Time consistency in Risk measures



Ivan Romanovski
St Cross College
University of Oxford

A thesis submitted in partial fulfilment for the degree of
Master of Science in Mathematical and Computational Finance

Trinity 2012

Acknowledgements

I would like to thank my supervisor, Samuel N. Cohen, for his guidance, support and patience as well as for all his valuable suggestions and advice. His valuable insights and encouragement have certainly helped me greatly throughout my work.

Abstract

An approach to construct time consistent dynamic risk measures from rescaled one-period risk measures is presented. By composing one-period time risk measure (with suitable scaling) a limit is obtained corresponding to the solution of a BSDE with a particular driver. For composed versions of risk measures Value-at-Risk and Expected Shortfall particular drivers of corresponding BSDEs are proposed. What happens to composed VaR and composed ES when the size of the discrete time grid goes to zero is investigated numerically. In particular, example with smooth terminal payoff when both composed VaR and composed ES converge to the solutions of BSDEs with proposed drivers is presented. Solutions corresponding to those BSDEs depend on the initial choice of risk level α, which gives a natural extension of discrete time risk measures to continuous time.

Keywords: dynamic risk measures, time consistency, g-expectation, discretisation, convergence, special drivers

Contents

1 Background **1**

1.1 Introduction 1

1.2 Risk measures 2

1.2.1 Static risk measures 3

1.2.2 Static risk measures via g-expectations 4

1.2.3 Dynamic risk measures 7

2 The model **10**

2.1 Value-at-Risk 10

2.1.1 Composed VaR 12

2.1.2 Composed VaR via g-expectations 13

2.2 Expected Shortfall 14

2.2.1 Composed ES via g-expectations 15

2.3 Research field overview 17

3 Numerical implementation **18**

3.1 Regression-based Monte Carlo method to solve BSDEs 18

3.1.1 Numerical scheme 20

3.1.2 Numerical experiment 21

3.2 Approximation results 23

3.2.1 Non-smooth payoff 24

3.2.2 Smooth payoff 25

4 Conclusions **26**

A MATLAB implementation **27**

A.1 BSDE solver 27

A.1.1 Additional functions 28

A.2 Composed VaR and composed ES 30

A.2.1 Additional functions 32

Bibliography **33**

Chapter 1

Background

1.1 Introduction

Consider a financial portfolio containing assets and liabilities. For a holder of such a portfolio it is important to estimate the risk of his position. The purpose of a risk measure is usually to summarise risk-relevant information about the position in a single number which should, in some form, relate to the potential losses of this portfolio. This might seem to be an oversimplification, however as pointed out by Delbaen [9] risk analysis often comes down to a yes-no question and therefore giving a single numerical value is a reasonable intermediate step. What exactly we want this number to represent depends on our application. Most commonly this would be used for the determination of risk capital and capital adequacy, or as a management tool. For example, one of the most common measures of risk, Value-at-Risk (VaR), is part of the Basel II capital-adequacy framework and has also been used as a management tool (the JPMorgan 4:15 report) [5] . VaR corresponds to the smallest amount of capital a bank or an insurance company needs to add to its position such that the probability of a negative outcome is kept below the desired risk level α.

Coherent risk measures (which were later generalised to convex risk measures) are a particular axiomatic class of risk measures which assign a number to a portfolio which can then be interpreted as minimal capital reserve. Coherent static (i.e. one-period) risk measures were introduced by Artzner et al. [1, 2]; they were inspired by the capital adequacy rules laid out in the Basel Accord. The more general concept of a convex risk measure was later simultaneously developed by Föllmer and Schied [12, 13] and Fritelli and Rosazza Gianin [14].

Static risk measures can be generalised to dynamic risk measures, where information is updated in each time period. In a dynamic context time consistency is a natural assumption used to link static risk measures together. This means that the

same risk is assigned to a financial position regardless of whether it is calculated over two time periods at once or in two steps backward in time. This is in fact equivalent to the property that if portfolio X is riskier than portfolio Y at some time t in the future, then it is guaranteed to be riskier at any time before t.

Dynamic risk measures have often been considered in a discrete time setting (see Stadje [26] and references therein); however, in many situations information arrives continuously and it seems natural to assume that the agent is allowed to asses his risk at any time. An elegant approach to obtain continuous time consistent convex risk measures is the use of an operator given by the solution of a backward stochastic differential equation (BSDE), the so-called g-expectation (here g is the driver of the BSDE) introduced by Peng [24]. Recently he has shown [23] the deep connection between BSDEs and dynamic risk measures, proving that any dynamic risk measure (satisfying some axiomatic conditions) is necessarily associated with a BSDE (the converse having been known for years). For a proof that *(almost) any dynamic coherent or convex risk measure comes from a conditional g-expectation* based on Coquet et. al. [8] see Rosazza Gianin [25].

The problem with modelling risk measures as solutions to BSDEs is that the drivers defining the underlying BSDEs are difficult to interpret. In this dissertation we will show that two particular static risk measures, VaR and Expected Shortfall (ES), when applied recursively on a discrete-time grid with appropriate scaling both converge to the solution of a particular type of BSDE.

In the remainder of this chapter we will present some necessary background material on risk measures based on the article by Rosazza Gianin [25]. In Chapter 2 we will present measures VaR and ES and argue for the choice of particular drivers of the associated BSDEs. In Chapter 3 we will explain the method used for numerical approximation of BSDEs, compare the numerical results given by recursively composing VaR and ES with numerical results given by a solution to the associated BSDEs and present our findings.

1.2 Risk measures

In this section we will present some results that show the connection between g-expectations and risk measures. In the first subsection we will review two classes of static risk measures, i.e. coherent and convex risk measures. The distinction between these two classes of risk is completely captured by the differences between the sets of characterising axioms. In the next subsection we will state the connection

between those risk measures and g-expectations. In particular we will state the results that show that, given a functional g (satisfying some suitable assumptions) and its corresponding g-expectation $\mathcal{E}_g$, the risk measure defined (on a suitable space) by

$$\rho_g(X) \triangleq \mathcal{E}_g(-X)$$

is coherent if g is sublinear and it is convex if g is convex. The final section concerns dynamic risk measures and the problem of a dynamic risk measure being induced by a conditional g-expectation.

1.2.1 Static risk measures

Risk measures were introduced to evaluate future losses in order to give some criteria on the acceptability of risk exposures. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, T a fixed future date and $\mathcal{X}$ the space of all financial positions in which we are interested. For instance, an element of $\mathcal{X}$ may be the net worth at maturity T of a financial contract.

A *static risk measure* is a functional $\rho : \mathcal{X} \to \mathbb{R}$, which can be interpreted as follows. Given a financial position X, the quantity $\rho(X)$ represents the riskiness of X and, by convention, X is acceptable when $\rho(X) \leq 0$ and unacceptable otherwise.

Such a map could satisfy some of the following axioms, many of which can be traced back to the premium principles in insurance literature:

- convexity: $\rho(\alpha X + (1-\alpha)Y) \leq \alpha\rho(X) + (1-\alpha)\rho(Y),\ \forall X, Y \in \mathcal{X}$,
- positivity: $X \geq 0 \Rightarrow \rho(X) \leq \rho(0)$,
- constancy: $\rho(\alpha) = -\alpha,\ \forall \alpha \in \mathbb{R}$,
- translability: $\rho(X + \alpha) = \rho(X) - \alpha,\ \forall \alpha \in \mathbb{R},\ \forall X \in \mathcal{X}$,
- sublinearity:
  - positive homogeneity: $\rho(\alpha X) = \alpha\rho(X),\ \forall X \in \mathcal{X}$,
  - subadditivity: $\rho(X + Y) \leq \rho(X) + \rho(Y),\ \forall X, Y \in \mathcal{X}$,
- lower semi-continuity: $\{X \in \mathcal{X} : \rho(X) \leq \gamma\}$ is closed in $\mathcal{X}$ for any $\gamma \in \mathbb{R}$.

We can briefly explain these axioms in the following way. Positivity guarantees that a position which always yields a gain has riskiness which is lower than the risk of the null position, $X \equiv 0$. Constancy means that the riskiness of constant positions is simply the opposite of their net worth. Translability implies $\rho(X + \rho(X)) = 0$, hence

for an unacceptable position X, the value $\rho(X)$ represents the minimal additional capital needed to get an acceptable new position $X + \rho(X) > X$. Subadditivity encourages the diversification of a portfolio since, under subadditivity, the riskiness of a portfolio $X + Y$ is smaller than the sum of the riskiness of the single positions X and Y. On the contrary, convexity assures only diversification through portfolios originated by "ad hoc" weighted sums of single positions. The main difference between convexity and sublinearity is positive homogeneity, which was originally motivated by liquidity reasons and it basically means that if we scale our portfolio by a positive factor then we double scale our risk by the same factor. Furthermore, it is well known that sublinearity is stronger than convexity and that sublinearity is equivalent to the pair convexity & positive homogeneity.

Depending on the set of characterising axioms we distinguish between two families of risk measures: coherent and convex risk measures.

Definition 1.2.1 (Artzner et al. [2] and Delbaen [10])**.** A functional $\rho : \mathcal{P} \to \mathbb{R}$ is a coherent risk measure if it satisfies positivity, translability and sublinearity.

Since positive homogeneity does not model liquidity risk well, the notion of coherence has been subsequently relaxed and convex risk measures were proposed simultaneously by Föllmer and Schied [12] and Fritelli and Rosazza Gianin [14].

Definition 1.2.2 (Fritelli and Rosazza Gianin [14])**.** A functional $\rho : \mathcal{P} \to \mathbb{R}$ is a convex risk measure if it satisfies convexity, lower semi-continuity and $\rho(0) = 0$.

