Some contributions of Bayesian and computational learning methods to portfolio selection problems

Johann Nicolle

Université de Paris

École Doctorale de Sciences Mathématiques de Paris Centre ED 386
Laboratoire de Probabilités, Statistique et Modélisation (LPSM, UMR 8001)

Thèse de Doctorat
en Mathématiques Appliqueés

Présentée par
Johann Nicolle

Some contributions of Bayesian and computational learning methods to portfolio selection problems

---

Sous la direction de **Carmine de Franco** et **Huyên Pham**

Soutenue publiquement le 5 octobre 2020 devant le jury composé de :

Claudia Ceci	Università "G. d'Annunzio" Chieti-Pescara	Rapporteur
Carmine de Franco	OSSIAM	Directeur de thèse
Olivier Guéant	Université Paris I Panthéon Sorbonne	Rapporteur
Idris Kharroubi	Sorbonne Université	Examinateur
Huyên Pham	Université de Paris	Directeur de thèse
Marie-Claire Quenez	Université de Paris	Examinatrice
Agnès Sulem	INRIA	Présidente du jury

À ma famille,

Remerciements

Cette page est pour moi l'occasion de remercier toutes les personnes qui m'ont accompagné et soutenu durant cette thèse.

Je tiens avant tout à remercier mes deux directeurs de thèse: Carmine de Franco et Huyên Pham. Ce fût très enrichissant de travailler avec vous et j'ai beaucoup appris durant cette thèse à vos côtés. Nos rencontres à l'Institut Henri Poincaré resteront pour moi de très bons souvenirs. Carmine je te remercie pour ton soutien, ta disponibilité et pour toutes les discussions intéressantes que nous avons eues ces trois dernières années. Huyên, je te remercie beaucoup de m'avoir accompagné tout au long de cette thèse. Ton sens de la rigueur et ton exigence sont pour moi un exemple à suivre. Je vous suis très reconnaissant de m'avoir proposé de travailler sur des sujets passionnants qui ont été des défis intellectuels et qui ont aiguisé ma manière de réfléchir et de penser.

Je remercie sincèrement mes rapporteurs Claudia Ceci et Olivier Guéant pour leur relecture attentive et leurs remarques qui ont grandement contribué à améliorer le présent manuscrit. Je suis également honoré de la présence dans mon jury de Marie-Claire Quenez, d'Agnès Sulem et d'Idris Kharroubi.

Je souhaite aussi remercier chaleureusement tous mes collègues d'OSSIAM qui ont croisé ou qui croisent encore ma route au quotidien. Particulièrement, les personnes qui m'ont fait confiance et qui ont permis à cette thèse d'être une réalité: Bruno, Antoine et Fabien je vous suis très reconnaissant. Je souhaite aussi remercier Philippe, Aurélie et Emmanuelle qui ont assuré le suivi administratif de ma thèse et qui ont su gérer avec brio le dispositif CIFRE. Je remercie également tous les membres passé et présent de l'équipe Gestion - Recherche, en particulier: Mathieu, Tristan, Yannis et Imsad. Merci également

aux membres de l'IT, particulièrement Nabil et Florent, pour nos discussions en dehors des heures "CSA". Je souhaite aussi remercier: Monique, Sarah, Thomas, Malek, Hamza sans oublier Pierre et Elijo pour les bons moments et les discussions intéressantes que nous avons eus.

Mes remerciements vont aussi aux membres du LPSM que j'ai croisés au cours de ma thèse, en particulier, Côme dont l'aide fût très précieuse, Luca que j'ai rencontré à Zurich, Ouzhi et William. Merci aussi à Valérie et Nathalie qui assurent les fonctions administratives au laboratoire et Amina de l'école doctorale ED386 très à l'écoute et disponible.

Je remercie aussi chaleureusement ma famille, en particulier mes parents Françoise et Jean-Claude à qui je dois tout, et mes amis pour leur soutien indéfectible : François, Julien, Aurélien, Camille, Matthieu, Béate, Simon, Adrien, Nicolas et Cécile. Merci pour tous ces moments et pour m'avoir soutenu pendant ces années éprouvantes. Je tiens enfin à exprimer ma gratitude éternelle à Perrine, mon âme-soeur, pour son amour, son infinie patience et compréhension à mon égard. Merci d'être à mes côtés dans cette vie.

Johann NICOLLE

Résumé

La présente thèse est une étude de différents problèmes d'allocation optimale de portefeuilles dans le cas où le taux d'appréciation, appelé le *drift*, du mouvement brownien de la dynamique des actifs est incertain. Nous considérons un investisseur ayant une croyance sur le *drift* sous la forme d'une distribution de probabilité, appelée *a priori*. L'incertitude sur le *drift* est prise en compte par une approche d'apprentissage bayésien qui permet de mettre à jour la distribution de probabilité *a priori* du *drift*. La thèse est divisée en deux parties autonomes; la première partie contient deux chapitres: le premier développe les résultats théoriques, et le second contient une application détaillée de ces résultats sur des données de marché.

La première partie de la thèse est consacrée au problème de sélection de portefeuilles de Markowitz dans le cas multidimensionel avec incertitude de *drift*. Cette incertitude est modélisée via une loi arbitraire *a priori* qui est mise à jour à l'aide du filtrage bayésien. Nous avons d'abord transformé le problème de Markowitz bayésien en un problème auxiliaire standard de contrôle pour lequel la programmation dynamique est appliquée. Ensuite, nous montrons l'existence et l'unicité d'une solution régulière à l'équation aux dérivées partielles (EDP) semi-linéaire associée. Dans le cas d'une distribution *a priori* gaussienne, la solution multidimensionnelle est explicitement calculée. De plus, nous étudions l'impact quantitatif de l'apprentissage à partir des données progressivement observées, en comparant la stratégie qui met à jour l'estimation du *drift*, appelée stratégie apprenante, à celle qui la maintient constante, appelée stratégie non-apprenante. Pour finir, nous analysons la sensibilité du gain lié à l'apprentissage, appelé valeur d'information ou valeur informative, par rapport à différents paramètres.
Ensuite, nous illustrons la théorie avec une application détaillée des résultats

précédents à des données historiques de marché. Nous soulignons la robustesse de la valeur ajoutée de l'apprentissage en comparant les stratégies optimales apprenante et non-apprenante dans différents univers d'investissement: indices de différentes classes d'actifs, devises et stratégies *smart beta*.

La deuxième partie aborde un problème d'optimisation de portefeuilles en temps discret. Ici, l'objectif de l'investisseur est de maximiser l'espérance de l'utilité de la richesse terminale d'un portefeuille d'actifs risqués, en supposant un *drift* incertain et une contrainte de *maximum drawdown* satisfaite. Dans cette partie, nous formulons le problème dans le cas général, et nous résolvons numériquement le cas gaussien avec la fonction d'utilité de type *constant relative risk aversion* (CRRA), via un algorithme d'apprentissage profond. Finalement, nous étudions la sensibilité de la stratégie au degré d'incertitude entourant l'estimation du *drift* et nous illustrons empiriquement la convergence de la stratégie non-apprenante vers un problème de Merton contraint, sans vente à découvert.

Keywords: Apprentissage bayésien ; problème de Markowitz ; allocation de portefeuille ; *maximum drawdown* ; contrôle optimal stochastique ; apprentissage profond ; réseaux de neurones ; *drift* incertain ; incertitude des paramètres ; fonction d'utilité ; problème moyenne-variance ; équation de Hamilton-Jacobi-Bellman ; programmation dynamique ; valeur de l'information ; filtre de Kalman.

Liste des articles faisant partie de cette thèse:

- De Franco, C., J. Nicolle, and H. Pham (2019a). Bayesian learning for the Markowitz portfolio selection problem. *International Journal of Theoretical and Applied Finance 22(07)*.
- De Franco, C., J. Nicolle, and H. Pham (2019b) Dealing with drift uncertainty: a Bayesian learning approach. *Risks (2019) 7(1), 5*.
- C. De Franco, J. Nicolle, and H. Pham. "Discrete-time portfolio optimization under maximum drawdown constraint with partial information and deep learning resolution". Soumis pour contribuer au livre à paraître *Stochastic Analysis, Filtering, and Stochastic Optimization: A Commemorative Volume to Honor Mark H. A. Davis's Contributions*.

Abstract

The present thesis is a study of different optimal portfolio allocation problems in the case where the appreciation rate, named the drift, of the Brownian motion driving the dynamics of the assets is uncertain. We consider an investor having a belief on the drift in the form of a probability distribution, called a prior. The uncertainty about the drift is managed through a Bayesian learning approach which allows for the update of the drift's prior probability distribution. The thesis is divided into two self-contained parts; the first part being split into two chapters: the first develops the theory and the second contains a detailed application to actual market data.

The first part of the thesis is dedicated to the multidimensional Markowitz portfolio selection problem in the case of drift uncertainty. This uncertainty is modeled via an arbitrary prior law which is updated using Bayesian filtering. We first embed the Bayesian-Markowitz problem into an auxiliary standard control problem for which dynamic programming is applied. Then, we show existence and uniqueness of a smooth solution to the related semi-linear partial differential equation (PDE). In the case of a Gaussian prior probability distribution, the multidimensional solution is explicitly computed. Additionally, we study the quantitative impact of learning from the progressively observed data, by comparing the strategy which updates the initial estimate of the drift, i.e. the learning strategy, to the one that keeps it constant, named the non-learning strategy. Ultimately, we analyze the sensitivity of the gain from learning, called value of information or informative value, with respect to different parameters.
Next, we illustrate the theory with a detailed application of the previous results on actual market data. We emphasize the robustness of the value added of learning by comparing learning to non-learning optimal strategies in different

investment universes: indices of various asset classes, currencies, and smart beta strategies.

The second part tackles a discrete-time portfolio optimization problem. Here, the goal of the investor is to maximize the expected utility of the terminal wealth of a portfolio of risky assets, assuming an uncertain drift and a maximum drawdown constraint. In this part, we formulate the problem in the general case, and we solve numerically the Gaussian case with the Constant Relative Risk Aversion (CRRA) type utility function via a deep learning resolution. Ultimately, we study the sensitivity of the strategy to the degree of uncertainty of the drift and, as a by-product, give empirical evidence of the convergence of the non-learning strategy towards a no short-sale constrained Merton problem.

Keywords: Bayesian Learning; Markowitz problem; portfolio selection; maximum drawdown; optimal stochastic control; deep learning; neural networks; uncertain drift; parameter uncertainty; utility function; mean variance problem; Hamilton-Jacobi-Bellman equation; dynamic programming; value of information; Kalman filter.

List of the papers being part of this thesis:

- De Franco, C., J. Nicolle, and H. Pham (2019a). Bayesian learning for the Markowitz portfolio selection problem. *International Journal of Theoretical and Applied Finance 22(07)*.
- De Franco, C., J. Nicolle, and H. Pham (2019b) Dealing with drift uncertainty: a Bayesian learning approach. *Risks (2019) 7(1), 5.*
- De Franco, C., J. Nicolle, and H. Pham "Discrete-time portfolio optimization under maximum drawdown constraint with partial information and deep learning resolution". submitted to contribute to the upcoming book *Stochastic Analysis, Filtering, and Stochastic Optimization: A Commemorative Volume to Honor Mark H. A. Davis's Contributions.*

Contents

Résumé vii

Abstract ix

List of Figures xvi

Notations xvii

1 Introduction (in French) 1
- 1.1 Revue de la littérature . . . 2
  - 1.1.1 Problème de contrôle optimal stochastique : formulation et résolution . . . 2
  - 1.1.2 Formulation et solution en temps continu du problème de Markowitz 11
  - 1.1.3 Théorie du filtrage et apprentissage bayésien . . . 14
  - 1.1.4 Maximum Drawdown . . . 18
- 1.2 Optimisation de portefeuilles avec incertitude de paramètres - contributions . . . 20
  - 1.2.1 Chapitre 3 . . . 20
  - 1.2.2 Chapitre 4 . . . 25
- 1.3 Optimisation de portefeuilles avec contraintes de risques - contributions 28
  - 1.3.1 Chapitre 5 . . . 28

2 Introduction 33
- 2.1 Literature Review . . . 34
  - 2.1.1 Optimal stochastic control problems: formulation and resolution 34
  - 2.1.2 Markowitz problem: formulation and continuous-time solution . 42
  - 2.1.3 Filtering theory and the Bayesian learning approach . . . 45
  - 2.1.4 Maximum drawdown . . . 50
- 2.2 Portfolio optimization under parameters uncertainty - contributions . . 51

2.2.1 Chapter 3 . . . 51
2.2.2 Chapter 4 . . . 56
2.3 Portfolio optimization under risk measure constraints - contributions . . . 59
2.3.1 Chapter 5 . . . 59

I The Markowitz portfolio selection problem with drift uncertainty 65

3 Bayesian learning for the Markowitz portfolio selection problem 67
3.1 Introduction . . . 68
3.2 Markowitz problem with prior law on the uncertain drift . . . 71
3.3 Bayesian learning . . . 74
3.4 Solution to the Bayesian-Markowitz problem . . . 77
3.4.1 Main result . . . 77
3.4.2 On existence and smoothness of the Bayesian risk premium . . . 80
3.4.3 Examples . . . 81
3.4.3.1 Prior discrete law . . . 82
3.4.3.2 The Gaussian case . . . 83
3.5 Impact of learning on the Markowitz strategy . . . 85
3.5.1 Computation of the Sharpe ratios . . . 85
3.5.2 Value of Information . . . 88
3.5.2.1 Standard deviation of the drift . . . 89
3.5.2.2 Sharpe ratio of the asset . . . 91
3.5.2.3 Time . . . 92
3.5.2.4 Investment horizon . . . 93
3.6 Conclusion . . . 94
3.A Appendices . . . 95
3.A.1 Proof of Lemma 3.2.3 . . . 95
3.A.2 Proof of Lemma 3.2.4 . . . 97
3.A.3 Proof of Theorem 3.4.1 . . . 98
3.A.4 Proofs of Theorem 3.4.3 . . . 102
3.A.5 Proof of Lemma 3.4.5 . . . 106

4 Dealing with drift uncertainty: a Bayesian learning approach 109
4.1 Introduction . . . 110
4.2 The model . . . 112
4.3 Market data . . . 115
4.4 The Base Case result . . . 118
4.5 Sensitivity analysis . . . 119
4.5.1 Impact of uncertainty . . . 119

4.5.2 Impact of leverage . . . 121
4.5.3 Impact of the review frequency . . . 123
4.5.4 Impact of the rebalancing frequency . . . 125
4.6 Investing in foreign currencies . . . 126
4.7 Investing in factor strategies . . . 128
4.A Appendices . . . 130
4.A.1 Dataset of Indices . . . 131
4.A.1.1 Commodity . . . 131
4.A.1.2 Bond . . . 131
4.A.1.3 Equity . . . 133
4.A.1.4 Cash . . . 133
4.A.2 Dataset of currencies . . . 134
4.A.3 Dataset of Smart Beta strategies . . . 134
4.A.3.1 Smart Beta strategies . . . 135

II Discrete-time portfolio optimization under maximum drawdown constraint with drift uncertainty 137

5 Discrete-time portfolio optimization under maximum drawdown constraint with partial information and deep learning resolution 139
5.1 Introduction . . . 140
5.2 Problem setup . . . 142
5.3 Dynamic programming system . . . 143
5.3.1 Change of measure and Bayesian filtering . . . 143
5.3.2 The static set of admissible controls . . . 145
5.3.3 Derivation of the dynamic programming equation . . . 146
5.3.4 Special case: CRRA utility function . . . 147
5.4 The Gaussian case . . . 149
5.4.1 Bayesian Kalman filtering . . . 149
5.4.2 Finite-dimensional dynamic programming equation . . . 150
5.5 Deep learning numerical resolution . . . 152
5.5.1 Architectures of the deep neural networks . . . 152
5.5.2 *Hybrid-Now* algorithm . . . 154
5.5.3 Numerical results . . . 156
5.5.3.1 Learning and non-learning strategies . . . 156
5.5.3.2 Learning, non-learning and constrained equally-weighted strategies . . . 160
5.5.3.3 Non-learning and Merton strategies . . . 161
5.5.4 Sensitivities analysis . . . 163
5.6 Conclusion . . . 171

5.A Appendices 171
5.A.1 Proof of Proposition 5.3.1 171
5.A.2 Proof of Proposition 5.3.2 172
5.A.3 Proof of Lemma 5.3.3 172
5.A.4 Proof of Lemma 5.3.4 173
5.A.5 Proof of Lemma 5.3.5 174
5.A.6 Proof of Lemma 5.3.6 174

Bibliography **177**

List of Figures

1.1 Illustration de la frontière d'éfficience théorisée par H. Markowitz . . . 11
1.2 Illustration du maximum drawdown sur une trajectoire simulée. 20

2.1 Illustration of the efficient frontier theorized by H. Markowitz. 43
2.2 Illustration of the maximum drawdown on a simulated wealth trajectory. 51

3.1 Value of information as a function of σ_0 with parameters as in Table 3.1. 90
3.2 Value of information, Sh_T^L and Sh_T^{NL} of Asset 4 as a function of σ_0 with parameters as in Table 3.2. 90
3.3 Value of information as a function of the Sharpe ratio of the asset with parameters as in Table 3.3. 92
3.4 Value of information as a function of t with parameters as in Table 3.4. 93
3.5 Value of information as a function of T with parameters as in Table 3.5. 94

4.1 Historical values of the portfolios calculated with both $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ under the Base Case (a) and their ratio (b). 119
4.2 Annualized performances and excess returns of BL as a function of *unc* (a); Sharpe ratios and relative improvement (b). 120
4.3 Historical values of the BL portfolios calculated with $unc = 10, 100, 300$ and NL. 121
4.4 Annualized performances and excess returns of BL as a function of L (a); Sharpe ratios and relative improvement (b). 123
4.5 Annualized performances and excess returns of BL as a function of L (a); Sharpe ratios and relative improvement (b). 125
4.6 Historical values of the portfolios calculated with both $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ under the Base Case (a) and their ratio (b). 128
4.7 Historical values of the portfolios calculated with both $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ under the Base Case (a) and their ratio (b). 130

5.1 A_{NN} architecture with $d = 3$ assets . . . 153
5.2 VF_{NN} architecture with $d = 3$ assets . . . 153
5.3 Historical Learning and Non-Learning levels with a 95% confidence interval. . . . 157
5.4 Historical ratio of Learning over Non-Learning levels. . . . 157
5.5 Historical Learning and Non-Learning asset allocations. . . . 159
5.6 Historical Learning and Non-Learning total allocations. . . . 159
5.7 Historical Learning, Non-Learning and constrained EW (Const. EW) levels with a 95% confidence interval. . . . 160
5.8 Ratio Learning over constrained EW (Const. EW) according to time. . 161
5.9 Ratio Non-Learning over constrained EW (Const. EW) according to time. 161
5.10 Wealth curves resulting from the Merton strategy and the non-learning strategy for different values of q with a 95% confidence interval. . . . 162
5.11 Asset 3 average weights of the non-learning strategies with $q \in \{0.7, 0.4, 0.1\}$ and the Merton strategy. . . . 163
5.12 Excess return of Learning over Non-Learning with a 95% confidence interval for different levels of uncertainty. . . . 164
5.13 Average total performance of Learning (L) and Non-Learning (NL), and excess return, for $unc \in \{1/6, 1, 3, 6, 12\}$. . . . 165
5.14 Sharpe ratio of terminal wealth of Learning (L) and Non-Learning (NL), and relative improvement, for $unc \in \{1/6, 1, 3, 6, 12\}$. . . . 165
5.15 Average maximum drawdown of Learning (L) and Non-Learning (NL) and the gain from learning for $unc \in \{1/6, 1, 3, 6, 12\}$. . . . 166
5.16 Learning and Non-Learning historical assets' allocations with $unc = 1/6$. 167
5.17 Learning and Non-Learning historical assets' allocations with $unc = 1$. . 167
5.18 Learning and Non-Learning historical assets' allocations with $unc = 3$. . 168
5.19 Learning and Non-Learning historical assets' allocations with $unc = 6$. . 168
5.20 Learning and Non-Learning historical assets' allocations with $unc = 12$. 169
5.21 Historical total allocations of Learning and Non-Learning with $unc = 1/6$. 169
5.22 Historical total allocations of Learning and Non-Learning with $unc = 1$. 170
5.23 Historical total allocations of Learning and Non-Learning with $unc = 3$. 170
5.24 Historical total allocations of Learning and Non-Learning with $unc = 6$. 170
5.25 Historical total allocations of Learning and Non-Learning with $unc = 12$. 171

Notations

I. Sets

$\mathbb{R}^d$ denotes the d-dimensional Euclidian space. $\mathbb{R} = \mathbb{R}^1$, $\mathbb{R}_+$ is the set of nonnegative real numbers, $\mathbb{R}_+^* = \mathbb{R}_+ \backslash \{0\}$. For all $x = (x^1, \ldots, x^d)$, $y = (y^1, \ldots, y^d)$ in $\mathbb{R}^d$, we denote by $<.,.>$ the scalar product and by $|.|$ the Euclidian norm:

$$< x, y >= \sum_{i=1}^{d} x_i y_i \text{ and } |x| = \sqrt{< x, x >}.$$

The Manhattan norm is sometimes used and denoted $|.|_1$ with its usual definition:

$$|x|_1 = \sum_{i=1}^{d} |x_i|.$$

$\mathbb{R}^{n \times d}$ is the set of real-valued $n \times d$ matrices ($\mathbb{R}^{n \times 1} = \mathbb{R}^n$). I_n is the identity $n \times n$ matrix. For all $\sigma = (\sigma^{ij})_{1 \leq i \leq n, 1 \leq j \leq d} \in \mathbb{R}^{n \times d}$, we denote by $\sigma' = (\sigma^{ji})_{1 \leq i \leq n, 1 \leq j \leq d}$ the transpose matrix in $\mathbb{R}^{d \times n}$. We set $\text{tr}(A) = \sum_{i=1}^{n} a^{ii}$ the trace, and $|A|$ the determinant of a $n \times n$ matrix $A = (a^{ij})_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$.
$\mathbb{1}_d$ is the d-dimensional vector whose components are all equal to 1.
$\mathcal{S}_n$ the set of symmetric $n \times n$ matrices and $\mathcal{S}_n^+$ is the set of nonnegative definite A in $\mathcal{S}_n$.

II. Functions and functional spaces

1 is the constant function equal to 1.
For any set A, the indicator of A is denoted by

$$\mathbb{1}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

$C^k(\mathcal{O})$ is the space of all real-valued continuous functions f on $\mathcal{O}$ with continuous derivatives up to order k. Here $\mathcal{O}$ is an open set of $\mathbb{R}^n$.

$C^0([0,T]\times\mathcal{O};\mathbb{R}_+)$ is the space of positive real-valued functions f on $[0,T]\times\mathcal{O}$.
$C^{1,2}([0,T)\times\mathcal{O};\mathbb{R}_+)$ is the space of positive real-valued functions f on $[0,T)\times\mathcal{O}$ whose partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$, $\frac{\partial^2 f}{\partial x_i \partial x_j}$, $1\leq i,j\leq n$, exist and are continuous on $[0,T)$. If these partial derivatives of $f\in C^{1,2}([0,T)\times\mathcal{O};\mathbb{R}_+)$ can be extended by continuity on $[0,T]\times\mathcal{O}$, we write $f\in C^{1,2}([0,T]\times\mathcal{O};\mathbb{R}_+)$.
Given a function $f\in C^2(\mathcal{O})$, we denote by ∇f the gradient vector in $\mathbb{R}^n$ with components $\frac{\partial f}{\partial x_i}$, $1\leq i\leq n$, and $D^2 f$ the Hessian matrix in $\mathcal{S}_n$ with components $\frac{\partial f}{\partial x_i \partial x_j}$, $1\leq i,j\leq n$. When $\mathcal{O}$ is an open set in $\mathbb{R}$, we simply write f' and f''.
The gradient vector and the Hessian matrix of a function $x\mapsto f(t,x)\in C^2(\mathcal{O})$ are denoted by $\nabla_x f$ and $D_x^2 f$ and by $\partial_x f$ and $\partial_{xx}^2 f$ respectively, when $\mathcal{O}$ is an open set in $\mathbb{R}$.
If $x\mapsto f(t,x)\in\mathcal{C}^{1,2}([0,T)\times\mathbb{R}^n;\mathbb{R}^n)$ we extend the definition of $\nabla_x f$ to a $\mathbb{R}^{n\times n}$-valued function whose entries are $\left[\frac{\partial f}{\partial x_j}\right]_i$, with $1\leq i,j\leq n$.

III. Integration and probability

$(\Omega,\mathcal{F},\mathbb{P})$ is a probability space.
$\mathbb{P}$ a.s. denotes "almost surely for the probability $\mathbb{P}$"(we often omit the reference to $\mathbb{P}$ when there is no ambiguity).
$\mathcal{B}(\mathbb{R}^d)$ the Borel σ-algebra on $\mathbb{R}^d$. That is the σ-algebra generated by the open sets in $\mathbb{R}^d$.
$\sigma(\mathcal{G})$: the smallest σ-algebra containing $\mathcal{G}$, collection of subsets of Ω.
$\bigvee_{k=1}^N \mathcal{G}_k = \sigma\left(\bigcup_{k=1}^N \mathcal{G}_k\right)$ denotes the smallest σ-algebra containing $\bigcup_{k=1}^N \mathcal{G}_k$, the union of the $\mathcal{G}_k$, collection of subsets of Ω. We also note by $\bigvee_{t\in\mathbb{R}_+}\mathcal{G}_t = \sigma\left(\bigcup_{t\geq 0}\mathcal{G}_t\right)$, the smallest sigma-algebra generated by the infinite union of the $\mathcal{G}_t$'s.
$\mathbb{Q}\ll\mathbb{P}$: the measure $\mathbb{Q}$ is absolutely continuous with respect to the measure $\mathbb{P}$.
$\frac{d\mathbb{Q}}{d\mathbb{P}}$: Radon-Nikodym density of $\mathbb{Q}\ll\mathbb{P}$.
For two random variables X and Y, we note $X\perp\!\!\!\perp Y$ when X is independent of Y.
$\mathbb{E}[X]$ is the expectation of the random variable X with respect to a probability $\mathbb{P}$ initially fixed. $\mathbb{E}[X|\mathcal{G}]$ is the conditional expectation of X given the σ-algebra $\mathcal{G}$.
$\mathrm{Var}(X)=\mathbb{E}[(X-\mathbb{E}[X])(X-\mathbb{E}[X])']\in\mathcal{S}_n$ is the variance of X. It is sometimes written as $\mathrm{Cov}(X)$.
We write $X\sim\mathcal{N}(b_0,\Sigma_0)$ when X follows the multidimensional normal probability distribution with drift b_0 and covariance matrix Σ_0.
For $a<b\in\mathbb{R}_+^2$, we write $X\sim\mathcal{U}(a,b)$ when X follows the uniform probability distribution between a and b.
$\mathcal{M}_+$ is the set of all positive measures on $\mathbb{R}^d$.

IV. Abbreviations

English:
BL or L: Bayesian learning (strategy)

NL: non-learning (strategy)
ODE: ordinary differential equation
SDE: stochastic differential equation
PDE: partial differential equation
HJB: Hamilton-Jacobi-Bellman
DPP: dynamic programming principle
DNN: deep neural network
MD: maximum drawdown
CRRA: constant relative risk aversion
a.s.: almost surely
a.e.: almost everywhere

Français:
EDO : équation différentielle ordinaire
EDP : équation aux dérivées partielles
PPD : principe de programmation dynamique
p.s. : presque sûrement

Chapter 1

Introduction (in French)

Dans ce manuscrit, nous étudions l'optimisation de portefeuilles avec différentes contraintes de risque dans un contexte d'incertitude de *drift*. Il est divisé en deux parties principales qui peuvent être lues indépendamment. Dans la première partie, nous étudions le problème de sélection de portefeuilles de Markowitz en temps continu avec incertitude de *drift* dans le cadre multidimensionnel (Chapitre 3), et nous appliquons ces résultats théoriques à des données historiques de marché, afin d'étudier de manière réaliste l'effet d'apprentissage sur la performance (Chapitre 4). Dans la deuxième partie de la thèse, nous nous intéressons à un problème d'optimisation de portefeuilles à temps discret, dans le contexte d'incertitude de *drift* avec contrainte de *maximum drawdown*, que nous résolvons numériquement en utilisant un algorithme d'apprentissage profond (Chapitre 5).

Ce chapitre passe en revue la littérature sur les différents sujets et résume les principales contributions.

Contents

1.1 Revue de la littérature 2

1.1.1 Problème de contrôle optimal stochastique : formulation et résolution 2

1.1.2 Formulation et solution en temps continu du problème de Markowitz 11

1.1.3 Théorie du filtrage et apprentissage bayésien 14

1.1.4 Maximum Drawdown 18

1.2 Optimisation de portefeuilles avec incertitude de paramètres - contributions 20

1.2.1 Chapitre 3 20

1.2.2 Chapitre 4 25
1.3 Optimisation de portefeuilles avec contraintes de risques - contributions 28
1.3.1 Chapitre 5 28

1.1 Revue de la littérature

1.1.1 Problème de contrôle optimal stochastique : formulation et résolution

Nous commençons par décrire la structure de base des problèmes d'optimisation stochastique en temps continu et à horizon fini tout en décrivant l'approche principale pour les résoudre. Cette théorie et les preuves des résultats sont décrites dans Pham (2009).

Considérons un système dynamique formalisé par son état à chaque instant, et évoluant dans un environnement incertain, représenté par un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Habituellement, nous désignons par $X_t(\omega)$ l'état du système au moment t dans le cas où $\omega \in \Omega$ est l'état du monde. La dynamique du système $t \mapsto X_t(\omega)$ est décrite grâce à une équation différentielle stochastique (EDS).
La dynamique du système peut être influencée par un contrôle exprimé comme un processus $\alpha = (\alpha_t)_t$, dont la valeur est décidée à chaque temps t en fonction de l'information disponible.
Le contrôle doit respecter certaines contraintes supposées dans le problème et, à ce titre, il est considéré comme admissible.

Dans les problèmes d'optimisation qui nous intéressent, en notant x_0 la valeur initiale du processus X_t, l'objectif est de maximiser une fonctionelle $J(x_0, \alpha)$ sur tous les contrôles admissibles. Nous envisageons des fonctions objectif de la forme

$$\mathbb{E}\left[\int_0^T f(X_t, \omega, \alpha_t)dt + g(X_T, \omega)\right], \text{ sur un horizon fini } T < \infty.$$

avec f la fonction profit dépendant de t et g une fonction profit terminale. En notant,

$$J(x_0, \alpha) = \mathbb{E}\left[\int_0^T f(X_t, \omega, \alpha_t)dt + g(X_T, \omega)\right],$$

la valeur maximale est définie par

$$v = \sup_{\alpha} J(x_0, \alpha).$$

L'objectif principal d'un problème d'optimisation stochastique est de trouver le processus de contrôle qui maximise la fonction valeur.

Maintenant, nous introduisons la méthode de programmation dynamique qui est un outil important pour résoudre les problèmes de contrôle stochastique. Fondamentalement, l'approche consiste à considérer une famille de problèmes de contrôle en faisant varier les valeurs d'état initiales, et à dériver une relation entre les fonctions valeur associées. Ce principe a été initié par Bellman (1952). Il donne une équation aux dérivées partielles (EDP) non linéaire de second ordre, appelée équation de Hamilton-Jacobi-Bellman (HJB). Lorsque cette EDP peut être résolue de manière explicite ou théorique avec une solution régulière, le théorème de vérification valide l'optimalité de la solution candidate à l'équation de HJB. Cette approche suppose l'existence d'une solution régulière à l'équation de HJB, qui n'est pas vraie en général.

Nous considérons un modèle où la dynamique du système suit l'équation différentielle stochastique (EDS) contrôlée à valeur dans $\mathbb{R}^n$:

$$dX_s = b(X_s, \alpha_s)ds + \sigma(X_s, \alpha_s)dW_s \tag{1.1}$$

où W est un mouvement brownien m-dimensionnel sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfaisant les conditions habituelles. Le contrôle $\alpha = (\alpha_s)$ est un processus progressivement $\mathbb{F}$-mesurable, à valeur dans A, sous-espace de $\mathbb{R}^d$.
Les fonctions mesurables $b : \mathbb{R}^n \times A \to \mathbb{R}^n$ et $\sigma : \mathbb{R}^n \times A \to \mathbb{R}^{n\times m}$ satisfont une condition de Lipschitz uniforme $A : \exists K \geq 0,\ \forall x,\ y \in \mathbb{R}^n,\ \forall a \in A,$

$$|b(x, a) - b(y, a)| + |\sigma(x, a) - \sigma(y, a)| \leq K|x - y|. \tag{1.2}$$

Fixons un horizon fini $0 < T < \infty$ et notons par $\mathcal{A}$, l'espace des processus de contrôle α, tels que

$$\mathbb{E}\left[\int_0^T |b(0, \alpha_t)|^2 + \sigma(0, \alpha_t)|^2 dt\right] < \infty. \tag{1.3}$$

Les conditions (1.2) et (1.3) assurent pour tout $\alpha \in \mathcal{A}$ et pour toutes conditions initiales $(t, x) \in [0, T] \times \mathbb{R}^n$, l'existence et l'unicité d'une solution forte à l'EDS (1.1) commençant à x au temps $s = t$. Nous notons ensuite par $\{X_s^{t,x},\ t \leq s \leq T\}$ la solution presque sûrement (p.s.) continue.

Pour $f : [0, T] \times \mathbb{R}^n \times A \to \mathbb{R}$ et $g : \mathbb{R}^n \to \mathbb{R}$ deux fonctions mesurables. Nous supposons que :
- g est bornée inférieurement

ou

- g satisfait une condition de croissance quadratique : $|g(x)| \leq C(1+|x|^2)$, $\forall x \in \mathbb{R}^n$, avec C une constante independante de x.

Pour $(t,x) \in [0,T] \times \mathbb{R}^n$, nous notons $\mathcal{A}(t,x)$ le sous-espace des contrôles α dans $\mathcal{A}$ tels que

$$\mathbb{E}\left[\int_t^T |f(s, X_s^{t,x}, \alpha_s)|ds\right] < \infty,$$

et nous considérons que $\mathcal{A}(t,x)$ est non vide pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$. Nous pouvons ainsi définir sous les hypothèses précédentes la fonction de gain :

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^T f(s, X_s^{t,x}, \alpha_s)ds + g(X_s^{t,x})\right],$$

pour tout $(t,x) \in [0,T] \times \mathbb{R}^n$ et $\alpha \in \mathcal{A}(t,x)$. L'objectif est de maximiser sur les processus de contrôle la fonction de gain J, et nous introduisons la fonction valeur associée :

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} J(t,x,\alpha). \tag{1.4}$$

Étant donné une condition initiale $(t,x) \in [0,T) \times \mathbb{R}^n$, nous disons que $\hat{\alpha} \in \mathcal{A}(t,x)$ est un contrôle optimal si $v(t,x) = J(t,x,\hat{\alpha})$. Nous disons qu'un processus de contrôle α sous la forme $\alpha_s = a(s, X_s^{t,x})$ pour une fonction mesurable a de $[0,T] \times \mathbb{R}^n$ dans A, est un contrôle markovien. Dans la suite, nous supposerons que la fonction valeur v est mesurable par rapport à ses arguments, ce qui n'est pas trivial *a priori*.

Le principe de programmation dynamique (PPD) est un principe fondamental de la théorie du contrôle stochastique. Dans le cadre de processus markoviens contrôlés, il est formulé comme suit :

Theorem 1.1.1 (Principe de programmation dynamique à horizon fini)**.** *Pour $0 \leq t \leq T < \infty$, nous désignons par $\mathcal{T}_{t,T}$ l'espace des temps d'arrêt à valeur dans $[t,T]$. Pour $(t,x) \in [0,T] \times \mathbb{R}^n$, nous avons ainsi*

$$\begin{aligned} v(t,x) &= \sup_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})]\right] \\ &= \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})]\right] \end{aligned}$$

Une version plus forte du PPD peut s'écrire ainsi

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^\theta f\left(s, X_s^{t,x}, \alpha_s\right) ds + v\left(\theta, X_\theta^{t,x}\right)\right]. \tag{1.5}$$

Le PPD implique qu'un contrôle optimal sur tout l'intervalle $[t, T]$ peut être obtenu : premièrement, en trouvant un contrôle optimal à partir du temps θ étant donné la variable d'état $X_\theta^{t,x}$, ou de manière équivalente calculer $v(\theta, X_\theta^{t,x})$, et deuxièmement, en maximisant sur $[t, \theta]$ l'expression

$$\mathbb{E}\left[\int_t^\theta f\left(s, X_s^{t,x}, \alpha_s\right) ds + v(\theta, X_\theta^{t,x})\right].$$

Dans la théorie de la programmation dynamique, l'équation de HJB est la version infinitésimale du PPD et explique le comportement local de la fonction valeur lorsque le temps d'arrêt θ tend vers t. Dans un problème d'optimisation stochastique standard, les étapes suivantes sont généralement suivies :

1. Trouver l'équation de HJB,
2. Montrer l'existence d'une solution régulière en utilisant des techniques d'EDP,
3. Vérifier que la solution régulière est la fonction valeur en utilisant la formule d'Itô,
4. Obtenir un contrôle optimal, appelé *feedback*, qui dépend de la valeur courante du processus X.

Nous allons maintenant dériver formellement l'équation de HJB dans le cas d'un horizon fini $T < \infty$. Considérons le temps $\theta = t + h$ et un contrôle constant arbitraire $\alpha_s = a \in A$ tel que, grâce au Théorème 1.1.1,

$$v(t, x) \geq \mathbb{E}\left[\int_t^{t+h} f(s, X_s^{t,x}, a) ds + v(t + h, X_{t+h}^{t,x})\right]. \tag{1.6}$$

Si nous supposons maintenant que v est suffisamment régulière pour appliquer la formule d'Itô entre les instants t et $t + h$:

$$v\left(t + h, X_{t+h}^{t,x}\right) = v(t, x) + \int_t^{t+h} \left(\partial_t v + \mathcal{L}^a v\right)\left(s, X_s^{t,x}\right) ds + \text{martingale locale},$$

avec $\mathcal{L}^a$ l'opérateur associé à la diffusion (1.1) pour le contrôle constant a, défini par

$$\mathcal{L}^a v = < b(x, a), \nabla_x v > + \frac{1}{2}\text{tr}\left(\sigma(x, a)\sigma'(x, a)\mathcal{D}_x^2 v\right).$$

En substituant dans l'Equation (1.6), supposant que v satisfait une condition de croissance quadratique et utilisant une technique de localisation afin de supprimer

l'intégrale stochastique, c'est-à-dire le terme de martingale locale dans l'espérance, nous obtenons

$$0 \geq \mathbb{E}\left[\int_t^{t+h} \left(\partial_t v + \mathcal{L}^a v\right)\left(s, X_s^{t,x}\right) + f\left(s, X_s^{t,x}, a\right) ds\right].$$

Nous divisons ensuite par h que nous faisons tendre vers 0 pour obtenir, par le théorème de la moyenne,

$$0 \geq \partial_t v(t,x) + \mathcal{L}^a v(t,x) + f(t,x,a).$$

L'inégalité précédente étant vraie pour tout $a \in A$, nous en déduisons l'inégalité suivante,

$$-\partial_t v(t,x) - \sup_{a \in A}\left[\mathcal{L}^a v(t,x) + f(t,x,a)\right] \geq 0. \tag{1.7}$$

Supposons ensuite que le contrôle optimal soit α^* dans l'Equation (1.5),

$$v(t,x) = \mathbb{E}\left[\int_t^{t+h} f\left(s, X_s^*, \alpha_s^*\right) ds + v\left(t+h, X_{t+h}^*\right)\right]$$

où X^* est solution de l'EDS (1.1) commençant à x au temps t, avec le contrôle α^*. En suivant le même raisonnement que précedemment, nous obtenons

$$-\partial_t v(t,x) - \mathcal{L}^{\alpha_t^*} v(t,x) - f(t,x,\alpha_t^*) = 0.$$

Avec l'Equation (1.7), cela suggère que la fonction valeur v doit satisfaire

$$-\partial_t v(t,x) - \sup_{a \in A}\left[\mathcal{L}^a v(t,x) + f(t,x,a)\right] = 0, \quad \forall (t,x) \in [0,T) \times \mathbb{R}^n,$$

quand le supremum en a est fini. Des techniques existent lorsque le supremum est infini, ce qui peut généralement se produire lorsque l'espace de contrôle A est non-borné, et nous référons à Pham (2009) pour plus de détails. L'EDP précédente est souvent réécrite sous la forme

$$-\partial_t v(t,x) - H\left(t, x, \nabla_x v(t,x), \mathcal{D}_x^2 v(t,x)\right) = 0, \quad \forall (t,x) \in [0,T) \times \mathbb{R}^n, \tag{1.8}$$

où pour $(t,x,p,M) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n$,

$$H(t,x,p,M) = \sup_{a \in A}\left[< b(x,a), p > + \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x,a)M\right) + f(t,x,a)\right].$$

La fonction H est appelée opérateur hamiltonien du problème de contrôle associé et l'Equation (1.8) est l'équation de HJB. En se souvenant de l'Equation (1.4), la condition terminale régulière de l'EDP précédente est

$$v(T,x) = g(x), \quad \forall x \in \mathbb{R}^n.$$

Finalement, nous présentons un résultat théorique important dans la théorie de la programmation dynamique : le théorème de vérification. Il consiste à prouver que, étant donnée une solution régulière à l'équation de HJB, elle coïncide avec la fonction valeur recherchée. Ce théorème donne également un contrôle markovien optimal.

Avec les hypothèses et notations que nous avons faites précédemment, nous formulons une version générale du théorème de vérification.

Theorem 1.1.2. *Soit w une fonction dans $C^{1,2}([0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$, satisfaisant une condition de croissance quadratique de la forme, il existe une constante $C>0$ telle que*

$$|w(t,x)| \le C(1+|x|^2), \quad \forall (t,x) \in [0,T]\times\mathbb{R}^n.$$

1. *Supposons que*

$$-\partial_t w(t,x) - \sup_{a\in A}[\mathcal{L}^a w(t,x) + f(t,x,a)] \ge 0, \quad (t,x)\in[0,T)\times\mathbb{R}^n,$$
$$w(T,x) \ge g(x), \quad x\in\mathbb{R}^n.$$

   Alors $w \ge v$ sur $[0,T]\times\mathbb{R}^n$.

2. *De plus, si $w(T,.) = g$, et s'il existe une fonction mesurable $\hat{\alpha}(t,x)$, $(t,x)\in[0,T)\times\mathbb{R}^n$, à valeur dans A telle que*

$$\begin{aligned} &-\partial_t w(t,x) - \sup_{a\in A}[\mathcal{L}^a w(t,x) + f(t,x,a)] \\ &= -\partial_t w(t,x) - \mathcal{L}^{\hat{\alpha}(t,x)} w(t,x) - f(t,x,\hat{\alpha}(t,x)) = 0 \end{aligned}$$

   l'équation différentielle stochastique

$$dX_s = b\left(X_s, \hat{\alpha}(s,X_s)\right) ds + \sigma\left(X_s, \hat{\alpha}(s,X_s)\right) dW_s$$

   admet une solution unique, notée $\hat{X}_s^{t,x}$, étant donnés une condition initiale $X_t = x$, et un processus $\left\{\hat{\alpha}\left(s,\hat{X}_s^{t,x}\right), t\le s\le T\right\}$ appartenant à $\mathcal{A}(t,x)$. Alors

$$w = v \quad \text{on } [0,T]\times\mathbb{R}^n,$$

   et $\hat{\alpha}$ est un contrôle optimal markovien.

Le théorème ci-dessus suggère les étapes suivantes pour résoudre le problème de contrôle stochastique dans le cas horizon fini. Tout d'abord, nous essayons de résoudre l'équation non linéaire de HJB :

$$-\partial_t w(t,x) - \sup_{a\in A}[\mathcal{L}^a w(t,x) + f(t,x,a)] = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n,$$

avec la condition terminale $w(T,x) = g(x)$. Fixons $(t,x) \in [0,T) \times \mathbb{R}^n$, et résolvons $\sup_{a\in A} |\mathcal{L}^a w(t,x) + f(t,x,a)|$ comme un problème de maximisation avec $a \in A$. Nous notons $a^*(t,x)$ l'argument maximum de ce maximum. Si cette EDP non-linéaire avec une condition terminale admet comme solution régulière w, alors w est la fonction valeur du problème de contrôle stochastique, et a^* est un contrôle markovien optimal. Cette approche est pertinente dès que l'équation de HJB possède une solution $C^{1,2}$ satisfaisant les conditions d'application du théorème de vérification.

Les résultats sur l'existence de solutions régulières aux équations aux dérivées partielles paraboliques de HJB se trouvent dans Fleming and Rishel (1975), Gilbarg and Trudinger (1985) et Krylov (1987). L'hypothèse principale requise est une condition d'ellipticité uniforme :

- Il existe une constante $c > 0$ telle que $y'\sigma(x,a)\sigma'(x,a)y \geq c|y|^2$, $\forall x, y \in \mathbb{R}^n$, $\forall a \in A$.

Nous illustrons la théorie développée précédemment, par un exemple intéressant tiré de Pham (2009), qui constitue une bonne entrée en matière pour les problèmes plus sophistiqués développés dans la suite.

Example 1.1.3 (Problème de Merton d'allocation de portefeuille à horizon fini)**.** *Le cadre considéré est le modèle Black-Scholes-Merton sur un horizon d'investissement fini T. À tout moment t, un investisseur répartit une proportion α_t de sa richesse dans un actif risqué de prix S et $1-\alpha_t$ dans un actif sans risque de prix S^0 avec un taux d'intérêt r. La proportion de richesse prend sa valeur dans A, un sous-ensemble convexe fermé de $\mathbb{R}$. Le processus de richesse est donc supposé évoluer ainsi*

$$\begin{aligned} dX_t &= \frac{X_t\alpha_t}{S_t}dS_t + \frac{X_t(1-\alpha_t)}{S_t^0}dS_t^0 \\ &= X_t(\alpha_t\mu + (1-\alpha_t)r)dt + X_t\alpha_t\sigma dW_t. \end{aligned}$$

L'ensemble des processus progressivement mesurables $\alpha \in A$ tels que $\int_0^T |\alpha_s|^2 ds < \infty$ a.s. est noté $\mathcal{A}$. La condition d'intégrabilité précédente garantit l'existence et l'unicité d'une solution forte à l'équation différentielle stochastique régissant le processus de richesse contrôlé par $\alpha \in \mathcal{A}$. On note $X^{t,x}$, le processus de richesse partant de $X_t = x > 0$ à l'instant t, correspondant à la stratégie de portefeuille $\alpha \in \mathcal{A}$.

L'investisseur a pour objectif de maximiser la valeur attendue de la richesse terminale de son portefeuille à l'instant terminal T. La fonction valeur de ce problème de maximisation d'espérance d'utilité est alors formulée

$$v(t,x) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}\left[U\left(X_T^{t,x}\right)\right], \quad (t,x) \in [0,T] \times \mathbb{R}_+. \tag{1.9}$$

La fonction U est une fonction d'utilité. Elle est donc croissante et concave sur $\mathbb{R}_+$. Nous pouvons facilement vérifier que pour tout $t \in [0,T]$, $v(t,.)$ est également croissante et concave en x. De plus, si U est strictement concave et s'il existe un contrôle optimal, on peut montrer que v est également strictement convave en x. Nous dérivons l'équation de HJB pour le problème de contrôle stochastique (1.9) comme suit :

$$-\partial_t w - \sup_{a \in \mathcal{A}} [\mathcal{L}^a w(t,x)] = 0, \tag{1.10}$$

avec la condition terminale

$$w(T,x) = U(x), \quad x \in \mathbb{R}_+. \tag{1.11}$$

Dans cet exemple, l'opérateur de diffusion a la forme explicite

$$\mathcal{L}^a w(t,x) = x(a\mu + (1-a)r)\partial_x w + \frac{1}{2}x^2 a^2 \sigma^2 \partial^2_{xx} w.$$

Si nous choisissons la fonction d'utilité puissance de type constant relative risk aversion *(CRRA), telle qu'utilisée par Merton à l'origine,*

$$U(x) = \frac{x^p}{p}, \quad x \geq 0,\ 0 < p < 1,$$

nous pouvons trouver analytiquement une solution régulière au problème (2.10) - (2.11). Nous proposons comme solution candidate pour l'équation de HJB,

$$w(t,x) = \phi(t)U(x)$$

avec ϕ une fonction positive. En mettant la solution candidate dans l'équation de HJB, ϕ doit satisfaire l'EDO

$$\phi'(t) + \tilde{\rho}\phi(t) = 0, \quad \phi(T) = 1, \tag{1.12}$$

dans laquelle

$$\tilde{\rho} = p \sup_{a \in A} \left[a(\mu - r) + r - \frac{1}{2}a^2(1-p)\sigma^2\right].$$

La solution de l'EDO 1.12 est $\phi(t) = \exp(\tilde{\rho}(T-t))$. *On peut en déduire que la fonction donnée par*

$$w(t,x) = \exp(\tilde{\rho}(T-t))U(x), \quad (t,x) \in [0,T] \times \mathbb{R}_+, \tag{1.13}$$

est strictement croissante et concave en x *et est une solution régulière au problème* (2.10) - (2.11). *De plus, la fonction* $a \in A \mapsto a(\mu - r) + r - \frac{1}{2}a^2(1-p)\sigma^2$ *est strictement concave sur l'ensemble convexe fermé* A, *et atteint ainsi son maximum, à un certain* $\hat{a}$ *constant. Par construction,* $\hat{a}$ *atteint le supremum de* $\sup_{a \in A} [\mathcal{L}^a w(t,x)]$. *De plus, le processus de richesse associé au contrôle constant* $\hat{a}$

$$dX_t = X_t(\hat{a}\mu + (1-\hat{a})r)dt + X_t\hat{a}\sigma dW_t,$$

admet une solution unique étant donné une condition initiale. Par le théorème de vérification 1.1.2, la fonction valeur au problème de maximisation d'espérance d'utilité (2.9) *est donnée par l'Equation* (1.13), *et la proportion optimale de richesse à allouer dans l'actif risqué, est la constante* $\hat{a}$. *Finalement, notons que lorsque* $A = \mathbb{R}$, *nous pouvons explicitement écrire les valeurs de* $\hat{a}$ *et* $\tilde{\rho}$ *qui sont respectivement égales à*

$$\hat{a} = \frac{\mu - r}{\sigma^2(1-p)},$$

et

$$\tilde{\rho} = \frac{(\mu - r)^2}{2\sigma^2}\frac{p}{1-p} + rp.$$

A l'origine, le principe de la programmation dynamique a été inventé par Bellman (1952). Bien que le concept soit facile à comprendre, la démonstration rigoureuse se révèle complexe et a été étudiée par de nombreux auteurs à l'aide de différentes méthodes : Bensoussan and Lions (1978), Krylov (1980), Nisio (1981), Lions (1983), Borkar (1989), Yong and Zhou (2000) ou Fleming and Soner (2006). Une littérature abondante traite de l'approche classique du contrôle optimal des processus de diffusion via un théorème de vérification sur l'équation de HJB : Fleming and Rishel (1975), Krylov (1980), Yong and Zhou (2000) ou Fleming and Soner (2006). L'extension aux processus contrôlés de diffusion à sauts est étudiée dans Øksendal and Sulem (2004). Certains autres aspects sont examinés dans les notes de cours de St-Flour El Karoui (1981).
L'existence générale d'un problème de contrôle optimal est abordée dans Kushner (1975) ou El Karoui et al. (1987).

1.1.2 Formulation et solution en temps continu du problème de Markowitz

Dans les Chapitres 3 et 4, nous traitons une version innovante du problème de sélection de portefeuilles de Markowitz, ce qui implique que nous passions brièvement en revue le cadre standard et certains résultats. L'analyse de portefeuilles de Markowitz est une procédure mathématique pour déterminer les portefeuilles optimaux dans lesquels investir. Cette théorie, initiée par Markowitz (1952), est une avancée majeure en finance. L'analyse de portefeuilles a pour objectif, de trouver l'ensemble des portefeuilles efficients, c'est-à-dire ceux qui :

- ont le meilleur rendement attendu pour un niveau de risque donné, ou inversement,
- offrent le niveau de risque le plus faible pour un niveau donné de rendement attendu.

Les portefeuilles efficients sont situés sur la frontière efficiente, représentée par la partie croissante de la ligne pointillée de la Figure 1.1, partant du portefeuille de variance minimale.

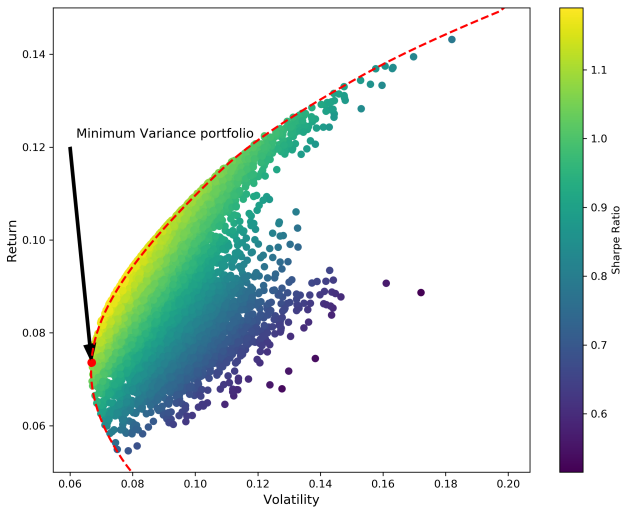


Figure 1.1: Illustration de la frontière d'éfficience théorisée par H. Markowitz

Intéressons-nous, maintenant, à la formulation et à la résolution du problème de Markowitz en temps continu. Pour simplifier et comme ce n'est pas le coeur de notre étude nous ne considérons pas l'actif sans-risque dans les aspects théoriques

du problème de Markowitz dans cette thèse. Néanmoins, dans les applications, il sera utilisé pour rendre les simulations plus réalistes. Pour être plus précis, il ne sera pas inclus dans le processus d'optimisation mais, comme on ne contraint pas les poids des actifs à sommer à 1, il sera utilisé comme taux sans risque pour rémunérer le *cash*.

On définit un marché financier sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ équipé d'une filtration $\{\mathcal{F}_t\}_{t\geq 0}$ qui remplit les conditions habituelles. Sur ce marché, se négocient en continu n actifs risqués. Le processus de prix des actifs S s'écrit

$$\begin{cases} dS_t &= \mathrm{diag}(S_t)\big(b_0 dt + \sigma dW_t\big), \quad t \in [0,T], \\ S_0 &= s_0 \in \mathbb{R}^n, \end{cases}$$

où W est un mouvement brownien standard de dimension n et T un horizon d'investissement fini.

Le drift n-dimensionnel b_0 est un paramètre connu et constant, et la matrice constante σ, de dimension $n \times n$, est connue et inversible. On note Σ la matrice carrée symétrique $\sigma\sigma'$, de dimension n.

Nous écrivons $\mathbb{F}^{\mathbb{S}} = \{\mathcal{F}_t^S\}_{t\geq 0}$, la filtration générée par le processus de prix S, augmentée des ensembles nuls de $\mathcal{F}$. La filtration $\mathbb{F}^S$ représente la seule source d'information disponible : le processus de prix S.

On note $\alpha = (\alpha_t)_{0\leq t\leq T}$, une stratégie d'investissement admissible représentant le montant investi dans chacun des actifs risqués. C'est un processus $\mathbb{F}^{\mathbb{S}}$-progressivement mesurable, à valeurs dans $\mathbb{R}^n$, satisfaisant une condition d'intégrabilité. Nous définissons $\mathcal{A}$, comme l'ensemble de toutes les stratégies d'investissement admissibles.

L'évolution du processus de richesse autofinancée X^α, étant donné $\alpha \in \mathcal{A}$ et une richesse initiale $x_0 \in \mathbb{R}$, est donnée par :

$$dX_t^\alpha = \alpha_t' \mathrm{diag}(S_t)^{-1} dS_t = \alpha_t'(b_0 dt + \sigma dW_t), \quad 0 \leq t \leq T, \quad X_0^\alpha = x_0.$$

Le problème de Markowitz est alors formulé comme suit :

$$\hat{U}_0(\vartheta) \;:=\; \sup_{\alpha\in\mathcal{A}} \Big\{ \mathbb{E}[X_T^\alpha] : \mathrm{Var}(X_T^\alpha) \leq \vartheta \Big\}, \tag{1.14}$$

où $\vartheta > 0$, est la variance maximale de la valeur terminale du portefeuille que l'investisseur autorise. La fonction valeur $\hat{U}_0$ est alors l'espérance de la valeur terminale optimale du portefeuille que l'investisseur peut atteindre, compte tenu de sa

contrainte de variance. Nous pouvons montrer que le problème de Markowitz (1.14) peut être transformé en un problème moyenne-variance, où la contrainte fait partie de la fonction objectif, qui s'écrit :

$$\hat{V}_0(\lambda) := \inf_{\alpha\in\mathcal{A}} \left[\lambda \mathrm{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha]\right], \quad \lambda > 0.$$

Grâce à Zhou and Li (2000), nous savons que ce problème incohérent en temps peut être transformé en un problème auxiliaire sur lequel nous pouvons appliquer le PPD. L'équation de HJB de ce problème auxiliaire, se résume à la résolution de l'EDO suivante :

$$\begin{cases} 0 = & R'(t) + |\sigma^{-1}b_0|^2, \\ 0 = & R(T). \end{cases}$$

qui a pour solution $R(t) = |\sigma^{-1}b_0|^2(T-t)$. Cela nous permet de résoudre le problème initial de variance moyenne

$$\hat{V}_0(\lambda) = -\frac{1}{4\lambda}\left(e^{|\sigma^{-1}b_0|^2 T} - 1\right) - x_0,$$

avec le contrôle optimal en forme *feedback*

$$\alpha^{*,\lambda} = a_0^{*,\lambda}(X_t^{\alpha^{*,\lambda}}),$$

où

$$a_0^{*,\lambda}(x) = \left(x_0 - x + \frac{e^{|\sigma^{-1}b_0|^2 T}}{2\lambda}\right)\Sigma^{-1}b_0, \quad x \in \mathbb{R}.$$

Enfin, nous déduisons l'espérance optimale du problème de Markowitz :

$$\hat{U}_0(\vartheta) = x_0 + \sqrt{\vartheta(e^{|\sigma^{-1}b_0|^2 T} - 1)}.$$

et son contrôle *feedback* optimal,

$$\alpha^{*,\vartheta} = a_0^{*,\vartheta}(X_t^{\alpha^{*,\vartheta}}),$$

où

$$a_0^{*,\vartheta}(x) = \left(x_0 - x + e^{|\sigma^{-1}b_0|^2 T}\sqrt{\frac{\vartheta}{e^{|\sigma^{-1}b_0|^2 T} - 1}}\right)\Sigma^{-1}b_0, \quad x \in \mathbb{R}.$$

Dans la présente thèse, nous souhaitons résoudre ce problème lorsque le drift est incertain et modélisé via une distribution de probabilité *a priori*.
Pour atteindre cet objectif, nous devons revenir brièvement sur quelques bases de la théorie du filtrage.

1.1.3 Théorie du filtrage et apprentissage bayésien

La présente thèse se concentre sur l'optimisation de portefeuilles avec incertitude des paramètres, et met donc l'accent sur le cas où l'investisseur ne peut pas observer directement le *drift* du mouvement brownien : il ne peut qu'observer les prix des actifs passés et présents. Nous supposons que le *drift* inconnu est une variable aléatoire inobservable, indépendante du mouvement brownien et possède une distribution de probabilité connue. Ce cas est appelé cas d'observations partielles.

Dans la littérature, étant donné que la théorie du filtrage est utilisée pour prendre en compte l'incertitude de *drift*, nous montrons dans cette section des résultats standards importants de cette théorie prouvés dans Bain and Crisan (2009).

L'approche de filtrage. Nous définissons un espace de probabilité $(\Omega,\ \mathcal{F},\ \mathbb{P})$ avec une filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ qui satisfait aux conditions habituelles. Considérons un processus inobservable adapté à $\mathbb{F}$, $X = \{X_t,\ t \geq 0\}$ qui prend ses valeurs dans $\mathbb{R}^n$. Nous appelons μ_0, la distribution de X_0.

