UNIVERSITÉ DU QUÉBEC À MONTRÉAL

L'UTILISATION DES MESURES COHÉRENTES DE RISQUE EN GESTION DE PORTEFEUILLE

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN MATHÉMATIQUES

PAR

SIMON BELZILE

JUIN 2008

REMERCIEMENTS

J'aimerais tout d'abord remercier mes parents et amis, qui m'ont supporté tout au long de mes études universitaires, votre support m'a été très précieux.

J'aimerais aussi remercier tous les étudiants avec qui j'ai eu la chance de discuter à maintes reprise de mathématiques, Marie-Claude Vachon, Vincent Benoliel, Nancy Tremblay et Gabriel Painchaud. Je vous remercie pour l'écoute que vous m'avez apportée et nos nombreuses discussions.

J'aimerais ensuite exprimer toute ma gratitude à mon directeur, René Ferland, pour son support dans ma rédaction, pour avoir nourri ma curiosité en répondant à mes innombrables questions qui, en général, n'avaient aucun lien avec mon sujet de recherche et finalement j'aimerais le remercier pour m'avoir permis de connaître les mathématiques financières. J'aimerais aussi remercier François Watier et Matthieu Dufour pour avoir gentiment accepté d'évaluer ce mémoire.

Pour leur support et encouragement tout au long de mes études académiques, en étant mes directeurs de stages de recherche, ou simplement professeur m'ayant aidé à progresser en mathématiques, merci à Pierre Bouchard pour nos nombreuses discussions et toutes les choses que vous m'avez apprises sur la géométrie algébrique, merci à Olivier Collin pour son support dans le cadre de mes recherches en théorie de Morse, Matthieu Dufour et Fernand Beaudet pour m'avoir permis d'assister au séminaire de mathématiques financières et finalement merci à Steven Lu pour le support à mes débuts de recherche en théorie de Hodge. Tout le temps passé à discuter avec vous a été grandement apprécié.

Pour finir, j'aimerais remercier Noémie et Émilie, pour avoir corrigé mes innombrables fautes de grammaire, d'orthographe et de syntaxe, ce qui a permis à ce mémoire d'être beaucoup plus lisible.

TABLE DES MATIÈRES

RÉSUMÉ

Dans ce mémoire, nous étudions les mesures cohérentes de risque et leurs applications en gestion de portefeuille.

Le premier chapitre est une introduction aux mesures cohérentes de risque et aux mesures convexes de risque. Nous motivons leur introduction et présentons les principales définitions et théorèmes lié à ces concepts. Le résultat le plus important du chapitre est le théorème de représentation des mesures convexes de risque qui permet de caractériser toutes ces mesures. Le lecteur trouvera à la fin de ce chapitre, un appendice sur les principaux résultats d'analyse fonctionnelle utilisés dans ce mémoire.

Dans le deuxième chapitre, nous introduisons les exemples les plus utilisés de mesures cohérentes de risque. Après avoir rappelé la définition de valeur au risque, une mesure qui n'est pas cohérente, on présente la valeur au risque conditionnelle qui, elle, est cohérente. Nous terminons ensuite par les mesures du type *shortfall risk.*

Le troisième chapitre expose le problème principal de ce mémoire. Nous nous plaçons dans un modèle de type Black-Scholes pour modéliser le prix d'un actif risqué. En expliquant comment d'autres auteurs ont réussi à minimiser la valeur au risque d'un portefeuille composé d'une action avec une option de vente, nous montrons comment l'on peut étendre leurs résultats à la valeur au risque conditionnelle et aux mesures du type *shortfall risk.* Nous terminons ensuite en expliquant, comment, en modifiant légèrement l'hypothèse, nous pouvons démontrer l'existence d'une solution à ce problème pour n'importe quelle mesure convexe de risque. Comme nous l'avons déjà mentionné, ce problème a été résolu dans le cas de la valeur au risque. Par contre, la généralisation à la valeur au risque conditionnelle et aux mesures du type *shortfall risk* sont des résultats nouveaux.

Mots-clefs : mesure cohérente de risque, mesure convexe de risque, valeur au risque, valeur au risque conditionnelle, *shortfall risk*

INTRODUCTION

Dans ce mémoire, nous nous intéresserons à la théorie des mesures de risque et nous ferons une démonstration de son utilisation dans le cadre d'un problème en théorie de la gestion des portefeuilles. Dans les marchés financiers, la valeur des actifs transigés varie dans le temps. Si nous possédons un portefeuille contenant plusieurs actifs transigés, nous sommes intéressés à savoir de quelle manière la valeur de ce dernier varie. Il y a une forme d'incertitude reliée à cela puisque nous ne savons pas quelle sera la valeur de ces actifs dans le futur. Ainsi nous appellerons flux d'investissement d'un actif transigé sur le marché, la valeur future de cet actif moins sa valeur présente pour une période donnée. On peut donc voir le flux d'investissement comme étant le gain/perte net associé à la possession de l'actif sur une certaine période donnée. Nous modéliserons l'incertitude réliée aux flux d'investissement futur par une variable aléatoire X, ce qui nous permettra d'appliquer la théorie des probabilités. Dans ce contexte, il existe plusieurs risques : le risque de crédit, le risque lié à la volatilité, le risque de taux d'intérêt, de taux de change, opérationnel, de pays, etc. Par conséquent, que l'on ait tenté de modéliser le risque de diverses manières depuis des décennies n'est pas étonnant.

Avec ces exemples d'incertitude en tête, nous voudrions avoir une définition en termes mathématiques qui nous permettra de développer une théorie du risque. Une première tentative de définition nous a été fournie par l'écart type de la variable aléatoire associée aux flux d'un investissement. En effet, nous pourrions dire que, plus une variable aléatoire a des flux qui s'éloigne de sa moyenne, plus elle est risquée car plus il est difficile de prévoir de combien sera le retour sur l'investissement ; c'est dans cette optique que la théorie de Markowitz, ou théorie moderne du portefeuille, s'est imposée dans les années 1950. Selon cette théorie, l'investisseur cherche à maximiser l'espérance de gain tout en minimisant l'écart type du portefeuille. Il s'agissait d'un bon point de départ et cette méthode est encore très utilisée aujourd'hui. Par contre, il y a un problème conceptuel avec la variance comme définition du risque associé à une position. Nous pourrions très bien imaginer une variable aléatoire avec un très grand écart type et des flux toujours positifs, tandis qu'une autre variable aléatoire pourrait avoir des flux toujours nuls. Dans ce cas, la variable aléatoire nulle serait considérée comme moins risquée mais dans les faits, tout investisseur voudrait investir dans la première variable aléatoire. Cela correspond bien à l'intuition première de risque que nous avons donnée car nous sommes certains que la variable nulle donnera toujours un rendement nul tandis que nous ne pouvons dire quel sera le rendement de la première variable aléatoire avec certitude.

Un second concept ayant fait son apparition dans le domaine financier est celui de la durée d'une obligation. Cette mesure donne une indication de la sensibilité du prix d'une obligation par rapport aux variations des taux d'intérêt. On peut essentiellement calculer celle-ci en effectuant la première dérivée de la fonction donnant le prix de l'obligation en fonction des taux d'intérêt. La durée d'une obligation est encore couramment utilisée en théorie de gestion de portefeuille de titres à revenu fixe. Cette notion n'est pas adéquate pour la plupart des classes d'actifs car pour la plupart de ces derniers, le prix ne dépend pas que d'un facteur unique et il n'est pas toujours possible d'avoir une formule analytique du prix d'un actif en fonction de plusieurs paramètres.

De nos jours, l'interprétation du risque n'est ni liée au rendement, ni à la sensibilité du mouvement du prix, mais plutôt à l'exposition de l'institution, cette dernière étant plus intéressée à connaître la perte financière que pourrait impliquer une position. À ce sujet, le domaine actuariel a apporté une solution en proposant la valeur au risque (*value at risk*) au niveau λ, notée VAR_λ. Celle-ci mesure le montant de la perte qui ne devrait pas être dépassé avec une probabilité donnée. En d'autres mots, la VAR_λ est le capital qui est à risque dans le pire des scénarios avec un degré de confiance de $(1-\lambda)\%$. Ce qui est bien avec cette définition de risque comme capital à risque est que cela s'applique à la plupart des situations. Malgré son côté un peu simpliste, la VAR_λ n'étant en fait qu'un quantile, elle s'est rapidement imposée depuis le début des années 1990 comme mesure de risque dans l'industrie, puisqu'elle est facile à calculer. Par contre, la VAR_λ ne possède pas certaines propriétés que l'on attend d'une mesure de risque. C'est pourquoi plusieurs autres mesures de risque, qui sont plus près de notre intuition, ont été développées par la suite. Par contre, ces mesures de risque ont plus de mal à percer l'industrie car elles sont en général plus complexes à calculer et l'interprétation que l'on peut leur donner est plus difficile à comprendre. Donc, pour toutes ces raisons, la VAR_λ est encore largement utilisée en ce moment car elle peut être comprise par quelqu'un qui détient même très peu de notions de probabilités, contrairement aux nouvelles mesures de risque. Elle fut donc importée sur les marchés financiers par la banque Bankers Trust et a été popularisée par la banque JP Morgan en 1993. Elle a même été adoptée par le comité de Bâle comme mesure de risque. Ce comité est celui chargé de donner des dispositifs prudentiels destinés à mieux appréhender le risque. Plusieurs recherches, utilisant la VAR_λ, ont déjà été effectuées en gestion de portefeuille.

Dans un premier temps (chapitre I de ce mémoire), nous introduirons la théorie générale des mesures convexes de risque. Pour cela, nous donnerons les principales définitions et les notions requises d'analyse fonctionnelle pour arriver à démontrer le théorème de représentation des mesures convexes de risque. Lors de l'introduction d'une nouvelle théorie, il est très important pour le lecteur de comprendre les champs d'applications de cette théorie. Dans le premier chapitre, quelques exemples seront mis en place pour faciliter cette compréhension, mais nous attendrons le deuxième chapitre pour expliciter pleinement la plupart des exemples et en donner de nouveaux.

Dans le chapitre I, nous aurons développé la théorie des mesures de risque pour des ensembles de variables aléatoires bornées. La restriction aux variables aléatoires bornées n'est pas contraignante du point de vue économique puisque dans la réalité, le rendement des actifs que l'on retrouvent sur le marché est borné, pour la simple raison qu'il n'y a qu'un nombre fini de ressources sur la Terre. Par contre, pour faciliter la modélisation, la plupart des actifs transigés sur le marché sont modélisés par des variables aléatoires non bornées. Dans cette perspective, au chapitre II, nous expliquerons de quelle façon il est possible d'adapter la théorie du chapitre I à n'importe quel espace de variables aléatoires, pour récupérer les résultats dans un cadre plus général. Nous compléterons ensuite par des exemples de mesures convexes de risque dans L^1.

En théorie de la gestion de portefeuille, on peut supposer que le prix d'un actif X est modélisé par une variable que nous appellerons aussi X. Puisque le prix de X au fil du temps est aléatoire, une question naturelle se pose : est-ce qu'il existe un moyen de minimiser le risque associé à la variation du prix d'un actif à l'aide d'une famille d'actifs $\{Y_i\}$? On peut aussi se demander s'il est possible de trouver un élément Y_i qui minimise ce risque. Puisque dans le premier chapitre nous aurons développé une manière de quantifier le risque et donné ensuite des exemples sur la façon dont nous pouvons calculer ce risque, nous appliquerons donc la théorie développée aux problèmes précédemment mentionnés. Dans notre cas, l'actif Y sera une option de vente sur notre actif X dont le prix d'exercice est K. Nous reprendrons premièrement les résultats obtenus par Ahn *et al.* (1999) qui montrent comment minimiser la mesure de risque VAR_λ à l'aide d'options sur un portefeuille d'actions. Ensuite, nous étendrons ces résultats à une mesure de risque appelée la moyenne de la valeur au risque (CVAR_λ). Finalement, nous nous attaquerons de nouveau à ce problème en considérant de la mesure dite *shortfall risk*. Contrairement à ce qui aura été fait précédemment, nous devrons modifier le contexte de notre problème pour être en mesure de démontrer l'existence d'une solution. Par ailleurs, dans le contexte que nous nous sommes donné, l'évaluation du risque pourrait se faire à plusieurs moments dans le temps. Par exemple, le gestionnaire pourrait décider de revoir ses positions à la fin de chaque journée. Mais dans notre cas, par souci de simplicité, l'évaluation se fera dans un modèle où il n'y a qu'une seule période d'investissement.

CHAPITRE I

MESURES DE RISQUE

1.1 Les besoins de la théorie

Dans ce chapitre, nous donnerons immédiatement quelques exemples de mesures de risque dans un cadre plutôt informel. Le but étant ici de motiver les définitions. Nous reviendrons sur ces exemples de façon beaucoup plus détaillée au chapitre II.

L'évaluation du risque financier a pris beaucoup d'importance ces dernières années dans le domaine de la finance. Les marchés financiers ont adopté presque unanimement une mesure nommée la VAR_λ malgré le fait que celle-ci se soit montrée déficiente à différents points de vue. C'est en réponse à ces lacunes qu'en 1997, un groupe de chercheurs a proposé un ensemble d'axiomes que devrait satisfaire une mesure de risque pour être conforme à l'intuition.

Au début des années 1990, la valeur au risque au niveau λ (*value at risk at level* λ) noté VAR_λ a été popularisée comme mesure de risque auprès des institutions financières. Cette mesure est naturelle du point de vue d'une agence de réglementation puisque dans ce cas, le risque associé à une position X est le plus petit montant de capital qui, s'il est ajouté à la position X et investi sans risque, rend la position acceptable. La VAR_λ au niveau $\lambda \in (0,1]$ d'une variable aléatoire X notée $\mathrm{VAR}_\lambda(X)$ se définit comme étant le quantile au niveau λ de X. En d'autres termes, si F_X est la fonction de répartition de X et F_X^{-1} est un inverse généralisée de X, alors

$$\mathrm{VAR}_\lambda(X) = -F_X^{-1}(1-\lambda).$$

Nous pouvons décrire la VAR_λ comme suit : avec une probabilité de $1-\lambda$, la perte ne dépassera pas $\mathrm{VAR}_\lambda(X)$. Dans ce cas, nous voyons que $\mathrm{VAR}_\lambda(X)$ est le plus petit niveau de capital à ajouter à la position pour que la perte soit nulle avec un degré de confiance de $1-\lambda$. Malheureusement, cette mesure s'est montrée déficiente à plusieurs points de vue. Premièrement, elle ne mesure pas le risque de grandes pertes. Par exemple, supposons que X et Y aient pour distribution :

$$F_X(x) = \begin{cases} 1\%, & \text{si } x = -1000000000000000000 \\ 1\%, & \text{si } x = 0 \\ 98\% & \text{si } x = 1 \end{cases}$$

$$F_Y(y) = \begin{cases} 1\%, & \text{si } y = -1 \\ 1\%, & \text{si } y = 0 \\ 98\% & \text{si } y = 1 \end{cases}$$

Alors, nous voyons que $\mathrm{VAR}_{.01}(X) = \mathrm{VAR}_{.01}(Y) = 0$, mais la variable aléatoire X semble beaucoup plus risquée que la variable aléatoire Y. De plus, intuitivement, une mesure de risque devrait être sous-additive, c'est-à-dire faire en sorte que le risque d'une somme de position soit plus petit que la somme individuelle des risques. Or pour la VAR_λ, comme nous le verrons au chapitre III, ce n'est pas le cas. Nous pouvons interpréter cela en disant que la VAR_λ n'encourage pas la diversification, ce qui est contraire à la pratique financière. Pour palier au fait que la VAR_λ ne tient pas compte de l'amplitude que peut prendre la perte au dessus de la VAR_λ, nous avons étendu de façon naturelle cette définition en considérant plutôt la perte moyenne dépassant la VAR_λ. C'est ainsi qu'est née la *Average Value at Risk* au niveau λ notée CVAR_λ, qui est définie par

$$\mathrm{CVAR}_\lambda(X) = \frac{1}{\lambda}\int_0^\lambda \mathrm{VAR}_\gamma(X)\mathrm{d}\gamma$$

où X est une variable aléatoire. Cette mesure de risque est meilleure que la VAR_λ à plusieurs titres. Elle est invariante par addition d'une constante :

$$\mathrm{CVAR}_\lambda(X+m) = \mathrm{CVAR}_\lambda(X) - m,$$

pour tout $m \in \mathbb{R}$. Une conséquence immédiate est que

$$\mathrm{CVAR}_\lambda(X + \mathrm{CVAR}_\lambda(X)) = 0.$$

Donc, si nous qualifions une position d'acceptable lorsque $\mathrm{CVAR}_\lambda(X) \leq 0$, alors $\mathrm{CVAR}_\lambda(X)$ est bien le capital minimum requis pour que X soit acceptable. De plus il est clair que si $X \leq Y$ alors $\mathrm{CVAR}_\lambda(Y) \leq \mathrm{CVAR}_\lambda(X)$ c'est-à-dire que si, peu importe le scénario, Y a des flux plus grands que X, alors le capital à ajouter à Y est plus petit que le capital à ajouter à X pour rendre ces dernières acceptables. De plus, CVAR_λ est positivement homogène : pour tout $c > 0$, $\mathrm{CVAR}_\lambda(cX) = c\,\mathrm{CVAR}_\lambda(X)$. Mais de toutes les propriétés de la CVAR_λ, la plus importante est certainement la sous-additivité :

$$\mathrm{CVAR}_\lambda(X+Y) \leq \mathrm{CVAR}_\lambda(X) + \mathrm{CVAR}_\lambda(Y).$$

Nous pouvons interpréter cette dernière en disant que la CVAR_λ encourage la diversification.