1.2.2 Static risk measures via g-expectations

A possible approach to obtain a large set of risk measures is through the use of an operator given by the solution to a backward stochastic differential equation (BSDE), the so-called g-expectation (here g is the driver of the BSDE). BSDEs are well known in finance because of their importance in the problem of derivative pricing; however, this is only their best known application, since many other problems in finance may be modelled by BSDEs.

Recall that, to evaluate a European vanilla option, one has to look for a self-financing strategy that replicates the payoff of such an option. Since the dynamics of such a replicating strategy and its final condition are known, the price of the option can be deduced by no arbitrage argument. Such case correspond to a linear BSDE, Pardoux and Peng [21] introduced non-linear BSDEs. The main definitions and results on BSDEs and g-expectations are now recalled.

Let $(W_t)_{0\leq t\leq T}$ be a standard one dimensional Brownian motion defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $(\mathcal{F}_t^W)_{0\leq t\leq T}$ be the filtration generated by it, i.e. $\mathcal{F}_t^B \triangleq \sigma\{W_t : 0 \leq t \leq T\}$ for any $0 \leq t \leq T$. Denote by $(\mathcal{F}_t)_{0\leq t\leq T}$ the augmented filtration associated with $(\mathcal{F}_t^W)_{0\leq t\leq T}$, i.e. $(\mathcal{F}_t) = \sigma\{\mathcal{F}_t^W \cup \mathcal{N}\}$ for any $0 \leq t \leq T$, where $\mathcal{N}$ is the collection of all $\mathbb{P}$-null sets.

Let $L^2(\mathcal{F}_t) = L^2(\Omega, \mathcal{F}, \mathbb{P})$ for $t \in [0, T]$ denote the space of all real valued, $\mathcal{F}_t$ measurable and square integrable random variables endowed with L^2-norm $\|\cdot\|_2$ topology, and let $L^2_{\mathcal{F}}(T; \mathbb{R})$ denote the space of all $\mathbb{R}$-valued, adapted processes $(V_t)_{t\in[0,T]}$ such that $\mathbb{E}[\int_0^T \|V_t\|^2 dt] < \infty$, where $\|\cdot\|$ stands for the Euclidean norm on $\mathbb{R}$.

Consider a functional $g : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, which maps (ω, t, y, z) to $g(\omega, t, y, z)$. Any equality involving stochastic processes has to be understood in a sense $\mathbb{P}$ almost surely but to simplify our notation we will omit "$\mathbb{P}$-a.s." and write $g(t, y, z)$ instead of $g(\omega, t, y, z)$.

The functional g is always supposed to satisfy the following assumptions.

1. g is Lipschitz in (y, z), i.e. there exists a constant $C > 0$ such that for any $t \in [0, T]$ and any $(y_0, z_0), (y_1, z_1) \in \mathbb{R} \times \mathbb{R}$,
   $$|g(t, y_0, z_0) - g(t, y_1, z_1)| \leq C(|y_0 - y_1| + |z_0 - z_1|),$$
2. $g(\cdot, y, z) \in L^2_{\mathcal{F}}(T; \mathbb{R})$ for any $y, z \in R$,
3. $\forall t \in [0, T]$ and $\forall y \in \mathbb{R}, g(t, y, 0) = 0$.

We will refer to the above assumptions as *usual assumptions*. Under these assumptions on g, for every $X \in L^2(\mathcal{F}_T)$ the BSDE

$$\begin{aligned} -dY_t &= g(t, Y_t, Z_t)dt - Z_t dW_t \\ Y_T &= X \end{aligned} \tag{1.1}$$

has a unique square integrable, adapted solution (see [21, 8]), i.e. there is a unique pair $(Y_t, Z_t)_{0\leq t\leq T} \in L^2_{\mathcal{F}}(T; \mathbb{R}) \times L^2_{\mathcal{F}}(T; \mathbb{R})$ that solves (1.1). In our case we are interested only in component $(Y_t)_{0\leq t\leq T}$ of the solution, i.e. the so-called "conditional g-expectation".

Definition 1.2.3 (Peng [22])**.** For any $X \in L^2(\mathcal{F}_T)$, let $(Y_t^X, Z_t^X)_{0\leq t\leq T} \in L^2_{\mathcal{F}}(T; \mathbb{R}) \times L^2_{\mathcal{F}}(T; \mathbb{R})$ be the solution of the BSDE (1.1) with terminal condition X. The *g-expectation* $\mathcal{E}_g$ of X is then defined by

$$\mathcal{E}_g[X] = Y_0^X,$$

while, for any $t \in [0, T]$, the *conditional g-expectation* of X under $\mathcal{F}_t$ is defined by

$$\mathcal{E}_g[X|\mathcal{F}_t] = Y_t^X.$$

The conditional g-expectation of X defined above is the unique random variable in $L^2(\mathcal{F}_t)$ satisfying

$$\mathcal{E}_g[1_A X] = \mathcal{E}_g[1_A \mathcal{E}_g[X|\mathcal{F}_t]],\ \forall A \in \mathcal{F}_t.$$

As shown by Peng [22], the classical (conditional) expectation corresponds to the case of $g \equiv 0$.

Let g satisfy the usual assumptions and set $\rho_g : L^2(\mathcal{F}_T) \to \mathbb{R}$ as follows:

$$\rho_g(X) \triangleq \mathcal{E}_g[-X],\ \forall X \in L^2(\mathcal{F}_T).$$

In order to make ρ_g a static risk measure some additional restrictions have to be imposed on the functional g. To make notation easier we introduce two properties that the functional g might possess.

Definition 1.2.4. Functional g is *sublinear* in (y, z) if it is

- *positively homogeneous* in (y, z):

  $\forall t \in [0, T],\ \forall \alpha \geq 0,\ \forall (y, z) \in \mathbb{R} \times \mathbb{R}, g(t, \alpha y, \alpha z) = \alpha g(t, y, z),$

- *subadditive* in (y, z):

  $\forall t \in [0, T],\ \forall (y_0, z_0), (y_1, z_1) \in \mathbb{R} \times \mathbb{R},$

  $g(t, y_0 + y_1, z_0 + z_1) \leq g(t, y_0, z_0) + g(t, y_1, z_1),$

Example 1.2.5. For instance, $g(z) = \mu|z|$ with $\mu > 0$ is sublinear.

Definition 1.2.6. Functional g is *convex* in (y, z) if
$\forall t \in [0, T],\ \forall (y_0, z_0), (y_1, z_1) \in \mathbb{R} \times \mathbb{R},\ \forall \alpha \in (0, 1)$
$g(t, \alpha y_0 + (1 - \alpha) y_1, \alpha z_0 + (1 - \alpha) z_1) \leq \alpha g(t, y_0, z_0) + (1 - \alpha) g(t, y_1, z_1),$

The following proposition holds (see [25]).

Proposition 1.2.7. *Assume that the usual assumptions on g hold.*

1. *If g satisfies sublinearity, then ρ_g is a coherent risk measure.*

2. *If g satisfies convexity, then ρ_g is a convex risk measure satisfying positivity, constancy and translability.*

As mentioned in the introduction, the problem with modelling risk measures as solutions to BSDEs is that the drivers defining the underlying BSDEs are difficult to interpret. Nevertheless, we can still deduce some financial interpretations of risk measures induced by g.

First of all, g-expectation is increasing with respect to g (see Coquet et al. [7]), i.e. if g and $\overline{g}$ satisfy the usual assumptions and are both continuous in t (i.e. for any $y \in R$ and any $z \in R, g(t, y, z)$ is continuous in t):

$$g \leq \overline{g} \Leftrightarrow \mathcal{E}_g[X] \leq \mathcal{E}_{\overline{g}}[X] \; X \in L^2(\mathcal{F}_T).$$

In other words, bigger g corresponds to a more conservative risk measure ρ_g.

Furthermore, we can consider dependence of the interpretation of g on dynamics of portfolio X. For example, X might be a function of a terminal value of a stochastic process $(S_t)_{0 \leq t \leq T}$, the dynamics of which are known, i.e. $X = \Phi(S_T)$. In order to illustrate this lets consider the framework of pricing of claims in a complete market.

Assume for simplicity that $(S_t)_{0 \leq t \leq T}$ are the price dynamics of an asset given by the Black-Scholes model with drift μ, volatility σ and risk-free rate r:

$$dS_t = S_t(\mu dt + \sigma dW_t).$$

If we set $g(t, y, z) = -(yr + z\theta)$, where $\theta = \frac{\mu - r}{\sigma}$, then the corresponding static risk measure $\rho_g(X) \triangleq \mathcal{E}_g[-X]$ corresponds to the initial value of the replicating strategy of $-X$ (see El Karoui [11], among others).

1.2.3 Dynamic risk measures

Since in many situations information arrives continuously it seems natural to assume that the agent is allowed to asses his risk at any time. Static risk measures can be generalised to dynamic risk measures as we shall now see.

First we will consider a general filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$, i.e. not necessarily the Brownian one. Let $L^0(\mathcal{F}_t) = L^0(\Omega, \mathcal{F}_t, \mathbb{P})$ denote the space of all finite valued $\mathcal{F}_t$-measurable random variables and let $\mathcal{X}$ be a set of risky positions with maturity time T. For simplicity, we will suppose that $\mathcal{X} = L^p(\Omega, \mathcal{F}_T, \mathbb{P})$ for $p \in [1, \infty]$. Any element of $\mathcal{X}$ is therefore $\mathcal{F}_T$-measurable.

The random variable $\rho_t(X)$ should now represent the riskiness at time t, taking into account information available up to time t. Moreover, two boundary conditions are imposed at times 0 and T on $(\rho_t)_{t \in [0,T]}$, i.e. ρ_0 has to be a static risk measure and ρ_T has to reduce to the opposite of our risky position.

