

HEC Montréal
École affiliée à l'Université de Montréal

OPTIMISATION ROBUSTE
D'UN PORTEFEUILLE D'OPTIONS

Par
Jean-François Fortin

Sciences de la gestion
(Ingénierie financière)

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences
(M.Sc.)

Avril 2015
© Jean-François Fortin, 2015

Remerciements

Je tiens d'abord et avant tout à remercier mon directeur de recherche, Erick Delage, pour sa disponibilité, sa générosité (tant scientifique que monétaire), son implication, de même que pour la patience infinie dont il a fait preuve tout au long de ce travail.

Je remercie également HEC Montréal ainsi que l'Institut de Finance Mathématique de Montréal (IFM2) pour l'aide financière qu'ils m'ont si généreusement apportée.

Merci beaucoup!

Table des matières

Remerciements	i
1 Introduction	1
2 Revue de la littérature	3
2.1 Analyse et mesure du risque	3
2.1.1 Cohérence d'une mesure de risque	4
2.1.2 Valeur-à-risque (VaR)	6
2.1.3 Valeur-à-risque conditionnelle (CVaR)	6
2.2 Recherche opérationnelle et optimisation robuste	8
2.2.1 Optimisation robuste de portefeuille	9
2.3 Mesures de performance	12
2.4 Contributions	13
3 Méthodologie et modélisation	14
3.1 Optimisation robuste et DRSP (Delage et Ye)	15
3.2 Valeur-à-risque conditionnelle (Rockafellar et Uryasev)	20
3.3 Minimisation robuste de la CVaR	24
3.4 Optimisation d'un portefeuille d'options	25
3.5 Optimisation robuste du ratio de Sharpe	31

4	Expérimentation et résultats	36
4.1	Expérience 1	36
4.2	Expérience 2	43
4.2.1	Méthodologie	43
4.2.2	Résultats	47
4.2.3	Ratio de Sharpe robuste	51
5	Conclusion	55
A	Démonstrations	57
A.1	Équation (3.31)	57
A.2	Proposition 3.1	59
A.3	Proposition 3.2	62
A.4	Proposition 3.3	64
B	Problème non robuste d'optimisation de la CVaR	67

Table des figures

3.1	VaR et CVaR de la distribution des pertes quotidiennes du S&P 500 entre janvier 1997 et janvier 2013.	22
4.1	Frontières efficientes robustes et non robustes obtenues à partir des portefeuilles optimaux robustes (R) et non robustes (NR), avec et sans options ($\beta = 0.65$).	41
4.2	Poids des actifs dans les différents portefeuilles optimaux en fonction du rendement espéré ($\beta = 0.65$).	42
4.3	Rendements cumulés des six indices entre janvier 1997 et janvier 2013.	44
4.4	Rendements cumulés des portefeuilles robustes (R) et non robustes (NR) avec et sans options.	47
4.5	Rendements cumulés des composantes « actions » et « options » des portefeuilles avec options robustes (R) et non robustes (NR).	49
4.6	Composition des portefeuilles robustes (R) et non robustes (NR), avec et sans options.	51
4.7	Rendements cumulés des portefeuilles « Sharpe-optimaux » robustes et non robustes, avec et sans options.	52
4.8	Ratios de Sharpe optimaux des portefeuilles avec et sans options, robustes (R) et non robustes (NR).	54

Liste des tableaux

4.1	Symboles et numéros de référence (Option ID) des différents indices.	43
4.2	Statistiques descriptives des rendements mensuels des six indices de 1997 à 2012.	46
4.3	Statistiques descriptives des rendements mensuels des six indices de 1997 à 2004.	46
4.4	Statistiques descriptives des rendements mensuels des six indices de 2004 à 2012.	46
4.6	Statistiques descriptives des rendements mensuels hors-échantillon des portefeuilles robustes (R) et non robustes (NR) avec options.	48
4.5	Statistiques descriptives des rendements mensuels hors-échantillon des portefeuilles robustes (R) et non robustes (NR) sans options.	48
4.7	CVaR réalisée hors-échantillon des portefeuilles robustes (R) et non robustes (NR), avec et sans options, en fonction de β	50
4.8	Statistiques descriptives des rendements mensuels des portefeuilles « Sharpe-optimaux » robustes (R) et non robustes (NR) sans options.	53
4.9	Statistiques descriptives des rendements mensuels des portefeuilles « Sharpe-optimaux » robustes (R) et non robustes (NR) avec options.	53

Chapitre 1

Introduction

L'effondrement des marchés boursiers à l'automne 2008 a malencontreusement rappelé l'importance de la gestion du risque de marché en gestion de portefeuille. Le risque de marché est défini comme le risque lié à la fluctuation de la valeur marchande des instruments financiers au sein d'un portefeuille. L'étude du risque de marché en gestion de portefeuille remonte au travail fondamental de Markowitz en 1952 [Mar52]. Celui-ci y postule un cadre d'optimisation moyenne-variance où le risque du portefeuille (c.-à-d. sa variance) est minimisé pour un certain rendement moyen. Malgré sa propagation et son adoption au sein de l'industrie financière, le cadre d'optimisation moyenne-variance a fait l'objet de critiques virulentes depuis sa parution.

L'un des reproches les plus fréquemment énoncés à l'encontre du cadre de Markowitz a pour objet son manque de robustesse. Entre autres, le cadre moyenne-variance est particulièrement sensible à l'estimation de ses deux paramètres, à savoir l'espérance et la matrice de variance-covariance. Bon nombre d'auteurs ont en effet démontré que les paramètres ont communément tendance à être instables lorsqu'ils sont estimés à partir d'échantillons de données historiques.

L'optimisation robuste est une discipline relativement jeune dont la raison d'être est justement d'atténuer l'impact de l'incertitude dans les paramètres d'entrée d'un problème d'optimisation de nature aléatoire par l'entremise de bornes sur les paramètres en question ou leur distribution. La solution optimale d'un problème d'optimisation robuste ne tient alors compte que de la « pire » réalisation de chacun des paramètres à l'intérieur de ces bornes.

L'objectif de ce mémoire est d'étudier le problème d'optimisation robuste de la valeur-à-risque conditionnelle d'un portefeuille d'options d'achat et de vente européennes lorsque l'incertitude sur la distribution des rendements des actifs est caractérisée par l'ensemble distributionnel introduit par Delage et Ye [DY10]. À l'instar de Zymler et al. [ZKR12], le rendement des options est défini par une fonction linéaire par morceaux des rendements des actifs sous-jacents. Le problème d'optimisation robuste ainsi obtenu se traduit par un programme semi-défini positif qui peut être résolu en temps polynomial pour tout niveau de précision.

Le travail de recherche est structuré comme suit. Au chapitre 2, nous survolons les publications académiques existantes en lien avec la gestion du risque et l'optimisation robuste de portefeuille. Dans le chapitre subséquent, nous étudions d'abord les travaux de Delage et Ye [DY10] et de Rockafellar et Uryasev [RU00], à partir desquels nous formulons une version robuste du problème d'optimisation de la valeur-à-risque conditionnelle (CVaR). Nous fournissons de plus une formulation robuste du « pire » ratio de Sharpe, où le risque du portefeuille est caractérisé par sa CVaR. Au chapitre 4, nous présentons les résultats d'expériences numériques sur les divers modèles d'optimisation robustes avec et sans options. Les résultats sont ensuite comparés à une version non robuste de ces mêmes problèmes d'optimisation.

Chapitre 2

Revue de la littérature

La revue de la littérature se découpe en trois volets. D’abord, nous discutons d’un bon nombre de travaux de premier plan ayant trait à la gestion et à la mesure du risque financier. Puis, nous nous penchons sur le sujet de l’optimisation robuste et sur les avancées théoriques et méthodologiques en lien avec l’optimisation de portefeuille. Enfin, nous traitons de divers travaux relatifs à l’optimisation de mesures de performance.

2.1 Analyse et mesure du risque

Markowitz a été l’un des premiers à formaliser le compromis entre le risque d’un portefeuille et son rendement. Dans l’article « *Portfolio Selection* », celui-ci met de l’avant un cadre d’optimisation moyenne-variance (MV) où le risque du portefeuille, en l’occurrence sa variance, est minimisé pour un certain rendement moyen [Mar52]. Le cadre de Markowitz est d’ailleurs au coeur de la théorie moderne de la gestion de portefeuille.

Malgré sa popularité, le cadre moyenne-variance a fait l'objet de vives critiques, notamment en raison de l'emploi de la variance comme mesure du risque agrégé du portefeuille. La variance a d'abord pour inconvénient majeur de pénaliser autant les déviations à la hausse qu'à la baisse (c.-à-d. autant les pertes que les profits). Markowitz lui-même relève le problème dans un ouvrage subséquent et suggère, dans un autre ouvrage, de remplacer la variance par la semi-variance [Mar59]. À la différence de la variance, la semi-variance ne considère que la part des rendements sous la moyenne.

Borch [Bor69] puis Chamberlain [Cha83] notent que le cadre MV sous-entend une fonction d'utilité quadratique ou une distribution jointe des rendements elliptique¹. Ces hypothèses, en plus d'être fortes, sont plutôt contre-intuitives. D'abord, l'utilité marginale dérivée d'une fonction d'utilité quadratique tend à décroître lorsque le niveau de richesse dépasse un certain seuil (c.-à-d. le maximum de la fonction). En outre, une telle fonction d'utilité se distingue par une aversion absolue au risque croissante. Autrement dit, plus l'investisseur est riche, moins il prend de risques, et vice-versa. En dernier lieu, la distribution des rendements observés sur les marchés tend à s'éloigner de la normalité (plus asymétrique et leptocurtique, par exemple), comme le remarque Fama [Fam65]. Ainsi, en réponse aux lacunes manifestes de la variance et du cadre d'optimisation de Markowitz, plusieurs autres mesures de risque ont, depuis, vu le jour.

2.1.1 Cohérence d'une mesure de risque

Artzner, Delbaen, Eber et Heath [ADEH99] introduisent la notion de cohérence d'une mesure de risque basée sur la satisfaction de quatre axiomes, à savoir l'invariance par translation, la sous-additivité, l'homogénéité positive et la monotonie. Une mesure dite « cohérente » au sens d'Artzner et al. a pour avantage, entre autres, d'être conforme à la logique et à l'intuition économique et financière.

1. Une distribution elliptique est entièrement caractérisée par ses deux premiers moments. La loi normale, par exemple, en est un cas particulier.

Notons par X et par Y deux variables aléatoires (p. ex. les pertes de deux portefeuilles à un point dans le futur). Une mesure de risque est simplement une fonction ρ qui joint un réel $\rho(X)$ à un risque X . Mathématiquement, les quatre axiomes d'Artzner et al. s'expriment ainsi :

$$\begin{aligned} \rho(X + a) &= \rho(X) + a, \quad \forall X, a \in \mathbb{R} && \text{(invariance par translation)} \\ \rho(X + Y) &\leq \rho(X) + \rho(Y), \quad \forall X, Y && \text{(sous-additivité)} \\ \rho(\lambda X) &= \lambda\rho(X), \quad \forall X, \lambda \in \mathbb{R}^+ && \text{(homogénéité positive)} \\ \forall X, Y, \text{ si } X &\leq Y, \text{ alors } \rho(Y) &\leq \rho(X). && \text{(monotonie)} \end{aligned}$$

Notons que l'égalité $\rho(X - \rho(X)) = 0$ découle naturellement de l'axiome d'invariance par translation. Pour cette raison, la quantité $\rho(X) > 0$ s'interprète plus souvent qu'autrement comme le plus petit montant de capital à ajouter pour rendre nul le risque du portefeuille. L'axiome de sous-additivité, de son côté, garantit que la diversification n'augmente pas le risque agrégé d'une position, de sorte que le risque total d'un portefeuille soit toujours borné par la somme des risques individuels qui le composent. Föllmer et Schied [FS02] rappellent que le risque d'une position n'augmente pas toujours linéairement avec la taille de la position. Par exemple, un risque de liquidité additionnel peut s'ajouter au risque agrégé du portefeuille lorsque la taille de la position s'accroît considérablement. Les deux auteurs proposent donc de relaxer les deux axiomes d'homogénéité positive et de sous-additivité et de les remplacer plutôt par l'axiome de convexité ci-dessous :

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y), \quad \forall \lambda \in [0, 1]. \quad \text{(convexité)}$$

Une mesure de risque est alors dite « convexe » si elle satisfait aux axiomes de convexité, de monotonie et d'invariance par translation. Notons que la variance n'est pas une mesure de risque cohérente ni convexe puisque qu'elle ne répond pas aux axiomes d'invariance par translation, d'homogénéité positive et de sous-additivité².

2. Par définition, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$, où $\text{Cov}(X, Y)$ peut être soit négatif, soit positif. Si $\text{Cov}(X, Y) > 0$, alors $\text{Var}(X + Y) > \text{Var}(X) + \text{Var}(Y)$.

2.1.2 Valeur-à-risque (VaR)

La notion de valeur-à-risque (VaR) a été démocratisée en 1994 par la banque d'affaires J.P. Morgan [Ris94]. Depuis, la VaR s'est imposée quasi unanimement au sein de l'industrie bancaire notamment en raison de son adoption par le Comité de Bâle (Bâle II). La valeur-à-risque correspond au niveau de perte maximal attendu à l'intérieur d'un horizon temporel, pour un certain niveau de confiance (habituellement entre 1 et 10%). La VaR n'est en fait qu'un quantile de la distribution de perte du portefeuille.

La popularité de la VaR n'est pas sans fondement. En effet, la VaR est une mesure scalaire qui réduit le risque d'un actif ou d'un portefeuille à un seul chiffre. Par le fait même, elle est à la fois relativement simple d'usage et d'interprétation. La valeur-à-risque est, de plus, une mesure asymétrique qui ne tient compte que de la queue défavorable de la distribution (contrairement à la variance, par exemple). Malgré tout, la VaR n'est malheureusement pas sans défaut. D'une part, la valeur-à-risque n'est ni convexe, ni cohérente, puisqu'elle ne répond pas à l'axiome de sous-additivité, sauf dans le cas où la distribution jointe des rendements est elliptique [ADEH99]. Gaivoronski et Pflug [GP05] remarquent, par le fait même, que l'optimisation de la VaR est une tâche plus complexe que l'optimisation de la variance ou de la CVaR en raison de son caractère non convexe. D'autre part, la VaR est également incapable de mesurer l'ampleur de la perte au-delà d'elle-même [YY02].

2.1.3 Valeur-à-risque conditionnelle (CVaR)

La valeur-à-risque conditionnelle (CVaR) s'est plus récemment retrouvée à l'avant-plan avec la parution d'un papier de Rockafellar et Uryasev [RU00]. Pour une distribution de perte quelconque, la CVaR correspond simplement à l'espérance conditionnelle des pertes plus grandes que la VaR. Intuitivement, la CVaR est une mesure de l'ampleur moyenne des pertes extrêmes. Rockafellar et Uryasev [RU00] reformulent le problème de minimisation de

la valeur-à-risque conditionnelle en un problème plus simple et plus fonctionnel qui n'exige pas a priori le calcul de la VaR, mais qui peut néanmoins la dériver comme sous-produit.

La valeur-à-risque conditionnelle se veut donc une réponse formelle aux lacunes explicites de la VaR. Pflug [Pfl00] puis Acerbi et Tasche [AT02] démontrent que la CVaR est notamment une mesure de risque convexe. Qui plus est, la CVaR tient compte de toute la queue défavorable de la distribution (et non seulement d'un quantile).

Exemple 2.1. Soit X et Y , deux variables aléatoires dont les fonctions de masse sont respectivement :

$$f_X(x) = \begin{cases} 1\% & \text{si } x = -1\,000\,000 \\ 1\% & \text{si } x = 0 \\ 98\% & \text{si } x = 1 \end{cases}$$

et

$$f_Y(x) = \begin{cases} 1\% & \text{si } x = -1 \\ 1\% & \text{si } x = 0 \\ 98\% & \text{si } x = 1 \end{cases} .$$

Nous remarquons d'emblée que la variable aléatoire X est à l'évidence plus risquée que la variable aléatoire Y , en raison de la queue plus prononcée de sa distribution. Or, la VaR de niveau 99% est la même dans les deux cas, à savoir $\text{VaR}_{0.99}(X) = \text{VaR}_{0.99}(Y) = 0$.

D'un autre côté, nous avons $\text{CVaR}_{0.99}(X) = -1$ et de $\text{CVaR}_{0.99}(Y) = -1\,000\,000$. Dans ce cas-ci, la CVaR peint donc un portrait plus représentatif du risque réel de ces deux variables aléatoires.

2.2 Recherche opérationnelle et optimisation robuste

La recherche opérationnelle est une discipline relativement jeune qui fait notamment appel à la modélisation, la simulation et l'optimisation pour aider le décideur ou le gestionnaire à prendre la décision la plus rationnelle possible. De fait, la recherche opérationnelle est aussi qualifiée de « science de la décision ».

L'optimisation robuste joue un rôle clef en recherche opérationnelle. Il n'est pas rare, par exemple, qu'un problème d'optimisation compte une ou plusieurs variables dont nous ignorons la valeur au moment de prendre la décision. L'objectif de l'optimisation robuste est alors de produire une solution optimale « robuste » compte tenu de l'incertitude dans les paramètres d'entrée du problème à optimiser. L'optimisation robuste n'est pas une classe de problèmes en soi, mais, plutôt, une manière d'approcher un problème pour prendre en compte l'incertitude en question.

L'incertitude peut se manifester tant dans la fonction objectif que dans les contraintes d'un problème d'optimisation. Dans tous les cas, l'objectif est de trouver une solution qui tient compte du pire cas possible, de sorte que la solution demeure « valable » quel que soit le scénario considéré. De façon générale, l'approche robuste se résume à borner chacun des paramètres incertains à l'intérieur d'un ensemble d'incertitude restreint, puis à optimiser sur le ou les pires scénarios possible au sein de cet ensemble. Sous certaines hypothèses sur les contraintes de départ (linéaires, quadratiques, etc.) et les ensembles d'incertitude (des intervalles ou des ellipsoïdes, par exemple), il est possible de formuler le problème d'optimisation robuste comme un problème d'optimisation convexe qui peut être résolu dans un temps « raisonnable » par programmation linéaire, conique (SOCP) ou semi-définie positive (SDP).

L'optimisation robuste a récemment connu un regain de popularité au sein de la communauté académique à la suite des travaux de Ben-Tal et Nemirovski [BTN98] et d'El Ghaoui, Oustry et Lebret [EOL98], entre autres, ainsi qu'à l'avènement de la programmation conique et semi-définie et à l'élaboration d'algorithmes de points intérieurs pour la résolution de problèmes d'optimisation convexes (voir Nesterov et Nemirovski [NN94], par exemple).

2.2.1 Optimisation robuste de portefeuille

La faible robustesse du cadre moyenne-variance a été soulevée par bon nombre d'auteurs au fil des ans. Michaud [Mic98], entre autres, remarque que le cadre d'optimisation de Markowitz est particulièrement sensible aux erreurs d'estimation de ses paramètres, à savoir la moyenne et la matrice de variance-covariance. Dans le même ordre d'idées, les travaux de Frankfurter, Phillips et Seagle [FPS71], Black et Litterman [BL92] et Chopra et Zeima [CZ93] montrent qu'une variation, aussi infime soit-elle, dans l'estimation de la moyenne ou de la variance peut avoir une incidence colossale sur la composition et la validité du portefeuille optimal.

Lobo et Boyd ont été parmi les premiers à étendre les notions d'optimisation convexe et de SDP (*Semidefinite Programming* ou programmation semi-définie) aux problèmes d'optimisation de portefeuille [LB99]. Dans leur papier, les auteurs s'intéressent au problème d'optimisation du risque maximal d'un portefeuille où l'incertitude est, entre autres, délimitée par des bornes sur la matrice de variance-covariance.

Pour leur part, Goldfarb et Iyengar analysent un modèle factoriel sous l'hypothèse de normalité où les rendements du portefeuille sont caractérisés par divers facteurs de marché [GI03]. Les auteurs formulent une version robuste du problème à l'aide d'un ensemble d'incertitude ellipsoïdal qui se traduit par un SOCP (*Second-Order Cone Programming* ou cône quadratique). Garlappi, Uppal et Wang [GUW07] se penchent sur le problème d'optimisation robuste d'un portefeuille dans le cadre moyenne-variance pour lequel l'in-

certitude ne se manifeste que dans le rendement espéré des actifs. Lorsque l'incertitude sur la moyenne est déterminée par un ellipsoïde, Garlappi et al. démontrent que le problème se réduit à un SOCP.

