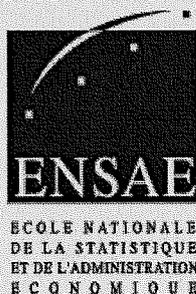


ÉCOLE NATIONALE DE LA STATISTIQUE ET DE L'ADMINISTRATION
ÉCONOMIQUE
3, avenue Pierre Larousse – 92245 Malakoff CEDEX, FRANCE



Mémoire d'actuariat – Promotion 2005

Mesures de risque et allocation optimale de capital

Jean-Christophe Bouëtté
Jean-François Chassagneux

MOTS-CLÉS: *Mesures de risque, fonds propres, utilité, sélection de portefeuille, copules*

KEYWORDS: *Risk Measures, Economic Capital, Utility, Portfolio Selection, Copulae*

ENCADREMENT: Arthur Charpentier (ENSAE – CREST)

CORRESPONDANT E.N.S.A.E.: Isabelle Nagot (Université Paris I)

MÉMOIRE NON CONFIDENTIEL

ÉCOLE NATIONALE DE LA STATISTIQUE ET DE L'ADMINISTRATION
ÉCONOMIQUE
3, avenue Pierre Larousse – 92245 Malakoff CEDEX, FRANCE

Actuarial Thesis – Year 2005
**Risk Measures and Optimal
Capital Allocation**

Jean-Christophe Bouëtté
Jean-François Chassagneux

KEYWORDS: *Risk Measures, Economic Capital, Utility,
Portfolio Selection, Copulae*

Due to the spectacular development of financial markets, banks and other financial institutions manage huge portfolios of correlated assets. This is also true in the insurance field. Therefore, these companies have to assess the risks they hold correctly, so as to properly choose their portfolio allocation.

The Expected Utility Theory can be used to evaluate the risk faced by the companies. However, it has shown some flaws and alternative theories have been developed, which derive from simple modifications of Von Neumann and Morgenstern's axioms: the Choquet capacity theory, the Gilboa-Schmeidler theory and, more recently, an axiomatic theory based on the notion of coherent risk measures. In the first chapter, we provide a short survey of these theories.

Once defined, risk measures can be used to perform portfolio selection or economic capital allocation. In the second chapter we formalize these matters in term of an optimization program for a company, and give solutions when possible.

At last, we introduce the problem of dependency between assets. We define dependency using the notions of elliptic distributions and copulae, and show how different correlation structures lead to different allocations.

ÉCOLE NATIONALE DE LA STATISTIQUE ET DE L'ADMINISTRATION
ÉCONOMIQUE

3, avenue Pierre Larousse – 92245 Malakoff CEDEX, FRANCE

Mémoire d'actuariat – Promotion 2005
Mesures de risque et allocation
optimale de capital

Jean-Christophe Bouëtté
Jean-François Chassagneux

MOTS-CLÉS : *Mesures de risque, fonds propres, utilité,
sélection de portefeuille, copules*

Le développement des marchés financiers ces trente dernières années fait que les banques gèrent aujourd'hui d'importants portefeuilles, composés d'actifs plus ou moins corrélés. Elles doivent donc pouvoir évaluer au mieux le risque généré par ces portefeuilles, ceci afin de pouvoir assurer leur solvabilité, et décider au mieux de l'allocation de fonds à telle ou telle composante d'un portefeuille.

La modélisation du comportement face au risque des agents s'est parallèlement affinée, la modification de certains axiomes de la théorie de l'utilité espérée ayant conduit à des représentations des préférences faisant appel à une intégrale de Choquet ou à des mesures additives. Plus récemment, une théorie axiomatique des mesures de risque adaptée à la régulation bancaire a été développée. Nous présentons ces différentes approches et montrons qu'elles peuvent être liées.

Les applications des mesures de risque sont multiples : gestion de portefeuille, allocation de capital économique au sein d'une entreprise. Nous donnons un programme général de maximisation pour une entreprise, et le résolvons pour le problème particulier de la gestion de portefeuille.

Enfin, la gestion de portefeuille peut être compliquée par le fait que l'agent connaisse mal la dépendance existant entre les composantes de son portefeuille. Après avoir défini la dépendance à travers la notion de copule, nous montrons comment les résultats sont modifiés lorsque cette information n'est pas entièrement disponible.

Table des matières

Introduction	7
1 Mesure du risque	9
1.1 Qu'est-ce que mesurer le risque?	9
1.1.1 L'espérance	10
1.1.2 La variance	10
1.1.3 La valeur en risque	11
1.1.4 Au-delà de la VaR	12
1.2 Approche axiomatique et scénarios de crise	13
1.2.1 Mesures de risque monétaires	13
1.2.2 Représentation des mesures de risque	16
1.3 Distorsion des préférences et mesures de Wang	18
1.3.1 La théorie duale de Yaari	18
1.3.2 Mesures de Wang et valeur en risque	19
2 Allocation de capital	23
2.1 Investissement et gestion de capital	23
2.1.1 Variables d'intérêt	23
2.1.2 Les programmes d'optimisation	24
2.2 Sélection de portefeuille: le cadre du modèle	25
2.2.1 Description	25
2.2.2 Fonction objectif et fonction de risque	26
2.2.3 Propriétés de la fonction de risque	26
2.2.4 Étude des contraintes	28
2.3 Résolution des programmes	28
2.3.1 Étude du programme général	29
2.3.2 Minimisation du risque sous contrainte	31
2.3.3 En présence d'actif sans risque	32
2.3.4 Quelques exemples	34
2.3.5 Maximisation de l'espérance de rendement	37
2.3.6 Application à la réassurance proportionnelle	37
2.4 Sélection de portefeuille et allocation de fonds propres	38
2.5 Allocation de capital économique	39

2.5.1	Définitions	39
2.5.2	Contributions au risque et différentiabilité	40
2.5.3	Exemples de contributions au risque	42
3	Impact de la structure de dépendance	45
3.1	Les distributions elliptiques	45
3.1.1	Définitions et propriétés	45
3.1.2	Expression des mesures de risque	46
3.2	Copule et dépendances	47
3.2.1	Motivation	47
3.2.2	Définition	48
3.2.3	Calculs des bornes pour la VaR	49
3.3	Structure de dépendance inconnue	50
3.3.1	Présentation du programme	50
3.3.2	Programme sur l'ensemble des copules	51
3.4	Structure de dépendance paramétrée	53
3.4.1	Structure gaussienne	53
3.4.2	La famille des copules de Clayton	56
3.4.3	Analyse de l'étude	58
	Conclusion	61
	Index	62
	Bibliographie	63