Definition 1.2.8. We call a *dynamic risk measure* any map such that:

- $\rho_t : \mathcal{X} \to L^0(\mathcal{F}_t)$, for all $t \in [0,T]$.
- ρ_0 is a static risk measure;
- $\rho_T(X) = -X$ for all $X \in \mathcal{X}$

Such map might satisfy some of the following axioms:

- dynamic convexity: $\forall t \in [0,T]$, ρ_t is convex,
- dynamic positivity: $X \geq 0 \Rightarrow \forall t \in [0,T]\ \rho_t(X) \leq \rho_t(0)$,
- dynamic constancy: $\forall t \in [0,T]$, $\forall \alpha \in \mathbb{R}$, $\rho_t(\alpha) = -\alpha$,
- dynamic translability: $\forall t \in [0,T]$, $\forall \alpha$ $\mathcal{F}_T$-measurable in $\mathcal{X}$, $\forall X \in \mathcal{X}$
  $\rho_t(X + \alpha) = \rho_t(X) - \alpha$,
- dynamic sublinearity: $\forall t \in [0,T], \forall \alpha > 0, \forall X, Y \in \mathcal{X}$,
  $\rho_t(\alpha X) = \alpha \rho_t(X),\ \rho_t(X+Y) \leq \rho_t(X) + \rho_t(Y)$.

The financial motivation of those axioms is analogous to that in the static case. Note that the translability axiom is stronger than that in the static case, since it requires translation invariance not only with respect to constants, but also with respect to any $\mathcal{F}_t$-measurable random variable.

Definition 1.2.9. A dynamic risk measure $(\rho_t)_{t\in[0,T]}$ is called

- *coherent* if it satisfies dynamic positivity, dynamic translability and dynamic sublinearity,
- *convex* if it satisfies dynamic convexity and $\rho_t(0) = 0$ for any $t \in [0,T]$,
- *time consistent* if $\forall t \in [0,T]$, $\forall X \in \mathcal{X}$, $\forall A \in \mathcal{F}_t$,
  $\rho_0(X1_A) = \rho_0(-\rho_t(X)1_A)$.

While the definitions of dynamic coherent and convex risk measures are dynamic versions of the static ones, time consistency is a new requirement. The financial interpretation of time consistency is the following. In order to quantify the riskiness of X at the initial time 0, the two approaches below are equivalent:

- using the static risk measure ρ_0 directly, i.e. computing $\rho_0(X)$,

- calculating $\rho_0(X)$ in two steps, i.e. evaluating first the riskiness $\rho_t(X)$ of X at an intermediate date t and then quantifying at time 0 the risk of $-\rho_t(X)$ through ρ_0.

Lets return from this general framework to the setup of the previous section. Consider Brownian filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ and set $\mathcal{X} = L^2(\mathcal{F}_T)$, and use the same assumptions as in previous section.

Let g satisfy the usual assumptions and set

$$\rho_t(X) \triangleq \mathcal{E}_g[-X|\mathcal{F}_t],\ \forall X \in L^2(\mathcal{F}_T),\ \forall t \in [0,T], \tag{1.2}$$

then the following proposition holds (see [25]).

Proposition 1.2.10. *$(\rho_t)_{t\in[0,T]}$ defined by (1.2) is a dynamic risk measure satisfying*

- continious-time recursivity*: for any $0 \leq s \leq t \leq T$*

  $\rho_0(X1_A) = \rho_0(-\rho_t(X)1_A)$

- *if, for some $t \in [0,T]$, $\rho_t^g(X) \leq \rho_t^g(Y) \Rightarrow \forall s \in [0,t],\ \rho_s^g(X) \leq \rho_s^g(Y)$,*

- *coherency: if g satisfies sublinearity, then $(\rho_t)_{t\in[0,T]}$ is a dynamic coherent and time consistent risk measure,*

- *convexity: if g satisfies convexity, then $(\rho_t)_{t\in[0,T]}$ is a dynamic convex and time consistent risk measure satisfying dynamic positivity, dynamic constancy and dynamic translability.*

Remark 1.2.11. The first two properties in the above proposition hold true when times s and t are replaced by the stopping times σ and τ such that $\sigma \leq \tau \leq T$.

Remark 1.2.12. In the above proposition, the second property is called time consistency by some authors (for example Epstein et al. [3]).

Chapter 2

The model

Based on the conclusions of the previous chapter we know that we can construct time consistent risk measures by choosing a suitable driver g for a BSDE

$$\begin{aligned} -dY_t &= g(t, Y_t, Z_t)dt - Z_t dW_t, \\ Y_T &= X. \end{aligned}$$

Many common measures of risk, such as Value-at-risk or Expected Shortfall, are not time consistent. However, if we apply one of these risk measures recursively on a discrete-time grid, the resultant measure is time consistent. In this chapter we will argue that various risk measures converge to the solutions of a particular type of BSDE when the time grid goes to zero.

2.1 Value-at-Risk

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, T a fixed future date and $\mathcal{X}$ the space of all financial positions in which we are interested. For instance, an element of $\mathcal{X}$ may be the net worth at maturity T of a financial contract. Time horizon T in market risk management is typically 1 or 10 days, while in credit risk it is usually one year [5].

As we know from the previous section, a static risk measure is a functional $\rho : \mathcal{X} \to \mathbb{R}$, which assigns a financial position X the quantity $\rho(X)$ that represents the riskiness of X and, by convention, X is acceptable when $\rho(X) \leq 0$ and unacceptable otherwise.

Value-at-Risk is one of the most common measures of risk, as it is part of the Basel II capital-adequacy framework. VaR_α corresponds to the smallest amount of capital a bank or an insurance needs to add to its position such that the probability of a negative outcome is kept below risk level α.

It is convenient so define the *loss* $L \triangleq -X$ of a given portfolio X. Positive values of L correspond to losses whereas negative values of L are considered to be gains. We can then define VaR in the following way.

Definition 2.1.1. *Value-at-Risk* (VaR) at risk level $\alpha \in (0, 1)$ is defined for portfolio X as

$$\mathrm{VaR}_\alpha = \inf\{l \in R : P(L > l) \leq 1 - \alpha\},$$

that is, VaR_α is the maximum loss l which will only be exceeded with small probability $1 - \alpha$.

If we know the loss distribution function $F_L(x) = P(L \leq x)$, then VaR is simply a quantile of the loss distribution. In the case when loss distribution F_L is continuous and has an inverse, then $\mathrm{VaR}_\alpha = F_L^{-1}(\alpha)$. Typical values for α are 0.95 and 0.99.

Example 2.1.2. VaR_α corresponding to a normally distributed $N(\mu, \sigma^2)$ loss $L = -X$ is

$$\mathrm{VaR}_\alpha(X) = \mu + \sigma\Phi^{-1}(\alpha),$$

where $\Phi^{-1}(x)$ is inverse cumulative distribution function (CDF) of the standard normal distribution.

One of the problems with VaR is that it does not give any indication of the scale of losses away from the quantile, i.e. VaR cannot tell the investor what he should expect under the worst scenario. With respect to the axioms from the previous chapter VaR satisfies positivity, constancy, translability as well as positive homogeneity, but not subadditivity. Non-subadditivity is probably this measure's most significant flaw as it basically means that VaR does not reward diversification (i.e. it is possible to find a portfolio $Z = X + Y$ such that $\mathrm{VaR}_{0.95}(X) + \mathrm{VaR}_{0.95}(Y) < \mathrm{VaR}_{0.95}(Z)$; for an example of such a portfolio see [17]). Since sublinearity is in fact equivalent to convexity & positive homogeneity, we can deduce that VaR is not convex either.

However, if the loss involved is normally distributed, then VaR is a coherent risk measure (and also a convex risk measure). If the loss involved is "nearly normal", it is then reasonable to believe that VaR will not be far from being coherent (convex). However, if we are considering financial derivatives, typical assets no longer have distributions which are similar to normal distributions.

Remark 2.1.3. It is easy to see that VaR is subadditive in the case of a normal distribution. Let $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ with correlation ρ, then

$$\begin{aligned} \mathrm{VaR}_\alpha(X+Y) &\overset{?}{\leq} \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y), \\ \mu_X + \mu_Y + \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y}\Phi^{-1}(\alpha) &\overset{?}{\leq} \mu_X + \mu_Y + (\sigma_X + \sigma_Y)\Phi^{-1}(\alpha), \\ \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y &\leq \sigma_X^2 + \sigma_Y^2 + 2\sigma_X\sigma_Y. \end{aligned}$$

2.1.1 Composed VaR

It is easy to construct a time consistent dynamic risk measure ρ^{com} on a discrete grid $(t_k = k\frac{T}{N})_{0\leq k\leq N}$ from a static one ρ. This is done by *composing* static (i.e. one-period) measures over time. This means that we initially set $\rho^{\mathrm{com}}_{t_{N-1}} \triangleq \rho_{t_{N-1}}$ and then continue in the recursive fashion setting for $t_k \leq t_{N-2}$:

$$\rho^{\mathrm{com}}_{t_k} \triangleq \rho_{t_k}(-\rho^{\mathrm{com}}_{t_{k+1}}).$$

Applying this method to VaR, we can construct time consistent *composed VaR* (ComVaR). Fix $\alpha \in (0,1)$ and start with $\mathrm{ComVaR}^\alpha_{t_{N-1}} \triangleq \mathrm{VaR}^\alpha_{t_{N-1}}$. For $t_k \leq t_{N-2}$ define recursively:

$$\mathrm{ComVaR}^\alpha_{t_k} \triangleq \mathrm{VaR}^\alpha_{t_k}(-\mathrm{ComVaR}^\alpha_{t_{k+1}}). \tag{2.1}$$

Then $(\mathrm{ComVaR}^\alpha_{t_k})_{0\leq k\leq N}$ is time consistent on the discrete grid by construction.

The choice of α should depend on the size of the time step, because the bigger the time step, the more conservative VaR should be. Precisely how we should scale α will become clear in the next section.