El Ghaoui, Oks et Oustry [EEO03], de leur côté, s'attellent à l'optimisation d'une version robuste de la valeur-à-risque qui ne postule qu'une connaissance partielle des deux premiers moments. Les auteurs démontrent que le problème peut être résolu à l'aide d'un SDP pour différents types d'ensembles d'incertitude.

Zymler, Kuhn et Rustem [ZKR12] présentent deux modèles d'optimisation robuste de la valeur-à-risque lorsque le portefeuille compte non seulement des actions, mais aussi des options d'achat et vente. Dans le premier cas, le rendement des options est caractérisé par une fonction convexe linéaire par morceaux du rendement des actifs sous-jacents. Dans le second, le rendement des options est déterminé par une approximation delta-gamma et se traduit par une fonction quadratique (possiblement concave) du rendement des actifs sous-jacents. Dans les deux cas, les auteurs présument que les deux premiers moments de la distribution des rendements des sous-jacents sont connus et finis. Notons que le second modèle rappelle notamment celui étudié par Alexander, Coleman et Li [ACL06] pour l'optimisation de la CVaR. Zymler, Kuhn et Rustem [ZKR13] confirment par la suite que les deux modèles d'optimisation de la VaR produisent les mêmes solutions optimales lorsque la VaR est remplacée par la CVaR dans la fonction objectif.

Zhu et Fukushima [ZF09] formulent trois versions robustes de la CVaR (*Worst-Case CVaR* ou WCVaR) pour lesquelles l'incertitude est liée à la distribution du portefeuille. La première version postule que l'ensemble d'incertitude est composé d'une mixture de distributions continues. Les deux auteurs représentent l'incertitude sur la distribution par une somme pondérée de scénarios possibles pour la distribution, et montrent que le cadre en question se réduit alors à un simple programme linéaire. Les deuxième et troisième versions abordées représentent l'ensemble d'incertitude par un intervalle et un ellipsoïde, respectivement. Dans les deux cas, le problème peut

être reformulé comme un SOCP. Zhu et Fukushima prouvent par la suite que la WCVaR demeure une mesure de risque cohérente. Nolet [Nol09], dans son mémoire de maîtrise, complète l'analyse de Zhu et Fukushima en fournissant des résultats numériques approfondis et en formulant un ratio de performance « rendement excédentaire-CVaR » sous le modèle en question.

Dans un tout autre registre, Popescu [Pop07] met sur pied un cadre d'optimisation général pour la résolution de problèmes stochastiques robustes de type max-min et min-max lorsque la fonction d'utilité est croissante et concave et lorsque ses dérivées sont convexes. L'approche de Popescu ne requiert qu'une connaissance primaire de l'espérance et de la variance des variables aléatoires, et se traduit par un programme paramétrique d'optimisation quadratique. L'auteur présente ensuite une application en gestion robuste de portefeuille où le risque du portefeuille est représenté par une fonction d'utilité concave.

De leur côté, Bertsimas, Doan, Natarajan et Teo [BDNT09] font l'étude du problème plus restrictif de programmation stochastique linéaire de type min-max pour des classes de distributions caractérisées par le premier et le deuxième moments. Bertsimas et al. montrent, notamment, comment reformuler le problème de façon à pouvoir le résoudre en temps polynomial à l'aide d'un programme semi-défini.

Delage et Ye [DY10] formulent un problème convexe d'optimisation robuste où l'incertitude est caractérisée par un intervalle de confiance autour de la moyenne et de la matrice de variance-covariance des paramètres aléatoires. Leur approche ne requiert d'ailleurs qu'une connaissance de base du support de la distribution. Lorsque la fonction objectif est convexe par rapport aux variables de décisions et concave par morceaux par rapport aux paramètres aléatoires, le problème peut être résolu en temps polynomial à l'aide de la méthode de l'ellipsoïde pour n'importe quel niveau de précision. Les auteurs présentent, de plus, une application en optimisation robuste de portefeuille qui se rapproche de celle de Popescu [Pop07].

2.3 Mesures de performance

Le ratio de Sharpe [Sha02] est traditionnellement défini comme le ratio du rendement excédentaire du portefeuille (c.-à-d. la différence entre son rendement espéré et un taux d'intérêt sans risque) et du risque du portefeuille (c.-à-d. l'écart-type). Intuitivement, le ratio de Sharpe mesure la performance marginale du portefeuille par unité de risque. Le ratio de Sharpe est fortement inspiré du cadre d'optimisation moyenne-variance, et partage malheureusement les mêmes faiblesses. En effet, en plus de faire appel à une mesure de risque non-cohérente, le ratio de Sharpe est, lui aussi, sujet aux erreurs d'estimation.

Le ratio Oméga, introduit par Keating et Shadwick [KS02], correspond au ratio entre les gains et les pertes d'un portefeuille par rapport à un seuil donné, pondérés par leurs probabilités. Par construction, le ratio Oméga tient compte de tous les moments de la distribution des rendements (contrairement au ratio de Sharpe, par exemple). L'optimisation du ratio Oméga est, par contre, un problème non convexe. Plus récemment, Kapsos, Christofides et Rustem [KCR14] formulent une version robuste du problème de maximisation du ratio Oméga sous l'hypothèse d'une connaissance partielle de la distribution des rendements. Le problème se traduit par un SOCP qui peut être résolu en temps polynomial pour plusieurs types d'ensembles d'incertitude.

Meskarian et Xu [MX13] présentent un cadre plus général d'optimisation robuste d'un couple rendement-risque où le risque est défini comme l'espérance des pertes au-delà d'un certain seuil. Les auteurs reformulent le ratio de façon à former un problème d'optimisation convexe.

2.4 Contributions

Les contributions de ce mémoire s'organisent comme suit. D'abord, au chapitre 3, nous nous inspirons du travail de Delage et Ye [DY10] pour formuler un cadre d'optimisation robuste d'un portefeuille d'options d'achat et de vente européennes qui tient compte de l'incertitude inhérente à l'estimation des moments de la distribution des rendements des actifs sous-jacents. Le modèle présenté ici fait l'hypothèse que l'incertitude sur la distribution des variables aléatoires du problème d'optimisation peut être caractérisée par des bornes ellipsoïdales autour des deux premiers moments de la distribution en question. Par la suite, nous développons pour la première fois une version robuste du ratio de Sharpe où le risque du portefeuille est caractérisé par sa CVaR. Enfin, au chapitre 4, nous fournissons des résultats numériques détaillés à partir de données simulées et historiques.

En cours de rédaction, nous avons pris connaissance du travail de Zymler et al. [ZKR12] qui s'effectuait au même moment et de façon indépendante. Les auteurs abordent un problème d'optimisation robuste de la valeur-à-risque d'un portefeuille avec options. Nous nous sommes d'ailleurs inspirés de leur notation au chapitre 3. Zymler et al. font l'hypothèse que les deux premiers moments de la distribution des rendements des sous-jacents sont connus et égaux à leurs valeurs empiriques et proposent des méthodes permettant d'identifier les portefeuilles qui sont robustes aux distributions satisfaisant ces moments. À la différence du modèle étudié ici, celui de Zymler et al. ne tient pas compte de l'incertitude associée à l'estimation des moments. Nous proposons, de plus, une analyse empirique beaucoup plus poussée et complète des modèles de minimisation de la CVaR et de maximisation du ratio de Sharpe.

Chapitre 3

Méthodologie et modélisation

L'objectif de ce chapitre est de jeter les bases de l'optimisation robuste de la valeur-à-risque conditionnelle d'un portefeuille d'options d'achat et de vente européennes.

À cet effet, nous introduisons dans un premier temps le cadre d'optimisation robuste et l'ensemble d'incertitude distributionnel mis de l'avant par Delage et Ye [DY10]. Nous survolons ensuite la méthodologie développée par Rockafellar et Uryasev [RU00] pour l'optimisation de la valeur-à-risque conditionnelle. À partir des précédents résultats, nous formulons enfin des versions robustes de la CVaR et du ratio de Sharpe par rapport au choix de la distribution.

3.1 Optimisation robuste et DRSP (Delage et Ye)

Delage et Ye [DY10] étudient un problème standard d'optimisation convexe du type

$$\min_{\mathbf{x} \in \mathcal{X}} c(\mathbf{x}, \boldsymbol{\xi}),$$

où $c(\mathbf{x}, \boldsymbol{\xi})$ est une fonction de coût (ou de perte) convexe par rapport au vecteur de décision $\mathbf{x} \in \mathbb{R}^m$, qui dépend d'un vecteur aléatoire (ou incertain) $\boldsymbol{\xi} \in \mathbb{R}^n$. Le vecteur \mathbf{x} est choisi à l'intérieur d'un sous-ensemble convexe $\mathcal{X} \subseteq \mathbb{R}^m$ défini préalablement. Rappelons par la même occasion qu'un problème d'optimisation convexe est un problème où l'objectif à minimiser de même que les contraintes sont convexes.

Dans bien des cas, le vecteur $\boldsymbol{\xi}$ ne s'est pas encore matérialisé au moment de la prise de décision (c.-à-d. qu'il n'est pas encore observable). Par exemple, le prix futur d'un actif risqué (une action, par exemple) est aléatoire dans la mesure où il est naturellement impossible de le prévoir avec certitude et précision. Par conséquent, si l'incertitude par rapport à la réalisation du vecteur $\boldsymbol{\xi}$ est représentée à travers sa distribution F , nous pouvons plutôt choisir de minimiser le coût espéré, soit

$$\min_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_F [c(\mathbf{x}, \boldsymbol{\xi})], \quad (3.1)$$

où $\mathbb{E}_F [\cdot]$ est l'opérateur d'espérance par rapport à la distribution F de $\boldsymbol{\xi}$. Dans ce cas, une valeur « moyenne » de $\boldsymbol{\xi}$ est implicitement inférée à partir de sa distribution.

Un inconvénient majeur du programme stochastique (3.1) est qu'il sous-tend une connaissance préalable complète de la distribution du vecteur aléatoire, ce qui est rarement le cas en pratique. Le plus souvent, la distribution sous-jacente doit être estimée à partir d'informations partielles et parfois peu fiables (un échantillon de données historiques, par exemple). Par le fait même, l'estimation est généralement sujette à plusieurs sources d'erreurs ou

de « bruit » qui peuvent malheureusement biaiser la solution optimale du problème d'optimisation. Une des façons d'aborder ce problème est plutôt de formuler une version robuste du problème (3.1) où l'objectif est minimisé par rapport à la « pire » distribution possible (c.-à-d. celle pour laquelle le coût espéré est le plus élevé) au sein d'un ensemble distributionnel prédéfini. L'idée est alors de choisir le vecteur \mathbf{x} optimal compte tenu du pire vecteur aléatoire $\boldsymbol{\xi}$ à l'intérieur de l'ensemble d'incertitude choisi.

Si \mathcal{D} est défini comme un ensemble de distributions possibles tel que $F \in \mathcal{D}$, alors nous obtenons le problème robuste quant au choix de la distribution ou DRSP (*Distributionally Robust Stochastic Program*) qui suit :

$$\min_{\mathbf{x} \in \mathcal{X}} \left(\max_{F \in \mathcal{D}} \mathbb{E}_F [c(\mathbf{x}, \boldsymbol{\xi})] \right). \quad (3.2)$$

Le type d'ensemble distributionnel le plus commun est celui où l'on impose des valeurs spécifiques à la moyenne et à la matrice de variance-covariance de la distribution des variables aléatoires (voir, par exemple, Scarf [Sca58], Bertsimas et Popescu [BP00] et Zymler et al. [ZKR12]). Plus souvent qu'autrement, on suppose donc que les moments en question sont connus de façon précise de manière à simplifier, voire permettre, la résolution de ce type de problèmes. Or, en pratique, cela peut mener à une fausse impression de sécurité dans la mesure où l'estimation des moments est elle-même sujette à l'erreur. Delage et Ye [DY10] proposent alors de caractériser l'incertitude par rapport à la distribution du vecteur aléatoire $\boldsymbol{\xi}$ par le biais de bornes sur son espérance $\boldsymbol{\mu}_0$ et sur sa matrice de variance-covariance $\boldsymbol{\Sigma}_0 \succ 0^1$, à savoir :

$$(\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0)^\top \boldsymbol{\Sigma}_0^{-1} (\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0) \leq \gamma_1 \quad (3.3a)$$

$$\mathbb{E}[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top] \preceq \gamma_2 \boldsymbol{\Sigma}_0, \quad (3.3b)$$

1. La notation $\boldsymbol{\Sigma}_0 \succ 0$ indique que la matrice $\boldsymbol{\Sigma}_0$ est définie positive. Formellement, $\boldsymbol{\Sigma}_0$ est définie positive si $\mathbf{u}^\top \boldsymbol{\Sigma}_0 \mathbf{u} > 0$ pour tout \mathbf{u} non nul dans \mathbb{R}^n . De façon analogue, l'expression $\boldsymbol{\Sigma}_0 \preceq 0$ signifie que $\boldsymbol{\Sigma}_0$ est semi-définie négative. Autrement dit, $\boldsymbol{\Sigma}_0 \preceq 0$ implique que $\mathbf{u}^\top \boldsymbol{\Sigma}_0 \mathbf{u} \leq 0$ pour tout \mathbf{u} non nul dans \mathbb{R}^n .

où $\gamma_1 \geq 0$ et $\gamma_2 \geq 1$ s'interprètent comme le niveau de confiance vis-à-vis chacun des paramètres $\boldsymbol{\mu}_0$ et $\boldsymbol{\Sigma}_0$, respectivement.

D'une part, (3.3a) suppose que la moyenne du vecteur $\boldsymbol{\xi}$ s'inscrit à l'intérieur d'un ellipsoïde de dimension γ_1 centré autour de $\boldsymbol{\mu}_0$. D'autre part, (3.3b) force la matrice de second moment centrée $\mathbb{E}[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top]$ à reposer à l'intérieur d'un cône positif semi-défini borné par une inégalité matricielle (c'est-à-dire $\gamma_2 \boldsymbol{\Sigma}_0$). Remarquons que l'expression de gauche dans (3.3b) correspond en fait à la matrice de variance-covariance lorsque $\boldsymbol{\mu}_0 = \mathbb{E}[\boldsymbol{\xi}]$. Intuitivement, (3.3a) et (3.3b) décrivent à quel point $\boldsymbol{\xi}$ est proche de $\boldsymbol{\mu}_0$ compte tenu des corrélations dans $\boldsymbol{\Sigma}_0$. Lorsque la matrice de variance-covariance est égale à la matrice identité (c.-à-d. $\boldsymbol{\Sigma}_0 = I$), (3.3a) n'est qu'une borne sur la distance euclidienne entre $\boldsymbol{\mu}_0$ et $\mathbb{E}[\boldsymbol{\xi}]$.

À partir de (3.3a) et (3.3b), nous pouvons donc formuler l'ensemble d'incertitude par rapport à la distribution de $\boldsymbol{\xi}$ qui suit :

$$\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2) = \left\{ F \in \mathcal{M} \left| \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \mathcal{S}) = 1 \\ (\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0)^\top \boldsymbol{\Sigma}_0^{-1} (\mathbb{E}[\boldsymbol{\xi}] - \boldsymbol{\mu}_0) \leq \gamma_1 \\ \mathbb{E}[(\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top] \preceq \gamma_2 \boldsymbol{\Sigma}_0 \end{array} \right. \right\}, \quad (3.4)$$

où \mathcal{M} est l'ensemble des mesures de probabilités sur l'espace mesurable $(\mathbb{R}^n, \mathcal{B})$, \mathcal{B} est la tribu de Borel sur \mathbb{R}^n , et \mathcal{S} est n'importe quel ensemble convexe et compact (c.-à-d. fermé et borné) qui renferme le support de la distribution F . L'ensemble $\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ ci-dessus (désormais noté \mathcal{D} , pour alléger la lecture) peut être vu comme une généralisation d'autres ensembles d'incertitude présentés dans la littérature. Par exemple, l'ensemble $\mathcal{D}(\mathbb{R}^n, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, 0, 1)$, qui sous-tend une connaissance exacte des deux premiers moments de F , rejoint en partie celui étudié par Scarf [Sca58] et Pospescu [Pop07], entre autres. Notons qu'en théorie, la distribution de $\boldsymbol{\xi}$ peut prendre une quelconque forme (fortement asymétrique, par exemple), pour autant qu'elle soit caractérisée par l'ensemble \mathcal{D} .

Pour n'importe quel $\mathbf{x} \in \mathcal{X}$ donné, puisque F est une mesure de probabilité sur $(\mathbb{R}^n, \mathcal{B})$, le problème d'optimisation $\max_{F \in \mathcal{D}} \mathbb{E}_F [c(\mathbf{x}, \boldsymbol{\xi})]$ pour lequel l'ensemble d'incertitude \mathcal{D} est celui défini à l'équation (3.4), peut être décrit comme un problème conique linéaire semi-infini, à savoir :

$$\max_F. \int_{\mathcal{S}} c(\mathbf{x}, \boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi} \quad (3.5a)$$

$$\text{sujet à } \int_{\mathcal{S}} p(\boldsymbol{\xi}) d\boldsymbol{\xi} = 1 \quad (3.5b)$$

$$\int_{\mathcal{S}} (\boldsymbol{\xi} - \boldsymbol{\mu}_0)(\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top p(\boldsymbol{\xi}) d\boldsymbol{\xi} \preceq \gamma_2 \boldsymbol{\Sigma}_0 \quad (3.5c)$$

$$\int_{\mathcal{S}} \begin{bmatrix} \boldsymbol{\Sigma}_0 & (\boldsymbol{\xi} - \boldsymbol{\mu}_0) \\ (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top & \gamma_1 \end{bmatrix} p(\boldsymbol{\xi}) d\boldsymbol{\xi} \succeq 0 \quad (3.5d)$$

$$F \in \mathcal{M}, \quad (3.5e)$$

où $p(\boldsymbol{\xi})$ est la mesure de probabilité conjointe du vecteur $\boldsymbol{\xi}$. L'équation (3.5b) fait en sorte que F soit bel et bien une mesure de probabilité (c.-à-d. que l'aire sous la fonction de densité soit égale à 1), tandis que (3.5c) et (3.5d) ne sont que les représentations intégrales de (3.3a) et (3.3b), respectivement. La formulation de la contrainte (3.5d) provient directement de l'application du complément de Schur sur l'inégalité (3.3a) (voir Boyd et Vandenberghe [BV04]). La solution optimale du problème (3.5) peut alors être trouvée à l'aide du programme semi-défini présenté au lemme 3.1.

Lemme 3.1. Pour un vecteur $\mathbf{x} \in \mathbb{R}^m$ donné, si $\gamma_1 \geq 0$, $\gamma_2 \geq 1$, $\boldsymbol{\Sigma}_0 \succ 0$ et $c(\mathbf{x}, \boldsymbol{\xi})$ est F -intégrable pour tout $F \in \mathcal{D}$, alors la valeur optimale du problème (3.5) est finie et égale à celle du problème dual

$$\min_{r, \mathbf{Q}, \mathbf{P}, \mathbf{p}, s} (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \mathbf{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \mathbf{P}) - 2\boldsymbol{\mu}_0^\top \mathbf{p} + \gamma_1 s \quad (3.6a)$$

$$\text{sujet à } \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - c(\mathbf{x}, \boldsymbol{\xi}) \geq 0, \quad \forall \boldsymbol{\xi} \in \mathcal{S} \quad (3.6b)$$

$$\mathbf{Q} \succeq 0, \quad (3.6c)$$

$$\begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^\top & s \end{bmatrix} \succeq 0, \quad (3.6d)$$

où $r \in \mathbb{R}$ et $\mathbf{Q} \in \mathbb{R}^{n \times n}$ sont les variables duales associées à (3.5b) et (3.5c), respectivement, et $\mathbf{P} \in \mathbb{R}^{n \times n}$, $\mathbf{p} \in \mathbb{R}^n$ et $s \in \mathbb{R}$ forment la matrice duale associée à (3.5d). La notation $(\mathbf{A} \bullet \mathbf{B}) = \text{tr}(\mathbf{A}\mathbf{B})$ correspond au produit de Frobenius entre \mathbf{A} et \mathbf{B} , à savoir $\sum_{i,j} A_{i,j} B_{i,j}$.

