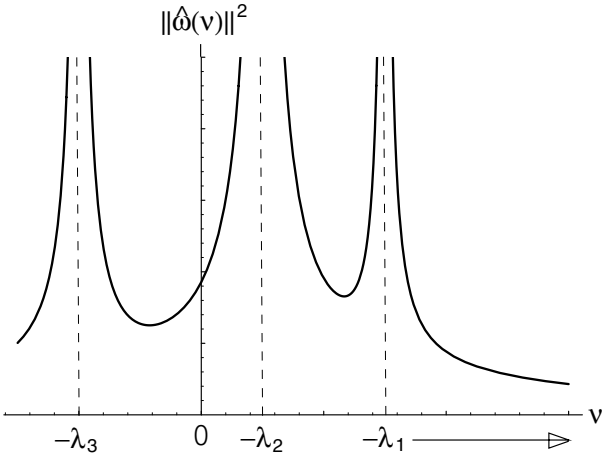

Figure 3.2: Analysis of $\|\hat{\omega}(\nu)\|^2$

1. $\nu^* = 0$: By condition (3.3), $\hat{G}$ is positive semidefinite, and from equation (3.4) we conclude that $\hat{\omega}^* = -\hat{G}^{-1}\hat{g}$ is the solution.

2. $\nu^* > 0$. Let $\nu_1 < \nu_2$ be two candidates for ν^*. Equation (3.4) implies that $\hat{\omega}_i = -(\hat{G} + \nu_i \mathrm{I})^{-1}\hat{g}, i = 1, 2$. From (3.7) we know that $\|\hat{\omega}_1\|^2 > \|\hat{\omega}_2\|^2$. But this means that we can apply a bisectionning method with starting values ν_1, ν_2 such that $\|\hat{\omega}(\nu_1)\|^2 \leq c_\alpha \leq \|\hat{\omega}(\nu_2)\|^2$ to find the optimal value ν^* which satisfies (3.5), i.e. $\|\hat{\omega}(\nu^*)\|^2 = c_\alpha$.

Obviously, neither the eigenvectors $\hat{e}_i$ of $\hat{G}$ nor its eigenvalues λ_i are required to implement such a bisectionning algorithm. Indeed, it is possible to implement the algorithm such that it solves the problem in polynomial time:

Theorem 3.21 *The number of arithmetic operations required to calculate an ϵ–approximation*[5] *of the quadratic ML problem $\min_{\hat{\omega}^T\hat{\omega} \leq c_\alpha} \hat{v}_q(\hat{\omega})$ is bounded by $O(M^3 \log(\frac{1}{\epsilon}))$.*

Proof: A rigorous proof is given in Fu et al. (1996). ∎

Technical details regarding an efficient implementation are described in Dennis and Schnabel (1996).

Remark 3.22 *It is important to note that the problem (3.2) is the only nonconvex global optimization problem for which a polynomial time algorithm is known.*

[5]Let $v_{\min}$ and $v_{\max}$ denote the exact minimum and maximum over the feasible domain. Then, ω^* is an ϵ–approximation to the minimization problem if $\frac{v(\omega^*) - v_{\min}}{v_{\max} - v_{\min}} \leq \epsilon$.

Example 3.23 *If the Levenberg–Marquardt algorithm is applied to the portfolio of Example 1.17, the worst case scenario ω^* for a confidence level of $\alpha = 90\%$ turns out to be*

$$\begin{aligned} \textit{Risk Factor 1} &= -2.796 \\ \textit{Risk Factor 2} &= -1.711 \\ \textit{Risk Factor 3} &= -3.395 \\ \textit{Risk Factor 4} &= -2.200 \\ \textit{Risk Factor 5} &= -2.494 \\ \textit{Risk Factor 6} &= -1.766 \\ \textit{Risk Factor 7} &= -0.844, \end{aligned}$$

and the Maximum Loss is 2786.36 units. If we look at Figure 1.2, we notice that the correlation structure forces risk factors 2, 6 and 7 to move into the profitable region in order to maximize the overall loss of the portfolio.
The worst case scenario ω^* does not only tell us for which changes of market rates the portfolio is most exposed, it is also an important information for the risk manager who has to reduce the overall risk of the portfolio. According to Section 3.1, the relation ML $\leq$ VaR holds for every portfolio. Hence, restructuring the portfolio for the worst case scenario ω^*, such that $v(\omega^*)$ becomes greater than VaR, will definitely reduce VaR.

3.4.2 Sensitivity Analysis

The effect of perturbations in the parameters G, g, Σ and c_α to the solution of the quadratic ML problem

$$\begin{aligned} \mathrm{ML}(\alpha) \quad &= \min \quad \frac{1}{2}\omega^T G\omega + g^T\omega \\ &\text{s.t.} \quad \omega^T\Sigma^{-1}\omega \leq c_\alpha \end{aligned} \tag{3.8}$$

can be estimated without having to resolve the new optimization problem.

Definition 3.24 *Let ω^* be the solution to the quadratic ML problem (3.8), then $\omega^*(\varepsilon)$ denotes the solution to the perturbed problem*

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad &\min \quad \frac{1}{2}\omega^T G\omega + g^T\omega \\ &\text{s.t.} \quad \omega^T\Sigma^{-1}\omega \leq c_\alpha + \varepsilon. \end{aligned} \tag{3.9}$$

Remark 3.25 *From the discussion of the Levenberg–Marquardt algorithm of Section 3.4.1 it follows that the one–dimensional variable ν depends continuously on ε (cf. Figure 3.2). Thus, Theorem 3.20 implies that the worst case scenario ω^* is a continuous function $\omega^*(\varepsilon)$.*

This allows us to state the following theorem:

Theorem 3.26 *The sensitivity of ML to the perturbed problem problem is*

$$\frac{\partial v_q(\omega^*(\varepsilon))}{\partial \varepsilon} = -\mu^*,$$

where the Lagrange multiplier is $\mu^* = \frac{\|G\omega^*+g\|}{2\|\Sigma^{-1}\omega^*\|}$ if $\omega^T\Sigma^{-1}\omega = c_\alpha$ and $\mu^* = 0$ otherwise.

Proof: Let $h(\omega) = \omega^T\Sigma^{-1}\omega - c_\alpha$ be the trust region constraint and $\mathcal{L}(\omega, \mu) = v(\omega) + \mu h(\omega)$ the Lagrangian function of problem (3.9). It is easy to verify that μ^* satisfies the Kuhn–Tucker conditions

$$\begin{aligned} \nabla_\omega \mathcal{L}(\omega^*, \mu^*) &= 0 \\ \mu^* h(\omega^*) &= 0 \\ h(\omega^*) &\leq 0 \\ \mu^* &\geq 0. \end{aligned}$$

If $\mu^*(\varepsilon)$ is the Lagrange multiplier of the perturbed problem, the second Kuhn–Tucker condition $\mu^* h(\omega^*) = 0$ implies that $v_q(\omega^*(\varepsilon)) = \mathcal{L}(\omega^*(\varepsilon), \mu^*(\varepsilon))$. It follows that

$$\frac{\partial v_q(\omega^*(\varepsilon))}{\partial \varepsilon} = \frac{\partial \mathcal{L}(\omega^*(\varepsilon), \mu^*(\varepsilon))}{\partial \varepsilon} = -\mu^*.$$

■

Theorem 3.26 allows to analyze the effect of small changes in Σ and c_α on the solution of the ML problem:

Corollary 3.27 *If the trust region is expanded by Δc_α to become $\mathcal{T} = \{\omega \mid \omega^T\Sigma^{-1}\omega \leq c_\alpha + \Delta c_\alpha\}$, then the quadratic ML is increased by $\Delta v_q(\omega^*) = -\mu^* \Delta c_\alpha$.*

Proof: It is sufficient to set $\varepsilon = \Delta c_\alpha$ in Theorem 3.26. ■

We will make use of this result later in Section 4.1.1 when we calculate ML for a series of expanding trust regions: Corollary 3.27 provides an efficient way to generate good interpolations.

Corollary 3.28 *If the covariance matrix Σ changes to $\tilde{\Sigma}$, the effect on ML is* $\Delta v_q(\omega^*) = \mu^*(\omega^*)^T[\tilde{\Sigma}^{-1} - \Sigma^{-1}]\omega^*$.

Proof: If the covariance matrix becomes $\tilde{\Sigma}$, the constraint function changes to $h(\omega) = \omega^T\tilde{\Sigma}^{-1}\omega - c_\alpha$. This corresponds to a perturbation of $\varepsilon = (\omega^*)^T(\Sigma^{-1} - \tilde{\Sigma}^{-1})\omega^*$ in Theorem 3.26. ∎

The effect of changes in the objective function $v_q(\omega)$ can be estimated by means of ordinary calculus:

Lemma 3.29 *If the quadratic function $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$ becomes $v_q'(\omega) = \frac{1}{2}\omega^T(G+\Delta G)\omega+(g+\Delta g)^T\omega$, then ML is augmented by $\Delta v_q(\omega^*) = \frac{1}{2}(\omega^*)^T\Delta G\omega^* + \Delta g^T\omega^*$.*

Proof: Differentiating the objective function $v_q(\omega) = \frac{1}{2}\omega^T G\omega+g^T\omega$ with respect to the coefficients gives $\frac{\partial v_q}{\partial(G)_{i,j}} = \frac{1}{2}\omega_i\omega_j$ and $\frac{\partial v_q}{\partial g_i} = \omega_i$. Since the global optimum is located at ω^*, the result follows immediately. ∎

This result can be used to update quickly ML after new instruments have been added to the portfolio. It allows also to examine the effect of candidate trades and to quickly check whether the risk remains within its limits.

Example 3.30 *If we have a linear portfolio $v_l(\omega) = a^T\omega$, we know from Theorem 3.15 that the worst case scenario is $\omega^* = -\frac{\sqrt{c_\alpha}}{\sqrt{a^T\Sigma a}}\Sigma a$. In accordance to Lemma 3.29, a new trade with risk factor sensitivities Δa affects ML by $\Delta a^T\omega^* = -\sqrt{c_\alpha}\frac{\Delta a^T\Sigma a}{\sqrt{a^T\Sigma a}}$.*
In Section 3.3 it was shown that for linear portfolios the relation $\mathrm{VaR}(\alpha) = \frac{z_\alpha}{\sqrt{c_\alpha}}\mathrm{ML}(\alpha)$ *holds. Hence, the new trade changes Value–at–Risk by*

$$\Delta\mathrm{VaR}(\alpha) = -z_\alpha\frac{\Delta a^T\Sigma a}{\sqrt{a^T\Sigma a}},$$

which coincides with the first order approximation of

$$-z_\alpha[\sqrt{(a+\Delta a)^T\Sigma(a+\Delta a)} - \sqrt{a^T\Sigma a}],$$

which is the difference of the analytic Delta–Normal VaRs of Theorem 1.16.[6]

3.5 Correlation Risk

What is the effect on Maximum Loss if there is a complete breakdown in the correlation structure? The following analysis can be understood as a 'correlation

[6]This result, although obtained from a sensitivity analysis of Maximum Loss, coincides exactly with the VaR-Delta formula of Garman (1997).

stress test'. In Section 3.2.1 it has been shown that the covariance matrix Σ can be decomposed into $\Sigma = VRV$, where V is the diagonal matrix of the standard deviations and R is the correlation matrix. The set of all correlations matrices $R \in \mathbb{R}^{M\times M}$ is denoted by $\mathcal{C}$.

Theorem 3.31 *Let $\omega \sim \mathcal{N}(0,\Sigma)$ and V the matrix of standard deviations of Definition 3.8. Then the sets $\mathcal{U}_1 = \{\omega \in \mathbb{R}^M \mid \exists R \in \mathcal{C} : \omega^T V^{-1}R^{-1}V^{-1}\omega \leq c_\alpha\}$ and $\mathcal{U}_2 = \{\omega \in \mathbb{R}^M \mid \ -\sqrt{c_\alpha}(V)_{i,i} \leq \omega_i \leq \sqrt{c_\alpha}(V)_{i,i}, i = 1,\dots,M\}$ are identical.*

Proof: First, for every fixed correlation matrix R the resulting ellipsoid

$$\mathcal{T}_R = \{\omega \in \mathbb{R}^M \mid \omega^T V^{-1}R^{-1}V^{-1}\omega \leq c_\alpha\}$$

is contained in the box $\mathcal{U}_2$: according to Theorem 3.15, the solution of the optimization problem $\min_{\omega^T V^{-1}R^{-1}V^{-1}\omega\leq c_\alpha} \omega_i$ is $\omega_i^* = -\sqrt{c_\alpha}\sqrt{(\Sigma)_{i,i}}, i = 1,\dots,M$. Hence, the interval $[-\sqrt{c_\alpha}(V)_{i,i}, \sqrt{c_\alpha}(V)_{i,i}]$ is the range in which risk factor ω_i takes its feasible values and consequently $\mathcal{T}_R \subseteq U_2$.
On the other hand, let ω be an arbitrary scenario of $\mathcal{U}_2$. Then, there exists a correlation matrix $R \in \mathcal{C}$ for which $\omega \in \mathcal{T}_R$: denote by $\tilde{\omega}$ the intersection of the vector ω with the box $\mathcal{U}_2$, i.e. $\tilde{\omega} = \tilde{\lambda}\omega$, where $\tilde{\lambda} = \max\{\lambda \geq 1 \mid \lambda\omega \in \mathcal{U}_2\}$. Let i be an active box constraint (i.e. $\tilde{\omega}_i = \pm\sqrt{c_\alpha}(V)_{i,i}$). Take any correlation matrix R whose ith row/column is equal to $(R)_i = -\frac{V^{-1}\tilde{\omega}}{\sqrt{c_\alpha}}$. This is a valid row of a correlation matrix and it is easy to verify that $\tilde{\omega}$ is a solution to the problem $\min_{\omega^T V^{-1}R^{-1}V^{-1}\omega\leq c_\alpha} \omega_i$. Hence, $\tilde{\omega} \in \mathcal{U}_1$ and therefore $\omega \in \mathcal{U}_1$. ∎

Ergo, the size of the box which contains all feasible scenarios depends only on the standard deviations of the risk factors and not on their correlations. Thus, if the correlations are allowed to vary freely, Maximum Loss is

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad v(\omega) \\ & \text{s.t.} \quad -\sqrt{c_\alpha}(V)_{i,i} \leq \omega_i \leq \sqrt{c_\alpha}(V)_{i,i}, \quad i = 1,\dots,M. \end{aligned} \tag{3.10}$$

It will be shown in Section 3.7 that, contrarily to the ellipsoidal ML problem $\min_{\omega^T\Sigma^{-1}\omega\leq c_\alpha} v_q(\omega)$, this box constrained quadratic programming problem is extremely difficult to solve.

3.6 Foreign Exchange: Consistency Restrictions

In Section 2.2 it has been shown that in order to model currency risk consistently, only risk factors related to *foreign exchange rates* and their volatilities have to be used. More precisely, Lemma 2.3 and Corollary 2.5 indicate how to derive the

value of the cross rate $\psi_{\frac{A}{B}}$ and its volatility $\psi_{\sigma,\frac{A}{B}}$ from the foreign exchange rates $\psi_{\frac{H}{A}}, \psi_{\frac{H}{B}}$ and their volatilities $\psi_{\sigma,\frac{H}{A}}, \psi_{\sigma,\frac{H}{B}}$.