It is interesting to investigate what happens with composed VaR as the size of the discrete time grid goes to zero. We know that (almost) any dynamic coherent or convex risk measure comes from a conditional g-expectation (see Rosazza Gianin [25]). Of course, VaR is not a coherent nor convex risk measure, therefore composed VaR on a discrete time grid is not either. However, as the size of the discrete time grid goes to zero, by the martingale representation theorem we know that martingale increments can be approximated by increments of the Brownian motion, which are normally distributed. Hence it is reasonable to believe that the loss involved is locally "nearly normal" on a small time grid, and VaR used for composing in (2.1) will not be far from being coherent (convex). Therefore, as the size of the discrete time grid goes to zero we believe that resulting composed VaR (with appropriate scaling) will be coherent (convex), even though VaR itself is not coherent (convex).

Furthermore, as we know from the theory of discrete BSDEs, any time consistent nonlinear expectation corresponds to a BSDE driver of the form $g(t,y,z) = a|z|$,

where a is some constant (see [6]). Taking into account conclusions from previous paragraph, i.e. that composed VaR (with appropriate scaling) converges to coherent (convex) risk measure as time step goes to zero, we conclude that composed VaR should converge to a risk measure corresponding to a BSDE with driver of the form $g(t, y, z) = a|z|$, where a is some constant.

2.1.2 Composed VaR via g-expectations

We know that we can construct time consistent risk measures by choosing a suitable driver g for a BSDE

$$\begin{aligned} -dY_t &= g(t, Y_t, Z_t)dt + Z_t dW_t, \\ Y_T &= X. \end{aligned} \tag{2.2}$$

If we discretise this BSDE on an equidistant discrete time grid $(t_k = kh = k\frac{T}{N})_{0 \leq k \leq N}$ we get the following approximation

$$Y^N_{t_{k+1}} - Y^N_{t_k} = g(t_k, Y^N_{t_k}, Z^N_{t_k})h - Z^N_{t_k}\Delta W_{k+1}, \tag{2.3}$$

where we used the following notation $\Delta W_{k+1} = W^N_{t_{k+1}} - W^N_{t_k}$.

Suppose we wish for this BSDE to represent ComVaR. Under this numerical approximation $Y^N_{t_{k+1}}$ conditional on $\mathcal{F}_{t_k}$ is normally distributed with mean $Y^N_{t_k} + g(t_k, Y^N_{t_k}, Z^N_{t_k})h$ and standard deviation $\sqrt{h}|Z^N_{t_k}|$. Therefore, if we set $Y^N_{t_N} = X$ we can calculate $\mathrm{ComVaR}^N_{t_{N-1}}$ at time step t_{N-1} as

$$\mathrm{ComVaR}^N_{t_{N-1}} = \mathbb{E}[Y^N_{t_N}|\mathcal{F}_{t_{N-1}}] + \Phi^{-1}(\alpha)\left|Z^N_{t_{N-1}}\right|\sqrt{h}, \tag{2.4}$$

where Φ^{-1} is the inverse CDF of a standard normal distribution.

On the other hand, we can rewrite numerical scheme (2.3) as

$$\begin{aligned} &Y^N_{t_N} = X,\ Z^N_{t_N} = 0, \\ &Z^N_{t_k} = \frac{1}{h}\mathbb{E}\left[Y^N_{t_{k+1}}\Delta W_{k+1}|\mathcal{F}_{t_k}\right], \\ &Y^N_{t_k} = \mathbb{E}\left[Y^N_{t_{k+1}}|\mathcal{F}_{t_k}\right] + hg(t_k, Y^N_{t_k}, Z^N_{t_k}), \end{aligned} \tag{2.5}$$

where equality for $Z^N_{t_k}$ was acquired by multiplying (2.3) by ΔW_{k+1} and taking the conditional expectation with respect to $\mathcal{F}_{t_k}$. Equality for $Y^N_{t_k}$ is then acquired by taking the conditional expectation of equation (2.3) with respect to $\mathcal{F}_{t_k}$.

Comparing (2.4) and (2.5) we conclude, that by choosing the driver

$$g(t, y, z) = \frac{\Phi^{-1}(\alpha)}{\sqrt{h}}|z| \tag{2.6}$$

we should get the continuous time version of the Composed VaR. We could interpret $\Phi^{-1}(\alpha)/\sqrt{h}$ in the equation above as the accumulated risk aversion in $\sqrt{h}$ time units. In order to make this constant regardless of time scaling, we have to scale the denominator accordingly, which puts a restriction on the choice of α. If we associate parameter α with the length of time step h denoting it as α_h, we get the following constraint on α_h

$$\frac{\Phi^{-1}(\alpha_h)}{\sqrt{h}} = \frac{\Phi^{-1}(\alpha_T)}{\sqrt{T}}.$$

Therefore the scaling of α_h is determined by the choice of parameter α_T in the following way:

$$\alpha_h = \Phi\left(\Phi^{-1}(\alpha_T)\sqrt{\frac{h}{T}}\right).$$

As mentioned before, a typical choice for α associated with static Value-at-Risk is 0.95 or 0.99, and therefore it is reasonable to use the same value for α_T.

In Chapter 3 we will investigate numerically what happens when the size of the discrete time grid goes to zero. In particular, we will show that ComVaR converges to the solution of a BSDE with driver (2.6) as time step goes to zero.

2.2 Expected Shortfall

As mentioned previously VaR has several drawbacks. It is not subadditive, i.e. it does not reward diversification, and it does not take into account how big the losses may be in case the threshold VaR is exceeded. Because of these shortcomings Expected Shortfall or Conditional VaR was introduced, which is a coherent risk measure (see [12] for proof). Expected Shortfall is the expected loss when the VaR is exceeded. Some people call expected shortfall Tail VaR or expected tail loss. In the insurance literature, expected shortfall is called Conditional Tail Expectation or Tail Conditional Expectation.

Definition 2.2.1. *Expected shortfall* (ES) at level $\alpha \in (0,1)$ is the average Value-at-Risk for levels $\alpha \leq u \leq 1$, that is

$$\mathrm{ES}_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u du.$$

When the loss is continuous, an even easier definition is possible:

$$\mathrm{ES}_\alpha = \mathbb{E}[L|L \geq \mathrm{VaR}_\alpha].$$

Example 2.2.2. Let loss $L = -X$ be normally distributed $N(\mu, \sigma^2)$. Using $Z \sim N(0,1)$ we can then rewrite it as $L \sim \mu + \sigma Z$. ES_α corresponding to loss L is then

$$\begin{aligned}\mathrm{ES}_\alpha(X) &= \mathbb{E}[L|L \geq \mathrm{VaR}_\alpha] \\ &= \mathbb{E}[\mu + \sigma Z|\mu + \sigma Z \geq \mu + \sigma\Phi^{-1}(\alpha)] \\ &= \mu + \sigma\mathbb{E}[Z|Z \geq \Phi^{-1}(\alpha)] \\ &= \mu + \frac{\sigma}{(1-\alpha)\sqrt{2\pi}} \exp\left(-\frac{\Phi^{-1}(\alpha)^2}{2}\right).\end{aligned}$$

As for VaR, we may compose ES recursively in the same way we composed VaR to obtain time consistent Expected Shortfall. Fix $\alpha \in (0,1)$ and define $\mathrm{ComES}^\alpha_{t_{N-1}} \triangleq \mathrm{ES}^\alpha_{t_{N-1}}$. For $t_k \leq t_{N-2}$ define recursively:

$$\mathrm{ComES}^\alpha_{t_k} \triangleq \mathrm{ES}^\alpha_{t_k}(-\mathrm{ComES}^\alpha_{t_{k+1}}).$$

Then $(\mathrm{ComVaR}^\alpha_{t_k})_{0\leq k\leq N}$ is time consistent by construction and it inherits all the coherency properties from ES. Since (almost) any dynamic coherent or convex risk measure comes from a conditional g-expectation, we expect composed ES to converge to a risk measure corresponding to a BSDE as the size of the discrete time grid goes to zero. Furthermore, because we know from the theory of discrete BSDEs that any time consistent nonlinear expectation corresponds to a BSDE driver of the form $g(t,y,z) = a|z|$, where a is some constant, we expect composed ES (with appropriate scaling) to converge to a coherent (convex) risk measure corresponding to a BSDE with a driver of the form $g(t,y,z) = a|z|$, where a is some constant.

2.2.1 Composed ES via g-expectations

As for ComVar, we seek a BSDE representation of ComES. Recall discretisation (2.3) of the BSDE (2.2). Under that numerical approximation $Y_{t_{k+1}}$ conditional on $\mathcal{F}_{t_k}$ is normally distributed with mean $Y^N_{t_k} + g(t_k, Y^N_{t_k}, Z^N_{t_k})h$ and standard deviation $\sqrt{h}|Z^N_{t_k}|$. Therefore, if we set $Y^N_{t_N} = X$ we can calculate $\mathrm{ComES}^N_{t_{N-1}}$ at time step t_{N-1} as

$$\mathrm{ComES}^N_{t_{N-1}} = \mathbb{E}[Y^N_{t_N}|\mathcal{F}_{t_{N-1}}] + \frac{1}{(1-\alpha)\sqrt{2\pi}} \exp\left(-\frac{\Phi^{-1}(\alpha)^2}{2}\right)|Z^N_{t_{N-1}}|\sqrt{h}, \tag{2.7}$$

where Φ^{-1} is the inverse CDF of a standard normal distribution.

Comparing (2.7) and (2.5) we conclude, that by choosing the driver

$$g(t,y,z) = \frac{1}{(1-\alpha)\sqrt{2\pi h}} \exp\left(-\frac{\Phi^{-1}(\alpha)^2}{2}\right)|Z^N_{t_k}| \tag{2.8}$$

we should get the continuous time version of the Composed ES. We could interpret $\exp\left(-0.5\Phi^{-1}(\alpha)^2\right)/((1-\alpha)\sqrt{2\pi h})$ in the equation above as the accumulated risk aversion in $\sqrt{h}$ time units. In order to make this constant regardless of time scaling, we have to scale α accordingly.