Le lemme 3.1 est en fait le résultat de l'application de la dualité lagrangienne au problème (3.5) (voir Shapiro [Sha00] et Calafiore et al. [CET09] pour le cas particulier du problème des moments). Si $\gamma_1 \geq 0$, $\gamma_2 \geq 0$ et $\Sigma_0 \succ 0$, il est possible de démontrer par le théorème de la dualité forte qu'il n'y a aucun saut de dualité entre les problèmes (3.5) et (3.6). Autrement dit, les problèmes (3.5) et (3.6) ont tous deux la même valeur optimale.

Par souci de complétude, nous portons à l'attention du lecteur les hypothèses 3.1 et 3.2 ci-dessous sur $c(\mathbf{x}, \boldsymbol{\xi})$, selon lesquelles le problème (3.6) peut être résolu en temps polynomial².

Hypothèse 3.1. La fonction de perte $c(\mathbf{x}, \boldsymbol{\xi})$ s'écrit sous la forme $c(\mathbf{x}, \boldsymbol{\xi}) = \max_{k \in \{1, \dots, K\}} c_k(\mathbf{x}, \boldsymbol{\xi})$, où pour chaque k , $c_k(\mathbf{x}, \boldsymbol{\xi})$ est concave par rapport à $\boldsymbol{\xi}$. De plus, pour une paire $(\mathbf{x}, \boldsymbol{\xi})$ quelconque, il est possible d'évaluer la fonction $c_k(\mathbf{x}, \boldsymbol{\xi})$ et de trouver un sur-gradient de $c_k(\mathbf{x}, \boldsymbol{\xi})$ par rapport à $\boldsymbol{\xi}$ en temps polynomial. Enfin, pour tout $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{q} \in \mathbb{R}^n$ et $\mathbf{Q} \in \mathbb{R}^{n \times n}$, où $\mathbf{Q} \succeq 0$, l'ensemble suivant est fermé : $\{y \in \mathbb{R} : \exists \boldsymbol{\xi} \in \mathcal{S}, y \leq c(\mathbf{x}, \boldsymbol{\xi}) - \mathbf{q}^\top \boldsymbol{\xi} - \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi}\}$.

Hypothèse 3.2. La fonction $c(\mathbf{x}, \boldsymbol{\xi})$ est convexe par rapport à \mathbf{x} . De plus, il est possible de trouver un sous-gradient de $c(\mathbf{x}, \boldsymbol{\xi})$ en \mathbf{x} en temps polynomial.

Delage et Ye [DY10] montrent bel et bien que sous ces deux hypothèses, le problème (3.6) peut être résolu pour n'importe quel degré de précision ε en $\log(1/\varepsilon)$, et proportionnellement au nombre de variables à optimiser.

². Intuitivement, un problème peut être résolu en temps polynomial si son temps d'exécution ou de résolution est proportionnel à un polynôme en fonction du nombre de variables d'entrée.

3.2 Valeur-à-risque conditionnelle (Rockafellar et Uryasev)

La valeur-à-risque de niveau β , rappelons-le, est le plus petit seuil α pour lequel la probabilité que la perte (d'une position ou d'un portefeuille, par exemple) soit plus grande ou égale à α est plus petite ou égale à $\beta \in (0, 1)$. Autrement dit, si $\beta = 0.95$, par exemple, la probabilité que la perte soit plus grande ou égale à la $\text{VaR}_{0.95}$ est d'au plus $1 - \beta$ (c.-à-d. 5%).

Plus formellement, soit $c(\mathbf{x}, \boldsymbol{\xi}) : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$, une fonction de perte caractérisée par le vecteur de décision $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^m$ et le vecteur aléatoire $\boldsymbol{\xi} \in \mathbb{R}^n$. La probabilité que la fonction de perte $c(\mathbf{x}, \boldsymbol{\xi})$ ne dépasse jamais le seuil α est donnée par

$$\psi [c(\mathbf{x}, \boldsymbol{\xi}), \alpha] = \mathbb{P} \{c(\mathbf{x}, \boldsymbol{\xi}) \leq \alpha\} \quad (3.7)$$

$$:= \int_{c(\mathbf{x}, \boldsymbol{\xi}) \leq \alpha} p(\boldsymbol{\xi}) d\boldsymbol{\xi}. \quad (3.8)$$

Définition 3.1. Pour un certain niveau de confiance β , la valeur-à-risque $\text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]$ est

$$\text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] := \inf \{\alpha \in \mathbb{R} : \psi [c(\mathbf{x}, \boldsymbol{\xi}), \alpha] \geq \beta\}. \quad (3.9)$$

Il ne s'agit ici que de la définition du quantile d'ordre β de la distribution de perte.

Hypothèse 3.3. La fonction de répartition de la perte, $\psi [c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$, est strictement croissante et continue en tout point.

L'hypothèse 3.3 a pour but de transformer l'inégalité de l'équation (3.9) en égalité (c.-à-d. $\psi [c(\mathbf{x}, \boldsymbol{\xi}), \alpha] = \beta$), et du même coup, de grandement simplifier la résolution du problème. Pour ce qui suit, nous nous restreignons donc aux distributions continues dont le domaine est \mathbb{R} .

Notons par ailleurs que le cas général (c.-à-d. pour des fonctions de réparti-
tions discontinues ou non strictement croissantes) est exploré en profondeur
par Rockafellar et Uryasev [RU02]. Ceux-ci confirment que la représenta-
tion présentée à l'équation (3.15) persiste lorsque l'hypothèse 3.3 n'est pas
satisfaite.

Définition 3.2. La valeur-à-risque conditionnelle $\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]$, de ni-
veau de confiance β , est

$$\begin{aligned} \text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] &= \mathbb{E} [c(\mathbf{x}, \boldsymbol{\xi}) \mid c(\mathbf{x}, \boldsymbol{\xi}) \geq \text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]] \\ &:= \frac{1}{1 - \beta} \int_{c(\mathbf{x}, \boldsymbol{\xi}) \geq \text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]} c(\mathbf{x}, \boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi}. \end{aligned} \quad (3.10)$$

Autrement dit, la valeur-à-risque conditionnelle est l'espérance condition-
nelle des pertes au-delà de la VaR. Par construction, pour un même seuil β ,
la CVaR est donc toujours plus conservatrice que la VaR. En effet,

$$\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] = \mathbb{E} [c(\mathbf{x}, \boldsymbol{\xi}) \mid c(\mathbf{x}, \boldsymbol{\xi}) \geq \text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]] \quad (3.11)$$

$$= \text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] \quad (3.12)$$

$$\begin{aligned} &+ \underbrace{\mathbb{E} [c(\mathbf{x}, \boldsymbol{\xi}) - \text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] \mid c(\mathbf{x}, \boldsymbol{\xi}) \geq \text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]]}_{\geq 0} \\ &\geq \text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]. \end{aligned} \quad (3.13)$$

Le concept de VaR et de CVaR est illustré à la figure 3.1. La portion colorée
(verte) correspond à l'ensemble des pertes supérieures à la VaR.

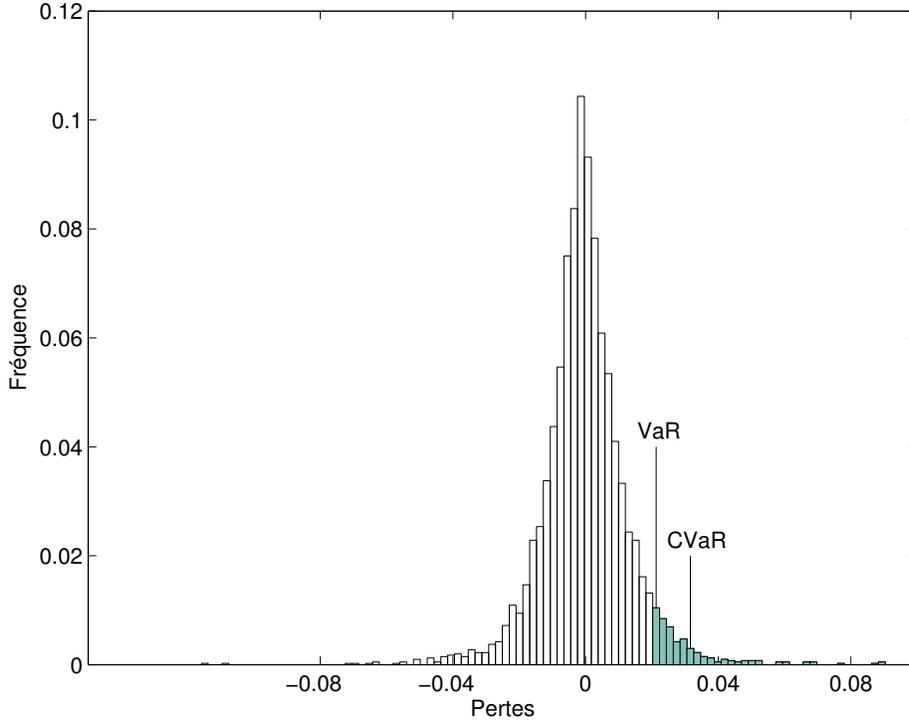


FIGURE 3.1 – VaR et CVaR de la distribution des pertes quotidiennes du S&P 500 entre janvier 1997 et janvier 2013.

Puisqu'elle dépend par construction de la fonction $\text{VaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]$ qui elle-même dépend de \mathbf{x} , l'optimisation de la CVaR peut parfois être difficile d'approche. À défaut d'avoir recours à une représentation analytique de la VaR, Rockafellar et Uryasev [RU00] formulent la fonction auxiliaire suivante :

$$\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] := \alpha + \frac{1}{1-\beta} \mathbb{E}[c(\mathbf{x}, \boldsymbol{\xi}) - \alpha]^+, \quad (3.14)$$

où $[x]^+ := \max\{x, 0\}$, puis démontrent ensuite l'équivalence qui suit :

$$\begin{aligned} \text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] &= \inf_{\alpha \in \mathbb{R}} \mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] \\ &\equiv \inf_{\alpha \in \mathbb{R}} \alpha + \frac{1}{1-\beta} \mathbb{E}[c(\mathbf{x}, \boldsymbol{\xi}) - \alpha]^+. \end{aligned} \quad (3.15)$$

La fonction $\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$ a l'avantage d'être conjointement convexe par rapport à α et \mathbf{x} , pour autant que la fonction de perte $c(\mathbf{x}, \boldsymbol{\xi})$ soit elle aussi convexe par rapport à \mathbf{x} . D'une part, le maximum de deux fonctions convexes est aussi convexe. D'autre part, puisque l'espérance est un opérateur linéaire positif, l'espérance d'une fonction convexe est également convexe. Par l'hypothèse 3.3, la fonction est de plus dérivable presque partout par rapport à α , et ce, peu importe la distribution de $\boldsymbol{\xi}$. L'équivalence (3.15) nous vient du fait que l'infimum de la fonction $\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$ par rapport à α est tel que $\psi[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] = \beta \Leftrightarrow \alpha = \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]$. En effet,

$$\frac{\partial}{\partial \alpha} (\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]) = \frac{\partial}{\partial \alpha} \left(\alpha + \frac{1}{1-\beta} \mathbb{E}[c(\mathbf{x}, \boldsymbol{\xi}) - \alpha]^+ \right) \quad (3.16)$$

$$= 1 - \frac{1}{1-\beta} \mathbb{P}\{c(\mathbf{x}, \boldsymbol{\xi}) \geq \alpha\} \quad (3.17)$$

$$= 1 - \frac{1}{1-\beta} (1 - \psi[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]). \quad (3.18)$$

Par définition, $\mathbb{P}\{c(\mathbf{x}, \boldsymbol{\xi}) \geq \alpha\} = 1 - \mathbb{P}\{c(\mathbf{x}, \boldsymbol{\xi}) \leq \alpha\} = 1 - \psi[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$. Par conséquent, puisque $\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$ est convexe par rapport à α , le minimum global n'est atteint que si et seulement si $\psi[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] = \beta$.

Puisque $\psi(\mathbf{x}, \alpha)$ est une fonction monotone et strictement croissante par rapport à α (par hypothèse), nous pouvons donc remplacer α par $\text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]$ dans l'équation (3.14). Dans ce cas,

$$\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]] \quad (3.19)$$

$$= \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})] + \frac{1}{1-\beta} \mathbb{E}[c(\mathbf{x}, \boldsymbol{\xi}) - \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]]^+ \quad (3.20)$$

$$= \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})] + \frac{1}{1-\beta} \mathbb{E}[c(\mathbf{x}, \boldsymbol{\xi}) - \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})] \mid c(\mathbf{x}, \boldsymbol{\xi}) \geq \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]] \cdot \underbrace{\mathbb{P}\{c(\mathbf{x}, \boldsymbol{\xi}) \geq \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]\}}_{1-\beta} \quad (3.21)$$

$$= \mathbb{E}[c(\mathbf{x}, \boldsymbol{\xi}) \mid c(\mathbf{x}, \boldsymbol{\xi}) \geq \text{VaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]] \quad (3.22)$$

$$= \text{CVaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})]. \quad (3.23)$$

Un avantage de la fonction auxiliaire $\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$ est que nous n'avons pas à choisir une valeur pour α au préalable, ce qui peut s'avérer difficile en pratique. En fait, celle-ci est dérivée naturellement lors du processus d'optimisation en fonction du niveau de confiance choisi.

Rockafellar et Uryasev [RU00] montrent enfin que minimiser $\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]$ par rapport à $\mathbf{x} \in \mathcal{X}$ est équivalent à minimiser $\mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$ par rapport à $(\mathbf{x}, \alpha) \in \mathcal{X} \times \mathbb{R}$. Plus formellement,

$$\min_{\mathbf{x} \in \mathcal{X}} \text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] = \min_{(\mathbf{x}, \alpha) \in \mathcal{X} \times \mathbb{R}} \mathcal{A}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] \quad (3.24)$$

$$= \min_{(\mathbf{x}, \alpha) \in \mathcal{X} \times \mathbb{R}} \alpha + \frac{1}{1 - \beta} \mathbb{E}[c(\mathbf{x}, \boldsymbol{\xi}) - \alpha]^+. \quad (3.25)$$

Puisque l'ensemble \mathcal{X} est convexe par définition, (3.24) est donc un problème d'optimisation convexe.

3.3 Minimisation robuste de la CVaR

Tel que démontré dans [RU02], le problème d'optimisation robuste de la CVaR, pour lequel $F \in \mathcal{D}$, s'écrit

$$\min_{\mathbf{x} \in \mathcal{X}} \left(\max_{F \in \mathcal{D}} \text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] \right),$$

ou, de façon équivalente,

$$\min_{\mathbf{x} \in \mathcal{X}} \left(\max_{F \in \mathcal{D}} \left(\min_{\alpha \in \mathbb{R}} \alpha + \frac{1}{1 - \beta} \mathbb{E}_F [c(\mathbf{x}, \boldsymbol{\xi}) - \alpha]^+ \right) \right).$$

Puisque la fonction $\alpha + (1 - \beta)^{-1} \mathbb{E}_F [c(\mathbf{x}, \boldsymbol{\xi}) - \alpha]^+$ est convexe par rapport à α et convave (en fait, linéaire) par rapport à $\boldsymbol{\xi}$, et puisque l'ensemble \mathcal{D} est faiblement compact, il nous est permis d'interchanger les opérateurs $\max_{F \in \mathcal{D}}$ et $\min_{\alpha \in \mathbb{R}}$ (voir Sion [Sio58]). Ainsi, nous pouvons formuler le problème équivalent qui suit :

$$\min_{(\mathbf{x}, \alpha) \in \mathcal{X} \times \mathbb{R}} \left(\max_{F \in \mathcal{D}} \mathbb{E}_F [\mathcal{Q}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]] \right), \quad (3.26)$$

où $\mathcal{Q}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] = \max\{(1 - (1 - \beta)^{-1})\alpha + (1 - \beta)^{-1}c(\mathbf{x}, \boldsymbol{\xi}), \alpha\}$. Si $c(\mathbf{x}, \boldsymbol{\xi})$ est convexe par rapport à \mathbf{x} et concave par rapport à $\boldsymbol{\xi}$, alors $\mathcal{Q}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha]$ répond aux hypothèses 3.1 et 3.2 et la solution du problème (3.26) peut être calculée en temps polynomial. Notons qu'à moins que la distribution la plus pessimiste de l'ensemble \mathcal{D} soit en fait la vraie distribution de $\boldsymbol{\xi}$, la vraie CVaR va toujours être plus faible que la CVaR robuste.

3.4 Optimisation d'un portefeuille d'options

Lorsqu'un événement rare et inattendu surgit (une crise financière, par exemple), le rendement du portefeuille robuste peut d'un coup se matérialiser à l'extérieur de l'ensemble d'incertitude, et celui-ci perd alors sa protection. Une façon simple mais néanmoins efficace de contourner le problème est d'élargir l'ensemble d'incertitude de manière à inclure les événements les plus rares. Malheureusement, cela peut mener à des portefeuilles robustes trop conservateurs qui performant plus ou moins bien en situation plus « normale ». Ainsi, nous proposons plutôt d'ajouter des options d'achat et de vente au sein du portefeuille. De façon générale, les produits dérivés tels les options ont deux rôles en gestion de portefeuille. Le premier est de spéculer sur la valeur future de l'actif sous-jacent, tandis que le second est de prémunir le portefeuille contre un mouvement non anticipé du prix du sous-jacent. Une option de vente, par exemple, protège le détenteur contre une baisse inattendue du prix de l'actif sous-jacent. L'utilisation d'options ajoute donc un certain niveau de robustesse puisqu'elle fournit une assurance additionnelle dans le cas où le rendement du sous-jacent se manifesterait à l'extérieur de l'ensemble d'incertitude.

Par définition, une option est un contrat par lequel le souscripteur a le droit, mais non l'obligation, d'acheter (option d'achat) ou de vendre (option de vente) un actif sous-jacent (une action, par exemple) à une date future (la date d'échéance ou de maturité) et à un prix convenu au préalable (le prix d'exercice). Intuitivement, le détenteur d'une option n'a intérêt à faire valoir son droit que si l'exercice lui est favorable, c'est-à-dire si le prix auquel il a le droit d'acheter (de vendre) le sous-jacent à l'échéance est inférieur (supérieur) au coût d'achat sur le marché. Une option est dite « européenne » si l'exercice n'est possible qu'à maturité. Notons que le cadre d'optimisation étudié ici est statique, et se porte donc mal à tout autre type d'options où l'exercice est permis entre deux dates de rebalancement (les options américaines ou bermudiennes, par exemple).

Nous présumons désormais que le marché est composé d'au plus m actifs, dont $n \leq m$ actifs sous-jacents et $m - n$ options d'achat et de vente européennes. L'élément x_i du vecteur de décision $\mathbf{x} \in \mathbb{R}^m$ correspond ici au poids en pourcentage investi aujourd'hui (c.-à-d. au temps $t = 0$) dans le i^{e} actif, tandis que l'élément ξ_i du vecteur aléatoire $\boldsymbol{\xi} \in \mathbb{R}^n$ correspond au rendement de ce même actif à la fin de l'horizon de placement (c.-à-d. $t \leq T$). Le vecteur de décision est choisi à l'intérieur du sous-ensemble convexe $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{1}^\top \mathbf{x} = 1\}$, où $\mathbf{1}$ est un vecteur de uns (c.-à-d. $\mathbf{1} = (1, \dots, 1)^\top$). La valeur d'une option à l'échéance est fonction du prix actuel (c.-à-d. à $t = 0$) de l'actif sous-jacent, s_i , et de son prix d'exercice, $k_{i,j}$ ³. À l'instar de Zymler et al. [ZKR12], nous exprimons le rendement d'une option d'achat ou de vente comme une fonction convexe linéaire par morceaux⁴ du rendement de l'actif sous-jacent. Le rendement à l'échéance d'une option d'achat de prix $c_{i,j}$ correspond donc au paiement qui lui est associé, divisé par son coût d'achat, moins 1.

3. La distinction entre $i \in (1, \dots, n)$ et $j \in (n + 1, \dots, m)$ permet à plusieurs options d'achat ou de vente d'avoir le même sous-jacent. L'option $c_{1,1}$, par exemple, est l'option d'achat écrite sur le sous-jacent de prix s_1 et de prix d'exercice $k_{1,1}$. De même, l'option $p_{1,2}$ est une option de vente écrite sur le même sous-jacent mais avec un prix d'exercice $k_{1,2}$.