Theorem 2.4 shows that if the relative returns of the foreign exchange rates $\frac{dX_t^{\frac{H}{A}}}{X_t^{\frac{H}{A}}}$ and $\frac{dX_t^{\frac{H}{B}}}{X_t^{\frac{H}{B}}}$ are normally distributed, then also $\frac{dX_t^{\frac{A}{B}}}{X_t^{\frac{A}{B}}} \sim \mathcal{N}(\mu_{\frac{A}{B}}dt, \sigma^2_{\frac{A}{B}}dt)$, where $\sigma^2_{\frac{A}{B}}$ can be calculated from the variances and the correlation of the foreign exchange rates. If we consider a holding period of length T we get

$$\mathrm{Var}\left(\frac{\psi_{\frac{A}{B}} - \psi^0_{\frac{A}{B}}}{\psi^0_{\frac{A}{B}}}\right) = \sigma^2_{\frac{A}{B}}T,$$

where $\psi^0_{\frac{A}{B}}$ is the value of the cross rate at the beginning of the period and $\psi_{\frac{A}{B}}$ the value at the end of the period. Since $\psi^0_{\frac{A}{B}}$ is a constant, it follows that

$$\mathrm{Var}(\psi_{\frac{A}{B}}) = (\psi^0_{\frac{A}{B}}\sigma_{\frac{A}{B}})^2T.$$

In the model developed in Section 3.2, the trust region $\mathcal{T}$ is an ellipsoid which restricts the moves of the risk factors $\omega_{\frac{H}{A}}, \omega_{\frac{H}{B}}, \omega_{\sigma,\frac{H}{A}}$ and $\omega_{\sigma,\frac{H}{B}}$. If we use the no–arbitrage relation $\bar{\psi}_{\frac{A}{B}} = \frac{\bar{\psi}_{\frac{H}{B}}}{\bar{\psi}_{\frac{H}{A}}}$, we can limit the moves in the cross rate $\psi_{\frac{A}{B}}$ by additional confidence bounds:

Theorem 3.32 *If, at the end of the holding period, the value of the cross rate is* $\psi_{\frac{A}{B}} \sim \mathcal{N}(\bar{\psi}_{\frac{A}{B}}, \hat{\sigma}^2_{\frac{A}{B}})$ *with* $\bar{\psi}_{\frac{A}{B}} = \frac{\bar{\psi}_{\frac{H}{B}}}{\bar{\psi}_{\frac{H}{A}}}$*, then the linear restrictions*

$$\begin{aligned}\left(-\frac{\bar{\psi}_{\frac{H}{B}}}{\bar{\psi}_{\frac{H}{A}}} - z_{\frac{1+\alpha}{2}}\hat{\sigma}_{\frac{A}{B}}\right)\omega_{\frac{H}{A}} + \omega_{\frac{H}{B}} &\leq \bar{\psi}_{\frac{H}{A}} z_{\frac{1+\alpha}{2}}\hat{\sigma}_{\frac{A}{B}} \\ \left(\frac{\bar{\psi}_{\frac{H}{B}}}{\bar{\psi}_{\frac{H}{A}}} - z_{\frac{1+\alpha}{2}}\hat{\sigma}_{\frac{A}{B}}\right)\omega_{\frac{H}{A}} - \omega_{\frac{H}{B}} &\leq \bar{\psi}_{\frac{H}{A}} z_{\frac{1+\alpha}{2}}\hat{\sigma}_{\frac{A}{B}},\end{aligned}$$

define a confidence region of level α *for the cross rate* $\frac{A}{B}$*.* $z_{\frac{1+\alpha}{2}}$ *is the* $(\frac{1+\alpha}{2})$*–quantile of the standard normal distribution.*

Proof: The deviation of the cross rate from its expected value is

$$\psi_{\frac{A}{B}} - \bar{\psi}_{\frac{A}{B}} = \frac{\omega_{\frac{H}{B}} + \bar{\psi}_{\frac{H}{B}}}{\omega_{\frac{H}{A}} + \bar{\psi}_{\frac{H}{A}}} - \frac{\bar{\psi}_{\frac{H}{B}}}{\bar{\psi}_{\frac{H}{A}}},$$

simply because $\psi_i = \omega_i + \bar{\psi}_i$ by Definition 1.1. Since the quantity $(\psi_{\frac{A}{B}} - \bar{\psi}_{\frac{A}{B}})$ is normally distributed with mean 0 and variance $\hat{\sigma}^2_{\frac{A}{B}}$, it follows that

$$\left| \frac{\omega_{\frac{H}{B}} + \bar{\psi}_{\frac{H}{B}}}{\omega_{\frac{H}{A}} + \bar{\psi}_{\frac{H}{A}}} - \frac{\bar{\psi}_{\frac{H}{B}}}{\bar{\psi}_{\frac{H}{A}}} \right| \leq z_{\frac{1+\alpha}{2}} \hat{\sigma}_{\frac{A}{B}}$$

is a confidence region of level α for $(\psi_{\frac{A}{B}} - \bar{\psi}_{\frac{A}{B}})$. The two linear restrictions are obtained by considering positive and the negative values separately. ■

This way, every combination of two foreign exchange rates gives rise to two additional linear consistency restrictions (cf. Figure 3.3).

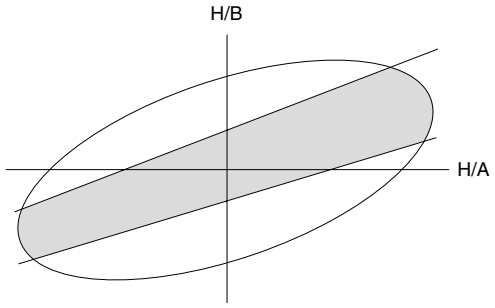

Figure 3.3: Effect of additional FX–restrictions

Consequently, if there are N foreign exchange rates in the model, $N(N-1)$ linear consistency restrictions can be added. Assuming a quadratic P&L function $v_q(\omega)$, the resulting ML problem has the following form:

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad v_q(\omega) \\ & \text{s.t.} \quad \omega^T \Sigma^{-1} \omega \leq c_\alpha \\ & \qquad A\omega \leq b, \end{aligned}$$

where $A\omega \leq b$ stands for the additional consistency restrictions. If we replace the ellipsoidal trust region restriction $\omega^T \Sigma^{-1} \omega \leq c_\alpha$ by lower and upper bounds[7] $-\sqrt{c_\alpha}(V)_{i,i} \leq \omega_i \leq \sqrt{c_\alpha}(V)_{i,i}$, we get a polytope constrained quadratic programming problem, for which even the calculation of an approximate solution is a difficult task, as will be shown in Section 3.7.

Hence, the introduction of additional consistency restrictions leads to less conservative and more realistic ML figures; however, the computational effort for their calculation increases considerably.

[7] See Section 3.5.

3.7 Complexity Issues

Vavasis (1991) proves that it is possible to calculate in polynomial time an ϵ–approximation to the ball constrained quadratic problem $\min_{\omega^T\omega\leq 1} \frac{1}{2}\omega^T G\omega + g^T\omega$,[8] which is the core problem of the Maximum Loss computations. In this section, we will show that this result is exceptional in the sense that quadratic problems with linearly restricted feasible domains are hard to solve. Linear constraints occur when introducing additional FX consistency–restrictions (cf. Section 3.6) or when determining an upper bound for the correlation risk by solving box constrained quadratic programming problems (cf. Section 3.5).

Definition 3.33 *Let $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$ be a quadratic P&L function. For $l_1 \leq l_2 \in \mathbb{R}^M$, problems of the type*

$$\begin{aligned} \text{(BQP)}: \quad & \min \quad v_q(\omega) \\ & \text{s.t.} \quad l_1 \leq \omega \leq l_2 \end{aligned}$$

are called box constrained quadratic programming problems (BQP).

The box constrained problems (BQP) are a subset of the general quadratic programming problems:

Definition 3.34 *Let $v_q(\omega)$ be a quadratic P&L function, $A \in \mathbb{R}^{N\times M}$ and $b \in \mathbb{R}^N$. The general quadratic programming problem (QP) is*

$$\begin{aligned} \text{(QP)}: \quad & \min \quad v_q(\omega) \\ & \text{s.t.} \quad A\omega \leq b. \end{aligned}$$

Both problems (QP) and (BQP) are known to be difficult to solve. However, before complexity issues can be properly addressed, it must be ensured that the problems have a finite encoding.

Assumption 3.35 *In the remaining part of Section 3.7, we assume that all data which specify a particular instance of a problem are rational numbers.*[9]

First of all, we will investigate the complexity of decision problems: 'Given a rational number ζ, does there exist $\omega \in \mathbb{R}^M$ satisfying $v_q(\omega) \leq \zeta, A\omega \leq b$?' The complexity class NP consists of the decision problems whose solutions can be verified efficiently. More precisely, a problem belongs to NP if for each of its

[8] See also Theorem 3.21.

[9] The quadratic programming problems are described by $G \in \mathbb{Q}^{M\times M}$ and $g \in \mathbb{Q}^M$, which define the objective function $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$, and $A \in \mathbb{Q}^{N\times M}, b \in \mathbb{Q}^N$ for (QP), respectively $l_1, l_2 \in \mathbb{Q}^M$ for (BQP), which specify the feasible regions.

yes–instances there exists a certificate which can be verified in polynomial time (i.e. $\exists \omega^*$ of polynomial length which satisfies $v_q(\omega^*) \leq \zeta$ and $A\omega^* \leq b$).
On the positive side, the decision problems (QP) and (BQP) do not lie in a more difficult complexity class than any other problem in NP:

Theorem 3.36 *The decision version of the quadratic programming problem (QP) is NP–complete.*

Proof: The proof that (QP) belongs to the class NP is given in Vavasis (1991), pp. 76. To show that each problem of NP can be polynomially transformed into an instance of (QP), we transform the satisfiability problem SAT[10], which is known to be NP–complete, into (QP): each boolean variable x_i is replaced by a variable ω_i and its negation by $(1 - \omega_i)$, the natural bounds being $0 \leq \omega_i \leq 1$. Then, each clause is transformed into a linear constraint, e.g. $(x_1 \vee x_2 \vee \neg x_3)$ becomes $\omega_1 + \omega_2 + (1 - \omega_3) \geq 1$. Finally, the objective function is constructed in such a way that each ω_i is forced to become either 0 or 1:

$$v_q(\omega) = \sum_{i=1}^{M} \omega_i(1 - \omega_i).$$

The result follows by observing that a yes–instance of SAT corresponds to an instance of (QP) with global minimum 0. ∎

Corollary 3.37 *The decision version of the box constrained quadratic programming problem (BQP) is NP–complete.*

Proof: This proof is based on the proof of Theorem 3.36 in the sense that an instance of (QP) is reduced to one of (BQP): each individual constraint $a_j^T \omega \geq b_j$ of (QP) is written as $a_j^T \omega = b_j + y_j$, where y_i is a new decision variable, for which lower and upper bounds can be specified. Then, each condition $a_j^T \omega = b_j + y_j$ is replaced by an additional term $(a_j^T \omega - b_j - y_j)^2$ in the objective function, which is nonnegative on the feasible region. It is easy to see that an objective value of 0 corresponds to a yes–instance of SAT and vice versa. ∎

Let us turn away from decision problems and look at the proper *optimization problems* (QP) and (BQP). A problem H is said to be NP–hard if any NP–complete problem C can be *solved* in polynomial time, provided that it can use an oracle which computes the solution of H in polynomial time. If a polynomial time algorithm for H were available, this would imply that P=NP. Obviously, both problems (QP) and (BQP) are NP–hard:

[10] A satisfiability problem consists of a set of boolean variables $\{x_1, \ldots, x_n\}$ and a set of clauses $S_1, \ldots, S_m$. A clause is an OR composition of variables and their negations, e.g. $S_1 = (x_1 \vee \neg x_3)$. An assignment $(x_1^*, \ldots, x_n^*)$ is feasible, if all clauses are satisfied simultaneously.

Theorem 3.38 *The general quadratic programming problem (QP) is NP–hard.*

Proof: Take an instance of SAT and apply the transformation used in the proof of Theorem 3.36. ∎

Corollary 3.39 *The box constrained quadratic Programming problem (BQP) is NP–hard.*

Proof: See the proof of Corollary 3.37. ∎

In fact, it is even impossible to compute efficiently an *approximate* solution as will be shown in Theorem 3.41. To understand the difficulties that can arise when solving the problems (QP) and (BQP), we will assume that the feasible region is bounded:

Definition 3.40 *Let $\omega \in \mathbb{R}^M$ and $v_q(\omega)$ be a quadratic P&L function and let $D = \{\omega \in \mathbb{R}^M \mid A\omega \leq b\}$ be a polytope, where $A \in \mathbb{R}^{N\times M}$ and $b \in \mathbb{R}^N$. The polytope constrained quadratic programming problem (PQP) is defined as*

$$\begin{aligned} (\text{PQP}): \quad & \min \quad v_q(\omega) \\ & \text{s.t.} \quad \omega \in D. \end{aligned}$$

When a polytope constrained quadratic programming problem $\min_{\omega\in D} \frac{1}{2}\omega^T G\omega + g^T\omega$ is solved, one of three cases can arise:

- if G is positive semidefinite, the optimum ω^* may lie in the interior of the feasible region
- if G is negative definite, the optimum ω^* is attained at a vertex
- if G is indefinite, the optimum ω^* lies at the boundary of the polytope.

Horst et al. (1995) propose an active set method for solving the problem (PQP). In this method, each of the 2^N possibly active sets[11] are examined, where N is the number of the linear constraints. Therefore, this *exact* algorithm has a rather poor performance.
Most discouraging is the fact that even the calculation of an ϵ–approximation of (PQP) is an intractable problem:

Theorem 3.41 *There exists a constant $\epsilon \in (0, 1)$ such that finding an ϵ–approximation to the polytope constrained quadratic programming problem (PQP) is NP–hard.*

[11] An active set is a subset of inequalities which are simultaneously satisfied as equalities, while the remaining constraints are disregarded.

Proof: A proof is presented in Bellare and Rogaway (1995).[12] ■

Hence, we can neither expect to find a 'good' algorithm to execute the correlation stress test of Section 3.5, nor an efficient method which would allow to consider the linear constraints induced by the additional FX restrictions of Section 3.6.

[12]Bellare and Rogaway (1995) show that the polytope constrained quadratic programming problem (PQP) is NP–hard as soon as $\epsilon \leq \frac{1}{3}$. Indeed, the best known polynomial time algorithm for the box constrained quadratic programming problem (BQP) calculates $\frac{4}{7}$ approximate solutions (cf. Ye (1997)).

Chapter 4

Advanced Applications

4.1 Portfolio Characterization

4.1.1 Maximum Loss Path

Repetitive calculations of Maximum Loss give insights which go far beyond a simple worst case identification. The Levenberg–Marquardt algorithm, which has been presented in Section 3.4.1, allows to determine the value of the maximal loss as well as to identify the worst case scenario for some given confidence level α. If this calculation is repeated for a sequence of increasing confidence levels α, a list of MLs and scenarios is obtained.

Example 4.1 *Table 4.1 shows the results of such a repetitive calculation for the test portfolio of Example 1.17.*

α	ML	Risk Factor 1	Risk Factor 2	...	Risk Factor 7
⋮	⋮	⋮	⋮	...	⋮
88 %	-2688.49	-2.727	-1.671	...	-0.825
89 %	-2735.55	-2.761	-1.690	...	-0.834
90 %	-2786.36	-2.796	-1.711	...	-0.844
91 %	-2841.72	-2.834	-1.733	...	-0.855
⋮	⋮	⋮	⋮	...	⋮

Table 4.1: Repetitive calculation of ML for test portfolio

The geometric interpretation of this procedure is the following: increasing α from 0 to some upper limit means expanding the trust region $\mathcal{T}$ from a single point, which represents the expected state $\bar{\psi}$ of the world at the end of the holding period, to the final ellipsoid. The sequence of the resulting ML scenarios defines a path which starts at the origin (i.e. $\omega = 0$) and follows the worst possible route.

By Remark 3.25, the worst case scenario ω^* depends continuously on the size c_α of the trust region. To obtain path which is smooth, the minimization problem has to be solved for many levels of α. However, the number of minimizations can be reduced by using the sensitivity analysis presented in Corollary 3.27 together with some interpolation scheme.

Remark 4.2 *The techniques which have been discussed in Chapter 3 can, of course, also be applied to the profit side of the portfolio. Indeed, the Maximum Profit (MP) problem* $\max_{\omega\in\mathcal{T}} v(\omega)$ *can be solved by inverting the sign of the objective function and calculating* $-\min_{\omega\in\mathcal{T}} -v(\omega)$.