Introduction

Les entreprises du secteur de la banque et de l'assurance gèrent aujourd'hui des risques importants et de natures diverses. Elles sont soumises à des obligations réglementaires de maîtrise du risque, et cherchent également par elles-mêmes à éviter la faillite. Un des outils intervenant dans cette gestion compliquée est la mesure du risque, qui doit permettre de fournir un indicateur pertinent permettant de générer du profit en limitant le risque de subir des pertes trop importantes. Il est donc important de pouvoir décider du choix de cette mesure, et de savoir comment elle peut être utilisée dans un programme général de maximisation du profit. Comment ces entreprises doivent-elles mesurer leur risque? Comment peuvent-elles utiliser ces mesures pour décider d'une allocation optimale? Que se passe-t-il si elles ne connaissent pas la manière dont les risques interagissent entre eux? Ce sont les questions auxquelles nous tenterons d'apporter une réponse.

L'identification et la gestion du risque est au centre des préoccupations des institutions financières. La mesure du risque passe par sa quantification. Nous étudions ce problème dans la première partie. Nous exposons en particulier le cas des mesures de risque cohérentes, qui ont été introduites pour remédier à certains défauts de la variance et de la valeur en risque, ou VaR.

La résolution de ce problème va permettre à une institution financière de pouvoir faire des choix mathématiquement cohérents en ce qui concerne ses investissements. La première application importante est la gestion de portefeuilles. En investissant d'une façon intelligente, l'institution cherche à minimiser son risque selon, bien sûr, sa propre mesure de risque. Nous discutons ces problèmes de sélection de portefeuilles dans la deuxième partie. Nous cherchons à allouer de façon optimale un capital entre deux risques. Pour ce faire, plusieurs programmes d'optimisation ont été considérés, ainsi que plusieurs mesures de risque. L'approche classique moyenne-variance peut, sous certaines hypothèses, être étendue à d'autres mesures de risque. Nous examinons ensuite la contribution au risque de chacun des actifs du portefeuille.

La troisième partie est consacrée à l'étude de l'impact de la structure de dépendance. En pratique, même quand la distribution des lois marginales des valeurs du portefeuille de risques est correctement connue ou estimée, la structure de dépendance est souvent mal connue. La notion de dépendance définie à l'aide des copules fournit des outils permettant de ne plus faire l'hypothèse, courante en finance ou en assurance, des rendements suivant un vecteur gaussien. Nous nous intéressons en particulier à l'impact de la structure de dépendance sur l'allocation optimale de richesse, pour les trois mesures VaR, Tail VaR et variance.

Chapitre 1

Mesure du risque

Ce groupe de travail s'intéresse à la manière dont une entreprise, par exemple une banque ou une compagnie d'assurance, perçoit et gère l'aléa attaché à certains flux financiers qu'elle doit verser ou recevoir. Ce problème est envisagé d'un point de vue probabiliste : nous supposons que ces flux sont représentés par les réalisations d'une variable aléatoire sur un espace probabilisé. Ce sont ces variables aléatoires que nous dénommons risque, le terme pouvant également désigner l'aléa lui-même, ou la possibilité de la survenance d'événements défavorables.

Ce sujet est à la frontière d'interaction entre plusieurs domaines. Pour ce qui touche à la perception du risque, il rejoint la théorie de la représentation des préférences dans le risque, qui est plus vaste que la théorie de Von Neumann et Morgenstern. Pour ce qui touche à la gestion du risque, et à son éventuelle couverture ou réduction, il se rapproche de la théorie financière et actuarielle.

Dans cette première partie, nous nous intéressons principalement à la quantification et à la comparaison des risques. Nous adopterons, selon les cas,

- soit la convention actuarielle qui veut que les variables aléatoires représentent des *pertes* lorsqu'elles sont positives,
- soit la convention financière qui veut que les variables aléatoires représentent des *gains* lorsqu'elles sont positives,

ces conventions ayant été fixées par l'usage.

1.1 Qu'est-ce que mesurer le risque ?

Lorsqu'un agent est placé devant une alternative se résumant au choix entre deux risques (par exemple jouer à la loterie ou s'en abstenir), il aimerait disposer d'un critère pouvant le guider dans sa décision, ce critère pouvant

varier d'un individu à l'autre. Pour cela, il peut par exemple réduire l'information associée au risque à une grandeur scalaire, appelée mesure de risque¹. Comparer les risques revient alors à comparer les dites mesures. La question fondamentale, qui a occupé des générations d'actuaire, financiers et économistes, est alors de savoir quelle mesure considérer, une question subsidiaire étant de donner un sens à la mesure du risque, qui soit plus que la traduction d'une relation ordinale.

1.1.1 L'espérance

Les actuaires ont utilisé, et continuent à utiliser, une mesure simple du risque : son espérance mathématique, dite aussi prime pure. Du point de vue d'une compagnie d'assurance qui mutualise un très grand nombre de risques homogènes indépendants, cette mesure semble pertinente : la loi des grands nombres assure en effet la convergence en probabilité, et même presque sûrement vers la prime pure.

Dans la théorie microéconomique, la représentation des préférences par l'espérance d'utilité repose sur la même idée, à la nuance près d'une transformation des niveaux de richesse appelée fonction d'utilité. Nous reviendrons dans la section 1.3.1 sur la représentation des préférences, mais remarquons dès à présent que Maurice Allais [1] critique la théorie de von Neumann et Morgenstern dans les termes suivants :

Si j'ai la possibilité effective de participer à une longue suite de coups, mais que je puisse être ruiné, c'est-à-dire nécessairement écarté du jeu dès les premiers coups, voire même au premier, il est manifeste que la justification par la loi des grands nombres de la règle de l'espérance mathématique ne vaut en aucune façon. Ce serait pour moi une bien piètre consolation que de savoir que, si j'avais pu tenir le coup, mon gain aurait finalement tendu en probabilité vers son espérance mathématique. En tout état de cause [...] il existera toujours une certaine dispersion du gain moyen dont j'aurai à tenir compte.

La théorie financière reprend implicitement à son compte cette critique, et considère l'espérance comme un indice de rendement moyen idéal et non comme une mesure de risque. Markowitz [23] propose d'utiliser plutôt la variance comme mesure de risque dans son modèle de gestion de portefeuille.