If we associate parameter α with the length of time step h denoting it as α_h, we get the following constraint on α_h

$$\frac{1}{(1-\alpha_h)\sqrt{2\pi h}}\exp\left(-\frac{\Phi^{-1}(\alpha_h)^2}{2}\right) = \frac{1}{(1-\alpha_T)\sqrt{2\pi T}}\exp\left(-\frac{\Phi^{-1}(\alpha_T)^2}{2}\right) \tag{2.9}$$

Thankfully this equation has a unique solution.

Lemma 2.2.3. *There is a unique rate $\alpha_h \in [0,1]$, which satisfies equation (2.9). Furthermore, such α_h is increasing in h.*

Proof. To show that this has a unique solution depending on parameter α_T we group all the constants on the right hand side:

$$\frac{1}{1-\alpha_h}\exp\left(-\frac{\Phi^{-1}(\alpha_h)^2}{2}\right) = \sqrt{\frac{h}{T}}\frac{1}{1-\alpha_T}\exp\left(-\frac{\Phi^{-1}(\alpha_T)^2}{2}\right) \tag{2.10}$$

Since right hand side is a positive constant, by showing that left hand side is a bijection between interval $[0,1]$ and $[0,\infty]$ we will show that the equation above has a unique solution.

Denote the left hand side of the equation (2.10) as $f(x)$. Since $f(0) = 0$ and $f(1) = \infty$ it is sufficient to show that f is monotone. The function f has first derivative

$$f'(x) = \frac{\exp(-\Phi^{-1}(x)^2/2)/\sqrt{\pi} + \sqrt{2}(x-1)\Phi^{-1}(x)}{\sqrt{2}(x-1)^2}.$$

Since its denominator is positive, we will show that its numerator is positive. Denoting the numerator as $h(x)$ and noting that $h(0) = \infty$ and $h(1) = 0$, we calculate its derivative,

$$h'(x) = 2\sqrt{\pi}(x-1)\exp(\Phi^{-1}(x)^2/2),$$

and note that it is negative on interval $[0,1]$. Hence f' is positive and so f is a bijection. □

In Chapter 3 we will investigate numerically what happens when the size of the discrete time grid goes to zero. In particular, we will show that ComES converges to the solution of a BSDE with driver (2.8) as time step goes to zero.

2.3 Research field overview

It is interesting to compare our conclusions with conclusion of Stadje [26]. Stadje assumes that the filtration in discrete time is generated by i.i.d. Bernoulli random variables. He then constructs dynamic risk measures from rescaled one-period risk measures, using a random walk converging to a Brownian motion. Finally, he obtains, under suitable conditions, convergence of the discrete risk measures to the solution of a BSDE. He derives limiting BSDE driver for the semi-deviation principle, VaR, ES and the Gini principle in closed form. He concludes that the limits depend on the specific discrete time markets and the specific discrete time filtrations.

For example, in a zero-one market where in every step there are only two possible scenarios (the price going up or down) each having probability 1/2, VaR in every step gives a lower possible price for every risk level α above 1/2. Hence, no matter how α is chosen, the continuous time risk measure is always the same. Consequently, the limit for a risk measure like VaR depends on the underlying discrete time market. He notes, that in a market in which increments are normally distributed different risk levels would also lead to different limits.

Our approach does not depend on a discrete time random walk. Therefore, while we also arrived at a conclusion that by composing discrete time VaR (with suitable scaling) we indeed obtain a limit corresponding to the solution of a BSDE, our solutions *do* depend on the initial choice of risk level α.

In this way, our approach gives a more natural extension of discrete time risk measures to continuous time. We think that this modelling approach is more appropriate for risk measures like VaR, since in Stadje's model limit for a risk measure depends on the underlying discrete time market.

Chapter 3

Numerical implementation

In order to solve BSDEs described in the previous section we have to numerically approximate the solution to a BSDE. There are several approaches presented in the literature. For instance, discretisation schemes based on an approximation of the Brownian motion, where the Brownian motion is replaced by a scaled random walk, has been analysed by Ma et al. [19] and Briand et al [4].

In contrast with that approach, we will use a method developed by Gobet et al. [15] which approximates the solution to a BSDE based on Monte Carlo approximation of the underlying Brownian motion. It is a simple yet efficient approach, based on Monte Carlo regression on function bases. The regression operator results from the L^2-projection on a finite basis of functions, which leads in practise to solving a least squares problem. We should also mention that this approach has been initially developed by Longstaff-Schwartz [18] for the pricing of Bermuda options.

Using this method we will confirm on some specific examples convergence of ComVaR and ComES to the solutions of a BSDE with corresponding drivers.

3.1 Regression-based Monte Carlo method to solve BSDEs

In this section we are interested in numerically approximating the solution of a decoupled forward backward stochastic differential equation (FBSDE henceforth)

$$\begin{aligned} dS_t &= \mu(t, S_t)dt + \sigma(t, S_t)dW_t \\ -dY_t &= g(t, S_t, Y_t, Z_t)dt - Z_t dW_t \end{aligned} \tag{3.1}$$

with the initial condition $S_0 = s$ at time $t = 0$ for the forward component and the terminal condition $Y_T = \Phi(S_T)$ at time $t = T$ for the backward component. Here, $W = \{W_t, t \geq 0\}$ is a Brownian motion defined on a filtered probability space

$(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{0\leq t\leq T})$, where $((\mathcal{F}_t)_{0\leq t\leq T})$ is the augmented natural filtration of W. The coefficients of the system (3.1) are given by deterministic functions $\mu : [0,T] \times \mathbb{R} \to \mathbb{R}, \sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ for the forward component and $g : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for the backward component. Function $\Phi : \mathbb{R} \to \mathbb{R}$ is also a real valued function. A solution to the forward backward system is the triplet (S_t, Y_t, Z_t).

The following assumptions are sufficient to ensure the existence and uniqueness of a triplet (S, Y, Z) solution to (3.1) (see [20] and references therein):

- The function $(t,x) \to \mu(t,s)$ and $(t,s) \to \sigma(t,x)$ are uniformly Lipshitz continious w.r.t. $(t,s) \in [0,T] \times \mathbb{R}$.

- The driver g satisfies the following continuity estimate:

  $$|g(t_2,s_2,y_2,z_2) - g(t_1,s_1,y_1,z_1)| \leq C_g(|t_2 - t_1|^{1/2} + |s_2 - s_1| + |y_2 - y_1| + |z_2 - z_1|).$$

  for any $(t_1,s_1,y_1,z_1), (t_2,s_2,y_2,z_2) \in [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. The terminal condition Φ is a Lipshitz function of the state process $X = \Phi(S_T)$ with the same Lipshitz constant as above.

The forward component of (3.1) can be easily approximated using a standard Euler scheme with time step h associated with the equidistant discrete times grid $(t_k = kh = k\frac{T}{N})_{0\leq k\leq N}$. This approximation is defined by $S_0^N = S_0$ and

$$S_{t_{k+1}}^N = S_{t_k}^N + \mu(t_k, S_{t_k}^N)h + \sigma(t_k, S_{t_k}^N)(W_{t_{k+1}} - W_{t_k}).$$

We will henceforth denote discretisation of the Brownian motion as $\Delta W_{k+1} = W_{t_{k+1}} - W_{t_k}$.

The backward component of the FBSDE (3.1) is evaluated in a backward manner. First, we set $Y_{t_N}^N = \Phi(S_{t_N}^N)$. Then, $(Y_{t_k}^N, Z_{t_k}^N)_{0\leq k\leq N-1}$ are defined by

$$Z_{t_k}^N = \frac{1}{h}\mathbb{E}\left[Y_{t_{k+1}}^N \Delta W_{k+1} | \mathcal{F}_{t_k}\right], \tag{3.2}$$

$$Y_{t_k}^N = \mathbb{E}\left[Y_{t_{k+1}}^N | \mathcal{F}_{t_k}\right] + hg(t_k, S_{t_k}^N, Y_{t_k}^N, Z_{t_k}^N), \tag{3.3}$$

It is also equivalent to assert that $(Y_{t_k}^N, Z_{t_k}^N)$ minimise the quantity

$$\mathbb{E}\left[(Y_{t_{k+1}}^N - Y + hg(t_k, X_{t_k}^N, Y, Z) - Y\Delta W_{k+1})^2\right] \tag{3.4}$$

over $L^2(\mathcal{F}_{t_k})$ random variables (Y, Z). Note that Y_{t_k} in 3.3 is well defined, because the mapping $Y \to \mathbb{E}[Y_{t_{k+1}}^N | \mathcal{F}_{t_k}] + hg(t_k, X_{t_k}^N, Y, Z_{t_k}^N)$, is a contraction in $L^2(\mathcal{F}_{t_k})$ for h small enough.

Since in our case the terminal conditional is dependent only on the terminal value of the process S_t, the random variables $Y_{t_k}^N$, $Z_{t_k}^N$ are deterministic functions of the underlying $S_{t_k}^N$ and the above conditional expectations are actually regressions

$$\mathbb{E}\left[Y_{t_{k+1}}^N \Delta W_{k+1} | \mathcal{F}_{t_k}\right] = \mathbb{E}\left[Y_{t_{k+1}}^N \Delta W_{k+1} | S_{t_k}^N\right],\ \mathbb{E}\left[Y_{t_{k+1}}^N | \mathcal{F}_{t_k}\right] = \mathbb{E}\left[Y_{t_{k+1}}^N | S_{t_k}^N\right].$$

We can therefore replace $Y_{t_k}^N$ and $Z_{t_k}^N$ as they appear in the definitions (3.3) and(3.2) by a $L^2(\Omega, \mathcal{P})$ projection on a finite basis of functions. Specifically, denoting values of basis functions evaluated at $S_{t_k}^N$ with $p_k = p_k(S_{t_k}^N)$, we can represent $(Y_{t_k}^N, Z_{t_k}^N)$ at a given time t_k via some projection coefficents (α_k, β_k) by

$$Y_{t_k}^N = \alpha_k \cdot p_k,\ Z_{t_k}^N = \beta_k \cdot p_k.$$

See section 3.1.2 for a specific example of a function base.