Example 4.3 *Figure 4.1 shows the ML and MP paths for the test portfolio of Example 1.17.*

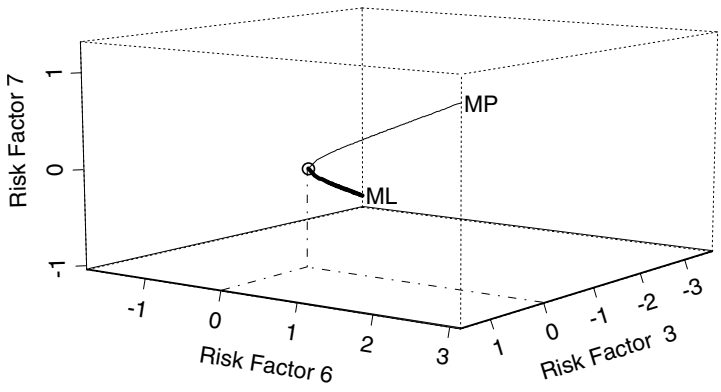

Figure 4.1: ML and MP paths of test portfolio

If the matrix G of the quadratic ML problem $\min_{\omega^T\Sigma^{-1}\omega\le c_\alpha} \frac{1}{2}\omega^T G\omega + g^T\omega$ is positive definite (i.e. all curvatures are positive), then the unconstrained optimization problem $\min v_q(\omega)$ has a *strict* global minimum ω^*. Consequently, the ML path will stop at ω^* as soon as this point lies in the interior of the trust region $\mathcal{T}$. It might be interesting, however, to see how the path would evolve if we restricted the solution to lie *on the surface* of the expanding trust region. This is equivalent to solve the problem

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad \frac{1}{2}\omega^T G\omega + g^T\omega \\ & \text{s.t.} \quad \omega^T\Sigma^{-1}\omega = c_\alpha, \end{aligned}$$

where the inequality sign has been replaced by an equality sign. Before we can solve this problem, the ellipsoid has first to be transformed into a sphere. This

can be achieved by the linear transformation of Definition 3.16. The resulting problem is

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad \frac{1}{2}\hat{\omega}^T \hat{G}\hat{\omega} + \hat{g}^T\hat{\omega} \\ & \text{s.t.} \quad \hat{\omega}^T\hat{\omega} = c_\alpha. \end{aligned} \tag{4.1}$$

The following theorem gives a hint about how to solve problem (4.1) with the already known Levenberg–Marquardt algorithm:

Theorem 4.4 *Let $\kappa > \max(\lambda_1, 0)$, where λ_1 is the lowest eigenvalue of $\hat{G}$. The scenario $\hat{\omega}^*$ is a solution to*

$$\begin{aligned} \min \quad & \frac{1}{2}\hat{\omega}^T(\hat{G} - \kappa\mathrm{I})\hat{\omega} + \hat{g}^T\hat{\omega} \\ \text{s.t.} \quad & \hat{\omega}^T\hat{\omega} = c_\alpha \end{aligned}$$

if and only if $\hat{\omega}^*$ is a solution to

$$\begin{aligned} \min \quad & \frac{1}{2}\hat{\omega}^T(\hat{G} - \kappa\mathrm{I})\hat{\omega} + \hat{g}^T\hat{\omega} \\ \text{s.t.} \quad & \hat{\omega}^T\hat{\omega} \leq c_\alpha. \end{aligned} \tag{4.2}$$

Proof: The matrix $(\hat{G} - \kappa\mathrm{I})$ is not positive semidefinite. Theorem 3.20 characterizes solutions to the problem (4.2): condition (3.3) requires that $(\hat{G} - \kappa\mathrm{I} + \nu^*\mathrm{I})$ has to be positive semidefinite, hence $\nu^* \geq \kappa > 0$ must hold. Consequently, we need that $\|\hat{\omega}^*\|^2 = c_\alpha$ in order to fulfil condition (3.5), which says that $\nu[c_\alpha - (\hat{\omega}^*)^T\hat{\omega}^*] = 0$ has to be satisfied. Since ν^* is unique by Theorem 3.20, it follows that $\hat{\omega}^*$ is a solution to both problems. ∎

Corollary 4.5 *Let $\kappa > \max(\lambda_1, 0)$, where λ_1 is the lowest eigenvalue of $\hat{G}$. If $\hat{\omega}^*$ is solution to the problem $\min_{\hat{\omega}^T\hat{\omega} \leq c_\alpha} \frac{1}{2}\hat{\omega}^T(\hat{G} - \kappa\mathrm{I})\hat{\omega} + \hat{g}^T\hat{\omega}$, then it is also solution to $\min_{\hat{\omega}^T\hat{\omega} = c_\alpha} \frac{1}{2}\hat{\omega}^T\hat{G}\hat{\omega} + \hat{g}^T\hat{\omega}$.*

Proof: In the proof of the last theorem we have shown that $\|\hat{\omega}^*\|^2 = c_\alpha$. Hence,

$$\frac{1}{2}(\hat{\omega}^*)^T(\hat{G} - \kappa\mathrm{I})\hat{\omega}^* + \hat{g}^T\hat{\omega}^* = \frac{1}{2}(\hat{\omega}^*)^T\hat{G}\hat{\omega}^* + \hat{g}^T\hat{\omega}^* - \frac{1}{2}\kappa c_\alpha,$$

which shows that the objective function of the two problems differ only by the constant $\frac{1}{2}\kappa c_\alpha$. ∎

Therefore, the worst case scenario ω^* of the original problem (4.1) is identical to the solution of problem (4.2). However, before problem (4.2) can be solved with the Levenberg–Marquardt algorithm, a feasible value for κ is needed. Such a value can be gained by estimating a lower bound of λ_1, for example, using Gerschgorin circles[1].

4.1.2 Expected Profit and Loss

The knowledge of only Maximum Loss and Maximum Profit is not sufficient to judge about the quality of a portfolio, as the following example shows:

Example 4.6 *Suppose that we have a portfolio which depends on the risk factor* $\omega = (\omega_1, \ldots, \omega_{10})$ *with* $\omega \sim \mathcal{N}(0, \mathrm{I})$. *If the P&L function is* $v(\omega) = \sum_{i=1}^{9} \omega_i^2 - 2\omega_{10}^2$, *then we have* $\mathrm{MP}(\alpha) = c_\alpha$ *and* $\mathrm{ML}(\alpha) = -2c_\alpha$. *The analysis of the extrema gives the impression that the portfolio is very risky, although the chances to make a profit are much higher than the chance to be hit by a loss.*

What is missing, is some information about the average P&L. Therefore, we will calculate the conditional expected value (EV) of the portfolio, given that we are *on the surface* of the trust region.

Theorem 4.7 *Let* $\hat{\omega} \sim \mathcal{N}(0, \mathrm{I})$ *and* $\hat{v}_q(\hat{\omega}) = \frac{1}{2}\hat{\omega}^T \hat{G} \hat{\omega} + \hat{g}^T \hat{\omega}$. *Then* $\mathrm{E}(\hat{v}_q(\hat{\omega}) \mid \hat{\omega}^T \hat{\omega} = c_\alpha) = \frac{c_\alpha}{2} \frac{\mathrm{Tr}(\hat{G})}{M}$, *where* $\mathrm{Tr}(\hat{G})$ *denotes the trace of matrix* $\hat{G}$.

Proof: By abuse of notation, the density of risk factor $\hat{\omega}$ is

$$f(\hat{\omega}) = (2\pi)^{-M/2} \exp(-\frac{1}{2} \sum_{i=1}^{M} \hat{\omega}_i^2).$$

Consequently, the conditional density $f(\hat{\omega} \mid \hat{\omega}^T \hat{\omega} = c_\alpha)$ is constant. The quadratic function $\hat{v}_q(\hat{\omega})$ is made up of three different kinds of terms:

- linear terms: for each component $\hat{\omega}_i$ we have

$$\begin{aligned} \mathrm{E}(\hat{\omega}_i \mid \hat{\omega}^T \hat{\omega} = c_\alpha) &= \frac{1}{2}\mathrm{E}(\hat{\omega}_i \mid \hat{\omega}^T \hat{\omega} = c_\alpha, \hat{\omega}_i \geq 0) \\ &\quad + \frac{1}{2}\mathrm{E}(\hat{\omega}_i \mid \hat{\omega}^T \hat{\omega} = c_\alpha, \hat{\omega}_i < 0) \\ &= \frac{1}{2}\mathrm{E}(\hat{\omega}_i \mid \hat{\omega}^T \hat{\omega} = c_\alpha, \hat{\omega}_i \geq 0) \\ &\quad + \frac{1}{2}\mathrm{E}(-\hat{\omega}_i \mid \hat{\omega}^T \hat{\omega} = c_\alpha, \hat{\omega}_i > 0) \\ &= 0. \end{aligned} \tag{4.3}$$

[1] See Noble and Daniel (1988), pp. 317.

- mixed terms: by the same reasoning as before we get for the expectation of the product $\hat{\omega}_i\hat{\omega}_j$ for $i \neq j$:

$$\mathrm{E}(\hat{\omega}_i\hat{\omega}_j \mid \hat{\omega}^T\hat{\omega} = c_\alpha) = 0. \tag{4.4}$$

- quadratic terms: since $\sum_{j=1}^{M} \hat{\omega}_j^2 = c_\alpha$, it follows

$$\begin{aligned}
\mathrm{E}(\hat{\omega}_i^2 \mid \hat{\omega}^T\hat{\omega} = c_\alpha) &= \mathrm{E}(c_\alpha - \sum_{j\neq i} \hat{\omega}_j^2 \mid \hat{\omega}^T\hat{\omega} = c_\alpha) \\
&= \mathrm{E}(c_\alpha \mid \hat{\omega}^T\hat{\omega} = c_\alpha) - \sum_{j\neq i} \mathrm{E}(\hat{\omega}_j^2 \mid \hat{\omega}^T\hat{\omega} = c_\alpha) \\
&= c_\alpha - (M-1)\mathrm{E}(\hat{\omega}_i^2 \mid \hat{\omega}^T\hat{\omega} = c_\alpha),
\end{aligned}$$

  which implies that

$$\mathrm{E}(\hat{\omega}_i^2 \mid \hat{\omega}^T\hat{\omega} = c_\alpha) = \frac{c_\alpha}{M}. \tag{4.5}$$

Putting the results of (4.3), (4.4) and (4.5) together, the conditional expectation becomes

$$\begin{aligned}
\mathrm{E}(\hat{v}_q(\hat{\omega}) \mid \hat{\omega}^T\hat{\omega} = c_\alpha) &= \mathrm{E}(\frac{1}{2}\sum_{i,j=1}^{M} (\hat{G})_{i,j}\hat{\omega}_i\hat{\omega}_j + \sum_{i=1}^{M} \hat{g}_i\hat{\omega}_i \mid \hat{\omega}^T\hat{\omega} = c_\alpha) \\
&= \mathrm{E}(\frac{1}{2}\sum_{i=1}^{M} (\hat{G})_{i,i}\hat{\omega}_i^2 \mid \hat{\omega}^T\hat{\omega} = c_\alpha) \\
&= \frac{1}{2}\sum_{i=1}^{M} (\hat{G})_{i,i}\frac{c_\alpha}{M} \\
&= \frac{c_\alpha}{2}\frac{\mathrm{Tr}(\hat{G})}{M}.
\end{aligned} \tag{4.6}$$

■

Hence, the expected P&L on a sphere depends only on the trace of $\hat{G}$, the radius $\sqrt{c_\alpha}$ of the sphere and the dimension M of the problem. Since $\mathrm{Tr}(\hat{G})$ is identical to the sum of all eigenvalues of $\hat{G}$, we get that the expectation of the P&L on a sphere with radius $\sqrt{c_\alpha}$ is $\frac{c_\alpha}{2}$ times the mean curvature of the quadratic function $\hat{v}_q(\hat{\omega}) = \frac{1}{2}\hat{\omega}^T\hat{G}\hat{\omega} + \hat{g}^T\hat{\omega}$.

Example 4.8 *The analysis for the test portfolio of Example 1.17 is displayed in Figure 4.2, which shows the P&L along the ML and MP paths as well as the values of EV on the surface of the growing trust region. Since the absolute value of ML highly exceeds MP for every α and EV is negative and decreasing, it may be concluded that changes in the market rates affect the portfolio value negatively most of the time.*

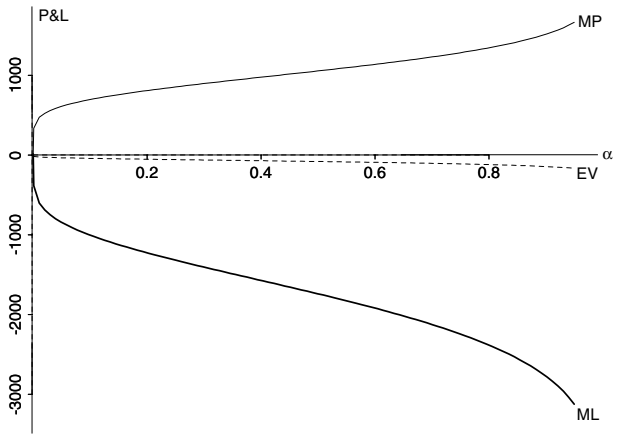

Figure 4.2: ML, MP and EV paths of test portfolio

Once an analytic expression for $\mathrm{E}(\hat{v}_q(\hat{\omega}) \mid \hat{\omega}^T\hat{\omega} = c_\alpha)$ is known, it is possible to calculate the expected value of the P&L *in the interior* of the trust region:

Corollary 4.9 *Let $\hat{\omega} \sim \mathcal{N}(0, \mathrm{I})$ and $\hat{v}_q(\hat{\omega}) = \frac{1}{2}\hat{\omega}^T\hat{G}\hat{\omega} + \hat{g}^T\hat{\omega}$. Then*

$$\mathrm{E}(\hat{v}_q(\hat{\omega}) \mid \hat{\omega}^T\hat{\omega} \leq c_\alpha) = \frac{\mathrm{Tr}(\hat{G})}{\alpha M} \frac{\Gamma_{\frac{c_\alpha}{2}}(\frac{M}{2}+1)}{\Gamma(\frac{M}{2})},$$

where $\Gamma_c(y) = \int_0^c y^{y-1}\exp(-y)dy$ denotes the incomplete gamma function.

Proof: From Theorem 4.7 we know that $\mathrm{E}(\hat{v}_q(\hat{\omega}) \mid \hat{\omega}^T\hat{\omega} = c_\alpha) = \frac{c_\alpha}{2}\frac{\mathrm{Tr}(\hat{G})}{M}$. Since c_α is the α–quantile of a χ^2 distribution with M degrees of freedom, the squared radius $y = \hat{\omega}^T\hat{\omega}$ can be seen as a random variable with density $f(y) = \frac{y^{\frac{M}{2}-1}}{\Gamma(\frac{M}{2})2^{\frac{M}{2}}}\exp(-\frac{y}{2})$. Integrating y from 0 to c_α leads to

$$
\begin{aligned}
\mathrm{E}(\hat{v}_q(\hat{\omega}) \mid \hat{\omega}^T\hat{\omega} \le c_\alpha) &= \int_0^{c_\alpha} \frac{y}{2}\frac{\mathrm{Tr}(\hat{G})}{M}\frac{1}{\alpha}\frac{y^{\frac{M}{2}-1}}{\Gamma(\frac{M}{2})2^{\frac{M}{2}}}\exp(-\frac{y}{2})dy \\
&= \frac{\mathrm{Tr}(\hat{G})}{2\alpha M}\frac{1}{\Gamma(\frac{M}{2})}\int_0^{c_\alpha}\frac{y^{\frac{M}{2}}}{2^{\frac{M}{2}}}\exp(-\frac{y}{2})dy \\
&= \frac{\mathrm{Tr}(\hat{G})}{\alpha M}\frac{1}{\Gamma(\frac{M}{2})}\int_0^{\frac{c_\alpha}{2}} z^{\frac{M}{2}}\exp(-z)dz \\
&= \frac{\mathrm{Tr}(\hat{G})}{\alpha M}\frac{\Gamma_{\frac{c_\alpha}{2}}(\frac{M}{2}+1)}{\Gamma(\frac{M}{2})},
\end{aligned}
$$

where $\frac{y}{2}$ has been substituted by z. ■

4.2 Coping with Nonlinearity

4.2.1 Quadratic Approximations

So far, we have mainly worked with quadratic P&L functions. As already mentioned, such functions can, for example, be obtained from local δ–Γ approximations $v_q(\omega) = \delta^T\omega + \frac{1}{2}\omega^T\Gamma\omega + o(\|\omega\|^2)$.[2] However, such local approximations are only valid for small changes of ω — in our model for short holding periods T and small confidence levels α. For large moves of the risk factors, local approximations can lead to tremendous errors. In practice, so–called risk profiles as those shown in Figure 1.2 are used to analyse the structure of the P&L functions:

Definition 4.10 *The set of scenarios $\mathcal{S} = \{\omega^{(1)}, \ldots, \omega^{(n)}\}$ holds n different M-dimensional scenarios. The profits and losses $\xi^{(i)} = v(\omega^{(i)})$ of the individual scenarios define the set $\mathcal{P} = \{\xi^{(1)}, \ldots, \xi^{(n)}\}$.*

From the sets $\mathcal{P}$ and $\mathcal{S}$, a quadratic approximation $v(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega + c$ can be constructed by minimizing the sum of the squared errors, i.e. by the method of least squares:

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{n}\left(\frac{1}{2}\omega^{(i)T}G\omega^{(i)} + g^T\omega^{(i)} + c - \xi^{(i)}\right)^2 \\
\text{s.t.} \quad & G \in \mathbb{R}^{M\times M}, \text{ symmetric} \\
& g \in \mathbb{R}^M \\
& c \in \mathbb{R}.
\end{aligned} \tag{4.7}
$$

[2] See Definition 1.8.

In this problem, the values $\omega^{(i)}$ and $\xi^{(i)}$ for $i = 1, \ldots, n$ are given and the matrix G, the vector g and the scalar c are the unknowns.