4. Une fonction est dite convexe par morceaux si elle peut être découpée en un nombre fini de sous-intervalles convexes.

Autrement dit,

$$\begin{aligned} & \frac{1}{c_{i,j}} (\max \{0, s_i(1 + \xi_i) - k_{i,j}\}) - 1 \\ &= \max \{-1, b_{i,j} + a_{i,j}\xi_i - 1\}, \quad \text{où } b_{i,j} = \frac{s_i - k_{i,j}}{c_{i,j}} \quad \text{et } a_{i,j} = \frac{s_i}{c_{i,j}}. \end{aligned} \quad (3.27)$$

De façon analogue, le rendement d'une option de vente de prix $p_{i,j}$ s'écrit :

$$\begin{aligned} & \frac{1}{p_{i,j}} (\max \{0, k_{i,j} - s_i(1 + \xi_i)\}) - 1 \\ &= \max \{-1, b_{i,j} + a_{i,j}\xi_i - 1\}, \quad \text{où } b_{i,j} = \frac{k_{i,j} - s_i}{p_{i,j}} \quad \text{et } a_{i,j} = -\frac{s_i}{p_{i,j}}. \end{aligned} \quad (3.28)$$

Pour chaque i (c.-à-d. pour chaque sous-jacent), les fonctions de rendements des options d'achat et de vente correspondent au maximum entre un scalaire et une fonction linéaire par rapport à ξ_i , et sont par conséquent convexes par morceaux.

À partir de (3.27) et (3.28), nous pouvons désormais caractériser la fonction de perte $c(\mathbf{x}, \boldsymbol{\xi})$ comme la somme des pertes attribuables aux actions et aux options, à savoir :

$$c(\mathbf{x}, \boldsymbol{\xi}) = -(\mathbf{x}^e)^\top \boldsymbol{\xi} - (\mathbf{x}^d)^\top (\max \{-\mathbf{1}, \mathbf{b} + \mathbf{A}\boldsymbol{\xi} - \mathbf{1}\}), \quad (3.29)$$

où $\mathbf{b} \in \mathbb{R}^{m-n}$, $\mathbf{A} \in \mathbb{R}^{(m-n) \times n}$, $\mathbf{1} \in \mathbb{R}^{m-n}$, et $\mathbf{x} = (\mathbf{x}^e, \mathbf{x}^d)$ lorsque $\mathbf{x}^d \geq 0$. Les vecteurs de décision $\mathbf{x}^e \in \mathbb{R}^n$ et $\mathbf{x}^d \in \mathbb{R}^{m-n}$ correspondent aux poids en pourcentage alloués aux actifs sous-jacents et aux options, respectivement. La matrice \mathbf{A} et le vecteur \mathbf{b} sont bâtis respectivement à partir des éléments $a_{i,j}$ et $b_{i,j}$ des équations (3.27) et (3.28). Notons que l'opérateur $\max(\cdot)$ désigne ici le maximum élément par élément.

Pour tout $\boldsymbol{\xi} \in \mathcal{S}$ et \boldsymbol{x}^d non-négatif, $(\boldsymbol{x}^d)^\top (\max\{-\mathbf{1}, \mathbf{b} + \mathbf{A}\boldsymbol{\xi} - \mathbf{1}\})$ est équivalent à

$$\min_{\boldsymbol{z} \in \mathbb{R}^{m-n}} \left\{ \boldsymbol{z}^\top \boldsymbol{x}^d : -\mathbf{1} \leq \boldsymbol{z}, \mathbf{b} + \mathbf{A}\boldsymbol{\xi} - \mathbf{1} \leq \boldsymbol{z} \right\} \quad (3.30)$$

$$= \max_{\boldsymbol{y} \in \mathbb{R}^{m-n}} \left\{ \boldsymbol{y}^\top (\mathbf{b} + \mathbf{A}\boldsymbol{\xi}) - \mathbf{1}^\top \boldsymbol{x}^d : \boldsymbol{x}^d \geq \boldsymbol{y} \geq \mathbf{0} \right\}, \quad (3.31)$$

où la seconde équation est en fait la solution du problème dual associé au problème (3.30) (voir l'annexe A.1 pour plus de détails). Remarquons que la contrainte $\boldsymbol{x}^d \geq \boldsymbol{y} \geq \mathbf{0}$ du problème (3.31) proscriit la vente d'options à découvert. Tel que noté à l'Annexe A.1, cette contrainte est indispensable pour obtenir la forme de l'équation (3.31) et, par le fait même, pour rendre le problème d'optimisation beaucoup plus malléable. Elle n'est toutefois pas déraisonnable dans un contexte de minimisation du risque puisque la vente d'une option est normalement considérée beaucoup plus risquée que l'achat. La fonction de perte peut alors être reformulée de la façon suivante :

$$c(\boldsymbol{x}, \boldsymbol{\xi}) = -\boldsymbol{\xi}^\top \boldsymbol{x}^e - \max_{\boldsymbol{x}^d \geq \boldsymbol{y} \geq \mathbf{0}} \left\{ \boldsymbol{y}^\top (\mathbf{b} + \mathbf{A}\boldsymbol{\xi}) - \mathbf{1}^\top \boldsymbol{x}^d \right\} \quad (3.32)$$

$$= \min_{\boldsymbol{x}^d \geq \boldsymbol{y} \geq \mathbf{0}} \left\{ -\boldsymbol{\xi}^\top \boldsymbol{x}^e - \boldsymbol{y}^\top (\mathbf{b} + \mathbf{A}\boldsymbol{\xi}) + \mathbf{1}^\top \boldsymbol{x}^d \right\}. \quad (3.33)$$

Mentionnons que la fonction $c(\boldsymbol{x}, \boldsymbol{\xi})$ est toujours convexe par rapport à \boldsymbol{x} et concave par rapport à $\boldsymbol{\xi}$. Par conséquent, la fonction $\mathcal{Q}[c(\boldsymbol{x}, \boldsymbol{\xi}), \alpha]$ du problème (3.26) satisfait toujours les hypothèses 3.1 et 3.2 et le problème peut être résolu en temps polynomial. En remplaçant la fonction de perte $c(\boldsymbol{x}, \boldsymbol{\xi})$ du problème (3.6) par (3.33), nous sommes en mesure de formuler la proposition 3.1.

Proposition 3.1. La valeur optimale du problème d'optimisation robuste de la valeur-à-risque conditionnelle (3.26) d'un portefeuille d'options, pour lequel \mathcal{D} est l'ensemble $\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ et $\mathcal{S} = \mathbb{R}^n$, peut être trouvée en résolvant le problème d'optimisation semi-défini qui suit :

$$\min_{x, \mathbf{Q}, \mathbf{P}, \mathbf{p}, y, r, s, \alpha} (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \mathbf{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \mathbf{P}) - 2\boldsymbol{\mu}_0^\top \mathbf{p} + \gamma_1 s \quad (3.34a)$$

$$\text{sujet à} \quad \begin{bmatrix} \mathbf{Q} & \mathbf{q} \\ \mathbf{q}^\top & r - \alpha \end{bmatrix} \succeq 0 \quad (3.34b)$$

$$\begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{1}{2} \left(\frac{\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}}{1-\beta} \right) \\ \mathbf{q}^\top + \frac{1}{2} \left(\frac{\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}}{1-\beta} \right)^\top & r - \alpha + \left(\frac{\alpha + \mathbf{y}^\top \mathbf{b} - \mathbf{1}^\top \mathbf{x}^d}{1-\beta} \right) \end{bmatrix} \succeq 0 \quad (3.34c)$$

$$\mathbf{Q} \succeq 0, \quad (3.34d)$$

$$\begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^\top & s \end{bmatrix} \succeq 0 \quad (3.34e)$$

$$\mathbf{q} = -(\mathbf{p} + \mathbf{Q}\boldsymbol{\mu}_0) \quad (3.34f)$$

$$\mathbf{x}^d \geq \mathbf{0}, \quad \mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}, \quad \mathbf{x}^\top \mathbf{1} = 1. \quad (3.34g)$$

La démonstration est présentée à l'Annexe A.2.

Mentionnons que l'ensemble $\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^n : \mathbf{l} \leq \boldsymbol{\xi} \leq \mathbf{u}\}$, où $\mathbf{l} \in \mathbb{R}^n$ et $\mathbf{u} \in \mathbb{R}^n$ sont des bornes inférieures et supérieures, respectivement, est un choix plus naturel et moins large pour le support de $\boldsymbol{\xi}$. D'une part, le prix d'un actif transigé en bourse ne peut être inférieur (ou même égal) à zéro. Le rendement d'un actif boursier est, par conséquent, toujours borné inférieurement par -100% . D'autre part, tout dépendant de l'horizon de placement, une borne supérieure « raisonnable » peut être inférée sans trop de difficulté. La proposition 3.2 présente le problème d'optimisation robuste de la CVaR lorsque le support de la distribution des rendements des actifs est borné par \mathbf{l} et \mathbf{u} .

Proposition 3.2. La valeur optimale du problème d'optimisation robuste de la valeur-à-risque conditionnelle (3.26) d'un portefeuille d'options, pour lequel \mathcal{D} est l'ensemble $\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ et $\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^n : \boldsymbol{l} \leq \boldsymbol{\xi} \leq \boldsymbol{u}\}$, peut être trouvée en résolvant le problème d'optimisation semi-défini qui suit :

$$\begin{aligned}
& \min_{\substack{x, \boldsymbol{Q}, \boldsymbol{P}, \boldsymbol{p}, \boldsymbol{y}, r, s, \alpha \\ \lambda_1, \boldsymbol{v}_1, \lambda_2, \boldsymbol{v}_2}} & (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \boldsymbol{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \boldsymbol{P}) - 2\boldsymbol{\mu}_0^\top \boldsymbol{p} + \gamma_1 s \\
\text{sujet à} & \quad \boldsymbol{\Xi}_1 = \begin{bmatrix} \mathbf{0} & \left(\frac{\boldsymbol{v}_1 - \lambda_1}{2}\right) \\ \left(\frac{\boldsymbol{v}_1 - \lambda_1}{2}\right)^\top & \lambda_1^\top \boldsymbol{l} - \boldsymbol{v}_1^\top \boldsymbol{u} \end{bmatrix} \\
& \quad \boldsymbol{\Xi}_2 = \begin{bmatrix} \mathbf{0} & \left(\frac{\boldsymbol{v}_2 - \lambda_2}{2}\right) \\ \left(\frac{\boldsymbol{v}_2 - \lambda_2}{2}\right)^\top & \lambda_2^\top \boldsymbol{l} - \boldsymbol{v}_2^\top \boldsymbol{u} \end{bmatrix} \\
& \quad \boldsymbol{\Xi}_1 + \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q} \\ \boldsymbol{q}^\top & r - \alpha \end{bmatrix} \succeq 0 \\
& \quad \boldsymbol{\Xi}_2 + \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q} + \frac{1}{2} \left(\frac{\boldsymbol{x}^e + \boldsymbol{A}^\top \boldsymbol{y}}{1 - \beta}\right) \\ \boldsymbol{q}^\top + \frac{1}{2} \left(\frac{\boldsymbol{x}^e + \boldsymbol{A}^\top \boldsymbol{y}}{1 - \beta}\right)^\top & r - \alpha + \left(\frac{\alpha + \boldsymbol{y}^\top \boldsymbol{b} - \mathbf{1}^\top \boldsymbol{x}^d}{1 - \beta}\right) \end{bmatrix} \succeq 0 \\
& \quad \boldsymbol{Q} \succeq 0, \\
& \quad \begin{bmatrix} \boldsymbol{P} & \boldsymbol{p} \\ \boldsymbol{p}^\top & s \end{bmatrix} \succeq 0 \\
& \quad \boldsymbol{q} = -(\boldsymbol{p} + \boldsymbol{Q} \boldsymbol{\mu}_0) \\
& \quad \lambda_1 \geq \mathbf{0}, \quad \boldsymbol{v}_1 \geq \mathbf{0}, \quad \lambda_2 \geq \mathbf{0}, \quad \boldsymbol{v}_2 \geq \mathbf{0} \\
& \quad \boldsymbol{x}^d \geq \mathbf{0}, \quad \boldsymbol{x}^d \geq \boldsymbol{y} \geq \mathbf{0}, \quad \boldsymbol{x}^\top \mathbf{1} = 1.
\end{aligned}$$

La démonstration est présentée à l'Annexe A.3.

3.5 Optimisation robuste du ratio de Sharpe

Dans sa forme générale, le ratio de Sharpe correspond au ratio entre le rendement excédentaire espéré d'un portefeuille et son risque. Lorsque le risque en question est caractérisé par la CVaR, le ratio de Sharpe $\text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})]$ peut s'exprimer ainsi :

$$\text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})] := \frac{\mathbb{E}_F[-c(\mathbf{x}, \boldsymbol{\xi})] - r_f}{\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]}, \quad (3.35)$$

où $\mathbb{E}_F[-c(\mathbf{x}, \boldsymbol{\xi})]$ est l'espérance sous F du négatif de la fonction de perte (c.-à-d. la fonction de rendement), r_f est le taux d'intérêt sans risque, et $\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]$ est la valeur-à-risque conditionnelle (3.15) sous F . Nous formulons donc le problème d'optimisation robuste du ratio de Sharpe comme suit :

$$\max_{\mathbf{x} \in \mathcal{X}} \min_{F \in \mathcal{D}} \text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})]. \quad (3.36)$$

Le problème ci-dessus peut alors être reformulé ainsi :

$$\max_{(\mathbf{x}, \tau) \in \mathcal{X} \times \mathbb{R}} \tau \quad (3.37a)$$

$$\text{sujet à } \tau \leq \min_{F \in \mathcal{D}} \text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})], \quad (3.37b)$$

ou, de façon équivalente,

$$\max_{\tau \in \mathbb{R}} f(\tau), \quad (3.38)$$

où

$$f(\tau) := \begin{cases} \tau & \text{si } \exists \mathbf{x} \in \mathcal{X} \text{ tel que } \tau \leq \min_{F \in \mathcal{D}} \text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})], \\ -\infty & \text{sinon.} \end{cases}$$

La difficulté est alors de vérifier la faisabilité de l'expression :

$$\exists \mathbf{x} \in \mathcal{X} \text{ tel que } \tau \leq \min_{F \in \mathcal{D}} \text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})]. \quad (3.39)$$

Hypothèse 3.4. Pour tout $\mathbf{x} \in \mathcal{X}$ et $F \in \mathcal{D}$, $\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] > 0$ et la valeur optimale $\tau^* \geq 0$.

Sous l'hypothèse 3.4, nous sommes en mesure de reformuler la contrainte $\tau \leq \min_{F \in \mathcal{D}} \text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})]$ du problème de faisabilité (3.39) de la manière suivante :

$$\begin{aligned} \tau &\leq \text{SR}_F [c(\mathbf{x}, \boldsymbol{\xi})] && \forall F \in \mathcal{D} \\ \tau &\leq \frac{\mathbb{E}_F [-c(\mathbf{x}, \boldsymbol{\xi})] - r_f}{\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]} && \forall F \in \mathcal{D} \\ &\equiv \tau \cdot \text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] - \mathbb{E}_F [-c(\mathbf{x}, \boldsymbol{\xi})] + r_f \leq 0 && \forall F \in \mathcal{D}, \end{aligned}$$

Nous écrivons maintenant le problème de faisabilité (3.39) ainsi :

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{F \in \mathcal{D}} \{ \tau \cdot \text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] - \mathbb{E}_F [-c(\mathbf{x}, \boldsymbol{\xi})] + r_f \} \leq 0. \quad (3.40)$$

En remplaçant $\text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})]$ et $c(\mathbf{x}, \boldsymbol{\xi})$ dans l'équation $\tau \cdot \text{CVaR}_\beta [c(\mathbf{x}, \boldsymbol{\xi})] - \mathbb{E}_F [-c(\mathbf{x}, \boldsymbol{\xi})] + r_f$ par (3.15) et (3.33), nous obtenons :

$$\begin{aligned} &\tau \left(\inf_{\alpha \in \mathbb{R}} \alpha + \frac{1}{1-\beta} \mathbb{E}_F [c(\mathbf{x}, \boldsymbol{\xi}) - \alpha]^+ \right) - \mathbb{E}_F [-c(\mathbf{x}, \boldsymbol{\xi})] + r_f \\ &= \inf_{\alpha \in \mathbb{R}} \mathbb{E}_F \left[\max \left\{ \left(\tau - \frac{\tau}{1-\beta} \right) \alpha + \left(1 + \frac{\tau}{1-\beta} \right) c(\mathbf{x}, \boldsymbol{\xi}) + r_f, \right. \right. \\ &\quad \left. \left. c(\mathbf{x}, \boldsymbol{\xi}) + \alpha \tau + r_f \right\} \right]. \end{aligned}$$

Pour un $\tau \geq 0$ donné, nous pouvons donc reformuler le problème (3.40) comme suit :

$$\min_{(\mathbf{x}, \alpha) \in \mathcal{X} \times \mathbb{R}} \left(\max_{F \in \mathcal{D}} \mathbb{E}_F [\mathcal{G} [c(\mathbf{x}, \boldsymbol{\xi}), \alpha, \tau]] \right) \leq 0, \quad (3.41)$$

où

$$\mathcal{G}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha, \tau] := \max \left\{ \left(\tau - \frac{\tau}{1-\beta} \right) \alpha + \left(1 + \frac{\tau}{1-\beta} \right) c(\mathbf{x}, \boldsymbol{\xi}) + r_f, \right. \\ \left. c(\mathbf{x}, \boldsymbol{\xi}) + \alpha\tau + r_f \right\}$$

La proposition 3.3 nous permet désormais d'exprimer le problème (3.41) comme un programme d'optimisation semi-défini convexe et facile à résoudre.

Proposition 3.3. Pour un $\tau \geq 0$ fixe, si l'ensemble \mathcal{D} est l'ensemble $\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ et $\mathcal{S} = \mathbb{R}^n$, alors la valeur optimale du sous-problème d'optimisation robuste du ratio de Sharpe présenté dans l'inégalité (3.41) peut être trouvée en résolvant le problème dual qui suit :

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{Q}, \mathbf{P}, \mathbf{p}, \mathbf{y}, \mathbf{y}_1, r, s, \alpha} \quad & (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \mathbf{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \mathbf{P}) - 2\boldsymbol{\mu}_0^\top \mathbf{p} + \gamma_1 s \\ \text{sujet à} \quad & \begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{1}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}) \\ \mathbf{q}^\top + \frac{1}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y})^\top & r - \alpha\tau - r_f + \mathbf{y}^\top \mathbf{b} - \mathbf{1}^\top \mathbf{x}^d \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{\kappa}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}_1) \\ \mathbf{q}^\top + \frac{\kappa}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}_1)^\top & r - \left(\tau - \frac{\tau}{1-\beta} \right) \alpha - r_f + \kappa \mathbf{y}_1^\top \mathbf{b} - \kappa \mathbf{1}^\top \mathbf{x}^d \end{bmatrix} \succeq 0 \\ & \mathbf{Q} \succeq 0, \\ & \begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^\top & s \end{bmatrix} \succeq 0 \\ & \mathbf{q} = -(\mathbf{p} + \mathbf{Q}\boldsymbol{\mu}_0), \quad \kappa = \left(1 + \frac{\tau}{1-\beta} \right) \\ & \mathbf{x}^d \geq \mathbf{0}, \quad \mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}, \quad \mathbf{x}^d \geq \mathbf{y}_1 \geq \mathbf{0}, \quad \mathbf{x}^\top \mathbf{1} = 1. \end{aligned}$$

La démonstration est présentée à l'annexe A.4.

Le problème robuste de maximisation du ratio de Sharpe (3.40) peut alors être reformulé comme suit :

$$\min_{(\mathbf{x}, \alpha, \tau) \in \mathcal{X} \times \mathbb{R} \times \mathbb{R}^+} -\tau \quad (3.42a)$$

$$\text{sujet à } \max_{F \in \mathcal{D}} \mathbb{E}_F [\mathcal{G}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha, \tau]] \leq 0. \quad (3.42b)$$

Notons que la fonction $\mathcal{G}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha, \tau]$ n'est pas conjointement convexe par rapport à (\mathbf{x}, τ) du fait que τ soit multiplié par α . Elle le devient toutefois lorsque l'on fixe τ . En effet, la région admissible pour \mathbf{x} n'est convexe que lorsque τ est fixé. De plus, celle-ci rétrécit au fur et à mesure que τ augmente. Intuitivement, il y a donc un τ maximal pour lequel la région admissible est non vide, et un τ minimal pour lequel la région admissible est vide. La valeur optimale de τ peut alors être trouvée à l'aide de la méthode de la bisection (voir Boyd et Vandenberghe [BV04]).