Approximation of $Y_{t_k}^N$ in (3.3) is usually obtained as a fixed point, which can be done by combining the projection on the function basis and I Picard iterations. According to [15] a fixed parameter $I = 3$ is a good choice.

We will denote by $Y_{t_k}^{N,i,I}$ the approximation of $Y_{t_k}^N$, where i Picard iterations with projections have been performed at time t_k and I Picard iterations with projections at any time after t_k. $Z_{t_k}^{N,i,I}$ is defined analogously. Approximation $(Y_{t_k}^{N,i,I}, Z_{t_k}^{N,i,I})$ corresponds to its projection $(\alpha_k^{i,I} \cdot p_k, \beta_k^{i,I} \cdot p_k)$.

The iteration proceeds in the following way. We set $Y_{t_N}^N = \Phi(S_{t_N}^N)$, independently of i and I. Assume that $Y_{t_{k+1}}^{N,I,I}$ is obtained and define $Y_{t_k}^{N,i,I}, Z_{t_k}^{N,i,I}$ for $i = 0, \ldots, I$. Begin with $Y_{t_k}^{N,0,I} = 0$ and $Z_{t_k}^{N,0,I} = 0$ corresponding to $\alpha_k^{0,I} = 0$ and $\beta_k^{0,I} = 0$. By analogy with (3.4), set $\alpha_k^{i,I}$ and $\beta_k^{i,I}$ as the arg min in (α, β) of the quantity

$$\mathbb{E}\left[(Y_{t_{k+1}}^{N,I,I} - \alpha \cdot p_k + hg(t_k, S_{t_k}^N, \alpha_k^{i-1,I} \cdot p_k, \beta_k^{i-1,I} \cdot p_k) - \beta \cdot p_k \Delta W_{k+1})^2\right].$$

Iterating with $i = 1, \ldots, I$, at the end we get $(\alpha_k^{I,I}, \beta_k^{I,I})$, thus, $Y_{t_k}^{N,i,I} = \alpha_k^{I,I} \cdot p_k$ and $Z_{t_k}^{N,i,I} = \beta_k^{I,I} \cdot p_k$. In practise, the above is a least squares problem.

3.1.1 Numerical scheme

Based on the above ideas the following simulation-based approximation of the BSDE (3.1) can be defined. The procedure combines a backward in time evaluation (from time $T_N = T$ to time $t_0 = 0$), a fixed point argument (using $i = 1, 2, \ldots, I$ Picard iterations), and a least squares problem on M simulated paths (using some function bases).

In the following, M independent simulations of $(S^N_{t_k})_{0\le k\le N}$, $(\Delta W_k)_{1\le k\le N}$ are used. We denote them as $((S^{N,m}_{t_k})_{0\le k\le N})_{1\le m\le M}$, $((\Delta W^m_k)_{1\le k\le N})_{1\le m\le M}$. The values of basis functions along these simulations are denoted $(p^m_k = p_k(S^{N,m}_{t_k}))_{1\le k\le N-1, 1\le m\le M}$ and to simplify we write $f^m_k(\alpha_k, \beta_k)$ for $f(t_k, S^{N,m}_{t_k}, \alpha_k \cdot p^m_k, \beta_k \cdot p^m_k)$.

Initialisation. The algorithm is initialised with $Y^{N,i,I,M}_{t_N} = \Phi^N(S^N_{t_N})$. Then the solution (Y_{t_k}, Z_{t_k}) at a given time t_k is represented via some projection coefficients $\alpha^{i,I,M}_k$ and $\beta^{i,I,M}_k$ by

$$Y^{N,i,I,M}_{t_k} = \alpha^{i,I,M}_k \cdot p_k,\ Z^{N,i,I,M}_{t_k} = \beta^{i,I,M}_k \cdot p_k.$$

Backward in time iteration at time $t_k < T$. Assume that an approximation $Y^{N,I,I,M}_{t_{k+1}} \triangleq \alpha^{I,I,M}_{k+1} \cdot p_{k+1}$ is built and denote $Y^{N,I,I,M,m}_{t_{k+1}} = \alpha^{I,I,M}_{k+1} \cdot p^m_{k+1}$ its realisation along the mth simulation.

1. For the initialisation $i = 0$ of Picard iterations, set $Y^{N,0,I,M}_{t_k} = 0$ and $Z^{N,0,I,M}_{t_k} = 0$, that is, $\alpha^{0,I,M}_k = 0$ and $\beta^{0,I,M}_k = 0$.

2. For $i = 1, 2, \dots, I$, the coefficients $\alpha^{i,I,M}_k$ and $\beta^{i,I,M}_k$ are iteratively obtained as the arg min in (α, β) of the quantity

$$\frac{1}{M}\sum_{m=1}^{M}\left(Y^{N,I,I,M,m}_{t_{k+1}} - \alpha \cdot p^m_k + h f^m_k(\alpha^{i-1,I,M}_k, \beta^{i-1,I,M}_k) - \beta \cdot p^m_k \Delta W^m_{k+1}\right).$$

If the above least squares problem has multiple solutions, we may choose, for instance, the unique solution of a minimal norm.

The convergence parameters of this scheme are the time step h ($h \to 0$), the function bases and the number of simulated paths M ($M \to \infty$). This is fully analysed in [15].

3.1.2 Numerical experiment

To check the accuracy of the method, a numerical example is given next. Assume for simplicity that $(S_t)_{0\le t\le T}$ is the price dynamics of an asset given by the Black-Scholes model with drift μ, volatility σ and risk-free rate r:

$$dS_t = S_t(\mu dt + \sigma dW_t).$$

For driver $g(t, y, z) = -(yr + z\theta)$, where $\theta = \frac{\mu - r}{\sigma}$, Y_t is the time t value of a self-financing portfolio replicating the payoff $\Phi(S_T)$ (see El Karoui [11], among others).

For this specific example we will choose the following parameters

$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$	$\boldsymbol{r}$	$\boldsymbol{T}$	$\boldsymbol{S_0}$	$\boldsymbol{K}$
0.06	0.2	0.04	0.5	100	100

.

Here K is the strike of the call option $\Phi(S_T) = (S_T - K)^+$. The exact price Y_0 is given by the Black-Scholes formula: 6.63 (see [27], among others).

One of the possible choices of a function base in the numerical scheme defined in section 3.1.1 is a global polynomial base (see [15] for examples of some non-polynomial bases, such as the hypercube base and Voronoi partition base), in which case we define p_k as the polynomial base of degree less than d. A specific example of global polynomial base is for instance Laguerre or Hermite polynomial base (see [18] or [16] respectively). For this numerical experiment we will use Hermite polynomial base of degree 9.

We test the accuracy of the above algorithm by launching it for different values of M, the number of Monte Carlo simulation paths, and different values of N, the number of time steps. More precisely, for each choice of M and N, we launch the algorithm 50 times and each time collect the value $Y_{t_0}^{N,I,I,M}$, The set of collected values is denoted $(Y_{t_0,i}^{N,I,I,M})_{1\leq i\leq 50}$. Then, we compute the empirical mean $\overline{Y}_{t_0}^{N,I,I,M} = \frac{1}{50}\sum_{i=1}^{50} Y_{t_0,i}^{N,I,I,M}$ and the empirical standard deviation

$$\sigma_{t_0}^{N,I,I,M} = \sqrt{\frac{1}{49}\sum_{i=1}^{50}\left|Y_{t_0,i}^{N,I,I,M} - \overline{Y}_{t_0}^{N,I,I,M}\right|^2}.$$

These two statistics provide an insight into the accuracy of the method.

Results $\overline{Y}_{t_0}^{N,I,I,M}$ are given in the table below with $\sigma_{t_0}^{N,I,I,M}$ in parenthesis. Clearly, $\overline{Y}_{t_0}^{N,I,I,M}$ converges towards 6.63.

$\boldsymbol{M}$	$\boldsymbol{N = 2500}$	$\boldsymbol{N = 5000}$	$\boldsymbol{N = 10000}$
64	6.54 (0.02)	6.55 (0.01)	6.56 (0.01)
128	6.56 (0.01)	6.58 (0.01)	6.59 (0.01)
256	6.59 (0.01)	6.60 (0.01)	6.60 (0.01)
512	6.60 (0.01)	6.61 (0.01)	6.62 (0.01)
1024	6.61 (0.01)	6.62 (0.01)	6.62 (0.01)

After plotting these result it becomes apparent that convergence is sublinear in M.