Definition 4.11 *The scenario matrix S is defined by the set of scenarios $\mathcal{S} = \{\omega^{(1)}, \ldots, \omega^{(n)}\}$ in the following way:*

$$S = \begin{array}{c} \\ \omega^{(1)} \\ \omega^{(2)} \\ \vdots \\ \omega^{(n)} \end{array} \begin{array}{c} \begin{array}{ccccccccccc} \frac{1}{2}\omega_1^2 & \omega_1\omega_2 & \ldots & \omega_1\omega_M & \frac{1}{2}\omega_2^2 & \omega_2\omega_3 & \ldots & \omega_1 & \ldots & \omega_M & 1 \end{array} \\ \begin{pmatrix} \ldots & & & & & & & & & \ldots & 1 \\ \ldots & & & & & & & & & \ldots & 1 \\ \vdots & & & & \ddots & & & & & & \vdots \\ \ldots & & & & & & & & & \ldots & 1 \end{pmatrix} \end{array}.$$

This means that S is an $n \times (\frac{M(M+3)}{2} + 1)$ matrix whose ith row contains, for scenario $\omega^{(i)}$, the various products indicated in the line above the matrix. The P&L vector $\xi \in \mathbb{R}^n$ holds the P&Ls of each scenario

$$\xi = \left(\xi^{(1)}, \ldots, \xi^{(n)}\right).$$

The vector of the unknowns x has an entry for each unknown of problem (4.7):

$$x = (G_{1,1}, G_{1,2}, \ldots, G_{1,M}, G_{2,2}, G_{2,3}, \ldots, G_{M,M}, g_1, \ldots, g_M, c),$$

in total $(\frac{M(M+3)}{2} + 1)$ elements.

With these new notations, problem (4.7) can be restated as

$$\min \|Sx - \xi\|^2.$$

The solution of this unconstrained minimization problem can be determined by solving a linear system:

Theorem 4.12 *x^* is solution to the minimization problem $\min \|Sx - \xi\|^2$ if and only if it is solution to the normal equation $S^T Sx = S^T \xi$.*

Proof: Since $S^T S$ is positive semidefinite, the function $f(x) = (Sx - \xi)^T (Sx - \xi)$ is convex and its gradient is $\nabla f(x) = 2S^T Sx - 2S^T \xi$. Thus, x^* is solution to $\nabla f(x^*) = 0$ if and only if $S^T Sx^* = S^T \xi$. ∎

The quadratic approximation $v_q(\omega)$ is well defined if the scenario matrix S is chosen appropriately:

Lemma 4.13 *Let $S \in \mathbb{R}^{n \times m}$ with $n \geq m$. Then, the normal equation $S^T Sx = y$ has a unique solution if S is of full rank.*

Proof: If we premultiply the equation $S^TSx = 0$ by x^T, we get $x^TS^TSx = \|Sx\|^2 = 0$, which implies that $x = 0$. ∎

Assumption 4.14 *We will assume in the remainder of this chapter that the set $\mathcal{S}$ of scenarios has been chosen such that the scenario matrix S is of rank $\frac{M(M+3)}{2} + 1$.*[3]

Note that the newly introduced scalar c can simply be added as a constant to all the previously discussed quantities ML, MP and EV.

4.2.2 Dynamic Weighting

The least squares method of Section 4.2.1 leads to quadratic approximations $v_q(\omega)$ with best fit with respect to the *entire* set $\mathcal{S}$ of scenarios. However, if the Maximum Loss is calculated *on the surface* of an ellipsoid,[4] we can get better quadratic approximations by weighting the scenarios adequately.

Definition 4.15 *Let $\mathcal{S} = \{\omega^{(1)}, \ldots, \omega^{(n)}\}$ be the set of scenarios. A weighting matrix for $\mathcal{S}$ is a diagonal matrix $W \in \mathbb{R}^{n\times n}$ with $(W)_{i,i} > 0, i = 1, \ldots, n$.*

The basic idea of the dynamic weighting procedure is to give a higher weight to those scenarios $\omega^{(i)}$, which lie close to the the surface of the ellipsoid, i.e. for which $\omega^{(i)T}\Sigma^{-1}\omega^{(i)} \approx c_\alpha$. In this case, problem (4.7) becomes

$$\min \|W^{\frac{1}{2}}(Sx - \xi)\|^2.$$

Theorem 4.16 *x^* is solution to the minimization problem $\min \|W^{\frac{1}{2}}(Sx - \xi)\|^2$ if and only if x^* is solution to $S^TWSx = S^TW\xi$.*

Proof: Since S^TWS is positive semidefinite, it follows that the function

$$f(x) = [W^{\frac{1}{2}}(Sx - \xi)]^T W^{\frac{1}{2}}(Sx - \xi)$$

is convex and its gradient is $\nabla f(x) = 2S^TWSx - 2S^TW\xi$. Therefore, x^* is solution to $\nabla f(x^*) = 0$ if and only if $S^TWSx^* = S^TW\xi$. ∎

In practice, good results have been obtained by using weights of the form

$$(W)_{i,i}(\alpha) = \frac{1}{1 + \beta \mid \omega^{(i)T}\Sigma_t^{-1}\omega^{(i)} - c_\alpha \mid^\gamma}, \tag{4.8}$$

[3] In Section 4.2.3 it will be explained how to choose a meaningful set of scenarios $\mathcal{S}$, which guarantees that the quadratic approximation $v_q(\omega)$ is well defined.

[4] See Section 4.1.1.

where $\beta, \gamma > 0$ are parameters which control the smoothness of the weighting function $W(\alpha)$ with respect to α. This notation stresses the fact that the weighting matrix W depends on the size of the ellipsoid, i.e. on the confidence level α. Hence, the unknown x of the normal equation depends also on α:

$$S^T W(\alpha) S x(\alpha) = S^T W(\alpha) \xi,$$

and the elements of $x(\alpha)$ define the matrix $G(\alpha)$, the vector $g(\alpha)$ and the scalar $c(\alpha)$ of the best fitting quadratic P&L function for confidence level α:

$$v_q(\omega, \alpha) = \frac{1}{2} \omega^T G(\alpha) \omega + g(\alpha)^T \omega + c(\alpha).$$

For fixed α, the paths of ML, MP and EV can be calculated applying the methods described in Section 4.1. This enables to analyze portfolios, which are not necessarily quadratic, by working with a sequence of quadratic approximations.

Example 4.17 *Figure 4.3 shows a two–dimensional, highly nonlinear P&L function. The bold lines represent the paths of the Maximum Loss scenarios and the Maximum Profit scenarios, which have been determined using a weighting function of type (4.8).*

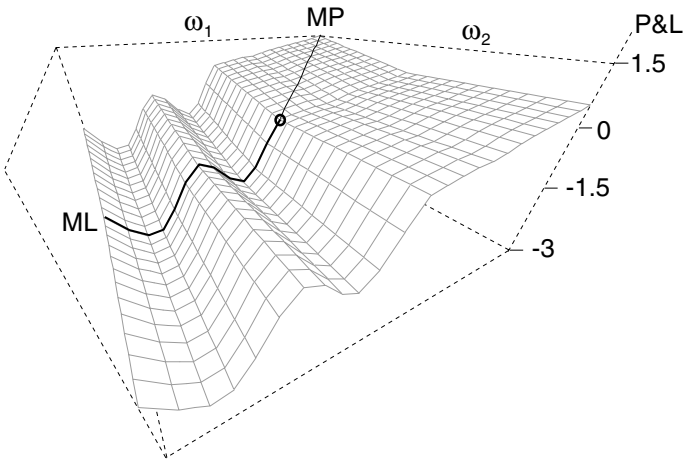

Figure 4.3: Nonquadratic portfolio with ML and MP paths

Figure 4.4 displays the P&Ls along these paths, together with the Expected Value EV on the surface of the ellipsoid.

Remark 4.18 *For quadratic portfolios, the technique of dynamic approximations produces exact results as there exists an x^* such that $Sx^* - \xi = 0$. Hence, $\|W^{\frac{1}{2}}(Sx^* - \xi)\|^2 = 0$, independently of the positive definite weighting matrix W.*

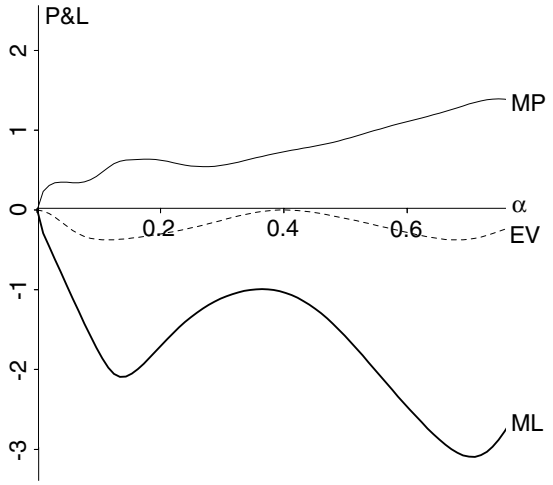

Figure 4.4: Characteristics of nonquadratic portfolio

4.2.3 Implementation

To implement the dynamic approximation technique, the following steps need to be executed:[5]

1. Given are a set of scenarios $\mathcal{S} = \{\omega^{(1)}, \ldots, \omega^{(n)}\}$ with corresponding P&Ls $\mathcal{P} = \{\xi^{(1)}, \ldots, \xi^{(n)}\}$, as well as the risk factor covariance matrix Σ for a holding period of T days.

2. Calculate the Cholesky decomposition $\Sigma = U^T U$.

3. Calculate the scenario matrix S and the P&L vector ξ according to Definition 4.11.

4. For $\alpha = 0\%, \ldots, 99\%$ do

   (a) Calculate the weight $(W)_{i,i}(\alpha) > 0$ of each scenario $\omega^{(i)}$, for example, using formula (4.8):

   $$(W)_{i,i}(\alpha) = \frac{1}{1 + \beta \mid \omega^{(i)T} \Sigma_t^{-1} \omega^{(i)} - c_\alpha \mid^\gamma},$$

   where $\beta, \gamma > 0$ are parameters which control the smoothness. All off–diagonal elements of $W(\alpha) \in \mathbb{R}^{n \times n}$ are set to 0.

[5]This is just an outline of the general structure of the algorithm. Practical implementations should take into account the various possibilities to reduce the computational costs. For example, it is advisable to use sparse matrix structures for the scenario matrix S and the weighting matrix $W(\alpha)$. Moreover, the quantity $\omega^{(i)T} \Sigma^{-1} \omega^{(i)}$ needs only to be evaluated in the first iteration of step 4(a).

(b) Solve the normal equation $S^T W(\alpha) S x(\alpha) = S^T W(\alpha)\xi$ for $x(\alpha)$.

(c) In accordance to Definition 4.11, the vector $x(\alpha)$ holds the elements of the symmetric matrix $G(\alpha) \in \mathbb{R}^{M\times M}$, the vector $g(\alpha) \in \mathbb{R}^M$, as well as the scalar $c(\alpha)$, which define the quadratic function

$$v_q(\omega, \alpha) = \frac{1}{2}\omega^T G(\alpha)\omega + g(\alpha)^T\omega + c(\alpha).$$

(d) Use the Cholesky decomposition of the covariance matrix to transform the ellipsoidal problem $\min_{\omega^T\Sigma^{-1}\omega\le c_\alpha} v_q(\omega, \alpha)$ to a spherical one: $\min_{\hat\omega^T\hat\omega\le c_\alpha} \frac{1}{2}\hat\omega^T\hat G\hat\omega + \hat g^T\hat\omega + c(\alpha)$, where $\hat g = Ug(\alpha)$ and $\hat G = UG(\alpha)U^T$.

(e) Determine the Maximum Loss

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad \frac{1}{2}\hat\omega^T\hat G\hat\omega + \hat g^T\hat\omega + c(\alpha) \\ & \text{s.t.} \quad \hat\omega^T\hat\omega = c_\alpha \end{aligned}$$

with the Levenberg–Marquardt algorithm of Section 3.4.1 (c_α is the α–quantile of a χ^2 distribution with M degrees of freedom, where M is the number of risk factors).

(f) Transform the result $\hat\omega^*$ back into the original coordinate system: $\omega^* = U^T\hat\omega^*$.

(g) Compute $\mathrm{MP}(\alpha)$ and the corresponding scenario by inverting the sign of the objective function[6] and applying steps (e) and (f).

(h) Calculate the expected value EV as shown in Section 4.1.2:

$$\mathrm{EV}(\alpha) = \frac{c_\alpha}{2}\frac{\mathrm{Tr}(\hat G)}{M}.$$

Choice of the scenario set $\mathcal{S}$

It remains to specify what scenarios $\omega^{(1)}, \ldots, \omega^{(n)}$ the set $\mathcal{S}$ should contain. A straightforward approach would be to cover the scenario space with a regular grid of points. However, if k points were chosen on each of the M axes, k^M scenarios would have to be evaluated in order to determine the set $\mathcal{P}$ of P&Ls. For portfolios with a large number M of risk factors (as observed in practice), this leads to a tremendous computational effort. Since the set $\mathcal{S}$ is used to construct the M–dimensional quadratic function

$$v(\omega, \alpha) = \frac{1}{2}\omega^T G(\alpha)\omega + g(\alpha)^T\omega + c(\alpha),$$

[6] See Remark 4.2.

with the unknowns being the elements of the $(M \times M)$ symmetric matrix $G(\alpha)$, the M–dimensional vector $g(\alpha)$ as well as the scalar $c(\alpha)$, the set $\mathcal{S}$ needs at least $(\frac{M(M+3)}{2} + 1)$ scenarios to determine a unique solution.[7]
In order to approximate the mixed term coefficients $G_{i,j}, i \neq j$, the set $\mathcal{S}$ must include scenarios whose components ω_i and ω_j are not simultaneously 0. If we restrict the values of each risk factor ω_i to an interval $[-l_i, l_i]$,[8] then the product $\omega_i\omega_j$ reaches its extremes at the corner points with coordinates $(\pm l_i, \pm l_j)$. Therefore, it seems reasonable to choose scenarios which lie on the diagonals in order to estimate the effect of the mixed terms (cf. Figure 4.5).
Choosing scenarios with at most two nonzero components $\omega_i, \omega_j \neq 0$ bears the advantage that one and the same set $\mathcal{S}$ can be used for analysing the effects of various subsets of risk factors and at different organizational levels, without having to reprice the instruments.

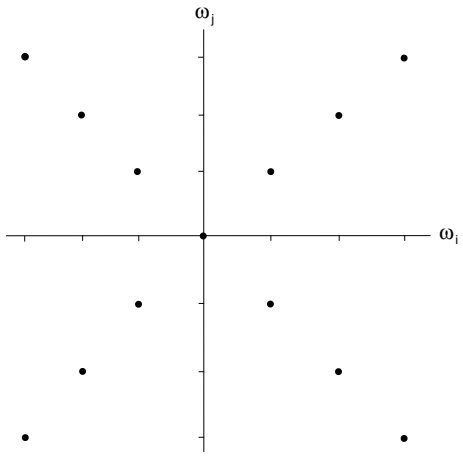

Figure 4.5: Choice of scenarios

Since the elements of the diagonal weighting matrix W are all positive, we conclude from Lemma 4.13 that the normal equation $S^TWSx = S^TW\xi$ of Theorem 4.16 has a unique solution if the scenario matrix S is of rank $(\frac{M(M+3)}{2} + 1)$.

Theorem 4.19 *Let the number of risk factors $M \geq 3$. If the set of scenarios $\mathcal{S}$ contains for each two by two combination of risk factors at least $k \geq 2$ symmetrical scenarios per diagonal, then the scenario matrix S is of rank $(\frac{M(M+3)}{2} + 1)$.*

Proof: If we include the scenario $\omega = 0$ into $\mathcal{S}$, this construction leads to a total number of $n = 2k\binom{M}{2} + 1 = kM(M-1) + 1$ scenarios. Since $kM(M-1) + 1 \geq \frac{M(M+3)}{2} + 1$ for $k \geq 2$ and $M \geq 3$, the matrix S has more rows than columns and

[7] See Lemma 4.13.