En considérant ces quatres propriétés nous comprenons que toute bonne mesure de risque devrait les posséder. D'où la justification de développer une théorie qui donne un ensemble d'axiomes que devrait satisfaie une mesure de risque. C'est ce qu'ont fait Artzner *et al.* (1999) en proposant une théorie des mesures cohérentes de risque, dans laquelle ils exposent les axiomes auxquels devrait répondre une mesure de risque et qui sont justement les quatres propriétés dont nous avons fait état concernant la CVAR_λ. De plus, ils ont caractérisé toutes les mesures de risque ρ en terme d'ensemble de scénarios. En effet, si l'on interprète une mesure de probabilité comme un scénario, ils ont réussi à montrer que, pour toute mesure qui satisfait les propriétés précédentes, il existe un sous-ensemble $\mathscr{Q}$ de l'ensemble de toutes les mesures de probabilité sur notre espace de référence tel que

$$\rho(X) = \sup_{P \in \mathscr{Q}} E_P[-X].$$

Nous pouvons interpréter les éléments de $\mathscr{Q}$ comme des scénarios pouvant survenir.

Plusieurs personnes ont critiqué les mesures cohérentes de risque en indiquant que le risque, en général, ne grandit pas linéairement en fonction de la grandeur de la position, ce qui est le cas avec la CVAR_λ. C'est pourquoi par la suite, Follmer et Schied (2002) ont proposé par la suite, de nouveaux axiomes donnant lieu aux mesures convexes de risque. Comme nous le verrons plus tard, avec ces nouvelles mesures, le risque ne grandit plus linéairement mais plutôt de manière convexe. Nous pouvons illustrer cela par l'exemple suivant : un investisseur qui possède plusieurs millions d'actions de la compagnie IBM aura plus de difficulté à se départir complètement de cette position qu'un investisseur qui n'en possède que 100 actions. Le prix de vente de ses actions risque donc d'être plus bas.

1.2 Définitions de mesures de risque

Dans la suite du chapitre, nous donnerons tous les détails de la théorie des mesures de risque. Nous attendrons par contre le chapitre II, pour donner des exemples de mesures cohérentes de risque.

Soit $(\Omega, \Im)$ un espace mesurable où Ω est un ensemble de scénario fixé. Nous noterons par $L_B(\Omega, \Im)$ l'espace des fonctions bornées mesurables à valeurs réelles sur $(\Omega, \Im)$. Comme nous l'avons mentionné précédemment, nous sommes intéressés à quantifier l'incertitude du flux financier d'un actif sur un horizon de temps fini. Par exemple, si l'on achète 100 actions d'IBM à une date fixée, disons t_0, la valeur de ces actions à une date ultérieure T est indéterminée. Nous modéliserons cette forme d'incertitude par une variable aléatoire X. Nous dirons dans ce cas que la perte associée à la position « 100 actions d'IBM » sera de $X_0 - X$ où X_0 est la valeur des 100 actions au temps t_0. Dans notre cadre, nous interpréterons les éléments de $L_B(\Omega, \Im)$ comme dans notre exemple, c'est-à-dire qu'une variable aléatoire X bornée représente la valeur future d'un actif par rapport à sa valeur initiale. Pour $\omega \in \Omega$ et $X \in L_B(\Omega, \Im)$, $X(\omega)$ représente donc la valeur de la perte si le scénario ω s'est réalisé. Nous remarquons qu'un nombre réel négatif pour $X(\omega)$ signifie que la valeur de l'actif est plus grande que la valeur d'acquisition.

Nous pourrions raffiner notre définition. En effet, la valeur de l'argent investi au temps 0 vaut en réalité plus que ce même argent au temps T. C'est le principe de valeur du temps de l'argent. Donc, plus généralement, nous interpréterons X comme la valeur future actualisée de l'actif au temps T. Dans le cas où l'on considère une économie dans laquelle il existe un compte bancaire rapportant un rendement composé continuellement de r, alors la variable aléatoire X représentera la différence entre $X_0 e^{r(T-t_0)}$ qui est la valeur future de notre investissement initial et la valeur au temps T de notre actif. Nous dirons que X, dans ce cas, est une variable de perte, ou plus simplement, une perte.

Nous aimerions trouver une manière de déterminer si un risque sur perte associé à un actif financier est acceptable dans un sens qui reste à déterminer. Cette question de savoir si nous sommes à risque nous pousse alors à définir une manière de quantifier ce risque. Cette quantification, nous l'appellerons mesure de risque. Cette mesure ne sera pas unique. Dans certaines situations, il sera plus convenable d'utiliser une mesure de risque plutôt qu'une autre. Évidemment toutes ces mesures partageront un ensemble de propriétés qui leur sont imposées pour satisfaire à notre intuition. C'est sur ces propriétés que nous bâtirons notre théorie.

Premièrement, nous interpréterons cette mesure comme le plus petit montant de capital à investir dans un actif sans-risque à ajouter à une position pour que celle-ci devienne acceptable. Cela correspond tout à fait au cas de la VAR_λ que nous avons vue précédemment. En effet, la VAR_λ nous indique qu'avec un intervalle de confiance de $(1-\lambda)\%$, la perte ne dépassera pas un certain montant c. Supposons que l'on ait une telle mesure $\rho : L_B(\Omega, \mathfrak{I}) \longrightarrow \mathbb{R}$ et que l'on dise qu'une perte X est acceptable si $\rho(X) \leq 0$. Puisque $\rho(X)$ est le plus petit montant à ajouter pour qu'une perte X soit acceptable, nous arrivons alors au résultat que $\rho(X)$ est le plus petit montant tel que $\rho(X+\rho(X)) \leq 0$. De plus, nous pouvons introduire un ordre partiel sur les positions en posant que $X \leq Y$ si pour tout $X(\omega) \leq Y(\omega)$ pour tout $\omega \in \Omega$. Nous voudrions aussi que ρ soit monotone pour la relation $X \leq Y$ si l'on munit $\mathbb{R}$ de la relation $\geq$. Nous voudrions également que cette mesure encourage la diversification.

Définition 1. Une **mesure de risque** sur $L_B(\Omega, \mathfrak{I})$ est une application $\rho : L_B(\Omega, \mathfrak{I}) \longrightarrow \mathbb{R}$ telle que :

1. (Invariance par translation) Pour tout $L \in L_B(\Omega, \mathfrak{I})$ et $\ell \in \mathbb{R}$, nous avons
$$\rho(L+\ell) = \rho(L) - \ell.$$
2. (Monotonicité) Si $X \leq Y$, alors $\rho(Y) \leq \rho(X)$.

Nous dirons que cette mesure est **convexe** si :

3. (Convexité) Pour tout $\lambda \in [0,1]$ et $X, Y \in L_B(\Omega, \mathfrak{I})$ nous avons
$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y).$$

Nous dirons qu'elle est **cohérente** si :

4. (Positivité homogène) Pour tout $\lambda \geq 0$ et $X \in L_B(\Omega, \Im)$ nous avons $\rho(\lambda X) = \lambda \rho(X)$.
5. (Sous-additivité) Pour tout $X, Y \in L_B(\Omega, \Im)$ nous avons $\rho(X+Y) \leq \rho(X) + \rho(Y)$.

Le deuxième critère est intuitivement évident du point de vue économique. En effet, en se rappelant qu'une variable aléatoire X représente la perte par rapport à l'investissement dans un actif sans risque, il est clair que si une variable a une perte toujours moins grande qu'un autre investissement Y, alors ce dernier est nécessairement moins risqué. Le premier axiome est quant à lui un peu moins intuitif. Ainsi si l'on veut interpréter la mesure de risque comme un capital requis pour rendre une position acceptable, par exemple en disant qu'une position est acceptable si $\rho(X) \leq 0$, alors cet axiome rend l'interprétation de l'acceptabilité concrète puisque nous obtenons dans ce cas $\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0$. Nous remarquons alors que $\rho(X)$ est le plus petit montant de capital à ajouter pour rendre notre position acceptable. Nous verrons un peu plus loin la relation qu'il est possible d'établir entre ces deux concepts : l'acceptabilité et les mesures de risque.

Nous avons mentionné plus tôt, lors de notre brève introduction historique que les axiomes de cohérence ont précédé les axiomes de convexité. Rien d'étonnant à cela puisqu'il n'est pas rare, en mathématiques, que nous faisions une première définition d'un concept avant de réaliser que celui-ci peut-être vue dans un cadre beaucoup plus général. Toutefois, ce n'est pas de cette manière qu'est survenue la définition de mesure convexe. Tel que mentionné précédemment, celle-ci fut plutôt introduite parce que les axiomes de cohérence étaient trop restrictifs. En effet, dans plusieurs situations, le risque ne grandit pas de manière linéaire lorsque l'on augmente la grandeur de la position car plus la position est grande, plus il peut être difficile de la vendre en entier. C'est ce que l'on appelle un risque lié à la liquidité. Dans le cas d'une mesure convexe ρ, si $\rho(0) = 0$ nous obtenons alors que, puisque ρ est convexe,

$$\rho(\lambda X) = \rho(\lambda X + (1-\lambda)0) \leq \lambda \rho(X)$$

d'où

$$\frac{1}{\lambda}\rho(\lambda X) \leq \rho(X).$$

Cette dernière équation peut s'interpréter comme suit : le montant à ajouter à la position λX pour rendre cette dernière acceptable est proportionnellement plus petit que le montant qu'il faut ajouter à la position X pour rendre ce dernier acceptable.

Remarque 1. Il est clair qu'une mesure cohérente est convexe. La réciproque est fausse. La cohérence est la notion qui se rapproche le plus de notre intuition. C'est pourquoi cette définition est apparue en premier historiquement. On s'est aperçu ensuite que les principaux résultats de la théorie des mesures cohérentes de risque pouvaient être étendus de manière similaire aux mesures convexes.

Remarque 2. Un lecteur avisé pourrait remarquer que la définition de mesure convexe de risque est analogue à celle d'une fonction d'utilité. C'est une des raisons pour laquelle le sujet s'est développé aussi rapidement, plusieurs énoncés étant duaux à ceux de la théorie de l'utilité.

1.3 Ensemble de positions acceptables

Dans cette section, nous montrerons comment construire, avec certains sous-ensembles de $L_B(\Omega,\Im)$ (dits ensembles de positions acceptables), une mesure de risque convexe et de quelle manière à partir d'une mesure convexe, il est possible de construire un ensemble de positions acceptables. Nous verrons alors qu'il y a équivalence entre ces deux concepts. Cette construction nous servira à définir des mesures de risque que l'on appellera « *shortfall risk* ».

Définition 2. Nous dirons qu'un sous-ensemble $A \subset L_B(\Omega,\Im)$ est un **ensemble de positions acceptables** si :

1. A est convexe et non vide ;
2. si $X \in A$ et $Y \in L_B(\Omega,\Im)$ sont tels que $Y \geq X$ alors $Y \in A$.

Si l'on pose $\rho_A : L_B(\Omega,\Im) \longrightarrow \mathbb{R}$ par $\rho_A(X) = \inf\{\, m \in \mathbb{R} \mid m+X \in A \,\}$, alors ρ, ainsi défini, est une mesure convexe de risque. On a premièrement que cette fonction est bien définie car $\{\, m \in \mathbb{R} \mid m+X \in A\}$ est non vide. En effet, puisque A est non vide, il existe une variable aléatoire bornée Y dans A. On a donc que $\sup(Y) < \infty$ et $\inf(X) > -\infty$, alors $-\inf(X)+X \geq 0$ et $-\inf(X)+\sup(Y)+X \geq Y$ qui est équivalent à $-\inf(X)+\sup(Y)+X \in A$, d'où l'on tire

$-\inf(X)+\sup(Y) \in \{m \in \mathbb{R} \mid m+X \in A\}$. Nous avons bien la monotonicité car si $X \in A$ et $Y \in L_B(\Omega, \Im)$ avec $Y \geq X$, nous avons que $m+Y \in A$ lorsque $m+X \in A$. Par conséquent,

$$\{m \in \mathbb{R} \mid m+X \in A\} \subseteq \{m \in \mathbb{R} \mid m+Y \in A\}$$

et nous avons $\rho(X) \geq \rho(Y)$. L'invariance par translation s'obtient en observant que

$$\inf\{m \in \mathbb{R} \mid m+(X+r) \in A\} = \inf\{m \in \mathbb{R} \mid m+X \in A\} - r.$$

Finalement, on voit que la mesure est convexe puisque si $\rho(X) = x$ et $\rho(Y) = y$, alors

$$\lambda(x+X)+(1-\lambda)(y+Y) \in A$$

puisque A est convexe et $x+X, y+Y \in A$. Donc

$$\begin{aligned} \rho(\lambda X+(1-\lambda)Y) &= \inf\{m \in \mathbb{R} \mid m+\lambda X+(1-\lambda)Y \in A\} \\ &\leq \lambda\rho(X)+(1-\lambda)\rho(Y). \end{aligned}$$

Dans ce qui suit, nous construirons à l'aide d'une mesure de risque ρ un ensemble de positions acceptables. Nous procéderons comme suit : un placement L sera dit acceptable si $\rho(L) \leq 0$. Nous remarquons alors que le premier axiome vient expliquer cette terminologie. En effet, que ρ soit invariante par translation nous donne que $\rho(L+\rho(L)) = \rho(L)-\rho(L) = 0$. De plus, il est clair que c'est le plus petit montant à ajouter pour rendre la position acceptable.

Si $\rho : L_B(\Omega, \Im) \longrightarrow \mathbb{R}$ est une mesure convexe de risque, nous posons alors

$$A = A_\rho = \{X \in L_B(\Omega, \Im) \mid \rho(X) \leq 0\}$$

Nous obtenons alors la proposition suivante, qui nous donne une bijection entre les ensembles de positions acceptables et les mesures convexes de risque.