Notons par p^* , la valeur optimale du problème d'optimisation quasi convexe (3.42), et par $\epsilon > 0$, un certain niveau de tolérance. Pour $a \leq p^*$ et $b \geq p^*$, l'algorithme de la bisection est :

```

répéter
  |  $\tau \leftarrow (a + b)/2$ 
  | résoudre le problème de faisabilité (3.41)
  | si (3.41) est réalisable alors
  |   |  $b \leftarrow \tau$ 
  |   | sinon  $a \leftarrow \tau$ ;
  | fin
jusqu'à  $b - a \leq \epsilon$ ;

```

Algorithme 1 : Méthode de la bisection.

La méthode de la bisection consiste ainsi à évaluer le problème de faisabilité (3.41) pour un certain τ au milieu de l'intervalle $[a, b]$. L'intervalle doit, bien sûr, être suffisamment grand pour contenir la valeur optimale du problème (3.42). Si le problème (3.41) est réalisable, l'algorithme en déduit que $\tau \geq p^*$ et la borne supérieure b de l'intervalle est remplacée par τ .

Si (3.41) n'est pas réalisable, alors $\tau \leq p^*$. L'algorithme en question divise alors le nouvel intervalle ainsi formé jusqu'à ce que la taille de l'intervalle soit inférieure ou égale au niveau de précision ϵ désiré. Une particularité de cette méthode est qu'elle converge linéairement et requiert exactement $\lceil \log_2((b-a)/\epsilon) \rceil$ itérations.

Remarque 3.1. Si $\tau \geq 0$ et $\text{CVaR}_\beta[c(\mathbf{x}, \boldsymbol{\xi})] > 0$ (hypothèse 3.4), alors la fonction $\text{SR}_F[c(\mathbf{x}, \boldsymbol{\xi})]$ du problème de faisabilité (3.39) peut s'exprimer comme la somme de deux fonctions convexes, et ce dernier peut alors être résolu à l'aide du problème d'optimisation convexe (3.41). L'hypothèse 3.4 nous permet donc de garantir la convexité du problème en question, en plus de le rendre plus facile à travailler. En pratique, tout investissement comporte une part de risque, aussi faible soit-elle (risque de marché, de crédit, de contre-partie, etc.). Ainsi, on peut penser qu'il existe un certain β pour lequel la CVaR qui lui associée est positive pour tout $\mathbf{x} \in \mathcal{X}$. Remarquons enfin que le ratio de Sharpe (3.35) n'est pas défini lorsque la CVaR est nulle.

Chapitre 4

Expérimentation et résultats

Dans ce chapitre, nous exposons les résultats d'expériences numériques sur le modèle d'optimisation robuste de la CVaR, et les comparons avec ceux obtenus à partir d'un modèle non robuste. Dans un premier temps, nous construisons une version robuste de la frontière efficiente à partir de données synthétiques. Puis, dans un deuxième temps, nous évaluons la performance des portefeuilles robustes avec et sans options sur un jeu de données historiques.

4.1 Expérience 1

Nous nous penchons d'abord sur le cas purement théorique d'un portefeuille composé de 14 actifs, dont deux actions (notées A et B), 3 options d'achat sur A, 3 options d'achat sur B, 3 options de vente sur A, et 3 options de vente sur B. Le prix initial des deux actions est de 100\$, tandis que les prix d'exercice des trois options d'achat et de vente sur A et B sont de 105\$, 100\$ et 95\$. De cette façon, il y a donc au moins une option d'achat et de vente « à la monnaie », « dans la monnaie » et « hors de la monnaie, et ce, tant sur A que sur B. Une option d'achat est dite « à la monnaie » si

$s = k$, « dans la monnaie » si $s > k$ et « hors de la monnaie » si $s < k$. À l'inverse, une option de vente est dite « dans la monnaie » si $k > s$ et « hors de la monnaie » si $k < s$. Par conséquent, l'ensemble d'incertitude est ici représenté par $\mathcal{D}(\mathbb{R}^2, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, 0, 1)$. Nous supposons à présent que le prix de chacune des deux actions suit un mouvement brownien géométrique, dont les coefficients de dérive (« *drift* ») et de volatilité sont respectivement $\mu_1 = 20\%$ et $\sigma_1 = 30\%$ pour A, et $\mu_2 = 10\%$ et $\sigma_2 = 20\%$ pour B. La corrélation instantanée entre A et B est $\rho = 10\%$, le taux sans risque est de 3% , et la date d'échéance de toutes les options est d'un an ($t = 1$). En conséquence, nous pouvons exprimer le prix d'une action $i = 1, 2$ comme suit :

$$s_{i,t} = s_{i,0} \exp \left\{ \left(\mu_i - \frac{\sigma_i^2}{2} \right) t + \sigma_i W_{i,t} \right\},$$

où $s_{i,0} = 100\$$ est le prix initial des deux actions, $W_{1,t} = B_{1,t}$, $W_{2,t} = \rho B_{1,t} + \sqrt{1 - \rho^2} B_{2,t}$, et $(B_{1,t}, B_{2,t})^\top$ est un mouvement brownien standard bivarié. Le rendement à l'échéance $\xi_{i,t}$ de l'action $i = 1, 2$ est défini par $s_{i,t}/s_{i,0} - 1$. Dans la circonstance, l'espérance et la matrice de variance-covariance des rendements sont :

$$\boldsymbol{\mu}_0 = \begin{bmatrix} e^{\mu_1 t} - 1 \\ e^{\mu_2 t} - 1 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} e^{2\mu_1 t} (e^{\sigma_1^2 t} - 1) & e^{(\mu_1 + \mu_2)t} (e^{\sigma_1 \sigma_2 \rho t} - 1) \\ e^{(\mu_1 + \mu_2)t} (e^{\sigma_1 \sigma_2 \rho t} - 1) & e^{2\mu_2 t} (e^{\sigma_2^2 t} - 1) \end{bmatrix}.$$

Les prix d'options sont ensuite déterminés tout simplement par la formule de Black-Scholes¹ (voir Hull [Hul14]).

1. Notons que le cadre Black-Scholes suppose l'existence d'un actif sans risque connu et constant. Bien que le modèle étudié ici ne contient aucun actif de la sorte, nous présumons que le vendeur des options en question est bel et bien en mesure d'investir dans un tel actif.

Nous obtenons donc :

$$\boldsymbol{\mu}_0 = \begin{bmatrix} 0.2214 \\ 0.1052 \end{bmatrix}, \quad \boldsymbol{\Sigma}_0 = \begin{bmatrix} 0.1405 & 0.0081 \\ 0.0081 & 0.0498 \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} 100/11.11 & 0 \\ 100/13.28 & 0 \\ 100/15.78 & 0 \\ -100/13.01 & 0 \\ -100/10.33 & 0 \\ -100/7.97 & 0 \\ 0 & 7.12 \\ 0 & 9.41 \\ 0 & 12.18 \\ 0 & -100/9.02 \\ 0 & -100/6.46 \\ 0 & -100/4.37 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} (100 - 105)/11.11 \\ (100 - 100)/13.28 \\ (100 - 95)/15.78 \\ (105 - 100)/13.01 \\ (100 - 100)/10.33 \\ (95 - 100)/7.97 \\ (100 - 105)/7.12 \\ (100 - 100)/9.41 \\ (100 - 95)/12.18 \\ (105 - 100)/9.02 \\ (100 - 100)/6.46 \\ (95 - 100)/4.37 \end{bmatrix}.$$

Pour tracer la frontière efficiente robuste, nous ajoutons tout bonnement la contrainte de rendement minimal espéré $\mathbb{E}[-c(\mathbf{x}, \boldsymbol{\xi})] \geq \varphi$ au problème d'optimisation (4.1), où $-c(\mathbf{x}, \boldsymbol{\xi})$ est le négatif de la fonction de perte présentée à l'équation (3.29) (c.-à-d. la fonction de rendement), et φ est le rendement minimal espéré. La frontière efficiente est ensuite dessinée pour chaque couple (φ, CVaR) . Notons que pour fins de simplification, nous incluons la contrainte $\mathbf{x}^e \geq 0$ qui proscrit la vente à découvert.

Par la suite, le modèle robuste est comparé à un modèle non robuste linéaire discret fortement inspiré de Rockafellar et Uryasev [RU00]. Plus spécifiquement, le modèle non robuste utilisé est :

$$\begin{aligned}
& \min_{\mathbf{x}, \mathbf{u}, \alpha} \alpha + \frac{1}{k(1-\beta)} \sum_{i=1}^k u_i \\
& \text{sujet à } u_i \geq c(\mathbf{x}, \boldsymbol{\xi}_i) - \alpha \\
& \quad u_i \geq 0 \\
& \quad \mathbf{x} \in \mathcal{X} \\
& \quad i = 1, \dots, k,
\end{aligned} \tag{4.1}$$

où $c(\mathbf{x}_i, \boldsymbol{\xi}_i)$ est la fonction de perte (3.33) et $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_k$ sont des rendements échantillonnés à partir de la distribution connue de $\boldsymbol{\xi}$. Le modèle d'optimisation non robuste en question est présenté plus en détail à l'annexe B. Le modèle non robuste est calibré à l'aide de 1,000 scénarios² tirés d'une loi normale multidimensionnelle de moyenne $\boldsymbol{\mu}_0$ et de variance $\boldsymbol{\Sigma}_0$. Les problèmes d'optimisation robustes et non robustes sont résolus à l'aide du logiciel Matlab (version 7.11.0) et du progiciel CVX (version 2.1).

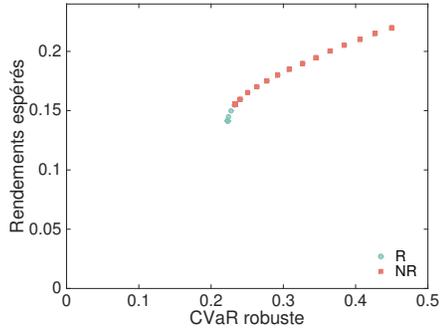
D'abord, la figure 4.1 trace les frontières efficaces robustes et non robustes, avec et sans options, obtenues à partir des portefeuilles optimaux générés par les modèles robustes et non robustes, respectivement. Par exemple, la figure 4.1c illustre les frontières efficaces robustes (c.-à-d. calculées à l'aide du modèle d'optimisation robuste) des portefeuilles optimaux générés par les modèles robustes et non robustes avec options, tandis que la figure (4.1b) présente les frontières efficaces non robustes (c.-à-d. calculées à l'aide du programme d'optimisation robuste) des portefeuilles optimaux produits par les modèles robustes et non robustes sans options. L'objectif ici est de présenter la composition optimale des portefeuilles robustes et non robustes sur le même dénominateur. Puis, la figure 4.2 présente la composition des portefeuilles optimaux robustes et non robustes, avec et sans options, en fonction du rendement espéré.

2. Le nombre de scénarios a été choisi de façon à obtenir un niveau de précision acceptable dans un temps raisonnable. En particulier, le calcul des portefeuilles « Sharpe-optimaux » prend plus de 20 heures avec un processeur Intel Core i5 de 2.3 GHz.

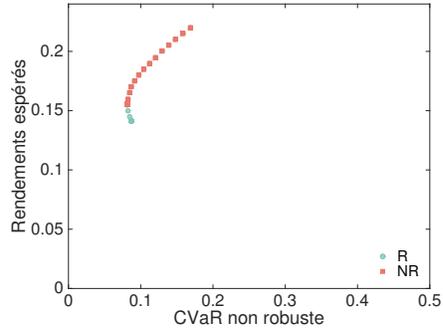
Notons que le choix d'un β de 0.65 pour les figures 4.1 et 4.2 n'est pas anodin. À vrai dire, plus le niveau de risque permis est faible (c.-à-d. plus le β choisi est élevé), plus les courbes sont proches et difficiles à distinguer. Autrement dit, les deux modèles sont difficiles à différencier lorsque β tend vers 1. Le modèle robuste se révèle alors plus utile lorsque le niveau de risque encouru par le portefeuille est plus important.

Nous remarquons en outre que l'ajout d'options permet, dans les deux cas (robustes et non robustes), de diminuer significativement le risque du portefeuille (c.-à-d. sa CVaR) pour un même niveau de rendement espéré. Si la composition des portefeuilles robustes et non robustes sans options est comparable, ce n'est toutefois pas le cas pour les portefeuilles avec options. Le portefeuille robuste alloue en effet un plus grand pourcentage aux options, en particulier lorsque le rendement minimal espéré requis est élevé, ce qui a pour effet d'abaisser la CVaR.

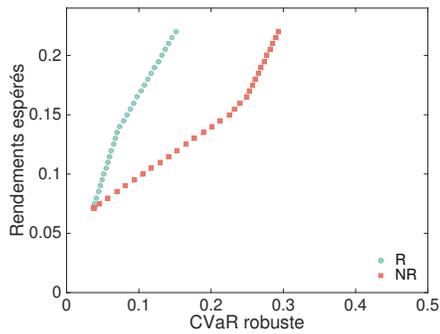
Qui plus est, la distinction entre les frontières robustes et non robustes augmente en fonction du rendement attendu. Autrement dit, le modèle robuste pénalise davantage la CVaR que le modèle non robuste pour toute augmentation du rendement minimal espéré. À cet effet, notons que la CVaR robuste du portefeuille optimal non robuste avec options est beaucoup plus grande que celle de son homologue robuste (figure 4.1c), tandis que la CVaR non robuste est similaire dans les deux cas (figure 4.1d).



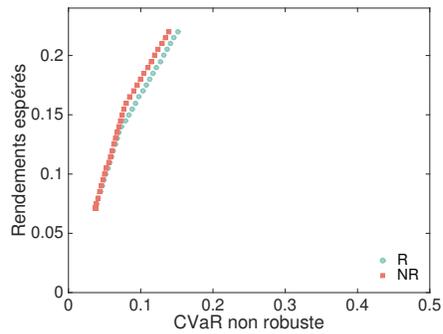
(a) Frontières efficaces robustes sans options.



(b) Frontières efficaces non robustes sans options.

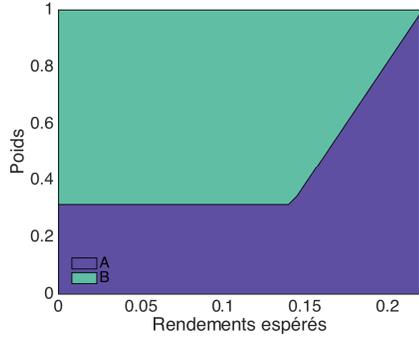


(c) Frontières efficaces robustes avec options.

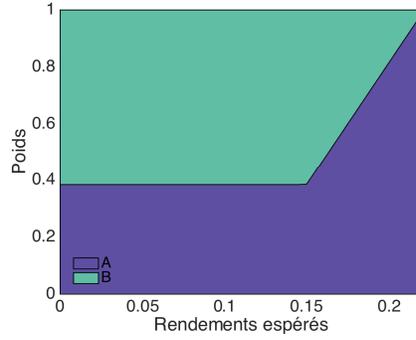


(d) Frontières efficaces non robuste avec options.

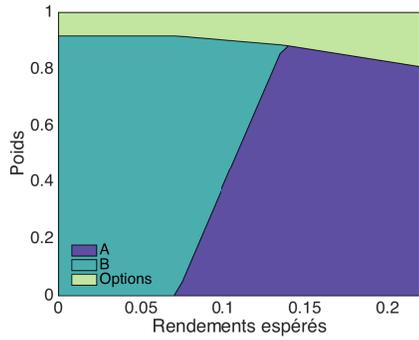
FIGURE 4.1 – Frontières efficaces robustes et non robustes obtenues à partir des portefeuilles optimaux robustes (R) et non robustes (NR), avec et sans options ($\beta = 0.65$).



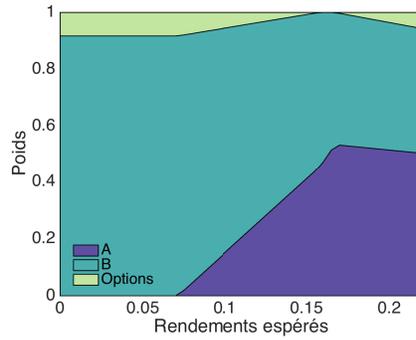
(a) Portefeuille robuste sans options.



(b) Portefeuille non robuste sans options.



(c) Portefeuille robuste avec options.



(d) Portefeuille non robuste avec options.

FIGURE 4.2 – Poids des actifs dans les différents portefeuilles optimaux en fonction du rendement espéré ($\beta = 0.65$).

4.2 Expérience 2

La seconde expérience porte sur l'optimisation des problèmes robustes de minimisation de la CVaR (3.26) et de maximisation du ratio de Sharpe (3.40) d'un portefeuille d'options à l'aide d'un jeu de données historiques mensuelles qui s'étend de janvier 1997 à janvier 2013. Contrairement à la première expérience, nous n'imposons ici aucune contrainte sur le rendement espéré.

4.2.1 Méthodologie

Nous concentrons cette étude sur les six indices américains présentés dans le tableau ci-dessous.

Symbole	Indice	ID
SPX	S&P 500 Index	108105
MID	S&P Midcap 400 Index	101507
SML	S&P Smallcap 600 Index	102442
RUT	Russell 2000 Index	102434
DJX	Dow Jones Industrial Average	102456
NDX	NASDAQ 100 Index	102480

TABEAU 4.1 – Symboles et numéros de référence (Option ID) des différents indices.

Les six indices en question ont au moins deux points en commun. D'une part, ils sont tous très liquides et, d'autre part, ils tiennent lieu de sous-jacents pour une foule d'options d'achat et de vente européennes toutes aussi liquides. Ce sont également les indices avec les historiques de prix d'options les plus longs. La figure 4.3 dépeint les rendements cumulés de chacun des indices sur la période étudiée. Les tableaux 4.2, 4.3 et 4.4, quant à eux, présentent les statistiques descriptives des six indices sur trois périodes. La période de 1997 à 2012 correspond à l'échantillon total, tandis que les

périodes de 1997 à 2004 et de 2004 à 2012 correspondent respectivement à la période de calibration (ou d'estimation) et à la période hors-échantillon.

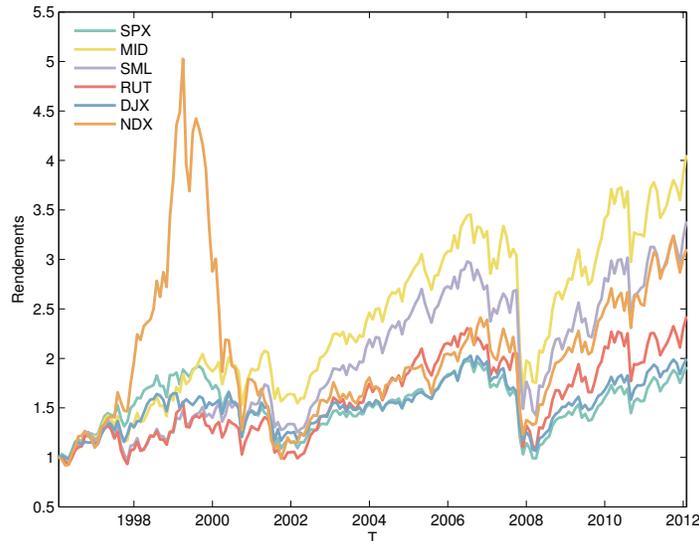


FIGURE 4.3 – Rendements cumulés des six indices entre janvier 1997 et janvier 2013.

Le prix des options et des indices ont été tirés de la base de données Option-Metrics Ivy Database et de la plateforme Bloomberg, respectivement. Pour chaque cote d'options, nous colligeons les informations suivantes : la date de maturité, le prix d'exercice, le prix offert (*bid*), le prix demandé (*ask*), le nombre de positions ouvertes (*open interest*), le volume de négociation quotidien, et la volatilité implicite.