[8] Theorem 3.31 has shown that if $\omega \sim \mathcal{N}(0, \Sigma)$, then all feasible values of risk factor ω_i with respect to the trust region $\omega^T\Sigma^{-1}\omega \leq c_\alpha$ belong to the interval $[-\sqrt{c_\alpha}\sqrt{(\Sigma)_{i,i}}, \sqrt{c_\alpha}\sqrt{(\Sigma)_{i,i}}]$.

it remains to proof that the $(\frac{M(M+3)}{2}+1)$ columns of S are linearly independent. We will proceed by induction:
For $M = 3$ risk factors, it can easily be verified that the 13×10 matrix S_3 of Definition 4.11 is of rank 10.
Assume that the columns of the scenario matrix S_M are linearly independent. If we add the $(M+1)$th risk factor, the new scenario matrix S_{M+1} can be written as

$$S_{M+1} = \begin{pmatrix} S_M & 0 \\ Q & \tilde{S} \end{pmatrix},$$

where $\tilde{S}$ has the structure

$$\tilde{S} = \begin{array}{c} \\ \omega^{(n+1)} \\ \omega^{(n+2)} \\ \omega^{(n+3)} \\ \omega^{(n+4)} \\ \omega^{(n+5)} \\ \omega^{(n+6)} \\ \omega^{(n+7)} \\ \omega^{(n+8)} \\ \vdots \\ \omega^{(n+4M)} \end{array} \begin{array}{c} \begin{array}{ccccccc} \frac{1}{2}\omega_{M+1}^2 & \omega_1\omega_{M+1} & \omega_2\omega_{M+1} & \omega_3\omega_{M+1} & \dots & \omega_M\omega_{M+1} & \omega_{M+1} \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & -1 & 0 & 0 & \dots & 0 & 1 \\ 1 & 1 & 0 & 0 & \dots & 0 & -1 \\ 1 & -1 & 0 & 0 & \dots & 0 & -1 \\ 1 & 0 & 1 & 0 & \dots & 0 & 1 \\ 1 & 0 & -1 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & 0 & \dots & 0 & -1 \\ 1 & 0 & -1 & 0 & \dots & 0 & -1 \\ \vdots & & & & & & \vdots \\ 1 & 0 & \dots & 0 & 0 & -1 & -1 \end{pmatrix} \end{array}.$$

Of course, the number of rows of $\tilde{S}$ exceeds the number of columns and all its columns are linearly independent. Thus, the columns of S_{M+1} itself are linearly independent. ■

In practice, the number n of scenarios might be reduced if it is known in advance that there are pairs of risk factors ω_i and ω_j without cross effects. In this case, the corresponding element $G_{i,j}$ can be set to 0 a priori. However, it may become necessary to choose scenarios which lie directly on the axis ω_i (and/or ω_j) in order to ensure that the scenario matrix S remains of rank $(\frac{M(M+3)}{2}+1)$.[9]

[9]It should be noticed that the parameters β and γ of the weighting function (4.8)

$$(W)_{i,i}(\alpha) = \frac{1}{1+\beta \mid \omega^{(i)T}\Sigma^{-1}\omega^{(i)} - c_\alpha \mid^\gamma}$$

need to be adjusted to the number k of scenarios per diagonal to assure that the P&L functions $v_q(\omega, \alpha)$ are reasonably smooth with respect to α. For example, $\beta = 100$ and $\gamma = 2$ turned out to be good choices for $k = 6$.

4.3 Factors–at–Risk

4.3.1 Importance of Risk Factors

In Section 3.1 it has been shown that the risk of a portfolio can be reduced by improving the P&L for the worst case scenario ω^*. Is there any indication about which risk factor $\omega_i, i = 1, \ldots, M$, is most promising for reducing the risk? To answer this question, we will first investigate the effect of fixing $\omega_j = y$, where y is a given value.

Definition 4.20 *For normally distributed risk factors $\omega \sim \mathcal{N}(0, \Sigma)$ we define the restricted ML problem as*

$$\begin{aligned} \mathrm{ML}_j(\alpha, y) = \quad & \min \quad v(\omega) \\ & \text{s.t.} \quad \omega^T \Sigma^{-1} \omega \leq c_\alpha \\ & \qquad \omega_j = y. \end{aligned}$$

For quadratic portfolios $v_q(\omega)$, this restricted ML problem can be solved with the Levenberg–Marquardt algorithm of Section 3.4.1 after some transformations:

Theorem 4.21 *Let $\omega \sim \mathcal{N}(0, \Sigma)$ be an M–dimensional risk factor and $v_q(\omega) = \frac{1}{2}\omega^T G \omega + g^T \omega$ a quadratic P&L function, where $G \in \mathbb{R}^{M \times M}$ is a symmetric matrix and $g \in \mathbb{R}^M$.*
For every $y \in \mathbb{R}$ there exist

- *a symmetric matrix $\tilde{G} \in \mathbb{R}^{(M-1)\times(M-1)}$*
- *a vector $\tilde{g} \in \mathbb{R}^{(M-1)}$*
- *a scalar $\tilde{c}$*
- *a positive definite matrix $\tilde{\Sigma} \in \mathbb{R}^{(M-1)\times(M-1)}$*
- *a scalar $\tilde{c}_\alpha$,*

such that the restricted ML problem $\mathrm{ML}_j(\alpha, y) = \min_{\omega^T \Sigma^{-1} \omega \leq c_\alpha, \omega_j = y} v_q(\omega)$ is equivalent to the transformed problem

$$\begin{aligned} \mathrm{ML}(\alpha) = \quad & \min \quad \frac{1}{2}\tilde{\omega}^T \tilde{G} \tilde{\omega} + \tilde{g}^T \tilde{\omega} + \tilde{c} \\ & \text{s.t.} \quad \tilde{\omega}^T \tilde{\Sigma}^{-1} \tilde{\omega} \leq \tilde{c}_\alpha, \end{aligned}$$

where $\tilde{\omega} \in \mathbb{R}^{(M-1)}$.

Proof: Without loss of generality we may assume that the first risk factor is fixed to $\omega_1 = y$. Then, the parameters of the quadratic ML problem can be written as

$$G = \begin{bmatrix} G_{1,1} & G_1^T \\ G_1 & \tilde{G} \end{bmatrix}, \quad \Sigma^{-1} = \begin{bmatrix} \Sigma_{1,1}^{-1} & \Sigma_1^{-T} \\ \Sigma_1^{-1} & \tilde{\Sigma}^{-1} \end{bmatrix}, \quad g = \begin{bmatrix} g_1 \\ \hat{g} \end{bmatrix}, \quad \omega = \begin{bmatrix} \omega_1 \\ \hat{\omega} \end{bmatrix},$$

which represents a partition into the first column/row and the remaining $(M-1)$ ones. We define

$$\begin{aligned} \tilde{\omega} &= \hat{\omega} + \omega_1 \tilde{\Sigma} \Sigma_1^{-1} \\ \tilde{g} &= \hat{g} + \omega_1 \left[G_1 - \tilde{G} \tilde{\Sigma} (\Sigma_1^{-1}) \right] \\ \tilde{c} &= \omega_1 \left[g_1 - \hat{g}^T \tilde{\Sigma} (\Sigma_1^{-1}) \right] + \frac{1}{2} \omega_1^2 \left[G_{1,1} + (\Sigma_1^{-1})^T \tilde{\Sigma} \tilde{G} \tilde{\Sigma} (\Sigma_1^{-1}) \right] \\ \tilde{c}_\alpha &= c_\alpha + \omega_1^2 \left[(\Sigma_1^{-1})^T \tilde{\Sigma} (\Sigma_1^{-1}) - \Sigma_{1,1}^{-1} \right], \end{aligned}$$

where $\tilde{\Sigma}$ is the inverse of $\tilde{\Sigma}^{-1}$. It is easy to verify that this choice satisfies

$$\frac{1}{2} \omega^T G \omega + g^T \omega = \frac{1}{2} \tilde{\omega}^T \tilde{G} \tilde{\omega} + \tilde{g}^T \tilde{\omega} + \tilde{c}$$

and

$$\omega^T \Sigma^{-1} \omega - c_\alpha = \tilde{\omega}^T \tilde{\Sigma}^{-1} \tilde{\omega} - \tilde{c}_\alpha,$$

for every scenario $\omega \in \mathbb{R}^M$. Hence, the two problems are equivalent. ∎

Geometrically, fixing $\omega_1 = y$ corresponds to cutting the M–dimensional ellipsoid by a plane orthogonal to ω_1, resulting in an $(M-1)$–dimensional trust region of ellipsoidal form. Before the minimization problem can be solved, the new ellipsoid needs to be recentered at the origin. This is achieved by the above transformation, which keeps the objective function quadratic by introducing a new constant $\tilde{c}$.
The same kind of analysis can be performed for the restricted maximum profit $\mathrm{MP}_j(\alpha, y) = \max_{\omega^T \Sigma^{-1} \omega \le c_\alpha, \omega_j = y} v_q(\omega)$. In both cases, fixing $\omega_j = 0$ corresponds to *eliminating the exposure* to risk factor ω_j since the P&L function $v_q(\omega)$ no longer depends on ω_j.

Definition 4.22 *Let ω^{ML} be the overall worst case scenario, i.e. the solution to the problem $\min_{\omega^T \Sigma^{-1} \omega \le c_\alpha} v_q(\omega)$, and ω^{MP} the overall best case scenario, i.e. the solution to $\max_{\omega^T \Sigma^{-1} \omega \le c_\alpha} v_q(\omega)$. For each risk factor $\omega_j, j = 1, \ldots, M$, we define*

- *marginal Maximum Loss:* $\Delta\mathrm{ML}_j = v_q(\omega^{\mathrm{ML}}) - \mathrm{ML}_j(\alpha, 0)$

- *marginal Maximum Profit:* $\Delta\mathrm{MP}_j = v_q(\omega^{\mathrm{MP}}) - \mathrm{MP}_j(\alpha, 0)$
- *marginal Expected Value:*[10]

$$\Delta\mathrm{EV}_j = \mathrm{E}(v_q(\omega) \mid \omega^T\Sigma^{-1}\omega \leq c_\alpha) - \mathrm{E}(v_q(\omega) \mid \omega^T\Sigma^{-1}\omega \leq c_\alpha, \omega_j = 0).$$

The comparison of these quantities for all risk factors $\omega_j, j = 1, \ldots, M$, gives information about which risk factors have the greatest potential for reducing risk (by eliminating positions which depend on these factors), without affecting the profit potential too much.

Example 4.23 *Figure 4.6 shows an example of the restricted Maximum Loss* $\mathrm{ML}_j(\alpha, y)$ *and the restricted Maximum Profit* $\mathrm{MP}_j(\alpha, y)$ *for all feasible fixings* $\omega_j = y$*. The scenario* ω_j^{ML}*, where the curve* $\mathrm{ML}_j(y)$ *attains its minimum, is identical to the* j*th component of the overall worst case scenario* ω^* *(i.e. the solution to* $\min_{\omega^T\Sigma^{-1}\omega \leq c_\alpha} v_q(\omega)$ *).*

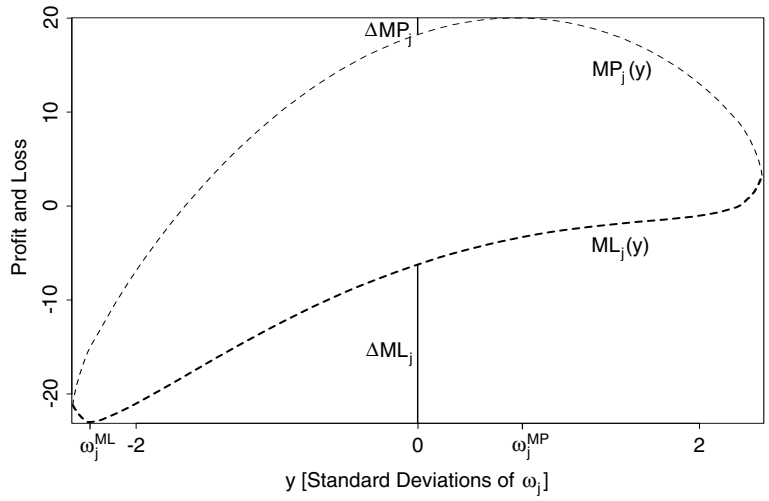

Figure 4.6: Marginal risk contributions of risk factor ω_j

For this portfolio, eliminating exposure in risk factor j *would reduce the overall Maximum Loss by* $\Delta\mathrm{ML}_j$*. Since the corresponding decrease* $\Delta\mathrm{MP}_j$ *in the maximum profit is considerably smaller, the risk of the portfolio can be reduced without influencing the profit potential substantially.*

[10] A formula to calculate the conditional expectations was presented in Corollary 4.9.

4.3.2 Risk Maps

Whereas Section 4.1.1 shows how to determine the path of the worst case scenarios by expansion of the trust region, this section addresses the question of how to extend a one–dimensional path to M–dimensional regions. Our goal is to identify regions, where the P&L is below (or above) an arbitrarily fixed value l, which is called *critical level*:

Definition 4.24 *Let $l \in \mathbb{R}$ be a critical level. A scenario $\omega \in \mathcal{O}$ is said to be dangerous if its P&L $v(\omega) < l$. Similarly, ω is a safe scenario if $v(\omega) > l$.*

A conceivable approach would be to determine the sets of all dangerous scenarios. For quadratic P&L functions $v_q(\omega)$, the level set $L(l) = \{\omega \in \mathbb{R}^M \mid v_q(\omega) = l\}$ is a second order surface (e.g. ellipsoid, cylinder, paraboloids, hyperboloids, cones, planes), which can be described parametrically: in principle, it is possible to use parametric equations to characterize the sets of safe and dangerous scenarios. In practice, however, this is not promising since such equations are very difficult to explicate and handle in high dimensional spaces.

Definition 4.25 *For a given critical level l and a trust region $\mathcal{T}$, the set $\mathcal{U} \subseteq \mathbb{R}^M$ is dangerous if $v(\omega) < l, \forall \omega \in \mathcal{T} \cap \mathcal{U}$. The set $\mathcal{U}$ is safe if $v(\omega) > l, \forall \omega \in \mathcal{T} \cap \mathcal{U}$.*

Sets with particularly simple form are *one–point intervals* $I_j(y) = \{\omega \in \mathbb{R}^M \mid \omega_j = y\}$; they represent an $(M-1)$–dimensional hyperplane in the risk factor space.

Theorem 4.26 *The one–point interval $I_j(y) = \{\omega \in \mathbb{R}^M \mid \omega_j = y\}$ is dangerous with respect to the critical level l and the trust region $\mathcal{T} = \{\omega \in \mathbb{R}^M \mid \omega^T \Sigma^{-1} \omega \leq c_\alpha\}$ if and only if $\mathrm{MP}_j(\alpha, y) < l$.*

Proof: Let $\mathcal{U} = \{\omega \in \mathbb{R}^M \mid \omega^T \Sigma^{-1} \omega \leq c_\alpha, \omega_j = y\}$. By Definition 4.20, the restricted Maximum Loss problem is $\mathrm{MP}_j(\alpha, y) = \max_{\omega \in \mathcal{U}} v(\omega)$. Obviously, we have $\mathrm{MP}_j(\alpha, y) < l$ if and only if $v(\omega) < l$ for every $\omega \in \mathcal{U}$. ∎

Corollary 4.27 *The one–point interval $I_j(y) = \{\omega \in \mathbb{R}^M \mid \omega_j = y\}$ is safe with respect to the critical level l and the trust region $\mathcal{T} = \{\omega \in \mathbb{R}^M \mid \omega^T \Sigma^{-1} \omega \leq c_\alpha\}$ if and only if $\mathrm{ML}_j(\alpha, y) > l$.*

Proof: As before, the condition $\mathrm{ML}_j(\alpha, y) = \min_{\omega^T \Sigma^{-1} \omega \leq c_\alpha, \omega_j = y} v(\omega) > l$ is equivalent to $v(\omega) > l, \forall \omega \in \{\omega \in \mathbb{R}^M \mid \omega^T \Sigma^{-1} \omega \leq c_\alpha, \omega_j = y\}$. ∎

Hence, the problem 'is the one–point interval $I_j(y)$ dangerous?' can be answered by solving the restricted maximum profit problem $\mathrm{MP}_j(\alpha, y)$. Next, we will generalize the concept of one–point intervals by introducing *one–dimensional intervals* $I_j(y_1, y_2) = \{\omega \in \mathbb{R}^M \mid y_1 \leq \omega_j \leq y_2\}$. By Definition 4.25, such sets are

dangerous with respect to the critical level l if $v(\omega) < l$ for every $\omega \in \mathcal{T}, \omega_j \in [y_1, y_2]$.
For quadratic P&L functions $v_q(\omega)$, dangerous and safe one–dimensional intervals can be determined numerically from the graphs $\mathrm{MP}_j(\alpha, y)$ and $\mathrm{ML}_j(\alpha, y)$:

Example 4.28 *Figure 4.7 shows the dangerous interval I_j^D and the safe interval I_j^S of risk factor ω_j for a critical level of $l = -10$. By Theorem 4.26 and Corollary 4.27, the dangerous interval corresponds to the segment where the iso-P&L line of level l lies above the function $\mathrm{MP}_j(y)$, and the safe interval is defined by the segment where the iso-P&L line lies below $\mathrm{ML}_j(y)$*