1.1.2 La variance

Le *Capital Asset Pricing Model* (CAPM), que nous aborderons dans le deuxième chapitre, propose en effet, pour choisir parmi un ensemble de por-

1. Des mesures vectorielles ont été proposées dans [21], mais il s'agit d'un cas isolé dans la littérature.

tefeuilles (caractérisés par leur rendement) de prendre à espérance fixée celui qui a la plus faible variance.

Les mathématiques actuarielles n'ignorent pas non plus la variance : le théorème de la limite centrale indique en effet que la dispersion autour de la prime pure est distribuée selon une loi de Gauss, dont l'écart-type est pris en compte dans le calcul des primes à travers les chargements de sécurité.

La justification de la variance comme mesure de risque est très liée au fait que l'on considère des risques gaussiens. En effet, dans ce cadre restreint, l'espérance et la variance suffisent à caractériser complètement le risque, ce qui n'est pas le cas dans un cadre plus général. Supposons par exemple le risque non symétrique, la variance cesse d'être pertinente : une variance élevée peut tout aussi bien indiquer des gains importants que des pertes importantes ! Les financiers ont donc proposé d'autres mesures, dont la valeur en risque (*Value at Risk*) est la plus célèbre.

1.1.3 La valeur en risque

La *Value at Risk* (VaR) n'est rien d'autre qu'un quantile, le plus souvent élevé, du risque *considéré comme une fonction de perte*. Un de ses intérêts est qu'elle donne une idée de la queue de distribution, qui peut ne pas être négligeable lorsqu'on s'extrait du cadre gaussien. C'est donc une mesure de risque extrême. C'est à ce titre qu'elle a été adoptée par la *Bank for International Settlements* dans le nouvel accord de Bâle comme outil privilégié de gestion du risque. Nous donnons ci-dessous sa définition en même temps que celle des quantiles. Nous supposons que la variable X représente une perte.

DÉFINITION 1.1

Soit $\alpha \in [0,1]$ et X une variable aléatoire. Alors q est un quantile de niveau α , ou α -quantile de X si :

$$P(X < q) \leq \alpha \leq P(X \leq q).$$

L'ensemble des quantiles de niveau α est un intervalle de \mathbb{R} .

Exactement, on définit comme dans [9] :

$$Q_p[X] = \inf\{x \in \mathbb{R} \mid F_X(x) \geq p\},$$

$$Q_p^+[X] = \sup\{x \in \mathbb{R} \mid F_X(x) \leq p\}.$$

Ces derniers diffèrent lorsque F_X est constante sur un segment et vaut p .

En finance, la *VaR* est généralement associé à une période : 1 jour, 10 jours, et correspond alors à un quantile de perte. Nous adopterons comme convention $VaR_p[X] = Q_p[X]$.

L'exemple suivant montre que la VaR n'est pas sous-additive, ce qui est gênant pour une gestion saine du risque, comme nous le verrons plus loin.

EXEMPLE 1.1

Soit X un sous-jacent, on considère les deux options digitales suivantes, $A = \mathbb{1}_{\{X > U\}}$ et $B = \mathbb{1}_{\{X < L\}}$. On choisit L et U tel que $\mathbb{P}[X > U] = \mathbb{P}[X < L] = 0.008$. La *Value at Risk* de deux traders vendant respectivement une option A et une option B est nulle. La *Value at Risk* d'un trader vendant une option A et une option B vaut un. On a donc bien contradiction de la sous-additivité :

$$VaR_{0,99}[A + B] > VaR_{0,99}[A] + VaR_{0,99}[B].$$

On peut de plus s'interroger, d'une part sur le choix du niveau α , d'autre part sur le niveau réel des pertes lorsque la VaR est dépassée, i.e. le risque résiduel.

1.1.4 Au-delà de la VaR

Pour remédier à ces défauts de la VaR, d'autres mesures ont été proposées. Nous en présentons trois très proches, dont l'idée commune est de quantifier la perte lorsque la VaR est dépassée :

- la *Tail Value at Risk*, ou TVaR est la moyenne des VaR de niveau supérieur à α ;
- la *Conditional Tail Expectation* ou CTE représente la perte attendue sachant que la VaR est dépassée ;
- l'*Expected Shortfall*, ou ES au niveau p , c'est la prime *stop-loss* dont la rétention (ou franchise, ou priorité) est fixée à VaR_p .

Nous définissons précisément ces mesures ci-dessous :

DÉFINITION 1.2

$\forall p \in]0,1[$, la TVaR, la CTE et l'ESF sont définies respectivement par :

$$\begin{aligned} TVaR_p(X) &= \frac{1}{1-p} \int_p^1 VaR_q(X) dq \\ CTE_p(X) &= \mathbf{E}_{\mathbb{P}}[X | X > VaR_p(X)] \\ ES_p(X) &= \mathbf{E}_{\mathbb{P}}[(X - VaR_p(X))_+]. \end{aligned}$$

Notons que la TVaR est plus grande que la VaR correspondante. En fait, ces trois mesures sont très liées, comme le montre le théorème suivant, prouvé par exemple dans [8] :

THÉORÈME 1.1

$$TVaR_p(X) = VaR_p(X) + \frac{1}{1-p} ESF_p(X)$$

$$CTE_p(X) = VaR_p(X) + \frac{1}{1 - F_X(VaR_p(X))} ESF_p(X)$$

$$CTE_p(X) = TVaR_{F_X(VaR_p(X))}(X), \forall p \in]0,1[.$$

La troisième égalité montre qu'en particulier, la TVaR et la CTE coïncident pour des risques de distributions continues.

Au vu de cet aperçu des mesures les plus utilisées, apparaît la nécessité d'introduire un cadre théorique plus général, de manière à pouvoir comparer les mesures de risque. C'est l'objet de la section suivante.

1.2 Approche axiomatique et scénarios de crise

L'adoption de la VaR par le Comité de Bâle n'a pas éteint les discussions sur les mesures de risque, bien au contraire. Les premiers auteurs à avoir proposé une définition spécifiquement adaptée aux problèmes de régulation dans le secteur bancaire sont Artzner, Delbaen, Eber et Heath [2]. Dans cette section, nous présentons tout d'abord leur approche en reprenant la présentation faite dans [15] et [28], puis faisons le lien avec une théorie de la représentation des préférences basée sur des scénarios.