Les prix d'options sont par la suite épurés de façon à obtenir un échantillon de prix d'options représentatif, fiable, et exempt d'erreurs. En premier lieu, nous éliminons toutes les options dont le prix offert est supérieur au prix demandé, de même que celles dont l'un ou l'autre des prix est nul. Nous excluons aussi les options dont la différence entre le prix offert et le prix demandé est inférieure à une unité de fluctuation. La valeur d'une unité de fluctuation est de 0.05\$ pour les options dont le prix est de 3\$ ou moins, et

de 0.10\$ pour les autres. Nous supprimons enfin toutes les options dont le volume ou le nombre de positions ouvertes est nul, de même que celles dont la volatilité implicite est inférieure à 3% ou supérieure à 200%.

Nous postulons pour cette expérience que le support de $\boldsymbol{\xi}$ est caractérisé par l'intervalle $\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^n : -0.15 \leq \boldsymbol{\xi} \leq 0.15\}$, et que le β de la CVaR est de 95%. Les bornes en question ($\pm 15\%$) ont été choisies de façon à englober la majorité des rendements observés durant de la période d'estimation (plus de 95% des rendements de la période tombent à l'intérieur de ces bornes). Mentionnons toutefois que les résultats présentés dans ce chapitre sont sensiblement les mêmes lorsque le support utilisé est $\mathcal{S} = \mathbb{R}^n$. Les paramètres $\boldsymbol{\mu}_0 = M^{-1} \sum_{i=1}^M \boldsymbol{\xi}_i$ et $\boldsymbol{\Sigma}_0 = M^{-1} \sum_{i=1}^M (\boldsymbol{\xi}_i - \boldsymbol{\mu}_0)(\boldsymbol{\xi}_i - \boldsymbol{\mu}_0)^\top$ sont estimés sur une fenêtre mobile de 18 mois ($M = 18$), soit un an et demi. Les paramètres γ_1 et γ_2 , de leur côté, sont trouvés au moyen d'une simple analyse statistique de la quantité de « bruit » détectée dans l'estimation des deux premiers moments au cours de la période de calibration. Plus précisément, pour n'importe quelle période de $2 \times M$ mois entre janvier 1997 et décembre 2004 (la période d'estimation), les paramètres γ_1 et γ_2 sont choisis de façon à ce que l'ensemble \mathcal{D} , centré autour des moments estimés au cours des M premiers mois, contiennent les moments des M mois suivants avec une probabilité $\delta = 99\%$. Nous obtenons ainsi les valeurs 6.20 et 28.40 pour γ_1 et γ_2 , respectivement (voir Delage et Ye [DY10] pour une approche empirique similaire). Notons que Delage et Ye présentent également une formule analytique pour le calcul de γ_1 et γ_2 , sous certaines conditions. Malheureusement, la fenêtre mobile utilisée ici est trop petite pour permettre l'estimation des deux paramètres de la façon proposée. Une fenêtre plus large serait de fait peu appropriée dans notre cas compte tenu de la nature du problème et des données disponibles.

1997-2012	SPX	MID	SML	RUT	DJX	NDX
#	189	189	189	189	189	189
Moyenne	0.49%	0.95%	0.88%	0.72%	0.50%	0.96%
Médiane	1.12%	1.81%	1.82%	0.72%	0.50%	1.45%
Écart-type	5.29%	6.31%	6.66%	6.98%	4.98%	8.33%
Aplatissement	5.91	7.05	6.10	5.65	5.23	4.73
Asymétrie	-0.98	-0.94	-0.88	-0.80	-1.01	-0.56
Minimum	5.29%	6.31%	6.66%	6.98%	4.98%	8.33%
Maximum	17.35%	24.63%	23.17%	24.36%	13.96%	23.82%

TABLEAU 4.2 – Statistiques descriptives des rendements mensuels des six indices de 1997 à 2012.

1997-2004	SPX	MID	SML	RUT	DJX	NDX
#	92	92	92	92	92	92
Moyenne	0.60%	1.12%	1.03%	0.80%	0.62%	1.14%
Médiane	0.55%	1.43%	1.81%	1.38%	1.07%	0.68%
Écart-type	5.03%	5.65%	6.22%	6.66%	5.10%	9.87%
Aplatissement	4.01	3.97	3.84	3.43	4.66	3.79
Asymétrie	-0.61	-0.54	-0.55	-0.44	-0.89	-0.36
Minimum	-16.68%	-17.92%	-19.26%	-19.52%	-19.54%	-29.36%
Maximum	11.03%	15.73%	18.21%	17.79%	10.06%	23.82%

TABLEAU 4.3 – Statistiques descriptives des rendements mensuels des six indices de 1997 à 2004.

2004-2012	SPX	MID	SML	RUT	DJX	NDX
#	97	97	97	97	97	97
Moyenne	0.39%	0.79%	0.73%	0.65%	0.38%	0.80%
Médiane	1.58%	1.90%	1.86%	1.64%	1.53%	1.91%
Écart-type	5.55%	6.91%	7.08%	7.30%	4.89%	6.59%
Aplatissement	6.97	7.91	7.16	7.00	5.85	5.69
Asymétrie	-1.22	-1.10	-1.08	-1.04	-1.14	-1.16
Minimum	-22.32%	-27.86%	-27.61%	-28.16%	-18.61%	-27.30%
Maximum	17.35%	24.63%	23.17%	24.36%	13.96%	15.57%

TABLEAU 4.4 – Statistiques descriptives des rendements mensuels des six indices de 2004 à 2012.

4.2.2 Résultats

Nous considérons dans un premier temps le problème robuste de minimisation de la CVaR. Notons que nous n'imposons ici aucune contrainte de rendement, mais interdisons toutefois la vente d'actions à découvert. La figure 4.4 montre l'évolution des rendements cumulés des portefeuilles robustes et non robustes avec et sans options. Les résultats sont également résumés dans les tableaux 4.5 et 4.6 sur trois différentes périodes entre 2004 et 2012. Les périodes de 2004 à 2008 et de 2008 à 2012 correspondent approximativement aux périodes avant et après la crise financière de 2008, respectivement.

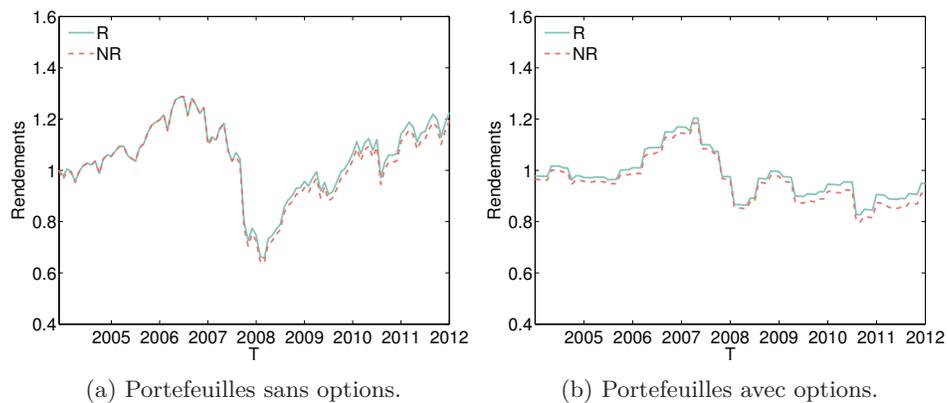


FIGURE 4.4 – Rendements cumulés des portefeuilles robustes (R) et non robustes (NR) avec et sans options.

Nous constatons au premier abord que la performance sur la période étudiée des deux portefeuilles avec options est beaucoup plus faible que celle de leur homologue sans options. À cet égard, il importe de rappeler que l'objectif de ces problèmes d'optimisation est la minimisation de la CVaR, et non la recherche de profit. Pour mieux comprendre la performance des portefeuilles avec options, nous séparons à la figure 4.5 le rendement des actions de celui des options au sein des différents portefeuilles.

	2004-2012		2004-2008		2008-2012	
	R	NR	R	NR	R	NR
#	97	97	48	48	49	49
Moyenne	-0.01%	-0.05%	-0.02%	-0.03%	0.00%	-0.08%
Médiane	-0.04%	-0.12%	-0.02%	-0.09%	-0.13%	-0.24%
Écart-type	3.03%	3.06%	2.63%	2.64%	3.39%	3.41%
Aplatissement	9.23	9.07	7.49	7.54	9.13	8.86
Asymétrie	-1.34	-1.32	-0.88	-0.92	-1.50	-1.44
Minimum	-13.32%	-13.32%	-9.10%	-9.10%	-13.32%	-13.32%
Maximum	8.72%	8.72%	7.03%	7.03%	8.72%	8.72%

TABLEAU 4.6 – Statistiques descriptives des rendements mensuels hors-échantillon des portefeuilles robustes (R) et non robustes (NR) avec options.

	2004-2012		2004-2008		2008-2012	
	R	NR	R	NR	R	NR
#	97	97	48	48	49	49
Moyenne	0.34%	0.32%	-0.39%	-0.44%	1.05%	1.06%
Médiane	1.51%	1.52%	0.84%	0.81%	2.29%	2.29%
Écart-type	4.91%	5.01%	5.16%	5.27%	4.55%	4.61%
Aplatissement	8.00	8.30	9.49	9.91	4.56	4.42
Asymétrie	-1.58	-1.64	-2.02	-2.12	-0.90	-0.91
Minimum	-23.54%	-25.70%	-23.54%	-24.36%	-13.32%	-13.32%
Maximum	11.72%	11.72%	7.03%	7.03%	11.72%	11.72%

TABLEAU 4.5 – Statistiques descriptives des rendements mensuels hors-échantillon des portefeuilles robustes (R) et non robustes (NR) sans options.

Nous remarquons d'emblée que les options des portefeuilles (robustes ou non) performant particulièrement bien durant la crise financière de 2008. Le gain attribuable aux options vient alors compenser la chute des prix des indices. Notons que la quasi-totalité des options « achetées » par les deux modèles sont des options de vente. Ce constat n'est guère surprenant. En effet, une option de vente protège contre la chute du prix d'un actif. L'achat d'un certain actif couplé avec l'achat d'une option de vente sur ce même actif est une stratégie éprouvée appelée « option de vente de protection » ou « *protective put* ».

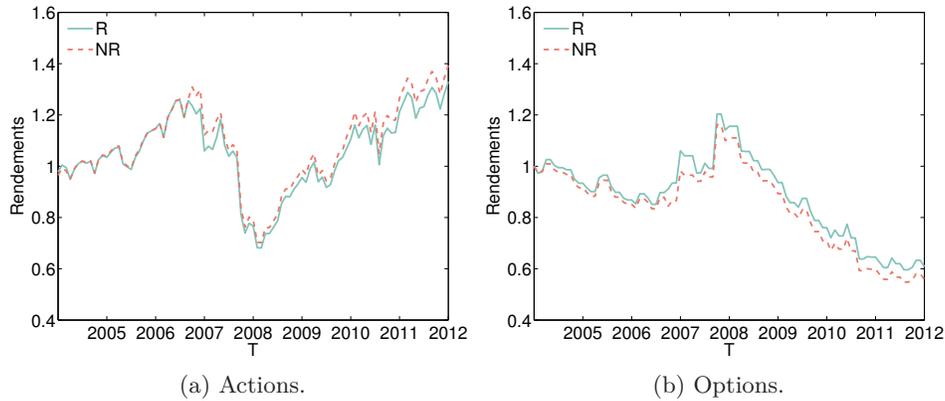


FIGURE 4.5 – Rendements cumulés des composantes « actions » et « options » des portefeuilles avec options robustes (R) et non robustes (NR).

L'utilisation d'options agit ainsi à titre d'assurance contre une baisse des prix des sous-jacents. Cette assurance a, bien sûr, un coût qui se traduit par un rendement négatif des options lorsque les indices sont en hausse. Comme le démontre la figure 4.6, la majorité des options sont de plus achetées aux alentours de la crise financière, lorsque la volatilité et le risque de pertes sont les plus élevés.

En dépit d'un niveau de risque plus faible, le rendement total cumulé à la fin de la période est plus élevé pour les portefeuilles robustes (avec et sans options) que pour les portefeuilles non robustes. Comme l'indiquent les tableaux 4.5 et 4.6, le rendement moyen des portefeuilles robustes avec et sans options est plus élevé, tandis que le risque, mesuré par l'écart-type des rendements, est légèrement plus faible.

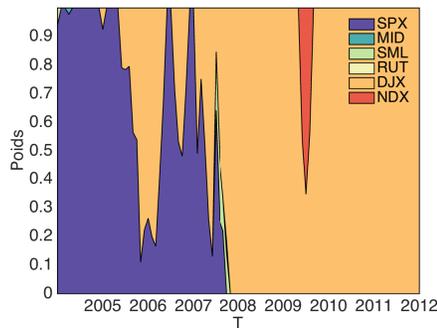
Le tableau 4.7 compare la CVaR réalisée des portefeuilles robustes et non robustes pour différents β . La CVaR réalisée est calculée à partir des rendements hors-échantillon réalisés par les différents portefeuilles optimaux. Nous avons également construit de façon plutôt élémentaire des intervalles de confiance de niveau 95% autour des valeurs estimées. Ces derniers ont été calculés sur les rendements hors-échantillon inférieurs à la VaR réalisée.

Plus précisément, les intervalles de confiance correspondent à plus ou moins 1.96 fois l'écart-type des rendements hors-échantillon qui se trouvent sous la VaR réalisée, où 1.96 est la valeur critique d'une loi normale centrée et réduite pour un seuil de 5%. Dans le cas où le β est de 0.99, puisque l'intervalle n'était calculé que sur un seul point, nous avons choisi de ne pas le présenter.

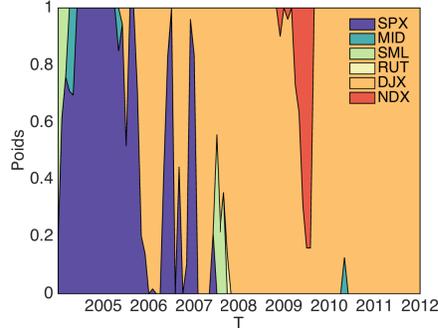
β	Sans options		Avec options	
	R	NR	R	NR
0.99	25.65%	25.70%	13.32%	13.32%
0.95	13.55% \pm 1.69%	13.73% \pm 1.62%	9.91% \pm 1.62%	9.98% \pm 1.62%
0.90	10.69% \pm 2.83%	10.75% \pm 2.22%	6.23% \pm 2.22%	6.31% \pm 2.22%
0.80	7.55% \pm 1.78%	7.62% \pm 1.55%	3.34% \pm 1.54%	3.40% \pm 1.55%
0.50	3.20% \pm 1.10%	3.30% \pm 0.73%	1.71% \pm 0.73%	1.71% \pm 0.73%

TABLEAU 4.7 – CVaR réalisée hors-échantillon des portefeuilles robustes (R) et non robustes (NR), avec et sans options, en fonction de β .

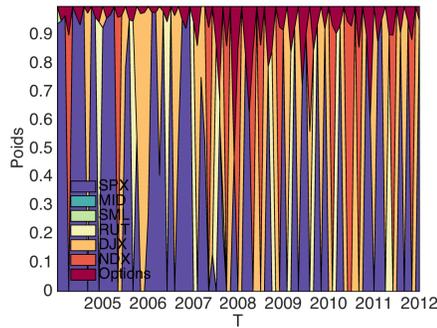
À première vue, la CVaR réalisée des portefeuilles avec options est toujours plus faible ou égale à celle des portefeuilles sans options. Qui plus est, la CVaR des portefeuilles robustes est toujours plus faible ou égale à celle des portefeuilles non robustes, et ce, peu importe le β choisi. Il importe toutefois d'apporter un léger bémol. Même si une tendance claire se dessine, il est néanmoins difficile de confirmer la supériorité du modèle robuste en raison de l'ampleur des intervalles de confiance. Les intervalles de confiance mesurés ici sont, en effet, très larges en raison notamment de la faible taille des échantillons.



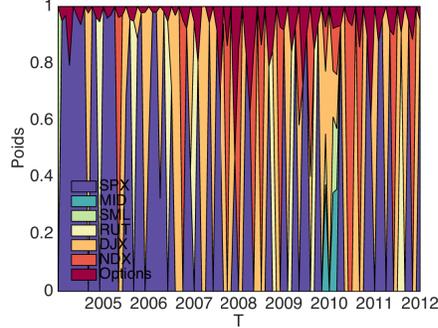
(a) Portefeuille robuste sans options.



(b) Portefeuille non robuste sans options.



(c) Portefeuille robuste avec options.



(d) Portefeuille non robuste avec options.

FIGURE 4.6 – Composition des portefeuilles robustes (R) et non robustes (NR), avec et sans options.

4.2.3 Ratio de Sharpe robuste

Le problème de maximisation du ratio de Sharpe robuste est examiné séparément. La raison est simple. D'abord, le problème d'optimisation (3.42) n'est réalisable que si $\tau \cdot \text{CVaR}_\beta(\mathbf{x}) \leq \mathbb{E}_F[-c(\mathbf{x}, \boldsymbol{\xi})] - r_f$. Puisque $\tau > 0$ et $\text{CVaR}_\beta(\mathbf{x}) > 0$ (par hypothèse), le rendement excédentaire espéré (c.-à-d. $\mathbb{E}_F[-c(\mathbf{x}, \boldsymbol{\xi})] - r_f$) doit alors assurément être positif. Malencontreusement, l'inégalité ci-dessus ne tient pas toujours, en particulier lorsque le rendement de tous les actifs du portefeuille est négatif. C'est le cas, entre autres, durant plusieurs mois au cours de la période étudiée.

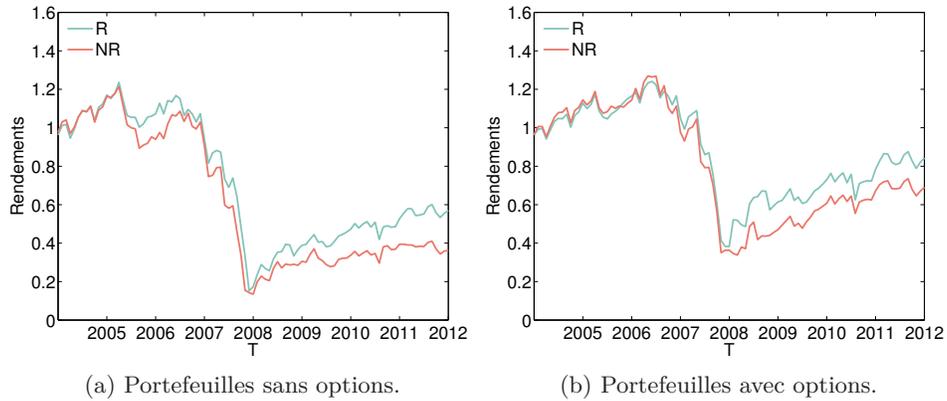


FIGURE 4.7 – Rendements cumulés des portefeuilles « Sharpe-optimaux » robustes et non robustes, avec et sans options.

Pour remédier à ce problème, nous permettons désormais la vente à découvert des actifs sous-jacents, jusqu'à un maximum de 150%. Mathématiquement, nous ajoutons simplement la contrainte $\mathbf{x}^e \geq -1.5$ au problème (3.42). Nous posons de plus $r_f = 0$ de façon à simplifier le problème et garantir que $\mathbb{E}_F[-c(\mathbf{x}, \boldsymbol{\xi})] > r_f$ soit possible.

La figure 4.7 présente les rendements cumulés des portefeuilles robustes et non robustes avec ratio de Sharpe optimal dans les cas avec et sans options. Les résultats sont également résumés dans les tableaux 4.8 et 4.9. De son côté, la figure 4.8 montre l'évolution du ratio de Sharpe optimal pour les différents portefeuilles.

Nous remarquons dans un premier temps que les portefeuilles robustes avec et sans options se démarquent de leurs homologues non robustes par un rendement total plus élevé à la fin de la période analysée. L'ajout d'options vient à la fois réduire la volatilité des portefeuilles et accroître le ratio de Sharpe (voir la figure 4.8), en plus d'augmenter le rendement moyen.

	2004-2012		2004-2008		2008-2012	
	R	NR	R	NR	R	NR
#	97	97	48	48	49	49
Moyenne	0.06%	-0.34%	-3.04%	-3.19%	3.11%	2.44%
Médiane	1.34%	0.46%	-0.10%	0.02%	2.44%	0.81%
Écart-type	10.45%	11.19%	10.82%	10.80%	9.10%	10.86%
Aplatissement	11.07	11.05	11.93	11.59	5.80	8.14
Asymétrie	-1.29	-0.29	-2.68	-2.63	0.97	1.90
Minimum	-54.21%	-54.21%	-54.21%	-54.21%	-17.57%	-14.70%
Maximum	36.35%	47.49%	7.03%	8.10%	36.35%	47.49%

TABLEAU 4.8 – Statistiques descriptives des rendements mensuels des portefeuilles « Sharpe-optimaux » robustes (R) et non robustes (NR) sans options.