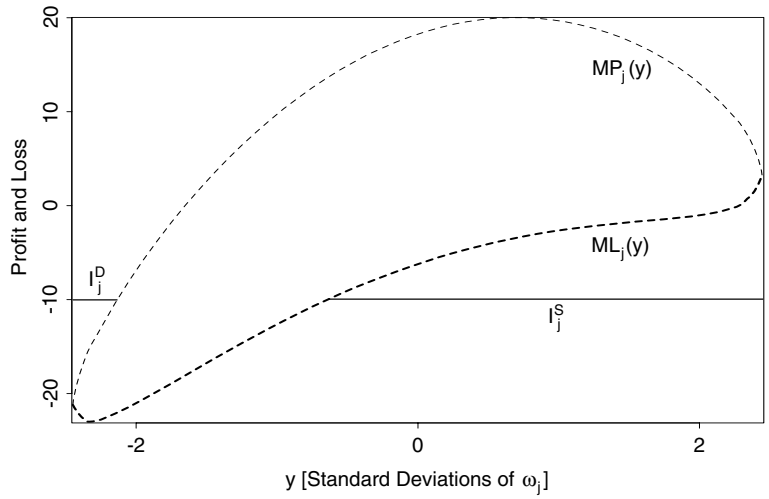

Figure 4.7: Dangerous and safe one–dimensional intervals

Theorem 4.29 *For a given critical level l, dangerous one–dimensional intervals $I_j^D(y_1, y_2)$ and safe intervals $I_j^S(y_3, y_4)$ do never intersect.*

Proof: Assume — ad absurdum — that there exists a $y \in \mathbb{R}$ which belongs to both intervals $[y_1, y_2]$ and $[y_3, y_4]$. Since $I_j^D(y_1, y_2)$ is a dangerous one–dimensional interval, Theorem 4.26 implies that $\mathrm{MP}_j(\alpha, y) < l$. On the other hand, $I_j^S(y_3, y_4)$ is safe and therefore $\mathrm{ML}_j(\alpha, y) > l$ by Corollary 4.27. This leads to the contradiction $\mathrm{MP}_j(\alpha, y) < l < \mathrm{ML}_j(\alpha, y)$. ■

Remark 4.30 *There may exist various disjoint dangerous (or safe) intervals $I_j^{D_1}, I_j^{D_2}, \ldots$ for each risk factor ω_j. An example can be found in Figure 4.7 if a critical level of $l = 10$ is chosen.*

One–dimensional intervals are particularly easy to interpret: a scenario $\omega \in \mathbb{R}^M$ is dangerous as soon as *one* of its components ω_j belongs to a one–dimensional dangerous interval $I_j^D, j = 1, \ldots, M$.[11]

Example 4.31 *Figure 4.8 shows the trust region $\mathcal{T}$ of a portfolio depending on two risk factors, together with the dangerous and safe regions which are implied by the one–dimensional intervals I_1^D, I_2^D and I_1^S.*

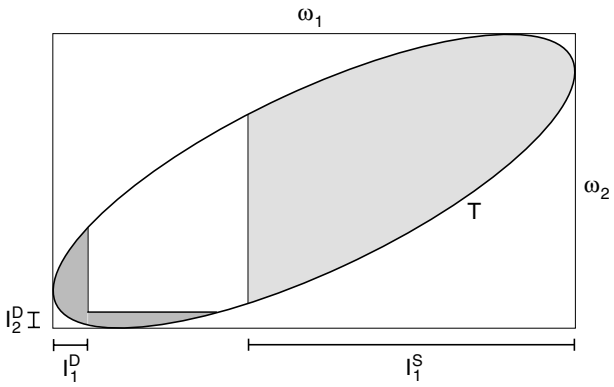

Figure 4.8: Dangerous and safe regions induced by one–dimensional intervals

It should be noticed that all the statements which have been made about regions are only valid for the feasible scenarios, i.e. those scenarios which belong to the trust region $\mathcal{T}$.[12] Since the set of the feasible scenarios grows with increasing confidence level α, the optimum of the problem $\mathrm{ML}_j(\alpha, y)$ decreases and the optimum of the problem $\mathrm{MP}_j(\alpha, y)$ increases. Therefore, the identifiable safe and dangerous regions might be *partially* reduced the greater α is chosen, i.e. the more scenarios are examined.

[11] Of course, an analogous statement holds for safe intervals: $\omega \in \mathbb{R}^M$ is safe as soon as one of its components ω_j belongs to a one–dimensional safe interval $I_j^S, j = 1, \ldots, M$.

[12] See Definition 4.25.

Chapter 5

Risk Adjusted Performance Measurement

5.1 Introduction

In the preceding part of this thesis, the problem of risk measurement has primarily been seen from a point of view of prudence: risk has been quantified to determine the amount of capital which has to be held as a cushion against potential future losses. However, in today's financial institutions, there is a rising demand for risk/reward assessments. Increasing market competition and more demanding shareholders compel to progress from a passive measurement/limitation of risks to an active management of risks: banks are forced to improve their *performance*, i.e. to optimize the relation between *returns* and risks.
Risk adjusted performance measurement (RAPM) is an attempt to provide a common yardstick which allows to contrast risk and return of portfolios, business units, business lines or whole companies. Such performance measures are the core elements for tackling various practical problems:

- valuation of companies (determining the shareholder value of a firm)
- financial governance (fixing financial goals and measuring the degree of their achievement)
- performance related compensation and incentive schemes
- allocation of capital.

The last item exemplifies the need for performance measures which are based on the same risk measures as those used for the risk limitation: on the one hand, capital can be viewed as a scarce resource whose minimal requirements are imposed by supervisory regulations. On the other hand, the more capital a financial institution holds, the more profit needs to be generated to satisfy the

shareholders' expectations. Hence, the question of how much capital (i.e. risk) to allocate to each business is recognized to be crucial for the management of a firm.
In the sequel, several methods for risk adjusted performance measurement will be discussed, and it will be shown that these methods are strongly related to each other. At this point, however, it should be mentioned that there is no standard terminology in the literature; the same name can be found for different methodologies, and different authors refer to one technique by different names. Nonetheless, all methods have obviously one thing in common: they all compare profits against risk.

5.2 Return on Capital

We will use the capital asset pricing model (CAPM)[1] as a motivation for a meaningful definition of a 'return on capital' ratio. If the asset returns $r_i, i = 1, \ldots, N$, are jointly normally distributed, then the CAPM postulates that in equilibrium the relation

$$\mathrm{E}(r_i) = r_f + \beta_i[\mathrm{E}(r_m) - r_f], \quad i = 1, \ldots, N \tag{5.1}$$

holds, where r_m is the return of the market portfolio m and r_f is the (deterministic) risk free rate of return. In Section 2.3.1 it has been shown that the optimal choice of the asset beta is $\beta_i = \frac{\rho_{i,m}\sigma_i}{\sigma_m}$, where σ_i denotes the standard deviation of asset i, σ_m the standard deviation of the market portfolio and $\rho_{i,m}$ is the correlation between asset i and the market portfolio. Thus, equation (5.1) can be reformulated as

$$\frac{\mathrm{E}(r_i) - r_f}{\sigma_i} = \rho_{i,m}\frac{\mathrm{E}(r_m) - r_f}{\sigma_m}, \quad i = 1, \ldots, N, \tag{5.2}$$

where the quantity $\mathrm{E}(r_i) - r_f$ is called *differential return* of asset i and $\frac{\mathrm{E}(r_i)-r_f}{\sigma_i}$ is referred to as *Sharpe ratio*.
Note that all elements of equation (5.2) represent relative amounts (i.e. they have no unit of measurement). Hence, we can replace returns r_i by profits R_i (i.e. net income minus opportunity costs of financing)[2] and asset volatilities by profit volatilities, without affecting the value of the quotients. If we assume that

[1] See Copeland and Weston (1995), pp. 193.

[2] In this context, profit is defined as net income minus opportunity costs of financing: let w be the value of an asset at the beginning of the period and $\tilde{w}$ its value at the end of the period. Then, the profit generated by this asset is $R = (\tilde{w} - w) - wr_f$, where r_f denotes the funding rate (which is supposed to be identical to the risk free rate).

the regulatory capital requirement C_i is a multiple ϑ of the profit volatility,[3] we get:

Theorem 5.1 *Let r_i be the return of asset i, r_f the risk free interest rate and σ_i the asset volatility. If the capital requirement C_i for the holdings in asset i is a multiple ϑ of the profit volatility, then $\frac{R_i}{C_i} = \frac{1}{\vartheta}\frac{r_i - r_f}{\sigma_i}$, where R_i denotes the profit generated by asset i.*

Proof: Let w_i be the value of asset i at the beginning of the period and $\tilde{w}_i$ its value at the end of the period. By definition, the return is equal to $r_i = \frac{\tilde{w}_i - w_i}{w_i}$. Since the profit generated by asset i is $R_i = (\tilde{w}_i - w_i) - w_i r_f$, we get

$$\frac{1}{\vartheta}\frac{r_i - r_f}{\sigma_i} = \frac{1}{\vartheta}\frac{(\tilde{w}_i - w_i) - w_i r_f}{w_i \sigma_i} = \frac{R_i}{C_i},$$

simply because $w_i\sigma_i$ is the volatility of the profit and therefore $C_i = \vartheta w_i \sigma_i$. ∎

Instead of considering a market of individual assets, we can alternatively interpret the market portfolio m as a firm with business units $i = 1, \ldots, N$. In this context, C_i corresponds to the capital requirement of business unit i, and the former asset profit R_i to the risk adjusted net return of business unit i, which comprises

+ revenues
– operating costs
– taxes
– financing costs
– expected losses.

The deduction of the financing costs from the revenues corresponds to the subtraction of the risk free interest rate r_f in formula (5.2). Moreover, taxes and expected losses are handled like an operating expense. Indeed, expected losses may be seen as the price which has to be paid to run a business and they should therefore be covered by provisions.[4]

Definition 5.2 *Let R_i be the net operating profit of business unit i and C_i the capital requirement for this business unit. The 'Return on Capital' of business unit i is defined as* $\mathrm{RoC}_i = \frac{R_i}{C_i}$.

[3] This is the case in the Basle Committee's regulations for internal (i.e. VaR based) models: for linear portfolios, VaR is proportional to the P&L volatility σ_v (cf. Theorem 1.16), and the regulatory capital requirement is $C = \vartheta z_{99\%}\sigma_v$, where typically $3 \leq \vartheta \leq 4$ and $z_{99\%}$ is the 99 percent quantile of the standard normal distribution.

[4] As explained in Chapter 1, risk (and hence capital) accounts only for the *variability* of the losses.

The coefficient $\frac{R_i}{C_i}$ is also called 'Risk Adjusted Return on Risk Adjusted Capital' (RARORAC)[5], which stresses the fact that there is an adjustment for the expected losses in the numerator and that the capital is a function of the risk.[6]

Example 5.3 *Suppose that a business unit holds assets with return* $r = 10\%$ *and volatility* $\sigma = 5\%$*, and that the risk free interest rate is* $r_f = 4\%$*. If the value of the assets is* 1000 *units at the beginning of the period, the profit turns out to be* $R = 1000(0.10 - 0.04) = 60$ *units. Assuming that the capital requirement is* $\vartheta = 3$ *times the profit volatility, it follows that* $C = 3 \cdot 1000 \cdot 0.05 = 150$ *units and consequently* $\mathrm{RoC} = \frac{R}{C} = \frac{60}{150} = 0.4$*, which is equivalent to* $\frac{1}{\vartheta}\frac{r-r_f}{\sigma} = \frac{1}{3}\frac{0.10-0.04}{0.05}$.

It should be noted that Definition 5.2 implies that a riskless asset leads to a RoC quotient with numerator and denominator identical to 0.
Finally, if we know the correlation $\rho_{i,m}$ between the profit R_i of business unit i and the profit R_m of the whole firm, we can formulate a relationship between the RoC of a business unit i and the RoC of the firm:

Corollary 5.4 *Let the profits* R_i *and* R_m *be normally distributed with correlation* $\rho_{i,m}$*. If the capital requirements* C_i *and* C_m *are a multiple* ϑ *of the profits' standard deviations, then* $\mathrm{E}(\mathrm{RoC}_i) = \rho_{i,m}\mathrm{E}(\mathrm{RoC}_m)$ *for* $i = 1, \ldots, N$.

Proof: Theorem 5.1 implies that $\frac{R_i}{C_i} = \frac{1}{\vartheta}\frac{r_i-r_f}{\sigma_i}$ and $\frac{R_m}{C_m} = \frac{1}{\vartheta}\frac{r_m-r_f}{\sigma_m}$. The result follows from formula (5.2) after taking the expectations. ■

5.3 Economic Value Added

It is important to understand that it is *not* sufficient for a firm to produce profits. In fact, the shareholders require to be adequately compensated for their investment risk. This can be seen in the CAPM model, where equation (5.1) states that the required return r_i increases with increasing volatility σ_i. Hence, the more capital C_i a business has to hold (from a regulatory perspective), the more profit R_i has to be generated to fulfil the shareholders' expectations. If we assume that the shareholders demand to be compensated by a return of κ, then we can quantify the wealth produced in one period by the concept of *Economic Value Added* (EVA):[7]

Definition 5.5 *Let* R *be the net operating profit of the firm,* C *the capital requirement and* κ *the cost of capital. The Economic Value Added is defined as* $\mathrm{EVA} = R - \kappa C$.

[5] See GARP (1996), p. 146.

[6] The terms 'Risk Adjusted Return on Capital' (RAROC) and 'Return on Risk Adjusted Capital' (RORAC) are also used for risk/return ratios, cf. Matten (1996), p. 59.

[7] See Hostettler (1997).

Hence, EVA is equal to the net profit less the costs for capital employed to produce this profit. Clearly, an EVA > 0 means that value for the shareholders is generated, whereas EVA < 0 indicates that economic value is destroyed.
In the economic literature, R is also called NOPAT (net operating profit after taxes), C is NOA (net operating assets) and κ is referred to as WACC (weighted average cost of capital). The meaning of these quantities is the following:

- R measures the net operating profit (minus taxes), without taking into consideration any accounting elements such as depreciation
- C is the capital which has to be held as a cushion against unexpected losses
- κ is the return required by the shareholders.

The performance measure EVA can be used to identify where value is created in a firm and where it is destroyed. However, the EVA of different companies cannot be compared directly since EVA is measured in nominal units and, thus, depends on the size of the firm. If EVA is divided by the capital C, we get the so–called *value spread*:

Definition 5.6 *Let R be the net operating profit, C the capital requirement and κ the cost of capital. The value spread is* $\text{VS} = \frac{R}{C} - \kappa$.

Using Definition 5.2, we can express the value spread in terms of return on capital: $\text{VS} = \text{RoC} - \kappa$, which leads to the hurdle rate $\text{RoC} \geq \kappa$ for profitable businesses.

Example 5.7 *In Example 5.3, the net profit was $R = 60$ units and the capital requirement was $C = 150$ units. If the shareholders require a return of $\kappa = 20\%$, we see that an economic profit of* $\text{EVA} = 60 - 0.2 \cdot 150 = 30$ *units is generated. The fact that this business is profitable is also reflected by the positive value spread of* $\text{VS} = \frac{60}{150} - 0.2 = 0.2$.

5.4 Shareholder Value Analysis

The concept of value spread cannot be used for capital allocation purposes because it is an ex–post, single period measure. This section discusses the application of EVA *forecasts* for defining a future oriented, multi–period risk adjusted performance measure, which we call *Market Value Added* (MVA).
In essence, MVA is the net present value of estimated future EVAs. The EVA forecasts are obtained from estimates of the net profit R_t and of the capital requirement C_t for the next periods $t = 1, \ldots, N$, assuming that the cost of capital κ is constant: $\text{EVA}_t = R_t - \kappa C_t$. The calculation of MVA is based on discounted future cash flows, using a discount rate commensurate with the risks

involved.[8] In our performance measurement framework, this concept translates into:

Definition 5.8 *Let* EVA_t *be the EVA forecast for each of the next periods* $t = 1, \ldots, N$*, with constant cost of capital* κ*. We assume that after period* N *the economic value added is constant and equal to* EVA_∞ *in each period* $t > N$*. Then, the market value added is defined as* $\mathrm{MVA} = \sum_{t=1}^{N} \frac{\mathrm{EVA}_t}{(1+\kappa)^t} + \frac{\mathrm{EVA}_\infty}{\kappa(1+\kappa)^N}$.