1.2.1 Mesures de risque monétaires

Nous supposons que X est une perte i.e. positive pour des pertes, convention actuarielle; la convention financière utiliserait les profits et pertes (P & L), valant $-X$. Un risque, ou portefeuille, X sera supposé appartenir à un ensemble χ de variables aléatoires bornées, par exemple $\chi = L_{\mathbb{R}}^{\infty}$. Il est à noter que, contrairement à ce qui est dit plus haut, ce n'est pas ici le point de vue de l'entreprise qui est envisagé, mais plutôt celui du régulateur.

Définitions

Le concept de mesure de risque monétaire (*monetary measure of risk*) donne une interprétation concrète aux mesures de risque. Nous allons tout d'abord définir cette notion de mesure de risque monétaire, puis nous montrerons qu'elle admet une interprétation simple: c'est la quantité d'actif sans risque qui, ajouté à une position, permet de rendre celle-ci acceptable par le régulateur.

DÉFINITION 1.3

Une application ρ de $\chi \rightarrow \mathbb{R}$ est dite mesure de risque monétaire si elle satisfait les propriétés suivantes:

Monotonie $\mathbb{P}(X \leq Y) = 1 \Rightarrow \rho(X) \leq \rho(Y)$ quels que soient les risques X et Y . Cette propriété traduit un constat tout à fait naturel : plus la perte est sévère, plus le capital nécessaire à la couverture est important.

Invariance au cash Pour tout risque X et toute constante c , $\rho(X + c) = \rho(X) + c$. Cette propriété entraîne $\rho(X - \rho(X)) = 0$ et donne aux mesures de risque monétaires leur interprétation comme un montant de capital.

On réclamera souvent en plus les deux propriétés suivantes :

Normalisation $\rho(0) = 0$ ce qui est le cas d'une mesure qui considère un risque nul associé à un rendement nul comme acceptable.

Neutralité Si X et Y ont même fonction de répartition, alors $\rho(X) = \rho(Y)$. Rappelons que cet axiome est le premier axiome menant à la théorie de l'espérance d'utilité.

Du point de vue du régulateur, il est souhaitable qu'une mesure de risque encourage la diversification. C'est pourquoi nous nous restreindrons dans la suite aux mesures vérifiant la définition suivante :

DÉFINITION 1.4

Une mesure de risque monétaire est appelée mesure de risque convexe si elle satisfait la propriété suivante :

Convexité Pour tout $\lambda \in [0,1]$, $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$. Cette propriété traduit directement le fait que la diversification n'accroît pas la mesure du risque.

On peut également se restreindre à une classe plus exigeante de mesures de risque, les mesures cohérentes introduites dans [2], et qui traduisent une croissance linéaire du risque :

DÉFINITION 1.5

Une mesure de risque convexe est appelée mesure de risque cohérente si elle satisfait la propriété suivante :

Positive homogénéité Pour tout risque X et toute constante λ positive, $\rho(\lambda X) = \lambda\rho(X)$.

Cependant, imposer une mesure du risque croissante linéairement avec celui-ci peut paraître trop restrictif, la classe plus large des mesures convexes aura donc notre préférence. Remarquons que si une mesure de risque monétaire est positivement homogène, alors il y a équivalence entre convexité et la propriété ci-dessous :

Sous-additivité Pour tous risques X et Y , $\rho(X + Y) \leq \rho(X) + \rho(Y)$. Cette propriété traduit également la réduction du risque par diversification.

La région des portefeuilles acceptables

Une notion intimement liée à celle de mesure de risque monétaire est celle de *région des portefeuilles acceptables*, qui est l'ensemble des portefeuilles acceptables par le régulateur :

$$\mathcal{A}_\rho = \{X \in \mathcal{X} \mid \rho(X) \leq 0\}.$$

La proposition suivante, tirée de [15, p.155], récapitule les propriétés de la région des portefeuilles acceptables en fonction de celles de la mesure de risque monétaire considérée :

PROPOSITION 1.1

Soit ρ une mesure de risque monétaire de région des portefeuilles acceptables \mathcal{A} . Alors :

1. \mathcal{A} est non-vide et satisfait :

$$\inf\{m \in \mathbb{R} \mid m \in \mathcal{A}\} > -\infty \quad (1.1)$$

$$X \in \mathcal{A}, Y \in \mathcal{X}, Y \geq X \Rightarrow Y \in \mathcal{A}; \quad (1.2)$$

2. ρ peut être reconstituée par la donnée de (\mathcal{A})

$$\rho(X) = \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}\};$$

3. ρ est convexe ssi \mathcal{A} est convexe, et ρ est cohérente ssi \mathcal{A} est un cône convexe.

Remarquons que dans [2], c'est d'abord la région des portefeuilles acceptables qui est définie, puis la mesure de risque comme étant *la plus petite quantité de cash à ajouter à un portefeuille pour le rendre acceptable* :

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

La proposition ci-dessous issue de [15, p.156] montre que les deux approches sont en fait équivalentes :

PROPOSITION 1.2

Soit \mathcal{A} un ensemble non-vide vérifiant (1.1) et (1.2), alors :

1. $\rho_{\mathcal{A}}$ est une mesure de risque monétaire, convexe si \mathcal{A} est convexe et cohérente si \mathcal{A} est un cône convexe;
2. \mathcal{A} est un sous-ensemble de $\mathcal{A}_{\rho_{\mathcal{A}}}$, avec égalité ssi pour tous $X \in \mathcal{A}$ et $Y \in \mathcal{X}$, $\{\lambda \in [0,1] \mid \lambda X + (1 - \lambda)Y \in \mathcal{A}\}$ est fermé dans $[0,1]$.

Cette définition concrète des mesures de risque monétaire comme une quantité de cash est primordiale. Ces mesures sont beaucoup plus que la simple traduction d'une relation d'ordre sur les variables aléatoires : elles quantifient le risque, lui donnent un prix.

Exemples de mesures de risque monétaires

La VaR, que nous avons présenté plus haut, est bien une mesure de risque monétaire : en effet, elle vérifie les propriétés de monotonie et d'invariance au cash. De plus, elle est positivement homogène. Cependant,

PROPOSITION 1.3

La VaR n'est pas une mesure de risque convexe, et a fortiori elle n'est pas cohérente.

Ceci est dû au fait qu'elle n'est pas sous-additive, comme nous l'avons montré dans l'exemple 1.1.

En revanche, la TVaR est cohérente :

PROPOSITION 1.4 (PROPRIÉTÉS DE LA TVaR)

1. La TVaR est invariante par translation et est homogène car la VaR l'est.
2. La TVaR est sous-additive.
3. La TVaR est monotone.