Soit $h = (h_i)_{i=1}^m :\ \mathbb{R}^n \mapsto \mathbb{R}^m$ une fonction mesurable telle que

$$\mathbb{P}\left(\int_0^t |h(X_s)|ds < \infty\right) = 1$$

pour tout $t \geq 0$. On note W, un mouvement brownien standard $\mathbb{F}$-adapté m-dimensionnel sur $(\Omega,\ \mathcal{F},\ \mathbb{P})$ indépendant de X.

Soit Y, le processus d'observation satisfaisant l'équation d'évolution suivante :

$$Y_t = Y_0 + \int_0^t h(X_s)ds + W_t.$$

Nous notons

$$\mathcal{Y} = \bigvee_{t\in\mathbb{R}_+} \mathcal{Y}_t,$$

où $\mathcal{Y}_t$ est l'augmentation habituelle avec les ensembles nuls de la filtration associée au processus Y.

Nous remarquons que, h étant mesurable, Y est $\mathbb{F}$-adapté et donc $\mathcal{Y}_t \subset \mathcal{F}_t$.

Remark 1. *La plupart des résultats peuvent être étendus dans le cas où h dépend du temps mais pour simplifier la notation, nous avons supprimé la dépendance temporelle.* ◊

Le problème de filtrage consiste à déterminer la distribution conditionnelle μ_t du processus X à l'instant t, compte tenu des informations accumulées par l'observation de Y dans l'intervalle $[0,t]$, c'est-à-dire pour une fonction mesurable ϕ, calculer

$$\mu_t(\phi) = \mathbb{E}[\phi(X_t)|\mathcal{Y}_t].$$

La variable Y_0 est considéré comme nulle car aucune information n'est disponible au temps 0. Ainsi, μ_0 la distribution initiale de X_0, est identique à la distribution conditionnelle de X_0 sachant $\mathcal{Y}_0$ et nous avons

$$\mu_0(\phi) = \int_{\mathbb{R}^d} \phi(x)\mathbb{P}X_0^{-1}(dx).$$

Deux approches sont disponibles dans la littérature pour déduire l'évolution de μ.

La méthode de changement de mesure. Cette technique, utilisée dans le Chapitre 5, consiste à construire une nouvelle méthode dans laquelle Y devient un mouvement brownien et μ a une représentation non-normalisée, π. Il est ensuite démontré que π satisfait une équation d'évolution linéaire, qui conduit à l'équation d'évolution de μ, par une application de la formule d'Itô.

Cette méthode consiste à modifier la mesure de probabilité sur Ω, afin de transformer le processus Y en un mouvement brownien, au moyen du théorème de Girsanov. Soit $Z = \{Z_t,\ t \geq 0\}$, le processus défini par

$$Z_t = e^{-\sum_{i=1}^m \int_0^t h^i(X_s)dW_s^i - \frac{1}{2}\sum_{i=1}^m \int_0^t h^i(X_s)^2 ds}, \quad t \geq 0.$$

Ce processus est une martingale, si la condition de Novikov suivante est vérifiée

$$\mathbb{E}\left[e^{\frac{1}{2}\sum_{i=1}^m \int_0^t h^i(X_s)^2 ds}\right] < \infty,$$

pout tout $t > 0$.

Remark 2. *Notons qu'avec des conditions appropriées, il suffit de montrer que*

$$\mathbb{E}\left[\sum_{i=1}^m \int_0^t h^i(X_s)^2 ds\right] < \infty,$$

pour que Z soit une martingale. $\Diamond$

Pour construire le changement de probabilité, Z doit être une martingale, ce qui est vrai si la condition sur h suivante est vérifiée

$$\mathbb{E}\left[\int_0^t |h(X_s)|^2 ds\right] < \infty, \;\; \mathbb{E}\left[\int_0^t Z_s |h(X_s)|^2 ds\right] < \infty, \quad t \geq 0. \tag{1.15}$$

Nous pouvons montrer que si la condition précédente (1.15) est satisfaite, le processus $Z = \{Z_t,\ t \geq 0\}$ est une martingale adaptée à $\mathbb{F}$.

Pour un $t \geq 0$ fixé, $Z_t > 0$ introduit une mesure de probabilité $\overline{\mathbb{P}}^t$ sur $\mathcal{F}_t$ en spécifiant sa dérivée de Radon-Nikodym par rapport à $\mathbb{P}$ égale à Z_t,

$$\left.\frac{d\overline{\mathbb{P}}^t}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t.$$

C'est à partir de la propriété martingale de Z que les mesures $\overline{\mathbb{P}}^t$ forme une famille cohérente. Par conséquent, nous pouvons définir une mesure de probabilité $\overline{\mathbb{P}}$ équivalente à $\mathbb{P}$ sur $\cup_{0\leq t\leq\infty}\mathcal{F}_t$.
Nous pouvons également prouver que, si la condition (1.15) est satisfaite, alors, sous $\overline{\mathbb{P}}$, le processus d'observation Y est un mouvement brownien indépendant de X. De plus, la loi du processus X sous $\overline{\mathbb{P}}$ est la même que celle sous $\mathbb{P}$.

On note $\overline{Z} = \{\overline{Z}_t,\ t \geq 0\}$, le processus défini comme $\overline{Z}_t = Z_t^{-1}$ pour $t \geq 0$. Sous $\overline{\mathbb{P}}$, $\overline{Z}_t$ satisfait l'EDS suivante :

$$d\overline{Z}_t = \sum_{i=1}^m \overline{Z}_t h^i(X_t) dY_t^i,$$

et, puisque $\overline{Z}_0 = 1$,

$$\overline{Z}_t = e^{\sum_{i=1}^m \int_0^t h^i(X_s) dY_s^i - \frac{1}{2}\sum_{i=1}^m \int_0^t h^i(X_s)^2 ds}.$$

Comme $\overline{\mathbb{E}}[\overline{Z}_t] = \mathbb{E}[\overline{Z}_t Z_t] = 1$, $\overline{Z}_t$ est une martingale adaptée à $\mathbb{F}$ sous $\overline{\mathbb{P}}$, et nous avons

$$\left.\frac{d\mathbb{P}}{d\overline{\mathbb{P}}^t}\right|_{\mathcal{F}_t} = \overline{Z}_t \quad \text{pour } t \geq 0.$$

Sous $\overline{\mathbb{P}}$, le processus d'observation Y est un mouvement brownien adapté à $\mathcal{Y}_t$. Puisque le mouvement brownien est un processus de Markov, pour une variable aléatoire intégrable $\mathcal{F}_t$-mesurable U, nous avons

$$\overline{\mathbb{E}}[U|\mathcal{Y}_t] = \overline{\mathbb{E}}[U|\mathcal{Y}].$$

Cela nous permet de remplacer la famille de tribus dépendante du temps $\mathcal{Y}_t$ dans les espérances conditionnelles, par la tribu fixe $\mathcal{Y}$. Cela signifie que la solution du problème de filtrage pour une variable aléatoire $\mathcal{F}_t$-adaptée U sachant toutes les observations, futures, présentes et passées, est égale à $\overline{\mathbb{E}}[U|\mathcal{Y}_t]$, ce qui signifie que les observations futures n'influenceront pas l'estimateur.

La méthode du processus d'innovation. La deuxième approche, utilisée dans le Chapitre 3, isole le mouvement brownien de l'équation d'évolution de μ, appelé le processus d'innovation, puis identifie le terme correspondant dans la décomposition de Doob-Meyer de μ.

Afin de définir le processus de distribution conditionnelle non-normalisé, nous énonçons d'abord la formule de Kallianpur-Striebel. Supposons que la condition (1.15) soit vérifiée pour chaque $\phi \in \mathcal{B}(\mathbb{R}^n)$, nous pouvons prouver que, à t fixé,

$$\mu_t(\phi) = \frac{\overline{\mathbb{E}}[\overline{Z}_t \phi(X_t)|\mathcal{Y}]}{\overline{\mathbb{E}}[\overline{Z}_t|\mathcal{Y}]}, \quad \overline{\mathbb{P}} \text{ et } \mathbb{P}\text{-p.s.}$$

Soit $\xi = \{\xi_t,\ t \geq 0\}$, le processus défini par

$$\xi_t = \overline{\mathbb{E}}[\overline{Z}_t|\mathcal{Y}_t],$$

alors, comme $\overline{Z}_t$ est une $\mathcal{F}_t$-martingale sous $\overline{P}$ et $\mathcal{Y}_s \subseteq \mathcal{F}_s$, il s'ensuit que pour $0 \leq s \leq t$,

$$\overline{\mathbb{E}}[\xi_t|\mathcal{Y}_s] = \overline{\mathbb{E}}[\overline{Z}_t|\mathcal{Y}_s] = \overline{\mathbb{E}}\left[\overline{\mathbb{E}}[\overline{Z}_t|\mathcal{F}_s]|\mathcal{Y}_s\right] = \overline{\mathbb{E}}[\overline{Z}_s|\mathcal{Y}_s] = \xi_s.$$

Nous pouvons prouver qu'il est possible de choisir une version càdlàg de ξ_t qui est une $\mathcal{Y}_t$-martingale.
Étant donné un tel processus ξ_t, nous définissons la distribution conditionnelle non-normalisée de X comme étant le processus $\pi = \{\pi_t,\ t \geq 0\}$, déterminé par les valeurs de $\pi_t(\phi)$ pour $\phi \in \mathcal{B}(\mathbb{R}^n)$ qui sont donnés pour $t \geq 0$ par

$$\pi_t(\phi) := \mu_t(\phi)\xi_t.$$

Nous pouvons montré que le processus $\{\pi_t,\ t \geq 0\}$ est càdlàg et $\mathcal{Y}_t$-adapté et pour tout $t \geq 0$,

$$\pi_t(\phi) = \overline{\mathbb{E}}[\overline{Z}_t \phi(X)_t|\mathcal{Y}_t] \quad \overline{\mathbb{P}} \text{ and } \mathbb{P}\text{-p.s.}$$

De plus, si nous supposons à nouveau la condition (1.15) satisfaite, pour chaque $\phi \in \mathcal{B}(\mathbb{R}^n)$,

$$\mu_t(\phi) = \frac{\pi_t(\phi)}{\pi_t(\mathbf{1})}, \quad \forall\, t \in [0, \infty), \quad \overline{\mathbb{P}} \text{ and } \mathbb{P}\text{-p.s.}$$

La formule de Kallinapur-Striebel justifie l'utilisation du terme non-normalisé dans la définition de π_t car le dénominateur $\pi_t(\mathbf{1})$ peut être considéré comme le facteur de normalisation. Ce résultat peut également être considéré comme la version abstraite de l'identité de Bayes dans ce cadre de filtrage.

Plus généralement, les modèles avec observations partielles ont été étudiés par Detemple (1986), Dothan and Feldman (1986) et Gennotte (1986) dans un cadre de filtrage gaussien linéaire. Karatzas and Xue (1991) ont introduit une approche bayésienne pour les problèmes de maximisation d'espérance d'utilité, en utilisant le filtrage et la théorie de la représentation martingale. Dans ce cadre Lakner (1995), Lakner (1998) et Zohar (2001) ont résolu les problèmes d'optimisation via l'approche martingale, Kuwana (1995) a étudié les conditions nécessaires et suffisantes pour que le principe d'équivalent-certain se vérifie, et Karatzas (1997) a abordé le problème de maximisation de la probabilité d'atteindre un objectif donné pendant un certain horizon temporel. Pour un processus de *drift* inobservable entraîné par un mouvement brownien indépendant, le problème d'optimisation a été étudié par Rishel (1999) pour les fonctions d'utilité de type puissance. Le cas particulier de la fonction d'utilité logarithmique et de la distribution *a priori* gaussienne a été étudié par Browne and Whitt (1996) sur un horizon infini.
Nous pouvons également nous référer aux articles de Rogers (2001), Karatzas and Zhao (2001) et Cvitanić et al. (2006) qui utilisent des techniques de filtrage et d'apprentissage, pour résoudre des problèmes d'optimisation stochastique avec diverses contraintes dans un cadre d'observation partielle. Les stratégies de couverture, minimisant les risques pour les créances contingentes dans un cadre d'observation partielle, ont été abordées dans Ceci (2006), tandis que Ceci and Gerardi (2006) ont considéré un modèle pour les mouvements de cours boursiers intra-journaliers où l'intensité du saut dépend d'une variable d'état cachée non-observable. Le problème de filtrage pour un processus général de diffusion à saut a été traité par Ceci and Colaneri (2012), et la prise en compte d'informations partielles dans le contexte des équations différentielles stochastiques *backward* ont été examinées dans Ceci et al. (2014). Récemment, Bismuth et al. (2019) ont étudié le choix optimal de portefeuilles en tenant compte, dans le processus de décision, à la fois de la liquidité des actifs et de l'incertitude concernant leurs rendements attendus, et De Franco, Nicolle, and Pham (2019a) ont traité le problème de sélection de portefeuilles de Markowitz avec information partielle.

1.1.4 Maximum Drawdown

Le Chapitre 5 se concentre sur un problème de choix de portefeuille, plus précisément de maximisation d'espérance d'utilité satisfaisant une contrainte de *maximum draw-*

down. En finance, d'importants problèmes de maximisation d'espérance d'utilité ont été introduits par Merton (1971) dans le contexte de coefficients constants, et ont été traités par des méthodes markoviennes de contrôle stochastique en temps continu, voir Karatzas et al. (1998). Pour des processus à paramètres généraux sur des marchés complets, des méthodologies basées sur la théorie des martingales et la dualité convexe ont été développées par Pliska (1986), Karatzas et al. (1987) et Cox and Huang (1989). Ils ont été étendus dans le cadre de marchés incomplets et/ou contraints par Karatzas et al. (1991), He and Pearson (1991) et Cvitanić and Karatzas (1992).

En ce qui concerne la littérature sur les contraintes de risque dans un cadre d'allocation de portefeuille, nous nous référons à l'article de Redeker and Wunderlich (2018) qui traite de contraintes de risque dynamiques, en comparant le cas continu et le cas discret, et aux articles de Grossman and Zhou (1993) et Cvitanić and Karatzas (1994) qui se concentrent, en particulier, sur une contrainte de type *maximum drawdown*. Plus récemment, Elie and Touzi (2008) étudient le problème optimal de consommation-investissement à horizon infini en temps continu, et Boyd et al. (2019) utilisent des prévisions de la moyenne et de la covariance des rendements financiers à partir d'un modèle de Markov caché multivarié avec paramètres variant dans le temps, pour construire les contrôles optimaux.

La contrainte de risque qui nous intéresse dans le Chapitre 5 est le *maximum drawdown* (MD), c'est pourquoi nous nous attardons brièvement sur cet indicateur.

Le MD, illustré sur la figure 1.2, est une mesure de risque largement considérée dans l'industrie et est définie comme la perte maximale observée d'un portefeuille, d'un pic à un creux, avant qu'un nouveau pic ne soit atteint. Il s'agit d'un indicateur du risque de baisse sur une période donnée. Exprimé en termes de pourcentage et prenant ses valeurs dans l'intervalle $[0, -100\%]$, il peut servir en tant que tel ou bien dans d'autres mesures, telles que le Calmar ratio qui se calcule en prenant la valeur absolue du ratio rendement attendu sur MD.

Pour une trajectoire X_t, $t \in [0, T]$, la formule pour calculer le MD au temps t est :

$$MD_t = \min_{0 \le s \le t} \left[\frac{\min_{s \le u \le t} X_u}{\max_{0 \le u \le s} X_u} - 1 \right].$$

Il convient de noter que le MD ne mesure que la taille de la plus grande perte, sans tenir compte de la fréquence des pertes. Étant donné qu'il ne mesure que le plus grand *drawdown*, le MD n'indique pas combien de temps il a fallu à un investisseur

pour se remettre de cette perte.

Un MD proche de zéro est préférable, car cela indique que les pertes du portefeuille sont faibles. Si un portefeuille voit constamment sa valeur augmenter, le MD sera nul. Le pire MD possible est -100%, ce qui signifie que la valeur du portefeuille est nulle.

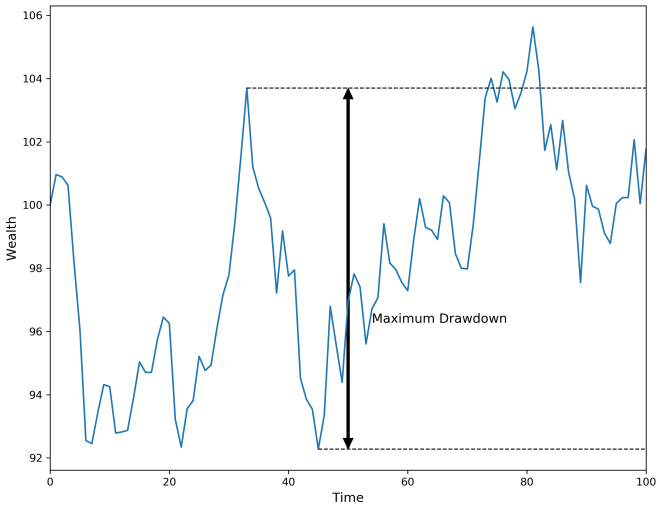


Figure 1.2: Illustration du maximum drawdown sur une trajectoire simulée.

1.2 Optimisation de portefeuilles avec incertitude de paramètres - contributions

1.2.1 Chapitre 3

Motivation. Dans sa forme standard, les paramètres du modèle de Markowitz, *drift* et volatilités, sont connus et constants. Cette hypothèse pose la question de la valorisation de ces paramètres. En règle générale, ils sont estimés à partir de données antérieures ou calculés sur la base d'analyses prédictives, et sont considérés constants ensuite. Cette solution ne semble pas appropriée en pratique car elle ne s'adapte pas à l'évolution des régimes de marché, matérialisée par l'information contenue dans les changements de prix. Ne pas considérer cette information est préjudiciable au processus d'optimisation, car elle ne permet pas une mise à jour du modèle, en tenant compte des informations disponibles les plus récentes.

Dans notre approche, nous incorporons l'incertitude dans le paramètre de *drift* tout en gardant le paramètre de covariance connu et constant. Ce choix est soutenu, par exemple par l'article de Merton (1980), qui montre que les estimations des variances et des covariances des rendements des actifs sont plus précises que les estimations de leurs rendements attendus. De plus, une erreur dans l'estimation des rendements attendus peut entraîner un portefeuille très sous-optimal *a posteriori*. Best and Grauer (1991) ont, en effet, montré que les portefeuilles optimaux moyenne-variance sont très sensibles au niveau des rendements attendus. Précisément, notre approche considère que le *drift* possède une distribution de probabilité connue qui sera mise à jour en utilisant l'apprentissage bayésien.

Description du problème. Nous modélisons un marché financier sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ équipé d'une filtration $\{\mathcal{F}_t\}_{t\geq 0}$ qui satisfait aux conditions habituelles. Nous définissons un mouvement brownien standard n-dimensionnel W, qui pilote la dynamique des n actifs risqués continuellement échangés sur ce marché. Le processus de prix des actifs S est défini comme suit :

$$\begin{cases} dS_t &= \operatorname{diag}(S_t)\big(Bdt + \sigma dW_t\big), \quad t \in [0,T], \\ S_0 &= s_0 \in \mathbb{R}^n, \end{cases} \tag{1.16}$$

où T est un horizon d'investissement fini, $B \in \mathbb{R}^n$ est un vecteur aléatoire indépendant de W, avec une distribution μ et tel que $\mathbb{E}[|B|^2] < \infty$. La distribution de probabilité *a priori* de B représente les croyances subjectives de l'investisseur quant à la probabilité des différentes valeurs que B pourrait prendre.

Dans notre cadre, nous considérons les hypothèses suivantes satisfaites :

- $\operatorname{Cov}(B) = \mathbb{E}\left[(B - \mathbb{E}[B])(B - \mathbb{E}[B])'\right]$ est définie positive,
- σ est connue et inversible.

Comme l'investisseur observe le processus de prix S mais pas le *drift* B, la filtration $\mathbb{F}^{\mathbb{S}} := \{\mathcal{F}_t^S\}_{t\geq 0}$ générée par le processus de prix S, augmenté des ensembles nuls de $\mathcal{F}$, représente la seule source d'information disponible.

Le processus à valeurs dans $\mathbb{R}^n$ et $\mathbb{F}^{\mathbb{S}}$-progressivement mesurable $\alpha = (\alpha_t)_{0\leq t\leq T}$, est une stratégie d'investissement admissible représentant le montant investi dans chacun des actifs risqués et satisfait une condition d'intégrabilité appropriée[1].

On note $\mathcal{A}$ l'ensemble de toutes les stratégies admissibles d'investissement.

[1] Détaillée dans le Chapitre 3.

La dynamique du processus de richesse autofinancée X^α, contrôlée par $\alpha \in \mathcal{A}$ et commençant à une richesse initiale $x_0 \in \mathbb{R}$, s'écrit

$$dX_t^\alpha = \alpha_t' \text{diag}(S_t)^{-1} dS_t = \alpha_t'(Bdt + \sigma dW_t), \quad 0 \leq t \leq T, \quad X_0^\alpha = x_0. \tag{1.17}$$

Nous définissons notre problème de la manière suivante,

$$U_0(\vartheta) \;:=\; \sup_{\alpha \in \mathcal{A}} \Big\{ \mathbb{E}[X_T^\alpha] : \text{Var}(X_T^\alpha) \leq \vartheta \Big\}, \tag{1.18}$$

où $\vartheta > 0$ est la variance maximale que l'investisseur peut supporter sur la valeur terminale de son portefeuille. Bien qu'identique dans sa formulation avec le problème de Markowitz standard, notre problème se distingue en considérant que l'incertitude de *drift* des actifs est mise à jour grâce au filtrage bayésien. Nous l'appelons le problème Bayésien-Markowitz dans la suite.

Nouveaux résultats. Nous formulons dans un cadre multidimensionnel, le problème de Markowitz dans le cas d'un *drift* incertain modélisée par une distribution *a priori*.
Partant du problème de Markowitz (1.18), nous écrivons le problème moyenne-variance associé, voir Lemma 3.2.3. Ce problème, *a priori* inconsistant en temps, ne peut pas être résolu par dualité convexe ou par une méthode martingale comme dans Karatzas and Zhao (2001). Au lieu de cela, en utilisant la méthodologie de Zhou and Li (2000), nous montrons comment le problème moyenne-variance associé, peut être intégré dans un problème de contrôle standard auxiliaire qui se résout par une approche de programmation dynamique, voir Lemma 3.2.4.

Bien que les lois *a priori* conjuguées soient largement utilisés dans la littérature, dans notre cadre, nous pouvons considérer toute distribution *a priori* pour modéliser le *drift* incertain. Cela vient de l'extension au cas multidimensionnel d'un résultat récent d'Ekstrom and Vaicenavicius (2016), que nous établissons dans le Lemme 3.3.1.

Décrivons les principaux résultats théoriques du Chapitre 3. Tout d'abord, nous transformons le modèle non observable (1.16)-(1.17) en un modèle observable. Pour ce faire, nous définissons $\hat{B}_t = \mathbb{E}[B|\mathcal{F}_t^S]$ et introduisons le processus d'innovation

$$\hat{W}_t \;:=\; \sigma^{-1} \int_0^t (B - \hat{B}_s) ds + W_t, \quad 0 \leq t \leq T.$$

Cela nous amène au modèle observable suivant

$$\begin{cases} dS_t &= \text{diag}(S_t) \left(\hat{B}_t dt + \sigma d\hat{W}_t \right), \quad t \in [0, T], \\ S_0 &= s_0 \in \mathbb{R}^n, \end{cases}$$

avec la dynamique correspondante pour le processus de richesse

$$dX_t^\alpha \;=\; \alpha_t' \mathrm{diag}(S_t)^{-1} dS_t = \alpha_t'(\hat{B}_t dt + \sigma d\hat{W}_t), \quad 0 \le t \le T, \quad X_0^\alpha = x_0.$$

Afin d'écrire la dynamique de $\hat{B}_t$, nous remarquons que le processus d'observation

$$Y_t \;:=\; \sigma^{-1} \int_0^t \hat{B}_s ds + \hat{W}_t, \tag{1.19}$$

peut s'écrire $Y_t = \sigma^{-1} Bt + W_t$ qui s'avère être une fonction mesurable de S_t. Ceci implique que Y_t génère la même filtration que S_t, et par le Lemme 3.3.1, nous pouvons écrire

$$\hat{B}_t = \mathbb{E}[B|Y_t] := f(t, Y_t).$$

En appliquant la formule d'Itô à $\hat{B}_t = f(t, Y_t)$, avec l'Equation (1.19) et en rappelant que $\hat{B}$ est une $(\mathbb{P}, \mathcal{F}^S)$-martingale, on voit que $d\hat{B}_t = \nabla_y f(t, Y_t) d\hat{W}_t$. En notant $f_t(y) := f(t, y)$ et en définissant la fonction matricielle ψ par

$$\psi(t, b) \;:=\; \nabla_y f(t, f_t^{-1}(b)), \quad t \in [0, T],\; b \in f_t(\mathbb{R}^n) \;=\; \mathcal{B}_t,$$

où $y = f_t^{-1}(b)$ est la valeur unique du processus d'observation Y_t qui donne $\hat{B}_t = b$, le Lemme 3.3.2 nous donne finalement la dynamique de $\hat{B}_t$,

$$d\hat{B}_t \;=\; \psi(t, \hat{B}_t) d\hat{W}_t,\; t \in [0, T], \quad \hat{B}_0 \;=\; b_0 \;= \mathbb{E}[B].$$

Ensuite, nous prouvons que notre problème revient à résoudre une équation de HJB non-linéaire qui peut être réduite à l'EDP semi-linéaire suivante sur $\mathcal{R} = \{(t, b) : t \in [0, T], b \in \mathcal{B}_t\}$:

$$\begin{cases} -\partial_t R - \dfrac{1}{2}\mathrm{tr}\big(\psi\psi' \mathcal{D}_b^2 R\big) + \tilde{F}(t, b, \nabla_b R) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad R(T, b) = 0, \quad b \in \mathcal{B}_t, \end{cases} \tag{1.20}$$

avec

$$\tilde{F}(t, b, p) := 2\left(\psi\sigma^{-1} b\right)' p - \frac{1}{2}|\psi' p|^2 - |\sigma^{-1} b|^2. \tag{1.21}$$

De plus, nous prouvons dans le Théorème 3.4.3 que, si les conditions suivantes sont vérifiées,

- $\forall (t, b) \in \mathcal{R}$, $\psi(t, b)$ et $\psi^{-1}(t, b)$ sont bornées, et
- la matrice $\nabla_b \psi$ existe et est bornée,

alors il existe une solution classique $R \in \mathcal{C}^{1,2}(\mathcal{R})$ à l'EDP (1.20). De plus, $R(t,b)$ croît au plus quadratiquement en b et $\nabla_b R(t,b)$ croît au plus linéairement en b.

Si les conditions précédentes sont remplies, nous montrons dans le Théorème 3.4.1 que nous pouvons résoudre le problème Bayésien-Markowitz avec son contrôle optimal associé.

Pour tout $\vartheta > 0$, la solution du problème Bayésien-Markowitz est

$$U_0(\vartheta) \;=\; x_0 + \sqrt{\vartheta\,(e^{R(0,b_0)} - 1)}, \quad \vartheta > 0,$$

avec la stratégie optimale donnée par

$$\hat{\alpha}_t^{\vartheta} \;=\; a_0^{\vartheta}(t, X_t^{\hat{\alpha}^{\vartheta}}, \hat{B}_t),$$

où

$$\begin{aligned} a_0^{\vartheta}(t,x,b) :=& \Big(x_0 - x + e^{R(0,b_0)}\sqrt{\frac{\vartheta}{e^{R(0,b_0)} - 1}}\Big) \\ &\times \Big(\Sigma^{-1} b - \big(\psi\sigma^{-1}\big)' \nabla_b R(t,b)\Big), \quad (t,b) \in \mathcal{R},\ x \in \mathbb{R}, \end{aligned}$$

et en se rappelant que $\hat{B}_t = \mathbb{E}[B|\mathcal{F}_t^S]$.

Dans le cas particulier d'une distribution *a priori* gaussienne multidimensionnelle, nous trouvons une solution explicite au problème Bayésien-Markowitz (1.18) dans le Lemme 3.4.5.

En plus des résultats théoriques précédents, nous fournissons une analyse de sensibilité complète dans la Section 3.5, qui conclut que l'apprentissage apporte de la valeur à la stratégie optimale du problème de Markowitz avec *drift* incertain.

Difficultés. Voici les difficultés techniques auxquelles nous avons été confrontés.
(1) Le problème initial est transformé en un problème moyenne-variance qui n'est pas consistant en temps en raison du terme d'espérance au carré dans la variance. Par conséquent, nous effectuons une deuxième transformation pour obtenir un problème standard qui peut être résolu en utilisant l'approche de programmation dynamique via l'équation de HJB.
(2) La preuve de l'existence d'une solution régulière à l'EDP (1.20) n'est pas triviale. En effet, les hypothèses habituelles d'existence et d'unicité que nous pouvons trouver, par exemple dans Ladyzhenskaia et al. (1968), ne s'appliquent pas ici puisque la fonction $\tilde{F}$ définie dans (1.21) n'est pas globalement Lipschitz en p. Cette difficulté vient de la croissance quadratique en p de la fonction $\tilde{F}$, qui implique que la solution

puisse croître quadratiquement, et du fait que son domaine de définition n'est pas borné.

Ouverture. On peut étendre les résultats précédents en étudiant les applications à des distributions de probabilité *a priori* autres que gaussiennes, et effectuer une analyse de la valeur de l'information.

1.2.2 Chapitre 4

Motivation. Introduire l'incertitude de *drift* dans le problème le rend plus réaliste et le rapproche d'une situation réelle, où non seulement l'investisseur ignore les paramètres du modèle, mais est également forcé d'admettre que ses estimations sont incertaines. Ainsi, notre objectif principal est de confronter aux données historiques de marché, la supériorité théorique de la stratégie apprenante (BL) par rapport à la stratégie non-apprenante (NL) qui considère le *drift* constant. Pour ce faire, nous appliquons les deux stratégies à différents univers d'investissement que nous détaillons dans l'Annexe, Partie 4.A.

Description du problème. Sur le marché financier considéré, sont négociés un actif sans risque, dont le rendement est désigné par r^f, et n actifs risqués dont les rendements r_t sont modélisés par

$$\begin{cases} r_t &= B + \sigma\xi_t \\ \xi_t &\sim N(0, I_n). \end{cases}$$

Nous supposons que les vecteurs aléatoires ξ_t sont indépendants pour tout t. $B \in \mathbb{R}^n$ est le vecteur des rendements attendus des actifs risqués, tandis que $\Sigma := \sigma\sigma' \in \mathbb{R}^{n\times n}$ est leur matrice de covariance. Nous supposons que σ^{-1} existe. Pour $A \subseteq \mathbb{R}^n$, l'ensemble des stratégies admissibles d'investissement est défini par

$$\mathcal{A} := \left\{\mathbf{w} = (\mathbf{w}_t)_{0\leq t\leq T} | \mathbf{w}_t \in \mathrm{A}\right\},$$

où une stratégie admissible $\mathbf{w} = (\mathbf{w}_t)_t$ représente la fraction de richesse investie dans les actifs à tout moment t. Nous écrivons la richesse à l'échéance $T > 0$ ainsi

$$X_T^{\mathbf{w}} = X_0 \prod_{s=0}^{T-1} \left(1 + \mathbf{w}_s' r_{s+1} + (1 - \mathbf{w}_\mathbf{s}' \mathbb{1}_n)\, r^f\right).$$

Nous considérons un investisseur qui vise à résoudre le problème de Markowitz :

$$\begin{cases} \max\limits_{\mathbf{w}\in\mathcal{A}} \mathbb{E}\left[X_T^{\mathbf{w}}\right] \\ \\ \mathrm{Var}\left(X_T^{\mathbf{w}}\right) \leq \vartheta, \end{cases}$$

où $\vartheta > 0$ est la tolérance au risque de l'investisseur.

Dans cette configuration, pour tenir compte de l'incertitude du *drift*, nous supposons raisonnablement que l'investisseur a une vision *a priori* sur les actifs risqués et une croyance sur leurs rendements attendus, mais il n'est pas certain de l'exactitude de ses prévisions. Cela se reflète dans le fait que l'investisseur n'observe pas B et suppose seulement que $B \sim \mu$, où μ est une distribution de probabilité sur $\mathbb{R}^n$, telle que $\mathbb{E}[B] = b_0$. Le paramètre b_0 est le vecteur des rendements attendus par l'investisseur tandis que μ traduit l'incertitude à son sujet.

Nouveaux résultats. La mise en œuvre des stratégies apprenante et non-apprenante sur des données historiques, implique d'apporter quelques ajustements aux résultats théoriques du Chapitre 3. Par exemple, dans la pratique, il y a des problèmes de liquidité et le *trading* se fait en temps discret. Par conséquent, nous discrétisons à la fois, la solution optimale du problème Bayésien-Markowitz en temps continu, et la stratégie non-apprenante. Ensuite nous transformons les montants en proportions en les plafonnant afin de répondre aux contraintes de gestion de portefeuilles.
Nous savons que ces ajustements rendent les solutions optimales en temps continu sous-optimales en temps discret. Néanmoins, nous traitons les stratégies BL et NL de manière égale et nous visons à mettre l'accent sur le gain de l'apprentissage plutôt que sur les performances des stratégies prises indépendamment. Ce faisant, nous acceptons de perdre une forme d'optimalité.

Dans notre étude, nous montrons numériquement la surperformance, mesurée en rendements annualisés et en ratios de Sharpe, de la stratégie apprenante par rapport à la stratégie non-apprenante, que nous considérions des univers d'investissement classiques : actions, devises, ou plus sophistiqués fondés sur des stratégies *smart beta*.
Non seulement, cela confirme les résultats théoriques du Chapitre 3, mais cela illustre également la robustesse de la supériorité de la stratégie apprenante, car elle est vérifiée pour différents univers d'investissement.

Nous effectuons également une analyse approfondie de la sensibilité de la performance de l'apprentissage par rapport au non-apprentissage selon différents paramètres. Nous détaillons les principales conséquences des variations de chacun des paramètres sur la performance relative.

- **Impact de l'incertitude.** Plus le niveau d'incertitude de *drift* est élevé, meilleur est le comportement de la stratégie apprenante par rapport à celle qui n'apprend pas. Néanmoins, un niveau d'incertitude (très) élevé autorise le *drift*

à parcourir une région très vaste qui ralentit tant le processus d'apprentissage, qu'il n'ajoute aucune valeur informative à la stratégie apprenante, ce qui implique que la stratégie non-apprenante peut surperformer.

- **Impact de l'effet de levier.** Dans notre cadre, le paramètre de levier est le niveau plafond des pondérations de chaque stratégie. Par conséquent, à mesure que ce paramètre augmente, les stratégies deviennent sans contraintes et sont plus sensibles à toute erreur dans les paramètres estimés. Comme attendu, plus l'effet de levier est élevé, plus les performances des deux stratégies sont élevées au prix d'un risque plus élevé. Les stratégies apprenantes sont plus performantes que celles qui n'apprennent pas, à tous les niveaux de levier pris en compte dans notre étude.

- **Impact de la fréquence de révision.** La fréquence de révision est la fréquence à laquelle nous recalibrons les paramètres du modèle. Notre principale conclusion est que, lorsque nous diminuons la fréquence, l'amélioration relative des ratios de Sharpe augmente considérablement. En effet, une faible fréquence de révision permet à la stratégie apprenante de tirer pleinement parti de sa fonction d'apprentissage et de surpasser la stratégie qui n'apprend pas, coincée avec son estimation initiale.

- **Impact de la fréquence de rebalancement.** Dans notre configuration, nous montrons que la stratégie d'apprentissage bayésien n'a aucune réelle incitation à augmenter sa fréquence de rebalancement, du mensuel au bimensuel. Pour des raisons détaillées dans le Chapitre 4, une fréquence plus élevée de rebalancement implique une augmentation de la surperformance relative de la stratégie d'apprentissage. Cela vient des mauvaises performances de la stratégie non-apprenante, plutôt que des performances plus élevées de la stratégie apprenante.

Difficultés. Les difficultés proviennent de la nécessité de procéder à des ajustements avant de mettre en œuvre les stratégies avec des données historiques. En particulier, nous avons dû définir une politique de gestion de portefeuilles sous la forme du *workflow* décrit dans la Section 4.3, et ajuster les paramètres associés. L'interprétation des résultats de sensibilité a également été une partie difficile de l'étude, car nous devons extraire une analyse significative souvent mélangée à du bruit.

Ouverture. Puisque la discrétisation des solutions optimales en temps continu les rend sous-optimales en temps discret, une extension naturelle de ce travail serait de dériver les solutions optimales discrètes des deux stratégies et de comparer les résultats.
Une autre extension intéressante pourrait être d'utiliser des distributions de probabilités *a priori* différentes de la gaussienne, éventuellement empirique sur une fenêtre de calcul donnée pour s'adapter au temps discret.

1.3 Optimisation de portefeuilles avec contraintes de risques - contributions

1.3.1 Chapitre 5

Motivation. Dans le Chapitre 5, nous étudions un problème d'optimisation de portefeuilles qui a pour objectif de maximiser l'espérance de l'utilité de la richesse terminale d'un portefeuille, compte tenu d'une contrainte de *maximum drawdown* avec des actifs dont la dynamique possède un *drift* inobservable. À notre connaissance, ce problème d'optimisation qui se pose en pratique, n'a jamais été traité, et nous proposons une manière originale de le résoudre numériquement. Nous utilisons une méthode avancée d'apprentissage profond, fondée sur des réseaux de neurones et un algorithme[2] spécifique, pour proposer une réponse numérique à ce problème qui n'a pas de solution explicite.

Description du problème. On considère un espace de probabilité $(\Omega,\ \mathcal{F},\ \mathbb{P})$ équipé d'une filtration discrète $\{\mathcal{F}_k\}_{k=0,\ ...,\ N}$ satisfaisant les conditions habituelles, sur lesquelles est considéré un modèle de marché financier avec un actif sans risque supposé normalisé à un et d actifs risqués. Le processus de prix $(S^i_k)_{k=0,\dots,N}$ de l'actif $i \in [\![1, d]\!]$ évolue ainsi

$$S^i_{k+1} \ = \ S^i_k e^{R^i_{k+1}}, \quad k = 0, \dots, N-1, \tag{1.22}$$

où $R_{k+1} = (R^1_{k+1}, \dots, R^N_{k+1})$ est le *log-return* des actifs, entre le temps k et $k+1$, dont la dynamique est :

$$R_{k+1} \ = \ B + \epsilon_{k+1}.$$

Le *drift* B est une variable aléatoire d-dimensionnelle avec une distribution de probabilité *a priori* μ_0, de moyenne connue $b_0 = \mathbb{E}[B]$ et un moment d'ordre deux fini.

[2] L'algorithme est détaillé dans Algorithm 1 p.155

Le bruit $\epsilon = (\epsilon_k)_k$ est une suite de variables vectorielles aléatoires i.i.d centrées, supposée indépendante de B, avec pour matrice de covariance $\Gamma = \mathbb{E}[\epsilon_k \epsilon_k']$. Nous supposons également que la distribution de probabilité ν de ϵ_k admet une fonction de densité strictement positive g sur $\mathbb{R}^d$ par rapport à la mesure de Lebesgue.

Le processus de prix S est observable, et grâce à (1.22), notons que R peut être déduit de S et vice-versa. Ainsi, nous notons $\mathbb{F}^o = \{\mathcal{F}_k^o\}_{k=0,\,\ldots,\,N}$, la filtration d'observation générée par S (donc de manière équivalente par R) augmentée des ensembles nuls de $\mathcal{F}$, avec la convention que pour $k = 0$, $\mathcal{F}_0^o$ est la tribu triviale.

Dans ce cadre, une stratégie d'investissement est un processus $\mathbb{F}^o$-progressivement mesurable $\alpha = (\alpha_k)_{k=0,\,\ldots,\,N-1}$, à valeur dans $\mathbb{R}^d$, représentant la proportion de la richesse actuelle investie dans chacun des d actifs risqués pour chaque $k = 0,\ \ldots,\ N-1$. Compte tenu d'une stratégie d'investissement α et d'une richesse initiale $x_0 > 0$, le processus de richesse (autofinancé) X^α évolue selon la dynamique

$$\begin{cases} X_{k+1}^\alpha = X_k^\alpha \left(1 + \alpha_k' \left(e^{R_{k+1}} - \mathbb{1}_d\right)\right), & k = 0,\ \ldots,\ N-1, \\ X_0^\alpha = x_0. \end{cases}$$

où $e^{R_{k+1}}$ est la variable aléatoire d-dimensionnelle dont les composantes $\left[e^{R_{k+1}}\right]_i = e^{R_{k+1}^i}$ pour $i \in [\![1, d]\!]$, et $\mathbb{1}_d$ est le vecteur d-dimensionnel dont les composantes sont toutes égales à 1.

Nous introduisons le processus Z_k^α, comme le maximum au temps k du processus de richesse X^α, formulé ainsi,

$$Z_k^\alpha \ := \ \max_{0 \le \ell \le k} X_\ell^\alpha, \qquad k = 0,\ \ldots,\ N.$$

La contrainte de *maximum drawdown* oblige la richesse X_k^α à rester au-dessus d'une fraction $q \in (0, 1)$ du maximum historique actuel Z_k^α.
Nous définissons ensuite l'ensemble des stratégies d'investissement admissible $\mathcal{A}_0^q$ comme l'ensemble des stratégies d'investissement α, telles que

$$X_k^\alpha \ \ge \ qZ_k^\alpha, \quad \text{p.s.}, \quad k = 0, \ldots, N.$$

Le problème de sélection de portefeuille est ainsi formulé :

$$V_0 \ := \ \sup_{\alpha \in \mathcal{A}_0^q} \mathbb{E}\left[U\left(X_N^\alpha\right)\right], \tag{1.23}$$

où U est une fonction d'utilité sur $(0, \infty)$ satisfaisant aux conditions standards d'Inada.

Nouveaux résultats. Nous incorporons l'incertitude de *drift* multidimensionnel dans un cadre bayésien, et établissons des résultats généraux en utilisant un changement de mesure approprié pour dériver l'équation de programmation dynamique.

Nous présentons maintenant les principaux résultats du Chapitre 5.

En suivant le raisonnement de la Section 5.3, nous pouvons écrire l'équation de programmation dynamique associée à (1.23) en forme *backward* ainsi

$$\left\{\begin{array}{rcl} v_N(x,z,\mu) & = & U(x)\mu(\mathbb{R}^d), \quad (x,z)\in\mathcal{S}^q,\ \mu\in\mathcal{M}_+, \\ v_k(x,z,\mu) & = & \sup\limits_{a\in A^q(x,z)} \overline{\mathbb{E}}\Big[v_{k+1}\Big(x\big(1+a'\big(e^{R_{k+1}}-\mathbb{1}_d\big)\big), \\ & & \qquad \max\big[z, x\big(1+a'\big(e^{R_{k+1}}-\mathbb{1}_d\big)\big)\big], \frac{g(R_{k+1}-\cdot)}{g(R_{k+1})}\mu\Big)\Big], \end{array}\right. \tag{1.24}$$

pour $k=0,\dots,N-1$, où

$$\mathcal{S}^q := \big\{(x,z)\in(0,\infty)^2 : qz\le x\le z\big\},$$

$\mathcal{M}_+$ est l'ensemble des mesures positives sur $\mathbb{R}^d$, et

$$A^q(x,z) \;=\; \Big\{a\in\mathbb{R}^d_+ : \sum_{i=1}^d |a_i| \le 1-q\frac{z}{x}\Big\}$$

est l'ensemble statique de contrôles admissibles.

Remarquons grâce au Lemme 5.3.1, que l'espérance dans la formule ci-dessus n'est prise que par rapport au bruit R_{k+1}, qui est sous $\overline{\mathbb{P}}$ distribué selon la loi de probabilité ν de densité g sur $\mathbb{R}^d$.

Dans le cas où la fonction d'utilité est de type CRRA (*Constant Relative Risk Aversion*), c'est à dire,

$$U(x) \;=\; \frac{x^p}{p}, \quad x>0, \quad \text{avec } p<1,\ p\neq 0, \tag{1.25}$$

on peut réduire la dimensionalité du problème et montrer que ce dernier revient à résoudre le système *backward* suivant :

$$\left\{\begin{array}{rcl} w_N(r,\mu) & = & \frac{r^p}{p}\mu(\mathbb{R}^d), \quad r\in[q,1],\ \mu\in\mathcal{M}_+, \\ w_k(r,\mu) & = & \sup\limits_{a\in A^q(r)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min\big[1, r\big(1+a'\big(e^{R_{k+1}}-\mathbb{1}_d\big)\big)\big], \dfrac{g(R_{k+1}-\cdot)}{g(R_{k+1})}\mu\big)\Big], \end{array}\right. \tag{1.26}$$

pour $k = 0, \dots, N-1$, où

$$A^q(r) \;=\; \Big\{ a \in \mathbb{R}^d_+ : a' \mathbb{1}_d \leq 1 - \frac{q}{r} \Big\}.$$

Dans le cadre gaussien, où le bruit et la croyance *a priori* sur le *drift* sont modélisés par une distribution gaussienne, le filtrage bayésien est simplement le filtrage de Kalman, ce qui transforme le système de programmation dynamique en un problème de dimension finie.

Nous prouvons que notre problème de sélection de portefeuilles initial (1.23) peut être reformulé comme un problème d'observation complet dont le système de programmation dynamique associé satisfait :

$$\begin{cases} \tilde{v}_N(x,z,b) &=& U(x), \quad (x,z) \in \mathcal{S}^q,\ b \in \mathbb{R}^d, \\ \tilde{v}_k(x,z,b) &=& \sup\limits_{a \in A^q(x,z)} \mathbb{E}\Big[\tilde{v}_{k+1}\Big(x\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big), \\ & & \qquad \max\big[z, x\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big)\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

pour $k = 0, \dots, N-1$, où K_{k+1} est la matrice de gain de Kalman.
Notons que dans la formule ci-dessus, l'espérance est prise par rapport au vecteur d'innovation $\tilde{\epsilon}_{k+1}$, qui est distribué selon $\mathcal{N}(0, \tilde{\Gamma}_{k+1})$[3].

De plus, dans le cas d'une fonction d'utilité comme dans (1.25), le problème se résume à résoudre le système *backward* suivant sur $[q,1] \times \mathbb{R}^d$:

$$\begin{cases} \tilde{w}_N(r,b) &=& \frac{r^p}{p}, \quad r \in [q,1],\ b \in \mathbb{R}^d, \\ \tilde{w}_k(r,b) &=& \sup\limits_{a \in A^q(r)} \mathbb{E}\Big[\tilde{w}_{k+1}\Big(\min\big[1, r\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big)\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

pour $k = 0, \dots, N-1$.

Nous appliquons ensuite l'étude théorique précédente à un modèle de marché à trois actifs risqués, portant la dimension du problème à cinq. Nous résolvons numériquement le problème, dans le cas gaussien avec fonction d'utilité de type CRRA, en utilisant l'algorithme d'apprentissage profond *Hybrid-Now*[4].
Ces résultats numériques nous permettent de fournir une analyse détaillée de la performance et des allocations des stratégies apprenante et non-apprenante comparées à une stratégie *equally weighted* appropriée. De plus, nous évaluons la valeur de

[3] La matrice $\tilde{\Gamma}_{k+1}$ est explicitement calculée dans la Section 5.4.1
[4] Bachouch et al. (2018a)

l'information par rapport au niveau d'incertitude de *drift*. Enfin, nous fournissons des preuves empiriques de la convergence de la stratégie non-apprenante vers la solution du problème classique de Merton, lorsque le paramètre contrôlant le *maximum drawdown* tend vers 0.

Difficultés. Les difficultés rencontrées dans ce problème sont triples :
1) Trouver un moyen d'incorporer la contrainte de *maximum drawdown* et l'incertitude de *drift* dans une configuration de programmation dynamique. Nous transformons la contrainte de *maximum drawdown* sur le processus de richesse en contrainte sur l'espace des contrôles, tout en gérant le *drift* incertain grâce à une approche d'apprentissage bayésien.
2) Résoudre numériquement le problème et trouver l'approche numérique appropriée. Nous choisissons d'utiliser un algorithme récemment développé, appelé *Hybrid-Now*, fondé sur des réseaux de neurones.
3) Réduire la dimensionalité du problème. Dans le cas gaussien et avec une fonction d'utilité de type CRRA, nous sommes capables de réduire la dimensionalité du problème afin d'obtenir des variables d'état aisément simulables. Ceci est important dans la phase d'apprentissage des réseaux de neurones, car nous pouvons facilement atteindre toutes les valeurs possibles des variables d'état via des simulations. Sans ce changement de variable, il aurait été plus difficile de simuler les variables d'état, par exemple celle représentant le maximum historique actuel.

Ouverture. Ce travail peut être étendu dans plusieurs directions. On peut penser à différentes lois *a priori* pour le *drift*, et étudier le cas d'autres contraintes de risque appliquées au portefeuille. On peut également rechercher un moyen de simuler le maximum historique actuel d'un processus aléatoire donné, afin de généraliser notre méthode numérique de résolution. Finalement, on pourrait tester ces méthodes avec des données historiques et mesurer la valeur informative de la stratégie apprenante par rapport à la non-apprenante, et sa sensibilité à un paramètre d'incertitude bien choisi, formalisé par la matrice de covariance du *drift*.

Chapter 2

Introduction

In this thesis, we investigate portfolio optimization with different risk constraints in the context of drift uncertainty. The manuscript is divided into two main parts that can be read independently. In the first part, we study the multidimensional Markowitz portfolio selection problem in continuous time with drift uncertainty (Chapter 3) and we apply these theoretical results to actual market data in order to study, realistically, the learning effect on performance (Chapter 4). In the second part of the thesis, we are interested in a discrete-time portfolio optimization problem in the context of drift uncertainty with a maximum drawdown constraint and its numerical resolution using deep learning (Chapter 5).

This chapter reviews the literature on the different topics and summarizes the main contributions.

Contents

2.1 Literature Review 34

2.1.1 Optimal stochastic control problems: formulation and resolution 34

2.1.2 Markowitz problem: formulation and continuous-time solution 42

2.1.3 Filtering theory and the Bayesian learning approach . . . 45

2.1.4 Maximum drawdown 50

2.2 Portfolio optimization under parameters uncertainty - contributions 51

2.2.1 Chapter 3 51

2.2.2 Chapter 4 56

2.3 Portfolio optimization under risk measure constraints - contributions 59
2.3.1 Chapter 5 59

2.1 Literature Review

2.1.1 Optimal stochastic control problems: formulation and resolution

We begin by outlining the basic structure of stochastic optimization problems in continuous time and finite horizon while describing the main approach to solve them. This theory and the proofs of the results can be found in Pham (2009).

Consider a dynamic system formalized by its state at any time and evolving in an uncertain environment represented by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Usually, we denote by $X_t(\omega)$ the state of the system at time t in case $\omega \in \Omega$ is the state of the world. The dynamics of the system $t \mapsto X_t(\omega)$ is described by a stochastic differential equation (SDE).
The dynamics of the system is influenced by a control, expressed as a process $\alpha = (\alpha_t)_t$, whose value is decided at each time t in function of the available information.
The control should respect some constraints assumed in the problem, and as such, is called admissible.

In the optimization problems we are interested in, noting x_0 the initial value of the process X_t, the purpose is to maximize a functional $J(x_0, \alpha)$ over all admissible controls. We are considering objective functionals of the form

$$\mathbb{E}\left[\int_0^T f(X_t, \omega, \alpha_t)dt + g(X_T, \omega)\right], \text{ on a finite horizon } T < \infty,$$

with f being a running profit function and g a terminal reward function.

Noting

$$J(x_0, \alpha) = \mathbb{E}\left[\int_0^T f(X_t, \omega, \alpha_t)dt + g(X_T, \omega)\right],$$

the maximum value is defined by

$$v = \sup_{\alpha} J(x_0, \alpha).$$

As a consequence, the main purpose of this stochastic optimization problem is to find the maximizing control process attaining the value function that needs to be determined.

Now, we introduce the dynamic programming method which is an important tool for solving stochastic control problems.
Basically, the approach is to consider a family of control problems by varying the initial state values and to derive a relation between the associated value functions. This principle was initiated by Bellman (1952). It yields a nonlinear PDE of second order called Hamilton-Jacobi-Bellman (HJB). When this PDE can be solved either explicitly or theoretically with a smooth solution, the verification theorem validates the optimality of the candidate solution to the HJB equation.
This approach supposes the existence of a smooth solution to the HJB equation which is not true in general.

We consider a model where the dynamics of the state of the system is a controlled SDE valued in $\mathbb{R}^n$:

$$dX_s = b(X_s, \alpha_s)ds + \sigma(X_s, \alpha_s)dW_s, \tag{2.1}$$

where W is a m-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions. The control $\alpha = (\alpha_s)$ is a progressively $\mathbb{F}$-measurable process, valued in $A \subset \mathbb{R}^d$.
The measurable functions b: $\mathbb{R}^n \times A \to \mathbb{R}^n$ and σ: $\mathbb{R}^n \times A \to \mathbb{R}^{n\times m}$ satisfy a uniform Lipschitz condition in A: $\exists K \geq 0$, $\forall x,\ y \in \mathbb{R}^n,\ \forall a \in A$,

$$|b(x, a) - b(y, a)| + |\sigma(x, a) - \sigma(y, a)| \leq K|x - y|. \tag{2.2}$$

Let us fix a finite horizon $0 < T < \infty$ and denote by $\mathcal{A}$ the set of control processes α, such that

$$\mathbb{E}\left[\int_0^T |b(0, \alpha_t)|^2 + \sigma(0, \alpha_t)|^2 dt\right] < \infty. \tag{2.3}$$

Conditions (2.2) and (2.3) ensure for all $\alpha \in \mathcal{A}$ and for any initial condition $(t, x) \in [0, T] \times \mathbb{R}^n$, the existence and uniqueness of a strong solution to SDE (2.1) starting from x at $s = t$. We then denote by $\{X_s^{t,x},\ t \leq s \leq T\}$ this solution with almost sure (a.s.) continuous paths.

Let f: $[0, T] \times \mathbb{R}^n \times A \to \mathbb{R}$ and g: $\mathbb{R}^n \to \mathbb{R}$ two measurable functions. We assume that:
- g is lower-bounded

or,

- g satisfies a quadratic growth condition: $|g(x)| \leq C(1+|x|^2)$, $\forall x \in \mathbb{R}^n$, for some constant C independent of x.

For $(t,x) \in [0,T] \times \mathbb{R}^n$, $\mathcal{A}(t,x)$ is the subset of controls α in $\mathcal{A}$ such that

$$\mathbb{E}\left[\int_t^T |f(s, X_s^{t,x}, \alpha_s)|ds\right] < \infty,$$

assuming that $\mathcal{A}(t,x)$ is not empty for all $(t,x) \in [0,T] \times \mathbb{R}^n$. We then define under the previous hypothesis the gain function:

$$J(t,x,\alpha) = \mathbb{E}\left[\int_t^T f(s, X_s^{t,x}, \alpha_s)ds + g(X_s^{t,x})\right],$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$ and $\alpha \in \mathcal{A}(t,x)$. The objective is to maximize the gain function J over the admissible control processes. Let's introduce the associated value function:

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} J(t,x,\alpha). \tag{2.4}$$

Given an initial condition $(t,x) \in [0,T) \times \mathbb{R}^n$, $\hat{\alpha} \in \mathcal{A}(t,x)$ is an optimal control if $v(t,x) = J(t,x,\hat{\alpha})$. We call a Markovian control, a control process α in the form $\alpha_s = a(s, X_s^{t,x})$ for some measurable function $a : [0,T] \times \mathbb{R}^n \mapsto A$. In the sequel, we shall assume that the value function v is measurable with respect to its arguments which is not trivial a priori.

We now describe the dynamic programming principle (DPP) which is a fundamental principle in the theory of stochastic control. In the context of controlled Markov processes, we formulate it as follows:

Theorem 2.1.1 (Dynamic programming principle for finite horizon)**.** *For $0 \leq t \leq T < \infty$, we denote by $\mathcal{T}_{t,T}$ the set of stopping times valued in $[t,T]$. Let $(t,x) \in [0,T] \times \mathbb{R}^n$, we then have*

$$\begin{aligned} v(t,x) &= \sup_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})]\right] \\ &= \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\left[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)ds + v(\theta, X_\theta^{t,x})]\right] \end{aligned}$$

A stronger version than the previous DPP can be written in the finite horizon case as follows:

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}\left[\int_t^\theta f\left(s, X_s^{t,x}, \alpha_s\right)ds + v\left(\theta, X_\theta^{t,x}\right)\right], \tag{2.5}$$

for any stopping time $\theta \in \mathcal{T}_{t,T}$.

The DPP means that an optimal control on the whole interval $[t, T]$ can be obtained first, by finding an optimal control from time θ given the state value $X_\theta^{t,x}$, or equivalently compute $v(\theta, X_\theta^{t,x})$, and then by maximizing on $[t, \theta]$ the expression

$$\mathbb{E}\left[\int_t^\theta f\left(s, X_s^{t,x}, \alpha_s\right) ds + v(\theta, X_\theta^{t,x})\right].$$

In the theory of dynamic programming, the HJB equation is the infinitesimal version of the DPP and explicits the local behavior of the value function when sending the stopping time θ to t. In a standard stochastic optimization problem in continuous time, the following steps are usually followed:

1. Derive the HJB equation,
2. Show existence of a smooth solution using PDE techniques,
3. Verify that the smooth solution is the value function using Itô's formula,
4. Obtain an optimal feedback control as a by-product.

We will now derive formally the HJB equation in the case of a finite horizon $T < \infty$. Let's consider the time $\theta = t + h$ and an arbitrary constant control $\alpha_s = a \in A$ such that, from Theorem 2.1.1,

$$v(t, x) \geq \mathbb{E}\left[\int_t^{t+h} f(s, X_s^{t,x}, a) ds + v(t + h, X_{t+h}^{t,x})\right]. \tag{2.6}$$

If we now assume v is smooth enough to apply Itô's formula between times t and $t + h$:

$$v\left(t + h, X_{t+h}^{t,x}\right) = v(t, x) + \int_t^{t+h} \left(\partial_t v + \mathcal{L}^a v\right)\left(s, X_s^{t,x}\right) ds + \text{local martingale},$$

noting $\mathcal{L}^a$ the operator associated to the diffusion in Equation (2.1) for the constant control a, defined as

$$\mathcal{L}^a v \ = < b(x, a), \nabla_x v > \ + \frac{1}{2}\text{tr}\left(\sigma(x, a)\sigma'(x, a)\mathcal{D}_x^2 v\right).$$

After substituying into Equation (2.6), and assuming a quadratic growth condition satisfied by v and localization arguments to remove the stochastic integral, i.e. the local martingale term in the expectation, we obtain

$$0 \geq \mathbb{E}\left[\int_t^{t+h} \left(\partial_t v + \mathcal{L}^a v\right)\left(s, X_s^{t,x}\right) + f\left(s, X_s^{t,x}, a\right) ds\right].$$

We then divide by h and make it vanish to obtain, by the mean-value theorem,

$$0 \geq \partial_t v(t,x) + \mathcal{L}^a v(t,x) + f(t,x,a).$$

The previous inequality holds true for any $a \in A$, thus we get the following inequality,

$$-\partial_t v(t,x) - \sup_{a \in A} \left[\mathcal{L}^a v(t,x) + f(t,x,a)\right] \geq 0. \tag{2.7}$$

Next suppose that the optimal control is α^* in Equation (2.5),

$$v(t,x) = \mathbb{E}\left[\int_t^{t+h} f\left(s, X_s^*, \alpha_s^*\right) ds + v\left(t+h, X_{t+h}^*\right)\right]$$

where X^* is the solution to SDE (2.1) starting from x at t, with control α^*. Thanks to the same reasoning as before, we get

$$-\partial_t v(t,x) - \mathcal{L}^{\alpha_t^*} v(t,x) - f(t,x,\alpha_t^*) = 0.$$

With Equation (2.7), this suggests that the value function v should satisfy

$$-\partial_t v(t,x) - \sup_{a \in A} \left[\mathcal{L}^a v(t,x) + f(t,x,a)\right] = 0, \quad \forall (t,x) \in [0,T) \times \mathbb{R}^n$$

when the supremum in a is finite. Techniques exist when the supremum is infinite, which may typically arise when the control space A is unbounded and we refer to Pham (2009) for more details on this issue.