Note that this definition employs the cost of capital κ as 'risk adjusted' discount rate. Furthermore, the hypothesis that there exists a final value EVA_∞ implies that both R_t and C_t are assumed to be constant for $t > N$, i.e. that the firm has reached a steady state after period N; the term $\frac{\mathrm{EVA}_\infty}{\kappa(1+\kappa)^N}$ reflects the value of a perpetuity starting after period N.
Keeping in mind the meaning of EVA, it becomes obvious that MVA measures the value (in today's units) which will be created by a company in the future. MVA is just one constituent of the *total value* FV of a firm; indeed, FV can be decomposed into

$$\mathrm{FV} = \mathrm{MVA} + C_0 + \mathrm{VNO}, \tag{5.3}$$

where C_0 is the capital held in the present period and VNO denotes the value of all nonoperating assets. Since FV is equal to the value of equity plus debt plus reserves, formula (5.3) can be used to calculate the *shareholder value* of a firm (i.e. the value of equity). We can even show that MVA is identical to the results obtained with classical concepts, which are typically based on discounted free cash flows (i.e. the cash flows generated for the shareholders):

Definition 5.9 *Let* FCF_t *be the free cash flow in period* $t = 1, \ldots, N$*. We assume that the free cash flow after period* N *is constantly* FCF_∞*. Then, the shareholder value analysis (SVA) for a discount rate of* r *is defined as* $\mathrm{SVA} = \sum_{t=1}^{N} \frac{\mathrm{FCF}_t}{(1+r)^t} + \frac{\mathrm{FCF}_\infty}{r(1+r)^N}$.

This definition of SVA uses the same hypotheses regarding the final value FCF_∞ as Definition 5.8. The free cash flow FCF_t of period t is identical to the net profits R_t minus the net investments I_t (i.e. investments minus depreciation) of period t:

$$\mathrm{FCF}_t = R_t - I_t, \tag{5.4}$$

where the investments I_t are supposed to be made at the end of period t.

[8] This fundamental principle of finance is, for example, applied for pricing financial instruments or for evaluating investment projects.

Theorem 5.10 *Denote by C_0 the capital of the present period. If the discount rate in the shareholder value approach is identical to the cost of capital κ, then* $\mathrm{SVA} = \mathrm{MVA} + C_0$.

Proof: First, we note that the investments I_t, which are made at the end of period t, have to be financed by increasing the capital C_{t+1} of period $(i+1)$: $C_{t+1} = C_t + I_t$, which can be restated as $C_t = C_0 + \sum_{k=1}^{t-1} I_k$ for $t = 1, \ldots, N$. By Definition 5.8, we have a constant profit R_∞ and a constant capital requirement C_∞ after period N. Thus, $\mathrm{EVA}_\infty = R_\infty - \kappa C_\infty$ and $\mathrm{FCF}_\infty = R_\infty$, and it follows that

$$\begin{aligned} \mathrm{MVA} + C_0 &= \sum_{t=1}^{N} \frac{R_t}{(1+\kappa)^t} - \kappa \left\{ \sum_{t=1}^{N} \frac{C_0}{(1+\kappa)^t} + \sum_{t=1}^{N} \frac{\sum_{k=1}^{t-1} I_k}{(1+\kappa)^t} \right\} \\ &\quad + \frac{R_\infty}{\kappa(1+\kappa)^N} - \kappa \left\{ \frac{C_0}{\kappa(1+\kappa)^N} + \frac{\sum_{k=1}^{N} I_k}{\kappa(1+\kappa)^N} \right\} + C_0. \quad (5.5) \end{aligned}$$

Furthermore, equation (5.4) implies

$$\mathrm{SVA} = \sum_{t=1}^{N} \frac{R_t}{(1+\kappa)^t} - \sum_{t=1}^{N} \frac{I_t}{(1+\kappa)^t} + \frac{R_\infty}{\kappa(1+\kappa)^N}. \quad (5.6)$$

Since the expression $\sum_{t=1}^{N} \frac{R_t}{(1+\kappa)^t} + \frac{R_\infty}{\kappa(1+\kappa)^N}$ appears in both formulas (5.5) and (5.6), it remains to show that

$$-\kappa \left\{ \sum_{t=1}^{N} \frac{C_0}{(1+\kappa)^t} + \frac{C_0}{\kappa(1+\kappa)^N} \right\} + C_0 = 0 \quad (5.7)$$

and

$$\kappa \left\{ \sum_{t=1}^{N} \frac{\sum_{k=1}^{t-1} I_k}{(1+\kappa)^t} + \frac{\sum_{k=1}^{N} I_k}{\kappa(1+\kappa)^N} \right\} = \sum_{t=1}^{N} \frac{I_t}{(1+\kappa)^t}. \quad (5.8)$$

Let us start with (5.7): if $C_0 = 0$, the relation is satisfied. Otherwise, we divide by C_0 and get $\kappa \sum_{t=1}^{N} \frac{1}{(1+\kappa)^t} + \frac{1}{(1+\kappa)^N} - 1 = 0$. From $\sum_{t=0}^{N} \frac{1}{(1+\kappa)^t} = \frac{(1+\kappa)^{N+1}-1}{\kappa(1+\kappa)^N}$, the result follows immediately.
If we multiply equation (5.8) by $(1+\kappa)^N$, we obtain

$$\kappa \sum_{t=1}^{N} \sum_{k=1}^{t-1} I_k (1+\kappa)^{N-t} + \sum_{k=1}^{N} I_k = \sum_{k=1}^{N} I_k (1+\kappa)^{N-k},$$

which is satisfied if for every $k = 1, \dots, N$ the relation $\kappa I_k \sum_{t=k+1}^{N}(1+\kappa)^{N-t} + I_k = I_k(1+\kappa)^{N-k}$ holds. If $I_k = 0$, the relation is satisfied; else we divide by I_k and get $\kappa \sum_{t=0}^{N-k-1}(1+\kappa)^t + 1 = (1+\kappa)^{N-k}$, which is true since $\sum_{t=0}^{N-k-1}(1+\kappa)^t = \frac{1-(1+\kappa)^{N-k}}{1-(1+\kappa)}$. ■

Example 5.11 *We assume that in each future period the net profit will constantly be* $R = 60$ *units and the capital requirement* $C = 150$ *units. If the shareholders require a return of* $\kappa = 20\%$ *per period, it follows that the value generated for the shareholders is* $\mathrm{MVA} = \frac{60-0.2\cdot 150}{0.2(1.0+0.2)} = 125$ *units. Thus, the total shareholder value (including today's equity) is* $\mathrm{SVA} = 125 + 150 = 275$ *units.*

Hence, we can calculate the value created for the shareholders from forecasts of the net profit R_t and the capital requirement C_t, which is itself a function of the risk $\rho(v)$ of a portfolio.[9] This allows us, by equation (5.3), to estimate the value of a company or the value of the equity, which result from the future effects of today's decisions. In this sense, the quantities risk and return can be used as basic elements for management decisions at the firm level.
However, the availability of a risk adjusted performance measure provides no answer to the question of how much capital to allocate to which business. In fact, the relation between risk and return is only one element which may affect the decision: many other, non quantifiable aspects such as the state of a market, the strategy of the competitors, the structure of the firm, etc. need also to be respected in the decision process. Thus, the problem of capital allocation is still unresolved and offers a wide area for research.

[9] See also formula (1.8).

Appendix A

Stochastic Processes

This appendix presents the basics of stochastic calculus, which are referred to in Chapter 1. The core of risk management is the analysis of market rates. Mathematically, the evolution of a market rate can be modelled as a *stochastic process*:

Definition A.1 *A stochastic process is a family of random variables* $(X_t)_{t\geq 0}$ *defined on the probability space* $(\Omega, \mathcal{A}, \mathrm{P})$ *with values in a measurable space* $(E, \mathcal{E})$.

In practice, the index t stands for time and the process can be understood as a function $X : \mathbb{R}^+ \times \Omega \to E$. The filtration $\mathcal{F}_t$ represents all the knowledge of the process at time t:

Definition A.2 *Let* $(\Omega, \mathcal{A}, \mathrm{P})$ *be a probability space. A family* $(\mathcal{F}_t)_{t\geq 0}$ *of* σ*–algebras of* $\mathcal{A}$ *is a filtration if* $\mathcal{F}_s \subseteq \mathcal{F}_t \subset \mathcal{A}$ *for every* $0 \leq s < t$.

Definition A.3 *A process* $(X_t)_{t\geq 0}$ *is said to be adapted to* $(\mathcal{F}_t)_{t\geq 0}$ *if* X_t *is* $\mathcal{F}_t$*–measurable for every* t.

We can use the process $(X_t)_{t\geq 0}$ itself to define a filtration:

Assumption A.4 *In the sequel, we will always take the natural filtration of the process* $(X_t)_{t\geq 0}$*:*

$$\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t) \cup \{A \in \mathcal{A} \mid \mathrm{P}(A) = 0\}.^1$$

Among the stochastic processes, *Brownian motion*[2] plays a particularly important role:

Definition A.5 *The process* $(X_t)_{t\geq 0}$ *with* $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ *is called Brownian motion if it has independent and stationary increments and if its trajectories are continuous* (P a.s.).

[1] Of course, each process $(X_t)_{t\geq 0}$ is adapted to his natural filtration.

[2] Brownian motion is also referred to as Wiener process.

If the Brownian motion $(B_t)_{t\geq 0}$ satisfies the conditions $B_0 = 0$ (P a.s.), $\mathrm{E}(B_t) = 0$ and $\mathrm{E}(B_t^2) = t$, then this process is called *standard* Brownian motion. In this case, it can be shown that $B_t \sim \mathcal{N}(0,t)$ and indeed more generally $B_t - B_s \sim \mathcal{N}(0, t-s)$ for $0 \leq s < t$. Standard Brownian motion is used to define the class of *Itô processes*:

Definition A.6 *We say that $(X_t)_{0\leq t\leq \tilde{t}}$ is an Itô process if*

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dB_s, \quad \forall t \leq \tilde{t}, \quad (\mathrm{P\ a.s.}),$$

where X_0 is $\mathcal{F}_0$–measurable, $(K_t)_{0\leq t\leq \tilde{t}}, (H_t)_{0\leq t\leq \tilde{t}}$ are both $\mathcal{F}_t$–adapted and $\int_0^{\tilde{t}} |K_s| ds < \infty$, $\int_0^{\tilde{t}} |H_s|^2 ds < \infty$ (P a.s.).

Finally, *Itô's lemma* shows how to differentiate a function $g(X_t)$ of an Itô process:

Theorem A.7 *Let $(X_t)_{0\leq t\leq \tilde{t}}$ be an Itô process $X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dB_s$ and g twice continuously differentiable (i.e. $g \in C^2$). Then*

$$g(X_t) = g(X_0) + \int_0^t g'(X_s) dX_s + \frac{1}{2} \int_0^t g''(X_s) d\langle X, X\rangle_s,$$

where $\langle X, X\rangle_t = \int_0^t H_s^2 ds$ and $\int_0^t g'(X_s) dX_s = \int_0^t g'(X_s) K_s ds + \int_0^t g'(X_s) H_s dB_s$.

Proof: A proof of this theorem is presented in Karatzas and Shreve (1988), pp. 149. ■

It should be noted that this theorem is often written in differential form as $dg(X_t) = g'(X_t) dX_t + \frac{1}{2} g''(X_t) d\langle X, X\rangle_t$.

Appendix B

Distributions of Quadratic Functions

In Section 1.3.1 we discussed the calculation of VaR for quadratic functions, as described by Rouvinez (1997). Starting point was the P&L function $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$ of multinormally distributed risk factors $\omega \sim \mathcal{N}(0, \Sigma)$, which was transformed into the separated system

$$\breve{v}_q(\breve{\omega}) = \sum_{j\in I} \frac{1}{2}(\tilde{G})_{j,j}\breve{\omega}_j^2 + \sum_{j\in J} \tilde{g}_j\breve{\omega}_j - \frac{1}{2}\sum_{j\in I} \frac{\tilde{g}_j^2}{(\tilde{G})_{j,j}} \tag{B.1}$$

of independent random variables $\breve{\omega}_j, j = 1, \ldots, M$, where $I = \{j \in \{1, \ldots, M\} \mid (\tilde{G})_{j,j} \neq 0\}$ and $J = \{j \in \{1, \ldots, M\} \mid (\tilde{G})_{j,j} = 0\}$.
In this section, we will first develop an expression for the characteristic function $\varphi_{\breve{v}_q}(s)$ of the separated system $\breve{v}_q(\breve{\omega})$. Then, we will use the inversion theorem to calculate the probability $\mathrm{P}(\breve{v}_q(\breve{\omega}) \leq z)$ for a given fractile z. Finally, we can apply a bisectionning technique for determining that value z^* for which $\mathrm{P}(\breve{v}_q(\breve{\omega}) \leq z^*) = 1 - \alpha$. Since VaR is the $(1-\alpha)$–quantile of the P&L distribution, we get that $\mathrm{VaR}(\alpha) = z^*$.
In expression (B.1), the random variables $\breve{\omega}_j, j \in J$, are standard normally distributed whereas the quantities $\breve{\omega}_j^2, j \in I$, represent noncentral χ^2 variables with one degree of freedom and noncentrality parameter $(\frac{\tilde{g}_j}{(\tilde{G})_{j,j}})^2$. In order to simplify the notations, we define

Definition B.1 *For each $j \in I$ we set $Y_j = \frac{1}{2}(\tilde{G})_{j,j}\breve{\omega}_j^2$; the normal variables are aggregated into $Y_J = \sum_{j\in J} \tilde{g}_j\breve{\omega}_j$ and $c = -\frac{1}{2}\sum_{j\in I} \frac{\tilde{g}_j^2}{(\tilde{G})_{j,j}}$. The new P&L function $v(Y)$ is then defined as $v(Y) = \sum_{j\in I} Y_j + Y_J + c$.*

This new formulation has the same characteristics as the separated system $\breve{v}_q(\breve{\omega})$:

Lemma B.2 *The random variables Y_J and $Y_j, j \in I$, are mutually independent and the P&L functions are identical: $v(Y) = \check{v}_q(\check{\omega})$.*

Proof: The mutual independence of the random variables $\check{\omega}_j$ implies that Y_J and $Y_j, j \in I$, are independent since each $\check{\omega}_j$ is element of exactly one of the new variables. The fact that $\sum_{j\in I} Y_j = \sum_{j\in I} \frac{1}{2}(\tilde{G})_{j,j}\check{\omega}_j^2$ and the definitions of Y_J and c imply that

$$\check{v}_q(\check{\omega}) = \sum_{j\in I} \frac{1}{2}(\tilde{G})_{j,j}\check{\omega}_j^2 + \sum_{j\in J} \tilde{g}_j\check{\omega}_j - \frac{1}{2}\sum_{j\in I} \frac{\tilde{g}_j^2}{(\tilde{G})_{j,j}} = \sum_{j\in I} Y_j + Y_J + c = v(Y).$$

■

Lemma B.3 *The distribution of the random variable Y_J is $Y_J \sim \mathcal{N}(0, \sum_{j\in J} \tilde{g}_j^2)$.*

Proof: Y_J is a linear combination of independent standard normal variables with $\mathrm{E}(Y_J) = 0$ and $\mathrm{Var}(Y_J) = \sum_{j\in J} \tilde{g}_j^2$. ■

Since we know the distribution of the random variables Y_J and $Y_j, j \in I$, we know also their characteristic functions:

Theorem B.4 *The characteristic functions of the random variables are:*

$$\begin{aligned}
\varphi_{Y_j}(s) &= [1 - i(\tilde{G})_{j,j}s]^{-\frac{1}{2}} \exp\left(\frac{i\tilde{g}_j^2 s}{2(\tilde{G})_{j,j}[1 - i(\tilde{G})_{j,j}s]}\right), \quad j \in I \\
\varphi_{Y_J}(s) &= \exp\left(-\frac{\sum_{j\in J} \tilde{g}_j^2 s^2}{2}\right) \\
\varphi_c(s) &= \exp(ics),
\end{aligned}$$

where $i = \sqrt{-1}$.

Proof: From Definition B.1 it follows that for each $j \in I$ the random variable Y_j is a multiple of a noncentral χ^2 variable with one degree of freedom and noncentrality parameter $\zeta_j = (\frac{\tilde{g}_j}{(\tilde{G})_{j,j}})^2$. Since the characteristic function of such a variable is equal to $(1 - 2is)^{-\frac{1}{2}} \exp(\frac{i\zeta_j s}{1-2is})$,[1] the results for $\varphi_{Y_j}(s)$ are a direct consequence of the fact that $\varphi_{\theta Y}(s) = \varphi_Y(\theta s)$ for $\theta \in \mathbb{R}$.
Moreover, the characteristic function of a normal variable[2] with mean 0 and variance σ^2 is $\exp(-\frac{\sigma^2 s^2}{2})$, and the definition of the characteristic function of a random variable Y: $\varphi_Y(s) = \mathrm{E}(\exp(iYs))$ implies that $\varphi_c(s) = \exp(ics)$. ■

[1] See Johnson and Kotz. (1970), p. 134.
[2] See Johnson et al. (1994), p. 89.