Théorème 1.1. *Sous les hypothèses énoncées précédemment, nous obtenons que $\rho_A = \rho$. De plus, A_ρ est un ensemble de positions acceptables et il satisfait la propriété suivante : si $X \in A$ et $Y \in L_B(\Omega, \Im)$ alors $\{\lambda \in [0,1] \mid \lambda X+(1-\lambda)Y \in A\}$ est fermé dans $[0,1]$*

DÉMONSTRATION. Nous voyons par l'invariance par translation de ρ que $\rho_A(X) = \inf\{m \in \mathbb{R} | m + X \in A\} = \inf\{m \in \mathbb{R} | \rho(m+X) \leq 0\} = \inf\{m \in R | \rho(X) \leq m\} = \rho(X)$. Pour l'affirmation suivante, il est clair que A est non vide car $X + \rho(X) \in A$. De plus A est convexe car si $X, Y \in A$ et $\lambda \in [0,1]$ alors $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y) \leq \lambda 0 + (1-\lambda)0$ car ρ est convexe. Si $X \in A$ et $Y \in L_B(\Omega, \mathfrak{I})$ est tel que $Y \geq X$ alors $\rho(Y) \leq \rho(X) \leq 0$ car X est dans A. Pour la dernière propriété, nous remarquons que la fonctionnelle

$$\begin{aligned} \rho_\lambda : \quad [0,1] &\longrightarrow \mathbb{R} \\ \lambda &\mapsto \rho(\lambda X + (1-\lambda)Y) \end{aligned}$$

est convexe et prend seulement des valeurs finies. Elle est donc continue et on a alors que $\rho_\lambda^{-1}(-\infty, 0]$ est fermé. □

Remarque 3. Cette dernière proposition nous donne donc une bijection entre les ensembles de positions acceptables satisfaisant la propriété du dernier théorème et les mesures convexes de risque.

Exemple 1. La valeur au risque au niveau $\lambda > 0$, soit

$$\text{VAR}_\lambda(X) = \inf\{m \mid P(X + m < 0) \leq \lambda\}$$

n'est pas une mesure convexe de risque. En effet, l'ensemble de positions acceptables engendré par VAR_λ n'est pas convexe. Considérons les deux investissements suivants : supposons qu'un investissement X rapporte $-1\$$ si l'état du monde est ω et $0\$$ sinon. L'investissement Y quant à lui rapporte $-1\$$ si 'état du monde est ω' survient $0\$$ sinon. Supposons que les états ω et ω' surviennent chacun avec une probabilité de 4%. Dans ce cas, les deux positions, prises individuellement, ont des VAR_λ au niveau $\lambda = 5\%$ de 0, et sont donc toutes deux acceptables. Par contre, le portefeuille composé d'un demi actif X et d'un demi actif Y a une VAR_λ au niveau $\lambda = 5\%$ de $-1/2\$$. En effet la probabilité de perdre $1/2\$$ est de 8%. Par conséquent, la position $1/2X + 1/2Y$ n'est pas acceptable ce qui prouve que l'ensemble des positions acceptables engendré par la VAR_λ n'est pas convexe. Il est à remarquer que cela démontre en plus que VAR_λ n'est pas une mesure cohérente de risque.

1.4 Théorème de représentation

Comme nous le constaterons dans les prochains chapitres, le théorème de représentation que nous allons démontrer a plusieurs applications théoriques. Par contre, il ne nous permet généralement pas de calculer facilement le risque associé à une position. Pour le lecteur n'ayant pas en tête les résultats d'analyse fonctionnelle requis pour la démonstration du théorème de représentation de cette section, les définitions et résultats de ce dernier sujet sont donnés et démontrés dans l'appendice à la fin du présent chapitre.

On peut interpréter une mesure de probabilité Q sur un ensemble Ω comme étant un scénario. En effet, la mesure Q pondère les probabilités avec lesquelles un état de la nature $\omega \in \Omega$ se réalise. Dans ce contexte, le théorème de représentation des mesures de risque stipule que si ρ est une mesure convexe de risque, alors il existe une fonction de pénalité α sur l'ensemble des mesures de probabilités de Ω telle que la mesure ρ peut être interprétée comme la pire perte moyenne sur tous les scénarios possibles, ces pertes étant pondérées par la fonction de pénalité α. Nous commencerons la démonstration en rappelant l'intégrale des fonctions par rapport à une mesure à variation finie. Nous verrons que l'espace de ces mesures est en bijection avec le dual topologique de l'espace des fonctions bornées mesurables sur Ω. Cette dernière remarque nous permettra par la suite de démontrer le théorème de représentation des mesures de risque.

Définition 3. Une **mesure additive** sur un espace mesurable $(\Omega, \mathfrak{I})$ est une fonction $\mu : \mathfrak{I} \to \mathbb{R}$ telle que :

1. $\mu(\emptyset) = 0$;
2. si $(A_i)_{i \in I}$ est une famille finie d'éléments de $\mathfrak{I}$ disjoints deux à deux, alors

$$\mu\Big(\bigcup_{i \in I} A_i\Big) = \sum_{i \in I} \mu(A_i).$$

Nous noterons par $M_{1,f}$ l'ensemble des mesures additives telles que $\mu(\Omega) = 1$. La variation totale d'une mesure additive est définie par

$$||\mu|| := \sup\Big\{ \sum_{i \in I} |\mu(A_i)| \mid I \text{ est fini, } A_i \in \mathfrak{I},\ A_i \cap A_j = \emptyset, \text{ si } i \neq j \Big\}$$

L'espace de toutes les mesures additives à variation totale finie est noté $ba(\Omega, \Im)$ ou simplement ba si le contexte est clair.

Proposition 1.2. *Soit $(\Omega, \Im)$ un espace mesurable et soit $L_B(\Omega, \Im)$ l'ensemble des fonctions mesurables bornées de $(\Omega, \Im)$. Alors toute fonctionnelle linéaire continue l sur $L_B(\Omega, \Im)$ (pour la topologie de la norme du supremum) définit une mesure additive à variation totale finie sur $(\Omega, \Im)$. Nous avons alors une correspondance biunivoque entre les formes linéaires continues sur $L_B(\Omega, \Im)$ et l'ensemble ba des mesures additives à variation totale finie sur $(\Omega, \Im)$.*

DÉMONSTRATION. Il est clair qu'un élément de ba nous donne une forme linéaire sur $L_B(\Omega, \Im)$ par l'intégrale. Supposons maintenant que nous ayons une forme linéaire continue l sur $L_B(\Omega, \Im)$ et posons

$$\mu(A) = l(\mathbf{1}_A), \quad A \in \Im$$

où $\mathbf{1}_A$ est la fonction indicatrice de A qui vaut 1 si $x \in A$ et 0 sinon. Puisque l est continue, elle est bornée, c'est-à-dire, $||l|| < L$ et par conséquent la variation totale de μ est bornée par L, c'est-à-dire est un élément de ba. Il suffit maintenant de constater que l'intégrale définie par cette forme linéaire coïncide avec cette forme linéaire. Cela est clair pour l'ensemble des fonctions étagées. Puisque l'ensemble des fonctions étagées est dense par passage à la limite, nous obtenons le résultat voulu. $\square$

Avec ce qui précède, nous sommes en mesure de montrer le théorème de représentation des mesures convexes de risque.

Théorème 1.3. *Une fonctionnelle $\rho : L_B(\Omega, \Im) \to \mathbb{R}$ est une mesure convexe de risque si, et seulement si, il existe une « fonction de pénalité »*

$$\alpha_{min} : M_{1,f} \to (-\infty, \infty]$$

telle que

$$\rho(Z) = \sup_{Q \in M_{1,f}} (E_Q[-Z] - \alpha_{min}(Q)).$$

La fonction α_{min} est donnée par $\alpha_{min}(Q) = \sup_{X \in A_\rho} E_Q[-X]$. De plus, α_{min} est minimale au sens où, si α est une autre fonction de pénalité qui représente ρ, alors $\alpha(Q) \geq \alpha_{min}(Q)$, pour tout $Q \in M_{1,f}$.

DÉMONSTRATION. Premièrement, comme mentionné précédemment, il est possible d'identifier le dual topologique $L_B(\Omega,\Im)^*$ de $L_B(\Omega,\Im)$ avec ba. De plus, ρ est semi-continue inférieurement pour la topologie faible $\sigma(L_B(\Omega,\Im),L_B(\Omega,\Im)^*)$ car l'ensemble $\rho^{-1}((-\infty,c])$ est convexe et fermé sous la topologie forte (la topologie de la norme sup) de $L_B(\Omega,\Im)$. Nous pouvons donc en conclure que $M_{1,f}$ est fermé sous la topologie faible de $\sigma(L_B(\Omega,\Im),L_B(\Omega,\Im)^*)$. Donc, selon le théorème de dualité pour la transformée de Fenchel-Legendre

$$\rho^{**} = \rho$$

c'est-à-dire que

$$\rho(X) = \sup_{l\in ba}(l(X) - \rho^*(l)),$$

ce qui nous donne déjà une partie de la représentation. Puisque ρ est invariante par translation, nous avons que si $\rho^*(l) < \infty$, alors

$$\begin{aligned}
\rho^*(l) &= \sup_{X\in L_B(\Omega,\Im)} \Big(l(X+1) - \rho(X+1)\Big) \\
&= \sup_{X\in L_B(\Omega,\Im)} \Big(l(X) + l(1) - \rho(X) + 1\Big) \\
&= \rho^*(l) + l(1) + 1
\end{aligned}$$

et donc $l(1) + 1 = 0 \Rightarrow l(1) = -1$. De plus, puisque ρ est positivement homogène, nous obtenons que si $\rho^*(l) < \infty$, alors pour tout $\lambda > 0$ et $Y \geq 0$,

$$0 \geq \rho(\lambda Y) = \sup_{l\in ba}\big\{l(Y) - \rho^*(l)\big\} \geq \lambda l(Y) - \rho^*(l).$$

On en conclut que $\rho^*(l) \geq \lambda l(Y)$ pour tout $\lambda > 0$. Cette dernière affirmation est vraie si, et seulement si, $l(Y) \leq 0$. Puisque $-l(1) = 1$ et $l(Y) \leq 0$ si $Y \geq 0$, nous voyons que $-l$ définit en réalité un élément de $M_{1,f}$. Nous identifions alors $-l$ avec $Q \in M_{1,f}$, de sorte que la représentation se réduit à

$$\rho(Z) = \sup_{Q\in M_{1,f}} (E_Q[-Z] - \rho^*(Q)).$$

Il suffit maintenant de constater que

$$\rho^*(Q) = \sup_{X\in A_\rho} E_Q[-X].$$

Mais, pour tout Q, on a

$$\begin{aligned}\rho^*(Q) &= \sup_{X\in L_B(\Omega,\mathfrak{Z})} \left(E_Q[-X]-\rho(X)\right)\\ &\geq \sup_{X\in A_\rho} E_Q[-X]\\ &= \alpha_{min}(Q).\end{aligned}$$

Soit $X' = X+\rho(X)$, nous avons alors que $\alpha_{min}(Q) \geq E_Q[-X'] = E_Q[X]-\rho(X)$. Ce qui nous montre que $\alpha_{min}(Q) \geq \rho^*(Q)$, c'est-à-dire $\alpha_{min}(Q) = \rho^*(Q)$. □

À partir du dernier théorème et de son corollaire ci-dessous, nous déduisons le théorème de représentation des mesures cohérentes de risque tel que présenté dans l'article original de Artzner *et al.* (1999).

Corollaire 1.4. *Si ρ est une mesure cohérente, alors toute fonction de pénalité α_{min} est telle qu'elle prend les valeurs 0 ou ∞.*

DÉMONSTRATION. En effet, si ρ est cohérente, il est clair que A_ρ est un cône : $X \in A_\rho$, alors $\lambda\rho(X) = \rho(\lambda X) \leq 0$ pour tout $\lambda > 0$. Donc

$$\begin{aligned}\alpha_{min}(Q) &= \sup_{X\in A_\rho} E_Q[-X]\\ &= \sup_{\lambda X\in A_\rho} E_Q[-\lambda X]\\ &= \lambda\,\alpha_{min}(Q).\end{aligned} \tag{1.1}$$

Cette dernière égalité est vraie si et seulement si α_{min} ne prend que les valeurs 0 ou $+\infty$. □

On constate ainsi que si ρ est une mesure cohérente de risque, alors

$$\rho(X) = \sup_{P\in\Upsilon} E_P[-X]$$

où Υ est l'ensemble des mesures Q telles que $\alpha_{min}(Q) = 0$.

1.5 Appendice d'analyse fonctionnelle

Dans cette section, nous suivrons essentiellement les définitions de Bourbaki (1971) et de Grothendieck (1973). Dans la présente, K désigne un corps commutatif muni d'une topologie compatible avec les opérations sur le corps. Nous donnerons toutes les définitions dans leur cadre le plus général possible. Par simplicité, le lecteur peut seulement remplacer le corps des scalaires K par le corps des réels. Dans un souci d'allégement des démonstrations, dans la plupart des théorèmes, nous avons remplacé le corps K par le corps des réels. Le lecteur pourra trouver des énoncés plus généraux dans les références susmentionnées.

1.5.1 Généralités

Définition 4. Un espace vectoriel topologique (EVT) est la donnée d'un espace vectoriel $(E,K,+,\cdot,0)$ et d'une topologie τ sur E. Ces données satisfaisant aux axiomes suivants :

1. la multiplication par un scalaire est continue par rapport à la topologie produit sur $K \times E$.
2. l'addition est continue par rapport à la topologie sur $E \times E$.

Nous rappellerons brièvement qu'une structure uniforme sur un espace E est un filtre Φ sur $E \times E$ tel que

1. Φ contient la diagonale de $E \times E$
2. si U et V sont dans Φ, alors l'ensemble $U \circ V = \{ (x,y) \in E \times E | \exists (x,z) \in U, (z,y) \in V \} \in \Phi$
3. si $U \in \Phi$, alors le symétrique de U noté U^{-1} est dans Φ.

Si E est un EVT, nous pouvons introduire sur E une structure d'espace uniforme de manière naturelle. En effet, pour tout voisinage U de 0, posons $U_f = \{ (x,y) \in E \times E \mid x - y \in U \}$. Nous voyons alors que U_f contient la diagonale de $E \times E$ pour tout voisinage U de 0. Lorsque U parcourt les voisinages de 0, nous remarquons qu'une intersection quelconque de ces ensembles est non vide car ils contiennent tous la diagonale de E. Cette intersection est clairement égale à $(\bigcap U)_f$, car si $(x,y) \in \bigcap U_f$, alors $x - y$ est dans chacun des U de notre intersection. Réciproquement, si $x - y$ est dans l'intersection des U, il est aussi dans chacun des U_f. De plus, le symétrique de U_f est clairement égal à U_f. Enfin, puisque l'addition est continue, pour chaque voisinage V de 0 il existe un voisinage W tel que $W + W \subseteq V$. De plus, si $(x,y), (y,z) \in W_f$, alors $(x,z) \in W_f \circ W_f \subseteq (W + W)_f \subseteq V_f$. Ce qui démontre bien que les U_f forment une base d'entourage pour une structure uniforme sur E. Si V est un entourage de E, nous poserons pour tout sous-ensemble A de E : $V(A) = \{x \in E \mid (a,x) \in V,\ a \in A\}$. Ainsi lorsque nous mentionnerons la structure uniforme sur un EVT E, nous ferons allusion à la structure uniforme composée des U_f.

Définition 5. Soit F un sous-ensemble de E, un espace vectoriel topologique. Si F est un sous-espace vectoriel de E et que l'on munit F de la topologie induite par E, alors nous dirons que F est un sous-espace de l'espace vectoriel topologique E.

1.5.2 Théorèmes de séparation

Définition 6. Soit $(E,\mathbb{R},+,\cdot,0)$ un espace vectoriel. Un sous-ensemble A de E est dit convexe si

$$\forall \lambda \in [0,1],\ x,y \in A, \quad \lambda x + (1-\lambda)y \in A,$$

c'est-à-dire si le segment entre x et y est complètement inclus dans A.

Définition 7. Un sous-ensemble H d'un espace vectoriel E est appelé un hyperplan s'il existe une forme linéaire l sur E et un scalaire α tel que

$$H = \{x \in E \mid l(x) = \alpha\}$$

c'est-à-dire qu'un hyperplan est une variété linéaire affine de codimension 1.