The previous PDE is often rewritten in the form

$$-\partial_t v(t,x) - H\left(t,x,\nabla_x v(t,x), \mathcal{D}_x^2 v(t,x)\right) = 0, \quad \forall (t,x) \in [0,T) \times \mathbb{R}^n, \tag{2.8}$$

where for $(t,x,p,M) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_n$,

$$H(t,x,p,M) = \sup_{a \in A}\left[< b(x,a), p > + \frac{1}{2}\mathrm{tr}\left(\sigma\sigma'(x,a)M\right) + f(t,x,a)\right].$$

The function H is called the Hamiltonian of the associated control problem and Equation (2.8) is the so-called HJB equation. From Equation (2.4), the regular terminal condition of the previous PDE is

$$v(T,x) = g(x), \quad \forall x \in \mathbb{R}^n.$$

Eventually, we present an important theoretical result in the theory of dynamic programming: the verification theorem. It consists in proving that, given a smooth solution to the HJB equation, this candidate coincides with the sought value function. This theorem also exhibits an optimal Markovian control as a by-product.

With the assumptions and notations we have previously defined, we formulate a general version of the verification theorem in the finite horizon case.

Theorem 2.1.2. *Let w be a function in $C^{1,2}([0,T)\times\mathbb{R}^n)\cap C^0([0,T]\times\mathbb{R}^n)$, satisfying a quadratic growth condition, i.e. there exists a constant C such that*

$$|w(t,x)| \leq C(1+|x|^2), \quad \forall (t,x) \in [0,T]\times\mathbb{R}^n.$$

1. *Suppose that*

$$-\partial_t w(t,x) - \sup_{a\in A}[\mathcal{L}^a w(t,x) + f(t,x,a)] \geq 0, \quad (t,x)\in[0,T)\times\mathbb{R}^n,$$
$$w(T,x) \geq g(x), \quad x\in\mathbb{R}^n.$$

   Then $w \geq v$ on $[0,T]\times\mathbb{R}^n$.

2. *Moreover, if $w(T,.) = g$, and there exists a measurable function $\hat{\alpha}(t,x)$, $(t,x)\in[0,T)\times\mathbb{R}^n$, valued in A such that*

$$\begin{aligned} &-\partial_t w(t,x) - \sup_{a\in A}[\mathcal{L}^a w(t,x) + f(t,x,a)] \\ &= -\partial_t w(t,x) - \mathcal{L}^{\hat{\alpha}(t,x)} w(t,x) - f(t,x,\hat{\alpha}(t,x)) = 0 \end{aligned}$$

   the stochastic differential equation

$$dX_s = b\left(X_s, \hat{\alpha}(s,X_s)\right) ds + \sigma\left(X_s, \hat{\alpha}(s,X_s)\right) dW_s$$

   admits a unique solution, denoted by $\hat{X}_s^{t,x}$, given an initial condition $X_t = x$, and the process $\left\{\hat{\alpha}\left(s, \hat{X}_s^{t,x}\right), t\leq s\leq T\right\}$ lies in $\mathcal{A}(t,x)$. Then

$$w = v \quad \text{on } [0,T]\times\mathbb{R}^n,$$

   and $\hat{\alpha}$ is an optimal Markovian control.

The above theorem suggests the following steps for solving stochastic control problems in the finite horizon case. First, we try to solve the nonlinear HJB equation:

$$-\partial_t w(t,x) - \sup_{a\in A}[\mathcal{L}^a w(t,x) + f(t,x,a)] = 0, \quad (t,x)\in[0,T)\times\mathbb{R}^n,$$

with the terminal condition $w(T,x) = g(x)$. Then we fix $(t,x)\in[0,T)\times\mathbb{R}^n$, and solve $\sup_{a\in A}|\mathcal{L}^a w(t,x) + f(t,x,a)|$ as a maximum control problem in $a\in A$ and we note $a^*(t,x)$ the argmax of this maximum. If this nonlinear PDE with a terminal condition admits a smooth solution w, then w is the value function to the stochastic control problem and a^* is an optimal Markovian control. This approach is relevant as soon as the HJB equation has a $C^{1,2}$ solution satisfying the conditions to apply the verification theorem.

The results about existence of smooth solutions to HJB parabolic PDEs are found in Fleming and Rishel (1975), Gilbarg and Trudinger (1985) and Krylov (1987). The main required assumption is a uniform ellipticity condition of the form:

- There exists a constant $c > 0$ such that $y'\sigma(x,a)\sigma'(x,a)y \geq c|y|^2$, $\forall x, y \in \mathbb{R}^n$, $\forall a \in A$.

We illustrate the previous methodology by an interesting example taken from Pham (2009) which constitutes a good entry point to the more sophisticated problems tackled in the next chapters.

Example 2.1.3 (Merton portfolio allocation problem in finite horizon)**.** *The framework considered is the Black-Scholes-Merton model over a finite investment horizon T. At any time t, an investor allocates a proportion α_t of her wealth in a risky asset of price S and $1-\alpha_t$ in a riskless asset of price S^0 with interest rate r. The proportion of wealth is bound to be valued in A, a closed convex subset of $\mathbb{R}$. The wealth process is thus assumed to evolve as*

$$\begin{aligned} dX_t &= \frac{X_t\alpha_t}{S_t}dS_t + \frac{X_t(1-\alpha_t)}{S_t^0}dS_t^0 \\ &= X_t(\alpha_t\mu + (1-\alpha_t)r)dt + X_t\alpha_t\sigma dW_t. \end{aligned}$$

The set of progressively measurable processes $\alpha \in A$ such that $\int_0^T |\alpha_s|^2 ds < \infty$ a.s. is denoted $\mathcal{A}$. This integrability condition ensures the existence and uniqueness of a strong solution to the SDE governing the wealth process controlled by $\alpha \in \mathcal{A}$. We denote by $X^{t,x}$, the wealth process starting from an initial wealth $X_t = x > 0$ at time t, corresponding to the portfolio strategy $\alpha \in \mathcal{A}$. The purpose of the investor is to maximize the expected value of the wealth of her portfolio at terminal time T. The value function of this expected utility maximization problem is then formulated as

$$v(t,x) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}\left[U\left(X_T^{t,x}\right)\right], \quad (t,x) \in [0,T]\times\mathbb{R}_+. \tag{2.9}$$

The function U is a utility function and as such, is increasing and concave in $\mathbb{R}_+$. We can easily check that for all $t \in [0,T]$, $v(t,.)$ is also increasing and concave in x. Moreover, if U is strictly concave and if an optimal control exists, it can be shown that v is also strictly convave in x. We derive the HJB equation for stochastic control Problem (2.9) as follows:

$$-\partial_t w - \sup_{a\in\mathcal{A}} [\mathcal{L}^a w(t,x)] = 0, \tag{2.10}$$

with the terminal condition

$$w(T,x) = U(x), \quad x \in \mathbb{R}_+. \tag{2.11}$$

In this example, the diffusion operator has the explicit form

$$\mathcal{L}^a w(t,x) = x(a\mu + (1-a)r)\partial_x w + \frac{1}{2}x^2 a^2 \sigma^2 \partial^2_{xx} w.$$

If we choose the specific power utility function of the constant relative risk aversion (CRRA) type, as used by Merton originally,

$$U(x) = \frac{x^p}{p}, \quad x \geq 0,\ 0 < p < 1,$$

we can find analytically a smooth solution to Problem (2.10)-(2.11). As a candidate solution to the HJB equation, let's propose

$$w(t,x) = \phi(t)U(x)$$

with ϕ a positive function. Plugging the candidate in the HJB equation, ϕ must satisfy the ordinary differential equation (ODE)

$$\phi'(t) + \tilde{\rho}\phi(t) = 0, \quad \phi(T) = 1, \tag{2.12}$$

where

$$\tilde{\rho} = p \sup_{a \in A} \left[a(\mu - r) + r - \frac{1}{2}a^2(1-p)\sigma^2 \right].$$

The solution of ODE (2.12) is $\phi(t) = \mathrm{e}^{\tilde{\rho}(T-t)}$. We can deduce that the function given by

$$w(t,x) = \mathrm{e}^{\tilde{\rho}(T-t)}U(x), \quad (t,x) \in [0,T] \times \mathbb{R}_+, \tag{2.13}$$

is strictly increasing and concave in x and is a smooth solution to Problem (2.10)-(2.11). Moreover, the function $a \in A \mapsto a(\mu - r) + r - \frac{1}{2}a^2(1-p)\sigma^2$ is strictly concave on the closed convex set A, and thus attains its maximum at some constant $\hat{a}$. By construction, $\hat{a}$ attains the supremum of $\sup_{a \in A}[\mathcal{L}^a w(t,x)]$.

In addition, the wealth process associated to the constant control $\hat{a}$,

$$dX_t = X_t(\hat{a}\mu + (1-\hat{a})r)dt + X_t \hat{a}\sigma dW_t,$$

admits a unique solution given an initial condition. From the verification Theorem 2.1.2, the value function of expected utility maximization Problem (2.9) is equal to Equation (2.13) and the optimal proportion of wealth to allocate in the risky asset is the constant $\hat{a}$.

Eventually, notice that when $A = \mathbb{R}$, the values of $\hat{a}$ and $\tilde{\rho}$ are explicit and respectively equal to

$$\hat{a} = \frac{\mu - r}{\sigma^2(1-p)},$$

and

$$\tilde{\rho} = \frac{(\mu - r)^2}{2\sigma^2}\frac{p}{1-p} + rp.$$

Originally, the principle of dynamic programming was invented by Bellman (1952). Although the concept is easy to understand, the rigorous demonstration proves to be complex and was studied by diferent authors using various methods: Bensoussan and Lions (1978), Krylov (1980), Nisio (1981), Lions (1983), Borkar (1989), Yong and Zhou (2000), or Fleming and Soner (2006). An extensive literature deals with the standard approach to optimal control for diffusion processes via a verification theorem on the HJB equation: Fleming and Rishel (1975), Krylov (1980), Yong and Zhou (2000) or Fleming and Soner (2006). The extension to controlled jump-diffusion processes is studied in Øksendal and Sulem (2004). Some other aspects are considered in the lecture notes of St-Flour El Karoui (1981).
The general existence of an optimal control problem is tackled in Kushner (1975) or El Karoui et al. (1987).

2.1.2 Markowitz problem: formulation and continuous-time solution

In Chapters 3 and 4, we deal with an innovative version of the Markowitz portfolio selection problem which implies that we review briefly the standard framework and some results.
Markowitz portfolio analysis is a mathematical procedure to determine the optimal portfolios in which to invest. The theory, initiated by Markowitz (1952), is a major step forward in finance. The objective of portfolio analysis is to find the set of efficient portfolios, i.e. the ones that:

- have the greatest expected return for a given level of risk, or conversely,
- offer the lowest level of risk for a given level of expected return.

These efficient portfolios are located on the so-called efficient frontier, represented by the dotted line in Figure 2.1, starting at the minimum variance portfolio and going upwards.

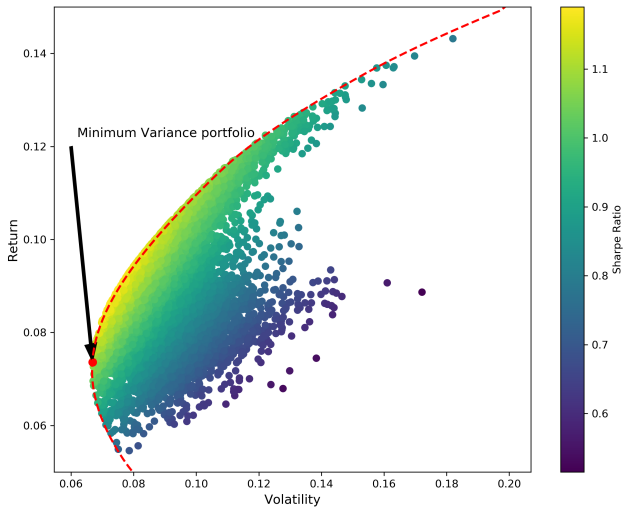


Figure 2.1: Illustration of the efficient frontier theorized by H. Markowitz.

Presently, we are interested in the formulation and resolution of the multidimensional Markowitz problem in continuous time. For the sake of simplicity and as it is not at the core of our study, we will not consider the riskless asset in the theoretical aspects of the Markowits problem in this thesis. Nonetheless, in the applications, it will be used to make the simulations more realistic. To be more precise, it will not be included in the optimisation process but, as we do not constraint the weights of the assets to add up to 1, it will be used as a riskless rate to borrow or lend the cash.

The framework consists in a financial market on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ that satisfies the usual conditions. On this market, are continuously traded n risky assets. The assets' price process S is defined as:

$$\begin{cases} dS_t &= \operatorname{diag}(S_t)\big(b_0 dt + \sigma dW_t\big), \quad t \in [0,T], \\ S_0 &= s_0 \in \mathbb{R}^n, \end{cases}$$

where W is a standard n-dimensional Brownian motion and T is a finite investment horizon.

The n-dimensional drift b_0 is a known and constant parameter and the $n \times n$ constant matrix σ is known and invertible. We denote by Σ the symmetric square matrix $\sigma\sigma'$.

We write $\mathbb{F}^{\mathbb{S}} = \{\mathcal{F}_t^S\}_{t\geq 0}$ the filtration generated by the price process S augmented by the null sets of $\mathcal{F}$. The filtration $\mathbb{F}^S$ represents the only available source of infor-

mation, those coming from the price process S.

We denote by $\alpha = (\alpha_t)_{0\leq t\leq T}$, an admissible investment strategy representing the amount invested in each risky asset. It is a $\mathbb{R}^n$-valued, $\mathbb{F}^{\mathbb{S}}$-progressively measurable process, satisfying some standard integrability condition. We define $\mathcal{A}$, the set of all admissible investment strategies.

The evolution of the self-financing wealth process X^α, given $\alpha \in \mathcal{A}$ and an initial wealth $x_0 \in \mathbb{R}$, is given by:

$$dX_t^\alpha = \alpha_t' \text{diag}(S_t)^{-1} dS_t = \alpha_t'(b_0 dt + \sigma dW_t), \quad 0 \leq t \leq T, \quad X_0^\alpha = x_0.$$

The Markowitz problem is then formulated as :

$$\hat{U}_0(\vartheta) \;:=\; \sup_{\alpha\in\mathcal{A}} \Big\{ \mathbb{E}[X_T^\alpha] : \text{Var}(X_T^\alpha) \leq \vartheta \Big\}, \tag{2.14}$$

where $\vartheta > 0$ is the variance budget the investor can allow on her portfolio. The value function $\hat{U}_0$ is then the expected optimal terminal wealth value the investor can achieve given her variance constraint. We can show that Markowitz Problem (2.14) can be turned into a so called mean-variance problem, where the constraint becomes part of the objective function:

$$\hat{V}_0(\lambda) \;:=\; \inf_{\alpha\in\mathcal{A}} \big[\lambda \text{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha]\big], \quad \lambda > 0.$$

From Zhou and Li (2000), we know this time-inconsistent problem can be turned into an auxiliary problem on which we can apply the DPP. The HJB equation of this auxiliary problem comes down to solving the simple following ODE:

$$\begin{cases} 0 = R'(t) + |\sigma^{-1} b_0|^2, \\ 0 = R(T). \end{cases}$$

for which the solution is $R(t) = |\sigma^{-1} b_0|^2 (T - t)$. This allows us to solve the initial mean-variance problem

$$\hat{V}_0(\lambda) = -\frac{1}{4\lambda}\left(e^{|\sigma^{-1}b_0|^2 T} - 1\right) - x_0,$$

with the optimal control in feedback form

$$\alpha^{*,\lambda} = a_0^{*,\lambda}(X_t^{\alpha^{*,\lambda}}),$$

where

$$a_0^{*,\lambda}(x) \;:=\; \left(x_0 - x + \frac{e^{|\sigma^{-1}b_0|^2 T}}{2\lambda}\right) \Sigma^{-1} b_0, \quad x \in \mathbb{R}.$$

Finally, we deduce the optimal performance of the Markowitz problem:

$$\hat{U}_0(\vartheta) = x_0 + \sqrt{\vartheta(e^{|\sigma^{-1}b_0|^2T} - 1)}.$$

with the optimal feedback control of the Markowitz problem being

$$\alpha^{*,\vartheta} = a_0^{*,\vartheta}(X_t^{\alpha^{*,\vartheta}}),$$

where

$$a_0^{*,\vartheta}(x) \ := \ \left(x_0 - x + e^{|\sigma^{-1}b_0|^2T} \sqrt{\frac{\vartheta}{e^{|\sigma^{-1}b_0|^2T} - 1}} \right) \Sigma^{-1} b_0, \quad x \in \mathbb{R}.$$

In the present thesis, we are interested in solving this problem when the drift is uncertain and modeled via a prior probability distribution.
Consequently, we need to review briefly the literature on filtering theory.

2.1.3 Filtering theory and the Bayesian learning approach

The present thesis focuses on portfolio optimization with parameters uncertainty and thus, concentrates on a framework in which the investor cannot observe directly the drift of the Brownian motion driving the dynamics of the assets: she can only observe past and present assets' prices. We assume the unknown drift is an unobservable random variable, independent of the driving Brownian motion and with known probability distribution. This case is refered to as the partial observations case.

Since we use filtering theory to manage drift uncertainty, we exhibit in this section important standard results of this theory proved in Bain and Crisan (2009).

The filtering approach. We define a probability space $(\Omega,\ \mathcal{F},\ \mathbb{P})$ together with a filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ which satisfies the usual conditions. Let's consider an $\mathbb{F}$-adapted unobservable process $X = \{X_t,\ t \geq 0\}$ valued in $\mathbb{R}^n$. The distribution of X_0 is assumed to be μ_0.

Let $h = (h_i)_{i=1}^m :\ \mathbb{R}^n \mapsto \mathbb{R}^m$ be a measurable function such that

$$\mathbb{P}\left(\int_0^t |h(X_s)| ds < \infty \right) = 1$$

for all $t \geq 0$. We denote by W, a standard $\mathbb{F}$-adapted m-dimensional Brownian motion on $(\Omega,\ \mathcal{F},\ \mathbb{P})$ independent of X.

Let $Y = \{Y_t,\ t \geq 0\}$, the observation process satisfying the following evolution equation:

$$Y_t = Y_0 + \int_0^t h(X_s)ds + W_t.$$

We denote

$$\mathcal{Y} = \bigvee_{t\in\mathbb{R}_+} \mathcal{Y}_t,$$

where $\mathcal{Y}_t$ is the usual augmentation with null sets of the filtration associated to the process Y.

We note that, since h is measurable, Y is $\mathbb{F}$-adapted and thus $\mathcal{Y}_t \subset \mathcal{F}_t$.

Remark 3. *Most results can be extended in the case where h is time-dependent but to simplify notation, we have removed the time dependence.* ◊

The filtering problem consists in determining the conditional distribution μ_t of the process X at time t, given the information accumulated from observing Y in the interval $[0, t]$, i.e. for a bounded measurable function ϕ, computing

$$\mu_t(\phi) = \mathbb{E}[\phi(X_t)|\mathcal{Y}_t].$$

Y_0 is considered to be zero since there is no information available initially. Hence, the initial distribution μ_0 of X_0 is identical to the conditional distribution of X_0 given $\mathcal{Y}_0$ and we have

$$\mu_0(\phi) = \int_{\mathbb{R}^n} \phi(x)\mathbb{P}X_0^{-1}(dx),$$

Two approaches are available in the literature to deduce the evolution of μ.

The change of measure method. This technique, used in Chapter 5, consists in constructing a new method in which Y becomes a Brownian motion and μ has a representation in terms of an associated unnormalised version, π. This measure π, is then shown to satisfy a linear evolution equation which leads to the evolution equation for μ by an application of Itô's formula.

This method consists in modifying the probability measure on Ω, in order to transform the process Y, into a Brownian motion by means of Girsanov's theorem. Let $Z = \{Z_t,\ t \geq 0\}$, be the process defined by

$$Z_t = e^{-\sum_{i=1}^m \int_0^t h^i(X_s)dW_s^i - \frac{1}{2}\sum_{i=1}^m \int_0^t h^i(X_s)^2 ds}, \quad t \geq 0.$$

The process Z is a martingale if the following Novikov condition holds

$$\mathbb{E}\left[e^{\frac{1}{2}\sum_{i=1}^{m}\int_0^t h^i(X_s)^2 ds}\right] < \infty,$$

for all $t > 0$.

Remark 4. *Note that, under suitable conditions, it suffices to show that*

$$\mathbb{E}\left[\sum_{i=1}^{m}\int_0^t h^i(X_s)^2 ds\right] < \infty,$$

for Z to be a martingale. ◊

To construct the change of probability, Z needs to be a martingale. This is true if the following condition on h is verified

$$\mathbb{E}\left[\int_0^t |h(X_s)|^2 ds\right] < \infty, \quad \mathbb{E}\left[\int_0^t Z_s|h(X_s)|^2 ds\right] < \infty, \quad t \geq 0. \tag{2.15}$$

We can show that if the previous condition (2.15) is satisfied, Z is an $\mathbb{F}$-adapted martingale.

For fixed $t \geq 0$, $Z_t > 0$ introduces a probability measure $\overline{\mathbb{P}}^t$ on $\mathcal{F}_t$, by specifying its Radon-Nikodym derivative with respect to $\mathbb{P}$ to be given by Z_t,

$$\left.\frac{d\overline{\mathbb{P}}^t}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t.$$

It is immediate from the martingale property of Z that the measures $\overline{\mathbb{P}}^t$ form a consistent family. Therefore, we can define a probability measure $\overline{\mathbb{P}}$ equivalent to $\mathbb{P}$ on $\cup_{0\leq t\leq\infty}\mathcal{F}_t$.

We can also prove that, if condition (2.15) is satisfied, then, under $\overline{\mathbb{P}}$, the observation process Y is a Brownian motion independent of X. Additionally the law of the process X under $\overline{\mathbb{P}}$ is the same as its law under $\mathbb{P}$.

We denote $\overline{Z} = \{\overline{Z}_t,\ t \geq 0\}$ the process defined as $\overline{Z}_t = Z_t^{-1}$ for $t \geq 0$. Under $\overline{\mathbb{P}}$, $\overline{Z}_t$ satisfies the following SDE,

$$d\overline{Z}_t = \sum_{i=1}^{m}\overline{Z}_t h^i(X_t)dY_t^i,$$

and since $\overline{Z}_0 = 1$,

$$\overline{Z}_t = e^{\sum_{i=1}^m \int_0^t h^i(X_s) dY_s^i - \frac{1}{2}\sum_{i=1}^m \int_0^t h^i(X_s)^2 ds}.$$

Since $\overline{\mathbb{E}}[\overline{Z}_t] = \mathbb{E}[\overline{Z}_t Z_t] = 1$, $\overline{Z}_t$ is an $\mathbb{F}$-adapted martingale under $\overline{P}$, and we have

$$\left.\frac{d\mathbb{P}}{d\overline{\mathbb{P}}^t}\right|_{\mathcal{F}_t} = \overline{Z}_t \quad \text{for } t \geq 0.$$

Under $\overline{\mathbb{P}}$, the observation process Y is a $\mathcal{Y}_t$-adapted Brownian motion. Since the Brownian motion is a Markov process, for an integrable $\mathcal{F}_t$-measurable random variable U, we have

$$\overline{\mathbb{E}}[U|\mathcal{Y}_t] = \overline{\mathbb{E}}[U|\mathcal{Y}].$$

This allows us to replace the time-dependent family of σ-algebras $\mathcal{Y}_t$, in the conditional expectations, with the fixed σ-algebra $\mathcal{Y}$. It means that the solution of the filtering problem for an $\mathcal{F}_t$-adapted random variable U given all observations, future, present and past, is equal to $\overline{\mathbb{E}}[U|\mathcal{Y}_t]$, meaning that the future observations will not influence the estimator.

The innovation process method. The second approach, used in Chapter 3, isolates the Brownian motion driving the evolution equation for μ, called the innovation process, then identifies the corresponding term in the Doob-Meyer decomposition of μ.

So as to define the unnormalised conditional distribution process, we first state the Kallianpur-Striebel formula. Assume that condition (2.15) holds for every $\phi \in \mathcal{B}(\mathbb{R}^n)$, for fixed $t \in [0, \infty)$, we can prove that

$$\mu_t(\phi) = \frac{\overline{\mathbb{E}}[\overline{Z}_t \phi(X_t)|\mathcal{Y}]}{\overline{\mathbb{E}}[\overline{Z}_t|\mathcal{Y}]}, \quad \overline{\mathbb{P}} \text{ and } \mathbb{P}\text{-a.s.}$$

Let $\xi = \{\xi_t,\ t \geq 0\}$ be the process defined by

$$\xi_t = \overline{\mathbb{E}}[\overline{Z}_t|\mathcal{Y}_t],$$

then, as $\overline{Z}_t$ is an $\mathcal{F}_t$-martingale under $\overline{P}$ and $\mathcal{Y}_s \subseteq \mathcal{F}_s$, it follows that for $0 \leq s \leq t$,

$$\overline{\mathbb{E}}[\xi_t|\mathcal{Y}_s] = \overline{\mathbb{E}}[\overline{Z}_t|\mathcal{Y}_s] = \overline{\mathbb{E}}\left[\overline{\mathbb{E}}[\overline{Z}_t|\mathcal{F}_s]|\mathcal{Y}_s\right] = \overline{\mathbb{E}}[\overline{Z}_s|\mathcal{Y}_s] = \xi_s.$$

We can prove that it is possible to choose a càdlàg version of ξ_t which is a $\mathcal{Y}_t$-martingale.

Given such a process ξ_t, we define the unnormalised conditional distribution of X to be the measure-valued process $\pi = \{\pi_t,\ t \geq 0\}$ which is determined by the values of $\pi_t(\phi)$ for $\phi \in \mathcal{B}(\mathbb{R}^n)$ which are given for $t \geq 0$ by

$$\pi_t(\phi) := \mu_t(\phi)\xi_t.$$

It has been proved that the process $\{\pi_t,\ t \geq 0\}$ is càdlàg and $\mathcal{Y}_t$-adapted and for any $t \geq 0$,

$$\pi_t(\phi) = \overline{\mathbb{E}}[\overline{Z}_t\phi(X)_t|\mathcal{Y}_t] \quad \overline{\mathbb{P}} \text{ and } \mathbb{P}\text{-a.s.}$$

Moreover, if we assume again condition (2.15) satisfied, for every $\phi \in \mathcal{B}(\mathbb{R}^n)$,

$$\mu_t(\phi) = \frac{\pi_t(\phi)}{\pi_t(\mathbf{1})} \quad \forall\, t \in [0,\infty) \quad \overline{\mathbb{P}} \text{ and } \mathbb{P}\text{-a.s.}$$

The Kallianpur-Striebel formula justifies the usage of the term unnormalised in the definition of π_t, as the denominator $\pi_t(\mathbf{1})$ can be viewed as the normalising factor. This result can also be viewed as the abstract version of Bayes' identity in this filtering framework.

More generally, models with partial observations were studied by Detemple (1986), Dothan and Feldman (1986) and Gennotte (1986) in a linear Gaussian filtering setting. Karatzas and Xue (1991) introduced a Bayesian approach for the expected utility maximization problems, using filtering and martingale representation theory. Within this framework Lakner (1995), Lakner (1998), and Zohar (2001) solved the optimization problems via the martingale approach, Kuwana (1995) studied necessary and sufficient conditions for the certainty-equivalence principle to hold, and Karatzas (1997) tackled the problem of maximizing the probability of reaching a given objective during some finite time-horizon. For an unobservable drift process driven by an independent Brownian motion, the optimization problem was studied by Rishel (1999) for utility functions of power type. The special case of logarithmic utility function and normal prior distribution was studied by Browne and Whitt (1996) on an infinite horizon.
We can also refer to the papers by Rogers (2001), Karatzas and Zhao (2001), and Cvitanić et al. (2006) which used filtering and learning techniques to solve stochastic optimization problems with various constraints in a partial observation framework. Risk minimizing hedging strategies for contingent claims in a partial observation framework was tackled in Ceci (2006), while Ceci and Gerardi (2006) considered a model for intraday stock price movements where the jump intensity depends upon an unobservable hidden state variable. The filtering problem for a general jump diffusion process was dealt with in Ceci and Colaneri (2012) and partial information in the context of backward stochastic differential equations was examined in Ceci

et al. (2014). Recently, Bismuth et al. (2019) studied optimal portfolio choice taking into account in the decision process both the liquidity of assets and the uncertainty regarding their expected returns, and De Franco, Nicolle, and Pham (2019a) dealt with the Markowitz portfolio selection problem with partial information.

2.1.4 Maximum drawdown

Chapter 5 focuses on a portfolio selection problem, more precisely an expected utility maximization problem satisfying a maximum drawdown constraint. Important expected utility maximization problems in finance were introduced by Merton (1971) in the context of constant coefficients, and were managed by the Markovian methods of continuous-time stochastic control, see Karatzas et al. (1998). For general parameter processes in complete markets, methodologies based on martingale theory and convex duality were developped by Pliska (1986), by Karatzas et al. (1987), and by Cox and Huang (1989). They were extended to the setting of incomplete and/or constrained markets by Karatzas et al. (1991), He and Pearson (1991) and Cvitanić and Karatzas (1992).

Regarding the literature on risk constraints in a portfolio allocation framework, we refer to Redeker and Wunderlich (2018) dealing with dynamic risk constraints and comparing the continuous and discrete time trading, and the papers by Grossman and Zhou (1993) and Cvitanić and Karatzas (1994) especially focusing on drawdown constraints. More recently, Elie and Touzi (2008) study infinite-horizon optimal consumption-investment problem in continuous-time, and Boyd et al. (2019) use forecasts of the mean and covariance of financial returns, from a multivariate hidden Markov model with time-varying parameters, to build the optimal controls.

The risk constraint that appeals to us in the present thesis is the maximum drawdown (MD), so we elaborate briefly on this indicator.

MD, illustrated in Figure 2.2, is a widely regarded risk measure in the industry, and is defined as the maximum observed loss of a portfolio, from a peak to a trough, before a new peak is attained. It is an indicator of downside risk over a specified time period. Expressed in percentage terms and taking values in the interval [0%,-100%], it can serve as a stand-alone measure or as an input for other metrics, such as the Calmar Ratio, calculated as the absolute value of the expected return over the MD.

For a continuous trajectory X_t, $t \in [0, T]$, the formula for MD at time t is:

$$MD_t = \min_{0 \le s \le t} \left[\frac{\min_{s \le u \le t} X_u}{\max_{0 \le u \le s} X_u} - 1 \right].$$

It's worth notifying that MD only measures the size of the largest loss, without taking into consideration the frequency of large losses. Because it measures only the largest drawdown, MD does not indicate how long it took an investor to recover from the loss, or if the investment even recovered at all.

A MD close to zero is preferred, as this indicates that losses from investment are small. If a portfolio constantly sees its value increasing, the maximum drawdown would be zero. The worst possible MD would be -100%, meaning the portfolio is completely worthless.

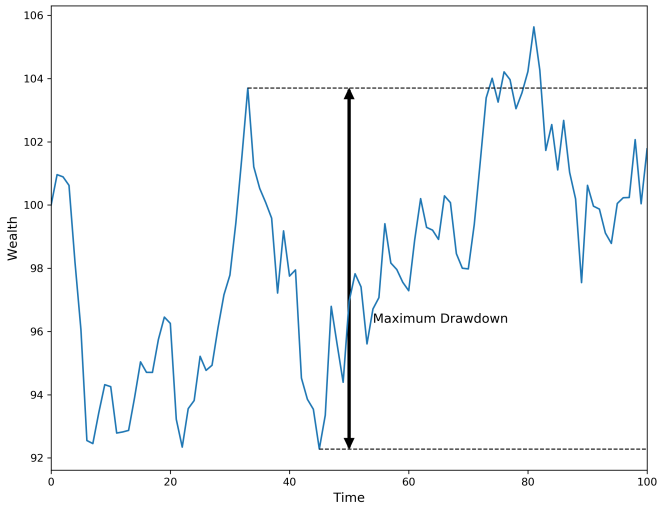


Figure 2.2: Illustration of the maximum drawdown on a simulated wealth trajectory.

2.2 Portfolio optimization under parameters uncertainty - contributions

2.2.1 Chapter 3

Motivation. In its standard form, the parameters of the Markowitz model, drifts and volatilities, are known and constant. This assumption raises the question of

giving a value to these parameters. Typically, they are estimated from past data or calculated based on predictive analyses at inception, and are considered constant afterwards. This convenient solution does not look appropriate in practice since it does not adapt to evolving market regimes, materialized through the information contained in the changing prices. Not considering this information is detrimental in the optimization process, because it does not allow for an update of the model, given the most recent available information.

In our approach, we incorporate uncertainty in the drift parameter while keeping the covariance parameter known and constant. This choice is supported, for instance by Merton (1980), who showed that estimates of variances and covariances of assets' returns are more accurate than estimates of their expected returns. Furthermore, being wrong in the estimation of expected returns can result in a very suboptimal portfolio *a posteriori*. Indeed, Best and Grauer (1991) argued that mean-variance optimal portfolios are very sensitive to the level of expected returns.
Precisely, our approach considers that the drift has a known probability distribution which will be updated using the Bayesian learning approach.

Problem description. We model a financial market on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ equipped with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ that satisfies the usual conditions. We define a standard n-dimensional Brownian motion W which drives the dynamics of the n risky assets that are continuously traded on this market.

The assets prices process S is defined as:

$$\begin{cases} dS_t &= \operatorname{diag}(S_t)\big(Bdt + \sigma dW_t\big), \quad t \in [0,T], \\ S_0 &= s_0 \in \mathbb{R}^n, \end{cases} \tag{2.16}$$

where T is a finite investment horizon, $B \in \mathbb{R}^n$ is a random vector independent of W, with distribution μ and such that $\mathbb{E}[|B|^2] < \infty$. B has a prior probability distribution that represents the subjective beliefs of the investor about the likelihood of the different values that it might take.

In our framework, we consider the following assumptions satisfied:

- $\operatorname{Cov}(B) = \mathbb{E}\left[(B - \mathbb{E}[B])(B - \mathbb{E}[B])'\right]$ is positive definite,
- σ is known and invertible.

As the investor observes the price process S but not the drift B, the filtration $\mathbb{F}^{\mathbb{S}} := \{\mathcal{F}_t^S\}_{t\geq 0}$ generated by the price process S, augmented by the null sets of $\mathcal{F}$, represents the only available source of information.

The $\mathbb{R}^n$-valued and $\mathbb{F}^{\mathbb{S}}$-progressively measurable process $\alpha = (\alpha_t)_{0\leq t\leq T}$, is an admissible investment strategy representing the amount invested in each of the n risky assets and satisfies an appropriate integrability condition.[1]

We denote by $\mathcal{A}$ the set of all admissible investment strategies.

The dynamics of the self-financing wealth process X^α, controlled by $\alpha \in \mathcal{A}$ and starting at an initial wealth $x_0 \in \mathbb{R}$, is written as

$$dX_t^\alpha = \alpha_t' \text{diag}(S_t)^{-1} dS_t = \alpha_t'(Bdt + \sigma dW_t), \quad 0 \leq t \leq T, \quad X_0^\alpha = x_0. \tag{2.17}$$

We define our problem as

$$U_0(\vartheta) \;:=\; \sup_{\alpha\in\mathcal{A}} \Big\{ \mathbb{E}[X_T^\alpha] : \text{Var}(X_T^\alpha) \leq \vartheta \Big\}, \tag{2.18}$$

where $\vartheta > 0$ is the maximal level of terminal wealth's variance the investor is able to bear. Although identical in its formulation to the standard Markowitz problem, our problem differentiates itself by considering assets' drift uncertainty updated thanks to Bayesian filtering. We call it the Bayesian-Markowitz problem in the sequel.

New results. We formulate in a multidimensional framework, the Markowitz problem in the case of an uncertain drift modeled by a prior distribution.
We first turn Markowitz Problem (2.18) into its associated mean-variance problem, see Lemma 3.2.3. This a priori time inconsistent problem cannot be tackled by convex duality and martingale methods as in Karatzas and Zhao (2001). Instead, using the methodology of Zhou and Li (2000) we show how the associated mean-variance problem can be embedded into an auxiliary standard control problem that can be tackled by the dynamic programming approach, see Lemma 3.2.4.

Although prior conjugates are widely used in the literature, in our framework we are able to consider any prior distribution to model the uncertain drift. This comes from the extension to the multidimensional case of a recent result by Ekstrom and Vaicenavicius (2016), we make in Lemma 3.3.1.

Let's state the main theoretical results of Chapter 3. First, we turn the unobservable model (2.16)-(2.17) into an observable model. To do so, we define $\hat{B}_t = \mathbb{E}[B|\mathcal{F}_t^S]$ and introduce the innovation process

$$\hat{W}_t \;:=\; \sigma^{-1} \int_0^t (B - \hat{B}_s)ds + W_t, \quad 0 \leq t \leq T.$$

[1] Detailed in Chapter 3.

It leads us to the following observable model

$$\begin{cases} dS_t &= \mathrm{diag}(S_t)\left(\hat{B}_t dt + \sigma d\hat{W}_t\right), \quad t \in [0,T], \\ S_0 &= s_0 \in \mathbb{R}^n, \end{cases}$$

with the corresponding dynamics for the wealth process

$$dX_t^\alpha \;=\; \alpha_t' \mathrm{diag}(S_t)^{-1} dS_t = \alpha_t'(\hat{B}_t dt + \sigma d\hat{W}_t), \quad 0 \le t \le T, \quad X_0^\alpha = x_0.$$

In order to derive the dynamics of $\hat{B}_t$, we notice that the observation process

$$Y_t \;:=\; \sigma^{-1} \int_0^t \hat{B}_s ds + \hat{W}_t, \tag{2.19}$$

can be written as $Y_t = \sigma^{-1} Bt + W_t$ which turns out to be a measurable function of S_t. This implies that Y_t generates the same filtration as S_t, and by Lemma 3.3.1, we can write

$$\hat{B}_t = \mathbb{E}[B|Y_t] := f(t, Y_t).$$

Applying Itô's formula to $\hat{B}_t = f(t, Y_t)$ with (2.19), and recalling that $\hat{B}$ is a $(\mathbb{P}, \mathcal{F}^S)$-martingale, we see that $d\hat{B}_t = \nabla_y f(t, Y_t) d\hat{W}_t$. Noting $f_t(y) := f(t,y)$ and defining the matrix-valued function ψ by

$$\psi(t,b) \;:=\; \nabla_y f(t, f_t^{-1}(b)), \quad t \in [0,T], \;\; b \in f_t(\mathbb{R}^n) \;:=\; \mathcal{B}_t,$$

where $y = f_t^{-1}(b)$ is the unique value of the observation process Y_t that yields $\hat{B}_t = b$, by Lemma 3.3.2 we finally have the dynamics of $\hat{B}_t$,

$$d\hat{B}_t \;=\; \psi(t, \hat{B}_t) d\hat{W}_t, \; t \in [0,T], \quad \hat{B}_0 \;=\; b_0 \;= \mathbb{E}[B].$$

Next, we prove that our problem comes down to solving a fully nonlinear HJB equation which can be reduced to the following semi-linear PDE on $\mathcal{R} = \{(t,b) : t \in [0,T], \; b \in \mathcal{B}_t\}$:

$$\begin{cases} -\partial_t R - \dfrac{1}{2}\mathrm{tr}\big(\psi\psi' \mathcal{D}_b^2 R\big) + \tilde{F}(t, b, \nabla_b R) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad R(T,b) = 0, \quad b \in \mathcal{B}_t, \end{cases} \tag{2.20}$$

with

$$\tilde{F}(t,b,p) := 2\left(\psi\sigma^{-1}b\right)' p - \frac{1}{2}|\psi' p|^2 - |\sigma^{-1}b|^2. \tag{2.21}$$

Furthermore, we prove in Theorem 3.4.3 that if the following conditions hold true,

- $\forall (t,b) \in \mathcal{R}$, $\psi(t,b)$ and $\psi^{-1}(t,b)$ are bounded, and
- the matrix $\nabla_b \psi$ exists and is bounded,

there exists a classical solution $R \in \mathcal{C}^{1,2}(\mathcal{R})$ to PDE (2.20). In addition, R is at most quadrically growing in b and $\nabla_b R$ is at most linearly growing in b.

If the previous conditions are fulfilled, we prove in Theorem 3.4.1 that we can solve the Bayesian-Markowitz problem with its associated optimal control.

For any $\vartheta > 0$, the solution to the Bayesian-Markowitz problem is

$$U_0(\vartheta) \;=\; x_0 + \sqrt{\vartheta\,(e^{R(0,b_0)} - 1)}, \quad \vartheta > 0,$$

with the associated optimal Bayesian-Markowitz strategy given by

$$\hat{\alpha}_t^{\vartheta} \;=\; a_0^{\vartheta}(t, X_t^{\hat{\alpha}^{\vartheta}}, \hat{B}_t),$$

where

$$\begin{aligned} a_0^{\vartheta}(t,x,b) :=& \left(x_0 - x + e^{R(0,b_0)} \sqrt{\frac{\vartheta}{e^{R(0,b_0)} - 1}}\right) \\ & \times \left(\Sigma^{-1} b - \left(\psi \sigma^{-1}\right)' \nabla_b R(t,b)\right), \quad (t,b) \in \mathcal{R},\ x \in \mathbb{R}, \end{aligned}$$

and remembering that $\hat{B}_t = \mathbb{E}[B|\mathcal{F}_t^S]$.

In the particular case of a multidimensional Gaussian prior conjugate, we solve explicitly the problem and we exhibit a closed-form analytical solution to Bayesian-Markowitz Problem (2.18) in Lemma 3.4.5.

In addition to the previous theoretical results, we provide a comprehensive sensitivity analysis in Section 3.5, which comes to the conclusion that learning brings value to the optimal strategy that solve the Markowitz problem with uncertain drift.

Difficulty. Here are the technical difficulties we have faced.
(1) The initial problem is turned into a mean-variance problem which is not time-consistent due to the squared expectation term in the variance. Hence, we make a second transformation to obtain a standard problem that can be solved using the dynamic programming approach via the HJB equation.
(2) The proof of existence of a smooth solution to PDE (2.20) is not trivial. Indeed, the standard assumptions of existence and uniqueness that we can find, for instance, in Ladyzhenskaia et al. (1968) do not apply here since the function $\tilde{F}$ defined in Equation (2.21) is not globally Lipschitz in p. This difficulty comes from

the unboundedness of the domain and the quadratic growth in p of the function $\tilde{F}$ which entail that the solution may grow quadratically.

Further research. We can extend the previous results studying the applications to other than Gaussian prior probability distributions and perform an analysis of the value of information.

2.2.2 Chapter 4

Motivation. Introducing uncertainty in the problem makes it more realistic and brings it closer to a real-life situation, where not only the investor ignores the parameters of the model, but also is forced to admit that her estimates are uncertain. Thus, our main objective is to confront to historical market data, the theoretical prevalence of the Bayesian learning (BL) strategy above the non-learning (NL) one which considers the drift constant. To do so, we apply both strategies to different investment universes that we detail in Appendix 4.A.

Problem description. In the financial market considered, are traded one risk-free asset, whose return is denoted by r^f, and n risky assets whose returns r_t are modeled by

$$\begin{cases} r_t &= B + \sigma\xi_t \\ \xi_t &\sim N(0, I_n). \end{cases}$$

We assume that the random vectors ξ_t are independent for all t. $B \in \mathbb{R}^n$ is the vector of the expected returns of the risky assets, while $\Sigma := \sigma\sigma' \in \mathbb{R}^{n\times n}$ is their covariance matrix. We assume that σ^{-1} exists. For $A \subseteq \mathbb{R}^n$, the set of admissible investment strategies is defined as

$$\mathcal{A} := \{\mathbf{w} = (\mathbf{w}_t)_{0\leq t\leq T} | \mathbf{w}_t \in \mathrm{A}\},$$

where an admissible strategy $\mathbf{w} = (\mathbf{w}_t)_t$ represents the fraction of wealth invested in the assets at any time t. We write the wealth at maturity $T > 0$ as

$$X_T^{\mathbf{w}} = X_0 \prod_{s=0}^{T-1} \left(1 + \mathbf{w}_s' r_{s+1} + (1 - \mathbf{w}_\mathbf{s}' \mathbb{1}_n)\, r^f\right).$$

We consider an investor who is aiming to solve the Markowitz problem:

$$\begin{cases} \max\limits_{\mathbf{w}\in\mathcal{A}} \mathbb{E}\left[X_T^{\mathbf{w}}\right] \\ \\ \mathrm{Var}\left(X_T^{\mathbf{w}}\right) \leq \vartheta, \end{cases}$$

where $\vartheta > 0$ is the risk tolerance for the investor.

In this setup, to account for drift uncertainty, we assume reasonably that the investor has an *a priori* view on the risky assets and a belief on their expected returns, but she is uncertain about the accuracy of her forecast. This is reflected in the fact that the investor does not observe B, and only assumes that $B \sim \mu$, where μ is a probability distribution in $\mathbb{R}^n$ such as $\mathbb{E}[B] = b_0$. The parameter b_0 is the vector of returns the investor is expecting, while μ translates her uncertainty about it.

New results. Implementing the Bayesian learning and non-learning strategies with historical data, implies making some adjustments to the theoretical results of Chapter 3. For instance, in practice, there are liquidity issues and the trading is in discrete time. Hence, we discretize both the optimal solution of the continuous Bayesian-Markowitz problem and the non-learning strategy. Then, we turn amounts into proportions and cap the optimal proportions to be invested, in order to fit portfolio management constraints.
We aknowledge that these adjustments make the continuous-time optimal solutions, suboptimal in discrete-time. Nonetheless, we treat both BL and NL strategies equally and we aim at emphasizing the gain from learning rather than the pure performance of the strategies independently. Doing so, we accept to loose a form of optimality.

In our study we show numerically the outperformance, measured as annualized returns and Sharpe ratios, of the learning strategy versus the non-learning one whether we consider classic investment universes: equities, currencies, or more sophisticated ones comprising smart beta strategies.
Not only, this confirms the theoretical results of Chapter 3, but also illustrates the robustness of the prevalence of the learning strategy since it is verified for different investment universes.

We also perform a thorough sensitivity analysis of the performance of learning versus non-learning according to different parameters. We detail the main consequences of the variations of each parameter on the relative performance.

- **Impact of uncertainty.** The higher the level of drift uncertainty, the better the learning strategy behaves compared to the non-learning one. Nonetheless, a (very) high level of uncertainty allows the drift to span a very vast region which slow down the learning process so that it does not add any informative value to the learning strategy, implying that the non-learning strategy can outperform.

– **Impact of leverage.** In our framework, the leverage parameter is the cap level of the weights of each strategies. Consequently, as this parameter increases, the strategies become unconstrained and are more sensitive to any error in the estimated parameter. As expected, the higher the leverage, the higher the performance of both strategies at the cost of higher risk. The learning strategies outperforms the non-learning one at any level of leverage considered in our study.

– **Impact of the review frequency.** The review frequency is the frequency at which we recalibrate the parameters of the model. Our main conclusion is that, as we lower the frequency, the relative improvement in Sharpe ratios increases dramatically. Indeed, a low review frequency allows the learning strategy to take full advantage of its learning feature and outperform the non-learning strategy, stuck with the initial estimate.

– **Impact of the rebalancing frequency.** In our setup, we show that the Bayesian learning strategy has no real incentive to increase its rebalancing frequency from monthly to bi-weekly. For reasons detailed in Chapter 4, a higher frequency of rebalancing implies an increase in relative outperformance of the learning strategy. This comes from the poor performance of the non-learning strategy at this frequency rather than a higher performance of the learning one.

Difficulty. The difficulties come from the need to make adjustments before implementing the strategies with historical data. In particular, we had to define a portfolio management policy in the form of the workflow described in section 4.3, and tune the associated parameters. The interpretation of the sensitivity results was also a difficult part of the study since we need to extract meaningful analysis often mixed with noise.

Further research. Since the discretization of continuous-time optimal solutions make them suboptimal in discrete-time, a natural extension of this work would be to derive the discrete optimal solutions for both strategies and compare the results. Another interesting extension could be to use probability distributions *a priori* different from the Gaussian, possibly empirical on a given calculation window to adapt to discrete time.

2.3 Portfolio optimization under risk measure constraints - contributions

2.3.1 Chapter 5

Motivation. In Chapter 5, we study a portfolio optimization problem whose goal is to maximize the expected utility of the terminal wealth of a portfolio given a maximum drawdown constraint and whose assets have an unobservable dynamics' drift. To our knowledge, this optimization problem which arises in practice, has never been approached this way, and we propose an original way to solve it numerically. We use an advanced deep learning methodology based on neural networks and a specifically designed algorithm[2] to propose a numerical answer to this problem for which no analytical solution is possible.

Problem description. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a discrete filtration $\{\mathcal{F}_k\}_{k=0,\,\ldots,\,N}$ satisfying the usual conditions, where lives a financial market model with one riskless asset assumed normalized to one, and d risky assets. The price process $(S^i_k)_{k=0,\ldots,N}$ of asset $i \in [\![1, d]\!]$ evolves as

$$S^i_{k+1} \;=\; S^i_k e^{R^i_{k+1}}, \quad k = 0, \ldots, N-1, \tag{2.22}$$

where $R_{k+1} = (R^1_{k+1}, \ldots, R^N_{k+1})$ is the assets' log-return between time k and $k+1$, and evolves as:

$$R_{k+1} \;=\; B + \epsilon_{k+1}.$$

The drift B is a d-dimensional random variable with probability distribution (prior) μ_0 of known mean $b_0 = \mathbb{E}[B]$ and finite second order moment. The noise $\epsilon = (\epsilon_k)_k$ is a sequence of centered i.i.d. random vector variables with covariance matrice $\Gamma = \mathbb{E}[\epsilon_k \epsilon_k']$, and assumed to be independent of B. We also assume that the probability distribution ν of ϵ_k admits a strictly positive density function g on $\mathbb{R}^d$ with respect to the Lebesgue measure.

The price process S is observable, and notice by Equation (2.22) that R can be deduced from S, and vice-versa. Thus, we denote by $\mathbb{F}^o = \{\mathcal{F}^o_k\}_{k=0,\,\ldots,\,N}$ the observation filtration generated by S (hence equivalently by R) augmented by the null sets of $\mathcal{F}$, with the convention that for $k = 0$, $\mathcal{F}^o_0$ is the trivial algebra.

In this framework, an investment strategy is an $\mathbb{F}^o$-progressively measurable process $\alpha = (\alpha_k)_{k=0,\,\ldots,\,N-1}$, valued in $\mathbb{R}^d$, which represents the proportion of the current

[2] The algorithm is detailed in Algorithm 1 p.155

wealth invested in each of the d risky assets at each time $k = 0, \ldots, N-1$. Given an investment strategy α and an initial wealth $x_0 > 0$, the (self-financed) wealth process X^α evolves according to

$$\begin{cases} X^\alpha_{k+1} = X^\alpha_k \left(1 + \alpha'_k \left(e^{R_{k+1}} - \mathbb{1}_d\right)\right), & k = 0, \ldots, N-1, \\ X^\alpha_0 = x_0, \end{cases}$$

where $e^{R_{k+1}}$ is the d-dimensional random variable with components $\left[e^{R_{k+1}}\right]_i = e^{R^i_{k+1}}$ for $i \in [\![1, d]\!]$, and $\mathbb{1}_d$ is the vector in $\mathbb{R}^d$ with all components equal to 1.
We introduce the process Z^α_k, as the maximum up to time k of the wealth process X^α, i.e.,

$$Z^\alpha_k := \max_{0 \le \ell \le k} X^\alpha_\ell, \qquad k = 0, \ldots, N.$$

The maximum drawdown constraints the wealth X^α_k to remain above a fraction $q \in (0, 1)$ of the current historical maximum Z^α_k. We then define the set of admissible investment strategies $\mathcal{A}^q_0$, as the set of investment strategies α such that

$$X^\alpha_k \geq q Z^\alpha_k \quad \text{a.s.}, \quad k = 0, \ldots, N.$$

The portfolio selection problem is formulated as

$$V_0 := \sup_{\alpha \in \mathcal{A}^q_0} \mathbb{E}\left[U\left(X^\alpha_N\right)\right], \tag{2.23}$$

where U is a utility function on $(0, \infty)$ satisfying the standard Inada conditions.

New results. We embed the multidimensional drift uncertainty into a Bayesian framework, and establish general results using an appropriate change of measure to derive the dynamic programming equation.

We now state the main results of Chapter 5. In the general case, following the reasoning of Section 5.3, we can write the dynamic programming equation associated to (2.23) in backward induction as,

$$\begin{cases} v_N(x, z, \mu) = U(x)\mu(\mathbb{R}^d), \quad (x, z) \in \mathcal{S}^q, \ \mu \in \mathcal{M}_+, \\ v_k(x, z, \mu) = \sup\limits_{a \in A^q(x,z)} \overline{\mathbb{E}}\Big[v_{k+1}\Big(x\big(1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big)\big), \\ \qquad\qquad \max\left[z, x\big(1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big)\big)\right], \frac{g(R_{k+1} - \cdot)}{g(R_{k+1})}\mu\Big)\Big], \end{cases} \tag{2.24}$$

for $k = 0, \ldots, N-1$, where

$$\mathcal{S}^q := \left\{(x, z) \in (0, \infty)^2 : qz \le x \le z\right\},$$

$\mathcal{M}_+$ is the set of nonnegative measures on $\mathbb{R}^d$, and

$$A^q(x,z) \;=\; \Big\{a \in \mathbb{R}^d_+ : \sum_{i=1}^d |a_i| \le 1 - q\frac{z}{x}\Big\}$$

is the static set of admissible controls.

Notice from Lemma 5.3.1, that the expectation in the above formula is only taken with respect to the noise R_{k+1}, which is distributed under $\overline{\mathbb{P}}$ according to the probability distribution ν with density g on $\mathbb{R}^d$.

In the case where the utility function is of CRRA (Constant Relative Risk Aversion) type, i.e.,

$$U(x) \;=\; \frac{x^p}{p}, \quad x > 0, \quad \text{for some } p < 1,\ p \neq 0, \tag{2.25}$$

we can reduce the dimensionality of the problem and find that the problem comes down to solving the following backward system

$$\begin{cases} w_N(r,\mu) &=& \frac{r^p}{p}\mu(\mathbb{R}^d), \quad r \in [q,1],\ \mu \in \mathcal{M}_+, \\ w_k(r,\mu) &=& \sup\limits_{a \in A^q(r)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min\big[1, r\big(1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big)\big)\big], \dfrac{g(R_{k+1} - \cdot)}{g(R_{k+1})}\mu\big)\Big], \end{cases} \tag{2.26}$$

for $k = 0, \dots, N-1$, where

$$A^q(r) \;=\; \Big\{a \in \mathbb{R}^d_+ : a'\mathbb{1}_d \le 1 - \frac{q}{r}\Big\}.$$

In the Gaussian framework, where the noise and the prior belief on the drift are modeled according to a Gaussian distribution, the Bayesian filtering is simplified into the Kalman filtering, which allows the dynamic programming system to be reduced to a finite-dimensional problem.

In this case, we prove that our initial portfolio selection Problem (2.23) can be reformulated as a full observation problem whose associated dynamic programming system satisfies

$$\begin{cases} \tilde{v}_N(x,z,b) &=& U(x), \quad (x,z) \in \mathcal{S}^q,\ b \in \mathbb{R}^d, \\ \tilde{v}_k(x,z,b) &=& \sup\limits_{a \in A^q(x,z)} \mathbb{E}\Big[\tilde{v}_{k+1}\Big(x\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big), \\ & & \qquad \max\big[z, x\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big)\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

for $k = 0, \dots, N-1$, where K_{k+1} is the Kalman gain matrix.

Notice that in the above formula, the expectation is taken with respect to the innovation vector $\tilde{\epsilon}_{k+1}$, which is distributed according to $\mathcal{N}(0, \tilde{\Gamma}_{k+1})$.[3]

Moreover, with a utility function as in Equation (2.25), the problem is simplified and consists in solving the following backward system:

$$\begin{cases} \tilde{w}_N(r,b) &= \frac{r^p}{p}, \quad r \in [q,1],\ b \in \mathbb{R}^d, \\ \tilde{w}_k(r,b) &= \sup\limits_{a \in A^q(r)} \mathbb{E}\Big[\tilde{w}_{k+1}\Big(\min\big[1, r\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big)\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

for $k = 0, \dots, N-1$.

Next, we apply the previous theoretical study to a three-risky asset market model, leading to a five-dimensional problem. We solve numerically the problem in the Gaussian case with CRRA utility functions using the deep learning *Hybrid-Now* algorithm.[4] The numerical results allow us to show the outperformance of the learning strategy over the non-learning one. We also detail the portfolio allocations of each strategy benchmarked with a comparable equally weighted strategy.

Finally, we assess the outperformance of the learning compared to the non-learning strategy with respect to the sensitivity of the drift uncertainty. Additionally, we provide empirical evidence of convergence of the non-learning strategy towards the solution of the classical Merton problem when the parameter controlling the maximum drawdown vanishes.

Difficulty. The difficulties we have faced in this problem are threefold:
1) Find a way to incorporate the maximum drawdown constraint and the drift uncertainty into a dynamic programming setup. We turn the maximum drawdown constraint on the wealth process into a constraint of the control domain, while managing the uncertain drift through a Bayesian learning approach.
2) Solve numerically the problem and find the appropriate numerical approach. We choose to use a recently developped algorithm, called *Hybrid-Now*, based on neural networks.
3) Reduce the dimensionality of the problem. In the Gaussian case and with a CRRA type utility function, we are able to reduce the dimensionality of the problem and remain with state variables that we can easily simulate. This is important in the training phase of the neural networks, since we can easily span the possible

[3] Explicitely derived in Section 5.4.1
[4] Sated in Bachouch et al. (2018a)

values of the state variables via simulations. Without this change of variable, it would have been more challenging to simulate the state variables, for instance the one representing the current historical maximum.

Further research. This work can be extended in multiple directions. One can think of different priors for the drift, and study other constraints that the portfolio could satisfy. One can also investigate a way to simulate the current historical maximum of a given random process and generalize our numerical method of resolution. Eventually, one could test these methods with historical data and measure the informative value of the learning strategy compared to the non-learning one and its sensitivity to a well-chosen uncertainty parameter formalized by the covariance matrix of the drift.

Part I

The Markowitz portfolio selection problem with drift uncertainty

Chapter 3

Bayesian learning for the Markowitz portfolio selection problem

Abstract. We study the Markowitz portfolio selection problem with unknown drift vector in the multidimensional framework. The prior belief on the uncertain expected rate of return is modeled by an arbitrary probability law, and a Bayesian approach from filtering theory is used to learn the posterior distribution about the drift given the observed market data of the assets. The Bayesian Markowitz problem is then embedded into an auxiliary standard control problem that we characterize by a dynamic programming method and prove the existence and uniqueness of a smooth solution to the related semi-linear partial differential equation (PDE). The optimal Markowitz portfolio strategy is explicitly computed in the case of a Gaussian prior distribution. Finally, we measure the quantitative impact of learning, updating the strategy from observed data, compared to non-learning, using a constant drift in an uncertain context, and analyze the sensitivity of the value of information with respect to various relevant parameters of our model.

Key Words: Bayesian learning, optimal portfolio, Markowitz problem, portfolio selection.