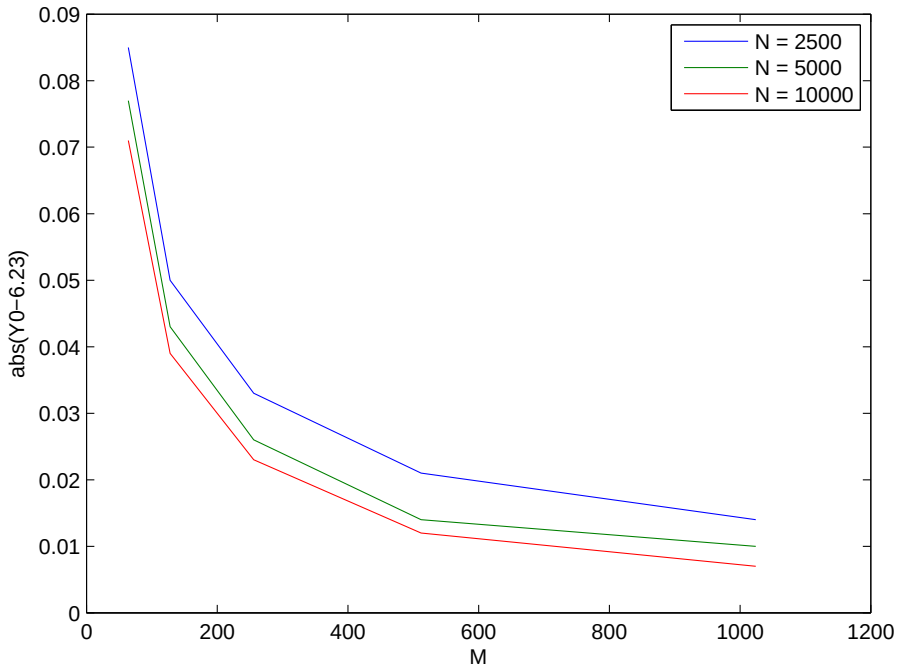


Figure 1: Convergence rate of call option price

3.2 Approximation results

We will use the same setup as in the previous section, i.e. we will assume that $(S_t)_{0\leq t\leq T}$ are the price dynamics of an asset given by the Black-Scholes model with drift μ, volatility σ and risk-free rate r: $dS_t = S_t(\mu dt + \sigma dW_t)$. We will use following parameters:

$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$	$\boldsymbol{r}$	$\boldsymbol{T}$	$\boldsymbol{S_0}$
0.06	0.2	0.04	0.5	100

.

We will investigate on some specific examples convergence of ComVaR and ComES to the solutions of a BSDE with corresponding drivers as time step goes to zero. To do this we will recursively calculate ComVaR and ComES using rescaled risk level α on a discrete-time grid as described in Chapter 2. We will denote risk measures corresponding to respective BSDE drivers as gVaR and gES (underlying the fact that those are g-expectaions).

We use Hermite polynomial base of degree 9 for all the computations in this section. We simulate ComVaR and ComES recursively on a discrete time grid, interpolating at each step using same base as we use for approximation of a BSDE. We test the accuracy of both algorithms (i.e. algorithm for approximating solution to a BSDE and algorithm to calculate composed risk measures) by launching them for 10000 Monte Carlo simulation paths and different values of N, i.e. the number of

time steps. For each choice of N, we launch 50 times each algorithm and collect each time the value returned by each algorithms. We then compute both empirical means and the empirical standard deviations. Empirical means are reported in tables below with standard deviations given in parenthesis. These two statistics provide an insight into the accuracy of both methods.

3.2.1 Non-smooth payoff

Let $K = 100$ be the strike of a call option and consider holding a portfolio consisting of a short position in such a call option. Terminal value such portfolio is then:

$$X \triangleq -(S_T - K)^+.$$

Corresponding dynamic risk measures via g-expectations are then calculated as:

$$\rho_g(X) \triangleq \mathcal{E}_g(-X)$$

Calculating comVaR (composed VaR), gVaR (VaR given by a BSDE with driver (2.6)), comES (composed ES) and gES (ES given by a BSDE with driver (2.8)) gives us following results for different number of time steps N:

N	ComVaR	gVaR	ComES	gES
32	29.9 (0.12)	29.8 (0.22)	37.6 (0.19)	37.3 (0.47)
64	29.9 (0.12)	30.3 (0.49)	38.1 (0.68)	38.6 (0.92)
128	29.8 (0.06)	30.3 (0.40)	38.7 (0.94)	38.9 (1.01)
256	29.8 (0.09)	30.5 (0.27)	39.0 (1.04)	38.7 (1.10)
512	29.7 (0.11)	31.3 (0.86)	48.4 (6.65)	42.8 (3.84)

Convergence is not as accurate as one would hope, nevertheless these estimates are well within standard deviation distance of each other and they behave particularly well when there is a small number of time steps. As the number of time steps increase, the error accumulates and standard deviation grows out of proportions.

The problem seems to be twofold. On one hand there is accumulated error from the projections, which is seen as size of time step goes to zero, and on the other hand there is a lack of smoothness in the terminal payoff. The problem with call payoff is that the function is not analytic, and the flat section creates an error which doesn't propagate well during backward iteration. This particular phenomenon is called Runge's phenomenon in numerical analysis, and it is a problem of oscillation at the edges of an interval that occurs when using polynomial interpolation with polynomials of high degree.

We should also take into account the error due to the algorithm used to solve a BSDE. Numerical approximation of BSDEs is a difficult problem and a very active area of research, with methods being far from mature. To reassure ourselves that poor behaviour is mostly due to a numerical error, we will explore a smooth payoff in next section, which gives far more stable results.

3.2.2 Smooth payoff

Consider a portfolio with following terminal value:

$$X \triangleq -\log(S_T + 1).$$

Corresponding dynamic risk measures via g-expectations are then calculated as:

$$\rho_g(X) \triangleq \mathcal{E}_g(-X)$$

Calculating comVaR (composed VaR), gVaR (VaR given by a BSDE with driver (2.6)), comES (composed ES) and gES (ES given by a BSDE with driver (2.8)) gives us following results for different number of time steps N:

N	ComVaR	gVaR	ComES	gES
32	4.87	4.87	4.93	4.92
64	4.87	4.87	4.93	4.92
128	4.88	4.87	4.93	4.92
256	4.88	4.87	4.92	4.92
512	4.88	4.87	4.93	4.92

The reason why there is no standard deviation given in parenthesis is due to its insignificant order of magnitude and general stability of the method for smooth payoffs.

In this case convergence is much better, which is consistent with our expectations.

Chapter 4

Conclusions

In this thesis we have presented an approach to construct time consistent dynamic risk measures from rescaled one-period risk measures. We arrived at a conclusion that by composing one-period time risk measure (with suitable scaling) we obtain a limit corresponding to the solution of a BSDE with a particular driver.

For composed versions of risk measures Value-at-Risk and Expected Shortfall, we have proposed particular drivers of corresponding BSDEs. We investigated numerically what happens to composed VaR and composed ES when the size of the discrete time grid goes to zero. In particular, we have presented an example with smooth terminal payoff when both composed VaR and composed ES converge to the solutions of BSDEs with proposed drivers. In the case of a non-smooth payoff (i.e. call option) the results where not conclusive due to numerical errors associated with methods used. It is interesting to note that resulting composed VaR (with apropriate scaling) converges to a coherent (and convex) risk measure as the size of the discrete time grid goes to zero, even though VaR itself is not coherent (and convex).

Solutions corresponding to those BSDEs depend on the initial choice of risk level α, which gives a natural extension of discrete time risk measures to continuous time. In this way, our approach gives a more natural extension of discrete time risk measures to continuous time, than the approach presented by Stadje [26], where model limit for a risk measure depends on the underlying discrete time market.

Appendix A

MATLAB implementation

A.1 BSDE solver

The following code approximates the BSDE solution using the regression based MC method described in Chapter 3.1.1. Input arguments are `R` (polynomial degree), `M` (number of Monte Carlo paths of forward process), `N` (number of time steps), `T` (terminal time) and options `polBase` (`lag` (laguerre poly base), `her` (hermite poly base)) and `opt call` (driver for the risk-neutral price of an option), `var` (driver for VaR), `es` (driver for ES)). The last two options give user a choice of polynomial basis and BSDE driver, respectively. The algorithm returns the estimate of Y_0^N.

```
function est = BSDESolver(R,M,N,T,polBase,opt)

    [X,W]=ForwardProcess(N+1,M,T,100);
    YT=Phi(X(N+1,:));

    h=T/N;
    Y=zeros(N+1,M);
    Z=zeros(N,M);
    Beta=zeros(N,R+1);
    Alpha=zeros(N,R+1);
    k=2; r=0.5; I = 3;
    Y(N+1,:)=YT;

    for m=N:-1:1
        p=base(R,X(m,:), k*((m+1)/T)^r,polBase)';
        for i= 1:I;
            b = Y(m+1,:) + h*F(m*h,X(m,:),Alpha(m,:)*p',
                                Beta(m,:)*p',opt,T);
            A = [p,repmat(W(m+1,:)',1,R+1).*p];
            coeff = A\b';
```

```
21              Alpha(m,:) = coeff(1:R+1);
22              Beta(m,:) = coeff(R+2:end);
23          end
24          Y(m,:)=(Alpha(m,:)*p');
25          Z(m,:)=(Beta(m,:)*p');
26      end
27
28      est = Y(1,1);
29
30  end
```

The above code is fairly straightforward and uses several additional functions which will be described later. It first simulates paths of the forward process (3) and defines the final payoff (4). On lines (6)-(10) some self-explanatory variables are introduced. On line (11) scaling rates of polynomials are set (see function `base` later) and on line (12) terminal values of the BSDE are set for the variable Y_T^N. In the loop on lines (14) through (26) we process backwards from terminal time T finding the coefficients for the basis functions as described in Chapter 3.1.1. Finally, estimate Y_0^N is reported.