	2004-2012		2004-2008		2008-2012	
	R	NR	R	NR	R	NR
#	97	97	48	48	49	49
Moyenne	0.13%	-0.05%	-1.69%	-1.70%	1.92%	1.56%
Médiane	1.34%	1.55%	1.10%	0.55%	2.15%	2.69%
Écart-type	7.81%	7.77%	7.11%	8.09%	8.05%	7.07%
Aplatissement	9.81	9.93	8.02	9.45	9.34	8.21
Asymétrie	0.15	-1.07	-2.01	-2.27	1.49	0.74
Minimum	-31.25%	-37.35%	-31.25%	-37.35%	-17.57%	-17.95%
Maximum	36.35%	30.92%	6.50%	7.88%	36.35%	30.92%

TABLEAU 4.9 – Statistiques descriptives des rendements mensuels des portefeuilles « Sharpe-optimaux » robustes (R) et non robustes (NR) avec options.

Mentionnons que la CVaR réalisée (c.-à-d. calculée sur les rendements hors-échantillon réalisés) est toujours plus faible pour les portefeuilles robustes. En effet, elle est de -20.02% et -28.78% respectivement pour les portefeuilles robustes avec et sans options, contre -22.22% et -29.15% respectivement pour les portefeuilles non robustes. De son côté, le ratio de Sharpe hors-échantillon réalisé sur la période considérée est de 0.225% et -0.778% respectivement pour les portefeuilles robustes et non robustes sans options, et de 0.671% et -0.237% respectivement pour les portefeuilles avec options.

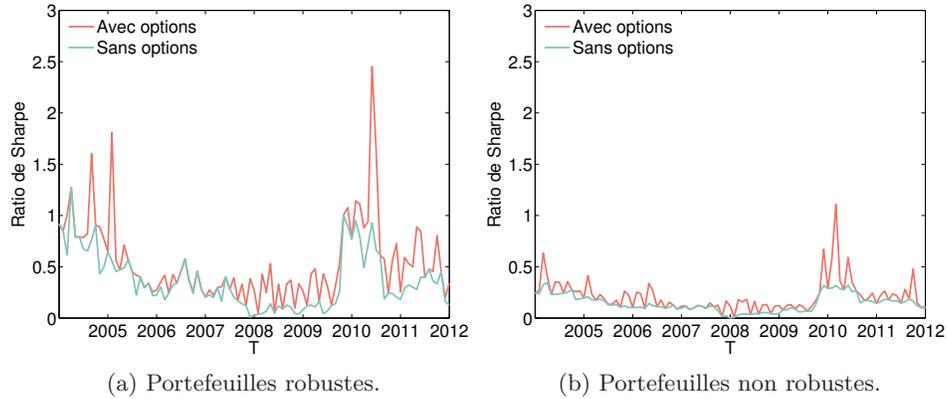


FIGURE 4.8 – Ratios de Sharpe optimaux des portefeuilles avec et sans options, robustes (R) et non robustes (NR).

L'ajout d'options permet d'augmenter le ratio de Sharpe, tant pour le portefeuille robuste que non robuste. Les ratios de Sharpe réalisés des portefeuilles robustes sont également nettement plus élevés que ceux des portefeuilles non robustes.

Chapitre 5

Conclusion

Dans ce mémoire, nous avons approfondi l'étude d'un cadre d'optimisation robuste de la valeur-à-risque conditionnelle d'un portefeuille avec et sans options, où l'incertitude est caractérisée par l'ensemble d'incertitude distributionnel introduit par Delage et Ye [DY10]. Le modèle étudié ici tient compte de l'incertitude inhérente à l'estimation des moments des rendements du portefeuille.

D'abord, nous avons formulé un cadre d'optimisation robuste de la valeur-à-risque conditionnelle avec et sans options. Puis, nous avons mis sur pied une version robuste du ratio de Sharpe, où le rendement excédentaire escompté est mis en relation avec la valeur-à-risque conditionnelle robuste. Nous avons ensuite mis ces modèles à l'épreuve à l'aide de données théoriques et historiques.

Si les deux portefeuilles sans options, robustes et non robustes, s'en tirent plutôt bien sur la période étudiée, on ne peut en dire autant des portefeuilles avec options. La raison est simple. L'ajout d'options, tant dans le portefeuille robuste que non robuste, fournit une « protection » appréciable contre une chute des prix des sous-jacents. Néanmoins, cette protection a un coût qui, ultimement, gruge la rentabilité des portefeuilles robustes et non robustes.

Malgré une performance plus intéressante durant la crise financière, cette « assurance » finit pas coûter cher aux deux portefeuilles. Le résultat est un portefeuille moins volatile, mais, également, moins performant. Ainsi, avant d'utiliser un tel modèle en pratique, il importe donc de bien comprendre ses avantages et ses limites.

Les résultats obtenus hors-échantillon favorisent légèrement les différents portefeuilles robustes, dans la mesure où leur rendement moyen réalisé est plus élevé, et l'écart-type, plus faible. Néanmoins, il nous paraît prématuré de proclamer la supériorité du modèle robuste compte tenu de la taille des échantillons utilisés. À cet effet, davantage d'expériences sur différents échantillons de données devraient être réalisées.

Pour faire suite à ce mémoire, nous croyons qu'il serait intéressant de se pencher sur une forme dynamique du cadre d'optimisation robuste présenté dans ce mémoire. En effet, l'analyse de l'évolution des rendements sous-entend une politique d'optimisation dynamique, alors que le modèle étudié ici est statique. Un tel modèle dynamique serait également en mesure de profiter d'options plus complexes (américaines ou bermudiennes, par exemple), où l'exercice peut s'effectuer avant l'échéance.

Annexe A

Démonstrations

A.1 Équation (3.31)

Démonstration. Considérons l'expression ci-dessous :

$$(\mathbf{x}^d)^\top (\max \{-\mathbf{1}, \mathbf{b} + \mathbf{A}\boldsymbol{\xi} - \mathbf{1}\}), \quad (\text{A.1})$$

où $\mathbf{A} \in \mathbb{R}^{(m-n) \times n}$, $\mathbf{b} \in \mathbb{R}^{m-n}$, $\mathbf{1} \in \mathbb{R}^{m-n}$, et $\mathbf{x}^d \geq 0$. Avec un peu d'algèbre, nous reformulons l'expression ci-dessus comme un programme linéaire équivalent, à savoir :

$$\begin{aligned} \min_{z \in \mathbb{R}^{m-n}} \quad & z^\top \mathbf{x}^d \\ \text{sujet à} \quad & z \geq -\mathbf{1} \\ & z \geq \mathbf{b} + \mathbf{A}\boldsymbol{\xi} - \mathbf{1}. \end{aligned} \quad (\text{A.2})$$

Le dual lagrangien du problème (A.2) s'écrit alors

$$\max_{\mathbf{y} \geq 0, \boldsymbol{\lambda} \geq 0} \left\{ \min_{z \in \mathbb{R}^{m-n}} z^\top \mathbf{x}^d - \boldsymbol{\lambda}^\top (z + \mathbf{1}) + \mathbf{y}^\top (\mathbf{b} + \mathbf{A}\boldsymbol{\xi} - z - \mathbf{1}) \right\}, \quad (\text{A.3})$$

où $\boldsymbol{\lambda} \in \mathbb{R}^{m-n}$ et $\boldsymbol{y} \in \mathbb{R}^{m-n}$ sont des variables duales. En résolvant l'intérieur de l'accolade, nous obtenons ensuite :

$$\begin{aligned} & \min_{\boldsymbol{z} \in \mathbb{R}^{m-n}} \boldsymbol{z}^\top (\boldsymbol{x}^d - \boldsymbol{\lambda} - \boldsymbol{y}) - \boldsymbol{\lambda}^\top \mathbf{1} + \boldsymbol{y}^\top (\boldsymbol{b} + \boldsymbol{A}\boldsymbol{\xi} - \mathbf{1}) \\ & = \begin{cases} -\boldsymbol{\lambda}^\top \mathbf{1} + \boldsymbol{y}^\top (\boldsymbol{b} + \boldsymbol{A}\boldsymbol{\xi} - \mathbf{1}) & \text{si } \boldsymbol{x}^d - \boldsymbol{\lambda} - \boldsymbol{y} = \mathbf{0} \\ -\infty & \text{sinon} \end{cases} \end{aligned} \quad (\text{A.4})$$

Puis, en posant $\boldsymbol{\lambda} = \boldsymbol{x}^d - \boldsymbol{y}$, le problème (A.3) devient :

$$\max_{\boldsymbol{y} \in \mathbb{R}^{m-n}} \boldsymbol{y}^\top (\boldsymbol{b} + \boldsymbol{A}\boldsymbol{\xi}) - \mathbf{1}^\top \boldsymbol{x}^d, \quad (\text{A.5})$$

où $\boldsymbol{x}^d \geq \boldsymbol{y} \geq \mathbf{0}$. Cette contrainte est, bien sûr, due au fait que $\boldsymbol{\lambda} = \boldsymbol{x}^d - \boldsymbol{y} \geq \mathbf{0}$. □

A.2 Proposition 3.1

Proposition. La valeur optimale du problème d'optimisation robuste de la valeur-à-risque conditionnelle (3.26) d'un portefeuille d'options, pour lequel \mathcal{D} est l'ensemble $\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ et $\mathcal{S} = \mathbb{R}^n$, peut être trouvée en résolvant le problème d'optimisation semi-défini qui suit :

$$\min_{x, \mathbf{Q}, \mathbf{P}, \mathbf{p}, \mathbf{y}, r, s, \alpha} (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \mathbf{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \mathbf{P}) - 2\boldsymbol{\mu}_0^\top \mathbf{p} + \gamma_1 s \quad (\text{A.6a})$$

$$\text{sujet à } \begin{bmatrix} \mathbf{Q} & \mathbf{q} \\ \mathbf{q}^\top & r - \alpha \end{bmatrix} \succeq 0 \quad (\text{A.6b})$$

$$\begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{1}{2} \left(\frac{\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}}{1-\beta} \right) \\ \mathbf{q}^\top + \frac{1}{2} \left(\frac{\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}}{1-\beta} \right)^\top & r - \alpha + \left(\frac{\alpha + \mathbf{y}^\top \mathbf{b} - \mathbf{1}^\top \mathbf{x}^d}{1-\beta} \right) \end{bmatrix} \succeq 0 \quad (\text{A.6c})$$

$$\mathbf{Q} \succeq 0, \quad (\text{A.6d})$$

$$\begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^\top & s \end{bmatrix} \succeq 0 \quad (\text{A.6e})$$

$$\mathbf{q} = -(\mathbf{p} + \mathbf{Q}\boldsymbol{\mu}_0) \quad (\text{A.6f})$$

$$\mathbf{x}^d \geq \mathbf{0}, \quad \mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}, \quad \mathbf{x}^\top \mathbf{1} = 1. \quad (\text{A.6g})$$

Démonstration. À partir du lemme 3.1, nous reformulons le problème (3.26) comme le problème dual qui suit :

$$\min_{x, \mathbf{Q}, \mathbf{P}, \mathbf{p}, r, s, \alpha} (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \mathbf{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \mathbf{P}) - 2\boldsymbol{\mu}_0^\top \mathbf{p} + \gamma_1 s \quad (\text{A.7a})$$

$$\text{sujet à } \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q}\boldsymbol{\mu}_0) + r - \mathcal{Q}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] \geq 0, \quad \forall \boldsymbol{\xi} \in \mathcal{S} \quad (\text{A.7b})$$

$$\mathbf{Q} \succeq 0, \quad (\text{A.7c})$$

$$\begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^\top & s \end{bmatrix} \succeq 0, \quad (\text{A.7d})$$

où $\boldsymbol{\xi} \in \mathbb{R}^n \subseteq \mathcal{X}$, $\mathbf{x} \in \mathbb{R}^m$, $\boldsymbol{\Sigma}_0 \succ 0$, $\gamma_1 \geq 0$, $\gamma_2 \geq 0$, $r \in \mathbb{R}$, $\mathbf{Q} \in \mathbb{R}^{n \times n}$, $\mathbf{P} \in \mathbb{R}^{n \times n}$, $\mathbf{p} \in \mathbb{R}^n$, $s \in \mathbb{R}$, $\alpha \in \mathbb{R}$ et $\mathcal{Q}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha] = \max\{(1 - (1 - \beta)^{-1})\alpha +$

$(1 - \beta)^{-1} c(\mathbf{x}, \boldsymbol{\xi}), \alpha\}$. Pour tout $\boldsymbol{\xi} \in \mathcal{S}$, l'inégalité (A.7b) du problème ci-dessus est équivalente aux deux inégalités ci-dessous :

$$\boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - \alpha + \frac{\alpha}{1 - \beta} - \frac{c(\mathbf{x}, \boldsymbol{\xi})}{1 - \beta} \geq 0 \quad (\text{A.8})$$

$$\boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - \alpha \geq 0. \quad (\text{A.9})$$

Attaquons-nous, en premier lieu, à l'inégalité (A.9). Puisque $\mathcal{S} = \mathbb{R}^n$, l'équation $\boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - \alpha$ est non négative si et seulement si

$$\inf_{\boldsymbol{\xi}} \{ \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - \alpha \} \geq 0, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n. \quad (\text{A.10})$$

Puisque $\mathbf{Q} \succeq 0$, la fonction $\boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - \alpha$ est convexe par rapport à $\boldsymbol{\xi}$ et concave (linéaire) par rapport à α . Avec un peu d'algèbre, nous reformulons l'inéquation comme suit :

$$\begin{bmatrix} \boldsymbol{\xi}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{Q} & -(\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) \\ -(\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0)^\top & r - \alpha \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi} \\ 1 \end{bmatrix} \geq 0, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^n,$$

ou, de façon équivalente,

$$\begin{bmatrix} \mathbf{Q} & -(\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) \\ -(\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0)^\top & r - \alpha \end{bmatrix} \succeq 0. \quad (\text{A.11})$$

Dans un deuxième temps, puisque

$$c(\mathbf{x}, \boldsymbol{\xi}) = \min_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ -\boldsymbol{\xi}^\top \mathbf{x}^e - \mathbf{y}^\top (\mathbf{b} + \mathbf{A} \boldsymbol{\xi}) + \mathbf{1}^\top \mathbf{x}^d \right\},$$

l'inégalité (A.8) peut s'écrire

$$\begin{aligned} \forall \boldsymbol{\xi} \in \mathbb{R}^n, \quad \exists \mathbf{y} \in [0, \mathbf{x}^d] : \quad & \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - \alpha + \frac{\alpha}{1 - \beta} \\ & + \frac{\boldsymbol{\xi}^\top \mathbf{x}^e}{1 - \beta} + \frac{\mathbf{y}^\top (\mathbf{b} + \mathbf{A} \boldsymbol{\xi})}{1 - \beta} - \frac{\mathbf{1}^\top \mathbf{x}^d}{1 - \beta} \geq 0, \end{aligned}$$

ou, de façon équivalente,

$$\begin{aligned} \inf_{\boldsymbol{\xi} \in \mathbb{R}^n} \max_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \quad & \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0) + r - \alpha + \frac{\alpha}{1 - \beta} \\ & + \frac{\boldsymbol{\xi}^\top \mathbf{x}^e}{1 - \beta} + \frac{\mathbf{y}^\top \mathbf{b}}{1 - \beta} + \frac{\mathbf{y}^\top \mathbf{A} \boldsymbol{\xi}}{1 - \beta} - \frac{\mathbf{1}^\top \mathbf{x}^d}{1 - \beta} \geq 0. \end{aligned}$$

Le problème ci-dessus est convexe par rapport à $\boldsymbol{\xi}$ et linéaire par rapport à \mathbf{y} . Par conséquent, il est possible d'interchanger le min et le max dans l'équation précédente (voir Sion [Sio58]) afin d'obtenir l'inégalité matricielle qui suit :

$$\exists \mathbf{y} \in [0, \mathbf{x}^d] : \begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{1}{2} \left(\frac{\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}}{1 - \beta} \right) \\ \mathbf{q}^\top + \frac{1}{2} \left(\frac{\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}}{1 - \beta} \right)^\top & r - \alpha + \left(\frac{\alpha + \mathbf{y}^\top \mathbf{b} - \mathbf{1}^\top \mathbf{x}^d}{1 - \beta} \right) \end{bmatrix} \succeq 0, \quad (\text{A.12})$$

où $\mathbf{q} = -(\mathbf{p} + \mathbf{Q} \boldsymbol{\mu}_0)$. En remplaçant la contrainte (A.7b) du problème dual par (A.11) et (A.12), nous obtenons enfin la proposition 3.1. \square

A.3 Proposition 3.2

Proposition. La valeur optimale du problème d'optimisation robuste de la valeur-à-risque conditionnelle (3.26) d'un portefeuille d'options, pour lequel \mathcal{D} est l'ensemble $\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ et $\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^n : \boldsymbol{l} \leq \boldsymbol{\xi} \leq \boldsymbol{u}\}$, peut être trouvée en résolvant le problème d'optimisation semi-défini qui suit :

$$\begin{aligned}
& \min_{\boldsymbol{x}, \boldsymbol{Q}, \boldsymbol{P}, \boldsymbol{p}, \boldsymbol{y}, r, s, \alpha} \\
& \quad \lambda_1, \boldsymbol{v}_1, \lambda_2, \boldsymbol{v}_2} \quad (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \boldsymbol{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \boldsymbol{P}) - 2\boldsymbol{\mu}_0^\top \boldsymbol{p} + \gamma_1 s \\
& \text{sujet à} \quad \boldsymbol{\Xi}_1 = \begin{bmatrix} \mathbf{0} & \left(\frac{\boldsymbol{v}_1 - \lambda_1}{2}\right) \\ \left(\frac{\boldsymbol{v}_1 - \lambda_1}{2}\right)^\top & \boldsymbol{\lambda}_1^\top \boldsymbol{l} - \boldsymbol{v}_1^\top \boldsymbol{u} \end{bmatrix} \\
& \quad \boldsymbol{\Xi}_2 = \begin{bmatrix} \mathbf{0} & \left(\frac{\boldsymbol{v}_2 - \lambda_2}{2}\right) \\ \left(\frac{\boldsymbol{v}_2 - \lambda_2}{2}\right)^\top & \boldsymbol{\lambda}_2^\top \boldsymbol{l} - \boldsymbol{v}_2^\top \boldsymbol{u} \end{bmatrix} \\
& \quad \boldsymbol{\Xi}_1 + \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q} \\ \boldsymbol{q}^\top & r - \alpha \end{bmatrix} \succeq 0 \\
& \quad \boldsymbol{\Xi}_2 + \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q} + \frac{1}{2} \left(\frac{\boldsymbol{x}^e + \boldsymbol{A}^\top \boldsymbol{y}}{1 - \beta}\right) \\ \boldsymbol{q}^\top + \frac{1}{2} \left(\frac{\boldsymbol{x}^e + \boldsymbol{A}^\top \boldsymbol{y}}{1 - \beta}\right)^\top & r - \alpha + \left(\frac{\alpha + \boldsymbol{y}^\top \boldsymbol{b} - \mathbf{1}^\top \boldsymbol{x}^d}{1 - \beta}\right) \end{bmatrix} \succeq 0 \\
& \quad \boldsymbol{Q} \succeq 0, \\
& \quad \begin{bmatrix} \boldsymbol{P} & \boldsymbol{p} \\ \boldsymbol{p}^\top & s \end{bmatrix} \succeq 0 \\
& \quad \boldsymbol{q} = -(\boldsymbol{p} + \boldsymbol{Q} \boldsymbol{\mu}_0) \\
& \quad \lambda_1 \geq 0, \quad \boldsymbol{v}_1 \geq \mathbf{0}, \quad \lambda_2 \geq 0, \quad \boldsymbol{v}_2 \geq \mathbf{0} \\
& \quad \boldsymbol{x}^d \geq \mathbf{0}, \quad \boldsymbol{x}^d \geq \boldsymbol{y} \geq \mathbf{0}, \quad \boldsymbol{x}^\top \mathbf{1} = 1.
\end{aligned}$$

Démonstration. La démonstration est similaire à celle de la proposition 3.1, à une exception près.