Contents

3.1 Introduction **68**
3.2 Markowitz problem with prior law on the uncertain drift **71**
3.3 Bayesian learning **74**
3.4 Solution to the Bayesian-Markowitz problem **77**
3.4.1 Main result 77
3.4.2 On existence and smoothness of the Bayesian risk premium 80

3.4.3 Examples . . . 81
3.4.3.1 Prior discrete law . . . 82
3.4.3.2 The Gaussian case . . . 83
3.5 Impact of learning on the Markowitz strategy . . . 85
3.5.1 Computation of the Sharpe ratios . . . 85
3.5.2 Value of Information . . . 88
3.5.2.1 Standard deviation of the drift . . . 89
3.5.2.2 Sharpe ratio of the asset . . . 91
3.5.2.3 Time . . . 92
3.5.2.4 Investment horizon . . . 93
3.6 Conclusion . . . 94
3.A Appendices . . . 95
3.A.1 Proof of Lemma 3.2.3 . . . 95
3.A.2 Proof of Lemma 3.2.4 . . . 97
3.A.3 Proof of Theorem 3.4.1 . . . 98
3.A.4 Proofs of Theorem 3.4.3 . . . 102
3.A.5 Proof of Lemma 3.4.5 . . . 106

3.1 Introduction

Portfolio selection is a core problem in mathematical finance and investment management. Its purpose is to choose the best portfolio according to a criterion of optimality. The mean-variance optimization provides a criterion of optimality considering the best portfolio as the one that maximizes an expected level of return given a certain level of variance, measuring the risk the investor can bear, or conversely, minimizes the variance of a portfolio given an expected level of return. Markowitz (1952) pioneered modern portfolio theory by settling the basic concepts. This conceptual framework, still widely used in the industry, leads to the efficient frontier principle which exhibits an intuitive relationship between risk and return.
Later, portfolio selection theory was extended several times to encompass multi-period problems, in discrete time by Samuelson (1969) and in continuous time by Merton (1969). Karatzas et al. (1987) made a decisive step forward when they solved Merton's problem for a large class of smooth utility functions using dual martingale methods under the no-bankruptcy constraint.
Originally, the literature on portfolio selection assumed that the parameters of the model, drifts and volatilities, are known and constant. This assumption raises the

question of estimating these parameters. Typically, the parameters are estimated from past data and are fixed once and for all. This convenient solution does not look realistic in practice since it does not adapt to changing market conditions. Moreover, as Merton (1980) among others showed, estimates of variances and covariances of the assets are more accurate than the estimates of the means. Indeed, he demonstrated the slow convergence of his estimators of the instantaneous expected return in a log-normal diffusion price model. Later, Best and Grauer (1991) argued that mean-variance optimal portfolios are very sensitive to the level of expected returns. As a consequence, estimating the expected return is not only more complicated than estimating the variance/covariance, but also a wrong estimation of the expected return can result in a very suboptimal portfolio a posteriori.

To circumvent these issues, an extensive literature incorporating parameters uncertainty in portfolio analysis emerged, see for example Barry (1974) and Klein and Bawa (1976), and methods using Bayesian statistics were developed, see Frost and Savarino (1986), Aguilar and West (2000), Avramov and Zhou (2010), and Bodnar et al. (2017). In particular, the Black and Litterman (1992) model, based on economic arguments and equilibrium relations, provided more stable and diversified portfolios than simple mean-variance optimization. Based upon the Markowitz problem and the capital asset pricing model (CAPM), this model remains static and cannot benefit from the flow of information fed by the market prices of the assets. This loss of information is detrimental in the optimization process, since it does not allow for an update of the model given the most recent available information.

Consequently, research then focused on taking advantage of the latest information conveyed by the prices. It explains the subsequent growing literature on filtering and learning techniques in a partial information framework, see Lakner (1995), Lakner (1998), Rogers (2001), and Cvitanić et al. (2006). The most noticeable research involving optimization and Bayesian learning techniques is by Karatzas and Zhao (2001). Using martingale methods, they computed the optimal portfolio allocation for a large class of utility functions in the case of an unknown drift and Gaussian asset returns. Bismuth, Guéant, and Pu (2019) extended the previous results to take into account both the liquidity and the expected returns of the assets, coupling Bayesian learning and dynamic programming techniques while addressing optimal portfolio liquidation and transition problems. Recently, Bauder et al. (2020) suggest to deal with the Markowitz problem using the posterior predictive distribution. A Bayesian efficient frontier was also derived and proved to outperform the overoptimistic sample efficient frontier.

In this chapter, we consider an investor who is willing to invest in an optimal portfolio in the Markowitz sense for a finite given time horizon. The investment universe consists of assets for which we assume the covariance matrix known and constant. The drift vector is uncertain and supposed to have a prior probability distribution. We contribute to the literature of optimal portfolio selection by formulating

and solving, in the multidimensional case, the Markowitz problem in the case of an uncertain drift modeled by a prior distribution. This a priori time inconsistent problem cannot be tackled by convex duality and martingale method as in Karatzas and Zhao (2001). Instead, we adapt the methodology of Zhou and Li (2000) to our Bayesian learning framework, and show how the Bayesian-Markowitz problem can be embedded into an auxiliary standard control problem that we study by the dynamic programming approach. Optimal strategies are then characterized in terms of a smooth solution to a semi-linear partial differential equation, and we highlight the effect of Bayesian learning compared to classical strategies based on constant drift. Although prior conjugates are widely used in the literature, here we extend to the multidimensional case, a recent result by Ekstrom and Vaicenavicius (2016) which enables us to use any prior distribution with invertible covariance matrix to model the uncertain drift allowing for a wide range of investors' strategies. In the particular case of a multidimensional Gaussian prior conjugate, we are able to exhibit a closed-form analytical formula.
Next, we measure the benefit of learning on the Markowitz strategy. To do so, we compare the Bayesian learning strategy to the subsequently called non-learning strategy which considers the coefficients of the model, especially the drift, constant. The Bayesian learning strategy is characterized by the fact that it uses, as information, the updated market prices of the assets to adjust the investment strategy and reaches optimality in the Markowitz sense. On the other hand, the non-learning strategy is a typical constant parameter strategy that estimates once and for all its unknown parameters with past data at time 0, leading to suboptimal solutions to the Markowitz problem because it misses the most recent information. The benefit of learning, measured as the difference between the Sharpe ratios of the portfolios based on the Bayesian learning and the non-learning strategies, is called the value of information in the sequel. We then analyze in a one-dimensional model the sensitivities of the value of information with respect to four essential variables: drift volatility, Sharpe ratio of the asset, time and investment horizon.

The chapter is organized as follows: Section 3.2 depicts the Markowitz problem in our framework of drift uncertainty modeled by a prior distribution and the steps to turn it into a solvable problem. Section 3.3 introduces the Bayesian learning framework and the methodology to consider any prior with positive definite covariance matrices. Section 3.4 provides the main results and some specific examples where some computations are possible analytically. Section 3.5 concerns the value of information and its sensitivity to different parameters of an unidimensional model: the standard deviation of the drift, the Sharpe ratio of the asset, the time, and the investment horizon. Section 3.6 concludes this chapter.

3.2 Markowitz problem with prior law on the uncertain drift

We consider a financial market on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual conditions, and on which is defined a standard n-dimensional Brownian motion W. The market consists of n risky assets that are continuously traded. The assets prices process S is defined as:

$$\left\{ \begin{array}{rcl} dS_t & = & \mathrm{diag}(S_t)\big(Bdt + \sigma dW_t\big), \quad t \in [0,T], \\ S_0 & = & s_0 \in \mathbb{R}^n, \end{array} \right. \tag{3.1}$$

where T is the investment horizon, B is a random vector in $\mathbb{R}^n$, independent of W, with distribution μ and such that $\mathbb{E}[|B|^2] < \infty$. The prior law of B represents the subjective beliefs of the investor about the likelihood of the different values that B might take.

In the sequel, we consider the following assumptions satisfied:

Assumption 3.2.1. *The covariance matrix* $\mathrm{Cov}\,(B)$ *of the random vector* B *is positive definite,*

Assumption 3.2.2. *The* $n \times n$ *matrix* σ *is known and invertible, and we denote* $\Sigma = \sigma\sigma'$.

We write $\mathbb{F}^{\mathbb{S}} = \{\mathcal{F}_t^S\}_{t\geq 0}$ the filtration generated by the price process S augmented by the null sets of $\mathcal{F}$. The filtration $\mathbb{F}^S$ represents the only available source of information: the investor observes the price process S but not the drift.
We denote by $\alpha = (\alpha_t)_{0\leq t\leq T}$ an admissible investment strategy representing the amount invested in each of the n risky assets. It is a n-dimensional $\mathbb{F}^{\mathbb{S}}$-progressively measurable process, valued in $\mathbb{R}^n$, and satisfying the following integrability condition

$$\mathbb{E}\bigg[\int_0^T |\alpha_t|^2 dt\bigg] < \infty.$$

We define $\mathcal{A}$ as the set of all admissible investment strategies.
The evolution of the self-financing wealth process X^α, given $\alpha \in \mathcal{A}$ and an initial wealth $x_0 \in \mathbb{R}$, is given by:

$$dX_t^\alpha = \alpha_t' \mathrm{diag}(S_t)^{-1} dS_t = \alpha_t'(Bdt + \sigma dW_t), \quad 0 \leq t \leq T, \quad X_0^\alpha = x_0. \tag{3.2}$$

The Markowitz problem is then formulated as :

$$U_0(\vartheta) \;:=\; \sup_{\alpha\in\mathcal{A}} \Big\{ \mathbb{E}[X_T^\alpha] : \mathrm{Var}(X_T^\alpha) \leq \vartheta \Big\}, \tag{3.3}$$

where $\vartheta > 0$ is the variance budget the investor can allow on her portfolio. The value function U_0 is then the expected optimal terminal wealth value the investor can achieve given her variance constraint. It is very standard to transform the initial problem (3.3) into a so called mean-variance problem, where the constraint becomes part of the objective function:

$$V_0(\lambda) \ := \ \inf_{\alpha \in \mathcal{A}} \left[\lambda \mathrm{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha] \right], \quad \lambda > 0. \tag{3.4}$$

The well-known connection between the Markowitz and mean-variance problem is stated in the following Lemma.

Lemma 3.2.3. *We have the duality relation*

$$\begin{cases} V_0(\lambda) &= \inf_{\vartheta > 0} \left[\lambda \vartheta - U_0(\vartheta) \right], \quad \forall \lambda > 0 \\ U_0(\vartheta) &= \inf_{\lambda > 0} \left[\lambda \vartheta - V_0(\lambda) \right], \quad \forall \vartheta > 0. \end{cases}$$

Furthermore, if $\alpha^{*,\lambda}$ is an optimal control for Problem (3.4), then $\hat{\alpha}^\vartheta := \alpha^{*,\lambda(\vartheta)}$ is an optimal control for Problem (3.3), with $\lambda(\vartheta) := \arg\min_{\lambda > 0} \left[\lambda \vartheta - V_0(\lambda) \right]$, and $\mathrm{Var}(X_T^{\hat{\alpha}^\vartheta}) = \vartheta$.

Proof. See Appendix 3.A.1. □

Because of the variance term, mean-variance Problem (3.4) is not a classical time-consistent control problem. To circumvent this issue, we adopt a similar approach as in Zhou and Li (2000) in order to embed it into an auxiliary control problem, for which we can use the dynamic programming method.

Lemma 3.2.4. *The mean-variance optimization Problem (3.4) can be written as follows:*

$$V_0(\lambda) \ = \ \inf_{\gamma \in \mathbb{R}} \left[\tilde{V}_0 \left(\lambda, \gamma \right) + \lambda \gamma^2 \right] \tag{3.5}$$

where

$$\tilde{V}_0(\lambda, \gamma) \ := \ \inf_{\alpha \in \mathcal{A}} \left[\lambda \mathbb{E}[(X_T^\alpha)^2] - (1 + 2\lambda\gamma) \mathbb{E}[X_T^\alpha] \right]. \tag{3.6}$$

Moreover, if $\tilde{\alpha}^{\lambda,\gamma}$ is an optimal control for (3.6) and $\gamma^*(\lambda) := \mathrm{argmin}_{\gamma \in \mathbb{R}} \left[\tilde{V}_0(\lambda, \gamma) + \lambda \gamma^2 \right]$ then $\alpha^{*,\lambda} := \tilde{\alpha}^{\lambda, \gamma^*(\lambda)}$ is an optimal control for $V_0(\lambda)$ in (3.4).

Furthermore, the following equality holds true

$$\gamma^*(\lambda) \ = \ \mathbb{E}[X_T^{\alpha^{*,\lambda}}]. \tag{3.7}$$

Proof. See Appendix 3.A.2. □

As we shall see in Section 3.4, Problem (3.6) is well suited for dynamic programming.

We end this section by recalling the solution to the Markowitz problem in the case of constant known drifts and volatilities, which will serve as a benchmark to refer to.

Remark 3.2.5 (Case of constant known drift)**.** In the case where the drift vector b_0 and the matrix σ are known and constant, we have the following results:

- The value function of the mean-variance problem with initial wealth x_0 is given by

$$V_0(\lambda) \;=\; v^\lambda(0,x_0) = -\frac{1}{4\lambda}\left(e^{|\sigma^{-1}b_0|^2T}-1\right) - x_0$$

with

$$v^\lambda(t,x) \;=\; \lambda e^{-|\sigma^{-1}b_0|^2(T-t)}\,(x-x_0)^2 - \frac{1}{4\lambda}\left(e^{|\sigma^{-1}b_0|^2(T-t)}-1\right) - x.$$

The corresponding optimal portfolio strategy for $V_0(\lambda)$ is given in feedback form by

$$\alpha_t^{*,\lambda} \;=\; a_0^\lambda(t,X_t^*),$$

where

$$a_0^\lambda(t,x) \;:=\; \Big(x_0 - x + \frac{e^{|\sigma^{-1}b_0|^2T}}{2\lambda}\Big)\Sigma^{-1}b_0$$

and X_t^* is the optimal wealth obtained with the optimal feedback strategy.

- The optimal portfolio strategy of Markowitz Problem $U_0(\vartheta)$ is then given by

$$\hat{\alpha}_t^\vartheta \;=\; \alpha_t^{*,\lambda_0(\vartheta)}$$

where

$$\lambda_0(\vartheta) \;:=\; \sqrt{\frac{e^{|\sigma^{-1}b_0|^2T}-1}{4\vartheta}}.$$

Moreover, the solution to the Markowitz problem is

$$U_0(\vartheta) \;=\; x_0 + \sqrt{\vartheta\big(e^{|\sigma^{-1}b_0|^2T}-1\big)}.$$

See for instance Zhou and Li (2000) for more details. ◊

3.3 Bayesian learning

Since the investor does not observe the assets drift vector B, she needs to have a subjective belief on its potential value. It is represented as a prior probability distribution μ on $\mathbb{R}^n$, assumed to satisfy for some $a > 0$,

$$\int_{\mathbb{R}^n} e^{a|b|^2} \mu(db) < \infty.$$

The prior probability distribution will then learn and infer the value of the drift from observable samples of the assets prices. Using filtering theory, and following in particular the recent work by Ekstrom and Vaicenavicius (2016) that we extend to the multidimensional case, we compute the posterior probability distribution of B given the observed assets prices.

Let us first introduce the $\mathbb{R}^n$-valued process $Y_t = \sigma^{-1}Bt + W_t$, which clearly generates the same filtrations as S, and write the observation process

$$S_t^i \;=\; S_0^i \exp\Big(< \sigma^i, Y_t > - \frac{|\sigma^i|^2}{2} t \Big), \quad i = 1, \dots, n, \;\; 0 \le t \le T,$$

where σ^i denotes the i-th line of the matrix σ.

Equivalently, we can express Y in terms of S as: $Y_t = h(t, S_t)$ where $h : [0,T] \times (0,\infty)^n \to \mathbb{R}^n$ is defined by

$$h(t,s) \;:=\; \sigma^{-1} \begin{pmatrix} \ln\left(\frac{s^1}{S_0^1}\right) + \frac{|\sigma^1|^2}{2} t \\ \vdots \\ \ln\left(\frac{s^n}{S_0^n}\right) + \frac{|\sigma^n|^2}{2} t \end{pmatrix}, \quad s = (s^1, \dots, s^n) \in (0,\infty)^n.$$

The following result gives the conditional distribution of B given observations of the assets market prices in terms of the current value of Y. We refer to Proposition 3.16 in Bain and Crisan (2009) as well as Karatzas and Zhao (2001), and Pham (2011) for a proof.

Lemma 3.3.1. *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ satisfying $\int_{\mathbb{R}^n} |g(b)| \mu(db) < \infty$. Then*

$$\mathbb{E}\big[g(B) | \mathcal{F}_t^S\big] \;=\; \mathbb{E}[g(B)|Y_t] \;=\; \frac{\int_{\mathbb{R}^n} g(b) e^{<\sigma^{-1}b, Y_t> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}{\int_{\mathbb{R}^n} e^{<\sigma^{-1}b, Y_t> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}.$$

From Lemma 3.3.1, the conditional distribution $\mu_{t,y}$ of B given $Y_t = y$ is given by

$$\mu_{t,y}(db) \;=\; \frac{e^{<\sigma^{-1}b, y> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}{\int_{\mathbb{R}^n} e^{<\sigma^{-1}b, y> - \frac{|\sigma^{-1}b|^2}{2}t} \mu(db)}, \tag{3.8}$$

and the posterior predictive mean of the drift is

$$\hat{B}_t \;:=\; \mathbb{E}[B|\mathcal{F}_t^S] \;=\; \mathbb{E}[B|Y_t] \;=\; f(t,Y_t),$$

where

$$f(t,y) \;:=\; \int_{\mathbb{R}^n} b\mu_{t,y}(db) \;=\; \frac{\int_{\mathbb{R}^n} be^{<\sigma^{-1}b,y>-\frac{|\sigma^{-1}b|^2}{2}t}\mu(db)}{\int_{\mathbb{R}^n} e^{<\sigma^{-1}b,y>-\frac{|\sigma^{-1}b|^2}{2}t}\mu(db)}. \tag{3.9}$$

The following result shows some useful properties regarding the function f.

Lemma 3.3.2. *For each $t \in [0,T]$, the function $f_t : \mathbb{R}^n \to \mathbb{R}^n$ defined as $f_t(y) := f(t,y)$, where f is given in (3.9), is invertible when restricted to its image $\mathcal{B}_t := f_t(\mathbb{R}^n)$, and we have*

$$\nabla_y f(t,y) \;=\; \mathrm{Cov}\left(B\,|Y_t = y\right)\left(\sigma^{-1}\right)', \quad y \in \mathbb{R}^n.$$

and the conditional covariance matrix of B given $Y_t = y$ is positive definite.

Proof. The vector-valued function $y \mapsto f_t(y)$ is differentiable, hence we can compute the matrix function $\nabla_y f$ element-wise,

$$\begin{aligned}
\left[\partial_{y_j} f_t(y)\right]_i &= \int_{\mathbb{R}^n} [\sigma^{-1}b]_j b_i \mu_{t,y}(db) - \int_{\mathbb{R}^n} b_i\mu_{t,y}(db)\int_{\mathbb{R}^n}[\sigma^{-1}b]_j\mu_{t,y}(db)\\
&= \sum_{k=1}^n \left(\sigma^{-1}_{j,k}\int_{\mathbb{R}^n} b_k b_i \mu_{t,y}(db) - \int_{\mathbb{R}^n} b_i\mu_{t,y}(db)\int_{\mathbb{R}^n} b_k\mu_{t,y}(db)\right)\\
&= \sum_{k=1}^n \sigma^{-1}_{j,k}\mathrm{cov}(B_i, B_k|Y_t = y)\\
&= \left(\sigma^{-1}\mathrm{cov}\left(B\mid Y_t = y\right)\right)_{j,i}\\
&= \left(\mathrm{cov}\left(B\mid Y_t = y\right)(\sigma^{-1})'\right)_{i,j}.
\end{aligned}$$

If $\mathrm{cov}\left(B \mid Y_t\right)$ is not positive definite, then for any linear combination of B, $\sum_{j=1}^n q_j B_j$ with $q \in \mathbb{R}^n$, we would have $\mathrm{Var}\left(\sum_{j=1}^n q_j B_j \mid Y_t\right) = 0$, meaning that $\sum_{j=1}^n q_j B_j = C$, $C \in \mathbb{R}$, $\mu_{t,y}$-a.e.. But from (3.8), this would imply $\sum_{j=1}^n q_j B_j = C$, $C \in \mathbb{R}$, μ-a.e which contradicts Assumption 3.2.1.

Fix $t \geq 0$, the function $f_t : \mathbb{R}^n \to \mathbb{R}^n$ is obviously surjective on its image $\mathcal{B}_t$. To show that it is injective, we define $\tilde{f} = f_t\sigma'$ and take two $\mathbb{R}^n$-vectors $y_1 \neq y_2$. If we compute the Taylor expansion of $\tilde{f}$,

$$\tilde{f}(y_1) - \tilde{f}(y_2) = \nabla_y \tilde{f}(\eta)(y_1 - y_2),$$

where η is a convex combination of y_1 and y_2. Since $\mathrm{cov}(B|Y_t)$ is positive definite $\nabla_y\tilde{f}$ is positive definite so $\tilde{f}(y_1) \neq \tilde{f}(y_2)$ and f_t is injective. □

Let us now introduce the so-called innovation process

$$\hat{W}_t \;:=\; \sigma^{-1}\int_0^t (B-\hat{B}_s)ds + W_t, \quad 0\le t\le T,$$

which is a $(\mathbb{P},\mathcal{F}^S)$-Brownian motion, see Proposition 2.30 in Bain and Crisan (2009). The observation process Y is written in terms of the innovation process as

$$Y_t \;=\; \sigma^{-1}\int_0^t \hat{B}_s ds + \hat{W}_t. \tag{3.10}$$

Applying Itô's formula to $\hat{B}_t = f(t,Y_t)$ with (3.10), and recalling that $\hat{B}$ is a $(\mathbb{P},\mathcal{F}^S)$-martingale, we see that $d\hat{B}_t = \nabla_y f(t,Y_t)d\hat{W}_t$. By Lemma 3.3.2, and defining the matrix-valued function ψ by

$$\psi(t,b) \;:=\; \nabla_y f(t, f_t^{-1}(b)), \quad t\in[0,T],\; b\in\mathcal{B}_t, \tag{3.11}$$

where $y = f_t^{-1}(b)$ is the unique value of the observation process Y_t that yields $\hat{B}_t = b$, we have

$$d\hat{B}_t \;=\; \psi(t,\hat{B}_t)d\hat{W}_t. \tag{3.12}$$

Remark 3.3.3 (Dirac case)**.** *The Dirac prior distribution $\mu = \mathbb{1}_{b_0}$ does not verify Assumption (3.2.1). In this case, $\forall\ (t,b)$, $f(t,b) := b_0$, consequently we cannot properly define the function f_t^{-1} and the matrix-valued function ψ. A natural way to extend our framework to the Dirac case is by setting $\forall (t,b)$, $\psi(t,b) := 0$. Indeed when B is known then $cov(B) = 0$ and thus $\psi = 0$.* ♢

We can rewrite Model (3.1) in terms of observable variables:

$$\begin{cases} dS_t &= \mathrm{diag}(S_t)\left(\hat{B}_t dt + \sigma d\hat{W}_t\right), \quad t\in[0,T], \qquad S_0 \;=\; s_0,\\ d\hat{B}_t &= \psi(t,\hat{B}_t)d\hat{W}_t, \qquad\qquad\quad t\in[0,T], \quad \hat{B}_0 \;=\; b_0 \;=\mathbb{E}[B]. \end{cases}$$

Notice that the dynamics of the observable process $\hat{B}_t = f(t,Y_t) = f(t,h(t,S_t))$ is fully determined by the function ψ given in analytical form by (3.11) (explicit computations will be given later in the next sections). The dynamics of the wealth process in (3.2) becomes

$$dX_t^\alpha \;=\; \alpha_t'\mathrm{diag}(S_t)^{-1}dS_t = \alpha_t'(\hat{B}_t dt + \sigma d\hat{W}_t), \quad 0\le t\le T, \quad X_0^\alpha = x_0. \tag{3.13}$$

3.4 Solution to the Bayesian-Markowitz problem

3.4.1 Main result

In order to solve initial Markovitz Problem (3.3), we first solve problem (3.6), then use Lemma 3.2.4 to find the optimal γ and obtain the dual function $V_0(\lambda)$ stated in (3.5), and finally apply Lemma 3.2.3 to find the optimal Lagrange multiplier λ which gives us the original value function $U_0(\vartheta)$ and the associated optimal strategy.
Let us define the dynamic value function associated to Problem (3.6):

$$v^{\lambda,\gamma}(t,x,b) := \inf_{\alpha\in\mathcal{A}} J^{\lambda,\gamma}(t,x,b,\alpha), \qquad t\in[0,T],\ x\in\mathbb{R},\ b\in\mathcal{B}_t, \tag{3.14}$$

with

$$J^{\lambda,\gamma}(t,x,b,\alpha) := \mathbb{E}\Big[\lambda(X_T^{t,x,b,\alpha})^2 - (1+2\lambda\gamma)X_T^{t,x,b,\alpha}\Big],$$

where $X^{t,x,b,\alpha}$ is the solution to (3.13) on $[t,T]$, starting at $X_t^{t,x,b,\alpha} = x$ and $\hat{B}_t = b$ at time $t\in[0,T]$, controlled by $\alpha\in\mathcal{A}$, so that $\tilde{V}_0(\lambda,\gamma) = v^{\lambda,\gamma}(0,x_0,b_0)$, with $b_0 = \mathbb{E}[B]$.

Problem (3.14) is a standard stochastic control problem that can be characterized by the dynamic programming Hamilton-Jacobi-Bellman (HJB) equation, which is a fully nonlinear PDE. Actually, by exploiting the quadratic structure of this control problem (as detailed below), the HJB equation can be reduced to the following semi-linear PDE on $\mathcal{R} = \{(t,b) : t\in[0,T], b\in\mathcal{B}_t\}$:

$$-\partial_t R - \frac{1}{2}\mathrm{tr}\big(\psi\psi'\mathcal{D}_b^2 R\big) + \tilde{F}(t,b,\nabla_b R) = 0, \tag{3.15}$$

with

$$\tilde{F}(t,b,p) := 2\left(\psi\sigma^{-1}b\right)'p - \frac{1}{2}|\psi' p|^2 - |\sigma^{-1}b|^2, \tag{3.16}$$

and terminal condition

$$R(T,b) = 0, \quad b\in\mathcal{B}_t, \tag{3.17}$$

We can then state the main result of this chapter, which provides the analytic solution to the Bayesian Markowitz problem.

Theorem 3.4.1. *Suppose there exists a solution $R\in C^{1,2}([0,T)\times\mathcal{B}_t;\mathbb{R}_+)\cap C^0(\mathcal{R};\mathbb{R}_+)$ to the semi-linear Equation (3.15) with terminal condition (3.17). Then, for any $\lambda > 0$,*

$$V_0(\lambda) \;=\; -\frac{1}{4\lambda}\left(e^{R(0,b_0)} - 1\right) - x_0$$

with the associated optimal mean-variance control given in feedback form by

$$\alpha_t^{*,\lambda} \;=\; a_0^{Bayes,\lambda}(t, X_t^{\alpha^{*,\lambda}}, \hat{B}_t) \;=\; a_0^{Bayes,\lambda}\big(t, X_t^{\alpha^{*,\lambda}}, f(t, h(t, S_t))\big), \tag{3.18}$$

where

$$a_0^{Bayes,\lambda}(t,x,b) := \Big(x_0 - x + \frac{e^{R(0,b_0)}}{2\lambda}\Big)\Big(\Sigma^{-1}b - \big(\psi\sigma^{-1}\big)'\nabla_b R(t,b)\Big), \tag{3.19}$$

and the corresponding optimal terminal wealth is equal to

$$\mathbb{E}\big[X_T^{\alpha^{*,\lambda}}\big] \;=\; x_0 + \frac{1}{2\lambda}\big(e^{R(0,b_0)} - 1\big). \tag{3.20}$$

Moreover, for any $\vartheta > 0$,

$$U_0(\vartheta) \;=\; x_0 + \sqrt{\vartheta\,(e^{R(0,b_0)} - 1)}$$

with the associated optimal Bayesian-Markowitz strategy given by

$$\hat{\alpha}^{\vartheta} \;=\; \alpha^{*,\lambda(\vartheta)},$$

with

$$\lambda(\vartheta) \;=\; \sqrt{\frac{e^{R(0,b_0)} - 1}{4\vartheta}}. \tag{3.21}$$

Proof. The detailed proof is postponed in Appendix 3.A.3, and we only sketch here the main arguments.
For fixed (λ, γ), the HJB equation associated to standard stochastic control Problem (3.14) is written as

$$\partial_t v^{\lambda,\gamma} + \frac{1}{2}\mathrm{tr}\big(\psi\psi'\mathcal{D}_b^2 v^{\lambda,\gamma}\big) + \inf_{\alpha\in\mathcal{A}}\Big[\alpha' b\partial_x v^{\lambda,\gamma} + \frac{1}{2}|\sigma'\alpha|^2\partial_{xx}^2 v^{\lambda,\gamma} + \alpha'\sigma\psi'\partial_{xb}^2 v^{\lambda,\gamma}\Big] \;=\; 0, \tag{3.22}$$

with terminal condition

$$v^{\lambda,\gamma}(T,x,b) \;=\; \lambda x^2 - (1 + 2\lambda\gamma)x.$$

We look for a solution in the ansatz form: $v^{\lambda,\gamma}(t,x,b) = K(t,b)x^2 + \Gamma(t,b)x + \chi(t,b)$. By plugging into the above HJB equation, and identifying the terms in x^2, x, we find after some straightforward calculations that

$$v^{\lambda,\gamma}(t,x,b) \;=\; e^{-R(t,b)}\left[\lambda x^2 - (1+2\lambda\gamma)x + \frac{(1+2\lambda\gamma)^2}{4\lambda}\right] - \frac{(1+2\lambda\gamma)^2}{4\lambda},$$

where R has to satisfy the semi-linear PDE (3.15) (which does not depend on λ, γ). Since R is assumed to exist smooth, the optimal feedback control achieving the argmin in HJB Equation (3.22) is given by

$$\tilde{\alpha}_t^{\lambda,\gamma} = \tilde{a}^{\lambda,\gamma}(t, X_t^{\tilde{\alpha}^{\lambda,\gamma}}, \hat{B}_t), \tag{3.23}$$

with

$$\tilde{a}^{\lambda,\gamma}(t,x,b) \ := \ \Big(\frac{1}{2\lambda} + \gamma - x\Big)\big(\Sigma^{-1}b - (\psi\sigma^{-1})'\nabla_b R(t,b)\big), \tag{3.24}$$

while $\gamma^*(\lambda) := \arg\min_{\gamma\in\mathbb{R}}\left[\tilde{V}_0(\lambda,\gamma)+\lambda\gamma^2\right] = \arg\min_{\gamma\in\mathbb{R}}\left[\tilde{v}^{\lambda,\gamma}(0,x_0,b_0)+\lambda\gamma^2\right]$ is given by

$$\gamma^*(\lambda) \ = \ x_0 + \frac{1}{2\lambda}\big(e^{R(0,b_0)} - 1\big), \tag{3.25}$$

which leads to the expression (3.20) from (3.7). We deduce that

$$\begin{aligned} V_0(\lambda) &= \tilde{V}_0(\lambda,\gamma^*(\lambda)) + \lambda(\gamma^*(\lambda))^2 \\ &= v^{\lambda,\gamma^*(\lambda)}(0,x_0,b_0) + \lambda(\gamma^*(\lambda))^2 \\ &= -\frac{1}{4\lambda}\left(e^{R(0,b_0)} - 1\right) - x_0. \end{aligned} \tag{3.26}$$

From Lemma 3.2.4, we know $\alpha_t^{*,\lambda} = \tilde{\alpha}_t^{\lambda,\gamma^*(\lambda)}$ and the optimal control for $V_0(\lambda)$ is thus

$$\alpha_t^{*,\lambda} = \tilde{a}^{\lambda,\gamma^*(\lambda)}(t, X_t^{\tilde{a}^{\lambda,\gamma^*(\lambda)}}, \hat{B}_t) = a_0^{Bayes,\lambda}(t, X_t^{\alpha^{*,\lambda}}, \hat{B}_t), \quad 0 \le t \le T,$$

with the function $a_0^{Bayes,\lambda}$ as in (3.19). This leads to the expression in (3.18) from (3.23), (3.24) and (3.25). Finally, the Lagrange multiplier $\lambda(\vartheta) = \arg\min_{\lambda>0}\left[\lambda\vartheta - V_0(\lambda)\right]$ is explicitly computed from (3.26), and is equal to the expression in (3.21). The optimal performance of the Markowitz problem is then equal to

$$U_0(\vartheta) \ = \ \lambda(\vartheta)\vartheta - V_0\big(\lambda(\vartheta)\big) \ = \ x_0 + \sqrt{\vartheta\,(e^{R(0,b_0)} - 1)},$$

by (3.21) and (3.26). The optimal Bayesian-Markowitz strategy is then given by $\hat{\alpha}^\vartheta = \alpha^{*,\lambda(\vartheta)}$ according to Lemma 3.2.3. □

Remark 3.4.2 (Financial interpretation)**.** When the drift $B = b_0$ is known, the function R is simply given by $R(t) = |\sigma^{-1}b_0|^2(T-t)$, and we retrieve the results recalled in Remark 3.2.5. Under a prior probability distribution on the drift B, we see that the optimal performance of the Bayesian-Markowitz problem has a similar form as in the constant drift case, when substituting $|\sigma^{-1}b_0|^2T$ by $R(0,b_0)$. For the optimal Bayesian-Markowitz strategy, we have to substitute the vector term $\Sigma^{-1}b_0$ by $\Sigma^{-1}\hat{B}_t$, i.e. replacing b_0 by the posterior predictive mean $\hat{B}_t$, and correcting with the additional term $(\psi\sigma^{-1})'\nabla_b R(t,\hat{B}_t)$. We call R the Bayesian risk premium function. ♢

Our main result in Theorem 3.4.1 is stated under the condition that R exists sufficiently smooth. In the next paragraph, we give some sufficient conditions ensuring this regularity, thus the existence of R, and provide some examples.

3.4.2 On existence and smoothness of the Bayesian risk premium

In this section we provide sufficient conditions ensuring the existence of the Bayesian risk premium function R, solution to the semi-linear PDE (3.15). Note that the standard assumptions of existence and uniqueness that we can find in Ladyzhenskaia et al. (1968) do not apply here since the function $\tilde{F}$ defined in (3.16) is not globally Lipschitz in p. The difficulty of our framework comes from the unboundedness of the domain and the quadratic growth in p of the function $\tilde{F}$ which entails that the solution may grow quadratically.

Theorem 3.4.3. *Suppose the following conditions hold true:*

- $\forall (t,b) \in \mathcal{R}$, $\psi(t,b)$ *and* $\psi^{-1}(t,b)$ *are bounded,*
- *The matrix* $\nabla_b \psi$ *exists and is bounded,*

then there exists a classical solution $R \in \mathcal{C}^{1,2}(\mathcal{R})$ *to PDE (3.15). In addition,* $R(t,b)$ *is at most quadratically growing in* b *and* $\nabla_b R(t,b)$ *is at most linearly growing in* b.

Proof. We provide a sketch of the proof and detail in Appendix 3.A.4. We follow a similar approach as Pham (2002) and Benth and Karlsen (2005).Without loss of generality we consider the function $\tilde{R} := -R$. We rewrite the PDE characterizing $\tilde{R}$ as

$$-\partial_t \tilde{R} - \frac{1}{2}\mathrm{tr}(\psi\psi'\mathcal{D}_b^2 \tilde{R}) + F(t,b,\nabla_b \tilde{R}) \;=\; 0$$

with terminal condition $\tilde{R}(T,.) = 0$, where the function $F : \mathcal{R} \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$F(t,b,p) \;=\; \frac{1}{2}|\psi' p|^2 + 2\big(\psi\sigma^{-1}b\big)' p + |\sigma^{-1}b|^2,$$

is quadratic (hence convex in the gradient argument p). Let us introduce the Fenchel-Legendre transform of $F(t,b,.)$ as

$$L_t(b,q) \;:=\; \max_{p\in\mathbb{R}^n}\big[-q'\psi' p - F(t,b,p)\big], \quad b \in \mathcal{B}_t,\; q \in \mathbb{R}^n,$$

which explicitly yields

$$L_t(b,q) \;=\; \frac{1}{2}|q|^2 + 2(\sigma^{-1}b)'q + |\sigma^{-1}b|^2. \tag{3.27}$$

The explicit form shows that L_t only depends on t through the domain $\mathcal{B}_t$ of b. It is also known that the following duality relationship holds:

$$F(t,b,p) \;=\; \max_{q\in\mathbb{R}^n}\big[-q'\psi' p - L_t(b,q)\big].$$

From (3.27) we deduce the following estimates on L and its gradient with respect to b:

$$\left\{\begin{array}{rcl} |L_t(b,q)| & \leq & C(|q|^2+|b|^2), \\ |\nabla_b L_t(b,q)| & \leq & C(|q|+|b|), \quad \forall b\in\mathcal{B}_t,\ q\in\mathbb{R}^n, \end{array}\right. \tag{3.28}$$

for some independent of t positive constant C. Let us now consider the truncated auxiliary function

$$F^k(t,b,p) \;=\; \max_{|q|\leq k}\big[-q'\psi' p - L_t(b,q)\big], \quad b\in\mathcal{B}_t,\ p\in\mathbb{R}^n,$$

for $k\in\mathbb{N}$. By (3.28) and Theorem 4.3 p. 163 in Fleming and Soner (2006), there exists a unique smooth solution $\tilde{R}^k$ with quadratic growth condition to the truncated semi-linear PDE

$$-\partial_t\tilde{R}^k - \frac{1}{2}\mathrm{tr}(\psi\psi'\mathcal{D}_b^2\tilde{R}^k) + F^k(t,b,\nabla_b\tilde{R}^k) \;=\; 0,$$

with terminal condition $\tilde{R}^k(T,.)=0$.
The next step is to obtain estimates on $\tilde{R}^k$ and its gradient with respect to b, uniformly in k:

$$\left\{\begin{array}{rcl} |\tilde{R}^k(t,b)| & \leq & C(1+|b|^2), \\ |\nabla_b\tilde{R}^k(t,b)| & \leq & C(1+|b|), \quad \forall (t,b)\in\mathcal{R}, \end{array}\right.$$

for some constant C independent of k. Then, following similar arguments as in Pham (2002) and Benth and Karlsen (2005), we show that for k large enough, $R^k=-\tilde{R}^k$ solves PDE (3.15). This proves the existence of a smooth solution to this semi-linear PDE. □

3.4.3 Examples

We illustrate with some relevant examples the explicit computation of the diffusion coefficient ψ appearing in the dynamics of the posterior mean of the drift described in (3.12), as well as the computation of the Bayesian risk premium function R.

3.4.3.1 Prior discrete law

We consider the case when the drift vector B has a prior discrete distribution

$$\mu(db) \;=\; \sum_{i=1}^{N} \pi_i \delta_{V_i}(db), \tag{3.29}$$

where for $i = 1, \ldots, N$, V_i are vectors in $\mathbb{R}^n$, $\pi_i \in (0,1)$ and $\sum_{i=1}^N \pi_i = 1$. We denote by V the $n \times N$-matrix $V = (V_1 \ldots V_N)$, with rank(V) $= n < N$ and we assume that the vectors V_i are chosen such that

$$\mathrm{Cov}(B) = \sum_{i=1}^{N} \pi_i V_i V_i' - b_0 b_0' > 0,$$

remembering that $b_0 = \mathbb{E}[B]$.

From (3.29) and (3.8), we easily compute the conditional distribution of B with respect to $Y_t = y \in \mathbb{R}^n$, which is given by

$$\mu_{t,y}(db) \;=\; \sum_{i=1}^{N} p^i_{t,y} \delta_{V_i}(db) \quad t \in [0,T],$$

where $p_{t,y} = (p^1_{t,y} \ldots p^N_{t,y}) \in [0,1]^N$ is determined by

$$p^i_{t,y} \;=\; \frac{\pi_i e^{y'\sigma^{-1}V_i - \frac{1}{2}V_i'\Sigma^{-1}V_i t}}{\sum_{j=1}^N \pi_j e^{y'\sigma^{-1}V_j - \frac{1}{2}V_j'\Sigma^{-1}V_j t}}, \quad i = 1, \ldots, N.$$

It follows that the function f in (3.9) is equal to

$$f(t,y) \;=\; \int_{\mathbb{R}^n} b\mu_{t,y}(db) \;=\; \sum_{i=1}^{N} p^i_{t,y} V_i \;=\; V p_{t,y},$$

from which we calculate its gradient:

$$\nabla_y f(t,y) \;=\; \left(\sum_{i=1}^{N} p^i_{t,y} V_i V_i' - V p_{t,y} (V p_{t,y})' \right) (\sigma^{-1})'.$$

In this case the domain $\mathcal{B}_t$ of b is the convex hull of the vectors V_i, $i = 1, \ldots, N$ mathematically defined as

$$\mathcal{B}_t = \left\{ b : V p_{t,y^*} = b \text{ and } \forall i \in \{1, \ldots, N\}, \;\; p^i_{t,y^*} \in (0,1) \;\text{ and }\; \sum_{i=1}^{N} p^i_{t,y^*} = 1 \right\},$$

where for fixed $(t,b) \in \mathcal{R}$, $y^* = f_t^{-1}(b)$. More precisely, y^* is the solution to the equation $f(t,y^*) = Vp_{t,y^*} = b$. We note $\overline{\mathcal{B}}_t$ the closure of $\mathcal{B}_t$.
The function ψ defined in (3.11) is therefore equal to

$$\begin{aligned}\psi(t,b) &= \nabla_y f(t,y^*) \\ &= \left(\sum_{i=1}^{N} p^i_{t,y^*} V_i V_i' - bb'\right)(\sigma^{-1})'.\end{aligned} \tag{3.30}$$

Remark 3.4.4. In the discrete case we cannot apply Theorem 3.4.3 for two main reasons: $\mathcal{B}_t$ is a bounded non-smooth domain, and ψ can vanish on the boundary of $\mathcal{B}_t$, leading to ψ^{-1} being ill defined on $\overline{\mathcal{B}_t}$. It is, in particular, easy to see that ψ vanishes at the points V_i. Indeed, b equals V_i implies that $\forall\ j = 1,\dots,N$, $p^j_{t,y^*} = 0$ for $j \neq i$ and $p^i_{t,y^*} = 1$ leading to Equation (3.30) being zero. The existence theorem of a solution to Problem (3.15)-(3.17), in our context, can be found in Krylov (1987), especially chapter 6, section 3. ♢

3.4.3.2 The Gaussian case

We consider the case when the drift vector B follows a n-dimensional normal distribution with mean vector b_0 and covariance matrix Σ_0:

$$\mu \ \sim\ \mathcal{N}(b_0, \Sigma_0). \tag{3.31}$$

In this conjugate case, it is well-known (see Proposition 10 in Bismuth, Guéant, and Pu (2019)) that the posterior distribution of B, i.e., the conditional distribution of B given $Y_t = y$ is also Gaussian with mean

$$f(t,y) \ = \ \left(\Sigma_0^{-1} + \Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0 + (\sigma')^{-1}y\right), \tag{3.32}$$

and covariance

$$\Sigma(t,y) \ = \ \left(\Sigma_0^{-1} + \Sigma^{-1}t\right)^{-1}.$$

For the sake of completeness, we show this result. From (3.8) and (3.31), the conditional distribution of B with respect to $Y_t = y$ is given by

$$\begin{aligned}\mu_{t,y}(db) =& \frac{e^{y'\sigma^{-1}b - \frac{1}{2}b'\Sigma^{-1}bt - \frac{1}{2}(b-b_0)'\Sigma_0^{-1}(b-b_0)}db}{\int_{\mathbb{R}^n} e^{y'\sigma^{-1}b - \frac{1}{2}b'\Sigma^{-1}bt - \frac{1}{2}(b-b_0)'\Sigma_0^{-1}(b-b_0)}db} \\ =& \frac{e^{-\frac{1}{2}\left(b'\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)b - 2b'\left(\Sigma_0^{-1}b_0 + (\sigma^{-1})'y\right)\right)}db}{\int_{\mathbb{R}^n} e^{-\frac{1}{2}\left(b'\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)b - 2b'\left(\Sigma_0^{-1}b_0 + (\sigma^{-1})'y\right)\right)}db} \\ =& \frac{e^{-\frac{1}{2}\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}(\Sigma_0^{-1}b_0+(\sigma^{-1})'y)\right)'\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}(\Sigma_0^{-1}b_0+(\sigma^{-1})'y)\right)}db}{\int_{\mathbb{R}^n} e^{-\frac{1}{2}\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}(\Sigma_0^{-1}b_0+(\sigma^{-1})'y)\right)'\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}(\Sigma_0^{-1}b_0+(\sigma^{-1})'y)\right)}db}.\end{aligned}$$

Recognising the form of a Gaussian distribution, we find

$$\mu_{t,y}(db) = \frac{e^{-\frac{1}{2}\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})'y\right)\right)'\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)\left(b-\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\left(\Sigma_0^{-1}b_0+(\sigma^{-1})'y\right)\right)}}{(2\pi)^{\frac{n}{2}}\left|\left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}\right|^{\frac{1}{2}}}db,$$

which shows that

$$\mu_{t,y} \sim \mathcal{N}\left(f(t,y),\Sigma(t,y)\right).$$

Next, from (3.32), we have

$$\nabla_y f(t,y) = \left(\Sigma_0^{-1}+\Sigma^{-1}t\right)^{-1}(\sigma^{-1})',$$

Left multiplying by $\Sigma_0\Sigma_0^{-1}$ and right multiplying by $\sigma'\Sigma^{-1}\sigma$ the previous equation, we obtain the multidimensional independent from b familiar expression of the diffusion coefficient ψ stated in Ekstrom and Vaicenavicius (2016) in dimension one:

$$\psi(t) = \Sigma_0\left(\Sigma+\Sigma_0 t\right)^{-1}\sigma.$$

For the computation of the Bayesian risk premium function R solution to (3.15), we look for a solution of the form:

$$R(t,b) = b'M(t)b+\mathcal{V}(t), \tag{3.33}$$

for some $\mathbb{R}^{n\times n}$-valued function M, and $\mathbb{R}^n$-valued function $\mathcal{V}$ defined on $[0,T]$, to be determined. Plugging into (3.15), we see that M and $\mathcal{V}$ should satisfy the first order ODE system:

$$\begin{cases} -M'(t)-2M(t)'G(t)'\Sigma G(t)M(t)+4G(t)M(t)-\Sigma^{-1} = 0 \\ -\mathcal{V}'(t)-\mathrm{tr}\left[G(t)'\Sigma G(t)M(t)\right] = 0, \end{cases} \tag{3.34}$$

with $G(t)=(\Sigma+\Sigma_0 t)^{-1}\Sigma_0$ and terminal conditions:

$$M(T) = 0, \qquad \mathcal{V}(T) = 0. \tag{3.35}$$

The solution to (3.34)-(3.35) is given by the following Lemma, whose proof is detailed in Appendix 3.A.5.

Lemma 3.4.5. *The solution to the ODE system* (3.34) *with terminal condition* (3.35) *is:*

$$\begin{aligned} M(t) &= \Sigma_0^{-1}+\Sigma^{-1}t-\Bigg[\left(\Sigma_0^{-1}+\Sigma^{-1}T\right)^{-1} \\ &\quad +2\int_t^T \Sigma_0\left(\Sigma+\Sigma_0 s\right)^{-1}\Sigma\left(\Sigma+\Sigma_0 s\right)^{-1}\Sigma_0 ds\Bigg]^{-1}, \\ \mathcal{V}(t) &= \int_t^T \mathrm{tr}\Bigg(\Sigma_0(\Sigma+\Sigma_0 s)^{-1}-\Bigg[\Bigg((\Sigma_0^{-1}+\Sigma^{-1}T)^{-1} \\ &\quad +2\int_s^T G(u)'\Sigma G(u)du\Bigg)(G(s)'\Sigma G(s))^{-1}\Bigg]^{-1}\Bigg)ds. \end{aligned}$$

We conclude that the Bayesian risk premium function R is explicitly given by (3.33) with M and $\mathcal{V}$ as in Lemma 3.4.5. The optimal Bayesian mean-variance strategy (3.18) is explicitly written as

$$\alpha_t^{*,\lambda} = \left(x_0 - X_t^{\alpha^{*,\lambda}} + \frac{e^{R(0,b_0)}}{2\lambda}\right)\left[\Sigma^{-1} - \left(\Sigma + \Sigma_0 t\right)^{-1}\Sigma_0 M(t)\right]\hat{B}_t$$

where $\hat{B}_t = f(t, Y_t) = f(t, h(t, S_t))$.

In the one-dimensional case $n = 1$, hence with $\Sigma = \sigma^2$, $\Sigma_0 = \sigma_0^2$, the expressions for M and $\mathcal{V}$ are simplified into

$$\begin{aligned} M(t) &= \frac{\sigma^2 + \sigma_0^2 t}{\sigma^2\left(\sigma^2 + \sigma_0^2(2T - t)\right)}(T - t), \\ \mathcal{V}(t) &= \ln\left(\frac{\sigma^2 + \sigma_0^2 T}{\sqrt{(\sigma^2 + \sigma_0^2 t)(\sigma^2 + \sigma_0^2(2T - t))}}\right). \end{aligned}$$

3.5 Impact of learning on the Markowitz strategy

This section shows the benefit of using a Bayesian learning approach to solve the Markowitz problem. In contrast to classical approaches that estimate unknown parameters in a second step, leading to suboptimal solutions, the Bayesian learning approach uses the updated data from the most recent prices of the assets to adjust the controls of the investment strategy and reaches optimality in the Markowitz sense.

In a framework where the drift B is unknown, following a prior probability distribution μ, we will compare the performance, in terms of Sharpe ratio, of the Bayesian learning strategy to the non-learning strategy. Recall, the non-learning strategy considers the drift constant set at $b_0 = \mathbb{E}[B] = \int b\mu(db)$. This allows us to exhibit the benefit of integrating the most recent available information into the strategy.

3.5.1 Computation of the Sharpe ratios

The Sharpe ratio associated to a portfolio strategy α is defined by

$$Sh_T^\alpha := \frac{\mathbb{E}[X_T^\alpha] - x_0}{\sqrt{\mathrm{Var}(X_T^\alpha)}}.$$

We denote by $\hat{\alpha}^{\vartheta,L}$ the optimal Bayesian-Markowitz strategy obtained in Theorem 3.4.1, $X^L = X^{\hat{\alpha}^{\vartheta,L}}$ the associated wealth process and Sh_T^L the corresponding Sharpe ratio. We also write $\hat{\alpha}^{\vartheta,NL}$ the non-learning Markowitz strategy based on a constant

drift parameter b_0 as described in Remark 3.2.5, $X^{NL} = X^{\hat{\alpha}^{\vartheta,NL}}$ the associated wealth process and Sh_T^{NL} the corresponding Sharpe ratio.

Proposition 3.5.1. *The Sharpe ratios of the learning and non-learning Markowitz strategies are explicitly given by*

$$Sh_T^L = \sqrt{e^{R(0,b_0)} - 1}, \tag{3.36}$$

$$Sh_T^{NL} = \frac{1 - \int_{\mathbb{R}^n} e^{-b_0'\Sigma^{-1}bT}\mu(db)}{\sqrt{\int_{\mathbb{R}^n} e^{-b_0'\Sigma^{-1}(2b-b_0)T}\mu(db) - \left(\int_{\mathbb{R}^n} e^{-b_0'\Sigma^{-1}bT}\mu(db)\right)^2}}. \tag{3.37}$$

Proof. **1.** We first focus on the Bayesian learning strategy. From (3.20), we have $\mathbb{E}[X^L] = x_0 + \frac{1}{2\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right)$ with $\lambda(\vartheta) = \sqrt{\frac{e^{R(0,b_0)}-1}{4\vartheta}}$ binding the variance of the optimal terminal wealth to ϑ. We thus obtain

$$Sh_T^L = \frac{\mathbb{E}[X^L] - x_0}{\sqrt{\vartheta}} = \frac{\sqrt{\vartheta\left(e^{R(0,b_0)} - 1\right)}}{\sqrt{\vartheta}} = \sqrt{e^{R(0,b_0)} - 1}.$$

2. Let us now consider the non-learning strategy $\hat{\alpha}^{\vartheta,NL}$ for which we need to compute its expectation and variance. From the expression of $\hat{\alpha}^{\vartheta,NL}$ given in Remark 3.2.5, and the self-financed equation of the wealth process (3.2), we deduce the dynamics of the conditional expectation of X^{NL} given B:

$$\begin{aligned} d\mathbb{E}[X_t^{NL}|B] &= \mathbb{E}\Big[\big(x_0 - X_t^{NL} + C_1\big)C_2'B|B\Big]dt \\ &= \Big[\left(x_0 + C_1\right)C_2'B - \mathbb{E}\big[X_t^{NL}|B\big]C_2'B\Big]dt, \end{aligned}$$

where we set $C_1 = \frac{e^{|\sigma^{-1}b_0|^2T}}{2\lambda_0(\vartheta)}$, $C_2 = \Sigma^{-1}b_0$ to alleviate notations. We thus obtain an ordinary differential equation for $t \mapsto \mathbb{E}[X_t^{NL}|B]$ with initial condition $\mathbb{E}[X_0^{NL}|B] = x_0$, whose solution is explicitly given by

$$\mathbb{E}[X_t^{NL}|B] = x_0 + C_1\big(1 - e^{-C_2'Bt}\big), \quad 0 \le t \le T. \tag{3.38}$$

Integrating with respect to the law of B, we obtain the expression of the expectation:

$$\mathbb{E}[X_t^{NL}] = x_0 + \frac{e^{|\sigma^{-1}b_0|^2T}}{2\lambda_0(\vartheta)}\left(1 - \int_{\mathbb{R}^n} e^{-b_0'\Sigma^{-1}bt}\mu(db)\right). \tag{3.39}$$

Next, let us compute the variance of X_T^{NL}. Applying Itô's formula to $|X^{NL}|^2$, and taking conditional expectation with respect to B, we have

$$\begin{aligned} d\mathbb{E}[|X_t^{NL}|^2|B] &= \Big[C_2'\left(\Sigma C_2 - 2B\right)\mathbb{E}[|X_t^{NL}|^2|B] - 2C_2'\left(x_0 + C_1\right) \\ &\quad \times \left(\Sigma C_2 - B\right)\mathbb{E}[X_t^{NL}|B] + C_2'\Sigma C_2\left(x_0 + C_1\right)^2\Big]dt. \end{aligned}$$

Replacing $\mathbb{E}[X_t^{NL}|B]$ by its value calculated in (3.38), we obtain an ODE in $\mathbb{E}[|X_t^{NL}|^2|B]$ that we explicitly solve:

$$\mathbb{E}[|X_t^{NL}|^2|B] = (x_0+C_1)^2 - 2C_1(x_0+C_1)e^{C_2'Bt} + C_1^2 e^{C_2'(\Sigma C_2 - 2B)t}.$$

By integrating over B, we obtain

$$\begin{aligned}\mathbb{E}[|X_t^{NL}|^2] &= (x_0+C_1)^2 - 2C_1(x_0+C_1)\int_{\mathbb{R}^n} e^{C_2'bt}\mu(db) \\ &\quad + C_1^2 e^{C_2'\Sigma C_2 t}\int_{\mathbb{R}^n} e^{-2C_2'bt}\mu(db).\end{aligned}$$

From $\mathrm{Var}(X_t^{NL}) = \mathbb{E}[|X_t^{NL}|^2] - \left(\mathbb{E}[X_t^{NL}]\right)^2$, we obtain the expression of the variance

$$\mathrm{Var}(X_t^{NL}) = \frac{e^{2|\sigma^{-1}b_0|^2 T}}{4\lambda_0(\vartheta)^2}\left[\int_{\mathbb{R}^n} e^{-b_0'\Sigma^{-1}(2b-b_0)t}\mu(db) - \left(\int_{\mathbb{R}^n} e^{-b_0'\Sigma^{-1}bt}\mu(db)\right)^2\right]. \tag{3.40}$$

Finally, coupling (3.39) and (3.40) we infer the value of Sh_T^{NL} in (3.37). □

Next we prove that $Sh_T^{NL} \leq Sh_T^L$. From (3.2) we know that X^{NL} is linear in its control $\hat{\alpha}^{\vartheta,NL}$, so we can always find a leveraged strategy $\tilde{\alpha}^{\vartheta,NL} = \delta\hat{\alpha}^{\vartheta,NL}$ with $\delta > 0$ such that $\mathrm{Var}(X_T^{\tilde{\alpha}^{\vartheta,NL}}) = \vartheta$, simply by taking $\delta = \sqrt{\frac{\vartheta}{\mathrm{Var}(X_T^{NL})}}$. Thus, by invariance of the Sharpe ratio with respect to the leverage we obtain,

$$\begin{aligned}Sh_T^L - Sh_T^{NL} &= \frac{\mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,L}}] - x_0}{\sqrt{\vartheta}} - \frac{\mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,NL}}] - x_0}{\sqrt{\mathrm{Var}(X_T^{\hat{\alpha}^{\vartheta,NL}})}} \\ &= \frac{\mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,L}}] - x_0}{\sqrt{\vartheta}} - \frac{\mathbb{E}[X_T^{\tilde{\alpha}^{\vartheta,NL}}] - x_0}{\sqrt{\vartheta}} \\ &= \mathbb{E}[X_T^{\hat{\alpha}^{\vartheta,L}}] - \mathbb{E}[X_T^{\tilde{\alpha}^{\vartheta,NL}}] \\ &\geq 0\end{aligned}$$

The last inequality comes from $\hat{\alpha}^{\vartheta,L}$ being the optimal control.

Remark 3.5.2. We have an upper-bound for the Sharpe ratio of the non-learning strategy:

$$Sh_T^{NL} \leq \sqrt{e^{|\sigma^{-1}b_0|^2 T} - 1}.$$

Indeed, by Jensen's inequality, we have

$$\begin{aligned}\left(\int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db)\right)^2 &\leq \int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db),\\ \int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db) &\geq e^{-T\int_{\mathbb{R}^n} b_0\Sigma^{-1}b\mu(db)} = e^{-|\sigma^{-1}b_0|^2T},\\ \int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db) &\geq e^{-2|\sigma^{-1}b_0|^2T}.\end{aligned}$$

From (3.37), we thus deduce

$$\begin{aligned}Sh_T^{NL} &\leq \frac{1-\int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db)}{\sqrt{\int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}(2b-b_0)T}\mu(db) - \int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db)}}\\ &= \frac{1-\int_{\mathbb{R}^n} e^{-b_0\Sigma^{-1}bT}\mu(db)}{\sqrt{(e^{|\sigma^{-1}b_0|^2T}-1)\int_{\mathbb{R}^n} e^{-2b_0\Sigma^{-1}bT}\mu(db)}}\\ &\leq \frac{1-e^{-|\sigma^{-1}b_0|^2T}}{\sqrt{(e^{|\sigma^{-1}b_0|^2T}-1)\,e^{-2|\sigma^{-1}b_0|^2T}}} = \sqrt{e^{|\sigma^{-1}b_0|^2T}-1}\end{aligned}$$

We notice that when $\mu = \mathbb{1}_{b_0}$ the Sharpe ratio of the non-learning strategy coincides with $\sqrt{e^{|\sigma^{-1}b_0|^2T}-1}$. ◊

3.5.2 Value of Information

We illustrate our results in the case where B follows a prior one-dimensional Gaussian distribution $\mathcal{N}(b_0, \sigma_0^2)$. To assess the value of information (VI), we consider the Sharpe ratio Sh_T^L and Sh_T^{NL} of the learning and non-learning strategies as computed in Proposition 3.5.1. We define the value of information as the difference $Sh_T^L - Sh_T^{NL}$, and measure its sensitivity with respect to various parameters. So, in the sequel, we denote VI(Asset i) the value of information obtained using the set of parameters corresponding to Asset i.

The following explicit formulas for the Sharpe ratios learning and non-learning, used to create the graphs of this section, are computed from (3.36) and (3.37) with $\mu \sim \mathcal{N}(b_0, \sigma_0^2)$.

$$Gaussian\_Sh_T^L = \sqrt{\frac{\sigma^2+\sigma_0^2T}{\sigma\sqrt{\sigma^2+2\sigma_0^2T}}e^{\frac{b_0^2T}{\sigma^2+2\sigma_0^2T}} - 1} \tag{3.41}$$

and

$$Gaussian\_Sh_T^{NL} = \frac{e^{\frac{b_0^2}{\sigma^2}T\left(1-\frac{\sigma_0^2}{2\sigma^2}T\right)} - 1}{\sqrt{e^{\frac{b_0^2}{\sigma^2}T\left(1+\frac{\sigma_0^2}{\sigma^2}T\right)} - 1}} \tag{3.42}$$

3.5.2.1 Standard deviation of the drift

Intuitively, the parameter σ_0 measures the confidence the investor puts in her estimate of b_0. The higher σ_0, the more confident the investor is about her estimate. It is an important parameter since it legitimates the use of a learning strategy.

To estimate the sensitivity of the value of information with respect to the standard deviation of the drift, we choose an investment horizon of $T = 1$ and a sample of three different assets with the same mean $b_0 = 5\%$ and different volatilities, $\sigma = \{5\%, 10\%, 20\%\}$ resulting in realistic Sharpe ratios of 1, 0.5 and 0.25 respectively. From Formulas (3.41) and (3.42) we compute the value of information according to the standard deviation of the drift ranging from 0 to 100%.