Proposition 1.5. *Si E est un EVT et F un sous-espace, alors l'adhérence de F est un sous-espace.*

DÉMONSTRATION. Posons $\overline{F}$ comme étant l'adhérence de F. Puisque la multiplication par un scalaire non nul est un automorphisme de E, étant linéaire, continue et inversible, nous obtenons que pour tout $\lambda \in \mathbb{R}^*$, $\lambda\overline{F} \supseteq F$ est un fermé qui contient F, donc qui contient $\overline{F}$ car $\overline{F}$ est le plus petit fermé contenant F. Nous pouvons donc déduire que $F \subseteq \overline{F} \subseteq \frac{1}{\lambda}\overline{F}$ pour tout λ non nul, et par conséquent $\lambda F \subseteq \lambda\overline{F} \subseteq \overline{F}$. Donc $\lambda\overline{F} = \overline{F}$ puisque $\overline{F}$ est le plus petit fermé contenant F. L'application + applique $F \times F$ dans F. Puisque cette application est continue, nous avons que $+(\overline{F \times F}) = +(\overline{F} \times \overline{F}) \subseteq \overline{+(F \times F)} = \overline{F}$. D'où $\overline{F}$ est un sous-espace. $\square$

Corollaire 1.6. *Soit E un EVT, un hyperplan H est soit dense, soit fermé.*

DÉMONSTRATION. Sans perte de généralité, en translatant, nous pouvons toujours supposer que l'hyperplan contient l'origine et qu'il est donc un sous-espace de E. Par la proposition précédente, nous savons que l'adhérence de H est un sous-espace de E. Puisqu'il n'y a aucun sous-espace compris strictement entre H et E puisque H est de codimension 1, nous obtenons $\overline{H} = H$ ou $\overline{H} = E$. □

Un des premiers théorèmes importants en analyse fonctionnelle est celui de Hahn-Banach, ou le théorème de l'hyperplan séparant. Il indique essentiellement que pour toute variété linéaire V et ouvert U non vide de E ne rencontrant pas V, il existe un hyperplan V ne rencontrant pas U.

Théorème 1.7. *Soit E un espace vectoriel topologique sur le corps des réels ou des complexes et U un sous-ensemble convexe, ouvert et non vide de E. Supposons de plus que V soit une variété linéaire de E ne rencontrant pas U. Alors il existe un hyperplan fermé contenant V et ne rencontrant pas U.*

DÉMONSTRATION. Supposons que l'origine est contenue dans V (sans perte de généralité, nous pouvons translater sinon), nous pouvons donc supposer que V est un sous-espace vectoriel de E. De plus, nous pouvons nous restreindre au cas où les scalaires sont réels car si W est un hyperplan réel satisfaisant la condition, alors $W \cap iW$ est un hyperplan complexe satisfaisant lui aussi la condition. Soit M l'ensemble des sous-espaces de E contenant V et ne rencontrant pas U. C'est un ensemble ordonné inductif, donc par le lemme de Zorn, il existe un élément maximal W. Montrons que la codimension de W est 1. Supposons que W est de codimension plus grande que 2, alors l'image canonique U' de U dans E/W est un ouvert non vide ne contenant pas 0. Considérons le cône C engendré par U', c'est-à-dire

$$C = \bigcup_{\lambda > 0} \lambda U',$$

qui est ouvert, convexe et clairement pas égal au complémentaire de $\{0\}$ car $\{0\}^c$ n'est pas convexe. Puisque $\{0\}^c$ est connexe, il existe donc un point x, dans l'adhérence de C, qui est non nul. Puisque C est un cône, pour tout $\lambda > 0$, λx est aussi dans l'adhérence de C. De plus, si $\lambda \leq 0$, et λx est dans C, alors le segment $(\lambda x, x)$ moins x est compris dans C, ce qui contredit le fait que 0 ne soit pas dans C. Donc le sous-espace R engendré par x dans E/W ne rencontre pas C, et par le fait même ne rencontre pas U. Nous obtenons alors que l'image inverse de R par la projection canonique de E/W est un sous-espace qui contient W et ne rencontre pas U, ce qui contredit la maximalité de W. Donc W est de codimension 1. Nous concluons enfin que W est fermé car $U \subset W^c$, i.é. le complémentaire de W contient un point intérieur, donc W n'est pas dense et par la proposition précédente, est donc fermé. □

Définition 8. Soit E un EVT. Nous disons que E est un espace vectoriel topologique localement convexe (EVTLC) si pour chaque point x de E, il existe un système fondamental de voisinage convexe de x.

On remarque qu'un hyperplan H donné par une forme linéaire l sépare l'espace en deux parties ouvertes données par les équations $l(x) < \alpha$ et $l(x) > \alpha$ ayant pour frontière V. La fermeture de ces deux sous-ensembles est appellée les demi-espaces définis par la variété V.

Proposition 1.8. *Supposons qu'un sous-ensemble B d'un espace vectoriel topologique E sur un corps K est compact et disjoint d'un ensemble fermé* A, *alors il existe un ouvert U non vide tel que*

$$(A+U) \cap (B+U) = \emptyset.$$

DÉMONSTRATION. Il est équivalent de dire qu'il existe un entourage V de la structure uniforme sur E tel que $V(A) \cap V(B) = \emptyset$. On remarque premièrement que $V(A)$ est le translaté de A par l'ouvert $-V$. En effet, si $x \in V(A)$, c'est donc dire qu'il existe $a \in A$ tel que $x - a = y \in V$, donc $a - y = -x$, c'est-à-dire que $V(A)$ est le translaté de A par l'ouvert $-V$. Supposons qu'il n'existe pas d'entourage V, tel que $V(A) \cap V(B) = \emptyset$. C'est-à-dire qu'il existe $x \in E$ tel que $x - a = x - b \in V$ pour un certain $a \in A$ et $b \in B$. Puisque V est symétrique, on a que $a - b \in V^2$, ce qui nous donne que $a \in V^2(B)$. Donc, les $A \cap V^2(B)$ sont non vides et forment un filtre sur A qui doit avoir un point adhérant x_0 car A est fermé. Soit V encore un entourage symétrique de E. Nous obtenons alors $V(x_0) \cap V^2(B) \neq \emptyset$, c'est-à-dire qu'il existe $x \in E$ tel que $(x, x_0) \in V$ et

$(x,b) \in V^2$ pour un certain $b \in B$. Or $(x,b) \in V^2$ revient à dire qu'il existe des couples (x,y) et $(y,b) \in V$ pour un certain y et donc nous obtenons que $V^3(x_0) \neq \emptyset$ pour tout entourage symétrique de V. Donc $x_0 \in B$ puisque B est fermé, mais cela contredit le fait que A ne rencontre pas B. Il est à noter ici, que nous avons utilisé le fait que les entourages symétriques forment une base des voisinages de x_0. Cela se démontre comme suit : si V est un ouvert contenant x_0, alors $V_f \cap V_f^{-1}$ est un entourage symétrique de E tel que $V_f \cap V_f^{-1}(x_0) \subseteq V$ □

Si A et B sont deux sous-ensembles de E, nous dirons qu'un hyperplan H sépare A et B si ceux-ci se retrouvent dans deux demi-espaces différents engendrés par H. De plus, nous dirons que H sépare strictement A et B si H sépare ces derniers et s'ils sont compris dans l'intérieur des deux demi-espaces.

Proposition 1.9. *Soit E un EVT non trivial sur le corps des réels, soient A et B deux sous-ensembles convexes disjoints de E.*

1. *Supposons que A ou B soient ouverts, alors il existe une variété linéaire affine par l'équation $l(x) = \alpha$ qui sépare A et B.*
2. *Si A est fermé, B compact et E localement convexe, alors il existe une variété linéaire affine qui sépare strictement A et B.*

DÉMONSTRATION. 1) Puisque A ou B sont ouverts, nous obtenons que $A - B$ est ouvert. Soit x un point de $(A-B)^c$, alors par le théorème de Hanh-Banach, il existe une variété linéaire affine W de codimension 1 qui contient x et qui ne rencontre pas $A - B$. Supposons que W soit donné par la forme linéaire l sur E. Nous voulons trouver α tel que la variété linéaire affine donnée par l et α sépare A et B. Puisque l'image d'un convexe par une application linéaire est nécessairement convexe, nous obtenons que $l(A)$ et $l(B)$ sont deux sous-ensembles convexes de $\mathbb{R}$, i.e. des intervalles. Nous savons que $l(A)$ ne rencontre pas $l(B)$ par construction car l ne s'annule pas sur $A - B$. Donc il existe un α entre les intervalles $l(A)$ et $l(B)$ et nous obtenons alors que l'hyperplan engendré par l et α sépare A et B.
2) Par la proposition 1.8, nous savons qu'il existe un voisinage U de 0 tel que $A+U$ et $B+U$ ne se rencontrent pas. De plus, nous pouvons choisir U convexe puisque notre espace est localement convexe. Nous obtenons alors que $A+U$ et $B+U$ sont des ouverts convexes

auxquels nous pouvons appliquer la partie 1. Puisque les deux parties sont ouvertes, nous obtenons que notre hyperplan les sépare strictement. □

De ce dernier point, nous déduisons le corrolaire suivant.

Corollaire 1.10. *Soit A un ensemble convexe d'un EVTLC, alors A est fermé s'il est l'intersection d'une famille de demi-espaces fermés.*

DÉMONSTRATION. En effet, pour chaque point x dans le complémentaire de A, nous pouvons appliquer la partie 2 de la proposition précédente et nous obtenons un demi-espace fermé H_x qui contient A et ne contient pas x. Il est alors clair que $A = \bigcap_{x \in CA} H_x$. □

1.5.3 Système dual, topologie faible

Dans cette section, à moins d'un avis contraire, tous les EVT seront sur un corps commutatif K. Soit E et E' deux EVT et $u : E \times E' \longrightarrow K$ une forme bilinéaire. Nous dirons qu'un système dual est la donnée d'une paire d'espaces vectoriels (E, E') et d'une forme bilinéaire sur leur produit. Nous dirons que ce système dual est séparé en E' si l'application qui à $x' \in E'$ fait correspondre la forme linéaire $x \mapsto u(x, x')$, est injective. Nous obtenons de telle manière l'existence d'une bijection naturelle entre les sous-ensembles du dual de E et les système duaux (E, E') séparés en E'.

Définition 9. Soit $((E, E'), u)$ un système dual, nous définissons la topologie faible sur E relativement au système $((E, E'), u)$, noté $\sigma(E, E', u)$ ou simplement $\sigma(E, E')$ quand le contexte est clair, comme étant la topologie la moins fine qui rend continues les fonctionnelles linéaires $x \mapsto u(x, x')$ pour $x' \in E'$.

Théorème 1.11. *Soit E un EVTLC sur le corps des réels, E' son dual et soit A un sous-ensemble convexe de E. Pour que A soit fermé, il faut et il suffit que A soit fermé dans la topologie faible $\sigma(E, E')$.*

DÉMONSTRATION. Supposons que A soit fermé, alors A est l'intersection de demi-espaces fermés. Un tel demi-espace est donné par une fonctionnelle l et un réel α et il est l'intersection des $l^{-1}((\infty,\alpha])$ qui sont fermés dans $\sigma(E,E')$. Réciproquement, supposons que A soit fermé dans $\sigma(E,E')$. Puisque $\sigma(E,E')$ est la plus petite topologie rendant continnues les éléments de E', les fermés de cette topologie doivent nécessairement être des fermés de la topologie initiale puisque de telles fonctionnelles linéaires sont déjà continues pour la topologie initiale. $\square$

Le prochain théorème est présenté sans démonstration. Le lecteur pourra en trouver une dans la plupart des livres d'analyse fonctionnelle.

Théorème 1.12. *Soit E un espace de Banach et E' le dual de E. Alors la boule unité dans E' est faiblement compacte.*

1.5.4 Transformée de Fenchel-Legendre

Dans cette section, E désignera un EVTLC. Nous y ferons l'étude de la transformée de Fenchel-Legendre. Cette transformation est l'outil principal de la démonstration du théorème de représentation des mesures de risque.

Définition 10. Soit $f : E \to \mathbb{R} \cup \{\infty\}$. Alors la transformée de Fenchel-Legendre de f, noté f^* est la fonction sur E' définie par

$$f^*(l) = \sup_{x \in E}(l(x) - f(x)), \quad l \in E'.$$

Si f n'est pas identiquement égale à ∞, alors f^* n'est pas identiquement égale à ∞ et est convexe. La transformée de Fenchel-Legendre associe à une fonction f sur E une fonction convexe f^* sur E'. Il est donc clair que si f n'est pas convexe, alors $f^{**} \neq f$. Par contre, si f est convexe et semi-continue inférieurement relativement à la topologie faible $\sigma(E,E')$, alors nous avons :

Théorème 1.13. *Soit f une fonction convexe telle qu'il existe $x \in E$ avec $f(x) < \infty$. Si f est semi-continue inférieurement relativement à la topologie faible, alors $f = f^{**}$.*

DÉMONSTRATION. Il est clair que $f \geq f^{**}$ car

$$\begin{aligned} f^{**}(y) &= \sup_{l \in E'}(y(l) - f^*(l)) \\ &= \sup_{l \in E'}(l(y) - \sup_{x \in E}(l(x) - f(x))) \\ &\leq \sup_{l \in E'}(l(y) - (l(y) - f(y))) \\ &= f(y). \end{aligned}$$

Supposons qu'il existe un point x_0 tel que $f(x_0) > f^{**}(x_0)$. Considérons l'ensemble $C = f^{-1}((-\infty, f^{**}(x_0)])$ qui est convexe et fermé et prenons l'ensemble $B = \{x_0\}$ qui est convexe et compact. Alors par l'énoncé 2 de la proposition 1.9, il existe une variété linéaire affine donnée par h et α qui sépare strictement B et C. Si l'on pose $\beta = h(x_0)$, nous obtenons alors que

$$f^{**}(x_0) = \sup_{l \in E'}(x_0(l) - f^*(l)) \geq (x_0(h) - f^*(h)) = \beta - f^*(h).$$

Mais

$$f^*(h) = \sup_{x \in E}(h(x) - f(x)) \geq \beta - f(x_0).$$

Donc nous obtenons

$$f^{**}(x_0) \geq \beta - \beta + f(x_0) = f(x_0)$$

ce qui contredit le fait que $f \geq f^{**}$. □

CHAPITRE II

EXEMPLES DE MESURE DE RISQUE ET MESURES DE RISQUE SUR L^1

2.1 Mesures de risque sur des espaces plus généraux

Dans ce chapitre, $(\Omega, \Im, P)$ fera référence à notre espace probabilisé. Nous noterons par $M_1(P)$ l'ensemble des mesures de probabilités sur $(\Omega, \Im)$ qui sont absolument continues par rapport à P. Comme d'habitude, nous noterons par $L^p(P)$ l'ensemble des variables aléatoires X telless que $E\left[|X|^p\right] < \infty$ par rapport à la mesure P.

Dans le chapitre précédent, nous supposions que les variables aléatoires étaient bornées et nous avions muni $L_B(\Omega, \Im)$, l'ensemble des variables aléatoires bornées sur Ω, de la métrique induite par la norme du supremum. Nous avions montré que dans ce contexte, la fonction ρ était semi-continue inférieurement pour la topologie faible $\sigma(L_B(\Omega, \Im), ba)$, ce qui nous permettait d'utiliser le théorème général sur les fonctions conjuguées (théorème 1.13). Un lecteur attentif remarquera qu'il s'agit du seul endroit où nous avons utilisé la topologie faible, et que nulle part ailleurs nous nous servions de la structure d'espace normé de $L_B(\Omega, \Im)$. Tout ce dont nous avions besoin c'était que l'ensemble $\{\rho \leq c\}$ soit fermé pour la topologie faible. Ainsi nous pourrions remplacer $L_B(\Omega, \Im)$ par n'importe quel espace vectoriel normé de fonctions χ sur $(\Omega, \Im)$. Il suffit simplement de supposer que l'ensemble $\rho^{-1}((-\infty, 0]) \subseteq \chi$ est faiblement fermé. Nous pouvons ainsi, et c'est ce que nous ferons, remplacer l'ensemble des fonctions mesurables bornées par un espace L^p pour un certain $p \geq 1$. Encore une fois nous noterons cet ensemble par χ. De cette discussion, nous pouvons conclure plus généralement que si ρ est une mesure convexe de risque sur un espace vectoriel réel topologique E et que pour tout c, l'ensemble $\{\rho \leq c\}$ est fermé relativement à la topologie $\sigma(E,E')$ où E' est le dual de E, alors ρ admet une représentation de la forme :

$$\rho(Z) = \sup_{Q \in M_{1,f}} (E_Q[-Z] - \alpha(Q))$$

pour une certaine fonction de pénalité α.