1.2.2 Représentation des mesures de risque

Les mesures de risque monétaires ont une interprétation aisée en terme de cash grâce à la définition de la région des portefeuilles acceptables. On peut se demander si elles ont d'autres interprétations utiles. Nous allons voir qu'elles peuvent être interprétées comme une perte moyenne maximale sur un ensemble de scénarii.

Dans le cas cohérent, il existe un théorème de représentation de ρ sous la forme :

$$\rho(X) = \sup_Q \{E_Q[X]\}.$$

Dans le cas convexe, on peut donner plus ou moins d'importance à la vraisemblance des scénarii en les pénalisant par une fonction α , ce qui nous donne la représentation plus générale :

$$\rho(X) = \sup_Q \{E_Q[X] - \alpha(Q)\}.$$

Le théorème suivant, basé sur un résultat de dualité dans des espaces de Banach et tiré de [6] et [15], formalise de manière précise et rigoureuse ce qui vient d'être dit.

THÉORÈME 1.2

Soit (Ω, \mathcal{F}) un espace mesurable, et χ l'ensemble des fonctions bornées mesurables sur cet espace. Soit $\mathcal{M}_{1,f}$ l'ensemble des mesures additives sur (Ω, \mathcal{F})

qui sont de masse 1, et \mathcal{M}_1 l'ensemble des mesures de probabilité sur (Ω, \mathcal{F}) .

Cas général : Soit ρ une mesure de risque convexe. On peut alors l'écrire sous la forme :

$$\rho(X) = \max_{Q \in \mathcal{M}_{1,f}} (E_Q[X] - \alpha(Q)),$$

avec $\alpha(Q) = \sup_{X \in \mathcal{A}_p} E_Q[-X]$ la fonction de pénalisation. Dans le cas cohérent $\alpha(\cdot)$ est à valeurs dans $\{0, +\infty\}$, et il existe donc $\mathcal{Q} \subset \mathcal{M}_{1,f}$ tel que :

$$\rho(X) = \max_{Q \in \mathcal{Q}} E_Q[X].$$

Cas P fixée : Si ρ est convexe, vérifie l'axiome de neutralité pour une probabilité P fixée et est continue par le haut (i.e. si $X_n \searrow X$ alors $\rho(X_n) \searrow \rho(X)$) alors, il existe une fonction de pénalisation α' telle que :

$$\rho(X) = \sup_{Q \in \mathcal{M}_1(P)} (E_Q[X] - \alpha'(Q)),$$

où $\mathcal{M}_1(P)$ est l'ensemble des probabilités absolument continues par rapport à P . Dans le cas cohérent il existe $\mathcal{Q}' \subset \mathcal{M}_1(P)$ tel que :

$$\rho(X) = \max_{Q \in \mathcal{Q}'} E_Q[X].$$

La représentation obtenue est proche d'une pratique courante des banques : les *stress tests* [24]. Le *stress testing* est un exercice réglementaire consistant à évaluer l'impact d'événements rares sur les actifs d'une banque. En particulier, cette exigence conduit à élaborer des scénarios de crise macro-économiques, et à en évaluer l'impact sur le portefeuille. Ce qui peut être traduit en termes de mesures de risque cohérente ou convexe, d'après le théorème ci-dessus.

Les deux exemples donnés par Hans Föllmer au cours de [14] montrent que ce concept de scénarios correspond à d'autres notions connues.

Exemple 1 On peut définir les risques X acceptables comme étant ceux pour lesquels il existe une stratégie de couverture, i.e. Y tel que $X + Y \geq 0$. On a alors :

$$\rho(X) = \sup_Q \{E_Q[X]\},$$

où Q est à valeurs dans l'ensemble des mesures martingales.

Exemple 2 On peut se donner un scénario de référence P , que l'on prend très au sérieux. On va cependant examiner les autres scénarii, mais en les

pénalisant par leur entropie par rapport à P (leur distance de Kullback à P). En d'autres termes :

$$\alpha(Q) = \frac{1}{\beta} \log \frac{dQ}{dP} dQ.$$

On retrouve alors la formule de l'équivalent certain pour un investisseur ayant une utilité exponentielle :

$$\rho(X) = \frac{1}{\beta} \log \mathbb{E}_P[e^{-\beta X}].$$

En fait, ce concept de scénario n'est pas nouveau dans la littérature sur la représentation des préférences, et il avait déjà été étudié par Gilboa et Schmeidler [16]. Ceux-ci montrent en effet qu'en considérant des axiomes plus larges que ceux de von Neumann et Morgenstern, on aboutit à une représentation des préférences du type :

$$\min_{Q \in \mathcal{Q}} E_Q[u(X)] \quad \text{où } \mathcal{Q} \subset \mathcal{M}_{1,f} \text{ et } X \text{ est une fonction de gain,}$$

ce qui est exactement la représentation à laquelle nous avons abouti en partant des mesures de risque monétaire, à la fonction d'utilité u près.

On voit donc que la théorie des mesures de risque monétaires, développée pour le but particulier de la régulation bancaire, se ramène à un cadre déjà connu dans le domaine de la représentation des préférences. Or, la littérature sur le sujet [5] a également donné naissance à un autre modèle, développé sous trois formes différentes :

- la théorie de l'espérance d'utilité dépendante du rang (RDEU) ou utilité anticipée de Quiggin [25] ;
- le modèle d'utilité à la Choquet de Schmeidler [29] ;
- la théorie duale de Yaari [34].

Dans la section suivante, nous étudions le modèle de représentation des préférences de Yaari comparativement à celui de von Neumann et Morgenstern, et montrons comment il a été traduit en termes de mesures de risque par Wang et Young [33].

1.3 Distorsion des préférences et mesures de Wang

1.3.1 La théorie duale de Yaari

Nous présentons ici brièvement une théorie duale de la théorie de von Neumann et Morgenstern, dans laquelle le rôle des paiements et des probabilités est inversée, d'où son nom de théorie duale. Sa formalisation sous

cette forme est due à Yaari [34]. Nous nous plaçons sur un espace $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et nous noterons G_X la fonction de survie de X variable aléatoire réelle.