Figure 3.1 shows the value of information as a function of the standard deviation of the drift for the three assets described previously, which parameters are summed up in Table 3.1. As we can see on the graph, the value of information is a monotone increasing function of the standard deviation of the drift. When the standard deviation of the drift is zero, as in the Dirac case, the Sharpe ratios of the learning and non-learning strategies are equal. It is clear since a standard deviation of the drift equal to zero means the drift is simply constant, so updating it does not bring any additional value to the learning strategy. For all three assets, we notice that as σ_0 increases the value of information becomes more and more valuable depending on the level of the Sharpe ratio of each asset. The higher the Sharpe ratio, the bigger the value of information and its rate of increase. For instance, the value of information of the highest Sharpe ratio asset, Asset 1, increases rapidly to 1 when $\sigma_0 = 10\%$ and reaches rapidly 3.5 when $\sigma_0 = 100\%$ whereas the lowest Sharpe ratio asset, Asset 3, equals roughly 0.1 when $\sigma_0 = 0.1$ and nearly reaches 2 when $\sigma_0 = 100\%$ at a relatively slower pace.

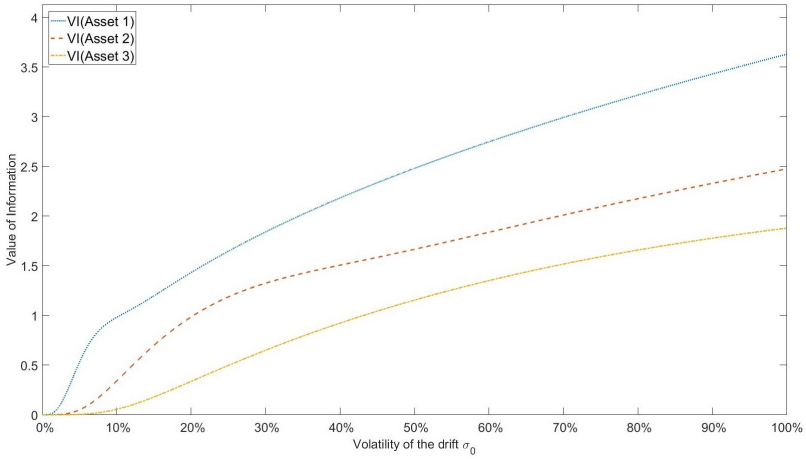


Figure 3.1: Value of information as a function of σ_0 with parameters as in Table 3.1.

Table 3.1: Parameter values used in Figure 3.1.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	**Sharpe ratio**	**T**	$\boldsymbol{\sigma_0}$
Asset 1	5%	5%	1	1	[0-100%]
Asset 2	5%	10%	0.5	1	[0-100%]
Asset 3	5%	20%	0.25	1	[0-100%]

An interesting fact happens in the case of an asset with a high Sharpe ratio. Figure 3.2 shows the value of information, the curves of the Sharpe ratio of the learning and the non-learning strategy based on Asset 4. This asset has a high Sharpe ratio of 2, a set of parameters as in Table 3.2 and a standard deviation of the drift ranging from 0 to 30%.

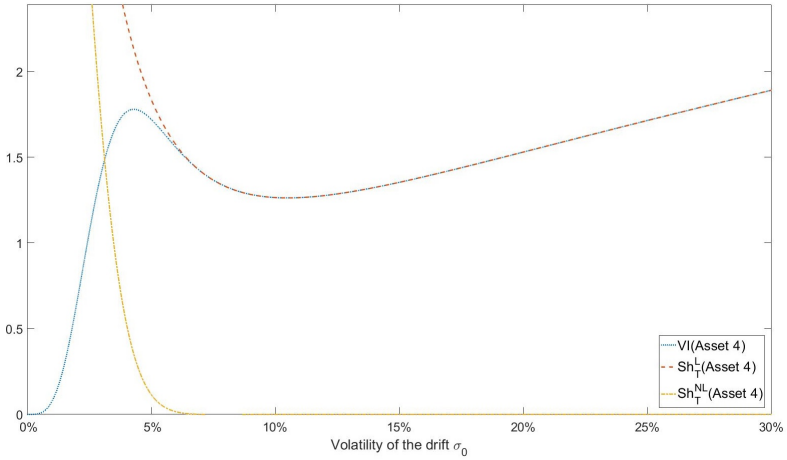


Figure 3.2: Value of information, Sh_T^L and Sh_T^{NL} of Asset 4 as a function of σ_0 with parameters as in Table 3.2.

Table 3.2: Parameter values used in Figure 3.2.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	**Sharpe ratio**	**T**	$\boldsymbol{\sigma_0}$
Asset 4	10%	5%	2	1	[0-30%]

As we can see on the graph in Figure 3.2, the value of information curve is no more monotonic and we observe a bump around $\sigma_0 = 5\%$. To explain this shape, we see that when σ_0 ranges between 0 and 5%, both Sharpe ratios decrease. It is understandable since the variance of the terminal wealth is an increasing function of σ and σ_0. Nevertheless, the decrease is much more rapid for the non-learning strategy reaching approximately zero when the drift volatility merely exceeds 5%, approximately the value of the volatility of the asset. The difference in the rate of decrease explains the bump we observe. Then, the Sharpe ratio of the learning strategy keeps decreasing and progressively recovers when the standard deviation of the drift approaches the double of the volatility of the asset, 10% . At this point, the learning effect really allows the strategy to exploit its advantage of updating the drift as the new information become available. As soon as $\sigma_0 > 2\sigma$, the signal coming from the standard deviation of the drift differentiates from the volatility of the asset and is captured by the learning strategy which allows the value of information to increase.

3.5.2.2 Sharpe ratio of the asset

Figure 3.3 exhibits the value of information as a function of the Sharpe ratio of the asset, defined as b_0/σ, for a sample of three assets with parameters set in Table 3.3. We see that the higher the Sharpe ratio of the asset the bigger the value of information for any level of σ_0. Obviously, the higher the standard deviation of the drift, the more value information has and the more necessary it is to update the strategy. We also notice that the slope of the curve is steeper when the Sharpe ratio of the asset is low, ranging roughly from 0 to 1. Intuitively, it shows that information has more value for assets with a low Sharpe ratio because assets with a high Sharpe ratio will perform well, no matter the learning.

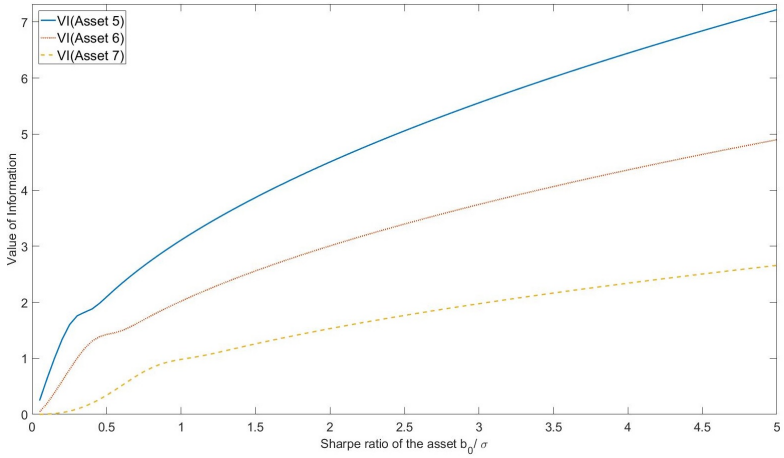


Figure 3.3: Value of information as a function of the Sharpe ratio of the asset with parameters as in Table 3.3.

Table 3.3: Parameter values used in Figure 3.3.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	**Sharpe ratio**	**T**	$\boldsymbol{\sigma_0}$
Asset 5	5%	[1-100]%	[0-5]	1	75%
Asset 6	5%	[1-100]%	[0-5]	1	35%
Asset 7	5%	[1-100]%	[0-5]	1	10%

Another explanation is for a fix b_0, a low Sharpe ratio means a high volatility which increases the need for learning and consequently the value of information. Furthermore, we clearly see that the curve with the higher σ_0 dominates the lower one, which confirms the intuition and understanding of Figure 3.1.

3.5.2.3 Time

Figure 3.4 displays the value of information along the life of the investment for time $t \in [0, T]$. To obtain this graph, we use the set of parameters shown in Table 3.4. We simulate $N = 1,000,000$ optimal wealth trajectories $(X^i)_{i\in[\![1,N]\!]}$ and at each point in time we compute the empirical Sharpe ratio of the learning strategy:

$$\hat{Sh}_t^L = \frac{\sqrt{N-1}}{N} \frac{\sum_{i=1}^N (X_t^i - x_0)}{\sqrt{\sum_{i=1}^N (X_t^i)^2 - \left(\sum_{i=1}^N X_t^i\right)^2}}$$

and use Formula (3.40) for the non-learning strategy.

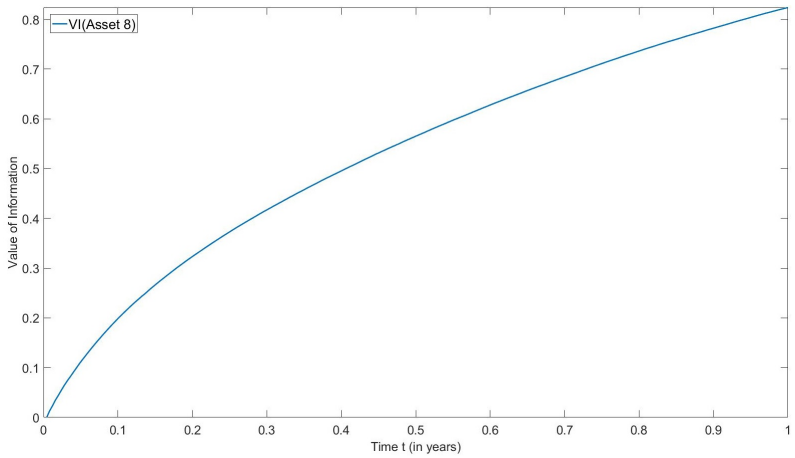


Figure 3.4: Value of information as a function of t with parameters as in Table 3.4.

Table 3.4: Parameter values used in Figure 3.4

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	$\boldsymbol{\sigma_0}$	**T**	$\boldsymbol{\vartheta}$	**Simulations**
Asset 8	5%	20%	40%	1	$(10\%)^2$	1,000,000

As Figure 3.4 shows, the value of information increases monotonically with time. Nonetheless, the speed of increase tends to slow as time goes by. The fact that the marginal gain on the value of information decreases with time is well known and analyzed in the recent article by Keppo et al. (2018) in the context of investment decisions and costs of data analytics.

3.5.2.4 Investment horizon

Figure 3.5 shows the value of information with respect to the investment time horizon T for a sample of three assets with parameters described in Table 3.5.

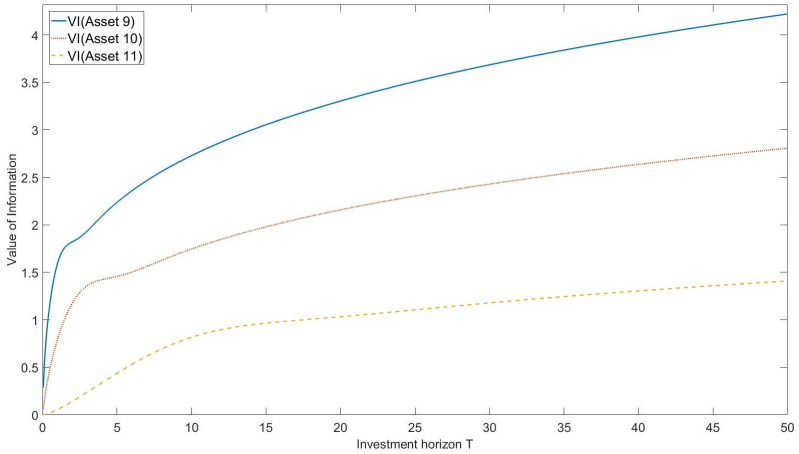


Figure 3.5: Value of information as a function of T with parameters as in Table 3.5.

Table 3.5: Parameter values used in Figure 3.5.

	$\mathbf{b_0}$	$\boldsymbol{\sigma}$	Sharpe ratio	T	$\boldsymbol{\sigma_0}$
Asset 9	5%	20%	25%	[0-50]	75%
Asset 10	5%	20%	25%	[0-50]	35%
Asset 11	5%	20%	25%	[0-50]	10%

Although the three curves are increasing no matter the value of the other parameters, the slope of these curves are mainly decreasing with the investment horizon. It suggests that the investor should consider that the marginal contribution to the value of information decreases with the investment horizon. It means that the spread between the optimal, Bayesian-Markowitz, and a suboptimal, non-learning, strategy tightens for long investment horizons. It makes sense since for long horizons, at some point, the posterior distribution will not move so much since the values of the parameters of the distribution will be marginally affected by more learning. In addition, we remark that the bigger the drift volatility parameter, the higher the level of the associated value of information curve. This confirms the analysis of Figure 3.1 and 3.3.

3.6 Conclusion

In this chapter, we have solved the Bayesian-Markowitz problem when the unknown drift is assumed to follow a probability distribution corresponding to a prior belief. The investor then updates her information on the drift from a predictive distribution based on observed data coming from assets prices. We have turned the non-standard

problem into a standard one to exhibit the HJB equation associated to the standardized problem and apply dynamic programming techniques. To illustrate our theoretical results, we have computed the key diffusion coefficient ψ in the case of the multidimensional discrete law and the Gaussian law, and provided the full solution to the problem in the multidimensional Gaussian case. We have described a way of measuring the performance of investment strategies, the Sharpe ratio of terminal wealth, and used it to assess the value of information. To exhibit the value added of implementing a Bayesian learning approach in our problem, we have illustrated the value of information sensitivity to various key parameters and concluded that learning brings value to the optimal strategy that solves the Markowitz problem in a framework of drift uncertainty, modeled by a prior distribution with a positive definite covariance matrix, and a constant known volatility.

3.A Appendices

3.A.1 Proof of Lemma 3.2.3

Proof. We use similar arguments as in Ismail and Pham (2017).

Fix $\vartheta > 0$, for any $\epsilon > 0$, there exists an ϵ-optimal control for $U_0(\vartheta)$, $\alpha^\epsilon \in \mathcal{A}$ such that

$$U_0(\vartheta) \leq \mathbb{E}[X_T^{\alpha^\epsilon}] + \epsilon \text{ and } \mathrm{Var}(X_T^{\alpha^\epsilon}) \leq \vartheta.$$

Then for any $\lambda > 0$,

$$V_0(\lambda) \leq \lambda \mathrm{Var}(X_T^{\alpha^\epsilon}) - \mathbb{E}[X_T^{\alpha^\epsilon}] \leq \lambda\vartheta - U_0(\vartheta) + \epsilon.$$

Since ϵ is arbitrary and the above relation holds for any ϑ we deduce,

$$V_0(\lambda) \leq \inf_{\vartheta>0}[\lambda\vartheta - U_0(\vartheta)], \quad \forall\lambda > 0.$$

Conversely, fix $\lambda > 0$ and $\epsilon > 0$ and consider the ϵ-optimal control $\alpha^{*,\lambda,\epsilon} \in \mathcal{A}$ for V_0 such that

$$V_0(\lambda) \geq \lambda \mathrm{Var}(X_T^{\alpha^{*,\lambda,\epsilon}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda,\epsilon}}] - \epsilon.$$

If we define

$$\vartheta^{\lambda,\epsilon} := \mathrm{Var}(X_T^{\alpha^{*,\lambda,\epsilon}}),$$

by definition of $U_0(\vartheta^{\lambda,\epsilon})$ the inequality $\mathbb{E}[X_T^{\alpha^{*,\lambda,\epsilon}}] \leq U_0(\vartheta^{\lambda,\epsilon})$ holds true.
Hence,

$$V_0(\lambda) \geq \lambda\vartheta^{\lambda,\epsilon} - \mathbb{E}[X_T^{\alpha^{*,\lambda,\epsilon}}] - \epsilon \geq \lambda\vartheta^{\lambda,\epsilon} - U_0(\vartheta^{\lambda,\epsilon}) - \epsilon$$

which holds for arbitrary $\epsilon > 0$. This shows the first duality relation, i.e. V_0 is the Fenchel Legendre transform of U_0, and ϑ^λ attains the infimum, namely

$$V_0(\lambda) = \inf_{\vartheta>0}[\lambda\vartheta - U_0(\vartheta)] = \lambda\vartheta^\lambda - U_0(\vartheta^\lambda).$$

We now prove the second duality relation. Fix $\vartheta > 0$ and for any $\lambda > 0$ consider the optimal control $\hat{\alpha}$ for U_0, we then have

$$\lambda\vartheta - U_0(\vartheta) \geq \lambda \mathrm{Var}(X_T^{\hat{\alpha}}) - \mathbb{E}[X_T^{\hat{\alpha}}] \geq V_0(\lambda).$$

The previous relation holds for any $\lambda > 0$, thus

$$U_0(\vartheta) \leq \inf_{\lambda>0}[\lambda\vartheta - V_0(\lambda)]. \tag{3.43}$$

Now fix $\vartheta > 0$ and consider $\alpha^{*,\lambda}$ the optimal control for V_0, we have

$$V_0(\lambda) = \lambda \mathrm{Var}(X_T^{\alpha^{*,\lambda}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda}}].$$

Choose $\lambda(\vartheta)$ such that $\mathrm{Var}(X_T^{\alpha^{*,\lambda(\vartheta)}}) = \vartheta$. We then have

$$\begin{aligned} V_0(\lambda(\vartheta)) &= \lambda(\vartheta)\mathrm{Var}(X_T^{\alpha^{*,\lambda(\vartheta)}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda(\vartheta)}}] \\ &= \lambda(\vartheta)\vartheta - \mathbb{E}[X_T^{\alpha^{*,\lambda(\vartheta)}}] \\ &\geq \lambda(\vartheta)\vartheta - U_0(\vartheta) \end{aligned}$$

by definition of U_0. Together with 3.43, this shows the second duality relation; U_0 is the Fenchel-Legendre transform of V_0 and $\lambda(\vartheta)$ attains the infimum in this transform

$$U_0(\vartheta) = \inf_{\lambda>0}[\lambda\vartheta - V_0(\lambda)] = \lambda(\vartheta)\vartheta - V_0(\lambda(\vartheta)).$$

Finally, for $\lambda > 0$, if $\alpha^{*,\lambda} \in \mathcal{A}$ is the optimal control for V_0 i.e. for Problem (3.4) then

$$V_0(\lambda) = \lambda \mathrm{Var}(X_T^{\alpha^{*,\lambda}}) - \mathbb{E}[X_T^{\alpha^{*,\lambda}}] = \lambda\vartheta^\lambda - U_0(\vartheta^\lambda)$$

with $\vartheta^\lambda := \mathrm{Var}(X_T^{\alpha^{*,\lambda}})$. Moreover, we have

$$\begin{aligned} U_0(\vartheta) = \inf_{\lambda>0}[\lambda\vartheta - V_0(\lambda)] &= \lambda(\vartheta)\vartheta - V_0(\lambda(\vartheta)) \\ &= \lambda(\vartheta)\vartheta - \lambda(\vartheta)\vartheta^{\lambda(\vartheta)} + \mathbb{E}[X_T^{\alpha^{*,\lambda(\vartheta)}}] \\ &= \mathbb{E}[X_T^{\hat{\alpha}^\vartheta}] \end{aligned}$$

since $\vartheta^{\lambda(\vartheta)} = \vartheta$, by definition of $\lambda(\vartheta)$. This means that $\hat{\alpha}^\vartheta$ is an optimal control for U_0 i.e. Problem (3.3). □

3.A.2 Proof of Lemma 3.2.4

Proof. We use a similar approach as in Zhou and Li (2000). We introduce the function $\tilde{f}$: $\gamma \in \mathbb{R} \mapsto \gamma^2 - 2\gamma\mathbb{E}[X_T^\alpha]$. $\tilde{f}$ is strictly concave in γ and reaches its infimum $-\mathbb{E}[X_T^\alpha]^2$ at a unique point $\gamma = \mathbb{E}[X_T^\alpha]$. Therefore, for all $\lambda > 0$, we can write

$$\begin{aligned} J^\lambda(\alpha) &:= \lambda \mathrm{Var}(X_T^\alpha) - \mathbb{E}[X_T^\alpha] \\ &= \lambda\big(\mathbb{E}[(X_T^\alpha)^2] - \mathbb{E}[X_T^\alpha]^2\big) - \mathbb{E}[X_T^\alpha] \\ &= \lambda\big(\mathbb{E}[(X_T^\alpha)^2] + \inf_{\gamma\in\mathbb{R}} \tilde{f}(\gamma)\big) - \mathbb{E}[X_T^\alpha]. \end{aligned}$$

Noticing that for all $\lambda > 0$, $V_0(\lambda) = \inf_{\alpha\in\mathcal{A}} J^\lambda(\alpha)$, we can then deduce that,

$$\begin{aligned} V_0(\lambda) &= \inf_{\alpha\in\mathcal{A}} \left\{ \lambda\big(\mathbb{E}[(X_T^\alpha)^2] + \inf_{\gamma\in\mathbb{R}} \tilde{f}(\gamma)\big) - \mathbb{E}[X_T^\alpha] \right\} \\ &= \inf_{\alpha\in\mathcal{A}} \left\{ \inf_{\gamma\in\mathbb{R}} \left\{ \lambda\mathbb{E}[(X_T^\alpha)^2] + \lambda\gamma^2 - (1+2\lambda\gamma)\mathbb{E}[X_T^\alpha] \right\} \right\} \\ &= \inf_{\gamma\in\mathbb{R}} \left\{ \tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2 \right\}, \end{aligned}$$

where

$$\tilde{V}_0(\lambda,\gamma) := \inf_{\alpha\in\mathcal{A}} \left\{ \lambda\mathbb{E}[(X_T^\alpha)^2] - (1+2\lambda\gamma)\mathbb{E}[X_T^\alpha] \right\}. \tag{3.44}$$

Now, the linear-quadratic function $\gamma \in \mathbb{R} \mapsto \tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2$ is strictly convex and admits an infimum at a unique point $\gamma^*(\lambda)$, noted γ^* here to alleviate the notation.

Consequently, assuming there exists $\tilde{\alpha}^{\lambda,\gamma} \in \mathcal{A}$ solution of Problem (3.44) and noting $\alpha^{*,\lambda} := \tilde{\alpha}^{\lambda,\gamma^*}$, we have

$$\begin{aligned} V_0(\lambda) &= \tilde{V}_0(\lambda,\gamma^*) + \lambda(\gamma^*)^2 \\ &= \lambda\left[\mathbb{E}[(X_T^{\alpha^{*,\lambda}})^2] + (\gamma^*)^2 - 2\gamma^*\mathbb{E}[X_T^{\alpha^{*,\lambda}}]\right] - E[X_T^{\alpha^{*,\lambda}}] \\ &\geq J^\lambda(\alpha^{*,\lambda}), \end{aligned}$$

which shows that $V_0(\lambda) = J^\lambda(\alpha^{*,\lambda})$ meaning $\alpha^{*,\lambda}$ is solution of V_0. Additionally, differentiating the strictly convex function $\gamma \in \mathbb{R} \mapsto \lambda\mathbb{E}[(X_T^{\alpha^{*,\lambda}})^2] - (1+2\lambda\gamma)\mathbb{E}[X_T^{\alpha^{*,\lambda}}] + \lambda\gamma^2$, we find that

$$\gamma^* = \mathbb{E}[X_T^{\alpha^{*,\lambda}}].$$

□

3.A.3 Proof of Theorem 3.4.1

Proof. Following the standard dynamic programming approach as in Pham (2009), we derive the HJB equation for the standard control Problem (3.14):

$$\begin{cases} \partial_t v^{\lambda,\gamma} + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 v^{\lambda,\gamma}\right) \\ \quad + \inf_{\alpha\in\mathcal{A}}\left\{\alpha_t' b\partial_x v^{\lambda,\gamma} + \frac{1}{2}|\sigma'\alpha_t|^2\partial_{xx}^2 v^{\lambda,\gamma} + \alpha_t'\sigma\psi'\partial_{xb}^2 v^{\lambda,\gamma}\right\} = 0, \\ v^{\lambda,\gamma}(T,x,b) = \lambda x^2 - (1+2\lambda\gamma)x, \end{cases} \tag{3.45}$$

where the matrix function $\psi(t,b)$ is noted ψ to alleviate notations.

Assuming for now that $\partial_{xx}^2 v^{\lambda,\gamma} \geq 0$, we find

$$\begin{aligned} &\operatorname*{argmin}_{\alpha\in\mathcal{A}}\left\{\alpha_t' b\partial_x v^{\lambda,\gamma} + \frac{1}{2}|\sigma'\alpha_t|^2\partial_{xx}^2 v^{\lambda,\gamma} + \alpha_t'\sigma\psi'\partial_{xb}^2 v^{\lambda,\gamma}\right\} \\ &\quad = -\Sigma^{-1}b\frac{\partial_x v^{\lambda,\gamma}}{\partial_{xx}^2 v^{\lambda,\gamma}} - (\psi\sigma^{-1})'\frac{\partial_{xb}^2 v^{\lambda,\gamma}}{\partial_{xx}^2 v^{\lambda,\gamma}} \end{aligned} \tag{3.46}$$

which turns the standard control Problem (3.45) into the same following problem

$$\begin{cases} \partial_t v^{\lambda,\gamma} + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 v^{\lambda,\gamma}\right) - \frac{1}{2\partial_{xx}^2 v^{\lambda,\gamma}}|\partial_x v^{\lambda,\gamma}\sigma^{-1}b + \psi'\partial_{xb}^2 v^{\lambda,\gamma}|^2 = 0 \\ v^{\lambda,\gamma}(T,x,b) = \lambda x^2 - (1+2\lambda\gamma)x. \end{cases} \tag{3.47}$$

We look for a solution in the ansatz form $v^{\lambda,\gamma}(t,x,b) = K(t,b)x^2 + \Gamma(t,b)x + \chi(t,b)$ with $K \geq 0$ which ensures $\partial_{xx}^2 v^{\lambda,\gamma} \geq 0$ and (3.46). Formally we derive,

$$\begin{aligned} \partial_t v^{\lambda,\gamma} &= \partial_t K x^2 + \partial_t\Gamma x + \partial_t\chi, & \partial_x v^{\lambda,\gamma} &= 2Kx + \Gamma, \\ \partial_{xx}^2 v^{\lambda,\gamma} &= 2K, & \nabla_b v^{\lambda,\gamma} &= \nabla_b K x^2 + \nabla_b\Gamma x + \nabla_b\chi, \\ \mathcal{D}_b^2 v^{\lambda,\gamma} &= \mathcal{D}_b^2 K x^2 + \mathcal{D}_b^2\Gamma x + \mathcal{D}_b^2\chi, & \partial_{xb}^2 v^{\lambda,\gamma} &= 2\nabla_b K x + \nabla_b\Gamma. \end{aligned}$$

Plugging the previous partial derivatives into (3.47), the problem becomes:

$$\begin{cases} 0 = \left(\partial_t K + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 K\right) - \left(2(\psi\sigma^{-1}b)'\nabla_b K + \frac{|\psi'\nabla_b K|^2}{K} + |\sigma^{-1}b|^2 K\right)\right)x^2 \\ \quad + \left(\partial_t\Gamma + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2\Gamma\right) - \left(\left((\psi\sigma^{-1}b)' + \frac{(\psi\psi'\nabla_b K)'}{K}\right)\nabla_b\Gamma \right.\right. \\ \quad \left.\left. + \left(|\sigma^{-1}b|^2 + \left(\psi\sigma^{-1}b\right)'\frac{\nabla_b K}{K}\right)\Gamma\right)\right)x \\ \quad + \partial_t\chi + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2\chi\right) - \frac{1}{4K}|\sigma^{-1}b\Gamma + \psi'\nabla_b\Gamma|^2, \\ v(T,x,b) = \lambda x^2 - (1+2\lambda\gamma)x, \end{cases}$$

and by identification we obtain the following system of PDEs:

$$\begin{cases} 0 = \partial_t K + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 K\right) - \frac{1}{K}|\sigma^{-1}bK + \psi'\nabla_b K|^2, \\ 0 = \partial_t \Gamma + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 \Gamma\right) - \left(\left((\psi\sigma^{-1}b)' + \frac{(\psi\psi'\nabla_b K)'}{K}\right)\nabla_b\Gamma \right. \\ \qquad \left. + \left(|\sigma^{-1}b|^2 + \left(\psi\sigma^{-1}b\right)'\frac{\nabla_b K}{K}\right)\Gamma\right), \\ 0 = \partial_t \chi + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 \chi\right) - \frac{1}{4K}|\sigma^{-1}b\Gamma + \psi'\nabla_b\Gamma|^2, \end{cases}$$

with terminal conditions

$$\begin{cases} K(T,b) = \lambda, \\ \Gamma(T,b) = -(1+2\lambda\gamma), \\ \chi(T,b) = 0. \end{cases}$$

We now introduce the functions $K = \lambda\tilde{K}$, $\Gamma = -(1+2\lambda\gamma)\tilde{\Gamma}$ and $\chi = (1+2\lambda\gamma)^2\tilde{\chi}/\lambda$, so that the previous system becomes

$$\begin{cases} 0 = \partial_t \tilde{K} + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 \tilde{K}\right) - \frac{1}{\tilde{K}}|\sigma^{-1}b\tilde{K} + \psi'\nabla_b\tilde{K}|^2, \\ 0 = \partial_t \tilde{\Gamma} + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 \tilde{\Gamma}\right) - \left(\left(\left(\psi\sigma^{-1}b\right)' + \frac{(\psi\psi'\nabla_b\tilde{K})'}{\tilde{K}}\right)\nabla_b\tilde{\Gamma} \right. \\ \qquad \left. + \left(|\sigma^{-1}b|^2 + (\psi\sigma^{-1}b)'\frac{\nabla_b\tilde{K}}{\tilde{K}}\right)\tilde{\Gamma}\right), \\ 0 = \partial_t \tilde{\chi} + \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 \tilde{\chi}\right) - \frac{1}{4\tilde{K}}|\sigma^{-1}b\tilde{\Gamma} + \psi'\nabla_b\tilde{\Gamma}|^2, \end{cases}$$

with terminal conditions

$$\begin{cases} \tilde{K}(T,b) = 1, \\ \tilde{\Gamma}(T,b) = 1, \\ \tilde{\chi}(T,b) = 0. \end{cases}$$

If there exists a unique solution of the equation for $\tilde{K}$, then $\tilde{\Gamma} = \tilde{K}$ is the unique solution that verifies the linear PDE for $\tilde{\Gamma}$ and $\tilde{\chi} = (\tilde{K}-1)/4$ is the unique solution of the linear PDE for $\tilde{\chi}$.

Since we look for positive $\tilde{K}$, we introduce $\tilde{K}(t,b) := e^{-R(t,b)}$ and rewrite the related PDE:

$$\begin{cases} 0 = \tilde{K}\big(- \partial_t R - \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 R - \psi\psi'\nabla_b R \nabla_b R'\right) \\ \qquad + 2\left(\psi\sigma^{-1}b\right)'\nabla_b R - |\psi'(\nabla_b R)|^2 - |\sigma^{-1}b|^2\big), \\ 0 = R(T,b). \end{cases}$$

From $\mathrm{tr}\left(\psi\psi'\nabla_b R\nabla_b R'\right) = |\psi'\nabla_b R|^2$ we finally obtain the expression of the semi-linear PDE in (3.15):

$$\begin{cases} 0 = -\partial_t R - \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 R\right) + 2\left(\psi\sigma^{-1}b\right)'\nabla_b R - \frac{1}{2}|\psi'(\nabla_b R)|^2 - |\sigma^{-1}b|^2, \\ 0 = R(T,b). \end{cases}$$

We obtain the following value function dependent upon (λ,γ):

$$\begin{aligned} v^{\lambda,\gamma}(t,x,b) =& \lambda\tilde{K}(t,b)x^2 - (1+2\lambda\gamma)\tilde{K}(t,b)x + \frac{(1+2\lambda\gamma)^2}{\lambda}\frac{\tilde{K}(t,b)-1}{4} \\ =& e^{-R(t,b)}\left(\lambda x^2 - (1+2\lambda\gamma)x + \frac{(1+2\lambda\gamma)^2}{4\lambda}\right) - \frac{(1+2\lambda\gamma)^2}{4\lambda}, \end{aligned} \tag{3.48}$$

and from Equation (3.46), the optimal feedback control $\tilde{\alpha}^{\lambda,\gamma}$ is:

$$\tilde{\alpha}^{\lambda,\gamma} = \tilde{a}^{\lambda,\gamma}(t, X_t^{\tilde{\alpha}^{\lambda,\gamma}}, b),$$

with

$$\tilde{a}^{\lambda,\gamma}(t,x,b) = \left(\frac{1}{2\lambda} + \gamma - x\right)\left(\Sigma^{-1}b - (\psi\sigma^{-1})'\nabla_b R(t,b)\right).$$

Moreover, from $\tilde{V}_0(\lambda,\gamma) = v^{\lambda,\gamma}(0,x_0,b_0)$ we compute

$$\begin{aligned} \gamma^*(\lambda) =& \underset{\gamma\in\mathbb{R}}{\mathrm{argmin}}\left[\tilde{V}_0(\lambda,\gamma) + \lambda\gamma^2\right] \\ =& \underset{\gamma\in\mathbb{R}}{\mathrm{argmin}}\left[e^{-R(0,b_0)}\left(\lambda x_0^2 - (1+2\lambda\gamma)x_0 + \frac{(1+2\lambda\gamma)^2}{4\lambda}\right) - \frac{(1+2\lambda\gamma)^2}{4\lambda} + \lambda\gamma^2\right] \\ =& \underset{\gamma\in\mathbb{R}}{\mathrm{argmin}}\Bigg[e^{-R(0,b_0)}\left(\lambda x_0^2 - x_0 + \frac{1}{4\lambda}\right) - \frac{1}{4\lambda} + \left((1-2\lambda x_0)e^{-R(0,b_0)} - 1\right)\gamma \\ & \qquad + \lambda e^{-R(0,b_0)}\gamma^2\Bigg] \\ =& x_0 + \frac{1}{2\lambda}\left(e^{R(0,b_0)} - 1\right). \end{aligned}$$

Remembering (3.48), the value function at time 0 with the optimal γ is computed as followed:

$$\begin{aligned} v^{\lambda,\gamma^*}(0,x_0,b_0) =& e^{-R(0,b_0)}\left(\lambda x_0^2 - (2\lambda x_0 + e^{R(0,b_0)}x_0)\right) \\ &+ \left(\lambda x_0^2 + \frac{e^{2R(0,b_0)}}{4\lambda} + x_0 e^{R(0,b_0)}\right)\left(e^{-R(0,b_0)} - 1\right) \\ =& -\lambda x_0^2 + \frac{e^{R(0,b_0)}}{4\lambda}\left(1 - e^{R(0,b_0)}\right) - x_0 e^{R(0,b_0)}, \end{aligned}$$

and using Equation (3.5), we find the expression in (3.26):

$$\begin{aligned} V_0(\lambda) =& \tilde{V}_0(\lambda,\gamma^*) + \lambda(\gamma^*)^2 = v^{\lambda,\gamma^*}(0,x_0,b_0) + \lambda(\gamma^*)^2 \\ =& \frac{e^{R(0,b_0)}}{4\lambda}\left(1 - e^{R(0,b_0)}\right) - x_0 e^{R(0,b_0)} + \lambda x_0^2 + \frac{e^{2R(0,b_0)}}{4\lambda} \\ &+ \frac{1}{4\lambda} + x_0 e^{R(0,b_0)} - x_0 - \frac{e^{R(0,b_0)}}{2\lambda} \\ =& -\frac{1}{4\lambda}\left(e^{R(0,b_0)} - 1\right) - x_0. \end{aligned}$$

From Lemma (3.2.4), we find the optimal terminal wealth in (3.20)

$$\mathbb{E}[X_T^{\alpha^*,\lambda}] = \gamma^*(\lambda) = x_0 + \frac{1}{2\lambda}\left(e^{R(0,b_0)} - 1\right),$$

obtained with the optimal feedback control:

$$\alpha_t^{*,\lambda} = \tilde{a}^{\lambda,\gamma^*(\lambda)}(t, X_t^{\alpha^*,\lambda}, \hat{B}_t) = a_0^{Bayes,\lambda}(t, X_t^{\alpha^*,\lambda}, \hat{B}_t),$$

with

$$a_0^{Bayes,\lambda}(t,x,b) = \left(x_0 - x + \frac{e^{R(0,b_0)}}{2\lambda}\right)\left(\Sigma^{-1}b - (\psi\sigma^{-1})'\nabla_b R(t,b)\right). \tag{3.49}$$

The Lagrange multiplier $\lambda(\vartheta)$ which makes the variance of the optimal wealth equals to the variance constraint ϑ is calculated as follows. We know that,

$$V_0(\lambda(\vartheta)) = -\frac{1}{4\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right) - x_0, \tag{3.50}$$

and from Equation (3.4), we explicitly compute

$$\begin{aligned} V_0(\lambda(\vartheta)) &= \lambda(\vartheta)\mathrm{Var}(X_T^{\hat{\alpha}^\vartheta}) - \mathbb{E}[X_T^{\hat{\alpha}^\vartheta}] \\ &= \lambda(\vartheta)\vartheta - \mathbb{E}[X_T^{\hat{\alpha}^\vartheta}] \\ &= \lambda(\vartheta)\vartheta - x_0 - \frac{1}{2\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right). \end{aligned} \tag{3.51}$$

From Equation (3.50) and (3.51) we obtain (3.21), namely

$$\lambda(\vartheta) = \sqrt{\frac{1}{4\vartheta}\,(e^{R(0,b_0)} - 1)}. \tag{3.52}$$

From the correspondance of the controls established in Lemma 3.2.3, we know that

$$\hat{\alpha}^{\vartheta} = \alpha^{*,\lambda(\vartheta)}.$$

We then compute the optimal controls of Bayesian-Markowitz Problem U_0 as

$$\hat{\alpha}^{\vartheta} = a_0^{Bayes,\lambda(\vartheta)}(t, X_t^{\hat{\alpha}^{\vartheta}}, \hat{B}_t),$$

where, from Equation (3.49) with $\lambda = \lambda(\vartheta)$ as in (3.52),

$$a_0^{Bayes,\lambda(\vartheta)}(t,x,b) \;=\; \left(x_0 - x + e^{R(0,b_0)}\sqrt{\frac{\vartheta}{e^{R(0,b_0)}-1}}\right)\left(\Sigma^{-1}b - (\psi\sigma^{-1})'\nabla_b R(t,b)\right).$$

Finally from (3.51) and (3.52), we deduce the optimal performance of the Bayesian-Markowitz problem:

$$\begin{aligned} U_0(\vartheta) \;&=\; \lambda(\vartheta)\vartheta - V_0(\lambda(\vartheta)) \\ &=\; x_0 + \frac{1}{2\lambda(\vartheta)}\left(e^{R(0,b_0)} - 1\right) \;\;=\; x_0 + \sqrt{\vartheta(e^{R(0,b_0)} - 1)}. \end{aligned}$$

□

3.A.4 Proofs of Theorem 3.4.3

Proof. Without loss of generality, we consider in this demonstration the function $\tilde{R} := -R$. The value of the positive constant C may change from line to line and the matrix function $\psi(t,b)$ will be simply noted ψ to alleviate notations.

We define the nonlinear function $F : \mathcal{R} \times \mathbb{R}^n \to \mathbb{R}$ by:

$$F(t,b,p) \;=\; \frac{1}{2}|\psi' p|^2 + 2(\psi\sigma^{-1}b)'p + |\sigma^{-1}b|^2,$$

and we introduce the function $L_t : \mathcal{B}_t \times \mathbb{R}^n \to \mathbb{R}$ by

$$L_t(b,q) \;:=\; \max_{p\in\mathbb{R}^n}\big[-q'\psi' p - F(t,b,p)\big].$$

We notice that F is quadratic hence convex in p, so it is easy to see that for fixed $(t,b) \in \mathcal{R}$ the function $p \mapsto -q'\psi' p - F(t,b,p)$ reaches a maximum $p^* = -(\psi')^{-1}(q+$

$2\sigma^{-1}b$). This gives us the explicit form of the function L_t which depends on t only through $\mathcal{B}_t$:

$$\begin{aligned} L_t(b,q) &= -q'\psi' p^* - F(t,b,p^*) \\ &= \frac{1}{2}|q|^2 + 2(\sigma^{-1}b)'q + |\sigma^{-1}b|^2. \end{aligned}$$

Conversely, the function L_t is convex in q so for fixed $b \in \mathcal{B}_t$ the function $q \mapsto -q'\psi' p - L_t(b,q)$ reaches its maximum for $q^* = -(\psi' p + 2\sigma^{-1}b)$. It shows that

$$\begin{aligned} \max_{q\in\mathbb{R}^n}\big[-q'\psi' p - L_t(b,q)\big] &= -(q^*)'\psi' p - L_t(b,q^*) \\ &= \frac{1}{2}|\psi' p|^2 + 2(\psi\sigma^{-1}b)'p + |\sigma^{-1}b|^2 \\ &= F(t,b,p), \end{aligned}$$

and it establishes the duality relation between the functions F and L. Let us know consider the truncated function $F^k : \mathcal{R} \times \mathbb{R}^n \to \mathbb{R}$ defined for each $k \in \mathbb{N}$ by

$$F^k(t,b,p) = \max_{q\in\mathcal{A}_k}\big[-q'\psi' p - L_t(b,q)\big].$$

We observe from the explicit form of L_t that $L_t \in C^1(\mathcal{B}_t \times \mathbb{R}^n)$ and L, $\nabla_b L$ satisfy a polynomial growth condition in b. Namely, we see that the following estimates hold true for $(b,q) \in \mathcal{B}_t \times \mathbb{R}^n$:

$$\begin{aligned} |L_t(b,q)| &\le \frac{1}{2}|q|^2 + 2|\sigma^{-1}||b||q| + |\sigma^{-1}|^2|b|^2 \\ &\le \frac{1}{2}|q|^2 + |\sigma^{-1}|(|q|^2 + |b|^2) + |\sigma^{-1}|^2|b|^2 \\ &\le C(|q|^2 + |b|^2) \end{aligned}$$

and

$$\begin{aligned} |\nabla_b L_t(b,q)| &\le 2(|(\sigma^{-1})||q| + |\Sigma^{-1}||b|) \\ &\le C(|q| + |b|). \end{aligned}$$

So, $|L|$ and $|\nabla_b L|$ are of polynomial growth in b uniformly in $|q|$ when $|q| \le k$. By classical theory (see Theorem 4.3 p 163 in Fleming and Soner (2006)), the previous estimates and the assumptions of the theorem tell us that there exists a unique quadratically growing smooth solution

$$\tilde{R}^k \in C^{1,2}([0,T) \times \mathcal{B}_t) \cap C(\mathcal{R}),$$

to the truncated semi-linear PDE,

$$-\partial_t \tilde{R}^k - \frac{1}{2}\mathrm{tr}(\psi\psi'\mathcal{D}_b^2\tilde{R}^k) + F^k(t,b,\nabla_b\tilde{R}^k) \;=\; 0,$$

with terminal condition $\tilde{R}^k(T,.) = 0$.

We then know from the Feynman-Kac formula and standard arguments that $\tilde{R}^k$ can be represented as the solution of the stochastic control problem

$$\tilde{R}^k(t,b) = \inf_{q\in\mathcal{A}_k} \mathbb{E}\left[\int_t^{'} L_t(\tilde{B}_s, q_s)ds|\tilde{B}_t = b\right] \tag{3.53}$$

where $\mathcal{A}_k$ is the compact set $\mathcal{A}_k = \{q \in \mathbb{R}^n : |q| \leq k\}, k > 0$, and the dynamic of $\tilde{B}$ is

$$d\tilde{B}_s = \psi(s,\tilde{B}_s)\left(q_s ds + dW_s\right). \tag{3.54}$$

Moreover, an optimal control for (3.53) is Markovian and given by

$$q_k^*(t,b) = \underset{q\in\mathcal{A}_k}{\mathrm{argmin}}\left\{q'\psi'\nabla_b\tilde{R}^k(t,b) + L_t(b,q)\right\}.$$

We then deduce that,

$$\tilde{R}^k(t,b) = \mathbb{E}\left[\int_t^{'} L_s\left(\tilde{B}_s^*, q_k^*\left(s,\tilde{B}_s^*\right)\right)ds|\tilde{B}_t^* = b\right]$$

where $\tilde{B}_s^*$ solves stochastic differential Equation (SDE) (3.54) with controls $q_k^*\left(s,\tilde{B}_s^*\right)$. From the theorem of differentiation under the expectation and the integral sign we know that

$$\nabla_b\tilde{R}^k(t,b) = \mathbb{E}\left[\int_t^{'} \nabla_b L_s\left(\tilde{B}_s^*, q_k^*\left(s,\tilde{B}_s^*\right)\right)ds|\tilde{B}_t^* = b\right].$$

To prove the linear growth condition of $\nabla_b\tilde{R}^k$, we will need the following inequality:

$$|\nabla_b L_t(b,q)|^2 \leq C_1 L_t(b,q) + C_2|b|^2,$$

with $C_1 > 0$ and $C_2 \geq 2C_1|\sigma^{-1}b|^2 > 0$.
Having in mind that the function L_t can be written as $L_t(b,q) = \frac{1}{2}|q + 2\sigma^{-1}|^2 - |\sigma^{-1}b|^2$, the previous inequality comes from:

$$\begin{aligned}
|\nabla_b L_t(b,q)|^2 &= |2(\sigma^{-1})'q + 2\Sigma^{-1}b|^2 \\
&\leq 4|\sigma^{-1}|^2|q + 2\sigma^{-1}b - \sigma^{-1}b|^2 \\
&\leq 8|\sigma^{-1}|^2\left(\frac{1}{2}|q + 2\sigma^{-1}b|^2 + |\sigma^{-1}b|^2\right) \\
&\leq C_1 L_t(b,q) + 2C_1|\sigma^{-1}|^2|b|^2 \\
&\leq C_1 L_t(b,q) + C_2|b|^2.
\end{aligned}$$

Then by the Cauchy-Schwartz inequality and the fact that $L_t(b,0) \leq C|b|^2$ for some positive constant C independent of k and t, we have

$$\begin{aligned}
|\nabla_b \tilde{R}^k(t,b)| &\leq \mathbb{E}\left[\int_t^T \left|\nabla_b L_s\left(\tilde{B}_s^*, q_k^*\left(s,\tilde{B}_s^*\right)\right)\right| ds|\tilde{B}_t^* = b\right] \\
&\leq C\mathbb{E}\left[\int_t^T \left|\nabla_b L_s\left(\tilde{B}_s^*, q_k^*\left(s,\tilde{B}_s^*\right)\right)\right|^2 ds|\tilde{B}_t^* = b\right]^{\frac{1}{2}} \\
&\leq C\Bigg(C_1\mathbb{E}\left[\int_t^T L_s\left(\tilde{B}_s^*, q_k^*\left(s,\tilde{B}_s^*\right)\right) ds|\tilde{B}_t^* = b\right] \\
&\quad + C_2\mathbb{E}\left[\int_t^T |\tilde{B}_s|^2 ds|\tilde{B}_t = b\right]\Bigg)^{\frac{1}{2}} \\
&\leq C\left(C_1\mathbb{E}\left[\int_t^T L_s\left(\tilde{B}_s, 0\right) ds|\tilde{B}_t = b\right] + C_2\mathbb{E}\left[\int_t^T |\tilde{B}_s|^2 ds|\tilde{B}_t = b\right]\right)^{\frac{1}{2}} \\
&\leq C\mathbb{E}\left[\int_t^T |\tilde{B}_s|^2 ds|\tilde{B}_t = b\right]^{\frac{1}{2}} \\
&\leq C(1+|b|).
\end{aligned}$$

We used that when $q := 0$, $\tilde{B}_s = \tilde{B}_t + \int_t^s \psi(u,\tilde{B}_u)dW_u$ and thanks to Ito's formula,

$$|\tilde{B}_s|^2 = |\tilde{B}_t|^2 + 2\tilde{B}_t'\int_t^s \psi(u,\tilde{B}_u)dW_u + \int_t^s \operatorname{tr}\left(\psi\psi'(u,\tilde{B}_u)\right) du.$$

Moreover,

$$\begin{aligned}
\mathbb{E}\left[\int_t^T |\tilde{B}_s|^2 ds|\tilde{B}_t = b\right]^{\frac{1}{2}} &= \mathbb{E}\Bigg[\int_t^T \Bigg(|\tilde{B}_t|^2 + 2\tilde{B}_t'\int_t^s \psi(u,\tilde{B}_u)dW_u \\
&\qquad + \int_t^s \operatorname{tr}\left(\psi\psi'(u,\tilde{B}_u)\right) du\Bigg) ds|\tilde{B}_t = b\Bigg]^{\frac{1}{2}} \\
&\leq \left((T-t)|b|^2 + n|\psi|^2(T-t)\right)^{\frac{1}{2}} \\
&\leq C(1+|b|^2)^{\frac{1}{2}} \\
&\leq C\left(1+|b|\right)
\end{aligned}$$

Consequently, since the function $q \to -q'\psi'\nabla_b R^k(t,b) - L_t(b,q)$ attains its maximum on $\mathbb{R}^n$ for

$$\tilde{q}_k(t,b) = -\left(\psi'\nabla_b R^k(t,b) + 2\sigma^{-1}b\right),$$

and knowing that by assumption $|\psi| < \infty$, it is easy to see that

$$|\tilde{q}_k(t,b)| \leq C(1+|b|).$$

As a consequence, for an arbitrarily large constant $\hat{C} > 0$, there exists a positive constant C independent of k such that,

$$|\tilde{q}_k(t,b)| \leq C, \quad t \in [0,T], \quad b \in \mathcal{B}_t \cap \{|b| \leq \tilde{C}\}.$$

Hence, for $k \geq C$, we have

$$\begin{aligned} F^k(t,b,\nabla_b \tilde{R}^k) &= \max_{q \in \mathcal{A}_k} \big[-q'\psi' \nabla_b \tilde{R}^k - L_t(b,q) \big], \\ &= \max_{q \in \mathbb{R}^n} \big[-q'\psi' \nabla_b \tilde{R}^k - L_t(b,q) \big], \\ &= F(t,b,\nabla_b \tilde{R}^k). \end{aligned}$$

for all $(t,b) \in [0,T] \times \mathcal{B}_t \cap \{|b| \leq \tilde{C}\}$. Letting $\tilde{C}$ goes to infinity implies that for k sufficiently large, $R^k = -\tilde{R}^k$ is a smooth solution satisfying a quadratic growth condition to (3.15) - (3.17). □

3.A.5 Proof of Lemma 3.4.5

Proof. To solve the system in the multidimensional Gaussian case, we solve the Riccati equation for M. We look for a symmetric solution M such that $GM = (GM)'$ which solves the following ODE

$$-M'(t) - 2M(t)'G(t)'\Sigma G(t)M(t) + 2G(t)M(t) + 2M(t)'G(t)' - \Sigma^{-1} = 0.$$

A particular solution to the ODE for M is $\hat{M}(t) = G^{-1}(t)\Sigma^{-1}$ and it is easy to see that $\hat{M}$ is symmetric since $\Sigma G = G'\Sigma$, and $G\hat{M} = (G\hat{M})'$.

Now we look for a function N such that $M = \hat{M} + N$. Note that N should be symmetric with $GN = (GN)'$. Plugging M into the ODE yields

$$-N'(t) - 2N'G(t)'\Sigma G(t)N(t) = 0.$$

Now, we change variable $N = \Theta^{-1}$ in the previous ODE to find

$$\Theta'(t) = 2G(t)'\Sigma G(t).$$

Noticing that $\Theta(T) = -\Sigma G(T) = -(\Sigma_0^{-1} + \Sigma^{-1}T)^{-1}$, we obtain after integration

$$\Theta(t) = -\left((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} + \int_t^T 2G(s)'\Sigma G(s) ds \right).$$

Finally, writing $G(t)$ as $(\Sigma + \Sigma_0 t)^{-1}\Sigma_0$, the solution to the original ODE is

$$\begin{aligned} M(t) =& (\hat{M} + \Theta^{-1})(t) \\ =& \Sigma_0^{-1} + \Sigma^{-1}t - \Bigg[(\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} \\ & + 2\int_t^T \Sigma_0 \left(\Sigma + \Sigma_0 s\right)^{-1} \Sigma \left(\Sigma + \Sigma_0 s\right)^{-1} \Sigma_0 ds\Bigg]^{-1}. \end{aligned}$$

To obtain $\mathcal{V}(t)$, we simply integrate and with some simplifications we obtain,

$$\begin{aligned} \mathcal{V}(t) = \int_t^T \operatorname{tr}\Bigg(& \Sigma_0 \left(\Sigma + \Sigma_0 s\right)^{-1} - \Bigg[\Bigg((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} \\ & + 2\int_s^T G(u)'\Sigma G(u) du\Bigg)(G(s)'\Sigma G(s))^{-1}\Bigg]^{-1}\Bigg) ds. \end{aligned}$$

It is easy to verify that $M(T) = \mathcal{V}(T) = 0$.

The next step is to check that N satisfies the conditions we have imposed to derive the solution M. The condition $N = N'$ is satisfied since it is easy to see that $\Theta = \Theta'$. It is straightforward to see that the second condition, $GN = (GN)'$ is equivalent to $\Theta\Sigma_0^{-1}\Sigma = \Sigma\Sigma_0^{-1}\Theta$, once G is written as $(\Sigma + \Sigma_0 t)^{-1}\Sigma_0$. Noticing the symmetry of $G\Sigma_0^{-1}$, we develop the left side of the equality,

$$\begin{aligned} \Theta\Sigma_0^{-1}\Sigma &= -\left((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1} + \int_t^{\intercal} 2G(s)'\Sigma G(s) ds\right)\Sigma_0^{-1}\Sigma \\ &= -\left((\Sigma_0^{-1} + \Sigma^{-1}T)^{-1}\Sigma_0^{-1}\Sigma + \int_t^{\intercal} 2G(s)'\Sigma G(s)\Sigma_0^{-1}\Sigma ds\right) \\ &= -\left(\Sigma\left[(\Sigma\Sigma_0^{-1} + T)\Sigma_0\right]^{-1}\Sigma + \int_t^{\intercal} 2\Sigma G(s)\Sigma_0^{-1}G(s)'\Sigma ds\right) \\ &= -\left(\Sigma\Sigma_0^{-1}\left(\Sigma_0^{-1} + \Sigma^{-1}T\right)^{-1} + \int_t^{\prime} 2\Sigma\Sigma_0^{-1}G'(s)\Sigma G(s) ds\right) \\ &= \Sigma\Sigma_0^{-1}\Theta. \end{aligned}$$

For the one-dimensional case n = 1, we check that the functions:

$$\begin{aligned} M(t) &= \frac{\sigma^2 + \sigma_0^2 t}{\sigma^2\left(\sigma^2 + \sigma_0^2(2T - t)\right)}(T - t) \quad \text{and} \\ \mathcal{V}(t) &= \log\left(\frac{\sigma^2 + \sigma_0^2 T}{\sqrt{(\sigma^2 + \sigma_0^2 t)(\sigma^2 + \sigma_0^2(2T - t))}}\right), \end{aligned}$$

satisfy the following unidimensional ODE system:

$$\begin{cases} M'(t) = -\dfrac{2\sigma^2\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}M(t)^2 + \dfrac{4\sigma_0^2}{\sigma^2+\sigma_0^2 t}M(t) - \sigma^{-2}, \\ \mathcal{V}'(t) = -\dfrac{\sigma^2\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}M(t), \end{cases}$$

with $M(T) = \mathcal{V}(T) = 0$.
We easily see that M and $\mathcal{V}$ satisfy the terminal conditions. We first check that M satisfies the first ODE of the system. We derive the function M with respect to t,

$$M'(t) = \frac{\sigma^{-2}}{(\sigma^2+\sigma_0^2(2T-t))^2}\left(\sigma_0^2(-2\sigma^2 t) + \sigma_0^4(2T^2-4Tt+t^2) - \sigma^4\right),$$

and notice that the right side of the equality for M yields:

$$\begin{aligned} &-\frac{2\sigma^2\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}M(t)^2 + \frac{4\sigma_0^2}{\sigma^2+\sigma_0^2 t}M(t) - \sigma^{-2} \\ &= \frac{\sigma^{-2}}{(\sigma^2+\sigma_0^2(2T-t))^2}(-2\sigma_0^4(T-t)^2 + 4\omega^2(T-t) \\ &\quad \times (\sigma^2+\sigma_0^2(2T-t)) - (\sigma^2+\sigma_0^2(2T-t))^2) \\ &= \frac{\sigma^{-2}}{(\sigma^2+\sigma_0^2(2T-t))^2}\left(\sigma_0^2(-2\sigma^2 t) + \sigma_0^4(2T^2-4Tt+t^2) - \sigma^4\right) \\ &= M'(t). \end{aligned}$$

We then verify the ODE for $\mathcal{V}$ by deriving the function $\mathcal{V}$ with respect to t:

$$\begin{aligned} \mathcal{V}'(t) &= (\sigma^2+\sigma_0^2 T)\frac{\omega^2(2\sigma_0^2(T-t))}{2\left((\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))\right)^{\frac{3}{2}}}\frac{\sqrt{(\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))}}{\sigma^2+\sigma_0^2 T} \\ &= -\frac{\sigma_0^4(T-t)}{(\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))}, \end{aligned}$$

and from the right part of the ODE, we find the equality:

$$\begin{aligned} -\frac{\sigma^2\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}M(t) &= -\frac{\sigma_0^4}{(\sigma^2+\sigma_0^2 t)^2}\frac{\sigma^2+\sigma_0^2 t}{\sigma^2+\sigma_0^2(2T-t)}(T-t) \\ &= -\frac{\sigma_0^4(T-t)}{(\sigma^2+\sigma_0^2 t)(\sigma^2+\sigma_0^2(2T-t))} \\ &= \mathcal{V}'(t). \end{aligned}$$

□

Chapter 4

Dealing with drift uncertainty: a Bayesian learning approach

Abstract. One of the main challenges investors have to face is model uncertainty. Typically, the dynamics of the assets are modeled using two parameters: the drift vector and the covariance matrix, which are both uncertain. Since the variance/covariance parameter is assumed to be estimated with a certain level of confidence, we focus on drift uncertainty in this paper. Building on filtering techniques and learning methods, we use a Bayesian learning approach to solve the Markowitz problem and provide a simple and practical procedure to implement optimal strategy. To illustrate the value added of using the optimal Bayesian learning strategy, we compare it with an optimal non-learning strategy that keeps the drift constant at all times. In order to emphasize the prevalence of the Bayesian learning strategy above the non-learning one in different situations, we experiment three different investment universes: indices of various asset classes, currencies, and smart beta strategies.

Key Words: Bayesian learning, optimal portfolio, Markowitz problem, portfolio selection.