2.2 Mesures de risque sur L^1

Nous nous attarderons maintenant aux mesures de risque sur L^1, c'est-à-dire les mesures de risque convexes ρ sur l'ensemble des variables aléatoires intégrables telles que $\rho(X) = \rho(Y)$ si X et Y sont dans la même classe d'équivalence selon la norme définie par le supremum essentiel. Ce cas nous permet de constater que si Q est une mesure de probabilité qui n'est pas absolument continue par rapport à P, alors $\alpha_{\min}(Q) = \infty$. En effet, puisque Q n'est pas absolument continue par rapport à P, il existe un ensemble $A \in \mathfrak{I}$ tel que $Q(A) > 0$ et $P(A) = 0$. Posons donc $X_n := X - n\mathbf{1}_A$, et nous obtenons que $\rho(X_n) = \rho(X)$ car les deux variables sont presque partout égales. On a alors que

$$\alpha_{\min}(Q) \geq E_Q[-X_n] = E_Q[-x] + nE_Q[\mathbf{1}_A]$$

et le membre de droite est aussi grand que nous le voulons. D'où les deux théorèmes ci-dessous.

Théorème 2.1. *Si ρ est une mesure convexe de risque sur L^1 représentée par une fonction de pénalité α comme dans le théorème de représentation. Alors $\alpha(Q) = +\infty$ pour tout $Q \in M_{1,f}$ qui n'est pas absolument continue par rapport à P.*

Théorème 2.2. *Soit $\rho : L^p \to \mathbb{R}$ une mesure convexe de risque, alors ρ admet une représentation de la forme*

$$\rho(Z) = \sup_{Q \in M_1(P)} (E_Q[-Z] - \alpha(Q))$$

si et seulement si l'ensemble des positions acceptables de ρ est $\sigma(L^p, L^q)$-fermé (avec q le conjugué de p, i.e. $q = (1 - p^{-1})^{-1}$).

2.3 Exemples de mesures de risque sur L^1

Plus particulièrement, nous nous attarderons aux mesures sur certains sous-ensembles de L^1. Nous remarquons que le dernier théorème demeure vrai si on y remplace $\sigma(L^p, L^q)$-fermé par $\sigma(L^1, L^\infty)$-fermé. Le premier exemple sera celui de la mesure du pire scénario, qui consiste à prendre pour perte la plus grande valeur c, pour une variable aléatoire X de sorte que $P(X < -c) > 0$. En d'autres termes, cela revient à considérer la plus petite valeur que prend X avec probabilité non nulle.

Exemple 1. *Worst case scenario risk measure.* On définit la mesure de risque du pire scénario comme étant

$$\rho_{\max}(X) = -\operatorname{ess\,sup}(-X) = -\operatorname{ess\,sup}(X)$$

Proposition 2.3. *Nous avons que* $\rho_{\max}(X) = \sup_{Q \in M_1(P)} E^Q[-X]$

DÉMONSTRATION. L'ensemble des positions acceptables associé à la mesure $\rho_{\max}$ est le cône de $L^1(P)$ constitué des fonctions P-presque sûrement positives. Donc

$$\begin{aligned} \alpha_{\min}(Q) &= 0 \\ &\leq \sup_{X \in A_{\rho_{\max}}} E_Q[-X] \\ &= -\inf_{X \in A_{\rho_{\max}}} E_Q[X] \leq E_Q[0] \\ &= 0 \end{aligned} \tag{2.1}$$

Donc $\alpha_{\min}(Q) = 0$ pour tout $Q \in M_1(P)$, c'est-à-dire

$$\rho_{\max}(X) = \sup_{Q \in M_1(P)} E_Q[-X].$$

□

2.3.1 Définition de la CVAR$_\lambda$

Dans l'introduction du chapitre I, nous avons expliqué comment définir la valeur au risque (VAR$_\lambda$) pour une variable aléatoire X au niveau λ. De plus, nous avons expliqué en quoi cette mesure de risque n'était pas satisfaisante dans le sens où celle-ci ne correspondait pas nécessairement à notre intuition. Nous allons dans le prochain exemple, montrer comment nous pouvons définir la VAR_λ au niveau λ et comment à partir de cette définition, nous pouvons construire une mesure de risque qui elle sera cohérente . Pour les besoins de la démonstration, nous débuterons par quelques notions préliminaires sur les inverse généralisée et les quantiles.

Définition 11. Soit $F : (a,b) \to \mathbb{R}$ est une fonction monotone croissante et posons

$$c = \lim_{x \to a} F(x), \quad d = \lim_{x \to b} F(x).$$

Une fonction q est appelée un **inverse généralisée** de F ou quantile de F si

$$F(q(s)-) \leq s \leq F(q(s)+), \quad s \in (c,d).$$

Dans le cas de $F(q(s)-)$ et $F(q(s)+)$, on les appelle quantile inférieur et supérieur respectivement de la fonction F. Si λ est une valeur dans l'image de la fonction F et q un quantile de F, on appelle λ-quantile de F la valeur $q(\lambda)$ qui n'est pas uniquement déterminée. Posons pour F, comme précédemment

$$q^-(s) = \inf\{x \in \mathbb{R} \mid F(x) > s\}, \quad q^+(s) = \sup\{x \in \mathbb{R} \mid F(x) < s\}.$$

Ces fonctions sont appelées les inverses généralisés à gauche et à droite de F.

Proposition 2.4. *Une fonction $q : (c,d) \to (a,b)$ est un inverse généralisée de F si*

$$q^-(s) \leq q(s) \leq q^+(s).$$

En particulier, q^- et q^+ sont des inverses généralisées.

Nous nous rappelons qu'au départ, nous avions défini une mesure de risque sur la variation du prix d'un actif par rapport à son coût initial investi au taux sans risque. Cela motive la définition suivante pour la VAR_λ.

Définition 12. Si X définit le prix d'un actif à un temps ultérieur, disons τ, nous définissons la **valeur à risque** au niveau $\lambda \in (0,1]$ pour X comme étant le $(1-\lambda)$-quantile inférieur de $X_0 \exp(r\tau) - X$ où r est le taux sans risque et τ la période d'investissement. Nous la noterons $\mathrm{VAR}_\lambda(X)$.

Il faut noter ici, comme nous l'avons vu dans l'exemple 1 du chapitre premier, que la VAR_λ n'est pas une mesure de risque au sens de notre définition.

Définition 13. Nous définissons la **valeur au risque conditionnelle**, notée CVAR_λ (pour *conditionnal value at risk*) au niveau $\lambda \in (0,1]$ pour une variable aléatoire X par la formule ci-dessous :

$$\mathrm{CVAR}_\lambda(X) = \frac{1}{\lambda}\int_0^\lambda \mathrm{VAR}_\gamma(X)\mathrm{d}\gamma \tag{2.2}$$

Proposition 2.5. *Pour tout $\lambda \in (0,1)$ et q un λ-quantile de X, on a*

$$\mathrm{CVAR}_\lambda(X) = \frac{1}{\lambda}E[(q-X)^+] - q.$$

DÉMONSTRATION. Soit q_X une fonction de quantile de X telle que $q_X(\lambda) = q$. Nous avons alors que

$$\begin{aligned}
\frac{1}{\lambda}E[(q-X)^+] - q &= \frac{1}{\lambda}\int_0^1 (q - q_X(t))^+ dt - q \\
&= \frac{1}{\lambda}\int_0^\lambda (q - q_X(t)) dt - q \\
&= \frac{1}{\lambda}\int_0^\lambda q_X(t) dt \\
&= \mathrm{CVAR}_\lambda(X).
\end{aligned}$$

□

Théorème 2.6. *Pour $\lambda \in (0,1]$, CVAR_λ est une mesure cohérente qui est semi-continue inférieurement. Elle a pour représentation*

$$\mathrm{CVAR}_\lambda(X) = \max_{Q \in \mathcal{Q}_\lambda} E_Q[-X]$$

où $\mathcal{Q}_\lambda$ est l'ensemble des mesures de probabilité Q absolument continues par rapport à P et telles que dQ/dP soit borné par $1/\lambda$. De plus, $\mathcal{Q}_\lambda$ est maximal dans le sens où c'est le plus grand ensemble de mesures de probabilité pour lequel le théorème de représentation soit vérifié.

DÉMONSTRATION. Pour $\lambda \in (0,1]$, considérons la mesure de risque $\rho_\lambda(X) = \sup_{Q \in \mathcal{Q}_\lambda} E_Q[-X]$. Supposons de plus, pour le moment, que $X < 0$ et définissons la mesure R équivalente à P par

$dR/dP = X/E[X]$. Alors

$$\begin{aligned}
\rho_\lambda(X) &= \sup_{Q\in\mathcal{Q}_\lambda} E_Q[-X] \\
&= \sup_{Q\in\mathcal{Q}_\lambda} E_R\left[-X\frac{dQ}{dR}\right] \\
&= \sup_{Q\in\mathcal{Q}_\lambda} E_R\left[-X\frac{dP}{dR}\frac{dQ}{dP}\right] \\
&= \sup_{Q\in\mathcal{Q}_\lambda} E_R\left[-X\frac{E[X]}{X}\frac{dQ}{dP}\right] \\
&= E[-X]\sup_{Q\in\mathcal{Q}_\lambda} E_R\left[\frac{dQ}{dP}\right]
\end{aligned}$$

tels que $0 \leq dQ/dP \leq 1/\lambda$ et $E[dQ/dP] = 1$. En multipliant la dernière condition par λ nous avons que

$$\rho_\lambda(X) = \frac{E[-X]}{\lambda}\sup_{Q\in\mathcal{Q}_\lambda} E_R\left[\frac{dQ}{dP}\right]$$

avec la condition que $0 \leq dQ/dP \leq 1$ et $E[dQ/dP] = \lambda$. Nous pouvons remplacer la condition $E[dQ/dP] = \lambda$ par $E[dQ/dP] \leq \lambda$ puisque le supremum sera assurément atteint lorsque $E[dQ/dP] = \lambda$. Cette remarque nous permet d'appliquer le lemme de Neyman-Pearson et de conclure que le supremum est atteint par la dérivée de Radon-Nykodim :

$$\frac{dQ_0}{dP} = I_{\{X<q\}} + \kappa I_{\{X=q\}}$$

pour un certain λ-quantile q de X et un certain $\kappa \in [0,1]$. Donc,

$$\begin{aligned}
\rho_\lambda(X) &= \frac{E[-X]}{\lambda}E_R\left[\frac{dQ_0}{dP}\right] \\
&= \frac{1}{\lambda}E\left[-X\frac{dQ_0}{dR}\right]
\end{aligned}$$

et puisque $\frac{1}{\lambda}\frac{dQ_0}{dR}\frac{dR}{dP} \in \mathcal{Q}_\lambda$, on en conclut que

$$\begin{aligned}
\rho_\lambda(X) &= \max_{Q\in\mathcal{Q}_\lambda} E_Q[-X] \\
&= E_{Q_0}[-X] \\
&= \frac{1}{\lambda}(E[-X\cap X<q] - q\lambda + qP(X<q)) \\
&= \frac{1}{\lambda}E[(q-X)^+] - q \\
&= \mathrm{CVAR}_\lambda(X)
\end{aligned}$$

par la proposition précédente. Ce qui démontre la proposition pour une variable aléatoire X telle que $X < 0$. Si X est quelconque, nous savons que $\operatorname{ess\,sup}(X) = c$ pour un certain $c \in \mathbb{R}$. Posons Y comme étant la variable aléatoire égale à X si $X < c$ et 0 sinon. Nous avons alors que :

$$\mathrm{CVAR}_\lambda(Y) = \mathrm{CVAR}_\lambda(X)$$

$$\rho_\lambda(Y) = \rho_\lambda(X)$$

et que

$$Y - c < 0.$$

D'où l'on tire :

$$\rho_\lambda(Y-c) = \rho_\lambda(Y) + c = \mathrm{CVAR}_\lambda(Y) + c = \mathrm{CVAR}_\lambda(Y-c)$$

ce qui nous permet de conclure que

$$\mathrm{CVAR}_\lambda(X) = \mathrm{CVAR}_\lambda(Y) = \rho_\lambda(Y) = \rho_\lambda(X).$$

Il faut maintenant voir que Q_λ est maximal. On se rappelle que le théorème de représentation des mesures convexes de risque nous permet d'écrire

$$\mathrm{CVAR}_\lambda(X) = \max_{Q \in M_{1,f}} (E_Q[-X] - \alpha_{\min}(Q))$$

avec

$$\alpha_{\min}(Q) = \sup_{X \in \chi}(E_Q[-X] - \mathrm{CVAR}_\lambda(X)).$$

Dans le même contexte, nous avions conclu que si CVAR_λ était une mesure cohérente, alors $\alpha_{\min}(Q) \in \{0, \infty\}$. Donc pour montrer que Q_λ est l'ensemble maximal où $\alpha_{\min}(Q) = 0$, il suffit de prouver que $\sup_{X \in \chi}(E_Q[-X] - \mathrm{CVAR}_\lambda(X)) = \infty$ pour $Q \notin Q_\lambda$. Soit ϕ la densité d'une telle mesure Q. Donc il existe un λ' et un $k > \frac{1}{\lambda'}$ tel que $P\left\{ \phi \wedge k > \frac{1}{\lambda'} \right\} > 0$. Posons pour $c > 0$,

$$X^{(c)} = -c(\phi \wedge k)\mathbf{1}_{\left\{\phi \geq \frac{1}{\lambda'}\right\}}.$$

Puisque

$$P\left\{ X^{(c)} < 0 \right\} = P\left\{ -c(\phi \wedge k)\mathbf{1}_{\{\phi \geq \frac{1}{\lambda'}\}} < 0 \right\} = P\left\{ \phi > \frac{1}{\lambda'} \right\} \leq \lambda' < \lambda$$

car $X^{(c)} \leq 0$ si $\phi < 1/\lambda'$. Nous avons donc que $\mathrm{VAR}_\lambda\left(X^{(c)}\right) = 0$ et par la proposition 2.5 nous obtenons que

$$\mathrm{CVAR}_\lambda\left(X^{(c)}\right) = \frac{1}{\lambda} E\left[-X^{(c)}\right] = \frac{c}{\lambda} E\left[\phi \wedge k; \phi \geq \frac{1}{\lambda'}\right].$$

D'autre part, nous avons que

$$E_Q\left[-X^{(c)}\right] = cE\left[\phi(\phi \wedge k); \phi \geq \frac{1}{\lambda'}\right] \geq \frac{c}{\lambda'} E\left[(\phi \wedge k); \phi \geq \frac{1}{\lambda'}\right].$$

Donc la différence entre $E_Q\left[-X^{(c)}\right]$ et $\mathrm{CVAR}_\lambda\left(X^{(c)}\right)$ devient arbitrairement grande quand c est grand. □

2.3.2 Le cas de la mesure *shortfall risk*

Un autre type de mesure qu'il est possible de représenter facilement sous la forme du théorème de représentation des mesures de risque est celle des mesures de type *shortfall risk*.

Définition 14. Une fonction $l : \mathbb{R} \to \mathbb{R}$ sera appelée une **fonction de perte** si elle est croissante et non identiquement constante.