Yaari remplace l'axiome d'indépendance de la théorie de l'espérance d'utilité par celui d'indépendance duale directe. Il obtient alors le théorème ci-dessous, tiré de [34]:

THÉORÈME 1.3

Une relation de préférence satisfait les cinq axiomes cités dans [34] si et seulement si il existe une fonction de distorsion continue f telle que, pour tous risques X et Y :

$$X \succeq Y \Leftrightarrow \int_0^1 f(G_X(t))dt \geq \int_0^1 f(G_Y(t))dt.$$

La fonction f vérifie la définition suivante :

DÉFINITION 1.6 (FONCTION DE DISTORSION)

Une fonction de distorsion g est une fonction croissante de $[0,1]$ dans $[0,1]$, telle que $g(0) = 0$ et $g(1) = 1$.

1.3.2 Mesures de Wang et valeur en risque

Cette théorie duale de celle de l'utilité espérée a une traduction immédiate en termes de mesures de risque :

DÉFINITION 1.7 (MESURE DE RISQUE DE WANG)

À toute fonction de distorsion g , on associe la mesure de risque de Wang – supposée bien définie – donnée par :

$$\rho_g[X] = - \int_{-\infty}^0 [1 - g(G_X(x))]dx + \int_0^{\infty} g(G_X(x))dx. \quad (1.3)$$

Cette définition appelle les remarques suivantes :

- Lorsque la fonction de distorsion est l'identité, l'expression précédente correspond à l'espérance de X ;
- Si $\forall p \in [0,1], g(p) \geq p$ alors la mesure de risque de Wang comporte un chargement de sécurité, c'est-à-dire que $\rho_g[X] \geq \mathbb{E}[X]$;
- enfin, si $\forall p \in [0,1], g_1(p) \leq g_2(p)$ alors $\rho_{g_1}[X] \leq \rho_{g_2}[X]$.

Une autre représentation

Il est possible d'exprimer toute mesure de risque de Wang sous forme de mélange de *VaR*.

$$\rho_g[X] = \int_0^1 Q_p[X]dg(p).$$

Et en considérant la fonction de distorsion $g_\alpha : p \longrightarrow \mathbb{1}_{\{p \geq 1-\alpha\}}$, on obtient :

$$\rho_{g_\alpha}[X] = VaR_\alpha[X].$$

De la même manière il est possible de représenter de nombreuses mesures de risque sous forme de mesure de Wang.

Exemples de mesures de Wang

La TVaR La *TVaR* peut s'écrire sous forme de mesure de risque de Wang. En effet, en considérant la fonction de distorsion :

$$f_\alpha = \min \left\{ \frac{x}{1-\alpha}, 1 \right\}$$

pour $\alpha \in [0,1]$ fixé, on obtient :

$$\rho_{f_\alpha}[X] = \frac{1}{1-\alpha} \int_0^{1-\alpha} VaR_{1-p}[X] dp = TVaR_\alpha[X].$$

Mesure de risque *Dual-power* La mesure de risque *Dual-power* est obtenue en considérant la fonction de distorsion :

$$\forall \xi \in [0,1], g_\xi(x) = 1 - (1-x)^\xi.$$

Nous obtenons alors la mesure de Wang :

$$\rho_{g_\xi}[X] = \int (1 - [F_X(x)]^\xi) dx.$$

Lorsque ξ est entier, $\rho_{g_\xi}[X]$ s'interprète comme la valeur attendue du maximum $M_\xi = \max\{X_1, \dots, X_\xi\}$ de ξ variables aléatoires indépendantes et identiquement distribuées. En effet, on a :

$$\mathbb{P}[M_\xi > x] = 1 - \mathbb{P}[X_1 \leq x, \dots, X_\xi \leq x] = 1 - [F_X(x)]^\xi$$

et donc

$$\rho_{g_\xi}[X] = \mathbb{E}[M_\xi].$$

Mesure de risque PH En considérant,

$$\forall \xi \geq 1, g_\xi(x) = x^{\frac{1}{\xi}}$$

nous obtenons la mesure de risque PH :

$$\rho_{g_\xi}[X] = \int [G_X(x)]^\xi dx$$

qui est bien une mesure de Wang. Et pour $\xi = 1$, $\rho_{g_1}[X] = \mathbb{E}[X]$.

Mesure de risque et transformée de Wang

DÉFINITION 1.8 (TRANSFORMÉE DE WANG)

Nous appelons transformée de Wang au niveau p (pour $p \in]0,1[$) la fonction de distorsion définie par :

$$\forall x \in [0,1], g_p(x) = \Phi(\Phi^{-1}(x) + \Phi^{-1}(p))$$

avec Φ fonction de répartition de la loi normale centrée réduite.

DÉFINITION 1.9 (WANG TRANSFORM RISK MEASURE)

En considérant les fonctions de distorsion précédentes, cette mesure est définie par :

$$WT_p[X] = \rho_{g_p}[X].$$

PROPOSITION 1.5

Lorsque la variable X est gaussienne, on obtient :

$$1 - g_p(\bar{F}_X(x)) = \Phi \left[\frac{x - Q_p(X)}{\sigma} \right]$$

et donc

$$WT_p[X] = VaR_p[X].$$

En revanche, on peut montrer que l'ES et la CTE ne sont pas des mesures de risque de Wang.

Propriétés générales

Les mesures de risques de Wang sont invariantes par translation, positivement homogènes, monotones et additives pour les sommes de risque comonotones, i.e. :

Additivité pour des risques comonotones Pour tous risques X_1, \dots, X_n comonotones, $\rho[\sum_{i=1}^n X_i^c] = \sum_{i=1}^n \rho[X_i]$.

Fonction de distorsion concave Lorsque la fonction de distorsion est concave (le risque est représenté par une perte) alors la mesure de risque est sous-additive. Dans ce cas là, la mesure de risque de Wang est une mesure de risque cohérente. Néanmoins, il existe des mesures de risque cohérentes qui ne sont pas des mesures de risques de Wang.

Il existe donc une grande variété de théories permettant de mesurer le risque, chacune d'entre elles permettant d'expliquer un aspect des comportements observés chez les agents. En pratique, il est cependant difficile de mesurer les paramètres d'aversion au risque des agents, ou en d'autres termes de donner leur degré de tolérance au risque. Les cadres théoriques présentés ici ne sont utilisés en pratique en finance qu'à travers trois représentants : la variance, la VaR et la Tail-VaR. Ces mesures sont utilisées en particulier pour l'allocation de capital, dans le cas par exemple de la gestion de portefeuille.

Chapitre 2

Allocation de capital

Nous avons vu dans la partie précédente comment pouvait être évalué le risque, notamment lorsqu'il s'agit d'investissement financier. Dans cette partie, nous allons nous intéresser à la sélection de portefeuille de risques financiers et à l'allocation de capital économique. Ces problèmes sont liés à la gestion optimale de capital et à la recherche du meilleur investissement possible, selon certains critères que nous préciserons.