Corollary B.5 *The characteristic function $\varphi_v(s)$ of the P&L function $v(Y)$ is*

$$\begin{aligned}\varphi_v(s) &= \prod_{j\in I}\left\{[1-i(\tilde{G})_{j,j}s]^{-\frac{1}{2}}\exp\left(\frac{i\tilde{g}_j^2 s}{2(\tilde{G})_{j,j}[1-i(\tilde{G})_{j,j}s]}\right)\right\} \\ &\quad\cdot\exp\left(ics-\frac{\sum_{j\in J}\tilde{g}_j^2 s^2}{2}\right).\end{aligned}$$

Proof: This result is simply due to the fact that the characteristic function of a sum of independent variables is the product of the characteristic functions of the individual variables. ■

Knowing the characteristic function of $v(Y)$, we can use the inversion theorem to calculate the probability $\mathrm{P}(v(Y)\leq z)$ for a given fractile z:

Theorem B.6 *Let Y be a random variable with characteristic function $\varphi_Y(s)$. If $\mathrm{E}(|Y|)<\infty$ and if there exist $\theta_1,\theta_2>0$ such that $|\varphi_Y(s)|<\theta_1 s^{-\theta_2},\forall s>1$, then*

$$\mathrm{P}(Y\leq z)=\frac{1}{2}+\frac{1}{2\pi}\int_0^\infty\frac{\exp(izs)\varphi_Y(-s)-\exp(-izs)\varphi_Y(s)}{is}ds.$$

Proof: The elements of the proof are given in Gil–Pelaez (1951) and Davies (1973). ■

Both conditions of the theorem are satisfied for the P&L function $v(Y)$: first, $\mathrm{E}(|v|)\leq\sum_{j\in I}\mathrm{E}(|Y_j|)+\mathrm{E}(|Y_J|)+|c|<\infty$. Second, the condition $|\varphi_v(s)|<\theta_1 s^{-\theta_2},\forall s>1$ is satisfied because $\varphi_v(s)$ is the product of characteristic functions of normal and noncentral χ^2 variables, which all satisfy this condition individually.[3] Hence, we can apply Theorem B.6 to calculate $\mathrm{P}(v(Y)\leq z)$ for a given fractile z. Since VaR is the $(1-\alpha)$–quantile of the P&L distribution $v(Y)$, we can use a bisectionning technique to determine that value z^* for which $\mathrm{P}(v(Y)\leq z^*)=1-\alpha$.

[3]The constant c has no influence since $|\varphi_c(s)|\equiv 1$ and, excluding the case where we have no risk factors in the model, there is at least one normal or χ^2 variable.

Glossary of Notation

a	M–dimensional vector of linear P&L function $v_l(\omega) = a^T\omega$
b	Vector of linear constraints: $A\omega \leq b$
c	Constant in P&L function, e.g. $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega + c$
c_α	α–quantile of χ^2 distribution with M degrees of freedom
e_i	ith eigenvector of matrix G
$f(\cdot)$	Density function of a random variable
g	M–dimensional vector of P&L function $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$
$h(\rho)$	Transformation of risk measure $\rho(v)$ into capital charge
k	Number of scenarios chosen on each diagonal to approximate quadratic P&L $v_q(\omega)$
l	Critical level of P&L
n	Number of scenarios in the set $\mathcal{S} = \{\omega^{(1)}, \ldots \omega^{(n)}\}$
$p(\bar{\psi})$	Profit/Loss incurred at time T due to expected outcome $\bar{\psi}$
$p_i^{(\cdot)}$	Position (market value) in risk factor ω_i
$r(v)$	Capital requirement for portfolio v
r_I	Return of equity index I
r_j	Return of stock j
t	Time
$u(\psi)$	Value of portfolio for market rates ψ
$v(\omega)$	Profit/Loss of scenario ω
$v_l(\omega)$	Linear P&L function $v_l(\omega) = a^T\omega, a \in \mathbb{R}^M$
$v_q(\omega)$	Quadratic P&L function $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$
x	Vector used for calculating the best fitting quadratic approximation
z_α	α–quantile of standard normal distribution
A	Matrix of linear constraints: $A\omega \leq b$
(B_t)	Standard Brownian motion
C_i	Capital requirement of business unit i
$\mathrm{Corr}(\cdot,\cdot)$	Correlation of random variables
$\mathrm{Cov}(\cdot,\cdot)$	Covariance of random variables
D	Polytope of feasible scenarios
$\mathrm{E}(\cdot)$	Expected value of a random variable
EV	Expected P&L on surface of trust region

EVA_i	Economic Value Added by business unit i
G	Symmetric matrix of $\mathbb{R}^{M\times M}$ of quadratic P&L function $v_q(\omega) = \frac{1}{2}\omega^T G\omega + g^T\omega$
I	Identity matrix
$I_j(y)$	One–point interval, i.e. subspace $I_j(y) = \{\omega \in \mathbb{R}^M \mid \omega_j = y\}$
M	Number of risk factors $\omega_1, \ldots, \omega_M$
$\mathrm{ML}(\alpha)$	Maximum Loss for confidence level α
$\mathrm{ML}_j(\alpha, y)$	Maximum Loss if risk factor ω_j is fixed to $\omega_j = y$
$\mathrm{MP}(\alpha)$	Maximum Profit for confidence level α
$\mathrm{MP}_j(\alpha, y)$	Maximum Profit if risk factor ω_j is fixed to $\omega_j = y$
MVA	Market Value Added
$\mathrm{P}(\cdot)$	Probability of a random variable
R	Correlation matrix
R_i	Net operating profit (after taxes) of business unit i
RoC_i	Return on Capital of business unit i
S	Scenario matrix used to calculate best fitting quadratic approximation $v_q(\omega)$
SVA	Shareholder Value Analysis
T	Length of holding period $[0, T]$
$\mathrm{Tr}(\cdot)$	Trace of a matrix
U	Cholesky decomposition of covariance matrix $\Sigma = U^T U$
V	Diagonal matrix of standard deviations: $(V)_{i,i} = \sqrt{(\Sigma)_{i,i}}$
$\mathrm{Var}(\cdot)$	Variance of a random variable
$\mathrm{VaR}(\alpha)$	Value–at–Risk for confidence level α
$W(\alpha)$	Weighting matrix for confidence level α, i.e. $(W)_{i,i}$ is the weight of scenario $\omega^{(i)}, i = 1, \ldots, n$
(X_t)	Stochastic process
α	Confidence level, $\alpha \in [0, 1]$
β, γ	Parameters used to control the smoothness of the weighting function $W(\alpha)$
δ	First order price sensitivity of an instrument
κ	Cost of capital, i.e. return required by the shareholders
λ_i	ith eigenvalue of matrix G
μ	Lagrange multiplier
ν	Variable used in the Levenberg–Marquardt algorithm (playing the role of a Lagrange multiplier)
ξ	Vector $\xi = (\xi^{(1)}, \ldots, \xi^{(n)})$ of scenario P&Ls, where $\xi^{(i)} = v(\omega^{(i)})$
$\rho(v)$	Risk measure of portfolio v
ψ	Market rates $\psi = (\psi_1, \ldots, \psi_M)$ at time T
ψ_i^0	Market rate i at time 0
$\psi_{\sigma,i}$	Volatility of asset i
$\bar{\psi}$	Expected value of market rates at time T, i.e. $\bar{\psi} = \mathrm{E}(\psi)$

ω	Risk factor $\omega = (\omega_1, \ldots, \omega_M)$; represents deviations in market rates, i.e. $\omega = \psi - \bar{\psi}$
$\omega^{(i)}$	Element of the scenario set $\mathcal{S} = \{\omega^{(1)}, \ldots \omega^{(n)}\}$
$\omega_{\sigma,i}$	Volatility risk factor of asset i
Γ	Second order price sensitivity of an instrument
$\Gamma_c(\cdot)$	Incomplete gamma function
Σ	Covariance matrix of risk factors: $\omega \sim \mathcal{N}(0, \Sigma)$
Ω	Probability space
$\mathcal{N}(0, \Sigma)$	Normal distribution with mean 0 and covariance matrix Σ
$\mathcal{O}$	Set of realizable scenarios, $\mathcal{O} \subseteq \mathbb{R}^M$
$\mathcal{P}$	Set of scenario P&Ls: $\mathcal{P} = \{\xi^{(1)}, \ldots, \xi^{(n)}\}$
$\mathcal{S}$	Set of scenarios $\{\omega^{(1)}, \ldots, \omega^{(n)}\}$ used to construct quadratic approximation $v_q(\omega)$
$\mathcal{T}$	Trust region with probability $\mathrm{P}(\mathcal{T}) = \alpha$
$\mathcal{U}$	Subset of risk factors, $\mathcal{U} \subseteq \mathbb{R}^M$
$\mathcal{V}$	Set of all P&L functions $v(\omega)$
y^+, y^-	Positive and negative parts of $y \in \mathbb{R}$, i.e. $y^+ = \max\{0, y\}$ and $y^- = \min\{0, y\}$

Bibliography

[Alexander (1996)] C. Alexander. Volatility and Correlation Forecasting. In: The Handbook of Risk Management and Analysis, edited by C. Alexander, John Wiley & Sons, Chichster, 1996.

[Artzner et al. (1996)] Ph. Artzner, F. Delbaen, J.–M. Eber, D. Heath. A Characterization of Measures of Risk. Working paper, February 1996.

[Baxter and Rennie (1996)] M. Baxter and A. Rennie. Financial Calculus, An introduction to derivative pricing. Cambridge University Press, Cambridge, 1996.

[Beckers (1996)] S. Beckers. A Survey of Risk Measurement, Theory and Practice. In: The Handbook of Risk Management and Analysis, edited by C. Alexander, John Wiley & Sons, Chichster, 1996.

[Beckström and Campbell (1995)] R. Beckström and A. Campbell. An Introduction to VAR. C·ATS Software Inc., Palo Alto, 1995.

[Bellare and Rogaway (1995)] M. Bellare and Ph. Rogaway. The complexity of approximating a nonlinear program. Mathematical Programming 69 (1995), 429–441.

[BIS (1988)] Basle Committee on Banking Supervision. International Convergence of Capital Measurement and Capital Standards. July, 1988.

[BIS (1995)] Basle Committee on Banking Supervision. An Internal Model–Based Approach to Market Risk Capital Requirements. April, 1995.

[BIS (1996a)] Basle Committee on Banking Supervision. Overview of the Amendment to the Capital Accord to Incorporate Market Risks. January, 1996.

[BIS (1996b)] Basle Committee on Banking Supervision. Amendment to the Capital Accord to Incorporate Market Risks. January, 1996.

[Chen (1996)] L. Chen. Interest Rate Dynamics, Derivatives Pricing, and Risk Management. Springer, Berlin, 1996.

[Copeland and Weston (1995)] T.E. Copeland and J.F. Weston. Financial Theory and Corporate Policy, Third Edition. Addison–Wesley Publishing Group, Reading, 1995.

[Cox and Rubinstein (1985)] J.C. Cox and M. Rubinstein. Options Markets. Prentice Hall, Englewood Cliffs, 1985.

[Cox et al. (1981)] J.C. Cox, J.E. Ingersoll, S.A. Ross. The Relationship between Forward Prices and Futures Prices. Journal of Financial Economics, December 1981.

[Davies (1973)] R.B. Davies. Numerical inversion of a characteristic function. Biometrika (60), Cambridge, 1973.

[Dembo et al. (1995)] R. Dembo, G. Pastrana, M. Zerbs. Stress Testing Portfolios, Conference Documentation. IIR Limited, London, June 1995.

[Dennis and Schnabel (1996)] J.E. Dennis and R.B. Schnabel. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice–Hall, Englewood Cliffs, 1996.

[Duffie and Pan (1997)] D. Duffie and J. Pan. An Overview of Value at Risk. Working Paper, January 1997.

[Fabozzi (1993)] F.J. Fabozzi. Bond Markets, Analysis and Strategies, Second Edition. Prentice–Hall, Englewood Cliffs, 1993.

[Fletcher (1987)] R. Fletcher. Practical Methods of Optimization, Second Edition. John Wiley & Sons, Chichester, 1987.

[Fu et al. (1996)] M. Fu, Z.–Q. Luo, Y. Ye. Approximation algorithms for quadratic programming. University of Newcastle, McMaster University, University of Iowa; August 1996.

[Garman (1997)] M.B. Garman. Ending the Search for Component VaR. Financial Engineering Associates, 1997.

[GARP (1996)] Coopers & Lybrand. Generally Accepted Risk Principles. 1996.

[Gil–Pelaez (1951)] J. Gil–Pelaez. Note on the inversion theorem. Biometrika (31), Cambridge, 1951.

[Heath et al. (1992)] D. Heath, R. Jarrow, A. Morton. Bond Pricing and the Term Structure of Interest Rates: a New Methodology for Contingent Claims Valuation. Econometrica (60), 1992.

[Horst et al. (1995)] R. Horst, P.M. Pardalos, N.V. Thoai. Introduction to Global Optimization. Kluwer Academic Publishers, Dordrecht, 1995.

[Hostettler (1997)] S. Hostettler. Economic Value Added (EVA): Darstellung und Anwendung auf Schweizer Aktiengesellschaften. Haupt, Bern, 1997.

[Hull (1993)] J.C. Hull. Options, Futures and other Derivative Securities. Prentice-Hall, Englewood Cliffs, 1993.

[Jackson (1991)] J.E. Jackson. A User's Guide to Principal Components. John Wiley & Sons, New York, 1991.

[Johnson and Kotz. (1970)] N.L. Johnson and S. Kotz. Continuous Univariate Distributions - 2. John Wiley & Sons, New York, 1970.

[Johnson et al. (1994)] N.L. Johnson, S. Kotz, N. Balakrishnan. Continuous Univariate Distributions, Second Edition, Volume 1. John Wiley & Sons, New York, 1994.

[Jorion (1997)] Ph. Jorion. Value at Risk: The New Benchmark for Controlling Derivatives Risk. Irwin, Chicago, 1997.

[Karatzas and Shreve (1988)] I. Karatzas and S.E. Shreve. Brownian Motion and Stochastic Calculus. Springer, New York, 1988.

[Lamberton and Lapeyre (1993)] D. Lamberton and B. Lapeyre. Introduction au Calcul Stochastique Appliqué à la Finance. Ellipses, Paris, 1993.

[Litterman and Scheinkman (1991)] R. Litterman and J. Scheinkman. Common Factors Affecting Bond Returns. Journal of Fixed Income, June 1991.

[Matten (1996)] Ch. Matten. Managing Bank Capital: Capital Allocation and Performance Measurement. John Wiley & Sons, Chichester, 1996.

[Moser and Quast (1994)] H. Moser and W. Quast. Organisation des Risikomanagements in einem Bankkonzern. In: Handbuch Bankcontrolling, edited by H. Schierenbeck & H. Moser, Gabler, Wiesbaden, 1994.

[Noble and Daniel (1988)] B. Noble and J. Daniel. Applied Linear Algebra. Prentice–Hall, Englewood Cliffs, 1988.

[Pritsker (1996)] M. Pritsker. Evaluating Value at Risk Methodologies: Accuracy versus Computational Time. Federal Reserve Board, Washington, 1996.

[RiskMetrics (1996)] J.P. Morgan & Co. RiskMetricsTM — Technical Document. New York, December 1996.

[Rouvinez (1997)] Ch. Rouvinez. Mastering Delta Gamma. Risk Magazine, London, February 1997.

[Schaefer (1995)] S. M. Schaefer. Risk Assessment and VAR Models. Documentation of the RI$K '95 Workshop on 'Implementing The Latest Practical Advances In Making Risk Assessment And VAR Models Work: Effective Risk Aggregation And Compliance With The Basle Proposals'. Paris, November, 1995.

[SPAN (1993)] Chicago Mercantile Exchange (CME). Standard Portfolio Analysis of Risk Overview. Chicago, 1993.

[SPAN (1997)] Chicago Mercantile Exchange (CME). The Standard Portfolio Analysis of Risk (SPAN) Performance Bond System at the Chicago Mercantile Exchange, Technical Specifications. Chicago, 1997.

[Taylor (1986)] S. Taylor. Modelling Financial Times Series. John Wiley & Sons, Chichester, 1986.

[Vasicek and Fong (1982)] O.A. Vasicek and G.H. Fong. Term Structure Modeling Using Exponential Splines. Journal of Finance, May 1982.

[Vavasis (1991)] S.A. Vavasis. Nonlinear Optimization, Complexity Issues. Oxford University Press, New York, 1991.

[Wilson (1996)] T. Wilson. Calculating Risk Capital. In: The Handbook of Risk Management and Analysis, edited by C. Alexander, John Wiley & Sons, Chichster, 1996.

[Ye (1997)] Yinyu Ye. Approximating Quadratic Programming with Bound Constraints. Working Paper, University of Iowa, March 1997.