Ici, nous ne considérerons que les fonctions de perte l convexes. Notons encore une fois par $L_B(\Omega, \Im)$ l'ensemble des fonctions mesurables bornées sur $(\Omega, \Im)$. Soit P une mesure de probabilité sur cet espace et soit x_0 un point intérieur dans l'image de l. Définissons un ensemble de positions acceptables comme suit :

$$A := \left\{X \in L_B(\Omega, \Im) \mid E_P[l(-X)] \leq x_0\right\}.$$

Puisque la fonction l est convexe et croissante, nous constatons que plus les valeurs de x sont grandes, plus les valeurs de $l(x)$ grandissent de manière non linéaire. Dans ce cas, plus x est petit, plus $l(-x)$ sera grand, ce qui fait en quelque sorte augmenter le poids des valeurs plus petites par rapport à celles qui sont plus grandes dans l'espérance. Si nous revenons à notre ensemble A ci-haut, nous voyons qu'il est non vide car il existe X tel que $l(-x) = x_0$. La variable aléatoire X identiquement égale à x est clairement bornée et celle-ci est donc dans A car

$$E[l(-x)] = E[x_0] = x_0 \leq x_0.$$

De plus cet ensemble est convexe car si $\lambda \in [0,1]$ et $X, Y \in L_B(\Omega, \Im)$, alors

$$\begin{aligned} E[l(-\lambda X + (1-\lambda)Y)] &\leq E[\lambda l(-X) + (1-\lambda)l(-Y))] \\ &\leq \lambda x_0 + (1-\lambda)x_0 \\ &= x_0. \end{aligned}$$

Maintenant, si $X \in A$ et $Y \in L_B(\Omega, \Im)$ avec $X \leq Y$, alors $-Y \leq -X$, ce qui nous donne que $l(-Y) \leq l(-X)$ et donc $E[l(-Y)] \leq E[l(-X)] \leq x_0$. On a alors par la discussion en section 3 du chapitre I que $\rho_A : L_B(\Omega, \Im) \longrightarrow \mathbb{R}$ définie par

$$\rho_A(X) = \inf\{ x \in \mathbb{R} \mid X + x \in A \}$$

est une mesure de risque convexe.

De plus, A est fermé et convexe, donc fermé dans la topologie faible, ce qui nous permet de conclure, avec le théorème 13, que ρ_A a une représentation de la forme $\rho(Z) = \sup_{Q \in M_1(P)}(E_Q[-Z] - \alpha(Q))$.

Exemple 2. Si $l(x) = e^{\beta x}$, alors

$$\begin{aligned} \rho(X) &= \inf\{ m \in \mathbb{R} \mid E[l(-m-X)] \leq x_0 \} \\ &= \inf\{ m \in \mathbb{R} \mid E[e^{-\beta(m+X)}] \leq x_0 \} \\ &= \inf\left\{ m \in \mathbb{R} \mid \frac{1}{x_0} E[e^{-\beta(X)}] \leq e^{\beta m} \right\} \\ &= \inf\left\{ m \in \mathbb{R} \mid \frac{1}{\beta}(\ln(E[e^{-\beta(X)}]) - \ln(x_0)) \leq m \right\} \end{aligned}$$

donc $\rho(X) = \frac{1}{\beta}\Big(\ln(E[e^{-\beta(X)}]) - \ln(x_0)\Big)$. Avec cette dernière formule, nous pouvons calculer $\alpha_{\min}(Q)$ pour notre mesure et on obtient que

$$\begin{aligned} \alpha_{\min}(Q) &= \sup_{X \in \chi}\Big(E_Q[-X] - \rho(X)\Big) \\ &= \sup_{X \in \chi}\Big(E_Q[-X] - \frac{1}{\beta}(\ln(E[e^{-\beta(X)}]) - \ln(x_0)\Big) \\ &= \sup_{X \in \chi}\Big(E_Q[-X] - \frac{1}{\beta}\ln(E[e^{-\beta(X)}])\Big) - \frac{1}{\beta}\ln(x_0) \\ &= \frac{1}{\beta}\Big(H(Q \mid P) - \ln(x_0)\Big) \end{aligned}$$

où $H(Q \mid P)$ est l'entropie relative de Q par rapport à P.

2.4 Conclusion

On a pu remarquer, à travers les exemples que nous avons donnés, que les mesures de risque étaient définies sur l'espace L^1. Par contre, au premier chapitre, nos variables aléatoires étaient celles qui étaient bornées. Du point de vue financier, la première approche semble plus naturelle. Le simple argument qu'il n'y ait qu'un nombre fini de ressource sur la planète l'indique. Par contre du point de vue de la modélisation, il semble plus naturel de se placer dans le cadre où nos variables peuvent avoir des rendements non bornés, comme dans le cas du modèle Black-Scholes. C'est d'ailleurs ce qui motive, dans ce chapitre, les définitions et exemples que nous avons donnés.

CHAPITRE III

LE PROBLÈME DE MINIMISATION DU RISQUE

Dans le chapitre qui suit, S_t est un processus stochastique qui modélise le prix d'un actif quelconque. Nous supposerons que notre modèle de marché financier est de type Black-Scholes. Donc qu'il existe un actif $B(t)$ que l'on appelle actif sans risque dont le prix au temps t, est déterminé par

$$B(t) = e^{rt},$$

d'un actif risqué S_t dont le prix satisfait à l'équation différentielle stochastique

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

où W_t est un mouvement brownien standard et μ et σ sont des réels positifs appelés moyenne et variance de S_t.

Dans un marché financier, il existe ce qu'on appelle des droits contingents. Ce sont des contrats qui donnent le pouvoir, ou l'obligation, d'exécuter une certaine transaction financière à un temps donné. Par exemple un contrat à terme est un contrat entre deux parties, A et B, donnant l'obligation à A d'acheter un certain actif de B à une date déterminée et à un prix donné. Nous dirons dans ce cas que la partie A est en position longue et que B est en position courte, que la date où la transaction doit être effectuée est la date d'échéance du contrat et que le prix est le prix d'exercice du contrat. Par exemple, un contrat à terme sur un baril de pétrole avec un prix d'exercice de 75\$ à une date T garantit que la partie A achètera le baril à 75\$ du vendeur B.

Un autre exemple, celui qui nous intéressera, est celui d'une option d'achat ou de vente de type européen. Une option d'achat sur un actif X donne le droit, mais non l'obligation à la partie A d'acheter ou de vendre le cas échéant l'actif X à une date déterminée et à un prix donné d'avance. Par exemple, supposons que nous achetons une option de vente sur l'action de IBM avec un prix d'exercice de 30$ et une date de maturité qui serait le 30 décembre 2007. Alors nous avons le droit, mais non l'obligation, de vendre à la partie B l'action de IBM à 30$. Si, au 30 décembre 2007, le prix de l'action est plus haut que notre prix d'exercice, alors nous n'exercerons pas le droit de vente puisque que nous ne pouvons vendre plus cher sur le marché. Par contre, si le prix de l'action est plus bas que le prix d'exercice, alors nous exerçons l'option et nous pouvons vendre les actions à un prix plus haut que sur le marché.

Ce type de contrat fut inventé, à la base, dans un but de protection. Par exemple, un producteur de blé sait qu'il devra vendre le blé qu'il aura produit à une certaine date, alors dans le but de protéger son prix de vente, il pourrait être intéressé à entrer dans un contrat de ce type. Le producteur limitera son exposition au risque de fluctuation du prix du blé, ce que nous appellerons couvrir notre risque.

Dans notre cas, nous nous limiterons aux options de ventes sur l'actif S_t. Une option de vente donne le droit, mais non l'obligation, à son détenteur de vendre un actif (l'actif sous-jacent) à un prix déterminé à l'avance (le prix d'exercice) à une date donnée (date de maturité). Le flux d'un tel contrat est $\max(0, K - S_T)$ où K est le prix d'exercice et S_T est la valeur de l'actif au temps T. En effet, si le prix de l'actif est plus grand que le prix d'exercice, le détenteur n'exercera pas son droit puisqu'il peut vendre l'actif sur le marché à un meilleur prix. Dans la situation inverse le détenteur exercera son droit de vendre les actions à un prix K puisque c'est un prix plus élevé sur le marché, donc il gagne $K - S_T$\$. Sans vouloir entrer dans les détails de la théorie de l'évaluation du prix des options, on peut mentionner que le prix dépend de 5 paramètres : l'échéance, le prix de l'actif au moment de l'achat, le taux d'intérêt sans risque, la volatilité du prix de l'actif et le prix d'exercice. Puisque dans notre cas nous regarderons des options de vente sur le même actif, les quatre premiers paramètres seront considérés constants. La formule du prix sera indiquée un peu plus tard dans le chapitre.

3.1 Le problème de minimisation de la VAR

Nous notons par $P_t^{X_i}$ le prix au temps t d'une option de vente sur S_t d'échéance T avec $T-t=\tau$. Étant donné une contrainte de budget C, ce que nous tenterons de faire dans cette section est de trouver un portefeuille $(h_1,\ldots,h_n)$ d'options de vente avec, respectivement prix d'exercices $P_t^{X_1},\ldots,P_t^{X_n}$, tels que la mesure de risque du portefeuille composé des options et de l'action soit minimisée. Cette technique qui consiste à utiliser un certain actif pour réduire la perte d'un autre actif est appelée une couverture. Le problème revient donc à trouver une série d'options de vente, telle que la couverture obtenue avec ces dernières pour notre actif S_t minimise le risque associé.

Nous commencerons en premier lieu par exposer une solution au problème de minimisation de la VAR_λ d'un portefeuille composé d'un actif risqué à l'aide d'une option de vente. Ce problème a déjà été résolu Ahn *et al.* (1999). À partir des calculs élaborés pour la VAR_λ, nous pourrons faire le même exercice avec la CVAR_λ. Dans ce dernier cas, les calculs seront sensiblement les mêmes. Nous poursuivrons ensuite avec le même problème, mais dans le cas où la mesure est de type *shortfall risk* et nous terminerons finalement en donnant l'existence d'une solution dans le cas d'une mesure convexe de risque générale. Dans ces deux derniers cas, nous n'aurons par contre d'autre choix que de changer certaines hypothèses pour s'assurer de l'existence d'une solution.

Posons V_t la valeur du portefeuille couvert au temps t ; V_t est donc un assemblage du titre S_t et des options $P_t^{X_1},\ldots,P_t^{X_n}$. La valeur de l'actif S_t au temps $t+\tau$ sera notée $S_{t+\tau}$ où τ est la durée de la position à risque. Par le lemme d'Itô, on peut trouver que $\ln(S_{t+\tau})$ est de loi normale centrée en $u=\ln(S_t)+\left(\mu-\frac{1}{2}\sigma^2\right)\tau$ et de variance $s=\sigma^2\tau$. La VAR_λ de l'actif est par définition le λ-quantile c tel que $P\{-S_{t+\tau}<c\}=1-\lambda$. Nous avons alors que

$$\text{VAR}_\lambda\,(S_{t+\tau})=-S_t\exp\left\{\Phi^{-1}(\lambda)s+(\mu-\frac{1}{2}\sigma^2)\tau\right\}$$

où Φ est la fonction de répartition de la loi normale centrée réduite. En effet, on a :

$$\begin{aligned} &P\{\,-S_{t+\tau}<c\,\}=1-\lambda \\ &\qquad\Leftrightarrow\quad P\{\,\ln(S_{t+\tau})<\ln(-c)\,\}=\lambda \end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow \quad P\left\{\frac{\ln(S_{t+\tau})-u}{s} < \frac{\ln(-c)-u}{s}\right\} = \lambda \\
&\Leftrightarrow \quad \Phi\left(\frac{\ln(-c)-u}{s}\right) = \lambda \\
&\Leftrightarrow \quad \frac{\ln(-c)-u}{s} = \Phi^{-1}(\lambda) \\
&\Leftrightarrow \quad c = -\exp\left(\Phi^{-1}(\lambda)s+u\right) = -S_t \exp\left\{\Phi^{-1}(\lambda)s + (\mu - \frac{1}{2}\sigma^2)\tau\right\}.
\end{aligned}$$

Nous pouvons remarquer que :

1. le coût des options affecte les flux monétaires dans tous les états du monde, donc augmentera la VAR_λ d'un montant égal au coût de la couverture. En effet, peu importe les événements qui surviendront dans le futur, nous savons avec certitude que l'effet de la couverture diminue le flux monétaire d'un montant égal à celui de la protection. Puisque le montant destiné à la couverture est dépensé en début de période, on doit capitaliser ce coût pour obtenir la valeur de la diminution du flux monétaire du portefeuille. Si par exemple, la couverture coûte C\$ alors ce montant que j'ai dépensé vaut, à la fin de la période, $C\exp(r\tau)$\$;

2. l'option est exercée seulement si le prix d'exercice est plus grand que le prix de l'action au temps T. Si le prix d'exercice a une valeur moindre que la VAR_λ de l'actif, alors le flux monétaire de l'option au niveau λ est 0 et n'affecte pas la VAR_λ du portefeuille total. Donc dans ce cas, le flux monétaire de l'option affecte la VAR_λ seulement si le prix de levée X de l'option est plus haut que la $S_{t+\tau}$. C'est à l'aide de cette dernière remarque que nous pourrons sans perte de généralité, supposer que le prix de levée des options est au-dessus de la $\mathrm{VAR}_\lambda(S_{t+\tau})$.

Avec ces deux remarques, nous pouvons voir que la $\mathrm{VAR}_\lambda(V_T)$ du portefeuille couvert est donc :

$$\mathrm{VAR}_\lambda(V_T) = S_t e^{r\tau} - S_t \exp(\theta(\lambda)) - \sum_{i=1}^{n} h_i(X_i - S_t\exp(\theta(\lambda)) + Ce^{r\tau}$$

où

$$\theta(\lambda) = \Phi^{-1}(\lambda)s + (\mu - \frac{1}{2}\sigma^2)\tau;$$

c'est-à-dire que $\mathrm{VAR}_\lambda(V_T)$ est le coût initial du portefeuille actualisé : $S_t e^{r\tau} + (C/P_t)\, e^{r\tau}$, duquel on soustrait la valeur des flux au niveau λ : $S_T + \sum_i h_i(X_i - S_t)$.

Nous pouvons ensuite simplifier le problème par quelques autres remarques préliminaires. Le prix d'une option de vente sur S_t est donné par la formule suivante :

$$P_t = Xe^{-r(T-t)}\Phi(d_1) - S_t\Phi(d_2),$$

où

$$d_1 = \frac{\ln(X/S_t) - (r - \frac{\sigma^2}{2})(\tau)}{\sigma\sqrt{\tau}}$$

et

$$d_2 = \frac{\ln(X/S_t) - (r + \frac{\sigma^2}{2})(\tau)}{\sigma\sqrt{\tau}}.$$

Supposons que nous ayons une stratégie qui consiste à utiliser différents prix de levée. En considérant une option dont le prix de levée est le prix moyen pondéré des prix de levée de la stratégie, nous obtenons alors que cette option engendrera le même flux que la stratégie au niveau critique λ. Par contre, puisque le prix d'une option est une fonction convexe du prix de levée, la stratégie consistant en une seule option de vente est moins coûteuse que la première stratégie. Pour montrer cette dernière affirmation, il suffit de voir que la deuxième dérivée de P_t par rapport à X est plus grande que 0. Ultérieurement, nous verrons que

$$\frac{\partial P_t}{\partial X} = e^{-r(\tau)}\Phi(d_1)$$

et

$$\frac{\partial \Phi}{\partial X} = \frac{1}{2\pi}\exp(-x^2/2)\frac{1}{X}$$

donc

$$\frac{\partial^2 P_t}{\partial X^2} = \frac{e^{r\tau}}{X}\frac{e^{(-d_1^2/2)}}{2\pi} \geq 0$$

pour $X > 0$. Par ailleurs, la contrainte de prix de couverture sera atteinte au sens où l'argent qui n'est pas dépensé dans la couverture offre le même rapport coût/bénéfice puisque la couverture augmente linéairement la protection. Autrement dit, chaque dollar dépensé réduit de façon proportionnelle le risque, indépendemment du nombre de dollars déjà dépensés. Par conséquent, l'institution dépensera toujours le montant maximal.