Considérons le problème quadratique suivant :

$$\begin{aligned} \min_{\xi \in \mathbb{R}^n} \quad & \xi^\top \mathbf{Q} \xi + \mathbf{q}^\top \xi + q_0 \geq 0 \\ \text{sujet à} \quad & \xi \leq \mathbf{u} \\ & \mathbf{l} \leq \xi. \end{aligned}$$

où $\xi \in \mathbb{R}^n$, $\mathbf{Q} \in \mathbb{R}^{n \times n}$, $\mathbf{Q} \succeq 0$, $\mathbf{q} \in \mathbb{R}^n$, $q_0 \in \mathbb{R}$, $\mathbf{u} \in \mathbb{R}^n$ et $\mathbf{l} \in \mathbb{R}^n$. La forme duale s'écrit :

$$\max_{\lambda, \mathbf{v} \geq \mathbf{0}} \min_{\xi \in \mathbb{R}^n} \xi^\top \mathbf{Q} \xi + \mathbf{q}^\top \xi + q_0 + \lambda^\top (\mathbf{l} - \xi) + \mathbf{v}^\top (\xi - \mathbf{u}) \geq 0,$$

ou, de façon équivalente,

$$\exists \lambda, \mathbf{v} \geq \mathbf{0}, \forall \xi \in \mathbb{R}^n : \xi^\top \mathbf{Q} \xi + \mathbf{q}^\top \xi + q_0 + \lambda^\top (\mathbf{l} - \xi) + \mathbf{v}^\top (\xi - \mathbf{u}) \geq 0.$$

Sous forme matricielle, nous obtenons enfin l'inégalité linéaire qui suit :

$$\exists \lambda, \mathbf{v} \geq \mathbf{0} : \begin{bmatrix} \mathbf{Q} & \frac{1}{2}(\mathbf{q} - \lambda + \mathbf{v}) \\ \frac{1}{2}(\mathbf{q} - \lambda + \mathbf{v})^\top & q_0 + \lambda^\top \mathbf{l} - \mathbf{v}^\top \mathbf{u} \end{bmatrix} \succeq 0.$$

Sans entrer dans le détail, nous employons la même démarche pour transformer deux des contraintes semi-définies de la proposition 3.1 afin qu'elles tiennent compte du nouveau support du vecteur ξ . \square

A.4 Proposition 3.3

Proposition. Pour un $\tau \geq 0$ fixe, si l'ensemble distributionnel \mathcal{D} est l'ensemble $\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \gamma_1, \gamma_2)$ et $\mathcal{S} = \mathbb{R}^n$, alors la valeur optimale du sous-problème d'optimisation robuste du ratio de Sharpe présenté dans l'inégalité (3.41) peut être trouvée en résolvant le problème dual qui suit :

$$\begin{aligned}
& \min_{x, \mathbf{Q}, \mathbf{P}, \mathbf{p}, \mathbf{y}, \mathbf{y}_1, r, s, \alpha} (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \mathbf{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \mathbf{P}) - 2\boldsymbol{\mu}_0^\top \mathbf{p} + \gamma_1 s \\
& \text{sujet à } \begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{1}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}) \\ \mathbf{q}^\top + \frac{1}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y})^\top & r - \alpha\tau - r_f + \mathbf{y}^\top \mathbf{b} - \mathbf{1}^\top \mathbf{x}^d \end{bmatrix} \succeq 0 \\
& \begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{\kappa}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}_1) \\ \mathbf{q}^\top + \frac{\kappa}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}_1)^\top & r - \left(\tau - \frac{\tau}{1-\beta}\right)\alpha - r_f + \kappa \mathbf{y}_1^\top \mathbf{b} - \kappa \mathbf{1}^\top \mathbf{x}^d \end{bmatrix} \succeq 0 \\
& \mathbf{Q} \succeq 0, \\
& \begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^\top & s \end{bmatrix} \succeq 0 \\
& \mathbf{q} = -(\mathbf{p} + \mathbf{Q}\boldsymbol{\mu}_0), \quad \kappa = \left(1 + \frac{\tau}{1-\beta}\right) \\
& \mathbf{x}^d \geq \mathbf{0}, \quad \mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}, \quad \mathbf{x}^d \geq \mathbf{y}_1 \geq \mathbf{0}, \quad \mathbf{x}^\top \mathbf{1} = 1.
\end{aligned}$$

Démonstration. À partir du lemme 3.1, nous reformulons donc le problème (3.39) comme le problème dual qui suit :

$$\min_{x, \mathbf{Q}, \mathbf{P}, \mathbf{p}, r, s, \alpha} (\gamma_2 \boldsymbol{\Sigma}_0 - \boldsymbol{\mu}_0 \boldsymbol{\mu}_0^\top) \bullet \mathbf{Q} + r + (\boldsymbol{\Sigma}_0 \bullet \mathbf{P}) - 2\boldsymbol{\mu}_0^\top \mathbf{p} + \gamma_1 s \quad (\text{A.13a})$$

$$\text{sujet à } \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} - 2\boldsymbol{\xi}^\top (\mathbf{p} + \mathbf{Q}\boldsymbol{\mu}_0) + r - \mathcal{G}[c(\mathbf{x}, \boldsymbol{\xi}), \alpha, \tau] \geq 0, \quad \forall \boldsymbol{\xi} \in \mathcal{S} \quad (\text{A.13b})$$

$$\mathbf{Q} \succeq 0, \quad (\text{A.13c})$$

$$\begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^\top & s \end{bmatrix} \succeq 0. \quad (\text{A.13d})$$

Pour tout $\xi \in \mathcal{S}$, la contrainte (A.13b) est équivalente à :

$$\xi^\top \mathbf{Q} \xi + 2\xi^\top \mathbf{q} + r - \left(\tau - \frac{\tau}{1-\beta} \right) \alpha - \left(1 + \frac{\tau}{1-\beta} \right) c(\mathbf{x}, \xi) - r_f \geq 0 \quad (\text{A.14})$$

$$\xi^\top \mathbf{Q} \xi + 2\xi^\top \mathbf{q} + r - c(\mathbf{x}, \xi) - \alpha\tau - r_f \geq 0, \quad (\text{A.15})$$

où $\mathbf{q} = -(\mathbf{p} + \mathbf{Q}\boldsymbol{\mu}_0)$. Puisque que

$$c(\mathbf{x}, \xi) = \min_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ -\xi^\top \mathbf{x}^e - \mathbf{y}^\top (\mathbf{b} + \mathbf{A}\xi) + \mathbf{1}^\top \mathbf{x}^d \right\},$$

l'inégalité (A.15) est égale à :

$$\begin{aligned} & \xi^\top \mathbf{Q} \xi + 2\xi^\top \mathbf{q} + r - \alpha\tau - r_f - \min_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ -\xi^\top \mathbf{x}^e - \mathbf{y}^\top (\mathbf{b} + \mathbf{A}\xi) + \mathbf{1}^\top \mathbf{x}^d \right\} \\ &= \xi^\top \mathbf{Q} \xi + 2\xi^\top \mathbf{q} + r - \alpha\tau - r_f + \max_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ \xi^\top \mathbf{x}^e + \mathbf{y}^\top (\mathbf{b} + \mathbf{A}\xi) - \mathbf{1}^\top \mathbf{x}^d \right\} \\ &\geq 0. \end{aligned}$$

Par conséquent, la contrainte (A.15) peut être réécrite ainsi :

$$\begin{aligned} & \min_{\xi \in \mathcal{S}} \max_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ \xi^\top \mathbf{Q} \xi + 2\xi^\top \mathbf{q} + r - \alpha\tau - r_f \right. \\ & \quad \left. + \xi^\top \mathbf{x}^e + \mathbf{y}^\top (\mathbf{b} + \mathbf{A}\xi) - \mathbf{1}^\top \mathbf{x}^d \right\} \geq 0 \end{aligned} \quad (\text{A.16})$$

Si $\mathcal{S} = \mathbb{R}^n$, alors la contrainte (A.15) est équivalente à l'inégalité matricielle linéaire qui suit :

$$\begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{1}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}) \\ \mathbf{q}^\top + \frac{1}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y})^\top & r - \alpha\tau - r_f + \mathbf{y}^\top \mathbf{b} - \mathbf{1}^\top \mathbf{x}^d \end{bmatrix} \succeq 0. \quad (\text{A.17})$$

Nous procédons de façon analogue pour l'inégalité (A.14). Nous avons donc :

$$\begin{aligned}
& \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} + 2\boldsymbol{\xi}^\top \mathbf{q} + r - \left(\tau - \frac{\tau}{1-\beta} \right) \alpha - \left(1 + \frac{\tau}{1-\beta} \right) c(\mathbf{x}, \boldsymbol{\xi}) - r_f \\
= & \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} + 2\boldsymbol{\xi}^\top \mathbf{q} + r - \left(\tau - \frac{\tau}{1-\beta} \right) \alpha - r_f \\
& - \left(1 + \frac{\tau}{1-\beta} \right) \min_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ -\boldsymbol{\xi}^\top \mathbf{x}^e - \mathbf{y}^\top (\mathbf{b} + \mathbf{A} \boldsymbol{\xi}) + \mathbf{1}^\top \mathbf{x}^d \right\} \\
= & \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} + 2\boldsymbol{\xi}^\top \mathbf{q} + r - \left(\tau - \frac{\tau}{1-\beta} \right) \alpha - r_f \\
& + \left(1 + \frac{\tau}{1-\beta} \right) \max_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ \boldsymbol{\xi}^\top \mathbf{x}^e + \mathbf{y}^\top (\mathbf{b} + \mathbf{A} \boldsymbol{\xi}) - \mathbf{1}^\top \mathbf{x}^d \right\} \\
= & \max_{\mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}} \left\{ \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} + 2\boldsymbol{\xi}^\top \mathbf{q} + r - \left(\tau - \frac{\tau}{1-\beta} \right) \alpha - r_f + \left(1 + \frac{\tau}{1-\beta} \right) \boldsymbol{\xi}^\top \mathbf{x}^e \right. \\
& \left. + \left(1 + \frac{\tau}{1-\beta} \right) \mathbf{y}^\top (\mathbf{b} + \mathbf{A} \boldsymbol{\xi}) - \left(1 + \frac{\tau}{1-\beta} \right) \mathbf{1}^\top \mathbf{x}^d \right\} \\
\geq & 0.
\end{aligned}$$

À l'instar de (A.15), nous reformulons la contrainte (A.14) comme l'inégalité matricielle ci-dessous :

$$\begin{bmatrix} \mathbf{Q} & \mathbf{q} + \frac{\kappa}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y}) \\ \mathbf{q}^\top + \frac{\kappa}{2}(\mathbf{x}^e + \mathbf{A}^\top \mathbf{y})^\top & r - \left(\tau - \frac{\tau}{1-\beta} \right) \alpha - r_f + \kappa \mathbf{y}^\top \mathbf{b} - \kappa \mathbf{1}^\top \mathbf{x}^d \end{bmatrix} \succeq 0,$$

où $\kappa = \left(1 + \frac{\tau}{1-\beta} \right)$.

Le reste de la démonstration est similaire à ce qui a déjà été présenté précédemment. \square

Annexe B

Problème non robuste d'optimisation de la CVaR

Rockafellar et Uryasev [RU00] démontrent que la fonction auxiliaire (3.14) est approximativement équivalente à :

$$\min_{\mathbf{x}, \alpha} \alpha + \frac{1}{k(1-\beta)} \sum_{i=1}^k [c(\mathbf{x}, \boldsymbol{\xi}_i) - \alpha]^+, \quad (\text{B.1})$$

où $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_k$ sont des rendements échantillonnés à partir de la distribution connue de $\boldsymbol{\xi}$ (par exemple, $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$). Les auteurs reformulent le problème ci-dessus comme un problème d'optimisation linéaire équivalent, à savoir :

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}, \alpha} \quad & \alpha + \frac{1}{k(1-\beta)} \sum_{i=1}^k u_i \\ \text{sujet à} \quad & u_i \geq c(\mathbf{x}, \boldsymbol{\xi}_i) - \alpha \\ & u_i \geq 0 \\ & \mathbf{x} \in \mathcal{X} \\ & i = 1, \dots, k, \end{aligned} \quad (\text{B.2})$$

Pour un portefeuille d'options, le problème s'écrit donc :

$$\begin{aligned}
& \min_{\mathbf{x}, \mathbf{y}, \mathbf{u}, \alpha} \alpha + \frac{1}{(1-\beta)} \frac{1}{k} \sum_{i=1}^k u_i \\
& \text{sujet à } u_i \geq -\boldsymbol{\xi}_i^\top \mathbf{x}^d - \mathbf{y}^\top (\mathbf{b} + \mathbf{A}\boldsymbol{\xi}_i) + \mathbf{1}^\top \mathbf{x}^e - \alpha \\
& \quad u_i \geq 0 \\
& \quad \mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0} \\
& \quad \mathbf{x} \in \mathcal{X} \\
& \quad i = 1, \dots, k.
\end{aligned} \tag{B.3}$$

Le problème d'optimisation du ratio de Sharpe robuste, de son côté, est approximativement égal au problème ci-dessous :

$$\begin{aligned}
& \min_{(\mathbf{x}, \alpha) \in \mathcal{X} \times \mathbb{R}} \frac{1}{k} \sum_{i=1}^k \max \left\{ \left(\tau - \frac{\tau}{1-\beta} \right) \alpha + \left(1 + \frac{\tau}{1-\beta} \right) c(\mathbf{x}, \boldsymbol{\xi}_i) + r_f, \right. \\
& \quad \left. c(\mathbf{x}, \boldsymbol{\xi}_i) + \alpha \tau + r_f \right\}.
\end{aligned} \tag{B.4}$$

De façon équivalente, le problème (B.4) peut s'écrire ainsi :

$$\begin{aligned}
& \min_{\mathbf{x}, \mathbf{y}, \mathbf{y}_1, \mathbf{u}, \alpha} \frac{1}{k} \sum_{i=1}^k u_i \\
& \text{sujet à } u_i \geq \left(\tau - \frac{\tau}{1-\beta} \right) \alpha + r_f + \\
& \quad \left(1 + \frac{\tau}{1-\beta} \right) \left(\mathbf{1}^\top \mathbf{x}^e - \boldsymbol{\xi}_i^\top \mathbf{x}^d - \mathbf{y}^\top (\mathbf{b} + \mathbf{A}\boldsymbol{\xi}_i) \right) \\
& \quad u_i \geq \mathbf{1}^\top \mathbf{x}^e - \boldsymbol{\xi}_i^\top \mathbf{x}^d - \mathbf{y}_1^\top (\mathbf{b} + \mathbf{A}\boldsymbol{\xi}_i) + \alpha \tau + r_f \\
& \quad \mathbf{x}^d \geq \mathbf{y} \geq \mathbf{0}, \quad \mathbf{x}^d \geq \mathbf{y}_1 \geq \mathbf{0} \\
& \quad \mathbf{x} \in \mathcal{X} \\
& \quad i = 1, \dots, k.
\end{aligned} \tag{B.5}$$

Bibliographie

- [ACL06] S. Alexander, T. F. Coleman, and Y. Li. Minimizing var and cvar for a portfolio of derivatives. *Journal of Banking and Finance*, 30(2) :583–605, 2006.
- [ADEH99] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3), 1999.
- [AT02] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26(7) :1487–1503, 2002.
- [BDNT09] D. Bertsimas, X. V. Doan, K. Natarajan, and C. P. Teo. Models for minimax stochastic linear optimization problems with risk aversion. Submitted to *Mathematics of Operations Research*, 2009.
- [BL92] F. Black and R. Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48(5) :28–43, 1992.
- [Bor69] K. Borch. A note on uncertainty and indifference curves. *Review of Economic Studies*, 36(105) :1–4, January 1969.
- [BP00] D. Bertsimas and I. Popescu. Optimal inequalities in probability theory : A convex optimization approach. *SIAM J. Optim.*, 15 :780–804, 2000.
- [BTN98] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Math. Oper. Res.*, 23(4) :769–805, 1998.
- [BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

- [CET09] G. Calafiore, L. El Ghaoui, and U. Topcu. Parameter estimation with expected and residual-at-risk criteria. *Systems and Control Letters*, 58(1) :39–46, 2009.
- [Cha83] G. Chamberlain. A characterization of the distributions that imply mean–variance utility functions. *Journal of Economic Theory*, 29(1) :185–201, February 1983.
- [CZ93] V. K. Chopra and W. T. Zeima. The effet of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19(2) :6–11, 1993.
- [DY10] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Oper. Res.*, 58(3) :595–612, 2010.
- [EOL98] L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM J. Optim.*, 9(1) :33–52, 1998.
- [EOO03] L. El Ghaoui, M. Oks, and F. Oustry. Worst-case value-at-risk and robust portfolio optimization : A conic programming approach. *Oper. Res.*, 51(4) :543–556, 2003.
- [Fam65] E. Fama. The behavior of stock market prices. *Journal of Business*, 38(1) :34–105, 1965.
- [FPS71] G. M. Frankfurter, H. E. Phillips, and J. P. Seagle. Portfolio selection : The effects of uncertain means, variances and covariances. *The Journal of Financial and Quantitative Analysis*, 6 :1251–1262, 1971.
- [FS02] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4) :429–447, 2002.
- [GI03] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Math. Oper. Res.*, 28(1) :1–38, 2003.
- [GP05] A. A. Gaivoronski and Georg Pflug. Value-at-risk in portfolio optmization : Properties and computational approach. *Journal of Risk*, 7(2) :1–31, 2005.

- [GUW07] L. Garlappi, R. Uppal, and T. Wang. Portfolio selection with parameter and model uncertainty : A multi-prior approach. *Review of Financial Studies*, 20(1) :41–81, 2007.
- [Hul14] J. C. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, ninth edition, 2014.
- [KCR14] M. Kapsos, N. Christofides, and B. Rustem. Worst-case robust omega ratio. *European Journal of Operational Research*, 234 :499–507, 2014.
- [KS02] C. Keating and W. F. Shadwick. A universal performance measure. *Journal of Performance Measurement*, 36(3) :59–64, 2002.
- [LB99] M. S. Lobo and S. Boyd. The worst-case risk of a portfolio. Technical report, Stanford University, February 1999.
- [Mar52] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1) :77–91, 1952.
- [Mar59] H. Markowitz. *Portfolio Selection : Efficient Diversification of Investments*. John Wiley & Sons, 1959.
- [Mic98] R. Michaud. *Efficient Asset Management*. Harvard Business School Press, 1998.
- [MX13] R. Meskarian and H. Xu. Robust reward-risk ratio optimization : Modelling and numerical methods. Working paper, Singapore University of Technology and Design (SUTD), 2013.
- [NN94] Y. Nesterov and A. Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, 1994.
- [Nol09] P. Nolet. Optimisation robuste de portefeuille. Master’s thesis, HEC Montréal, Montréal, 2009.
- [Pfl00] G. Ch. Pflug. Some remarks on the value-at-risk and the conditional value-at-risk. In E. Uryasev, editor, *Probabilistic Constrained Optimization*, pages 272–281. Kluwer Academic Publishers, 2000.

- [Pop07] I. Popescu. Robust mean-covariance solutions for stochastic optimization. *Oper. Res.*, 55(1) :98–112, 2007.
- [Ris94] RiskMetricsTM. Technical document – second edition. Technical report, J.P. Morgan, 1994.
- [RU00] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3) :21–41, 2000.
- [RU02] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26 :1443–1471, 2002.
- [Sca58] H. Scarf. A min-max solution of an inventory problem. *Studies in The Mathematical Theory of Inventory and Production*, pages 201–209, 1958.
- [Sha00] Alexander Shapiro. On duality theory of conic linear problems. In *Semi-Infinite Programming*, pages 135–165. Kluwer Academic Publishers, 2000.
- [Sha02] W. F. Sharpe. Mutual fund performance. *Journal of Business*, 39(1) :119–138, 2002.
- [Sio58] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1) :171–176, 1958.
- [YY02] Y. Yamai and T. Yoshida. On the validity of value-at-risk : Comparative analyses with expected shortfall. *Monetary and Economic Studies*, 20(1) :57–85, January 2002.
- [ZF09] S. S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Oper. Res.*, 57 :1155–1168, 2009.
- [ZKR12] S. Zymler, D. Kuhn., and B. Rustem. Worst-case value-at-risk of non-linear portfolios. *Management Sci.*, 59(1) :172–188, 2012.
- [ZKR13] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Math. Programming*, 137(1) :167–198, 2013.