A.1.1 Additional functions

Generating forward process. The function below simulates M paths of the forward process $X_{t_k}^N$ using Euler scheme starting at initial value x. It uses two additional functions `mu` and `sigma` which return coefficients of the forward process $X_{t_k}^N$.

```
1  function [ X,dW ] = ForwardProcess(N,M,T,x)
2      h=T/N;
3      X=zeros(N,M);
4
5      NormInc=randn(N,M);
6      NormInc(1,:)=zeros(1,M);
7      dW=sqrt(h)*NormInc;
8
9      X(1,:) = x;
10     for m=2:N
11         X(m,:) = X(m-1,:) + mu((m-2)*h,X(m-1,:))*h +
12                  sigma((m-2)*h,X(m-1,:)).*dW(m,:);
13     end
14 end
```

Drift of forward process.

```
1  function [outmu]= mu(t, s)
2      outmu=0.06*s;
```

```
3 end
```

Volatility of forward process

```
1 function [outsigma]= sigma(t, s)
2     outsigma=0.2*s;
3 end
```

BSDE terminal function.

```
1 function [outphi]= Phi(XT)
2     outphi= max(XT-100,0);
3 end
```

Evaluating polynomial basis. The following code returns a $k+1$ by `length(x)` matrix with the first k base polynomials evaluated at x/d, multiplied by exponential weighting. It supports two different polynomial bases, Laguerre and Hermite polynomial bases, through the choice of parameter `opt`, accepting arguments `lag` (Laguerre polynomial base) or `her` (Hermite polynomial base).

```
1  function [ L ] = base(k,x,d,opt)
2      x=x/d;
3      L=zeros([k+1, length(x)]);
4      L(1,:)=1;
5      w=exp(-0.5*abs(x));
6
7      switch opt
8           case 'lag'
9             if k>=2
10                L(2,:)=(1-x);
11            end
12            for j=3:(k+1)
13                L(j,:) = w.*((2*(j-2)+1-x).*L(j-1,:)-
14                          (j-2)*L(j-2,:))/(j-1);
15            end
16          case 'her'
17            if k>=2
18                 L(2,:)=x;
19            end
20            for j=3:(k+1)
21                L(j,:) = w.*(x.*L(j-1,:)-(j-1)*L(j-2,:));
22            end
23     end
24 end
```

BSDE driver. Given a specific value of parameter `opt` a driver for the BSDE is returned, solution of which is the risk-neutral price of a call option (`call`), driver for continuous VaR (`call`), or the driver for continuous ES (`es`)

```
function [outF]= F(t,x,y,z,opt,T)
    mu = 0.06; sigma = 0.2; r = 0.04; R = 0.06;

    switch opt
      case 'call'
        theta = (mu-r)/sigma;
        outF=-(y*r + z*theta);
      case 'var'
        outF = norminv(0.95,0,1)*abs(z)/sqrt(T);
      case 'es'
        outF = exp(-norminv(0.95,0,1)^2/2)/
               (sqrt(2*pi*T)*0.05)*abs(z);
    end
end
```

A.2 Composed VaR and composed ES

The following code approximates composed VaR or ES along the simulated underlying asset paths. Input arguments are `R` (polynomial degree), `M` (number of Monte Carlo paths of forward process), `N` (number of time steps), `T` (terminal time) and options `polBase` (`lag` (laguerre poly base), `her` (hermite poly base)) and `opt` (`var` (calculates composed VaR), `es` (calculates composed ES)). The last two options give the user a choice of polynomial basis and what risk measure the algorithm should return. The algorithm returns estimate of the risk measure at the initial time.

```
function est = ComMeasure(R,M,K,N,T,polBase,opt)

    [X,W]=ForwardProcess(N+1,M,T,100);
    YT=Phi(X(N+1,:));

    h=T/N;
    Y=zeros(N+1,M);
    Alpha=zeros(N,R+1);
    Y(N+1,:)=YT;
    switch opt
      case 'var'
        alpha = normcdf(norminv(0.95,0,1)*sqrt(h/T));
      case 'es'
```

```
14          f = @(x) scale(x,h,T);
15          alpha = fminbnd(f,0,1);
16      end
17      k=2; r=0.5;
18
19      for m=N:-1:1
20          Paths=ForwardStepMat(K,m,h,X(m,:));
21          Values = Paths.*0;
22          switch opt
23            case 'var'
24              if m == N
25                  Values = Phi(Paths);
26              else
27                  for i=1:M
28                      p = base(R,Paths(:,i)',
29                          k*((m+2)/T)^r,polBase)';
30                      Values(:,i) = (Alpha(m+1,:)*p')';
31                  end
32              end
33              Y(m,:) = quantile(Values, alpha);
34              case 'es'
35                  if m == N
36                      Values = Phi(Paths);
37                      quantiles = quantile(Values, alpha);
38                      es = zeros(M,1);
39                      parfor i=1:M
40                          tmp = Values(Values(:,i)>=
41                                       quantiles(i),i);
42                          es(i) = sum(tmp)/length(tmp);
43                      end
44                  else
45                      a = Alpha(m+1,:);
46                      es = zeros(M,1);
47                      parfor i=1:M
48                          p = base(R,Paths(:,i)',
49                              k*((m+2)/T)^r,polBase)';
50                          vals = (a*p')';
51                          VaR = quantile(vals, alpha);
52                          tmp = vals(vals>=VaR);
53                          es(i) = sum(tmp)/length(tmp);
54                      end
55                  end
56                  Y(m,:) = es;
57          end
58          A = base(R,X(m,:), k*((m+1)/T)^r,polBase)';
```

```
59          Alpha(m, :)=(A\Y(m,:)')';
60          Y(m,:)=(Alpha(m,:)*A');
61      end
62
63      est = Y(1,1);
64
65  end
```

The above code is fairly straightforward and uses a couple of additional functions. It first simulates several paths of the forward process (3) and defines the terminal payoff (4). On lines (6)-(9) some self-explanatory variables are introduced and terminal values for the composed risk measure Y_T^N are set. On lines (10) through (16) risk level α is rescaled depending on the time scale used as explained in Chapters 2.1.2 and 2.2.1. On line (11) scaling rates of polynomials are set (see function `base` later). In the loop on lines (19) through (61) we process backwards from terminal time T finding the coefficients for the projection of composed risk measures VaR and ES on basis functions. Finally, estimate Y_0^N is reported.

A.2.1 Additional functions

Simulating paths from each point. This functions performs K one time step long Monte Carlo simulations of forward process X starting at x, where x can be a vector, in which case function `ForwardStepMat` performs K simulations from each point in x. A matrix of simulated values is returned with each column representing paths from one of those points in x. Other required input data is the time increment (`h`) and the current time step (`m`, i.e. current time value is mh).

```
1   function X = ForwardStepMat(K,m,h,x)
2       X=repmat(x,K,1);
3
4       NormInc=randn(K,length(x));
5       dW=sqrt(h)*NormInc;
6       muVal = repmat(mu(m*h,x),K,1);
7       sigmaVal = repmat(sigma(m*h,x),K,1);
8
9       X = X + muVal*h +  sigmaVal.*dW;
10  end
```

Bibliography

[1] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Thinking coherently. *RISK*, 10:68–71, 1997.

[2] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[3] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, and H. Ku. Coherent multiperiod risk adjusted values and bellmans principle. *Annals of Operations Research*, in press.

[4] P. Briand, B. Delyon, and J. Mémin. Donsker-type theorem for bsdes. *Electronic Communications in Probability*, 6:1–14, 2001.

[5] S. Cohen. Lecture notes on risk measures. Available at `https://www.maths.ox.ac.uk/system/files/attachments/M7_12_Cohen_Risk_Measures.pdf`, 2012.

[6] S.N. Cohen and R.J. Elliott. Backward stochastic difference equations and nearly-time-consistent nonlinear expectations. *SIAM Journal of Control and Optimization*, 49:125139, 2011.

[7] F. Coquet, Y. Hu, J. Mémin, and S. Peng. A general converse comparison theorem for backward stochastic differential equations. *Comptes Rendus de l Academie des Sciences Paris*, 333:577–581, 2001.

[8] F. Coquet, Y. Hu, J. Mémin, and S. Peng. Filtration-consistent nonlinear expectations and related g-expectations. *Probability Theory and Related Fields*, 123:1–27, 2002.

[9] F. Delbaen. Risk measures or measures that describe risk? Available at `http://www.math.ethz.ch/~delbaen/ftp/preprints/Risk-executives.pdf`.

[10] F Delbaen. *Coherent risk measure on general probability spaces.* Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann, Springer, 2002.

[11] N. El Karoui, S. Peng, and M.C. Quenez. Backward stochastic differential equation in finance. mathematical finance. *Mathematical Finance*, 7 (1):1–71, 1997.

[12] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastic*, 6:429–4447, 1999.

[13] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time.* De Gruyter, Berlin, 2nd edition, 2004.

[14] M. Fritelli and E. Rosazza Gianin. Putting order in risk measures. *Journal of Banking and Finance*, 26:1473–1486, 2002.

[15] E. Gobet, J.-P. Lemor, and X. Warin. A regression-based monte carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15:2172–2202, 2005.

[16] Hermite polynomials. Available at `http://en.wikipedia.org/wiki/Hermite_polynomials`.

[17] S. A. Klugman, H. H. Panjer, and G. E. Willmot. *Loss Models: From Data to Decisions.* Wiley, New York, 2003.

[18] F. Longstaff and E. S. Schwartz. Valuing american options by simulation: A simple least squares approach. *The Review of Financial Studies*, 14:113–147, 2001.

[19] J. Ma, P. Protter, and J. San Martin. Numerical methods for backward stochastic differential equations. *Annals of Applied Probability*, 12(1):302–316, 2002.

[20] J. Ma and J. Zhang. Path regularity for solutions of baskward stochastic differential equations. *Probability theory related fields*, 98:339–359, 2002.

[21] E. Pardoux and S. Peng. Adapted solution of a bacward stochastic differential equation. *System and Control Letters*, 14 (1):55–61, 1990.

[22] S. Peng. Backward sde and related g-expectations. *Pitman Res. Notes Math. Ser.*, 364:141159, 1997.

[23] S. Peng. *Dynamically consistent evaluations and expectations.* Technical report, Institute Mathematics, Shandong Univ, 2003.

[24] S. Peng. Nonlinear expectations, nonlinear evaluations and risk measures. *Lecture notes in Math.*, 1856:165–253, 2004.

[25] E. Rosazza Gianin. Risk measures via g-expectations. *Insurance: Mathematics and Economics*, 39:19–34, 2006.

[26] M. Stadje. Extending dynamic convex risk measures from discrete time to continuous time: A convergence approach. *Insurance: Mathematics and Economics*, 47:391–404, 2010.

[27] P. Wilmott, S. Howison, and J. Dewynne. *The Mathematics of Financial Derivatives: a Student Introduction.* Cambridge University Press, Oxford, 1995.