Contents

4.1 Introduction **110**
4.2 The model **112**
4.3 Market data **115**
4.4 The Base Case result **118**
4.5 Sensitivity analysis **119**
4.5.1 Impact of uncertainty 119

4.5.2 Impact of leverage . . . 121
4.5.3 Impact of the review frequency . . . 123
4.5.4 Impact of the rebalancing frequency . . . 125
4.6 Investing in foreign currencies . . . 126
4.7 Investing in factor strategies . . . 128
4.A Appendices . . . 130
4.A.1 Dataset of Indices . . . 131
4.A.1.1 Commodity . . . 131
4.A.1.2 Bond . . . 131
4.A.1.3 Equity . . . 133
4.A.1.4 Cash . . . 133
4.A.2 Dataset of currencies . . . 134
4.A.3 Dataset of Smart Beta strategies . . . 134
4.A.3.1 Smart Beta strategies . . . 135

4.1 Introduction

The seminal work by Markowitz (1952) initiated the modern portfolio theory and provided a solution to the portfolio selection problem according to a mean-variance criterion. The mean-variance optimal portfolio is the one that maximizes its expected return at a given level of risk, measured by its variance, or conversely the one that minimizes its risk at a given level of expected return. This one-period model has known multiple extensions, among them, the discrete time multi-period model by Samuelson (1969) and the continuous time model by Merton (1969).
Initially, the parameters of these models have been considered known and constant, especially the parameters that drive the behavior of the assets: drifts and covariances. Not only it oversimplifies the reality but also it raises the question of estimating these parameters. The most basic and widely used method consists in estimating drifts and covariances from past data and fix them once and for all. Estimating volatility in this way, appears to give relatively good results in practice while estimating the drift seems to be more difficult or even impossible, see Merton (1980). Moreover, optimal portfolios are very sensitive to the level of expected returns, as shown in Best and Grauer (1991), and a wrong estimation can result in very suboptimal portfolios a posteriori.
This explains the development of the literature on parameters uncertainty in portfolio analysis, see for instance Barry (1974) and Klein and Bawa (1976), and especially on Bayesian statistics, see Frost and Savarino (1986), Aguilar and West

(2000), Avramov and Zhou (2010), and Bodnar et al. (2017). Nonetheless, these models remain static and cannot benefit from the flow of information which results in a non-adaptive strategy, unable to process the most recent information conveyed by the assets market prices. It is one of the main reasons why a literature on filtering and learning techniques in a partial information framework has developed, see Lakner (1995), Lakner (1998), Rogers (2001), and Cvitanić et al. (2006). The Bayesian learning approach consists in modeling the uncertainty of a set of parameters by a prior distribution, representing the beliefs of the investor on the potential values of the parameters, which is updated with incoming information, for instance assets market prices. In particular, in our companion paper De Franco, Nicolle, and Pham (2019a), we have solved the multidimensional Markowitz problem in the case of an uncertain drift using Bayesian learning with a Gaussian prior.

In this paper, we adapt the results in De Franco, Nicolle, and Pham (2019a) to implement the strategy in practice. Indeed, the solution provided is in continuous time and the amounts invested in the assets are unconstrained. Of course, in reality, trading is discrete and amounts invested in assets are limited, so we discretize the optimal solution of the continuous Markowitz problem, turn amounts into proportions and cap the optimal proportions to be invested in order to fit the portfolio management constraints. Our purpose is to show the prevalence of the Bayesian learning strategy above the non-learning one which considers the drift constant. To do so, we illustrate our point by confronting both strategies to different datasets representing different investment universes. First, we use a panel of major indices in four different asset classes: corporate bonds, sovereign bonds, commodities and equities completed with cash. Then, we consider bank accounts in foreign banks that pay the local interest rate but are valued in EUR in order to study the performance of both strategies with respect to foreign exchange rates. Finally, we implement both strategies in an investment universe composed of smart beta strategies. Moreover, using the first dataset, we provide a sensitivity analysis of both strategies to various parameters: the uncertainty in the model, the impact of the leverage, the review frequency and the rebalancing frequency. We do not show this analysis for the two other datasets since we would find similar results.

The paper is organized as follows: Section 4.2 details the model and the discretized optimal strategies while Section 4.3 depicts the market data and the work flow. Section 4.4 shows the results in the case of the first dataset and Section 4.5 is about the sensitivity analysis of the Bayesian learning and the non-learning strategies applied to this dataset. Finally, Section 4.6 deals with foreign exchange rates and Section 4.7 with smart beta strategies.

4.2 The model

We consider a financial market consisting of one risk-free asset, whose return is denoted by r^f, and n risky assets whose returns r_t are modeled by

$$\begin{cases} r_t &= B + \sigma\xi_t \\ \xi_t &\sim N(0, I_n). \end{cases} \tag{4.1}$$

We shall assume that the random vectors ξ_t are independent for all t. Model (4.1) includes major linear models available in the financial literature such as CAPM (Sharpe, 1964; Lintner, 1965), discrete-time Black and Scholes (Black and Scholes, 1973) or Fama-French models (Fama and French, 1993, 2015, 2016). $B \in \mathbb{R}^n$ is the vector of the expected returns of the risky assets, while $\Sigma := \sigma\sigma' \in \mathbb{R}^{n\times n}$ is the covariance matrix of the risky assets. We assume that σ^{-1} exists. Within such financial markets, for $A \subseteq \mathbb{R}^n$, we define the set of admissible investment strategies as

$$\mathcal{A} := \{\mathbf{w} = (\mathbf{w}_t)_{0\leq t\leq T} | \mathbf{w}_t \in \mathrm{A}\},$$

where an admissible strategy $\mathbf{w} = (\mathbf{w}_t)_t$ represents the fraction of wealth invested in the assets at any time t. We write the wealth at maturity $T > 0$ as

$$X_T^{\mathbf{w}} = X_0 \prod_{s=0}^{T-1} \left(1 + \mathbf{w}_s' \mathbf{r_{s+1}} + (1 - \mathbf{w}_s' \mathbb{1}_n)\, r^f\right).$$

We consider an investor who is aiming to solve the Markowitz problem:

$$\begin{cases} \max\limits_{\mathbf{w}\in\mathcal{A}} \mathbb{E}\left[X_T^{\mathbf{w}}\right] \\ \\ \mathrm{Var}\left(X_T^{\mathbf{w}}\right) \leq \vartheta, \end{cases} \tag{4.2}$$

where $\vartheta > 0$ is the risk tolerance for the investor.

The initial version of the Markowitz problem (Markowitz, 1952), which was stated for a single period, has been widely studied and solutions in the multi-period framework (such ours) in both discrete and continuous time have been provided (see e.g. Merton (1969); Samuelson (1969); Merton (1971); Karatzas et al. (1987) among others). The common assumptions in previous works are that both expected return and volatility coefficients (B and Σ in our framework) are known. In practice, these parameters are not directly observable and must be estimated from the data or input by the investor at inception. In both cases, biased parameters can significantly affect ex-post performance of the optimal strategy. Although the parameter Σ can be estimated from the data with some degree of confidence, the estimation of B

turns out to be quite difficult if not possible at all. Because the optimal strategy strongly depends on both B and Σ, a wrong estimation could significantly affect the optimal strategy.
To get closer to reality and account for model uncertainty, we assume reasonably that the investor has an *a priori* view on the risky assets and their *expected* returns, but she is uncertain about how good her forecast is. Introducing uncertainty in the problem brings it closer to the real-life situation, where not only the investor does not know the parameters of the model, but also is forced to admit that her estimates are uncertain. More precisely, the investor does not observe B and only assumes that $B \sim \mu$, where μ is a probability distribution in $\mathbb{R}^n$, centered at b_0 ($\mathbb{E}[B] = b_0$). The parameter b_0 is the vector of returns the investor is expecting while μ translates her uncertainty about it.

Remark 5. *When $\mu = \delta_{b_0}$, the Dirac distribution at b_0, then the investor has no uncertainty about her forecast.* ♢

Remark 6 (Discretization)**.** *Implementing the optimal strategies in practice, leads to make adjustments to the optimal solutions in continuous time and make them suboptimal. Here, we discretize the continuous optimal solutions and controls, and we cap them. To propose a solution to Problem (4.2), we suggest to confront the discretized continuous optimal Bayesian learning and non-learning solutions since the investor can observe assets prices nearly continuously but trades in discrete time.* ♢

The paper by De Franco, Nicolle, and Pham (2019a) solved the continuous-time version of Problem (4.1) for a large class of distributions μ, in a Bayesian framework using dynamic programming techniques. We report here the discretized version of the results in the Gaussian case $\mu = N(b_0, \Sigma_0)$ and $A = \mathbb{R}^n$:

Application (Discretized-version of the Bayesian learning continuous optimal solution)**.** *Let $\mu = N(b_0, \Sigma_0)$ with $det(\Sigma_0) > 0$ and $A = \mathbb{R}^n$. Then, the discretized version of the continuous-time optimal solution of the Bayesian-Markowitz problem is given by*

$$\mathbf{w}_s^{BL} = \alpha\left(s, X_s, \hat{B}_s\right)/X_s \quad \textbf{\textit{Bayesian learning (BL)}}, \tag{4.3}$$

where X_s is the wealth process at time s and

- $\hat{B}_s = \left({\Sigma_0}^{-1} + s\Sigma^{-1}\right)\left({\Sigma_0}^{-1} b_0 + (\sigma')^{-1} Y_s\right)$,
- $[\sigma Y_s]^i = \prod_{u=0}^{s-1}(1 + r_u^i) + \frac{s}{2}\Sigma_{ii}$,
- $\alpha(s, x, b) = \left(x_0 - x + \sqrt{\vartheta}\frac{e^{R(0,b_0)}}{\sqrt{e^{R(0,b_0)}-1}}\right)\left(\Sigma^{-1} b - (\psi(s)\sigma^{-1})' \nabla_b R(s, b)\right)$,

- $\psi(s) = \Sigma_0 \left(\Sigma + s\Sigma_0\right)^{-1} \sigma$,
- *R is the unique solution to the following semi-linear parabolic PDE*

$$\begin{cases} 0 = -\partial_t R - \frac{1}{2}\mathrm{tr}\left(\psi\psi'\mathcal{D}_b^2 R\right) + 2\left(\psi\sigma^{-1}b\right)'\nabla_b R - \frac{1}{2}\left|\psi'\nabla_b R\right|^2 - \left|\sigma^{-1}b\right|^2, \\ 0 = R(T,b), \end{cases}$$

which in the case of a Gaussian prior proves to be of the form $R(s,b) = b'M(s)b + \mathcal{V}(s)$, where M is the solution to a multidimensional Riccati equation, for which the solution can be found explicitly, and r is the solution to a first-order linear differential equation depending on ψ and M, which can also be explicitly calculated. See De Franco, Nicolle, and Pham (2019a) for further details.

The process $\hat{B}$ is the conditional expectation of B given the current observation of the assets returns, which are given by Y. The matrix Σ_0 represents the uncertainty around b_0. As time goes by, we observe more returns which in turn improves our knowledge of B, as one can see from the fact that when $s \to T$, more weight is put on Σ^{-1} in the definition of $\hat{B}$. The matrix valued function ψ is linked to the conditional covariance of B given Y.

Remark 7. *When there is no uncertainty ($\mu = \delta_{b_0}$), or when the investor directly inputs her estimate b_0, the structure of the discretized-version of the optimal continuous solution is simplified as follows:*

$$\mathbf{w}_s^{NL} = \alpha^{NL}\left(s, X_s\right)/X_s \quad \textbf{\textit{non-learning (NL)}}, \tag{4.4}$$

where

- $\alpha^{NL}(s,x) = \left(x_0 - x + \sqrt{\vartheta}\frac{e^{R(0,b_0)}}{\sqrt{e^{R(0,b_0)}-1}}\right)\Sigma^{-1}b_0$,
- $R(s,b) = b'\Sigma^{-1}b(T-s)$.

$\Diamond$

The main differences between the Bayesian learning and the non-learning strategies arise from

- The *market risk premium* R in the leverage coefficient,
- The *correction* term $\left(\psi(s)\sigma^{-1}\right)'\nabla_b R(s,b)$ which is zero for the non-learning strategy.

As time goes by, we observe realized returns and we learn more about B. Indeed, with uncertainty, the Bayesian learning strategy is updated with $\hat{B}$, which is the conditional expectation of B given the current observation (*new knowledge*). The trade $\Sigma^{-1}\hat{B}$ is modified with the corrective term $(\psi(s)\sigma^{-1})' \nabla_b R(s, \hat{B})$.
Both $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ are unconstrained. In order to provide a realistic analysis on both strategies, we consider capped versions of them.

Definition 4.2.1. *The capping operator for an investment strategy* $\mathbf{w}$ *with a leverage* $l > 0$ *is*

$$c(\mathbf{w}, l) := \mathbf{w}\mathbb{1}_{|\mathbf{w}|_1 \leq l} + \frac{l}{|\mathbf{w}|_1}\mathbf{w}\mathbb{1}_{|\mathbf{w}|_1 > l}.$$

In Section 4.3, we will implement both the Bayesian learning and the non-learning strategies in the context of asset allocation with real market data, to get insights on the effect of learning and its value added (*value of information*).

4.3 Market data

We consider 4 asset classes across different regions, each of which represented by a well-known market index, and the EONIA rate as the risk-free rate for an investor whose base currency is the Euro.

Bloomberg Ticker	Name	Currency	Asset Class
SPDYCITR Index	S&P GSCI Dynamic Roll TR	USD	Commodity
GOLDLNPM Index	LBMA Gold Price PM	USD	Commodity
LECPTREU Index	Bloomberg Barclays Euro Aggregate Corporate TR Index	EUR	Corporate Bonds
IBOXHY Index	iBoxx USD Liquid High Yield Index	USD	Corporate Bonds
IBOXIG Index	iBoxx USD Liquid Investment Grade Index	USD	Corporate Bonds
IBOXXMJA Index	EUR Corporate Liquid Hight Yield	EUR	Corporate Bonds
SPTR500N Index	S&P 500 Net TR Index	USD	Equity
SX5T Index	Eurostoxx 50 Net TR Index	EUR	Equity
TUKXG Index	FTSE 100 TR Index	GBP	Equity
NDUEEGF Index	MSCI Emerging Net TR Index	USD	Equity
SPTPXN Index	S&P Topix 150 NR	JPY	Equity
LUATTRUU Index	Bloomberg Barclays US Treasury TR Unhedged Index	USD	Sovereign Bonds
LEATTREU Index	Bloomberg Barclays EurAgg Treasury TR Unhedged Index	EUR	Sovereign Bonds
JPEICORE Index	JP Morgan EMBI Global Core Index	USD	Sovereign Bonds
FTFIBGT Index	FTSE Actuaries UK Conv. GILTs All Stocks TR Index	GBP	Sovereign Bonds
EONIA Index	EMMI EURO Overnight Index average	EUR	Cash

Table 4.1: Major market indices relative to each currency and asset class.

Non Euro-denominated indices are hedged against the Euro simply by implementing a monthly rolled hedging overlay with 1-month forward contracts. We collect data from December 1998, except for the EUR Corporate Liquid High Yield for which we only have data since January 2006. Prices are sourced from Bloomberg while currency spot and forward rates come from Datastream, and both refer to the 4 p.m. London fixing. Our choice of market indices is motivated by their popularity

among investors, their liquidity, and the wide range of financial products available in the market that give exposures to these indices (such as listed futures and ETFs). While limited in number, they provide well-diversified exposures to major global asset classes. Therefore, this suits the underlying premises of the Markowitz problem which is less suited for asset allocation in presence of many underlyings.
The testing period begins in January 2000 and ends in June 2018. To implement both BL and NL, we iteratively follow the work flow outlined below, while Table 4.2 collects the parameters (in bold in the work flow) used for our test.

Work flow.

- *Let* $\mathbf{T} > 0$ *and consider the time-frame* $\{t_0 < t_1 < \cdots t_n = T\}$.
- *We call* t_0 *a Review date since at this date we estimate all the parameters and calculate the function* R *that define* $\mathbf{w}^{BL}$ *in* (4.3) *and* $\mathbf{w}^{NL}$ *in* (4.4). *To ensure a realistic implementation of both strategies, all data-based estimations are performed with data available before the Review date* t_0 *and ending on* $t_0 - \mathbf{Lag}$. *The first* ***Review*** *date is January 21, 2000.*
- *We call* t_k *a Rebalancing date if, at this date, we update the portfolio weights according to* (4.3) *and* (4.4). *To limit turnover and transaction costs, the subsets of Rebalancing dates in* $[0, T]$ *is relatively small. We assume that both Review and Rebalancing dates are Fridays, so that the number of Rebalancing dates in the* $[0, T]$ *period is given by the frequency* ***Freq***.
- b_0 *is estimated as the sample mean over the past* $r^{\mathbf{w}}$ *days ending on* $t_0 - \mathbf{Lag}$ *or over the maximum data available with a minimum of 30 data-points.* Σ *is estimated as the sample covariance matrix over* $s^{\mathbf{w}}$ *days ending on* $t_0 - \mathbf{Lag}$ *or over the maximum data available with a minimum of 30 data-points.*
- *We consider a parametric function for* ϑ *as follows:*

$$\vartheta = \vartheta(\mathbf{d}) := T\left(\mathbf{d} \times X_{t_0}\right)^2 / \left(1.96^2\right), \quad \mathbf{d} \in (0, 1).$$

The motivation behind our choice comes from Problem (4.2). *Indeed, we expect that*

$$X_T \succeq \mathbb{E}\left[X_T\right] - 1.96 \times \sqrt{\frac{Var\left(X_T\right)}{T}} \geq \mathbb{E}\left[X_T\right] - 1.96 \times \sqrt{\frac{\vartheta}{T}},$$

and we set ϑ *so that within our confidence interval, the gap to the expected value of* X_T *is a fraction of the initial wealth*

$$1.96 \times \sqrt{\frac{\vartheta}{T}} \sim \mathbf{d} \times X_{t_0}.$$

- *The uncertainty around the estimate b_0 is measured by Σ_0. We assume Σ_0 to be diagonal. Therefore, each diagonal entry measures the degree of confidence we have on the relative entry of b_0. Each diagonal entry is modeled as follows:*

$$\Sigma_0^{ii} = \Sigma_0^{ii}(\mathbf{unc}) := \mathbf{unc} \times \left(\frac{pct_{95}^i - pct_5^i}{2} \right)^2,$$

  where pct_{95}^i (pct_5^i) is the 95% (5%) quantile of the empirical distribution of the return r^i. To calculate these quantiles, we consider the time series of the returns r^i spanning for r^q days and ending on $t_0 - \mathbf{Lag}$. Up to the parameter **unc**, *the square root of Σ_0^{ii} is half of the segment, centered at the median, that contains 90% of the empirical distribution over r^q days.*

- *Strategies $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ are considered in their capped versions according to Definition 4.2.1 with a maximum leverage* **L**.

- *At time T, we rerun the* ***Work Flow*** *by setting $t_0 = T$ and update all parameters according to the new data available.*

Parameter	Value
X_0	100
T	3 months
Lag	1 day
Freq	Monthly, 3^{rd} Friday of each month.
$r^{\mathbf{w}}$	750 days
$s^{\mathbf{w}}$	125 days
d	10%
unc	100
r^q	125
L	200%

Table 4.2: Parameters used in the implementation of BL and NL as defined in the **Work Flow**.

The **Work Flow** and Table 4.2 represent the *Base Case*. In Section 4.4, we report the results for both the Bayesian learning (BL) and the non-learning (NL) strategies over the last 18 years. In Section 4.5, we will assess the sensitivity of both strategies to changes in the parameters of the *Base Case* to get a better grasp of the value added of learning in the context of portfolio construction.

4.4 The Base Case result

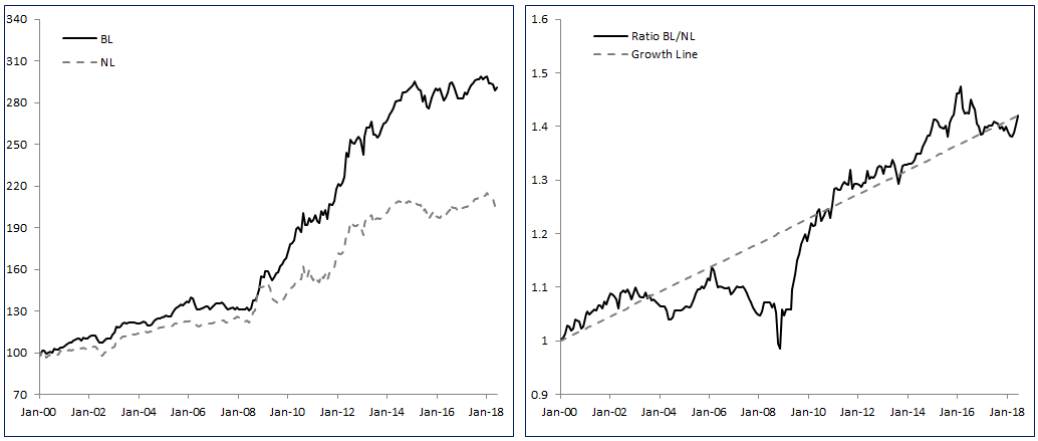
We implement the **Work Flow** to build the wealth processes X^{BL} and X^{NL} as in (4.2) related to the BL and NL strategies in the Base Case framework (Table 4.2). In the sequel, for the sake of simplicity, we refer to BL as either the strategy weights $\mathbf{w}^{BL}$ or the associated wealth process X^{BL} and we do the same for NL. Table 4.3 collects long term statistics for both strategies.

	BL	NL	Difference
Ann. Performance	5.96%	3.96%	2.00%
Ann. Volatility	5.51%	4.92%	0.59%
Max. Drawdown	-8.51%	-11.2%	2.70%
Sharpe ratio	1.08	0.80	34.26%
Information ratio	0.55	-	-

Table 4.3: Statistics for the BL and NL strategies. The Sharpe Ratio is calculated as the ratio between annualized performance and annualized volatility. For this one, the difference is given in relative terms. Data from January 2000 to June 2018.

Over the period starting from January 2000 to June 2018, BL has delivered an annualized performance of 5.96% while NL has reached 3.96%. Incorporating uncertainty and learning from the data yields an annualized 2% excess return. In terms of risk metrics, annualized volatility is slightly higher for BL (0.59% difference) but it also shows a better maximum drawdown figure (-8.51% for BL versus -11.20% for NL). Finally, the Sharpe ratio is 1.08 for BL versus 0.80 for NL, or 34.26% improvement in relative terms. The information ratio of BL over NL is about 0.55, meaning that the performance of BL comes from a better ability to process incoming information.

Figure 4.1a shows the historical levels of the wealth processes BL and NL, while Figure 4.1b provides the relative strength of BL over NL as well as the growth line. An increasing relative strength index signals outperformance of BL over NL, while a decreasing index shows underperformance.



(a) Levels (b) Ratio

Figure 4.1: Historical values of the portfolios calculated with both $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ under the Base Case (a) and their ratio (b).

Looking at Figure 4.1b, BL mostly delivers a robust and regular outperformance over NL. Integrating uncertainty in the drift, learning from the data and adjusting the strategy accordingly, clearly adds value over time. Specifically, we identify three distinct periods: since inception until February 2006, the ratio BL/NL is trending upward at moderate pace reaching the value 1.13 (or equivalently 13% cumulated outperformance), then BL underperforms NL until October 2008, and finally, from October 2008 to June 2018, BL strongly dominates NL with 40% of cumulated outperformance.

4.5 Sensitivity analysis

In the following paragraphs, we will stress-test the BL strategy by measuring the effects of the main parameters on long term results. Our stress-test methodology consists in fixing all but one of the parameters detailed in Table 4.2 and varying the remaining parameter across plausible values.

4.5.1 Impact of uncertainty

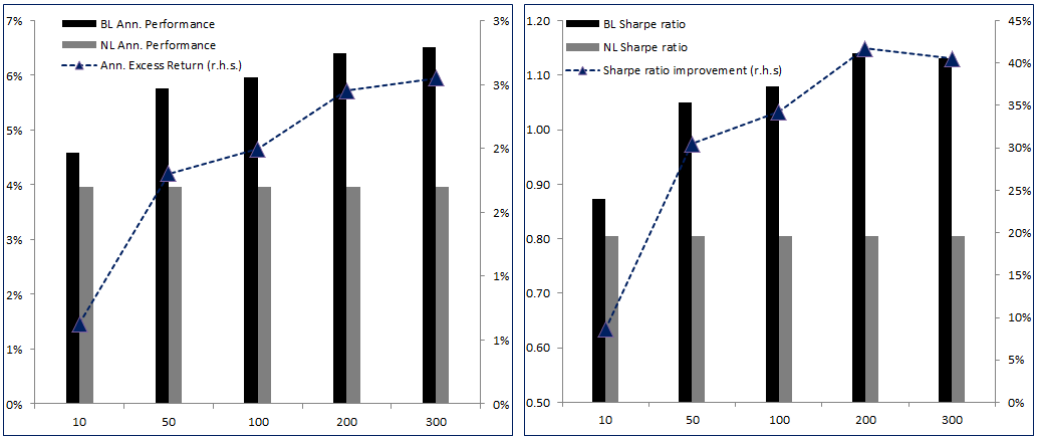
We study the effect of the uncertainty parameter *unc* in the strategy. Higher values of *unc* signal a higher standard deviation of the estimate b_0, hence a higher uncertainty on the estimate of the expected returns. We study the BL strategy for $unc \in \{10, 50, 100, 200, 300\}$, where the value $unc = 100$ corresponds to the Base Case detailed in Section 4.4. Table 4.4 collects the results.

	BL					NL
unc	10	50	100	200	300	
Ann. Performance	4.59%	5.76%	5.96%	6.41%	6.51%	3.96%
Ann. Volatility	5.25%	5.49%	5.51%	5.62%	5.76%	4.92%
Max. Drawdown	-10.21%	-8.56%	-8.51%	-8.27%	-8.58%	-11.2%
Sharpe ratio	0.87	1.05	1.08	1.14	1.13	0.8
Sharpe ratio impr.	8.64%	30.45%	34.26%	41.72%	40.54%	-
Information ratio	0.34	0.58	0.55	0.57	0.54	-

Table 4.4: Statistics for BL strategies with different values of *unc* and NL as defined in the **Work Flow**. The Sharpe ratio improvement is relative to NL.

Intuitively, among the BL strategies detailed in Table 4.4, the strategy with $unc = 10$ is the closest to NL and this is confirmed by the performance and volatility figures. Indeed, with such a low level of uncertainty, we are confident in our initial estimate b_0. As we increase *unc*, the confidence we place in b_0 decreases.
Very interestingly, as *unc* increases, the excess return of the corresponding BL strategy over NL increases, without significantly increasing the risk. Therefore, Sharpe ratios increase with *unc*: from 0.87 for $unc = 10$ to 1.13 for $unc = 300$. More striking is the relative increase in Sharpe ratios with respect to NL: for $unc = 10$ we only have an increase of 8.64%; for $unc = 300$, the relative increase is 40.54%. Finally, the Information ratio rapidly stabilizes around 0.5.

Figures 4.2a-4.2b graphically render the increase in both annualized performances, excess returns and Sharpe ratios with respect to NL as a function of *unc*.



(a) Performances (b) Sharpe ratios

Figure 4.2: Annualized performances and excess returns of BL as a function of *unc* (a); Sharpe ratios and relative improvement (b).

Figure 4.3 shows the historical levels of the strategy BL with $unc = 10$ (minimum uncertainty), $unc = 300$ (high uncertainty) and $unc = 100$ (Base Case) compared to NL. The BL strategies share the same profile, but the higher the *unc* parameter,

the higher the excess return we find in the long run.

Clearly, the standard solution to the Markowitz problem suffers from the poor estimate of b_0, while the BL strategy is able to adjust and react to new observable data. Moreover, the higher the uncertainty, the better BL behaves compared to NL. Unreported tests showed us that with this particular dataset, we can increase the *unc* parameter even further. At some point though, around 10,000, the BL strategy underperforms NL. Indeed, when uncertainty is extremely high, the matrix Σ_0 that controls the a priori knowledge we have on B is simply uninformative because we are allowing B to span a too vast region of potential values. Therefore, the learning process, in a relatively short period of time of three months, is slow and does not add any value.

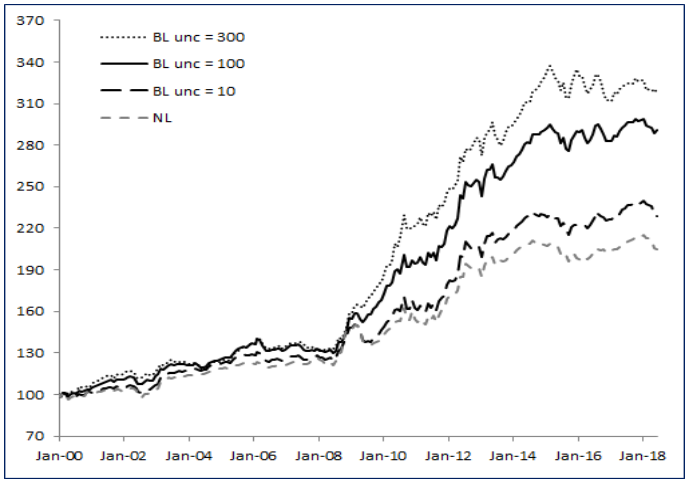


Figure 4.3: Historical values of the BL portfolios calculated with $unc = 10, 100, 300$ and NL.

4.5.2 Impact of leverage

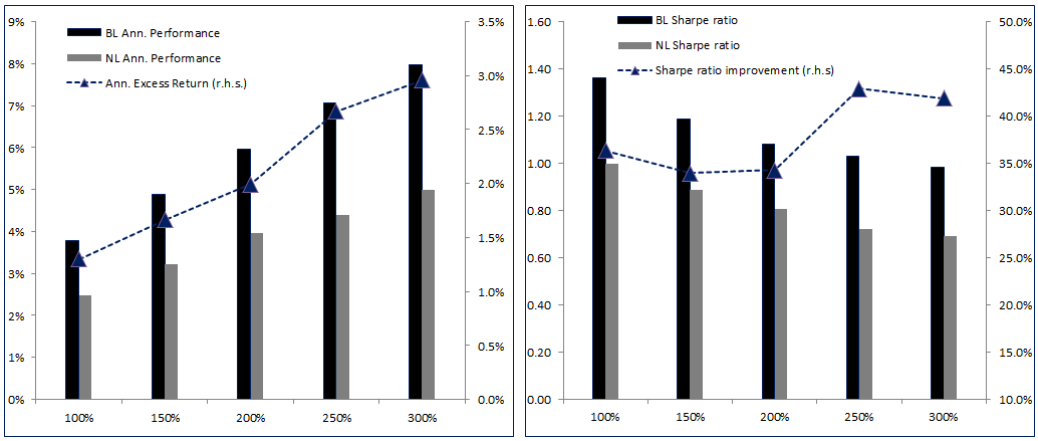
We now look at the maximum leverage parameter L in both strategies. Higher values of L make both $c(\mathbf{w}^{BL}, L)$ and $c(\mathbf{w}^{NL}, L)$ closer to the unconstrained weights. We test $L \in \{100\%, 150\%, 200\%, 250\%, 300\%\}$, where the value $L = 200\%$ corresponds to the Base Case detailed in Section 4.4. Table 4.5 collects the results, while Figures 4.4a-4.4b give a graphical overview of the impact on long term statistics.

L	100%		150%		200%		250%		300%	
	BL	NL	BL	NL	BL	NL	BL	NL	BL	NL
Ann. Performance	3.78%	2.48%	4.89%	3.23%	5.96%	3.96%	7.06%	4.39%	7.96%	5.00%
Ann. Volatility	2.78%	2.48%	4.12%	3.64%	5.51%	4.92%	6.85%	6.09%	8.11%	7.23%
Max. Drawdown	-3.96%	-5.59%	-5.92%	-8.43%	-8.51%	-11.2%	-12.49%	-13.01%	-16.33%	-14.07%
Sharpe ratio	1.36	1	1.19	0.89	1.08	0.8	1.03	0.72	0.98	0.69
Sharpe ratio impr.	36.36%	-	33.96%	-	34.26%	-	42.89%	-	41.88%	-
Information Ratio	0.69	-	0.65	-	0.55	-	0.58	-	0.53	-

Table 4.5: Statistics for BL and NL strategies with different levels of maximum leverage L. The Sharpe ratio improvement is calculated as the relative difference between BL and NL Sharpe ratios for the same level of maximum leverage.

As the maximum leverage L increases, the corresponding excess return of the BL strategy over NL increases. For $L = 100\%$, the excess return is 1.30% and goes up to 2.96% for $L = 300\%$.
Clearly, and as expected, the excess return increases with the leverage, and on average, we observe that adding 100% of leverage brings 0.80% in extra excess return. When we look at the Sharpe ratios, we see that BL always outperforms NL but the relation is more complex than the previous one. Both BL and NL Sharpe ratios decrease with L. Indeed, as L increases, we observe higher performances but also higher volatilities, and the volatility grows faster than the performance. This is obvious because when L increases, the strategy becomes more sensitive to any error in the estimated parameters and this is reflected in higher volatility. As $L \to \infty$, the strategies become unconstrained and bring all the instability that is well known to go alongside with price-based strategies. On a relative basis, the improvement in the Sharpe ratios first decreases as L reaches 200% and then increases.

To conclude, it appears that the impact of the maximum leverage L is as expected: it brings extra performance at the cost of more risk. Because we know real data do not need to follow model (4.1), unconstrained BL and NL (as any price-based strategies) tend to amplify any model error, making therefore the leverage constraint a wise feature to consider.



(a) Performances (b) Sharpe ratios

Figure 4.4: Annualized performances and excess returns of BL as a function of L (a); Sharpe ratios and relative improvement (b).

4.5.3 Impact of the review frequency

Probably one of the most important features of the learning effect on the portfolio is the frequency at which we recalibrate the parameters of the model. We have chosen a three-month frequency in the Base Case. In other words, every three months, we compute new estimates of b_0 and Σ and input a new matrix Σ_0 driving the uncertainty. We expect that the lower the frequency, the better BL compared to NL. The main reason behind this thesis is that the strategy NL will be stuck with parameters b_0 and Σ for a long period of time, during which we know that they will most likely become obsolete as the market evolves and new information is processed. As investors know very well, forecasting is difficult and outdated forecasts are badly suited for portfolio construction[1]. On the other side, BL can adapt over time because it embeds uncertainty about b_0.

Let us consider the frequency parameter *Freq* in both strategies varying from three to twelve months: $Freq \in \{3M, 6M, 9M, 12M\}$, where the value $Freq = 3M$ corresponds to the Base Case detailed in Section 4.4. Table 4.6 collects the results.

[1] *It's tough to make predictions, especially about the future.* The quote, which clearly applies in the context of portfolio construction, has been reported by many different people. According to https://quoteinvestigator.com/2013/10/20/no-predict/, it first appeared in Danish documents, but well known personalities such as Neils Bohr, Mark Kac, Stanislaw M. Ulam, and, probably misattributed, Mark Twain and Yogi Berra.

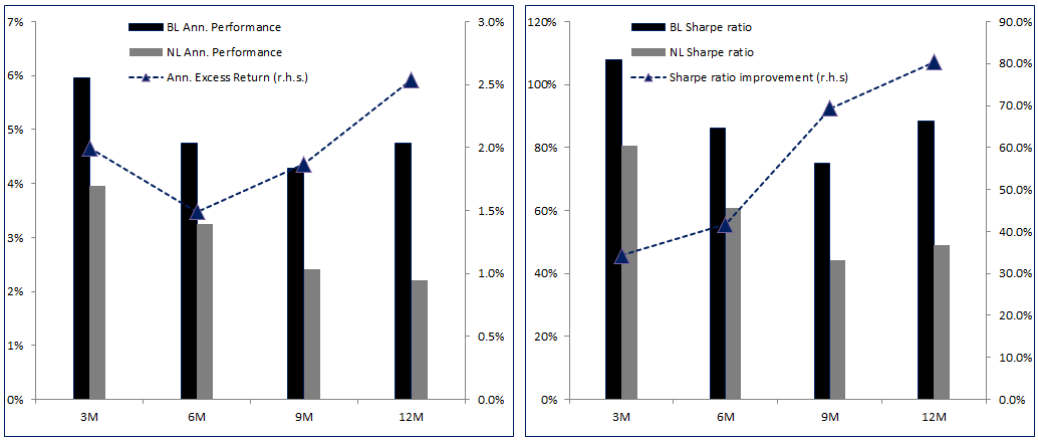
Freq	3M		6M		9M		12M	
	BL	NL	BL	NL	BL	NL	BL	NL
Ann. Performance	5.96%	3.96%	4.74%	3.25%	4.28%	2.41%	4.75%	2.21%
Ann. Volatility	5.51%	4.92%	5.5%	5.35%	5.71%	5.45%	5.36%	4.5%
Max. Drawdown	-8.51%	-11.2%	-11.62%	-13.71%	-12.44%	-16.15%	-8.9%	-16.05%
Sharpe ratio	1.08	0.8	0.86	0.61	0.75	0.44	0.89	0.49
Sharpe ratio impr.	34.26%	-	41.74%	-	69.38%	-	80.4%	-
Information Ratio	0.55	-	0.33	-	0.35	-	0.57	-

Table 4.6: Statistics for BL and NL strategies with different levels of the review frequency *Freq*. The Sharpe ratio improvement is calculated as the relative difference between BL and NL Sharpe ratios for the same review frequency.

As we lower the review frequency, we see the performances of both BL and NL going down. Nevertheless, the excess return of BL over NL actually increases at lower than 6M-frequencies. The fact that it is not monotonic is related to the timing effect. The best metric to assess the value added of the learning feature remains the relative improvement in Sharpe ratios, as shown in Figure 4.5b. Here we see that as the frequency lowers, more risk-adjusted value added comes from the learning effect: at 3M, the relative improvement in Sharpe ratio between BL and NL is 34.26%. At 6M it goes up to 41.47%, then 69.38% at 9M and finally 80.4% relative improvement at the 12M-frequency. Clearly, BL outperforms NL if we do not review the parameters of the model quite often, and this is clearly attributable to the fact that BL can adjust over time according to the data observed and the adjustments it can make on the a priori distribution of B.

For the sake of simplicity, we do not report the results here, but it is possible to increase the efficiency of BL compared to NL if we modify simultaneously the frequency *Freq* and the uncertainty parameter *unc* in Σ_0. Over long investment horizons, the investor definitely has more uncertainty on her a priori estimate b_0, therefore it makes perfectly sense to consider *Freq = 12* coupled with *unc = 1600 or higher*.[2]

[2]We suggest 1600 simply as a rule of thumb: the ratio of frequency between 3M and 12M is 4 so that we pass a $4^2 = 16$ factor in *unc*. Of course, we do not expect this parameter to be the best choice since, empirically, uncertainty grows more than linearly with time, but even this, provides slightly higher Sharpe ratio improvements.



(a) Performances (b) Sharpe ratios

Figure 4.5: Annualized performances and excess returns of BL as a function of L (a); Sharpe ratios and relative improvement (b).

4.5.4 Impact of the rebalancing frequency

We conclude this section with an overview on the impact of the rebalancing frequency on both BL and NL. This exercise is only theoretical since within the Base Case we only perform three trades (one each month before the next review). A lower rebalancing frequency of trading would mean that we do not fully exploit the power of learning by adjusting the portfolio. Higher rebalancing frequency carries a turnover (and transaction cost) issue, that would make any benefit only hypothetical. Furthermore, when we rebalance too often, we embark significant amount of noise coming from daily, short term, price movements. Usually, practitioners consider higher rebalancing frequencies only for a small portion of the portfolio, i.e. they only rebalance a fraction x of their portfolio while keeping the remaining $1-x$ unchanged, and they roll over. Because this goes beyond the scope of this paper, we limit ourselves to monthly versus bi-weekly rebalancing frequency. Table 4.7 collects summary statistics for BL and NL when the optimal weights in the Base Case are implemented on both monthly and bi-weekly frequencies.

Reb	Monthly		Bi-Weekly	
	BL	NL	BL	NL
Ann. Performance	5.96%	3.96%	5.42%	2.43%
Ann. Volatility	5.51%	4.92%	5.18%	4.93%
Max. Drawdown	-8.51%	-11.2%	-13.62%	-14.08%
Sharpe ratio	1.08	0.8	1.05	0.49
Sharpe ratio impr.	34.26%	-	112.01%	-
Information Ratio	0.55	-	0.77	-

Table 4.7: Statistics for BL and NL strategies with different rebalancing frequencies, monthly versus bi-weekly. The Sharpe ratio improvement is calculated as the relative difference between BL and NL Sharpe ratios for the same rebalancing frequency.

As we go from the monthly to the bi-weekly rebalancing frequency, we do observe a small decrease in annualized performance of BL, while the drop is more significant for NL. Although these numbers should not be taken as very informative due to the timing effect at work here, there is clearly not a strong incentive for BL to rebalance more often: performance goes slightly down as does volatility. In the end, the Sharpe ratio seems fairly stable. On the contrary, NL experiences a large drop in performance as well as an important increase in its maximum drawdown. Indeed, when we look at the structure of the strategy NL in (4.4), we see that the trade $\Sigma^{-1}b_0$ does not change at higher frequency. So, NL will implement the same trade (up to the maximum leverage). If for example the NL strategy was successful in the first two weeks, at the next rebalancing, it will most likely reverse this successful trade because of the $x_{t_0} - x_t$ part of $\mathbf{w}^{NL}$. Conversely, if the NL strategy was unsuccessful over a two-week period with a trade, at the next rebalancing it will increase this trade for the same reason. Therefore, NL tends to have a reversal feature at short horizon/high frequency.

As far as BL is concerned, because of uncertainty, the strategy can accommodate for larger deviation of the process B relatively to the forecast b_0, so that it does not necessarily have to reverse a successful trade or leverage an unsuccessful one. Indeed, when we look at (4.3), a successful trade will lower the leverage, but this can be compensated by the corrective term in the trade which depends on ψ and $\nabla_b R$, so that in theory BL has not, systematically, a reversal feature.

Finally, given our dataset and the fact that the momentum premium (see for example Carhart (1997), Blitz and Van Vliet (2008), Asness et al. (2013), and Geczy and Samonov (2017)) exists in the multi-asset framework, it is not surprising that NL experiences a drop in performance while BL is almost unaffected.

This is confirmed by the significant Sharpe ratio improvement: from a +34.26% in the Base Case, we improve the final Sharpe ratio at the bi-weekly frequency by 112.01% (mainly driven by the drop in NL Sharpe ratio). Furthermore, we see how BL is able to extract more information (or *alpha*) from the market because the Information ratio increases from 0.55 with the monthly rebalancing to 0.77 (a 40% increase) with the bi-weekly rebalancing.

4.6 Investing in foreign currencies

In our second example, we consider investing in different currencies: the Australian Dollar (AUD), the Canadian Dollar (CAD), the Euro (EUR), the British Pound (GBP), the Japanese Yen (JPY) and the U.S. Dollar (USD). Usually set by the central banks, the local risk-free interest rate, such as the federal funds rate in the U.S., together with the foreign exchange rate of currencies versus the Euro are the sources of performance for the investor. Therefore, the underlying assets available

to the investor are bank accounts in foreign banks that pay the local interest rate but are valued in EUR. Details are collected in Table 4.8.

Currency	Rate	Source
AUD	Thomson Reuters Australian Dollar Overnight Deposit	Thomson Reuters
CAD	Canada Money Market Overnight	Bank of Canada
EUR	Eonia (Euro OverNight Index Average)	European Banking Federation
GBP	United Kingdom Sonia	Wholesale Markets Brokers' Association
JPY	Japan Uncollateralized Overnight	Bank of Japan
USD	United States Federal Funds Effective Rate	Federal Reserve, United States

Table 4.8: Currency and their reference rates

The **Work Flow** detailed in Section 4.3 is implemented with the parameters listed in Table 4.9.

Parameter	Value
X_0	100
T	12 months
Lag	1 day
Freq	Weekly, Friday.
$r^{\mathbf{w}}$	60 days
$s^{\mathbf{w}}$	250 days
d	10%
unc	100
r^q	30
L	100%

Table 4.9: Parameters used in the implementation of BL and NL as defined in the **Work Flow** for the foreign currency strategy.

With respect to Table 4.2, here we consider a longer investment horizon (one year), we rebalance more often (weekly) because of reduced transaction costs in highly liquid foreign currencies, we consider short windows for estimating the drift parameter and we do not allow leverage.

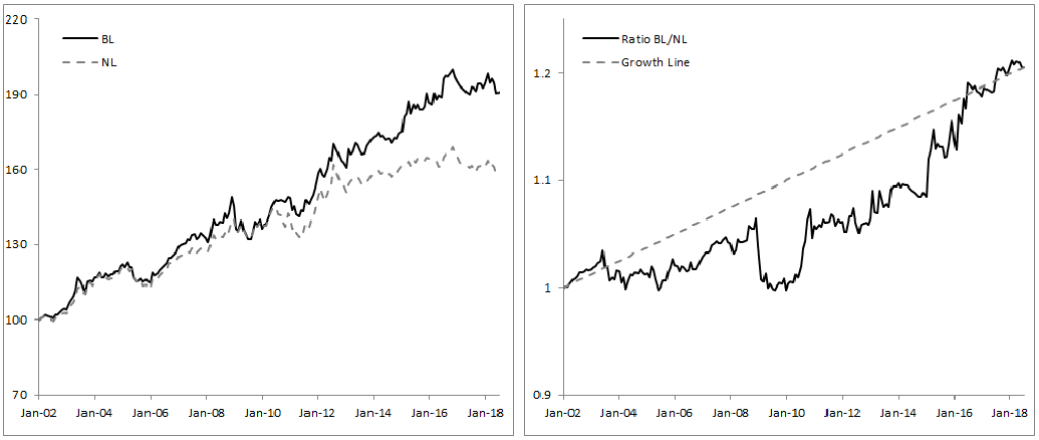
The result of the BL and NL strategies from January 2002 to June 2018 are reported in Table 4.10, while the historical performance is shown in Figure 4.6.

	BL	NL	Difference
Ann. Performance	4.00%	2.83%	1.17%
Ann. Volatility	5.97%	4.93%	1.04%
Max. Drawdown	-14.17%	-10.51%	-3.66%
Sharpe ratio	0.67	0.57	17.5%
Information Ratio	0.42	-	-

Table 4.10: Statistics for the BL and NL strategies. The Sharpe Ratio is calculated as the ratio between annualized performance and annualized volatility. For this one, the difference is given in relative terms. Data from January 2002 to June 2018.

Over the period, BL outperforms NL by 1.17% annualized. This is quite impressive given the low yields of developed countries currencies (mainly EUR and JPY) that

can usually be observed in the market. Therefore, the effect of learning clearly brings value to the investor by adding extra performance, although this comes at slightly higher risk (roughly 1% more volatile and 3.66% larger maximum drawdown). Nevertheless, the improvement in the Sharpe ratio is consistent. Figure 4.6b shows the relative strength of BL over NL and the growth line. We can see that, except for a few months in the second half of 2008 where BL strongly underperformed, it has since then outperformed NL regularly.



(a) Levels (b) Ratio

Figure 4.6: Historical values of the portfolios calculated with both $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ under the Base Case (a) and their ratio (b).

4.7 Investing in factor strategies

In the last few years, investors have embraced alternative strategies that target specific, well-established equity factors (such as Size, Value, Volatility or Momentum). These strategies offer an efficient and direct exposure to the main driving factors of equity markets and allow for an optimal allocation across factors. The main challenge is to invest in the right factor at the right time, as their performance usually shows cyclical patterns. Table 4.11 contains more details on the strategies we are considering. All of them are based on the U.S. large capitalization equity market.

Factor	Name	Source
Dividend	S&P 500 Dividend Aristocrats Net Total Return Index	S&P
Growth	MSCI USA Growth Net Total Return Index	MSCI
Momentum	MSCI USA Momentum USD Net Total Return Index	MSCI
Quality	MSCI USA Quality Net Total Return Index	MSCI
Size	MSCI USA Small Cap Net Total Return Index	MSCI
Value	Shiller Barclays CAPE US Sector Value Net TR Index	Barclays
Volatility	Ossiam US Minimum Variance Index Net Return/ ESG US Minimum Variance Index	S&P, Solactive
Cash	United States Federal Funds Effective Rate	Federal Reserve, United States

Table 4.11: Factor based strategies and the cash. For the volatility factor, we chained two strategies on December, 18 2009. Data in USD.

We follow the **Work Flow** of Section 4.3 with the parameters listed in Table 4.12 to derive both BL and NL strategies.

Parameter	Value
X_0	100
T	12 months
Lag	1 day
Freq	Monthly, 3^{rd} Friday of each month.
$r^{\mathbf{w}}$	90 days
$s^{\mathbf{w}}$	250 days
d	10%
unc	100
r^q	60
L	100%

Table 4.12: Parameters used in the implementation of the BL and NL factor rotation strategies as defined in the **Work Flow** .

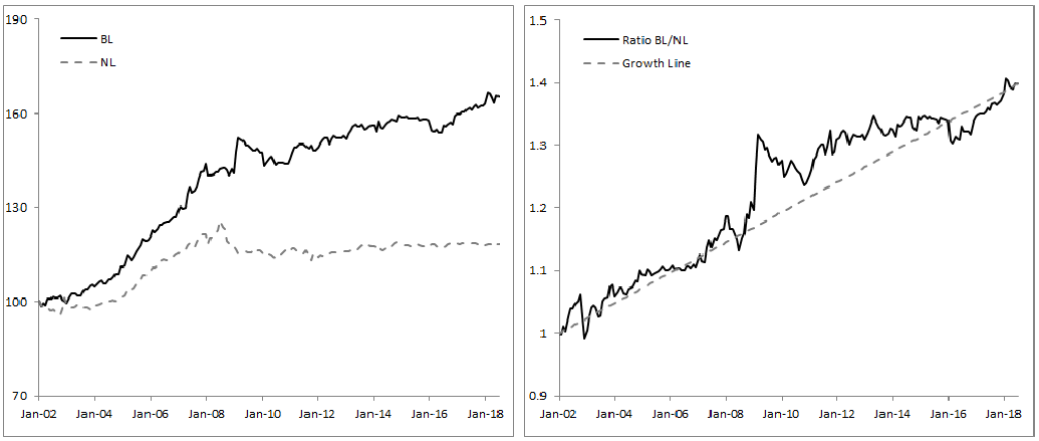
The results are shown in Table 4.13 from January 2002 to June 2018. In this example as well, the BL strategy has significantly outperformed the NL one. Here again, the learning effect clearly brings value to the investor. On an annualized basis, BL improves the performance by a bit more than 2%, with a lower maximum drawdown (-8.24% versus -10.43% for NL) but a slightly higher annualized volatility (3.61% for BL versus 2.26% for NL).

	BL	NL	Difference
Ann. Performance	3.1%	1.03%	2.07%
Ann. Volatility	3.61%	2.26%	1.34%
Max. Drawdown	-8.24%	-10.43%	2.2%
Sharpe ratio	0.86	0.45	91.11%
Information Ratio	0.55	-	-

Table 4.13: Statistics for the BL and NL strategies. The Sharpe Ratio is calculated as the ratio between annualized performance and annualized volatility. For this one, the difference is given in relative terms. Data from January 2002 to June 2018.

The improvement in the Sharpe ratio is particularly high: 0.86 for the BL factor rotation strategy versus 0.45 for the NL one (an improvement of 91.11%). Figure

4.7 shows the historical performance (4.7a) and the strength ratio (4.7b) of the BL strategy over NL. Clearly, learning over time produces regular outperformance, as it is shown by the regular upward trend of the ratio BL over NL.



(a) Levels (b) Ratio

Figure 4.7: Historical values of the portfolios calculated with both $\mathbf{w}^{BL}$ and $\mathbf{w}^{NL}$ under the Base Case (a) and their ratio (b).

4.A Appendices

We detail in this section, the indices we use in Chapter 4. A total return index measures the total return of the underlying index, combining both capital performance and income, mainly dividends. The net version of the previous index takes into account the different taxes associated to the income.

Contents

4.1 Introduction 110
4.2 The model 112
4.3 Market data 115
4.4 The Base Case result 118
4.5 Sensitivity analysis 119
4.5.1 Impact of uncertainty 119
4.5.2 Impact of leverage 121
4.5.3 Impact of the review frequency 123
4.5.4 Impact of the rebalancing frequency 125
4.6 Investing in foreign currencies 126

4.7 Investing in factor strategies 128
4.A Appendices 130
4.A.1 Dataset of Indices 131
4.A.1.1 Commodity 131
4.A.1.2 Bond 131
4.A.1.3 Equity 133
4.A.1.4 Cash 133
4.A.2 Dataset of currencies 134
4.A.3 Dataset of Smart Beta strategies 134
4.A.3.1 Smart Beta strategies 135

4.A.1 Dataset of Indices

4.A.1.1 Commodity

- *S&P GSCI Dynamic Roll Total Return Index.*[3] This index aims at reflecting the return of an investment in a world production-weighted portfolio comprised of the principal physical commodities that are the subject of active, liquid futures markets. The index employs a flexible and systematic futures contract rolling methodology which seeks to maximize yield from rolling long futures contracts in backwardated markets and minimize roll loss from rolling long futures contracts in contangoed markets.

- *LBMA Gold Price PM.*[4] The London Bullion Market Association (LBMA) publishes the gold price in U.S. dollars twice a day, at 10:30 AM and 3:00 PM (local time). Spot gold is traded for settlement two business days following the trade date, with a business day defined as a day when both the New York and London markets are open for business. LBMA Gold Price PM means the price per troy ounce of Gold stated in U.S. dollars as set via an electronic auction process and published at 3:00 PM.

4.A.1.2 Bond

1. **Corporate Bond Indices:**

[3] https://us.spindices.com/indices/commodities/sp-gsci-dynamic-roll
[4] http://www.lbma.org.uk/lbma-gold-price

- *Bloomberg Barclays Euro Aggregate Corporate Total Return Index.*[5] It is a benchmark that measures the corporate component of the Euro Aggregate Index. It includes investment grade, euro-denominated, fixed-rate securities.
- *iBoxx U.S. dollar Liquid High Yield Index.*[6] The Markit iBoxx U.S. dollar Liquid High Yield Index consists of liquid U.S. dollar high yield bonds, selected to provide a balanced representation of the broad U.S. dollar high yield corporate bond universe.
- *iBoxx U.S. dollar Liquid Investment Grade Index.*[7] The Markit iBoxx U.S. dollar Liquid Investment Grade Index is designed to provide a balanced representation of the U.S. dollar investment grade corporate market and to meet the investors demand for a U.S. dollar denominated, highly liquid and representative investment grade corporate index.
- *EUR Corporate Liquid High Yield.*[8] The Markit iBoxx EUR Liquid High Yield Index reflects the performance of the largest and most liquid corporate bonds, rated below investment grade denominated in Euro and issued by companies based in both within and outside Europe.

2. **Sovereign Bond Indices:**

- *Bloomberg Barclays U.S. Treasury Total Return Unhedged Index.*[9] The Bloomberg Barclays U.S. Treasury Index measures U.S. dollar-denominated, fixed-rate, nominal debt issued by the U.S. Treasury. Treasury bills are excluded by the maturity constraint.
- *Bloomberg Barclays EurAgg Treasury Total Return Unhedged Index.*[10] The Bloomberg Barclays Euro-Aggregate Treasury Index is a benchmark that measures the Treasury component of the Euro-Aggregate. The index consists of fixed-rate, investment grade public obligations of the sovereign countries in the Eurozone. This index currently contains euro-denominated issues from 17 countries.
- *JP Morgan EMBI Global Core Index.*[11] This index is a broad, diverse U.S. dollar-denominated emerging markets debt benchmark that tracks

[5] https://www.bloomberg.com/quote/LECPTREU:IND
[6] https://cdn.ihs.com/www/pdf/MKT-iBoxx-USD-Liquid-High-Yield-Index-factsheet.pdf
[7] https://cdn.ihs.com/www/pdf/MKT-iBoxx-USD-Liquid-Investment-Grade-Index-factsheet.pdf
[8] https://cdn.ihs.com/www/pdf/MKT-iBoxx-EUR-Liquid-High-Yield-Indices-factsheet.pdf
[9] https://www.bloomberg.com/quote/LUATTRUU:IND
[10] https://www.bloomberg.com/quote/LEATTRUU:IND
[11] https://www.jpmorgan.com/country/US/EN/jpmorgan/investbk/solutions/research/indices/product

the total return of actively traded debt instruments in emerging market countries.

- *FTSE Actuaries UK Conv. GILTs All Stocks TR Index.*[12] This is among the industry's most widely-used performance benchmarks for the UK Government bond market.

4.A.1.3 Equity

- *S&P 500 Net Total Return Index.*[13] The S&P 500® is a gauge of large-cap U.S. equities. There is over 9.9 trillion U.S. dollars indexed or benchmarked to the index, with indexed assets comprising approximately 3.4 trillion U.S. dollars of this total. The index includes 500 leading companies and covers approximately 80% of available market capitalization.

- *Euro Stoxx 50 Net Total Return Index.*[14] The Euro Stoxx 50 Index, Europe's leading blue-chip index for the Eurozone, provides a blue-chip representation of supersector leaders in the region. The index covers 50 stocks from 8 Eurozone countries: Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands and Spain.

- *FTSE 100 Total Return Index.*[15] FTSE 100 Index comprises the 100 most highly capitalised blue-chip companies listed on the London Stock Exchange and measures the performance of all capital and industry segments of the UK equity market.

- *MSCI Emerging Net Total Return Index.*[16] The MSCI Emerging Markets Index captures large and mid cap representation across 26 Emerging Markets countries. With 1,404 constituents, the index covers approximately 85% of the free float-adjusted market capitalization in each country.

4.A.1.4 Cash

- *EMMI EURO Overnight Index average.*[17] The euro overnight index average (EONIA) is a daily reference rate, published by the European Money Markets Institute (EMMI), representative of the rate at which credit institutions in the EU lend overnight funds to each other in the unsecured money market in euro.

[12] https://research.ftserussell.com/Analytics/FactSheets/temp/60927984-5c62-40c2-8211-17981b025e9a.pdf
[13] https://us.spindices.com/indices/equity/sp-500
[14] https://www.stoxx.com/index-details?symbol=sx5e
[15] https://www.ftserussell.com/products/indices/uk
[16] https://www.msci.com/emerging-markets
[17] https://www.emmi-benchmarks.eu/euribor-eonia-org/about-eonia.html

- *United States Federal Funds Effective Rate.*[18] The federal funds rate is the interest rate at which depository institutions trade federal funds (balances held at Federal Reserve Banks) with each other overnight. When a depository institution has surplus balances in its reserve account, it lends to other banks in need of larger balances. In simpler terms, a bank with excess cash, which is often referred to as liquidity, will lend to another bank that needs to quickly raise liquidity.

4.A.2 Dataset of currencies

We detail in this section, the currencies we use in Chapter 4. The local risk-free interest rates, usually set by central banks, together with the foreign exchange rate of currencies versus the Euro are the sources of performance for the investor. Consequently, the underlying assets considered are bank accounts in foreign banks that pay the local interest rate but are valued in Euro.

- *Australian Dollar Overnight Deposit*[19]
- *Canada Money Market Overnight*[20]
- *Eonia (Euro OverNight Index Average)*[21]
- *United Kingdom Sonia*[22]
- *Japan Uncollateralized Overnight*[23]
- *United States Federal Funds Effective Rate*[24]

4.A.3 Dataset of Smart Beta strategies

We detail in this section, the smart beta strategies we use in Chapter 4.

[18] https://fred.stlouisfed.org/series/FEDFUNDS
[19] https://www.rba.gov.au/statistics/cash-rate/
[20] https://www.bankofcanada.ca/rates/interest-rates/money-market-yields/
[21] https://www.euribor-rates.eu/en/eonia/
[22] https://www.bankofengland.co.uk/markets/sonia-benchmark
[23] https://www.boj.or.jp/en/announcements/education/oshiete/seisaku/b32.htm/
[24] https://fred.stlouisfed.org/series/FEDFUNDS

4.A.3.1 Smart Beta strategies

- *S&P 500 Dividend Aristocrats Net Total Return Index.*[25] S&P 500 Dividend Aristocrat index tracks the performance of well-known, mainly large cap, blue-chip companies with a float-adjusted market capitalization of at least 3 billion U.S dollars and an average trading volume of at least 5 million U.S dollars, in addition to consistently increasing dividend payments. The index typically contains 40 to 50 companies. The strength of the dividend aristocrats lies not just in the ability to continually increase dividend payments to shareholders, but also performance. These companies have historically outperformed the S&P 500 by a little more than 1% per year and exhibit slightly less volatility.

- *MSCI U.S.A. Growth Net Total Return Index.*[26] The MSCI U.S.A. Growth Index captures large and mid cap securities exhibiting overall growth style characteristics in the U.S. market. The growth investment style characteristics for index construction are defined using five variables: long-term forward EPS growth rate, short-term forward EPS growth rate, current internal growth rate, long-term historical EPS growth trend and long-term historical sales per share growth trend.

- *MSCI U.S.A. Momentum U.S. dollar Net Total Return Index.*[27] The MSCI U.S.A. Momentum Index is based on its parent index, the MSCI U.S.A. Index, which captures large and mid cap stocks of the U.S. market. It is designed to reflect the performance of an equity momentum strategy by emphasizing stocks with high price momentum, while maintaining reasonably high trading liquidity, investment capacity and moderate index turnover.

- *MSCI U.S.A. Quality Net Total Return Index.*[28] Based on the MSCI U.S.A. Index, this index includes large and mid cap stocks in the U.S. equity market. This index aims to capture the performance of quality growth stocks by identifying stocks with high quality scores based on three main fundamental variables: high return on equity (ROE), stable year over year earnings growth and low financial leverage.