Le problème de minimisation devient alors :

$$\min_{h,X} \mathrm{VAR}_\lambda = -S_t\exp(\theta(\lambda)) - h(X - S_t\exp(\theta(\lambda)) + hP_t\exp(r\tau) \qquad (3.1)$$

sous la contrainte

$$C = hP_t, \ 0 \leq h < 1.$$

En substituant la contrainte de budget dans VAR_λ, nous obtenons que :

$$\begin{aligned}
X^* &= \arg\min_X \left\{ -S_t \exp(\theta(\lambda)) - \frac{C}{P_t}(X - S_t \exp(\theta(\lambda)) + hP_t \exp(r\tau) \right\} \\
&= \arg\max_X \left\{ \frac{C}{P_t}(X - S_t \exp(\theta(\lambda)) + \exp(r\tau) \right\} \\
&= \arg\max_X \left\{ \frac{X - S_t \exp(\theta(\lambda)}{P_t} \right\}.
\end{aligned}$$

Nous remarquons que dans notre cas, la VAR_λ est une fonction affine du coût de couverture, et donc la contrainte de budget n'affecte pas le choix de X^*. Une autre remarque intéressante est que de trouver un tel X^* revient à maximiser la distance entre le ratio X/P_t du prix X sur le prix de l'option P_t et le ratio VAR_λ/P_t de la VAR_λ de l'actif non couvert et P_t.

La théorie du calcul différentiel et intégral nous donne que le X^* qui maximise (3.1) satisfait :

$$\frac{P_t - (X^* - S_t e^{\theta(\lambda)}) \partial P_t / \partial X}{P_t^2} = 0.$$

Donc un tel X^* satisfait à

$$P_t - (X^* - S_t e^{\theta(\lambda)}) \frac{\partial P_t}{\partial X} = 0. \tag{3.2}$$

Calculons $\partial P_t / \partial X$. Puisque $P_t = X e^{-r(T-t)} \Phi(d_1) - S_t \Phi(d_2)$, nous avons que

$$\frac{\partial P_t}{\partial X} = e^{-r\tau} \Phi(d_1) + X e^{-r\tau} \frac{\partial \Phi(d_1)}{\partial X} - S_t \frac{\partial \Phi(d_2)}{\partial X}. \tag{3.3}$$

Or $d_2 = d_1 - \sigma\sqrt{\tau}$, donc

$$\frac{\partial \Phi(d_2)}{\partial X} = \frac{\partial \Phi(d_1)}{\partial X} = \frac{1}{X}.$$

De plus,

$$\frac{\partial \Phi(x)}{\partial X} = \frac{1}{2\pi\sigma\sqrt{\tau}} \exp(-x^2/2),$$

ce qui nous donne que

$$\begin{aligned}
\frac{\partial\Phi(d_1)}{\partial X} &= \frac{1}{2\pi}\exp(-d_1^2/2)\frac{1}{\sigma\sqrt{\tau}X}, \\
\\
\frac{\partial\Phi(d_2)}{\partial X} &= \frac{1}{2\pi}\exp(-(d_1+\sigma\sqrt{\tau})^2/2)\frac{1}{X} \\
&= \frac{1}{2\pi}\exp(-(d_1)^2/2)\exp((d_1)\sigma\sqrt{\tau}-\sigma^2\tau/2)\frac{1}{X} \\
&= \frac{\partial\Phi(d_1)}{\partial X}\exp(-(d_1)\sigma\sqrt{\tau}-\sigma^2\tau/2) \\
&= \frac{\partial\Phi(d_1)}{\partial X}\exp(\ln(X/S)-r\tau+\sigma^2\tau/2-\sigma^2\tau/2)\frac{1}{X} \\
&= \frac{\partial\Phi(d_1)}{\partial X}\frac{X}{S}\exp(-r\tau)\frac{1}{X} \\
&= \frac{\partial\Phi(d_1)}{\partial X}\frac{1}{S}\exp(-r\tau).
\end{aligned}$$

En intégrant ces deux dernières équations avec (3.3) nous obtenons de P_t

$$\begin{aligned}
\frac{\partial P_t}{\partial X} &= e^{-r(\tau)}\Phi(d_1)+e^{-r(\tau)}\frac{1}{2\pi\sigma\sqrt{\tau}}\exp(-d_1^2/2)-e^{-r(\tau)}\frac{\partial\Phi(d_1)}{\partial X}\partial X \\
&= e^{-r(\tau)}\Phi(d_1.)
\end{aligned}$$

En remplaçant et en réorganisant les termes dans l'équation (3.2), nous obtenons qu'un tel X^* satisfait à :

$$\frac{P_t}{\frac{\partial P_t}{\partial X}} = X^* - S_t e^{\theta(\lambda)} \Leftrightarrow S_t\frac{\Phi(d_2)}{\Phi(d_1)} = S_t e^{\theta(\lambda)}. \tag{3.4}$$

Pour prouver l'existence d'une solution, nous pouvons utiliser le théorème des fonctions implicites appliqué à l'équation :

$$F(X,\beta) = \Phi(d_1)e^{\theta(\lambda)-r\tau} - \Phi(d_2)$$

où $\beta = (S_t,\mu,\sigma,r,\tau,\alpha)$. En dérivant par rapport à X, nous obtenons

$$\frac{\partial\Phi(d_1)}{\partial X}e^{\theta(\lambda)-r\tau} - \frac{\partial\Phi(d_1)}{\partial X}\frac{e^{(-r\tau)}}{S} = \frac{\partial\Phi(d_1)}{\partial X}\frac{e^{-r\tau}}{X}\left(e^{\theta(\lambda)}-\frac{1}{S}\right).$$

Nous voyons de cette dernière équation que, pour qu'une solution existe, il faut que

$$\theta(\lambda) \neq -\ln(S).$$

De plus, l'équation (3.4) nous indique que pour qu'une solution existe, puisque $\frac{\Phi(d_2)}{\Phi(d_1)} < 1$, il faut que $\theta(\lambda) < r\tau$. Donc sous ces deux conditions, nous savons par le théorème des fonctions implicites qu'une solution à notre problème de minimisation existe.

À l'issue de cette section, nous sommes donc en mesure de conclure qu'une solution au problème de minimisation de la VAR_λ existe, pourvu que $\theta(\lambda) < r\tau$ et $\theta(\lambda) \neq -\ln(S)$.

3.2 Le problème de minimisation de CVAR

Dans cette section, nous reprendrons l'analyse de la section précédente dans le but de résoudre le même problème mais en remplacant la mesure VAR_λ par la mesure cohérente de risque CVAR_λ. Donc nous voulons minimiser

$$\mathrm{CVAR}_\lambda(V_T(X)) = \frac{1}{\lambda}\int_0^\lambda \mathrm{VAR}_\gamma(V_T(X))\,d\gamma$$

où V_t est la valeur du portefeuille composé de l'actif S_t et de $h = C/P_t$ options de vente sur S_t à un prix d'exercice X. De plus, pour des raisons de simplicité, nous supposerons que $h < 1$, c'est-à-dire qu'en aucun cas le portefeuille est complètement couvert ou même sur-couvert. Comme précédemment, nous pouvons déjà simplifier le problème en remarquant que chaque dollar dépensé pour un prix d'exercice donné entraîne une diminution de la CVAR_λ. Ainsi l'institution dépensera toujours le plein budget C pour la couverture.

Avec l'analyse de la section précédente, nous savons que VAR_λ est donnée par

$$\begin{cases} S_t\exp(r\tau) - S_t\exp(\theta(\lambda)) - h(X - S_t\exp(\theta(\lambda)) + hP_t\exp(r\tau), & \text{si } X > S_t\exp(\theta(\lambda)); \\ S_t\exp(r\tau) - S_t\exp(\theta(\lambda)) + hP_t\exp(r\tau), & \text{si } X \leq S_t\exp(\theta(\lambda)). \end{cases}$$

Posons η tel que $X = Se^{\theta(\eta)}$. Nous avons deux cas possibles. Si $\eta < \lambda$ alors

$$\begin{aligned} \mathrm{CVAR}_\lambda(V_T(X)) &= \frac{1}{\lambda}\int_\eta^\lambda \left(S_te^{(r\tau)} - S_te^{\theta(\gamma)} + hP_te^{(r\tau)}\right)d\gamma \\ &\quad + \frac{1}{\lambda}\int_0^\eta \left(S_te^{(r\tau)} - S_te^{\theta(\lambda)} - h(X - S_te^{\theta(\gamma)}) + hP_te^{r\tau}\right)d\gamma \\ &= \frac{1}{\lambda}\int_0^\lambda \left(S_te^{r\tau} - S_te^{\theta(\gamma)} + hP_te^{(r\tau)}\right)d\gamma + \frac{1}{\lambda}\int_0^\eta h(X - S_te^{\theta(\gamma)})\,d\gamma. \end{aligned}$$

Nous remarquons alors que

$$\arg\min_X \mathrm{CVAR}_\lambda(V_T) = \arg\max_X \frac{1}{\lambda}\int_0^\eta h(X - S_te^{\theta(\gamma)})d\gamma.$$

Si $\eta \geq \lambda$, alors

$$\begin{aligned}
\mathrm{CVAR}_\lambda(V_T) &= \int_0^\lambda \left(S_t e^{(r\tau)} - S_t e^{\theta(\lambda)} - h(X - S_t e^{\theta(\gamma)}) + hP_t e^{r\tau}\right) d\gamma \\
&= \frac{1}{\lambda}\int_0^\lambda \left(S_t e^{r\tau} - S_t e^{\theta(\gamma)} + hP_t e^{(r\tau)}\right) d\gamma + \frac{1}{\lambda}\int_0^\lambda h(X - S_t e^{\theta(\gamma)}) d\gamma.
\end{aligned}$$

Comme précédemment, nous concluons que

$$\arg\min_X \mathrm{CVAR}_\lambda(V_T) = \arg\max_X \frac{1}{\lambda}\int_0^\lambda h(X - S_t e^{\theta(\gamma)})\, d\gamma.$$

c'est-à-dire, X satisfait à

$$\arg\min_X \mathrm{CVAR}_\lambda(V_T) = \arg\max_X \frac{1}{\lambda}\int_0^{\min(\eta,\lambda)} h(X - S_t e^{\theta(\gamma)})\, d\gamma.$$

Cette fonction n'est pas dérivable au point X lorsque $\eta = \lambda$, c'est-à-dire si $X = S_t e^{\theta(\lambda)}$. Par contre, puisque l'ensemble des points sur lesquels la fonction n'est pas dérivable est de mesure nulle, cette fonction est donc presque partout dérivable et nous pouvons utiliser le test de la dérivée première pour trouver les extremas. Premièrement, montrons que $\min(\eta,\lambda) = \eta$. Nous obtenons par la règle de chaîne que si $\min(\eta,\lambda) = \lambda$, alors :

$$\begin{aligned}
\frac{\partial}{\partial X}\mathrm{CVAR}_\lambda(V_T) &= \frac{1}{\lambda}\int_0^\lambda \frac{\partial}{\partial X} h(X - S_t e^{\theta(\gamma)})\, d\gamma \\
&= \frac{1}{\lambda}\int_0^\lambda h(X - S_t e^{\theta(\gamma)})\, d\gamma \\
&= \frac{C}{\lambda}\int_0^\lambda \frac{P_t - \frac{\partial P_t}{\partial X}(X - S_t e^{\theta(\gamma)})}{P_t^2}\, d\gamma
\end{aligned}$$

et si $\min(\eta,\lambda) = \eta$:

$$\frac{\partial}{\partial X}\mathrm{CVAR}_\lambda(V_T) = \frac{1}{\lambda}\int_0^\eta \frac{\partial}{\partial X} h(X - S_t e^{\theta(\gamma)}) d\gamma.$$

En posant cette dernière équation égale à 0, nous obtenons

$$\begin{aligned}
&\frac{C}{\lambda}\int_0^{\min(\eta,\lambda)} \frac{P_t - \frac{\partial P_t}{\partial X}(X - S_t e^{\theta(\gamma)})}{P_t^2} d\gamma = 0 \\
&\Leftrightarrow \int_0^{\min(\eta,\lambda)} P_t - \frac{\partial P_t}{\partial X}(X - S_t e^{\theta(\gamma)}) d\gamma = 0 \\
&\Leftrightarrow \int_0^{\min(\eta,\lambda)} \frac{P_t}{\frac{\partial P_t}{\partial X}} d\gamma - \int_0^{\min(\eta,\lambda)} (X - S_t e^{\theta(\gamma)}) d\gamma = 0 \\
&\Leftrightarrow \int_0^{\min(\eta,\lambda)} S_t e^{-r\tau}\frac{\Phi(d_2)}{\Phi(d_1)} d\gamma - \int_0^{\min(\eta,\lambda)} (S_t e^{\theta(\gamma)}) d\gamma = 0. \qquad (3.5)
\end{aligned}$$

Nous remarquons que cette dernière équation est la même que celle de la section précédente, outre le fait que nous intégrons sur tout le domaine, l'intervalle $[0,\eta]$. Supposons maintenant que $\min(\eta,\lambda)=\lambda$. Nous avons alors, avec l'équation (3.5), que pour tout $X>\theta(\lambda)$

$$\frac{\Phi(d_2)}{\Phi(d_1)}=\frac{S_te^{r\tau}}{\lambda}\int_0^{\lambda}(S_te^{\theta(\gamma)})d\gamma=c$$

où c est une constante positive. Donc, en supposant que $\min(\eta,\lambda)=\lambda$, $\Phi(d_2)/\Phi(d_1)$ est une constante qui ne dépend pas de X ce qui contredit le fait que $\Phi(d_2)/\Phi(d_1)$ dépend de X. Dans notre étude, nous pouvons donc supposer que $X\leq e^{\theta(\lambda)}$. Dans ce cas, nous voyons que la fonction

$$f(X,S_t,\mu,\sigma,r,\tau)=\int_0^{\eta}S_te^{-r\tau}\frac{\Phi(d_2)}{\Phi(d_1)}d\gamma-\int_0^{\eta}(S_te^{\theta(\gamma)})d\gamma$$

possède une dérivée partielle partout non nulle par rapport à X. En effet,

$$\begin{aligned}\frac{\partial f}{\partial X}&=\frac{\partial}{\partial X}\left(\int_0^{\eta}S_te^{-r\tau}\frac{\Phi(d_2)}{\Phi(d_1)}d\gamma-\int_0^{\eta}(S_te^{\theta(\gamma)})d\gamma\right)\\&=S_te^{-r\tau}\frac{\Phi(d_2)}{\Phi(d_1)}\frac{\partial\eta}{\partial X}+\eta\frac{\partial}{\partial X}S_te^{-r\tau}\frac{\Phi(d_2)}{\Phi(d_1)})+S_te^{\theta(\eta)})\frac{\partial\eta}{\partial X}\\&=\frac{S_te^{-r\tau}}{X\sigma^2\tau}\frac{\Phi(d_2)}{\Phi(d_1)}+\eta\Phi'(d_1)\end{aligned}$$

Donc le théorème des fonctions implicites nous garantit l'existence d'une solution X^* qui satisfait à (3.5). De plus, X^* satisfait à :

$$\int_0^{\min(\eta,\lambda)}S_te^{-r\tau}\frac{\Phi(d_2)}{\Phi(d_1)}d\gamma-\int_0^{\min(\eta,\lambda)}(S_te^{\theta(\gamma)})d\gamma=0.$$

En conclusion, dans le cas de la minimisation de CVAR_{λ}, le problème de trouver un portefeuille optimal composé d'une action et d'une option de vente sur cette action peut-être résolu.

3.3 Cas de la mesure *shortfall risk*

Dans cette section, nous reprennons encore une fois le problème initial, mais dans le cas de la mesure de type *shortfall risk*. Nous voulons donc trouver comment, à l'aide de h options de vente, on peut minimiser le risque de notre portefeuille avec la contrainte $hP_t^X\leq C$ où C est, comme auparavant, la contrainte de budget et P_t^X le prix de l'option de vente au temps t de prix d'exercice X. Dans le cas qui nous intéresse maintenant, nous voulons minimiser la mesure de risque de type *shortfall risk*. Sauf avis du contraire, nous utiliserons ici les mêmes notations que dans les sections précédentes. Nous devrons par contre ajouter des hypothèses pour s'assurer de l'existence d'une solution. Nous ne pourrons pas exhiber un minimum explicitement, mais l'existence de la solution nous sera assurée.