2.1 Investissement et gestion de capital

2.1.1 Variables d'intérêt

Dans tout ce chapitre, nous considérons un agent a et comme variable primordiale sa richesse W . Nous nous plaçons dans un cadre unipériodique $[0, T]$, W^0 est alors la richesse initiale de a qui est positive ou nulle et W^T sa richesse en fin de période. Dans la suite nous utiliserons la lettre t pour désigner indifféremment 0 ou T .

L'agent a peut investir dans $n + 1$ « activités », au sens large, auxquelles sont associées des facteurs de risque $(S_i)_{0 \leq i \leq n}$, éventuellement négatifs. Pour simplifier l'approche, nous supposons que ces facteurs représentent la valeur en argent des activités. En général ces facteurs de risque sont aléatoires dans le sens où leurs valeurs sont connues à l'instant 0 mais indéterminées en t , date de fin de période. Toutefois, nous nous autorisons par la suite à considérer des actifs représentant des emprunts sans risque, c'est-à-dire dont le taux de rémunération du capital est connu en début de période. Donnons quelques exemples d'activités ou d'actifs que l'on peut considérer : cash, actions, produits dérivés, portefeuilles, contrat d'assurance, filiales d'entreprises, etc. De manière générale, il s'agit de toute activité pouvant générer des pertes ou des revenus.

La grandeur d'intérêt est ici est la richesse totale à l'instant t : W^t , qui

résulte des choix investissements. Remarquons que dans tout investissement, il y a recherche d'un équilibre entre réalisation de profit (but de l'investissement) et la limitation des risques (mise en place de garanties).

Nous considérons donc une fonction *pay-off* de l'investissement dans l'activité i : Ψ_i . Cette fonction dépend évidemment de la part de la richesse investie w_i que nous séparons, pour simplifier, en deux types : q_i , lié à la recherche du profit, et k_i , garantie de l'investissement. Dans les exemples, nous verrons que bien souvent q_i donne la proportion de richesse investie dans un actif risqué et k_i peut s'interpréter comme un montant de fond propre. Exactement, nous considérons des fonctions d'investissement ayant la forme suivante : $\Psi_i(q_i, k_i, S_i^0, S_i^T)$. Généralement, les Ψ_i et S_i^0 sont connues, les S_i^T aléatoires, les q_i et k_i sont à déterminer.

La richesse totale W^T a la forme simple suivante :

$$W^T(\mathbf{q}, \mathbf{k}, \mathbf{S}_i^0, \mathbf{S}_i^T) = \sum_{i=1}^n \Psi_i(q_i, k_i, S_i^0, S_i^T),$$

qui exprime que la richesse totale est la somme des richesses accumulées dans chaque investissement.

La contrainte naturelle qui s'impose est la contrainte de richesse en début de période, i.e.

$$W^0 = \sum_{i=1}^n \Psi_i(q_i, k_i, S_i^0, S_i^0).$$

Selon l'investissement, Ψ_i peut prendre les formes suivantes :

- dans un portefeuille : $\Psi_i = \frac{q_i}{S_i^0} S_i^t$;
- dans un portefeuille avec coûts de transaction : $\Psi_i = a \frac{q_i}{S_i^0} S_i^t$, ($a < 1$) ;
- investissement avec garanties : $\Psi_i = g(q_i, S_i^0) S_i^t + k_i$;
- investissement avec cloisonnement des pertes : $\Psi_i = (\frac{q_i}{S_i^0} S_i^t + k_i)_+$, S_i^t est un gain ;
- investissement avec protection : $\Psi_i = (\frac{q_i}{S_i^0} S_i^t + k_i)_+ - c_i(k_i)$, c_i fonction de coût de la protection.

Nous traitons surtout dans la suite du premier problème, où Ψ à une forme simple et est linéaire en les q_i .

2.1.2 Les programmes d'optimisation

À l'agent a sont associées :

- une fonction « objectif » $\theta(W^t, W^0)$: par exemple une espérance d'utilité de la richesse, ou un rendement espéré ;

- une fonction « mesure de risque » $\rho(W^t)$, à laquelle nous associons plus loin une fonction de risque.

L'agent cherche à allouer son capital de manière juste. Pour cela, il utilise un programme d'optimisation parmi les deux suivants :

Maximisation de la fonction objectif

$$\boxed{\max \theta(W^t, w^0) \quad \text{s.c.} \quad \rho(W^t) \leq r}$$

dans ce cas la contrainte porte sur le niveau de risque.

Minimisation de la fonction de risque

$$\boxed{\min \rho(W^t) \quad \text{s.c.} \quad \theta(W^t, W^0) \geq l}$$

dans ce cas la contrainte porte sur l'objectif.

Ces programmes sont résolus avec la prise en compte de la contrainte de richesse. De plus, il est possible d'imposer d'autres contraintes sur les paramètres : certains peuvent être fixés, dans une plage de valeur ou positifs (pour représenter l'impossibilité de vente à découvert).

2.2 Sélection de portefeuille : le cadre du modèle

2.2.1 Description

Nous considérons un marché et un agent ayant les caractéristiques suivantes :

- une période, deux dates $t \in \{0, T\}$;
- $(S_i)_{0 \leq i \leq n}$ $n + 1$ actifs, variables aléatoires réelles intégrables, supposées linéairement indépendantes et de valeur à la date $t = 0$ strictement positive ;
- une richesse initiale $W^0 > 0$.

\mathcal{P} est l'ensemble des portefeuilles P réalisés par l'achat ou la vente d'une certaine quantité des actifs de base disponibles. En notant q_i ces quantités, \mathcal{P} s'identifie à l'ensemble des vecteurs $\mathbf{q} = (q_i)_{0 \leq i \leq n}$, i.e. \mathbb{R}^{n+1} . Mais on considère, $w_i = q_i S_i^0$, la richesse investie dans l'actifs i et on choisit de décrire \mathcal{P} par l'ensemble des vecteurs $\mathbf{w} = (w_i)_{0 \leq i \leq n}$.