- *MSCI U.S.A. Small Cap Net Total Return Index.*[29] The MSCI U.S.A. Small Cap Index is designed to measure the performance of the small cap segment of the U.S. equity market. With 1,747 constituents, the index represents ap-

[25] https://eu.spindices.com/indices/strategy/sp-500-dividend-aristocrats
[26] https://www.msci.com/documents/10199/5e550fca-978c-4ae2-a028-1117029cbe86
[27] https://www.msci.com/documents/10199/f3a22268-affd-478a-b7a7-50dc90fad923
[28] https://www.msci.com/documents/10199/4af921f5-0bbc-470b-ad69-19a177fad9cf
[29] https://www.msci.com/documents/10199/8038650a-0e6f-43d5-bdb0-1f8f3063e565

proximately 14% of the free float-adjusted market capitalization in the United-States.

- *Shiller Barclays CAPE U.S. Sector Value Net Total Return Index.*[30] The index reflects the performance of a long dynamic exposure to four American equity sectors chosen each month according to their relative CAPE, i.e. PER adjusted for cyclical variations, and their price variations during the twelve previous months. Exposure to the U.S. equity sectors is through the sectorial indices S&P Select, which reflect the performance of the U.S. companies included in the S&P 500 Index. The sector components of the index are selected each month according to the methodology of the Shiller Barclays CAPE series of indices. Each of the sub-indices chosen receives the same weighting during each rebalancing.

- *Ossiam U.S. Minimum Variance Index Net Return / ESG U.S. Minimum Variance Index.*[31] The U.S. ESG Minimum Variance NR Index reflects the performance of a dynamic selection of stocks which satisfy ESG, i.e. Environment, Social and Governance, criteria among the most liquid stocks from the Solactive U.S. Large Cap Index which tracks the performance of about 500 leading companies in major industries in the United-States. Constituents of the index will be selected on a monthly basis and weighted according to a minimum variance optimization procedure.

[30] https://www.ossiam.fr/produits/see/id/46#theTabs1
[31] https://www.ossiam.com/produits/see/id/28#theTabs3

Part II

Discrete-time portfolio optimization under maximum drawdown constraint with drift uncertainty

Chapter 5

Discrete-time portfolio optimization under maximum drawdown constraint with partial information and deep learning resolution

Abstract. We study a discrete-time portfolio selection problem with partial information and maximum drawdown constraint. Drift uncertainty in the multidimensional framework is modeled by a prior probability distribution. In this Bayesian framework, we derive the dynamic programming equation using an appropriate change of measure, and obtain semi-explicit results in the Gaussian case. The latter case, with a CRRA utility function is completely solved numerically using recent deep learning techniques for stochastic optimal control problems. We emphasize the informative value of the learning strategy versus the non-learning one by providing empirical performance and sensitivity analysis with respect to the uncertainty of the drift. Furthermore, we show numerical evidence of the close relationship between the non-learning strategy and a no short-sale constrained Merton problem, by illustrating the convergence of the former towards the latter as the maximum drawdown constraint vanishes.

Contents

5.1 Introduction **140**
5.2 Problem setup **142**
5.3 Dynamic programming system **143**
5.3.1 Change of measure and Bayesian filtering 143
5.3.2 The static set of admissible controls 145

5.3.3 Derivation of the dynamic programming equation 146
5.3.4 Special case: CRRA utility function 147
5.4 The Gaussian case 149
5.4.1 Bayesian Kalman filtering 149
5.4.2 Finite-dimensional dynamic programming equation 150
5.5 Deep learning numerical resolution 152
5.5.1 Architectures of the deep neural networks 152
5.5.2 *Hybrid-Now* algorithm 154
5.5.3 Numerical results 156
5.5.3.1 Learning and non-learning strategies 156
5.5.3.2 Learning, non-learning and constrained equally-weighted strategies 160
5.5.3.3 Non-learning and Merton strategies 161
5.5.4 Sensitivities analysis 163
5.6 Conclusion 171
5.A Appendices 171
5.A.1 Proof of Proposition 5.3.1 171
5.A.2 Proof of Proposition 5.3.2 172
5.A.3 Proof of Lemma 5.3.3 172
5.A.4 Proof of Lemma 5.3.4 173
5.A.5 Proof of Lemma 5.3.5 174
5.A.6 Proof of Lemma 5.3.6 174

5.1 Introduction

This paper is devoted to the study of a constrained allocation problem in discrete time with partial information. We consider an investor who is willing to maximize the expected utility of her terminal wealth over a given investment horizon. The risk-averse investor is looking for the optimal portfolio in financial assets under a maximum drawdown constraint. The maximum drawdown is a common metric in finance and represents the largest drop in the portfolio value. Our framework incorporates this constraint by setting a threshold representing the proportion of the current maximum of the wealth process that the investor is willing to keep.
The expected rate of assets' return (drift) is unknown but information can be learnt by progressive observation of the financial asset prices. The uncertainty about the

rate of return is modeled by a probability distribution, i.e., a prior belief on the drift. To take into account the information conveyed by the prices, this prior will be updated using a Bayesian learning approach.

An extensive literature exists on parameters uncertainty and especially on filtering and learning techniques in a partial information framework. To cite just a few, see Lakner (1998), Rogers (2001), Cvitanić et al. (2006), Karatzas and Zhao (2001), Bismuth et al. (2019), and De Franco, Nicolle, and Pham (2019a). Somme articles deal with risk constraints in a portfolio allocation framework. For instance, the paper by Redeker and Wunderlich (2018) tackles dynamic risk constraints and compares the continuous and discrete time trading while some papers especially focus on drawdown constraints, see in particular the seminal paper by Grossman and Zhou (1993) or Cvitanić and Karatzas (1994). More recently, the authors in Elie and Touzi (2008) study infinite-horizon optimal consumption-investment problem in continuous-time, and the paper Boyd et al. (2019) uses forecasts of the mean and covariance of financial returns from a multivariate hidden Markov model with time-varying parameters to build the optimal controls.

As it is not possible to solve analytically our constrained optimal allocation problem, we have applied a machine learning algorithm developed in Bachouch et al. (2018a) and Bachouch et al. (2018b). This algorithm, called *Hybrid-Now*, is particularly suited for solving stochastic control problems in high dimension using deep neural networks.

Our main contributions to the literature is twofold: a detailed theoretical study of a discrete-time portfolio selection problem including both drift uncertainty and maximum drawdown constraint, and a numerical resolution using a deep learning approach for an application to a model of three risky assets, leading to a five-dimensional problem. We derive the dynamic programming equation (DPE), which is in general of infinite-dimensional nature, following the change of measure suggested in Elliott et al. (2008). In the Gaussian case, the DPE is reduced to a finite-dimensional equation by exploiting the Kalman filter. In the particular case of constant relative risk aversion (CRRA) utility function, we reduce furthermore the dimensionality of the problem. Then, we solve numerically the problem in the Gaussian case with CRRA utility functions using the deep learning *Hybrid-Now* algorithm. Such numerical results allow us to provide a detailed analysis of the performance and allocations of both the learning and non-learning strategies benchmarked with a comparable equally-weighted strategy. Finally, we assess the performance of the learning compared to the non-learning strategy with respect to the sensitivity of the uncertainty of the drift. Additionally, we provide empirical evidence of convergence of the non-learning strategy to the solution of the classical Merton problem when the parameter controlling the maximum drawdown vanishes.

The paper is organized as follows: Section 5.2 sets up the financial market model and the associated optimization problem. Section 5.3 describes, in the general case, the change of measure and the Bayesian filtering, the derivation of the dynamic programming equation and details some properties of the value function. Section 5.4 focuses on the Gaussian case. Finally, Section 5.5 presents the neural network techniques used and shows the numerical results.

5.2 Problem setup

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a discrete filtration $(\mathcal{F}_k)_{k=0,\,\ldots,\,N}$ satisfying the usual conditions, we consider a financial market model with one riskless asset assumed normalized to one, and d risky assets. The price process $(S^i_k)_{k=0,\ldots,N}$ of asset $i \in [\![1, d]\!]$ is governed by the dynamics

$$S^i_{k+1} \;=\; S^i_k e^{R^i_{k+1}}, \quad k = 0, \ldots, N-1, \tag{5.1}$$

where $R_{k+1} = (R^1_{k+1}, \ldots, R^N_{k+1})$ is the vector of the assets log-return between time k and $k+1$, and modeled as:

$$R_{k+1} \;=\; B + \epsilon_{k+1}. \tag{5.2}$$

The drift vector B is a d-dimensional random variable with probability distribution (prior) μ_0 of known mean $b_0 = \mathbb{E}[B]$ and finite second order moment. Note that the case of known drift B means that μ_0 is a Dirac distribution. The noise $\epsilon = (\epsilon_k)_k$ is a sequence of centered i.i.d. random vector variables with covariance matrix $\Gamma = \mathbb{E}[\epsilon_k \epsilon_k']$, and assumed to be independent of B. We also assume the fundamental assumption that the probability distribution ν of ϵ_k admits a strictly positive density function g on $\mathbb{R}^d$ with respect to the Lebesgue measure.
The price process S is observable, and notice by relation (5.1) that R can be deduced from S, and vice-versa. We shall then denote by $\mathbb{F}^o = \{\mathcal{F}^o_k\}_{k=0,\,\ldots,\,N}$ the observation filtration generated by the process S (hence equivalently by R) augmented by the null sets of $\mathcal{F}$, with the convention that for $k = 0$, $\mathcal{F}^o_0$ is the trivial algebra.
An investment strategy is an $\mathbb{F}^o$-progressively measurable process $\alpha = (\alpha_k)_{k=0,\,\ldots,\,N-1}$, valued in $\mathbb{R}^d$, and representing the proportion of the current wealth invested in each of the d risky assets at each time $k = 0,\ \ldots,\ N-1$. Given an investment strategy α and an initial wealth $x_0 > 0$, the (self-financed) wealth process X^α evolves according to

$$\begin{cases} X^\alpha_{k+1} = X^\alpha_k \left(1 + \alpha_k' \left(e^{R_{k+1}} - \mathbb{1}_d\right)\right), & k = 0,\ \ldots,\ N-1, \\ X^\alpha_0 = x_0. \end{cases} \tag{5.3}$$

where $e^{R_{k+1}}$ is the d-dimensional random variable with components $\left[e^{R_{k+1}}\right]_i = e^{R^i_{k+1}}$ for $i \in [\![1, d]\!]$, and $\mathbb{1}_d$ is the vector in $\mathbb{R}^d$ with all components equal to 1.

Let us introduce the process Z_k^α, as the maximum up to time k of the wealth process X^α, i.e.,

$$Z_k^\alpha \;:=\; \max_{0\le \ell\le k} X_\ell^\alpha, \qquad k=0,\ \ldots,\ N.$$

The maximum drawdown constraints the wealth X_k^α to remain above a fraction $q\in(0,1)$ of the current historical maximum Z_k^α. We then define the set of *admissible* investment strategies $\mathcal{A}_0^q$ as the set of investment strategies α such that

$$X_k^\alpha \;\ge\; qZ_k^\alpha, \quad \text{a.s.}, \quad k=0,\ldots,N.$$

In this framework, the portfolio selection problem is formulated as

$$V_0 \;:=\; \sup_{\alpha\in\mathcal{A}_0^q} \mathbb{E}\left[U\left(X_N^\alpha\right)\right], \tag{5.4}$$

where U is a utility function on $(0,\infty)$ satisfying the standard Inada conditions: continuously differentiable, strictly increasing, concave on $(0,\infty)$ with $U'(0)=\infty$ and $U'(\infty)=0$.

5.3 Dynamic programming system

In this section, we show how Problem (5.4) can be characterized from dynamic programming in terms of a backward system of equations amenable for algorithms. In a first step, we shall update the prior on the drift uncertainty, and take advantage of the newest available information by adopting a Bayesian filtering approach. This relies on a suitable change of probability measure.

5.3.1 Change of measure and Bayesian filtering

We start by introducing a change of measure under which R_1,..., R_N are mutually independent, identically distributed random variables and independent from the drift B, hence behaving like a noise. Following the methodology detailed in (Elliott et al., 2008) we define the σ-algebras

$$\mathcal{G}_k^0 := \sigma(B,R_1,\ldots,R_k), \quad k=1,\ldots,N,$$

and $\mathbb{G}=(\mathcal{G}_k)_k$ the corresponding complete filtration. We then define a new probability measure $\overline{\mathbb{P}}$ on $(\Omega,\bigvee_{k=1}^N\mathcal{G}_k)$ by

$$\left.\frac{\mathrm{d}\overline{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_k} \;:=\; \Lambda_k, \quad k=1,\ldots,N,$$

with

$$\Lambda_k \;:=\; \prod_{\ell=1}^{k} \frac{g(R_\ell)}{g(\epsilon_\ell)}, \quad k=1,\dots,N, \;\; \Lambda_0 \;=\; 1.$$

The existence of $\overline{\mathbb{P}}$ comes from the Kolmogorov's theorem since Λ_k is a strictly positive martingale with expectation equal to one. Indeed, for all $k = 1, ..., N$,

- $\Lambda_k > 0$ since the probability density function g is strictly positive
- Λ_k is $\mathcal{G}_k$-adapted,
- As $\epsilon_k \perp\!\!\!\perp \mathcal{G}_{k-1}$, we have

$$\begin{aligned}\mathbb{E}[\Lambda_k|\mathcal{G}_{k-1}] \;&=\; \Lambda_{k-1}\mathbb{E}\Big[\frac{g(B+\epsilon_k)}{g(\epsilon_k)}\Big|\mathcal{G}_{k-1}\Big] \\ &=\; \Lambda_{k-1}\int_{\mathbb{R}^d}\frac{g(B+e)}{g(e)}g(e)de \;=\; \Lambda_{k-1}\int_{\mathbb{R}^d} g(z)dz \;=\; \Lambda_{k-1}.\end{aligned}$$

Proposition 5.3.1. *Under $\overline{\mathbb{P}}$, $(R_k)_{k=1,\dots,N}$, is a sequence of i.i.d. random variables, independent from B, having the same probability distribution ν as ϵ_k.*

Proof. See Appendix 5.A.1. □

Conversely, we recover the initial measure $\mathbb{P}$ under which $(\epsilon_k)_{k=1,\dots,N}$ is a sequence of independent and identically distributed random variables having probability density function g where $\epsilon_k = R_k - B$. Denoting by $\overline{\Lambda}_k$ the Radon-Nikodym derivative $\mathrm{d}\mathbb{P}/\mathrm{d}\overline{\mathbb{P}}$ restricted to the σ-algebra $\mathcal{G}_k$:

$$\left.\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\overline{\mathbb{P}}}\right|_{\mathcal{G}_k} = \overline{\Lambda}_k,$$

we have

$$\overline{\Lambda}_k \;=\; \prod_{\ell=1}^{k}\frac{g(R_\ell - B)}{g(R_\ell)}.$$

It is clear that, under $\mathbb{P}$, the return and wealth processes have the form stated in equations (5.2) and (5.3). Moreover, from Bayes formula, the posterior distribution of the drift, i.e. the conditional law of B given the asset price observation, is

$$\mu_k(db) \;:=\; \mathbb{P}\big[B \in db|\mathcal{F}_k^o\big] \;=\; \frac{\pi_k(db)}{\pi_k(\mathbb{R}^d)}, \quad k=0,\dots,N, \tag{5.5}$$

where π_k is the so-called unnormalized conditional law

$$\pi_k(db) := \overline{\mathbb{E}}\big[\overline{\Lambda}_k \mathbb{1}_{\{B\in db\}} | \mathcal{F}_k^o\big], \quad k = 0, \dots, N.$$

We then have the key recurrence linear relation on the unnormalized conditional law.

Proposition 5.3.2. *We have the recursive linear relation*

$$\pi_k \;=\; \bar{g}(R_k - \cdot)\pi_{k-1}, \quad k = 1, \dots, N, \tag{5.6}$$

with initial condition $\pi_0 = \mu_0$, *where*

$$\bar{g}(R_k - b) \;=\; \frac{g(R_k - b)}{g(R_k)}, \quad b \in \mathbb{R}^d,$$

and we recall that g *is the probability density function of the identically distributed* ϵ_k *under* $\mathbb{P}$.

Proof. See Appendix 5.A.2 . □

5.3.2 The static set of admissible controls

In this subsection, we derive some useful characteristics of the space of controls which will turn to be crucial in the derivation of the dynamic programming system. Given time $k \in [\![0, N]\!]$, a current wealth $x = X_k^\alpha > 0$, and current maximum wealth $z = Z_k^\alpha \geq x$ that satisfies the drawdown constraint $qz \leq x$ at time k for an admissible investment strategy $\alpha \in \mathcal{A}_0^q$, we denote by $A_k^q(x, z) \subset \mathbb{R}^d$ the set of static controls $a = \alpha_k$ such that the drawdown constraint is satisfied at next time $k+1$, i.e. $X_{k+1}^\alpha \geq qZ_{k+1}^\alpha$. From the relation (5.3), and noting that $Z_{k+1}^\alpha = \max[Z_k^\alpha, X_{k+1}^\alpha]$, this yields

$$\begin{aligned} &A_k^q(x, z) = \\ &\quad \Big\{ a \in \mathbb{R}^d : 1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big) \;\geq\; q \max\Big[\frac{z}{x}, 1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big)\Big] \text{ a.s.}\Big\}. \end{aligned} \tag{5.7}$$

Recalling from Proposition 5.3.1, that the random variables $R_1, ..., R_N$ are i.i.d. under $\overline{\mathbb{P}}$, we notice that the set $A_k^q(x, z)$ does not depend on the current time k, and we shall drop the subscript k in the sequel, and simply denote by $A^q(x, z)$.

Remembering that the support of ν, the probability distribution of ϵ_k, is $\mathbb{R}^d$, the following lemma characterizes more precisely the set $A^q(x, z)$.

Lemma 5.3.3. *For any* $(x, z) \in \mathcal{S}^q := \big\{(x, z) \in (0, \infty)^2 : qz \leq x \leq z\big\}$, *we have*

$$A^q(x, z) \;=\; \Big\{ a \in \mathbb{R}^d_+ : |a|_1 \leq 1 - q\frac{z}{x} \Big\},$$

where $|a|_1 = \sum_{i=1}^d |a_i|$ *for* $a = (a_1, \dots, a_d) \in \mathbb{R}^d_+$.

Proof. See Appendix 5.A.3. □

Let us prove some properties on the admissible set $A^q(x,z)$.

Lemma 5.3.4. *For any $(x,z) \in \mathcal{S}^q$, the set $A^q(x,z)$ satisfies the following properties:*

1. *It is decreasing in q: $\forall q_1 \leq q_2$, $A^{q_2}(x,z) \subseteq A^{q_1}(x,z)$,*
2. *It is continuous in q,*
3. *It is increasing in x: $\forall x_1 \leq x_2$, $A^q(x_1,z) \subseteq A^q(x_2,z)$,*
4. *It is a convex set,*
5. *It is homogeneous: $a \in A^q(x,z) \Leftrightarrow a \in A^q(\lambda x, \lambda z)$, for any $\lambda > 0$.*

Proof. See Appendix 5.A.4. □

5.3.3 Derivation of the dynamic programming equation

The change of probability detailed in Subsection 5.3.1 allows us to turn the initial partial information Problem (5.4) into a full observation problem as

$$\begin{aligned} V_0 := \sup_{\alpha \in \mathcal{A}_0^q} \mathbb{E}[U(X_N^\alpha)] &= \sup_{\alpha \in \mathcal{A}_0^q} \overline{\mathbb{E}}[\overline{\Lambda}_N U(X_N^\alpha)] \\ &= \sup_{\alpha \in \mathcal{A}_0^q} \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\big[\overline{\Lambda}_N U(X_N^\alpha)\big|\mathcal{F}_N^o\big]\Big] \\ &= \sup_{\alpha \in \mathcal{A}_0^q} \overline{\mathbb{E}}\Big[U(X_N^\alpha)\pi_N(\mathbb{R}^d)\Big], \end{aligned} \tag{5.8}$$

from Bayes formula, the law of conditional expectations, and the definition of the unnormalized filter π_N valued in $\mathcal{M}_+$, the set of nonnegative measures on $\mathbb{R}^d$. In view of Equation (5.3), Proposition 5.3.1, and Proposition 5.3.2, we then introduce the dynamic value function associated to Problem (5.8) as

$$v_k(x,z,\mu) = \sup_{\alpha \in \mathcal{A}_k^q(x,z)} J_k(x,z,\mu,\alpha), \quad k \in [\![0,N]\!],\ (x,z) \in \mathcal{S}^q,\ \mu \in \mathcal{M}_+,$$

with

$$J_k(x,z,\mu,\alpha) = \overline{\mathbb{E}}\Big[U\big(X_N^{k,x,\alpha}\big)\pi_N^{k,\mu}(\mathbb{R}^d)\Big],$$

where $X^{k,x,\alpha}$ is the solution to Equation (5.3) on $[\![k,N]\!]$, starting at $X_k^{k,x,\alpha} = x$ at time k, controlled by $\alpha \in \mathcal{A}_k^q(x,z)$, and $(\pi_\ell^{k,\mu})_{\ell=k,\dots,N}$ is the solution to (5.6)

on $\mathcal{M}_+$, starting from $\pi_k^{k,\mu} = \mu$, so that $V_0 = v_0(x_0, x_0, \mu_0)$. Here, $\mathcal{A}_k^q(x,z)$ is the set of admissible investment strategies embedding the drawdown constraint: $X_\ell^{k,x,\alpha} \geq qZ_\ell^{k,x,z,\alpha}$, $\ell = k, \ldots, N$, where the maximum wealth process $Z^{k,x,z,\alpha}$ follows the dynamics: $Z_{\ell+1}^{k,x,z,\alpha} = \max[Z_\ell^{k,x,z,\alpha}, X_{\ell+1}^{k,x,\alpha}]$, $\ell = k, \ldots, N-1$, starting from $Z_k^{k,x,z,\alpha} = z$ at time k. The dependence of the value function upon the unnormalized filter μ means that the probability distribution on the drift is updated at each time step from Bayesian learning by observing assets price.

The dynamic programming equation associated to (5.8) is then written in backward induction as

$$\left\{ \begin{array}{rcl} v_N(x,z,\mu) & = & U(x)\mu(\mathbb{R}^d), \\ v_k(x,z,\mu) & = & \sup\limits_{a \in A^q(x,z)} \overline{\mathbb{E}}\Big[v_{k+1}\big(X_{k+1}^{k,x,a}, Z_{k+1}^{k,z,a}, \pi_{k+1}^{k,\mu}\big)\Big], \quad k = 0, \ldots, N-1. \end{array} \right.$$

Recalling Proposition 5.3.2 and Lemma 5.3.3, this dynamic programming system is written more explicitly as

$$\left\{ \begin{array}{rcl} v_N(x,z,\mu) & = & U(x)\mu(\mathbb{R}^d), \quad (x,z) \in \mathcal{S}^q, \ \mu \in \mathcal{M}_+, \\ v_k(x,z,\mu) & = & \sup\limits_{a \in A^q(x,z)} \overline{\mathbb{E}}\Big[v_{k+1}\Big(x\big(1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big)\big), \\ & & \qquad \max\big[z, x\big(1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big)\big)\big], \bar{g}(R_{k+1} - \cdot)\mu\Big)\Big], \end{array} \right. \tag{5.9}$$

for $k = 0, \ldots, N-1$. Notice from Proposition 5.3.1 that the expectation in the above formula is only taken with respect to the noise R_{k+1}, which is distributed under $\overline{\mathbb{P}}$ according to the probability distribution ν with density g on $\mathbb{R}^d$.

5.3.4 Special case: CRRA utility function

In the case where the utility function is of CRRA (Constant Relative Risk Aversion) type, i.e.,

$$U(x) \ = \ \frac{x^p}{p}, \quad x > 0, \quad \text{for some } 0 < p < 1, \tag{5.10}$$

one can reduce the dimensionality of the problem. For this purpose, we introduce the process $\rho = (\rho_k)_k$ defined as the ratio of the wealth over its maximum up to current as:

$$\rho_k^\alpha \ = \ \frac{X_k^\alpha}{Z_k^\alpha}, \quad k = 0, \ldots, N.$$

This ratio process lies in the interval $[q, 1]$ due to the maximum drawdown constraint. Moreover, recalling (5.3), and observing that $Z_{k+1}^\alpha = \max[Z_k^\alpha, X_{k+1}^\alpha]$, together with

the fact that $\frac{1}{\max[z,x]} = \min[\frac{1}{z}, \frac{1}{x}]$, we notice that the ratio process ρ can be written in inductive form as

$$\rho^{\alpha}_{k+1} \quad = \quad \min\Big[1, \rho^{\alpha}_{k}\big(1 + \alpha'_k\big(e^{R_{k+1}} - \mathbb{1}_d\big)\big)\Big], \quad k = 0, \dots, N-1.$$

The following result states that the value function inherits the homogeneity property of the utility function.

Lemma 5.3.5. *For a utility function U as in (5.10), we have for all $(x, z) \in \mathcal{S}^q$, $\mu \in \mathcal{M}_+$, $k \in [\![0, N]\!]$,*

$$v_k(\lambda x, \lambda z, \mu) \quad = \quad \lambda^p v_k(x, z, \mu), \quad \lambda > 0.$$

Proof. See Appendix 5.A.5. □

In view of the above Lemma, we consider the sequence of functions w_k, $k \in [\![0, N]\!]$, defined by

$$w_k(r, \mu) \quad = \quad v_k(r, 1, \mu), \quad r \in [q, 1], \ \mu \in \mathcal{M}_+,$$

so that $v_k(x, z, \mu) = z^p w_k(\frac{x}{z}, \mu)$, and we call w_k the reduced value function. From the dynamic programming system satisfied by v_k, we immediately obtain the backward system for $(w_k)_k$ as

$$\left\{ \begin{array}{rcl} w_N(r, \mu) & = & \frac{r^p}{p}\mu(\mathbb{R}^d), \quad r \in [q, 1], \ \mu \in \mathcal{M}_+, \\ w_k(r, \mu) & = & \sup\limits_{a \in A^q(r)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min\big[1, r\big(1 + a'\big(e^{R_{k+1}} - \mathbb{1}_d\big)\big)\big], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big], \end{array} \right. \tag{5.11}$$

for $k = 0, \dots, N-1$, where

$$A^q(r) \quad = \quad \Big\{a \in \mathbb{R}^d_+ : a'\mathbb{1}_d \leq 1 - \frac{q}{r}\Big\}.$$

We end this section by stating some properties on the reduced value function.

Lemma 5.3.6. *For any $k \in [\![0, N]\!]$, the reduced value function w_k is nondecreasing and concave in $r \in [q, 1]$.*

Proof. See proof in Appendix 5.A.6. □

5.4 The Gaussian case

We consider in this section the Gaussian framework where the noise and the prior belief on the drift are modeled according to a Gaussian distribution. In this special case, the Bayesian filtering is simplified into the Kalman filtering, and the dynamic programming system is reduced to a finite-dimensional problem that will be solved numerically. It is convenient to deal directly with the posterior distribution of the drift, i.e. the conditional law of the drift B given the assets price observation, also called normalized filter. From (5.5) and Proposition 5.3.2, it is given by the inductive relation

$$\mu_k(db) \quad = \quad \frac{g(R_k - b)\mu_{k-1}(db)}{\int_{\mathbb{R}^d} g(R_k - b)\mu_{k-1}(db)}, \quad k = 1, \ldots, N. \tag{5.12}$$

5.4.1 Bayesian Kalman filtering

We assume that the probability law ν of the noise ϵ_k is Gaussian: $\mathcal{N}(0, \Gamma)$, and so with density function

$$g(r) \quad = \quad (2\pi)^{-\frac{d}{2}}|\Gamma|^{-\frac{1}{2}}e^{-\frac{1}{2}r'\Gamma^{-1}r}, \quad r \in \mathbb{R}^d. \tag{5.13}$$

Assuming also that the prior distribution μ_0 on the drift B is Gaussian with mean b_0, and invertible covariance matrix Σ_0, we deduce by induction from (5.12) that the posterior distribution μ_k is also Gaussian: $\mu_k \sim \mathcal{N}(\hat{B}_k, \Sigma_k)$, where $\hat{B}_k = \mathbb{E}[B|\mathcal{F}_k^o]$ and Σ_k satisfy the well-known inductive relations:

$$\hat{B}_{k+1} \quad = \quad \hat{B}_k + K_{k+1}(R_{k+1} - \hat{B}_k), \quad k = 0, \ldots, N-1 \tag{5.14}$$

$$\Sigma_{k+1} \quad = \quad \Sigma_k - \Sigma_k(\Sigma_k + \Gamma)^{-1}\Sigma_k, \tag{5.15}$$

where K_{k+1} is the so-called Kalman gain given by

$$K_{k+1} \quad = \quad \Sigma_k(\Sigma_k + \Gamma)^{-1}, \quad k = 0, \ldots, N-1. \tag{5.16}$$

We have the initialization $\hat{B}_0 = b_0$, and the notation for Σ_k is coherent at time $k = 0$ as it corresponds to the covariance matrice of B. While the Bayesian estimation $\hat{B}_k$ of B is updated from the current observation of the log-return R_k, notice that Σ_k (as well as K_k) is deterministic, and is then equal to the covariance matrix of the error between B and its Bayesian estimation, i.e. $\Sigma_k = \mathbb{E}[(B - \hat{B}_k)(B - \hat{B}_k)']$. Actually, we can explicitly compute Σ_k by noting from Equation (5.12) with g as in (5.13) and $\mu_0 \sim \mathcal{N}(b_0, \Sigma_0)$ that

$$\mu_k \sim \frac{e^{-\frac{1}{2}\left(b-\left(\Sigma_0^{-1}+\Gamma^{-1}k\right)^{-1}\left(\Gamma^{-1}\sum_{j=1}^k R_j+\Sigma_0^{-1}b_0\right)\right)\left(\Sigma_0^{-1}+\Gamma^{-1}k\right)\left(b-\left(\Sigma_0^{-1}+\Gamma^{-1}k\right)^{-1}\left(\Gamma^{-1}\sum_{j=1}^k R_j+\Sigma_0^{-1}b_0\right)\right)}}{(2\pi)^{\frac{d}{2}}|(\Sigma_0^{-1} + \Gamma^{-1}k)^{-1}|^{\frac{1}{2}}}.$$

By identification, we then get

$$\Sigma_k = (\Sigma_0^{-1} + \Gamma^{-1}k)^{-1} = \Sigma_0(\Gamma + \Sigma_0 k)^{-1}\Gamma. \tag{5.17}$$

Moreover, the innovation process $(\tilde{\epsilon}_k)_k$, defined as

$$\tilde{\epsilon}_{k+1} \;=\; R_{k+1} - \mathbb{E}[R_{k+1}|\mathcal{F}_k^o] \;=\; R_{k+1} - \hat{B}_k, \quad k = 0, \ldots, N-1, \tag{5.18}$$

is a $\mathbb{F}^o$-adapted Gaussian process. Each $\tilde{\epsilon}_{k+1}$ is independent of $\mathcal{F}_k^0$ (hence $\tilde{\epsilon}_k$, $k = 1, \ldots, N$ are mutually independent), and is a centered Gaussian vector with covariance matrix:

$$\tilde{\epsilon}_{k+1} \;\sim\; \mathcal{N}\big(0, \tilde{\Gamma}_{k+1}\big), \qquad \text{with } \tilde{\Gamma}_{k+1} \;=\; \Sigma_k + \Gamma.$$

We refer to Kalman (1960) and Kalman and Bucy (1961) for these classical properties about the Kalman filtering and the innovation process.

Remark 5.4.1. From (5.14), and (5.18), we see that the Bayesian estimator $\hat{B}_k$ follows the dynamics

$$\begin{cases} \hat{B}_{k+1} = \hat{B}_k + K_{k+1}\tilde{\epsilon}_{k+1}, & k = 0, \ldots, N-1 \\ \hat{B}_0 = b_0, \end{cases}$$

which implies in particular that $\hat{B}_k$ has a Gaussian distribution with mean b_0, and covariance matrix satisfying

$$\mathrm{Var}(\hat{B}_{k+1}) = \mathrm{Var}(\hat{B}_k) + K_{k+1}(\Sigma_k + \Gamma)K_{k+1}' \;=\; \mathrm{Var}(\hat{B}_k) + \Sigma_k(\Sigma_k + \Gamma)^{-1}\Sigma_k.$$

Recalling the inductive relation (5.15) on Σ_k, this shows that $\mathrm{Var}(\hat{B}_k) = \Sigma_0 - \Sigma_k$. Note that, from Equation (5.15), $(\Sigma_k)_k$ is a decreasing sequence which ensures that $\mathrm{Var}(\hat{B}_k)$ is positive semi-definite and is nondecreasing with time k. ♢

5.4.2 Finite-dimensional dynamic programming equation

From (5.18), we see that our initial portfolio selection Problem (5.4) can be reformulated as a full observation problem with state dynamics given by

$$\begin{cases} X_{k+1}^\alpha &= X_k^\alpha\Big(1 + \alpha_k'\big(e^{\hat{B}_k + \tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\Big), \\ \hat{B}_{k+1} &= \hat{B}_k + K_{k+1}\tilde{\epsilon}_{k+1}, \quad k = 0, \ldots, N-1. \end{cases} \tag{5.19}$$

We then define the value function on $[\![0, N]\!] \times \mathcal{S}^q \times \mathbb{R}^d$ by

$$\tilde{v}_k(x, z, b) \;=\; \sup_{\alpha \in \mathcal{A}_k^q(x,z)} \mathbb{E}\big[U(X_N^{k,x,b,\alpha})\big], \quad k \in [\![0, N]\!],\ (x, z) \in \mathcal{S}^q,\ b \in \mathbb{R}^d,$$

where the pair $(X^{k,x,b,\alpha}, \hat{B}^{k,b})$ is the process solution to (5.19) on $[\![k, N]\!]$, starting from (x, b) at time k, so that $V_0 = \tilde{v}_0(x_0, x_0, b_0)$. The associated dynamic programming system satisfied by the sequence $(\tilde{v}_k)_k$ is

$$\begin{cases} \tilde{v}_N(x,z,b) &= U(x), \quad (x,z) \in \mathcal{S}^q,\ b \in \mathbb{R}^d, \\ \tilde{v}_k(x,z,b) &= \sup\limits_{a \in A^q(x,z)} \mathbb{E}\Big[\tilde{v}_{k+1}\Big(x\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big), \\ & \qquad\qquad \max\big[z, x\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big)\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

for $k = 0, \ldots, N-1$. Notice that in the above formula, the expectation is taken with respect to the innovation vector $\tilde{\epsilon}_{k+1}$, which is distributed according to $\mathcal{N}(0, \tilde{\Gamma}_{k+1})$. Moreover, in the case of CRRA utility functions $U(x) = x^p/p$, and similarly as in Section 5.3.4, we have the dimension reduction with

$$\tilde{w}_k(r,b) \quad = \quad \tilde{v}_k(r,1,b), \quad r \in [q,1],\ b \in \mathbb{R}^d,$$

so that $\tilde{v}_k(x,z,b) = z^p \tilde{w}_k(\frac{x}{z}, b)$, and this reduced value function satisfies the backward system on $[q,1] \times \mathbb{R}^d$:

$$\begin{cases} \tilde{w}_N(r,b) &= \frac{r^p}{p}, \quad r \in [q,1],\ b \in \mathbb{R}^d, \\ \tilde{w}_k(r,b) &= \sup\limits_{a \in A^q(r)} \mathbb{E}\Big[\tilde{w}_{k+1}\Big(\min\big[1, r\big(1 + a'\big(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\big)\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

for $k = 0, \ldots, N-1$.

Remark 5.4.2 (No short-sale constrained Merton problem)**.** In the limiting case when $q = 0$, the drawdown constraint is reduced to a non-negativity constraint on the wealth process, and by Lemma 5.3.3, this means a no-short selling and no borrowing constraint on the portfolio strategies. When the drift B is also known, equal to b_0, and for a CRRA utility function, let us then consider the corresponding constrained Merton problem with value function denoted by v_k^M, $k = 0, \ldots, N$, which satisfies the standard backward recursion from dynamic programming:

$$\begin{cases} v_N^M(x) = \dfrac{x^p}{p}, \quad x > 0, \\ v_k^M(x) = \sup\limits_{\substack{a'\mathbb{1}_d \le 1 \\ a \in [0,1]^d}} \mathbb{E}\Big[v_{k+1}^M\big(x\big(1 + a'\big(e^{b_0+\epsilon_{k+1}} - \mathbb{1}_d\big)\big)\Big], \quad k = 0, \ldots, N-1. \end{cases} \tag{5.20}$$

By searching for a solution of the form $v_k^M(x) = K_k x^p/p$, with $K_k \ge 0$ for all $k \in [\![0, N]\!]$, we see that the sequence $(K_k)_k$ satisfies the recursive relation:

$$K_k \quad = \quad SK_{k+1}, \quad k = 0, \ldots, N-1,$$

starting from $K_N = 1$, where

$$S := \sup_{\substack{a'\mathbb{1}_d \leq 1 \\ a\in[0,1]^d}} \mathbb{E}\Big[\Big(1 + a'\big(e^{b_0+\epsilon_1} - \mathbb{1}_d\big)\Big)^p\Big],$$

by recalling that $\epsilon_1, \ldots, \epsilon_N$ are i.i.d. random variables. It follows that the value function of the constrained Merton problem, unique solution to the dynamic programming system (5.20), is equal to

$$v_k^M(x) = S^{N-k}\frac{x^p}{p}, \quad k = 0, \ldots, N,$$

and the constant optimal control is given by

$$a_k^M = \operatorname*{argmax}_{\substack{a'\mathbb{1}\leq 1 \\ a\in[0,1]^d}} \mathbb{E}\left[\left(1 + a'\left(e^{R_1} - \mathbb{1}_d\right)\right)^p\right] \quad k = 0, \ldots, N-1.$$

$\Diamond$

5.5 Deep learning numerical resolution

In this section, we exhibit numerical results to promote the benefits of learning from new information. To this end, we compare the learning strategy (Learning) to the non-learning one (Non-Learning) in the case of the CRRA utility function and the Gaussian distribution for the noise. The prior probability distribution of B is the Gaussian distribution $\mathcal{N}(b_0, \Sigma_0)$ for Learning while it is the Dirac distribution concentrated at b_0 for Non-Learning.
We use deep neural network techniques to compute numerically the optimal solutions for both Learning and Non-Learning. To broaden the analysis, in addition to the learning and non-learning strategies, we have computed an "admissible" equally weighted (EW) strategy. More precisely, this EW strategy will share the quantity $X_k - qZ_k$ equally among the d assets. Eventually, we show numerical evidence that the Non-Learning converges to the optimal strategy of the constrained Merton problem, when the loss aversion parameter q vanishes.

5.5.1 Architectures of the deep neural networks

Neural networks (NN) are able to approximate nonlinear continuous functions, typically the value function and controls of our problem. The principle is to use a large amount of data to train the NN so that it progressively comes close to the target function. It is an iterative process in which the NN is tuned on a training set, then

tested on a validation set to avoid over-fitting. For more details, see for instance Hornik (1991) and Géron (2019).

The algorithm we use, relies on two dense neural networks: the first one is dedicated to the controls (A_{NN}) and the second one to the value function (VF_{NN}). Each NN is composed of four layers: an input layer, two hidden layers and an output layer:

(i) The input layer is $d+1$-dimensional since it embeds the conditional expectations of each of the d assets and the ratio of the current wealth to the current historical maximum ρ.

(ii) The two hidden layers give the NN the flexibility to adjust its weights and biases to approximate the solution. From numerical experiments, we see that, given the complexity of our problem, a first hidden layer with $d+20$ neurons and a second one with $d+10$ are a good compromise between speed and accuracy.

(iii) The output layer is d-dimensional for the controls, one for each asset representing the weight of the instrument, and is one-dimensional for the value function. See Figures 5.1 and 5.2 for an overview of the NN architectures in the case of $d=3$ assets.

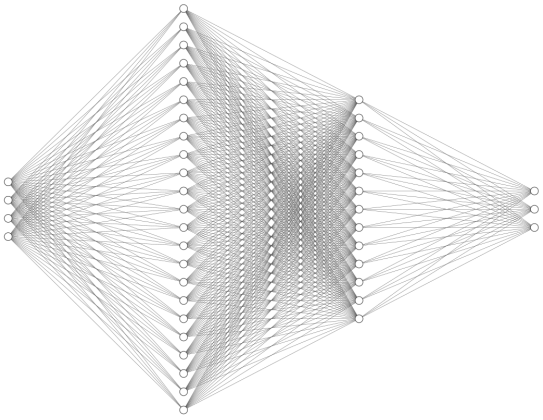


Figure 5.1: A_{NN} architecture with $d=3$ assets

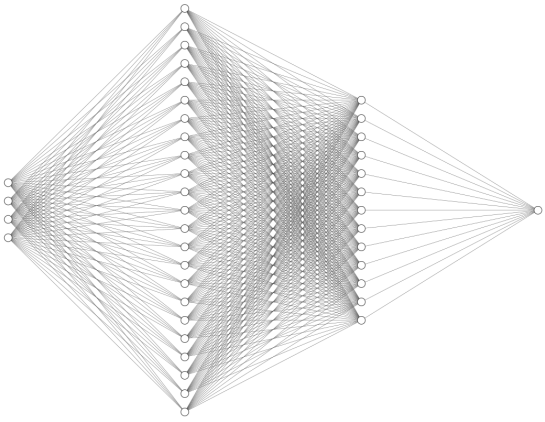


Figure 5.2: VF_{NN} architecture with $d=3$ assets

We follow the indications in Géron (2019) to setup and define the values of the various inputs of the neural networks which are listed in Table 5.1.

	$\mathbf{A_{NN}}$	$\mathbf{VF_{NN}}$
Initializer	uniform$(0,1)$	He_uniform
Regularizers	L2 norm	L2 norm
Activation functions	Elu/(Sigmoid for output layer)	Elu/(Sigmoid for output layer)
Optimizer	Adam	Adam
Learning rates : step N − 1	5e-3	1e-3
steps k = 0, ..., N − 2	6.25e-4	5e-4
Scale	1e-3	1e-3
Number of elements in a training batch	3e2	3e2
Number of training batches	1e2	1e2
Size of the validation batches	1e3	1e3
Penalty constant	3e-1	NA
Number of epochs : step N − 1	2e3	2e3
steps k = 0, ..., N − 2	5e2	5e2
Size of the training set : step N − 1	6e7	6e7
steps k = 0, ..., N − 2	1.5e7	1.5e7
Size of the validation set : step N − 1	2e6	2e6
steps k = 0, ..., N − 2	5e5	5e5

Table 5.1: Parameters for the neural networks of the controls $\mathbf{A_{NN}}$ and the value function $\mathbf{VF_{NN}}$.

To train the NN, we simulate the input data. For the conditional expectation $\hat{B}_k$, we use its time-dependent Gaussian distribution (see Remark 5.4.1): $\hat{B}_k \sim \mathcal{N}(b_0, \Sigma_0 - \Sigma_k)$, with Σ_k as in Equation (5.17). On the other hand, the training of ρ is drawn from the uniform distribution between q and 1, the interval where it lies according to the maximum drawdown constraint.

5.5.2 *Hybrid-Now* algorithm

We use the *Hybrid-Now* algorithm developped in Bachouch et al. (2018b) in order to solve numerically our problem. This algorithm combines optimal policy estimation by neural networks and dynamic programming principle which suits the approach we have developped in Section 5.4.

With the same notations as in Algorithm 1 detailed in the next insert, at time k, the algorithm computes the proxy of the optimal control $\hat{\alpha}_k$ with A_{NN}, using the known function $\hat{V}_{k+1}$ calculated the step before, and uses V_{NN} to obtain a proxy of the value function $\hat{V}_k$. Starting from the known function $\hat{V}_N := U$ at terminal time N, the algorithm computes sequentially $\hat{\alpha}_k$ and $\hat{V}_k$ with backward iteration till time 0. This way, the algorithm loops to build the optimal controls and the value function pointwise and gives as output the optimal strategy, namely the optimal controls from 0 to $N-1$ and the value function at each of the N time steps.

The maximum drawdown constraint is a time-dependent constraint on the maximal proportion of wealth to invest (recall Lemma 5.3.3). In practice, it is a constraint on the sum of weights on each asset or equivalently on the output of A_{NN}. For that reason, we have implemented an appropriate penalty function that will reject undesirable values:

$$G_{Penalty}(A, r) = K \max\left(|A|_1 \leq 1 - \frac{q}{r}, 0\right), \quad A \in [0,1]^d, \quad r \in [q, 1].$$

This penalty function ensures that the strategy respects the maximum drawdown constraint at each time step, when the parameter K is chosen sufficiently large.

Algorithm 1: *Hybrid-Now*

Input: the training distributions μ_{Unif} and μ^k_{Gauss};

▷ $\mu_{Unif} = \mathcal{U}(q, 1)$

▷ $\mu^k_{Gauss} = \mathcal{N}(b_0, \Sigma_0 - \Sigma_k)$

Output:

- estimate of the optimal strategy $(\hat{a}_k)_{k=0}^{N-1}$;
- estimate of the value function $\left(\hat{V}_k\right)_{k=0}^{N-1}$;

Set $\hat{V}_N = U$;

for $k = N-1, \dots, 0$ **do**

Compute:

$$\hat{\beta}_k \in \underset{\beta \in \mathbb{R}^{2d^2+56d+283}}{\operatorname{argmin}} \mathbb{E}\left[G_{Penalty}(A_{NN}(\rho_k, \hat{B}_k; \beta), \rho_k) - \hat{V}_{k+1}\left(\rho^{\beta}_{k+1}, \hat{B}_{k+1}\right)\right]$$

where $\rho_k \sim \mu_{Unif}$, $\hat{B}_k \sim \mu^k_{Gauss}$,
$\hat{B}_{k+1} = \tilde{H}_k(\hat{B}_k, \tilde{\epsilon}_{k+1})$ and $\rho^{\beta}_{k+1} = F\left(\rho_k, \hat{B}_k, A_{NN}\left(\rho_k, \hat{B}_k; \beta\right), \tilde{\epsilon}_{k+1}\right)$;

▷ $F(\rho, b, a, \epsilon) = \min\left(1, \rho\left(1 + \sum_{i=1}^d a^i\left(e^{b^i + \epsilon^i} - 1\right)\right)\right)$

▷ $\tilde{H}_k(b, \epsilon) = b + \Sigma_0(\Gamma + \Sigma_0 k)^{-1}\epsilon$

Set $\hat{a}_k = A_{NN}\left(.; \hat{\beta}_k\right)$;

▷ $\hat{a}_k$ is the estimate of the optimal control at time k.

Compute:

$$\hat{\theta}_k \in \underset{\theta \in \mathbb{R}^{2d^2+54d+261}}{\operatorname{argmin}} \mathbb{E}\left[\left(\hat{V}_{k+1}\left(\rho^{\hat{\beta}_k}_{k+1}, \hat{B}_{k+1}\right) - VF_{NN}\left(\rho_k, \hat{B}_k; \theta\right)\right)^2\right]$$

Set $\hat{V}_k = VF_{NN}\left(., \hat{\theta}_k\right)$;

▷ $\hat{V}_k$ is the estimate of the value function at time k.

A major argument behind the choice of this algorithm is that, it is particularly relevant for problems in which the neural network approximation of the controls and value function at time k, are close to the ones at time $k+1$. This is what we expect in our case. We can then take a small learning rate for the Adam optimizer which enforces the stability of the parameters' update during the gradient-descent based learning procedure.

5.5.3 Numerical results

In this section, we explain the setup of the simulation and exhibit the main results. We have used Tensorflow 2 and deep learning techniques for Python developped in Géron (2019). We consider $d = 3$ risky assets and a riskless asset whose return is assumed to be 0, on a 1-year investment horizon for the sake of simplicity. We consider 24 portfolio rebalancing during the 1-year period, i.e., one every 2 weeks. This means that we have $N = 24$ steps in the training of our neural networks. The parameters used in the simulation are detailed in Table 5.2.

Parameter	Value
Number of risky assets d	3
Investment horizon in years T	1
Number of steps/rebalancing N	24
Number of simulations/trajectories $\tilde{N}$	1000
Degree of the CRRA utility function p	0.8
Parameter of risk aversion q	0.7
Annualized expectation of the drift B	$\begin{bmatrix} 0.05 & 0.025 & 0.12 \end{bmatrix}$
Annualized covariance matrix of the drift B	$\begin{bmatrix} 0.2^2 & 0 & 0 \\ 0 & 0.15^2 & 0 \\ 0 & 0 & 0.1^2 \end{bmatrix}$
Annualized volatility of ϵ	$\begin{bmatrix} 0.08 & 0.04 & 0.22 \end{bmatrix}$
Correlation matrix of ϵ	$\begin{bmatrix} 1 & -0.1 & 0.2 \\ -0.1 & 1 & -0.25 \\ 0.2 & -0.25 & 1 \end{bmatrix}$
Annualized covariance matrix of the noise ϵ	$\begin{bmatrix} 0.0064 & -0.00032 & 0.00352 \\ -0.00032 & 0.0016 & -0.0022 \\ 0.00352 & -0.0022 & 0.0484 \end{bmatrix}$

Table 5.2: Values of the parameters used in the simulation.

First, we show the numerical results for the learning and the non-learning strategies by presenting a performance and an allocation analysis in Subsection 5.5.3.1. Then, we add the admissible constrained EW to the two previous ones and use this neutral strategy as a benchmark in Subsection 5.5.3.2. Ultimately, in Subsection 5.5.3.3, we illustrate numerically the convergence of the non-learning strategy to the constrained Merton problem when the loss aversion parameter q vanishes.

5.5.3.1 Learning and non-learning strategies

We simulate $\tilde{N} = 1000$ trajectories for each strategy and exhibit the performance results with an initial wealth $x_0 = 1$. Figures 5.3 illustrates the average historical level of the learning and non-learning strategies with a 95% confidence interval. Learning outperforms significantly Non-Learning with a narrower confidence interval revealing that less uncertainty surrounds Learning performance, thus yielding less risk.

An interesting phenomenon, visible in Figure 5.3, is the nearly flat curve for Learning between time 0 and time 1. Indeed, whereas Non-Learning starts investing immediately, Learning adopts a safer approach and needs a first time step before allocating a significant proportion of wealth. Given the level of uncertainty surrounding b_0, this first step allows Learning to fine-tune its allocation by updating the prior belief with the first return available at time 1. On the contrary, Non-Learning, which cannot update its prior, starts investing at time 0.
Figure 5.4 shows the ratio of Learning over Non-Learning. A ratio greater than one means that Learning outperforms Non-Learning and underperforms when less than one. It shows the significant outperformance of Learning over Non-Learning except during the first period where Learning was not significantly invested and Non-Learning had a positive return. Morevover, this graph reveals the typical increasing concave curve of the value of information described in Keppo et al. (2018), in the context of investment decisions and costs of data analytics, and in De Franco, Nicolle, and Pham (2019a) in the resolution of the Markowitz portfolio selection problem using a Bayesian learning approach.

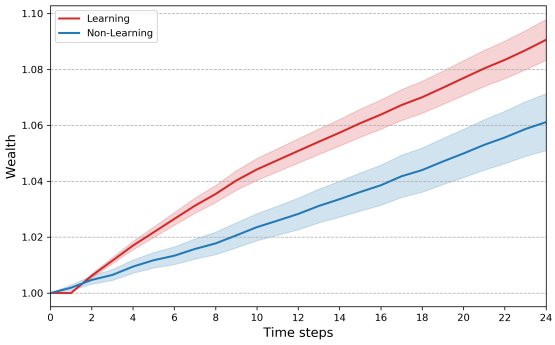


Figure 5.3: Historical Learning and Non-Learning levels with a 95% confidence interval.

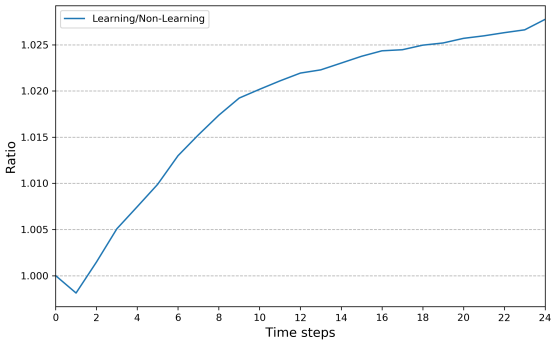


Figure 5.4: Historical ratio of Learning over Non-Learning levels.

Table 5.3 gathers relevant statistics for both Learning and Non-Learning such as: average total performance, standard deviation of the terminal wealth X_T, Sharpe ratio computed as average total performance over standard deviation of terminal wealth. The maximum drawdown (MD) is examined through two statistics: noting $MD_\ell^{\tilde{s}}$ the maximum drawdown of the ℓ-*th* trajectory of a strategy $\tilde{s}$, the average MD is defined as,

$$\text{Avg MD}^{\tilde{s}} = \frac{1}{\tilde{N}} \sum_{\ell=1}^{\tilde{N}} \text{MD}_\ell^{\tilde{s}},$$

for $\tilde{N}$ trajectories of the strategy $\tilde{s}$, and the worst MD is defined as,

$$\text{Worst MD}^{\tilde{s}} = \min\left(MD_1^{\tilde{s}}, \ldots, MD_{\tilde{N}}^{\tilde{s}}\right).$$

Finally, the Calmar ratio, computed as the ratio of the average total performance over the average maximum drawdown, is the last statistic exhibited.

With the simulated dataset, Learning delivered, on average, a total performance of 9.34% while Non-Learning only 6.40%. Integrating the most recent information yielded a 2.94% excess return. Moreover, risk metrics are significantly better for Learning than for Non-Learning. Learning exhibits a lower standard deviation of terminal wealth than Non-Learning (11.88% versus 16.67%), with a difference of 4.79%. More interestingly, the maximum drawdown is notably better controlled by Learning than by Non-Learning, on average (-1.53% versus -6.54%) and in the worst case (-11.74% versus -27.18%). This result suggests that learning from new observations helps the strategy to better handle the dual objective of maximizing total wealth while controlling the maximum drawdown. We also note that learning improves the Sharpe ratio by 104.95% and the Calmar ratio by 525.26%.

	Learning	**Non-Learning**	**Difference**
Avg total performance	9.34%	6.40%	2.94%
Std dev. of X_T	11.88%	16.67%	-4.79%
Sharpe ratio	0.79	0.38	104.95%
Avg MD	-1.53%	-6.54%	5.01%
Worst MD	-11.74%	-27.18%	15.44%
Calmar ratio	6.12	0.98	525.26%

Table 5.3: Performance metrics: Learning and Non-Learning. The difference for ratios are computed as relative improvement.

Figure 5.5 and 5.6 focus more precisely on the portfolio allocation. The graphs of Figure 5.5 show the historical average allocation for each of the three risky assets. First, none of the stategies invests in Asset 2 since it has the lowest expected return according to the prior, see Table 5.2. Whereas Non-Learning focuses on Asset 3, the one with the highest expected return, Learning performs an optimal allocation between Asset 1 and Asset 3 since this strategy is not stuck with the initial estimate given by the prior. Therefore, Learning invests little at time 0, then balances nearly equally both Assets 1 and 3, and then invests only in Asset 3 after time step 12. Instead, Non-Learning is investing only in Asset 3, from time 0 until the end of the investment horizon.

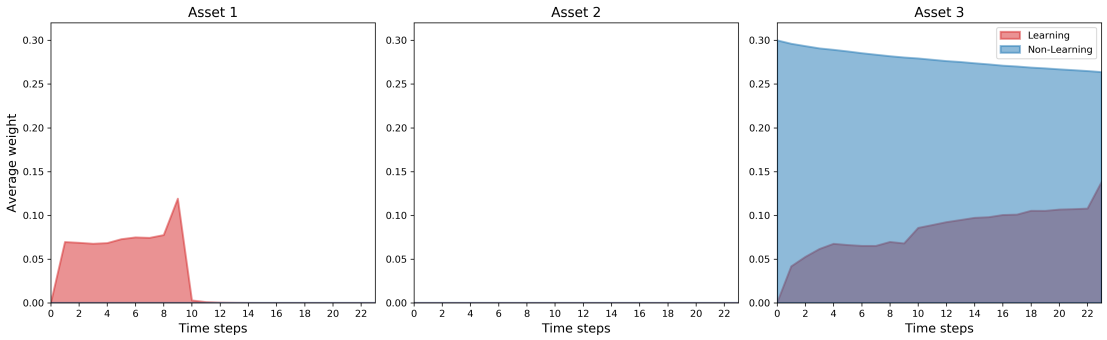


Figure 5.5: Historical Learning and Non-Learning asset allocations.

The curves in Figure 5.6 recall each asset's optimal weight, but the main features are the colored areas that represent the average historical total percentage of wealth invested by each strategy. The dotted line represents the total allocation constraint they should satisfy to be admissible. To satisfy the maximum drawdown constraint, admissible strategies can only invest in risky assets the proportion of wealth that, in theory, could be totally lost. This explains why the non-learning strategy invests at full capacity on the asset that has the maximum expected return according to the prior distribution.

We clearly see that both strategies satisfy their respective constraints. Indeed, looking at the left panel, Learning is far from saturating the constraint. It has invested, on average, roughly 10% of its wealth while its constraint was set around 30%. Non-learning invests at full capacity saturating its allocation constraint. Remark that this constraint is not a straight line since it depends on the value of the ratio: current wealth over current historical maximum, and evolves according to time.

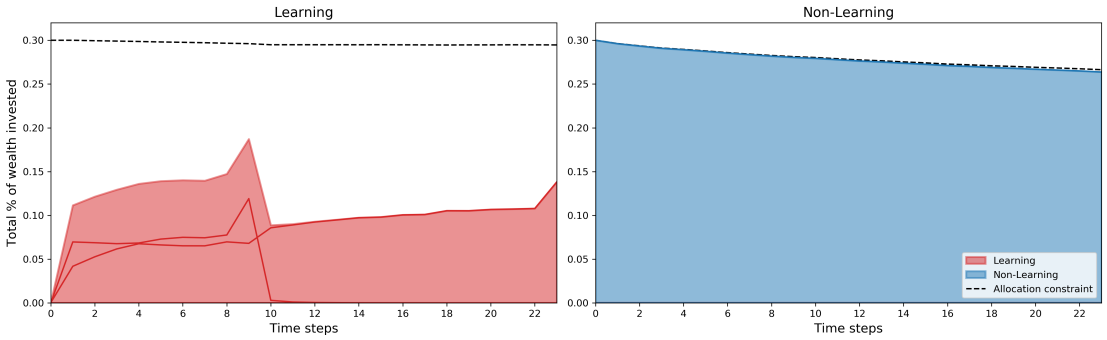


Figure 5.6: Historical Learning and Non-Learning total allocations.

5.5.3.2 Learning, non-learning and constrained equally-weighted strategies

In this section, we add a simple constrained equally-weighted (EW) strategy to serve as a benchmark for both Learning and Non-Learning. At each time step, the constrained EW strategy invests, equally accross the three assets, the proportion of wealth above the threshold q.

Figure 5.7 shows the average historical levels of the three strategies: Learning, Non-Learning and constrained EW. We notice Non-Learning outperforms constrained EW and both have similar confidence intervals. It is not surprising to see that Non-Learning outperforms constrained EW since Non-Learning always bets on Asset 3, the most performing, while constrained EW diversifies the risks equally among the three assets.

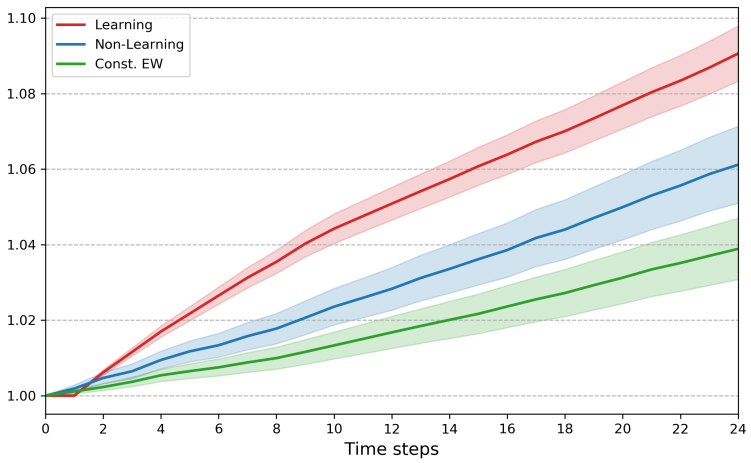


Figure 5.7: Historical Learning, Non-Learning and constrained EW (Const. EW) levels with a 95% confidence interval.