3.3.1 Préliminaires mathématiques

Premièrement, il nous faut une méthode pour calculer le *shortfall risk* puisque la mesure de risque nous est donnée en termes d'ensemble de positions acceptables. Nous démontrerons donc comment nous pouvons calculer la mesure en termes de zéro d'une certaine fonction. Nous pouvons voir que la translation est continue par rapport à la topologie faible. De plus, on a

$$\{\rho(X) \leq c\} = \{\rho(X+c) \leq 0\}.$$

Ce dernier ensemble étant fermé pour la topologie faible, nous avons donc :

Théorème 3.1. *Si ρ est une mesure convexe de risque telle que l'ensemble des positions acceptables est faiblement fermé, alors cette mesure de risque est semi-continue inférieurement pour rapport à la topologie faible.*

Théorème 3.2. *Pour une fonction de perte l strictement convexe, la mesure du* shortfall risk, *pour une variable aléatoire X, est donnée par l'unique solution de l'équation*

$$E\big[l(-z-X)\big] - x_0 = 0.$$

3.3.2 Existence d'une solution au problème de minimisation

Nous démontrons maintenant qu'il existe une solution aux problèmes de minimisation dans le cas de la mesure du *shortfall risk* et nous expliquons comment, à l'aide du dernier théorème, nous pourrons calculer ce minimum. Nous notons d'abord qu'il ne sera pas possible de résoudre le problème dans le cadre général imposé par notre problème. En effet, nous devrons nous assurer que le seuil h de couverture soit strictement plus grand que 0 pour pouvoir utiliser les outils d'analyses à notre disposition. Cela nous forcera à limiter l'espace dans lequel il existe des solutions. Le première étape sera de comprendre cette dernière assertion et de montrer que dans ce cas, il est possible de résoudre le prix d'exercice X en fonction de h. Nous tenterons ensuite de démontrer par des outils élémentaires d'analyse que dans ce cas, notre problème de minimisation revient à minimiser une fonction sur $\mathbb{R}$.

Premièrement, nous remarquons que si h représente, comme avant, le ratio $C/P_t(X)$, il est possible de considérer la fonction

$$F(h,X) = h - \frac{C}{P_t(X)}$$

qui est définie pour $h \geq 0$. Nous pouvons remarquer que F a une dérivée partielle par rapport à X partout plus grande que 0 car le prix d'une option de vente $P_t(X)$ est strictement croissant par rapport à X. Donc nous pouvons résoudre, une fois de plus par le théorème des fonctions implicites, X en fonction de h, c'est-à-dire qu'il existe localement, pour $h > 0$, une fonction $X(h)$ qui satisfait $F(h,X(h)) = 0$. De plus une telle fonction est continue puisqu'elle est différentiable. Toutefois le théorème des fonctions implicites ne nous permet pas de résoudre quand $h = 0$. Nous fixerons donc un $h_{\min} > 0$ qui sera notre seuil minimum. Nous pouvons interpréter ce $h_{\min}$ comme étant le seuil minimum de couverture que l'institution voudrait maintenir. Dans ce cas, l'application

$$\begin{aligned} r : [h_{min}, 1] &\rightarrow \mathbb{R} \\ h &\mapsto \rho((S_t + C)e^{r\tau} - S_T - h\max(X(h) - S_T, 0)) \end{aligned}$$

est continue. En effet, ρ est continue étant une fonction convexe car nous avons montré précédemment que la mesure du type *shortfall risk* est une mesure convexe de risque. De plus, $[h_{min}, 1]$ est compact, la fonction définie ci-haut atteint donc son maximum et son minimum sur ce compact. Évidemment, ce minimum est celui qui minimise la mesure de risque. Il est à noter que jusqu'à maintenant, nous n'avons pas utilisé la définition de la mesure du *shortfall risk*. La méthode de résolution que l'on vient d'exhiber peut s'appliquer à n'importe quelle mesure de risque. Maintenant, si nous remplaçons ρ par la définition du *shortfall risk*, nous obtenons donc que la mesure du *shortfall risk* est donnée par l'unique zéro de l'équation

$$E[l(-z - (S_t + C)e^{r\tau} - S_T - h\max(X(h) - S_T, 0))] - x_0.$$

Nous pourrons donc appliquer une procédure numérique pour trouver ce minimum en posant l'équation ci-haut égale à zéro. La technique que nous proposons s'inspire de Gundel et Weber (2008). Cette procédure se divise en trois étapes :

- pour trouver le minimum, nous choisirons une suite de points x_k qui converge vers la solution x cherchée. Pour cela nous pouvons utiliser la méthode de la sécante par exemple ou tout autre méthode qui n'utilise pas le calcul de la dérivée puisque nous ne sommes pas en mesure de calculer analytiquement la dérivée de notre mesure convexe de risque ;
- pour réaliser la première étape, nous devrons d'abord être en mesure de calculer la fonction $X(h)$ pour un certain h donné. Cela pourra se faire par une méthode de résolution d'équations non linéaires comme la méthode de Newton ;
- finalement, il faudra trouver une méthode pour calculer ρ. Puisque nous avions vu que la mesure de type *shortfall risk*, associée à une fonction de perte l, pouvait se calculer comme le zéro de la fonction

  $$E[l(-X-z)]-x_0$$

  nous pouvons utiliser une simulation de Monte Carlo pour évaluer l'intégrale et la méthode de Newton pour calculer le z qui annule la fonction précédente.

Exemple 3. Si nous considérons la fonction de perte $l(x)=\frac{1}{\beta}\exp(-\beta x)$, nous avons déjà montrer au chapitre II, que la mesure de risque associée à l au niveau x_0 est donné par la formule

$$\rho(X)=\frac{1}{\beta}(\ln(E[e^{-\beta(X)}])-\ln(x_0)).$$

Notre problème consiste donc à trouver un minimum à la fonction qui à $h\in[h_{min},1]$ fait correspondre

$$\frac{1}{\beta}(\ln(E[e^{-\beta((S_t+C)e^{r\tau}-S_T-h\max(X(h)-S_T,0))}])-\ln(x_0))$$

Notre programme pourrait fonctionner ainsi : pour chaque h, nous trouverons $X(h)$ avec la fonction *fzero* de **Matlab** qui utilise la méthode de Brent pour trouver les zéros d'une fonction. Ensuite, pour la valeur de X trouvée, nous utiliserons la méthode de Monte-Carlo pour résoudre l'intégrale. Maintenant que nous avons décrit comment calculer la valeur de la mesure de risque en chaque point, nous pouvons trouver le minimum à l'aide de la fonction *fminbnd* de **Matlab**.

3.4 Le cas général

Une fois de plus, nous reprennons le problème de minimisation de la mesure de risque avec l'option de vente des sections précédentes. Ici, nous ne nous bornerons pas à une mesure de risque en particulier. Nous expliquerons comment, en modifiant un peu les hypothèses, nous pouvons en toute généralité et sans spécifier la mesure convexe de risque, démontrer l'existence d'une solution au problème de minimisation de la mesure de risque du portefeuille composé d'une action S_t et de h options de vente sur S_t avec comme précédemment la contrainte $h < 1$ et $hP_t^X < C$.

Dans le cas qui nous concerne, notons par V, comme précédemment, l'ensemble des portefeuilles admissibles, c'est-à-dire l'ensemble des portefeuilles composés d'une action et de h options de vente sur S_t satisfaisant aux contraintes énoncées précédemment. L'existence d'un portefeuille qui minimise la mesure de risque repose sur le théorème ci-dessous.

Théorème 3.3. *Supposons que $\rho : \Lambda \longrightarrow \mathbb{R}$ soit une fonction convexe sur un ensemble de variables aléatoires Λ définies sur un espace probabilisé $(\Omega, \Im, P)$. Supposons de plus que*

1. *Λ est convexe et fermé par rapport à la topologie de la convergence P-presque partout.*
2. *Il existe une variable aléatoire $W \in L^0_+(\Omega, \Im, P)$ telle que $X \leq W < \infty$ P-presque partout pour tout $X \in \Lambda$.*
3. $\sup_{X \in \Lambda} \rho(X) < \infty$.
4. *ρ est semi-continue inférieurement pour la convergence P-presque sûre.*

Alors, sous ces conditions, il existe un X^* dans Λ qui maximise ρ. De plus ce X^* est unique si ρ est strictement convexe.

Avant d'entamer la démonstration, nous énoncerons sans démonstration la proposition suivante (pour une démonstration, voir Föllmer et Schied, 2002).

Proposition 3.4. *Soit (X_n) une suite dans $L^0(\Omega, \Im, P)$ telle que $\sup_n |X_n| < \infty$, P-presque partout. Alors il existe une suite de combinaisons linéaires convexes*

$$Y_n \in \text{conv}\{X_n, X_{n+1}, \dots\}$$

qui converge P-presque partout vers un certain $Y \in L^0(\Omega, \Im, P)$.

DÉMONSTRATION (du théorème 3.3). L'unicité est claire puisque la mesure de risque est convexe. Donc s'il existait deux variables aléatoires X et Y qui minimisaient la mesure de risque, n'importe quelle combinaison linéaire convexe de X et Y serait dans Λ et aurait une mesure de risque plus petite que celle de X ou de Y. Soit une suite X_n dans Λ qui converge vers l'infimum de la mesure convexe de risque. Puisque

$$\sup_n |X_n| \leq W < \infty,$$

il existe donc une suite

$$\overline{X_n} \in \operatorname{conv}\{X_n, X_{n+1}, \dots\}$$

qui converge P-presque partout vers un certain $\overline{X}$. Puisque Λ est convexe, chacun des $\overline{X_n}$ est dans Λ et comme ce dernier est fermé en convergence P-presque partout, on a $\overline{X} \in \Lambda$. Nous pouvons maintenant écrire chaque $\overline{X_n}$ comme combinaison linéaire convexe des X_{n_i} où n_i est plus grand que n. Par suite, on a

$$\rho(\overline{X_n}) \leq \sum_i \alpha_i^n \rho(X_{n_i})$$

et nous en concluons que

$$\rho(\overline{X_n}) \leq \sup_{m \geq n} \rho(X_n).$$

Le théorème de convergence dominée nous donne donc que

$$\rho(\overline{X}) = \lim_{n \uparrow \infty} \rho(X_n)$$

et le membre de droite converge vers l'infimum de la mesure de risque. □

Maintenant armé de ce théorème, posons $\overline{V}$ la fermeture de V par rapport à la convergence en probabilité. Cet ensemble est encore convexe, donc satisfait à la propriété 1 et si l'on prend pour W la variable aléatoire $S_0 + |S_T - S_0|$, nous remarquons que cette variable aléatoire satisfait à la propriété 2 du théorème. Pour la propriété 3, il suffit de remarquer que si $\rho(W)$ est fini alors, $\sup \rho(V_t)$ est aussi fini. La propriété 4 est trivialement satisfaite puisque ρ est une mesure convexe de risque. Donc sous ces conditions, il existe une variable aléatoire Z^* dans V qui minimise ρ. Ce Z^* est dans la fermeture de V mais pas nécessairement dans V. Donc, Z^* n'est peut-être pas un portefeuille de la forme recherchée. Néanmoins, nous savons que V est dense dans $\overline{V}$ et nous pouvons donc approcher autant que nous le voulons la solution Z^* par un portefeuille dans V.

Si nous récapitulons, nous avons donc démontré que si $\rho(W)$ est fini, nous pouvons trouver une suite de Cauchy dans Λ telle que sa limite sera un minimum de la fonction ρ dans un espace plus grand. Il suffit donc d'utiliser un algorithme qui cherche un minimum dans l'espace Λ et le résultat lors de l'arrêt de l'algorithme sera près d'une solution minimale.

3.5 Appendice

Tout au long du présent chapitre nous avons utilisé le théorème des fonctions implicites. Nous nous contenterons simplement ici d'énoncer ce résultat classique d'analyse ; en faire la démonstration impliquerait des pré-requis d'analyse dépassant le cadre de ce mémoire.

Théorème 3.5. *Soit V_1, V_2, W trois espaces de Banach et $\Omega \subset V_1 \times V_2$ un ouvert contenant un point (x_0, y_0). Supposons que nous nous donnions $F : \Omega \longrightarrow W$, une application différentiable et telle que*

$$F(x_0, y_0) = 0.$$

Supposons de plus que l'application linéaire

$$D_2F(x_0, y_0) : V_2 \longrightarrow W$$

soit inversible. Alors, dans ces conditions, il existe des voisinages Ω_1 de x_0 et Ω_2 de y_0 tels que $\Omega_1 \times \Omega_2 \subset \Omega$ et une fonction différentiable $g : \Omega_1 \longrightarrow \Omega_2$ telle que

$$F(x, g(x)) = 0$$

et

$$Dg(x) = -(D_2F(x, g(x)))^{-1} \circ D_1F(x, g(x))$$

pour tout $x \in \Omega_1$.

CONCLUSION

Dans ce mémoire, nous avons étudié la théorie des mesures cohérentes et convexes de risque et appliquer cette théorie à un problème concret de gestion des portefeuilles. Rappelons ici le parcours que nous avons fait au long de ce présent travail. Nous avons tout d'abord exposé la théorie générale des mesures convexes de risque sur l'espace des variables aléatoires bornés et nous avons remarqué que les mesures cohérentes de risque n'étaient qu'un cas particulier de ces dernières. Étudier les mesures convexes de risque nous a par ailleurs permis de faire un petit détour dans la théorie de l'analyse fonctionnelle. On a d'ailleurs pu se rendre compte que la démonstration de l'existence de la représentation de toute mesure convexe de risque par une fonction de pénalité utilisait presque uniquement des résultats de ce dernier sujet tels que le théorème de dualité pour la transformée de Fenchel-Legendre.

Au chapitre suivant, nous avons par la suite donné quelques exemples de mesures convexes de risque dans L^1 et nous avons expliqué comment les résultats du premier chapitre pouvaient être adaptés dans le cas où l'on remplaçait l'espace des variables aléatoires bornées par un espace L^p.

Enfin, nous avons terminé par la résolution du problème de la minimisation de certaine mesure de risque pour un portefeuille composé d'un actif S_t et de h option de ventes sur cet actif. Nous avons démontré l'existence d'une solution à ce problème dans le cas de la CVAR_λ, de la mesure du type *shortfall risk* et comment, en changeant légèrement les hypothèses de notre problème, on pouvait trouver une solution à notre problème de minimisation dans un cadre n'impliquant aucun choix d'une mesure de risque particulière.

Dans le cas de la mesure de type *shortfall risk*, nous avons énoncé un algorithme permettant de trouver le minimum basé sur des travaux de Stefan Weber. Une personne capable de programmer sur Matlab peut en faire un programme très rapidement.

Dans la résolution de notre problème de minimisation, en aucun cas nous avons pu fournir une formule fermée pour notre problème. Dans un problème semblable, celui de la maximisation de l'utilité sous la contrainte de la mesure *shortfall risk*, Gundel et Weber (2008) ont réussi à obtenir une formule explicite pour exprimer la variable aléatoire qui répondait au problème d'optimisation. Faute de temps, nous n'avons pu d'adapter ces résultats au problème qui nous concerne mais je reste convaincu qu'il doit être possible d'adapter sa démonstration.

RÉFÉRENCES

Ahn, D.-H., J. Boudoukh, M. Richardson et R. F. Whitelaw. 1999. « Optimal risk management using options ». *Journal of Finance*, vol. 54, no 1, p. 359–375.

Artzner, P., F. Delbaen, J.-M. Eber et D. Heath. 1999. « Coherent measures of risk ». *Mathematical Finance*, vol. 9, no 3, p. 203–228.

Bourbaki, Nicolas. 1971. *Topologie générale. Chapitres 1 à 4.* Paris : Dunod.

————. 1981. *Espaces vectoriels topologiques.* Paris : Dunod.

Föllmer, H., et A. Schied. 2002. « Convex measures of risk and trading constraints ». *Finance and Stochastics*, vol. 6, no 4, p. 429–447.

————. 2004. *Stochastic Finance. An Introduction in Discrete Time.* Berlin : Walter de Gruyter.

Grothendieck, Alexandre. 1973. *Topological Vector Spaces.* New York : Gordon and Breach.

Gundel, A., et S. Weber. 2008. « Utility maximization under a shortfall risk constraint ». À paraître dans *Journal of Mathematical Economics*.

Jost, Jürgen. 2005. *Postmodern Analysis*, 3e éd. Berlin : Springer.