Figure 5.8 shows the ratio of Learning over constrained EW: it depicts the same concave shape as Figure 5.4. The outperformance of Non-Learning with respect to constrained EW is plot in Figure 5.9 and confirms, on average, the similarity of the two strategies.

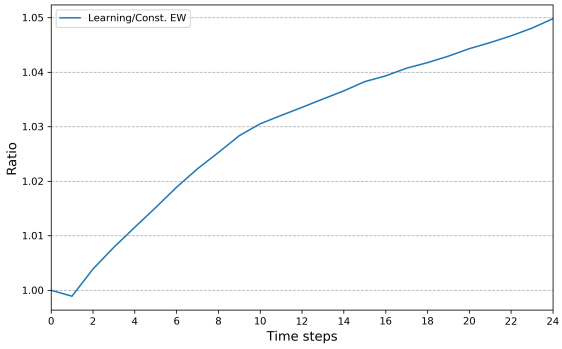


Figure 5.8: Ratio Learning over constrained EW (Const. EW) according to time.

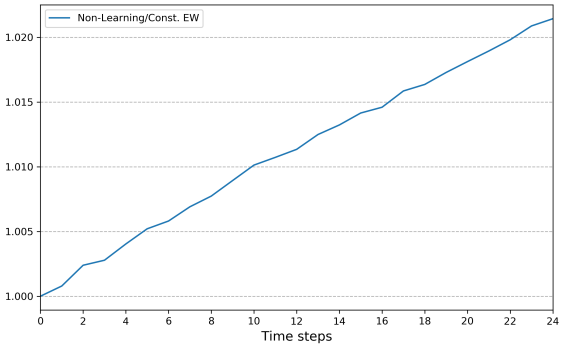


Figure 5.9: Ratio Non-Learning over constrained EW (Const. EW) according to time.

Table 5.4 collects relevant statistics for the three strategies. Learning clearly outstands constrained EW: it outperforms by 5.49% while reducing uncertainty on terminal wealth by 1.92% resulting in an improvement of 182.08% of the Sharpe ratio. Moreover, it better handles maximum drawdown regarding both the average and the worst case, exhibiting an improvement of 3.17% and 10.09% respectively, enhancing the Calmar ratio by 647.56%.
The Non-Learning and the constrained EW have similar profiles. Even if Non-Learning outperforms constrained EW by 2.5%, it has a higher uncertainty in terminal wealth (+2.87%). This results in similar Sharpe ratios. Maximum drawdown, both on average and considering the worst case are better handled by constrained EW (−4.70% and −21.83% respectively) than by Non-Learning (−6.54% and −27.18% respectively) thanks to the diversification capacity of constrained EW. The better performance of Non-Learning compensates the better maximum drawdown handling of constrained EW, entailing a better Calmar ratio for Non-Learning 0.98 versus 0.82 for constained EW.

	Const. EW	**L**	**NL**	**L - Const. EW**	**NL - Const. EW**
Avg total performance	3.85%	9.34%	6.40%	5.49%	2.55%
Std dev. of X_T	13.80%	11.88%	16.67%	-1.92%	2.87%
Sharpe ratio	0.28	0.79	0.38	182.08%	37.63%
Avg MD	-4.70%	-1.53%	-6.54%	3.17%	-1.84%
Worst MD	-21.83%	-11.74%	-27.18%	10.09%	-5.34%
Calmar ratio	0.82	6.12	0.98	647.56%	-19.56%

Table 5.4: Performance metrics: Constrained EW (Const. EW) vs Learning (L) and Non-Learning (NL). The difference for ratios are computed as relative improvement.

5.5.3.3 Non-learning and Merton strategies

We numerically analyze the impact of the drawdown parameter q, and compare the non-learning strategies (assuming that the drift is equal to b_0), with the constrained

Merton strategy as described in Remark 5.4.2. Figure 5.10 confirms that when the loss aversion parameter q goes to zero, the non-learning strategy approaches the Merton strategy.

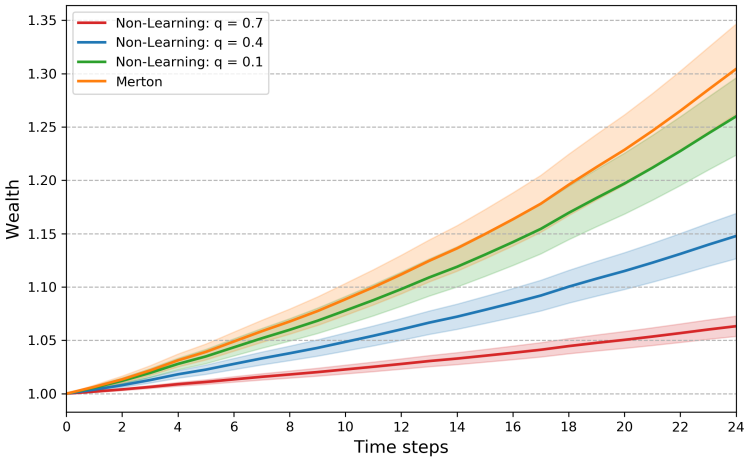


Figure 5.10: Wealth curves resulting from the Merton strategy and the non-learning strategy for different values of q with a 95% confidence interval.

In terms of assets' allocation, the Merton strategy saturates the constraint only by investing in the asset with the highest expected return, Asset 3, while the non-learning strategy adopts a similar approach and invests at full capacity in the same asset. To illustrate this point, we easily see that the areas at the top and bottom-left corner converge to the area at the bottom-right corner of Figure 5.11.

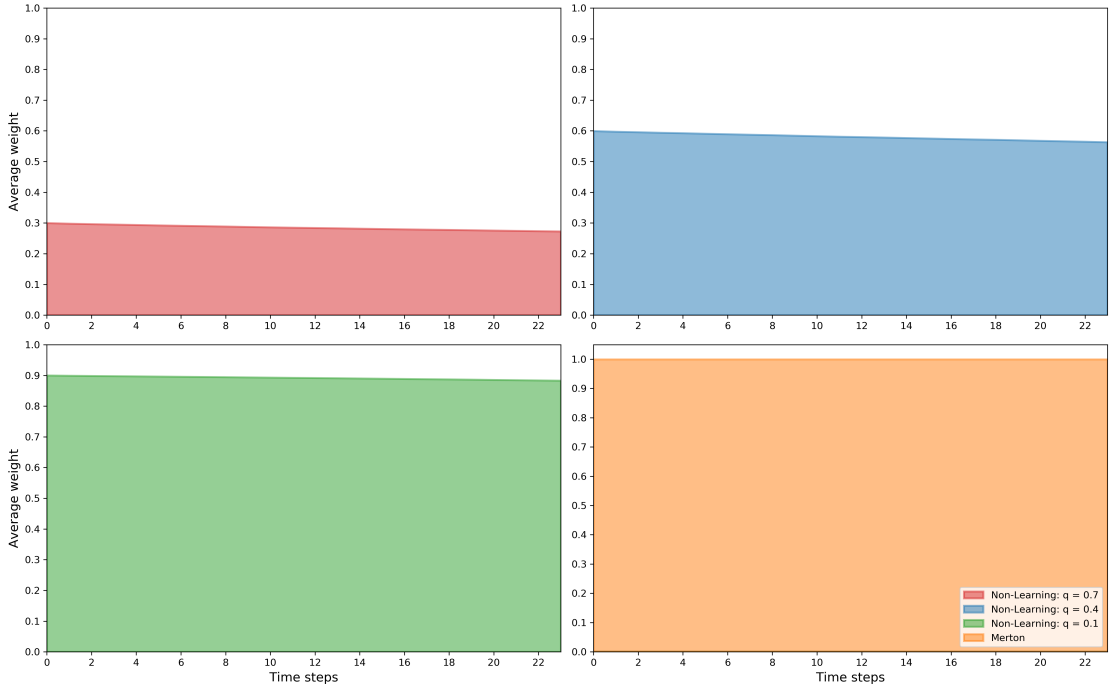


Figure 5.11: Asset 3 average weights of the non-learning strategies with $q \in \{0.7, 0.4, 0.1\}$ and the Merton strategy.

As q vanishes, we observe evidence of the convergence of the Merton and the non-learning strategies, materialized by a converging allocation pattern and resulting wealth trajectories. It should not be surprising since both have in common not to learn from incoming information conveyed by the prices.

5.5.4 Sensitivities analysis

In this subsection, we study the effect of changes in the uncertainty about the beliefs of B. These beliefs take the form of an estimate b_0 of B, and a degree of uncertainty about this estimate, the covariance of Σ_0 of B. For the sake of simplicity, we design Σ_0 as a diagonal matrix whose diagonal entries are variances representing the confidence the investor has in her beliefs about the drift. To easily model a change in Σ_0, we define the modified covariance matrix $\tilde{\Sigma}$ as

$$\tilde{\Sigma}_{unc} := unc * \Sigma_0,$$

where $unc > 0$. From now on, the prior of B is $\mathcal{N}(b_0, \tilde{\Sigma}_{unc})$.
A higher value of unc means a higher uncertainty materialized by a lower confidence in the prior estimate of the expected return of B, b_0. We consider learning strategies with values of $unc \in \{1/6,\ 1,\ 3,\ 6,\ 12\}$. The value $unc = 1$ was used for Learning in Subsection 5.5.3.

Equation (5.2) implies that the returns' probability distribution depends upon *unc*. It implies that for each value of *unc*, we need to compute both Learning and Non-Learning on the returns sample drawn from the same probability law to make relevant comparisons.
Therefore, from a sample of a thousand returns paths' draws, we plot in Figure 5.12 the average curves of the excess return of Learning over its associated Non-Learning, for different values of the uncertainty parameter *unc*.

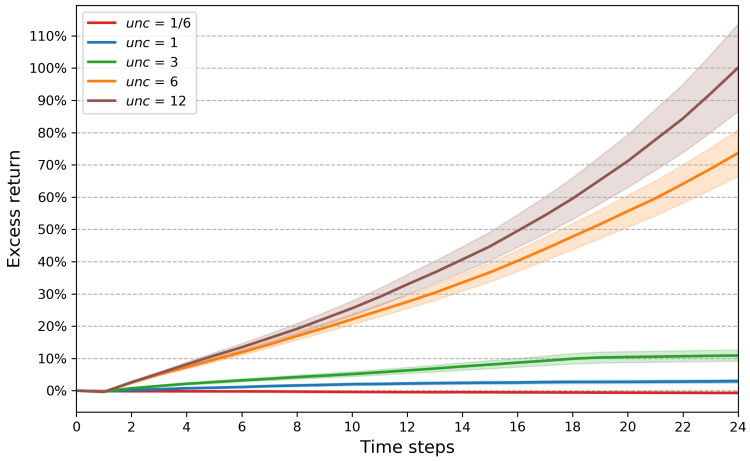


Figure 5.12: Excess return of Learning over Non-Learning with a 95% confidence interval for different levels of uncertainty.

Looking at Figure 5.12, we notice that when uncertainty about b_0 is low, i.e. $unc = 1/6$, Learning is close to Non-Learning and unsurprisingly the associated excess return is small. Then, as we increase the value of *unc* the curves steepen increasingly showing the effect of learning in generating excess return.

Table 5.5 summarises key statistics for the ten strategies computed in this section. When $unc = 1/6$, Learning underperforms Non-Learning. This is explained by the fact that Non-Learning has no doubt about b_0 and knows Asset 3 is the best performing asset acoording to its prior, whereas Learning, even with low uncertainty, needs to learn it generating a lag which explains the underperformance on average. For values of $unc \geq 1$ Learning outperforms Non-learning increasingly, as can be seen on Figure 5.13, at the cost of a growing standard deviation of terminal wealth.

The Sharpe ratio of terminal wealth is higher for Learning than for Non-Learning for any value of *unc*. Nevertheless, an interesting fact is that the ratio rises from $unc = 1/6$ to $unc = 1$, then reaches a level close to 0.8 for values of $unc = 1, 3, 6$ then decreases when $unc = 12$.

	$unc = 1/6$		$unc = 1$		$unc = 3$		$unc = 6$		$unc = 12$	
	L	**NL**	**L**	**NL**	**L**	**NL**	**L**	**NL**	**L**	**NL**
Avg total performance	3.87%	4.35%	9.45%	6.00%	19.96%	10.25%	90.03%	16.22%	130.07%	30.44%
Std dev. of X_T	5.81%	9.22%	12.10%	17.28%	25.01%	28.18 %	113.69%	41.24%	222.77%	70.84%
Sharpe ratio	0.67	0.47	0.78	0.35	0.80	0.36	0.79	0.39	0.58	0.43
Avg MD	-2.51%	-5.21%	-1.40%	-6.78%	-1.90%	-8.40%	-2.68%	-10.14%	-3.58%	-11.35%
Worst MD	-7.64%	-17.88%	-5.46%	-24.01%	-7.99%	-26.68%	-15.62%	-29.22%	-16.98%	-29.47%
Calmar ratio	1.54	0.83	6.77	0.89	10.49	1.22	33.65	1.60	36.32	2.68

Table 5.5: Performance and risk metrics: Learning (L) vs Non-Learning (NL) for different values of uncertainty *unc*.

This phenomenon is more visible on Figure 5.14 that displays the Sharpe ratio of terminal wealth of Learning and Non-Learning according to the values of *unc*, and the associated relative improvement. Clearly, looking at Figures 5.13 and 5.14, we remark that while increasing *unc* gives more excess return, too high values of *unc* in the model turn to be a drag as far as Sharpe ratio improvement is concerned.

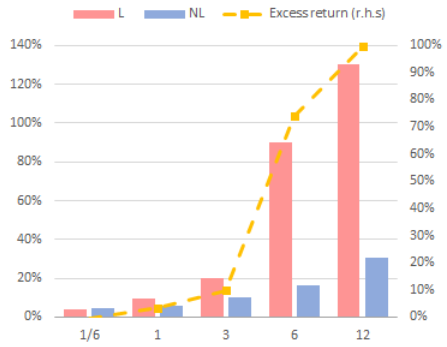


Figure 5.13: Average total performance of Learning (L) and Non-Learning (NL), and excess return, for $unc \in \{1/6, 1, 3, 6, 12\}$.

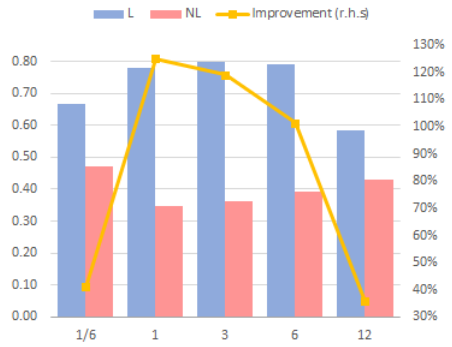


Figure 5.14: Sharpe ratio of terminal wealth of Learning (L) and Non-Learning (NL), and relative improvement, for $unc \in \{1/6, 1, 3, 6, 12\}$.

For any value of *unc*, Learning handles maximum drawdown significantly better than Non-Learning whatever it is the average or the worst. This results in a better performance per unit of average maximum drawdown (Calmar ratio), for Learning. We also see that the maximum drawdown constraint is satisfied for every strategies of the sample and for any value of *unc* since the worst maximum drawdown is always above -30%, the lowest admissible value with a loss aversion parameter q set at 0.7.

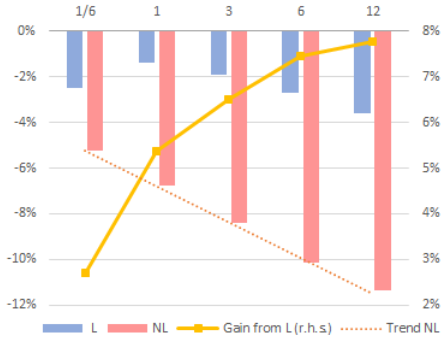


Figure 5.15: Average maximum drawdown of Learning (L) and Non-Learning (NL) and the gain from learning for $unc \in \{1/6, 1, 3, 6, 12\}$.

Figure 5.15 reveals how the average maximum drawdown behaves regarding the level of uncertainty. Non-Learning maximum drawdown behaves linearly with uncertainty: the wider the range of possible values of B the higher the maximum drawdown is on average. It emphasizes its inability to adapt to an environment in which the returns have different behaviors compared to their expectations. Learning instead, manages to keep a low maximum drawdown for any value of *unc*. Given the previous remarks, it is obvious that the gain in maximum drawdown from learning grows with the level of uncertainty.

Figures 5.16-5.20 represent portfolio allocations averaged over the simulations. They depict, for each value of the uncertainty parameter *unc*, the average proportion of wealth invested, in each of the three assets, by Learning and Non-Learning. The purpose is not to compare the graphs with different values of *unc* since the allocation is not performed on the same sample of returns. Rather, we can identify trends that are typically differentiating Learning from Non-Learning allocations.

Since the maximum drawdown constraint is satisfied by the capped sum of total weights that can be invested, the allocations of both Learning and Non-Learning are mainly based on the expected returns of the assets.
Non-Learning, by definition, does not depend on the value of the uncertainty parameter. Hence, no matter the value of *unc*, its allocation is easy to characterize since it saturates its constraint investing in the asset that has the best expected return according to the prior. In our setup, Asset 3 has the highest expected return, so Non-Learning invests only in it and saturates its constraint of roughly 30% during all the investment period. The slight change of the average weight in Asset 3 comes from ρ, the ratio wealth over maximum wealth, changing over time.

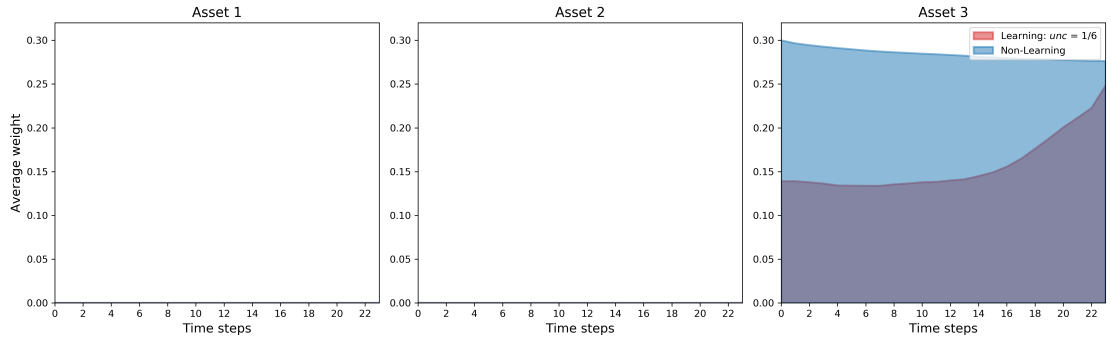


Figure 5.16: Learning and Non-Learning historical assets' allocations with $unc = 1/6$.

Unlike Non-Learning, depending of the value of *unc*, Learning can perform more sophisticated allocations because it can adjust the weights according to the incoming information. Nonetheless, in Figure 5.16, when *unc* is low, Learning and Non-Learning look similar regarding their weights allocation since both strategies invest, as of time 0, a significant proportion of their wealth only in Asset 3.
On the right panel of Figure 5.16, the progressive increase in the weight of Asset 3 illustrates the learning process. As time goes by, Learning progressively increases the weight in Asset 3 since it has the highest expected return. It also explains why Learning underperforms Non-Learning for low values of *unc*; contrary to Non-Learning which invests at full capacity in Asset 3, Learning needs to learn that Asset 3 is the optimal choice.

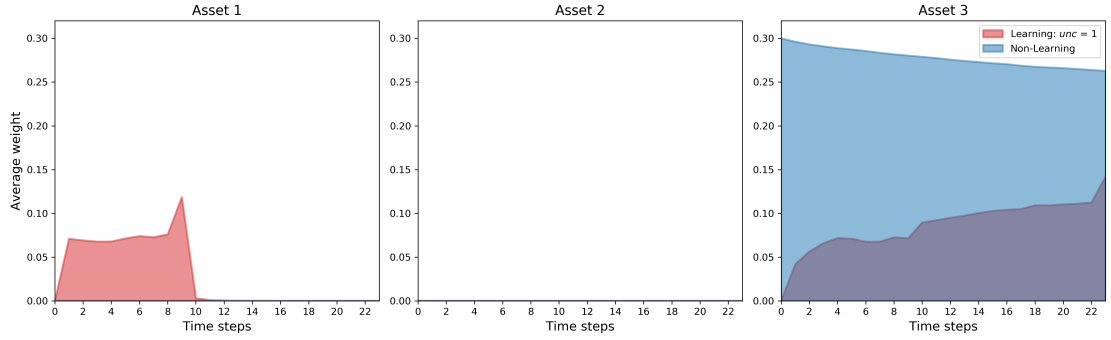


Figure 5.17: Learning and Non-Learning historical assets' allocations with $unc = 1$.

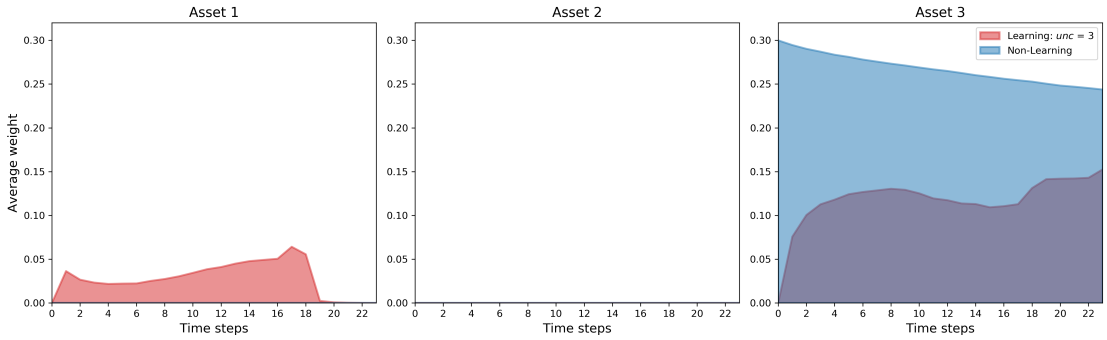


Figure 5.18: Learning and Non-Learning historical assets' allocations with $unc = 3$.

However, as uncertainty increases, Learning and Non-Learning strategies start differentiating. When $unc \geq 1$, Learning invests little, if any, at time 0. In addition, an increase in unc allows the inital drift to lie in a wider range and generates investment opportunities for Learning. This explains why Learning invests in Asset 1 when $unc = 1,\ 3,\ 6,\ 12$ although the estimate b_0 for this asset is lower than for Asset 3. In Figure 5.19, we see that Learning even invests in Asset 2 which has the lowest expected drift.

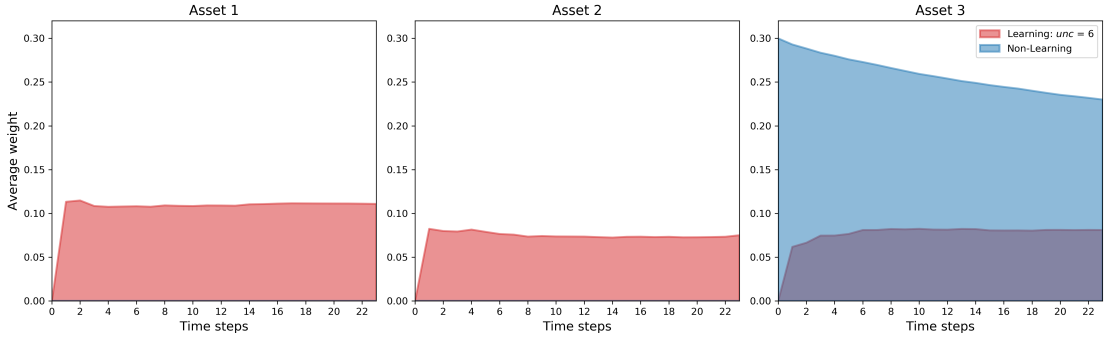


Figure 5.19: Learning and Non-Learning historical assets' allocations with $unc = 6$.

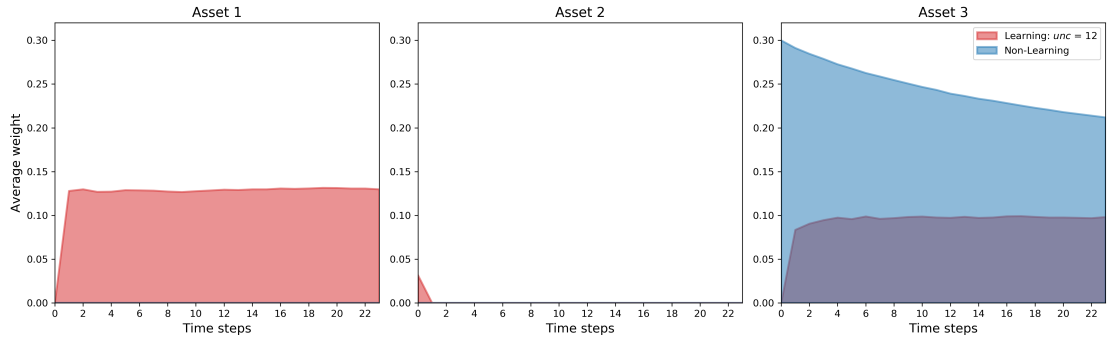


Figure 5.20: Learning and Non-Learning historical assets' allocations with $unc = 12$.

Figures 5.21-5.25 illustrate the historical total percentage of wealth allocated for Learning and Non-Learning with different levels of uncertainty. As seen previously, Non-Learning has fully investment in Asset 3 for any value of *unc*.

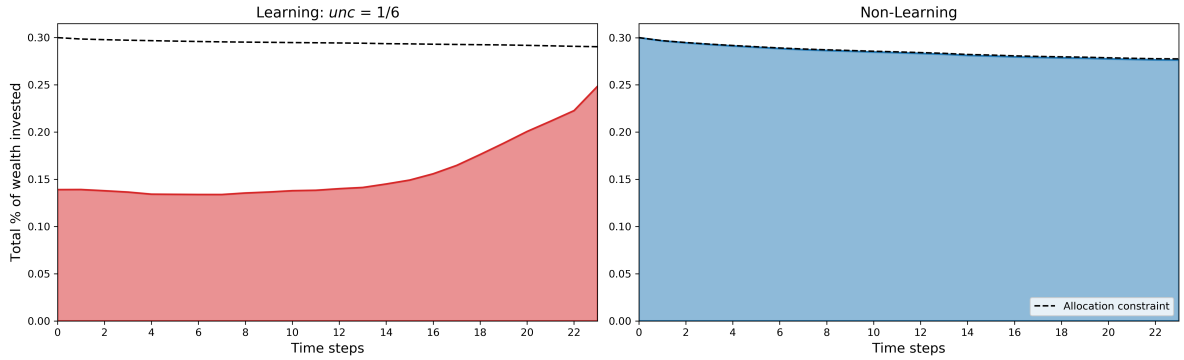


Figure 5.21: Historical total allocations of Learning and Non-Learning with $unc = 1/6$.

Moreover, Learning has always less investment that Non-Learning for any level of uncertainty. It suggests that Learning yields a more cautious strategy than Non-Learning. This fact, in addition to its wait-and-see approach at time 0 and its ability to better handle maximum drawdown, makes Learning a safer and more conservative strategy than Non-Learning. This can be seen in Figure 5.21, where both Learning and Non-Learning have invested in Asset 3, but not at the same pace. Non-Learning goes fully in Asset 3 at time 0, whereas Learning increments slowly its weight in Asset 3 reaching 25% at the final step. When *unc* is low, there is no value added to choose Learning over Non-Learning from a performance perspective. Nevertheless, Learning allows for a better management of risk as Table 5.5 exhibits.

As *unc* increases, in addition to being cautious, Learning mixes allocation in different assets, see Figures 5.22-5.25, while Non-Learning is stuck with the highest expected return asset.

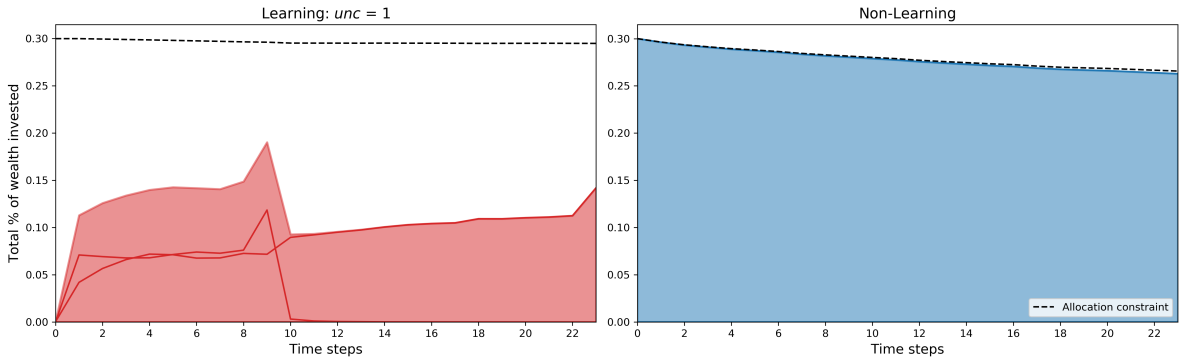


Figure 5.22: Historical total allocations of Learning and Non-Learning with $unc = 1$.

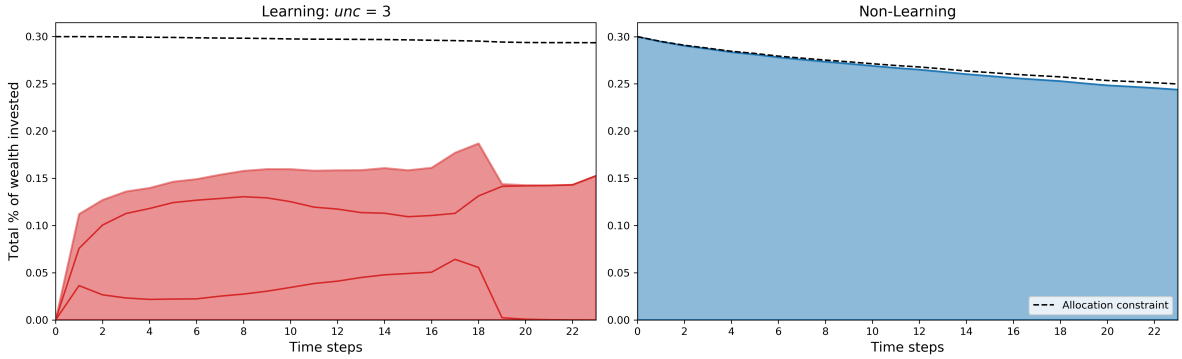


Figure 5.23: Historical total allocations of Learning and Non-Learning with $unc = 3$.

Learning is able to be opportunistic and changes its allocation given the prices observed. For example in Figure 5.22, Learning starts investing in Asset 1 and 3 at time 1 and stops at time 12 to weigh Asset 1 while keeping Asset 3. Similar remarks can be made for Figure 5.23, where Learning puts non negligeable weights in all three risky assets for $unc = 6$ in Figure 5.24.

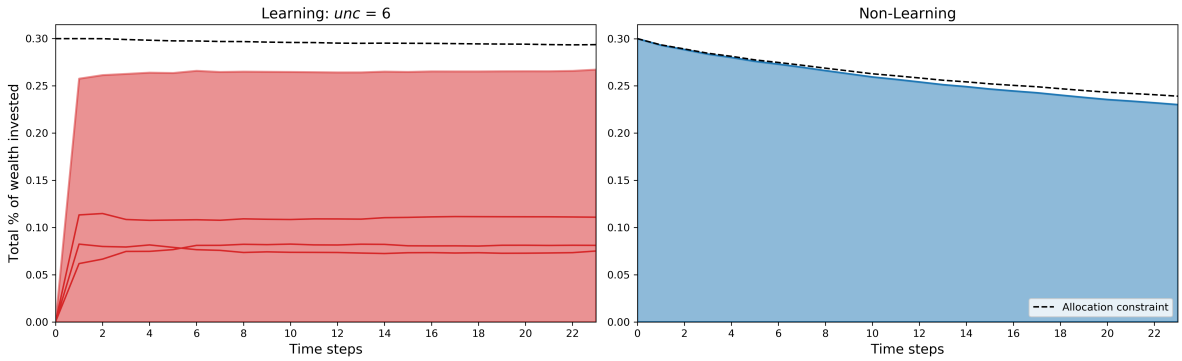


Figure 5.24: Historical total allocations of Learning and Non-Learning with $unc = 6$.

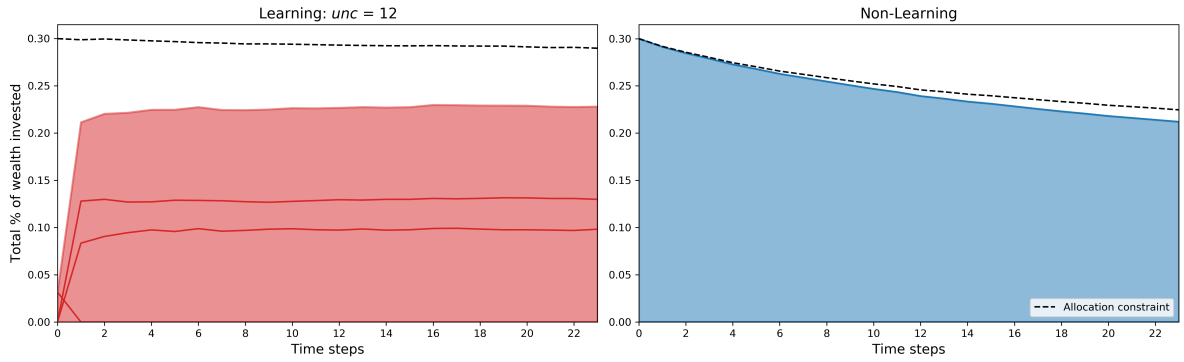


Figure 5.25: Historical total allocations of Learning and Non-Learning with $unc = 12$.

5.6 Conclusion

We have studied a discrete-time portfolio selection problem by taking into account both drift uncertainty and maximum drawdown constraint. The dynamic programming equation is derived in the general case, thanks to a specific change of measure, and more explicit results have been provided in the Gaussian case using the Kalman filter. Moreover, a change of variable has reduced the dimensionality of the problem in the case of CRRA utility functions. Furthermore, we have provided extensive numerical results in the Gaussian case with CRRA utility functions using recent deep neural network techniques. Our numerical analysis has clearly shown and quantified the better risk-return profile of the learning strategy versus the non-learning one. Indeed, besides outperforming the non-learning strategy, the learning one provides a significantly lower standard deviation of terminal wealth and a better controlled maximum drawdown. Confirming the results established in De Franco, Nicolle, and Pham (2019b), this study exhibits the benefits of learning in providing optimal portfolio allocations.

5.A Appendices

5.A.1 Proof of Proposition 5.3.1

Proof. For all $k = 1, ..., N$, the law under $\overline{\mathbb{P}}$, of R_k given the filtration $\mathcal{G}_{k-1}$ yields the unconditional law under $\mathbb{P}$ of ϵ_k. Indeed, since $(\Lambda_k)_k$ is a $(\mathbb{P}, \mathbb{G})$-martingale, we

have from Bayes formula, for all Borelian $F \subset \mathbb{R}^d$,

$$\begin{aligned}\overline{\mathbb{P}}[R_k \in F|\mathcal{G}_{k-1}] &= \overline{\mathbb{E}}[\mathbb{1}_{\{R_k\in F\}}|\mathcal{G}_{k-1}] = \frac{\mathbb{E}[\Lambda_k \mathbb{1}_{\{R_k\in F\}}|\mathcal{G}_{k-1}]}{\mathbb{E}[\Lambda_k|\mathcal{G}_{k-1}]} \\ &= \mathbb{E}[\frac{\Lambda_k}{\Lambda_{k-1}}\mathbb{1}_{\{R_k\in F\}}|\mathcal{G}_{k-1}] = \mathbb{E}\left[\frac{g(B+\epsilon_k)}{g(\epsilon_k)}\mathbb{1}_{\{R_k\in F\}}\Big|\mathcal{G}_{k-1}\right] \\ &= \int_{\mathbb{R}^d}\frac{g(B+e)}{g(e)}\mathbb{1}_{\{B+e\in F\}}g(e)de = \int_{\mathbb{R}^d} g(z)\mathbb{1}_{\{z\in F\}}dz \\ &= \mathbb{P}[\epsilon_k \in F].\end{aligned}$$

This means that, under $\overline{\mathbb{P}}$, R_k is independent from B and from $R_1, .., R_{k-1}$ and that R_k has the same probability distribution as ϵ_k. □

5.A.2 Proof of Proposition 5.3.2

Proof. For any borelian function $f : \mathbb{R}^d \mapsto \mathbb{R}$ we have, on one hand, by definition of π_{k+1}:

$$\overline{\mathbb{E}}\left[\overline{\Lambda}_{k+1}f(B)|\mathcal{F}^o_{k+1}\right] = \int_{\mathbb{R}^d} f(b)\pi_{k+1}(db),$$

and, on the other hand, by definition of $\overline{\Lambda}_k$:

$$\begin{aligned}\overline{\mathbb{E}}[\overline{\Lambda}_{k+1}f(B)|\mathcal{F}^o_{k+1}] &= \overline{\mathbb{E}}\left[\overline{\Lambda}_k f(B)\frac{g(R_{k+1}-B)}{g(R_{k+1})}\Big|\mathcal{F}^o_{k+1}\right] \\ &= \overline{\mathbb{E}}\left[\overline{\Lambda}_k f(B)g(R_{k+1}-B)\big|\mathcal{F}^o_{k+1}\right](g(R_{k+1}))^{-1} \\ &= \int_{\mathbb{R}^d} f(b)\frac{g(R_{k+1}-b)}{g(R_{k+1})}\pi_k(db),\end{aligned}$$

where we use in the last equality the fact that R_{k+1} is independent of B under $\overline{P}$ (recall Proposition 5.3.1). By identification, we obtain the expected relation. □

5.A.3 Proof of Lemma 5.3.3

Proof. Since the support of the probability distribution ν of ϵ_k is $\mathbb{R}^d$, we notice that the law of the random vector $Y_k := e^{R_k} - \mathbb{1}_d$ has support equal to $(-1,\infty)^d$. Recall from (5.7) that $a \in A^q_k(x,z)$ if and only if

$$1 + a'Y_{k+1} \geq q\max\left[\frac{z}{x}, 1 + a'Y_{k+1}\right], \quad a.s. \tag{5.21}$$

(i) Take some $a \in A^q_k(x,z)$, and assume that $a^i < 0$ for some $i \in [\![1,d]\!]$. Let us then define the event $\Omega^i_M = \{Y^i_{k+1} \geq M, Y^M_{k+1} \in [0,1], j \neq i\}$, for $M > 0$, and observe

that $\mathbb{P}[\Omega_M^i] > 0$. It follows from (5.21) that

$$1 + a_i M + \max_{j \neq i} |a_j| \;\geq\; q\frac{z}{x}, \quad \text{on } \Omega_M^i,$$

which leads to a contradiction for M large enough. This shows that $a^i \geq 0$ for all $i \in [\![1, d]\!]$, i.e. $A_k^q(x,z) \subset \mathbb{R}_+^d$.

(ii) For $\varepsilon \in (0,1)$, let us define the event $\Omega_\varepsilon = \{Y_{k+1}^i \leq -1 + \varepsilon, i = 1, \ldots, d\}$, which satisfies $\mathbb{P}[\Omega_\varepsilon] > 0$. For $a \in A^q(x,z)$, we get from (5.21), and since $a \in \mathbb{R}_+^d$ by Step (i):

$$1 - (1-\varepsilon)a'\mathbb{1}_d \;\geq\; q\frac{z}{x}, \quad \text{on } \Omega_\varepsilon.$$

By taking ε small enough, this shows by a contradiction argument that

$$A^q(x,z) \;\subset\; \Big\{a \in \mathbb{R}_+^d : 1 - a'\mathbb{1}_d \geq q\frac{z}{x}\Big\}. \;=:\; \tilde{A}^q(x,z). \tag{5.22}$$

(iii) Let us finally check the equality in (5.22). Fix some $a \in \tilde{A}^q(x,z)$. Since the random vector Y_{k+1} is valued in $(-1,\infty)^d$, it is clear that

$$1 + a'Y_{k+1} \;\geq\; 1 - a'\mathbb{1}_d \;\geq\; q\frac{z}{x} \;\geq\; 0, \quad a.s.,$$

and thus

$$1 + a'Y_{k+1} \;\geq\; q\big[1 + a'Y_{k+1}\big], \quad a.s.,$$

which proves (5.21), hence the equality $A^q(x,z) = \tilde{A}^q(x,z)$. □

5.A.4 Proof of Lemma 5.3.4

Proof. **1.** Fix $q_1 \leq q_2$ and $(x,z) \in \mathcal{S}^{q_2} \subset \mathcal{S}^{q_1}$. We then have

$$a \in A^{q_2}(x,z) \;\Rightarrow\; a \in \mathbb{R}_+^d \text{ and } a'\mathbb{1}_d \;\leq\; 1 - q_2\frac{z}{x} \;\leq\; 1 - q_1\frac{z}{x} \;\Longrightarrow\; a \in A^{q_1}(x,z),$$

which means that $A^{q_2}(x,z) \subseteq A^{q_1}(x,z)$.

2. Fix $q \in (0,1)$, and consider the decreasing sequence $q_n = q + \frac{1}{n}$, $n \in \mathbb{N}^*$. For any $(x,z) \in \mathcal{S}^{q_n}$, we then have $A^{q_n}(x,z) \subseteq A^{q_{n+1}}(x,z) \subset A^a(x,z)$, which implies that the sequence of increasing sets $A^{q_n}(x,z)$ admits a limit equal to

$$\lim_{n\to\infty} A^{q_n}(x,z) \;=\; \bigcup_{n\geq 1} A^{q_n}(x,z) \;=\; A^q(x,z),$$

since $\lim_{n\to\infty} q_n = q$. This shows the right continuity of $q \mapsto A^q(x,z)$. Similarly, by considering the increasing sequence $q_n = q - \frac{1}{n}$, $n \in \mathbb{N}^*$, we see that for any $(x,z) \in A^q(x,z)$, the sequence of decreasing sets $A^{q_n}(x,z)$ admits a limit equal to

$$\lim_{n\to\infty} A^{q_n}(x,z) \;=\; \bigcap_{n\geq 1} A^{q_n}(x,z) \;=\; A^q(x,z),$$

since $\lim_{n\to\infty} q_n = q$. This proves the continuity in q of the set $A^q(x,z)$.

3. Fix $q \in (0,1)$, and (x_1,z), $(x_2,z) \in \mathcal{S}^q$ s.t. $x_1 \leq x_2$. Then,

$$a \in A^q(x_1,z) \implies a \in \mathbb{R}^d_+ \text{ and } a'\mathbb{1}_d \leq 1 - q\frac{z}{x_1} \leq 1 - q\frac{z}{x_2} \implies a \in A^q(x_2,z),$$

which shows that $A^q(x_1,z) \subseteq A^q(x_2,z)$.

4. Fix $q \in (0,1)$, $(x,z) \in A^a(x,z)$. Then, for any a_1, a_2 of the set $A^q(x,z)$, and $\beta \in (0,1)]$, and denoting by $a_3 = \beta a_1 + (1-\beta)a_2 \in \mathbb{R}^d_+$, we have

$$a_3'\mathbb{1}_d \;=\; \beta a_1'\mathbb{1}_d + (1-\beta)a_2'\mathbb{1}_d \;\leq\; \beta\big(1 - q\frac{z}{x}\big) + (1-\beta)\big(1 - q\frac{z}{x}\big) \;=\; 1 - q\frac{z}{x}.$$

This proves the convexity of the set $A^q(x,z)$.

4. The homogeneity property of $A^q(x,z)$ is obvious from its very definition. □

5.A.5 Proof of Lemma 5.3.5

Proof. We prove the result by backward induction on time k from the dynamic programming equation for the value function.

• At time N, we have for all $\lambda > 0$,

$$v_N(\lambda x, \lambda z, \mu) \;=\; \frac{(\lambda x)^p}{p} = \lambda^p v_N(x,z,\mu),$$

which shows the required homogeneity property.

• Now, assume that the homogeneity property holds at time $k+1$, i.e $v_{k+1}(\lambda x, \lambda z, \mu) = \lambda^p v_{k+1}(x,z,\mu)$ for any $\lambda > 0$. Then, from the backward relation (5.9), and the homogeneity property of $A^q(x,z)$ in Lemma 5.3.4, it is clear that v_k inherits from v_{k+1} the homogeneity property. □

5.A.6 Proof of Lemma 5.3.6

Proof. **1.** We first show by backward induction that $r \mapsto w_k(r,\cdot)$ is nondecreasing in on $[q,1]$ for all $k \in [\![0,N]\!]$.

• For any $r_1, r_2 \in [q,1]$, with $r_1 \leq r_2$, and $\mu \in \mathcal{M}_+$, we have at time N

$$w_N(r_1,\mu) \;=\; U(r_1)\mu(\mathbb{R}^d) \;\leq\; U(r_2)\mu(\mathbb{R}^d) \;=\; w_N(r_2,\mu).$$

This shows that $w_N(r,\cdot)$ is nondecreasing on $[q,1]$.
• Now, suppose by induction hypothesis that $r \mapsto w_{k+1}(r,\cdot)$ is nondecreasing. Denoting by $Y_k := e^{R_k} - \mathbb{1}_d$ the random vector valued in $(-1,\infty)^d$, we see that for all $a \in A^q(r_1)$

$$\min\Big[1, r_1\big(1 + a'Y_{k+1}\big)\Big] \;\leq\; \min\Big[1, r_2\big(1 + a'Y_{k+1}\big)\Big], \quad a.s.$$

since $1 + a'Y_{k+1} \geq 1 - a'\mathbb{1}_d \geq q\frac{1}{r_1} \geq 0$. Therefore, from backward dynamic programming Equation (5.11), and noting that $A^q(r_1) \subset A^q(r_2)$, we have

$$\begin{aligned} w_k(r_1,\mu) &= \sup_{a \in A^q_{(}r_1)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min\big[1, r_1\big(1 + a'Y_{k+1}\big)\big], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\ &\leq \sup_{a \in A^q(r_2)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min\big[1, r_2\big(1 + a'Y_{k+1}\big)\big], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \;=\; w_k(r_2,\mu), \end{aligned}$$

which shows the required nondecreasing property at time k.

2. We prove the concavity of $r \in [q,1] \mapsto w_k(r,\cdot)$ by backward induction for all $k \in [\![0,N]\!]$. For $r_1, r_2 \in [q,1]$, and $\lambda \in (0,1)$, we set $r = \lambda r_1 + (1-\lambda)r_2$, and for $a_1 \in A^q(r_1)$, $a_2 \in A^q(r_2)$, we set $a = \big(\lambda r_1 a_1 + (1-\lambda)r_2 a_2\big)/r$ which belongs to $A^q(r)$. Indeed, since $a_1, a_2 \in \mathbb{R}^d_+$, we have $a \in \mathbb{R}^d_+$, and

$$a \;=\; \Big(\frac{\lambda r_1 a_1 + (1-\lambda)r_2 a_2}{r}\Big)'\mathbb{1}_d \;\leq\; \frac{\lambda r_1}{r}\big(1 - \frac{q}{r_1}\big) + \frac{(1-\lambda)r_2}{r}\big(1 - \frac{q}{r_2}\big) \;=\; 1 - \frac{q}{r}.$$

• At time N, for fixed $\mu \in \mathcal{M}_+$, we have

$$\begin{aligned} w_N\big(\lambda r_1 + (1-\lambda)r_2, \mu\big) &= U(\lambda r_1 + (1-\lambda)r_2) \\ &\geq \lambda U(r_1) + (1-\lambda)U(r_2) \;=\; \lambda w_N(r_1,\mu) + (1-\lambda)w_N(r_2,\mu), \end{aligned}$$

since U is concave. This shows that $w_N(r,\cdot)$ is concave on $[q,1]$.
• Suppose now the induction hypothesis holds true at time $k+1$: $w_{k+1}(r,\cdot)$ is concave on $[q,1]$. From the backward dynamic programming relation (5.11), we then have

$$\begin{aligned} &\lambda w_k(r_1,\mu) + (1-\lambda)w_k(r_2,\mu) \\ \leq\; &\lambda\mathbb{E}\Big[w_{k+1}\big(\min[1, r_1(1 + a_1'Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\ &\qquad +(1-\lambda)\mathbb{E}\Big[w_{k+1}\big(\min[1, r_2(1 + a_2'Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\ \leq\; &\mathbb{E}\Big[w_{k+1}\big(\lambda\min[1, r_1(1 + a_1'Y_{k+1})] + (1-\lambda)\min[1, r_2(1 + a_2'Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\ =\; &\mathbb{E}\Big[w_{k+1}\big(\min[1, r(1 + a'Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \;\leq\; w_k(r,\mu), \end{aligned}$$

where we used for the second inequality, the induction hypothesis joint with the concavity of $x \mapsto \min(1,x)$, and the nondecreasing monotonicity of $r \mapsto w_{k+1}(r,\cdot)$. This shows the required inductive concavity property of $r \mapsto w_k(r,\cdot)$ on $[q,1]$. □

Bibliography

Aguilar, O. and M. West (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics 18*(3), 338–357.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013). Value and momentum everywhere,. *Journal of Finance 68*(3), 929–986.

Avramov, D. and G. Zhou (2010). Bayesian portfolio analysis. *Annual Review of Financial Economics 2*(1), 25–47.

Bachouch, A., C. Huré, N. Langrené, and H. Pham (2018a). Deep neural networks algorithms for stochastic control problems on finite horizon, part 1: convergence analysis. *arXiv:1812.04300*.

Bachouch, A., C. Huré, N. Langrené, and H. Pham (2018b). Deep neural networks algorithms for stochastic control problems on finite horizon, part 2: numerical applications. *arXiv preprint arXiv:1812.05916, to appear in Methodology and Computing in Applied Probability*.

Bain, A. and D. Crisan (2009). *Fundamentals of Stochastic Filtering*. Springer.

Barry, C. B. (1974). Portfolio analysis under uncertain means, variances, and covariances. *The Journal of Finance 29*(2), 515–522.

Bauder, D., T. Bodnar, N. Parolya, and W. Schmid (2020). Bayesian mean-variance analysis: optimal portfolio selection under parameter uncertainty. *Quantitative Finance*, 1–22.

Bellman, R. (1952, 6). On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the U.S.A. 38*, 716–719.

Bensoussan, A. and J. L. Lions (1978). *Applications des Inéquations Variationelles en Contrôle Stochastique*. Dunod.

Benth, F. E. and K. H. Karlsen (2005). A PDE representation of the density of the minimal entropy martingale measure in stochastic volatility markets. *Stochastics an International Journal of Probability and Stochastic Processes 77*(2), 109–137.

Best, M. J. and R. R. Grauer (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The Review of Financial Studies 4*(2), 315–342.

Bismuth, A., O. Guéant, and J. Pu (2019). Portfolio choice, portfolio liquidation, and portfolio transition under drift uncertainty. *Mathematics and Financial Economics 13*(4), 661–719.

Black, F. and R. Litterman (1992). Global portfolio optimization. *Financial Analysts Journal 48*(5), 28–43.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*(3), 637–654.

Blitz, D. and P. Van Vliet (2008). Global tactical cross-asset allocation: applying value and momentum across asset classes. *Journal of Portfolio Management 35*, 23–28.

Bodnar, T., S. Mazur, and Y. Okhrin (2017). Bayesian estimation of the global minimum variance portfolio. *European Journal of Operational Research 256*(1), 292–307.

Borkar, V. (1989). Optimal control of diffusion processes. *Pitman Research Notes in Mathematics* (203).

Boyd, S., E. Lindström, H. Madsen, and P. Nystrup (2019). Multi-period portfolio selection with drawdown control. *Annals of Operations Research 282*(1-2), 245–271.

Browne, S. and W. Whitt (1996). Portfolio choice and the bayesian kelly criterion. *Advances in Applied Probability 28*(4), 1145–1176.

Carhart, M. (1997). On persistence in mutual fund performance. *The Journal of Finance 52*(1), 57–82.

Ceci, C. (2006). Risk minimizing hedging for a partially observed high frequency data model. *Stochastics: An International Journal of Probability and Stochastics Processes 78*(1), 13–31.

Ceci, C. and K. Colaneri (2012). Nonlinear filtering for jump diffusion observations. *Advances in Applied Probability 44*(3), 678–701.

Ceci, C., A. Cretarola, and F. Russo (2014). Bsdes under partial information and financial applications. *Stochastic Processes and their Applications 124*(8), 2628–2653.

Ceci, C. and A. Gerardi (2006). A model for high frequency data under partial information: a filtering approach. *International Journal of Theoretical and Applied Finance 9*(04), 555–576.

Cox, J. C. and C.-f. Huang (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory 49*(1), 33–83.

Cvitanić, J. and I. Karatzas (1992). Convex duality in constrained portfolio optimization. *The Annals of Applied Probability*, 767–818.

Cvitanić, J. and I. Karatzas (1994). On portfolio optimization under "drawdown" constraints. *Constraints, IMA Lecture Notes in Mathematics & Applications 65*.

Cvitanić, J., A. Lazrak, L. Martellini, and F. Zapatero (2006). Dynamic portfolio choice with parameter uncertainty and the economic value of analysts? recommendations. *The Review of Financial Studies 19*(4), 1113–1156.

De Franco, C., J. Nicolle, and H. Pham (2019a). Bayesian learning for the markowitz portfolio selection problem. *International Journal of Theoretical and Applied Finance 22*(07).

De Franco, C., J. Nicolle, and H. Pham (2019b). Dealing with drift uncertainty: a bayesian learning approach. *Risks 7*(1), 5.

Detemple, J. B. (1986). Asset pricing in a production economy with incomplete information. *The Journal of Finance 41*(2), 383–391.

Dothan, M. U. and D. Feldman (1986). Equilibrium interest rates and multiperiod bonds in a partially observable economy. *The Journal of Finance 41*(2), 369–382.

Ekstrom, E. and J. Vaicenavicius (2016). Optimal liquidation of an asset under drift uncertainty. *SIAM Journal on Financial Mathematics 7*(1), 357–381.

El Karoui, N. (1981). Les aspects probabilistes du contrôle stochastique. *Lecture Notes in Mathematics 876*, 73–238.

El Karoui, N., N. D. Huu, and J.-P. Monique (1987). Compactification methods in the control of degenerate diffusions: existence of optimal controls. *Stochastics and Stochastic Reports 20*(3), 169–219.

Elie, R. and N. Touzi (2008). Optimal lifetime consumption and investment under a drawdown constraint. *Finance and Stochastics 12*(3), 299.

Elliott, R. J., L. Aggoun, and J. B. Moore (2008). *Hidden Markov Models: Estimation and Control*. Springer.

Fama, E. and K. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. and K. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Fama, E. and K. French (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies 29*(1), 69–103.

Fleming, W. H. and Rishel (1975). *Deterministic and Stochastic Optimal Control*. Springer.

Fleming, W. H. and H. M. Soner (2006). *Controlled Markov Processes and Viscosity Solutions*. Springer.

Frost, P. A. and J. E. Savarino (1986). An empirical bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis 21*(3), 293–305.

Geczy, C. and M. Samonov (2017). Two centuries of multi-asset momentum (equities, bonds, currencies, commodities, sectors and stocks). *Available at SSRN: http://dx.doi.org/10.2139/ssrn.2607730*.

Gennotte, G. (1986). Optimal portfolio choice under incomplete information. *The Journal of Finance 41*(3), 733–746.

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media.

Gilbarg, D. and N. Trudinger (1985). *Elliptic Differential Equations of Second Order*. Springer.

Grossman, S. J. and Z. Zhou (1993). Optimal investment strategies for controlling drawdowns. *Mathematical finance 3*(3), 241–276.

He, H. and N. D. Pearson (1991). Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. *Journal of Economic Theory 54*(2), 259–304.

Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Networks 4*(2), 251–257.

Ismail, A. and H. Pham (2017). Robust markowitz mean-variance portfolio selection under ambiguous covariance matrix. *29*, 174–207.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering 82*, 35–45.

Kalman, R. E. and R. S. Bucy (1961, 03). New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering 83*(1), 95–108.

Karatzas, I. (1997). Adaptive control of a diffusion to a goal and a parabolic monge–ampere-type equation. *Asian Journal of Mathematics 1*(2), 295–313.

Karatzas, I., J. P. Lehoczky, and S. E. Shreve (1987). Optimal portfolio and consumption decisions for a "small investor" on a finite horizon. *SIAM Journal on Control and Optimization 25*(6), 1557–1586.

Karatzas, I., J. P. Lehoczky, S. E. Shreve, and G.-L. Xu (1991). Martingale and duality methods for utility maximization in an incomplete market. *SIAM Journal on Control and Optimization 29*(3), 702–730.

Karatzas, I., S. E. Shreve, I. Karatzas, and S. E. Shreve (1998). *Methods of Mathematical Finance*. Springer.

Karatzas, I. and X.-X. Xue (1991). A note on utility maximization under partial observations. *Mathematical Finance 1*(2), 57–70.

Karatzas, I. and X. Zhao (2001). Bayesian Adaptative Portfolio Optimization. In *Option Pricing, Interest Rates and Risk Management*. Cambridge University Press.

Keppo, J., H. M. Tan, and C. Zhou (2018). Investment decisions and falling cost of data analytics.

Klein, R. W. and V. S. Bawa (1976). The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics 3*(3), 215–231.

Krylov, N. V. (1980). *Controlled Diffusion Processes*. Springer.

Krylov, N. V. (1987). *Nonlinear Elliptic and Parabolic Equations of the Second Order*, Volume 7. Springer.

Kushner, H. (1975). Existence results for optimal stochastic controls. *Journal of Optimization Theory and Applications 15*(4), 347–359.

Kuwana, Y. (1995). Certainty equivalence and logarithmic utilities in consumption/investment problems. *Mathematical Finance 5*(4), 297–309.

Ladyzhenskaia, O. A., V. A. Solonnikov, and N. N. Ural'ceva (1968). *Linear and Quasi-linear Equations of Parabolic Type*. American Mathematical Society.

Lakner, P. (1995). Utility maximization with partial information. *Stochastic Processes and their Applications 56*(2), 247–273.

Lakner, P. (1998). Optimal trading strategy for an investor: the case of partial information. *Stochastic Processes and their Applications 76*(1), 77–97.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics 47*(1), 13–37.

Lions, P.-L. (1983). Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. *Communications in Partial Differential Equations 8*(11), Part I, 1101–1134, Part II, 1229–1276.

Markowitz, H. (1952, 3). Portfolio Selection. *The Journal of Finance 7*(1), 77–91.

Merton, R. (1971). Optimum consumption and portfolio rules in a continuous-time framework. *Journal of Economic Theory 6*, 213–214.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 247–257.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics 8*(4), 323–361.

Nisio, M. (1981). Lectures on stochastic control theory. *ISI Lecture Notes 9*.

Øksendal, B. and A. Sulem (2004). *Applied Stochastic Control of Jump Diffusions*. Springer.

Pham, H. (2002). Smooth solutions to optimal investment models with stochastic volatilities and portfolio constraints. *Applied Mathematics & Optimization 46*(1).

Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer.

Pham, H. (2011). Portfolio optimization under partial observation: theoretical and numerical aspects. In *The Oxford Handbook of Nonlinear Filtering*.

Pliska, S. R. (1986). A stochastic calculus model of continuous trading: optimal portfolios. *Mathematics of Operations Research 11*(2), 371–382.

Redeker, I. and R. Wunderlich (2018). Portfolio optimization under dynamic risk constraints: Continuous vs. discrete time trading. *Statistics & Risk Modeling 35*(1-2), 1–21.

Rishel, R. (1999). Optimal portfolio management with partial observations and power utility function. In *Stochastic Analysis, Control, Optimization and Applications*. Springer.

Rogers, L. C. G. (2001). The relaxed investor and parameter uncertainty. *Finance and Stochastics 5*(2), 131–154.

Samuelson, P. (1969). Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economics and Statistics 51*(3), 239–246.

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

Yong, J. and X. Y. Zhou (2000). *Stochastic Controls, Hamiltonian Systems and HJB Equations*. Springer.

Zhou, X. Y. and D. Li (2000, 1). Continuous-time mean-variance portfolio selection: a stochastic LQ framework. *Applied Mathematics & Optimization 42*(1), 19–33.

Zohar, G. (2001, 02). A generalized Cameron-Martin formula with applications to partially observed dynamic portfolio optimization. *Mathematical Finance 11*, 475–494.