Nous nous intéressons à la valeur du portefeuille :

$$V_P^t(\mathbf{w}) = \sum_{i=0}^n q_i S_i^t = \sum_{i=0}^n \frac{w_i}{S_i^0} S_i^t$$

ainsi qu'au rendement par rapport à la richesse W^0 :

$$r_P^t(\mathbf{a}, W^0) = \sum_{i=0}^n a_i r_i^t$$

avec $a_i = w_i/W^0$ la richesse investie en i , exprimée en proportion de la richesse initiale et $r_i^t = S_i^t/S_i^0$ la valeur de l'actif i en t par rapport à sa valeur initiale.

2.2.2 Fonction objectif et fonction de risque

Avant de passer à l'écriture puis à la résolution du problème d'optimisation, il nous faut choisir une fonction objectif et une mesure de risque, à laquelle va correspondre une fonction de risque.

Fonction objectif

La fonction objectif g choisie est l'espérance des rendements en $t = T$, c'est une fonction linéaire :

$$g(\mathbf{a}) = \sum_{i=0}^n a_i \mathbb{E}[r_i^T].$$

Nous ne prenons pas en compte ici une espérance d'utilité, qui inclurait une mesure du risque subjective.

Fonction de risque

Tout d'abord, nous considérons ρ une mesure de risque positivement homogène, invariante pour le *cash*, par exemple une mesure de risque cohérente. Nous nous plaçons dans le cadre de la partie précédente où la mesure de risque porte sur des pertes. Lorsque la mesure de risque s'applique à un rendement, sa valeur représente aussi un rendement.

Mais, il nous faut considérer aussi comme mesure de risque, la variance ou l'écart-type σ , qui ont été traditionnellement utilisés dans la gestion de portefeuille [23].

A partir de la mesure de risque, nous introduisons la fonction de risque associée, qui dépend aussi des actifs de base :

$$f_{\rho, \mathbf{S}}(\mathbf{a}) = \rho[-r_P^T(\mathbf{a}, W^0)] \mathbb{1}_{\{\rho[-r_P^T(\mathbf{a}, W^0)] \geq 0\}}.$$

Nous la notons désormais simplement f .

2.2.3 Propriétés de la fonction de risque

DÉFINITION 2.1 (*n*-HOMOGENÉITÉ)

f est *n*-homogène si

$$\forall \lambda \geq 0, f(\lambda \mathbf{a}) = \lambda^n f(\mathbf{a}).$$

$n = 1$ correspond à une mesure de risque monétaire ou l'écart-type, $n = 2$ correspond à la variance. Nous supposons désormais que f est 1-homogène. Remarquons que la fonction objectif l'est aussi, puisqu'elle est linéaire.

Hypothèses supplémentaires sur f

L'exemple de la VaR montre que les mesures de risque considérées ne sont pas toujours continues, ni sous-additives. La résolution des programmes d'optimisation dans le cas général peut poser des difficultés. Nous sommes donc amenés à supposer certaines propriétés pour f .

H0 : La fonction f est semi-continue inférieurement i.e

$$\forall x \in \mathbb{R}^{n+1}, \liminf_{x \rightarrow x'} f(x') \geq f(x),$$

l'inégalité étant prise dans $\mathbb{R} \cup +\infty$. De manière équivalente :

$$\forall \alpha \in \mathbb{R}, \{x \in \mathbb{R}^{n+1} \mid f(x) \leq \alpha\} \text{ est fermé.}$$

Nous employons par la suite l'abréviation s.c.i si nécessaire.

H1 : f est sous-additive.

Cette hypothèse est particulièrement intéressante puisque la propriété est vraie pour les mesures de risque cohérentes et les mesures de Wang convexes. De plus, f étant 1-homogène, **H1** implique la convexité de f .

Propriétés de f convexe semi-continue inférieurement

Nous définissons le sous-différentiel de f et donnons ses propriétés.

DÉFINITION 2.2 (SOUS-DIFFÉRENTIEL)

Le sous-différentiel de f en x , noté $\partial f(x)$ est l'ensemble des vecteurs s tels que :

$$\forall y \in \mathbb{R}^{n+1}, f(y) \geq f(x) + \langle s, y - x \rangle$$

ce qui équivaut à :

$$\forall h \in \mathbb{R}^{n+1}, f(x + h) \geq f(x) + \langle s, h \rangle .$$

En particulier, pour $h = x$, on obtient :

$$f(2x) \geq f(x) + \langle s, x \rangle$$

soit,

PROPOSITION 2.1

$$\forall s \in \partial f(x), f(x) \geq \langle s, x \rangle .$$

De plus,

PROPOSITION 2.2

Le sous-différentiel de f en x est constant sur la demi-droite ouverte $]0, x)$:

$$\forall \lambda > 0, \text{ si } s \in \partial f(\lambda x) \text{ alors } s \in \partial f(x).$$

Ces propriétés peuvent être affinées sous l'hypothèse :

H2 : f est différentiable.

En effet, d'après le théorème d'Euler, nous avons :

$$\forall x \in \mathbb{R}^{n+1}, \langle \nabla f(x), x \rangle = pf(x). \quad (2.1)$$

De plus $\nabla f(x)$ est constant sur $]0, x)$ i.e. « 0-homogène ».

Dans ce cadre (pour f p -homogène et différentiable), le théorème d'Euler fournit :

THÉORÈME 2.1 (EULER)

$$\forall x \in \mathbb{R}^{n+1}, pf(x) = \sum_{i=0}^n x_i \partial_i f(x) = \langle \nabla f(x), x \rangle .$$

La quantité $x_i \partial_i f(x)$ apparaît comme une contribution de la i -ème variable à la valeur $f(x)$. Dans la suite, nous parlons de contribution au risque.

Nous définissons aussi les *sensibilités* suivantes :

$$\beta_i(x, y) = \frac{\partial_i f(x)}{pf(x)}.$$

Enfin, une dernière propriété intéressante que peut posséder la fonction f est :

H3 : La fonction f est strictement convexe.

En effet, cette propriété est très utile pour obtenir des résultats d'unicité.

2.2.4 Étude des contraintes

Les contraintes pour les programmes d'optimisation sont représentées par un ensemble \mathcal{S} qui est toujours supposé fermé. Ces contraintes sont le plus souvent affines ou convexes. De plus, nous prenons en compte la plupart du temps la contrainte naturelle de richesse :

$$V_P^0 = W^0$$

ou encore $\sum_{i=0}^n a_i = 1$.

2.3 Résolution des programmes

La détermination d'une allocation optimale de richesse passe par la résolution d'un programme d'optimisation. Ce programme a la forme suivante :

$$\boxed{\min_{x \in \mathcal{S}} f(x)} \quad (2.2)$$

où \mathcal{S} est l'ensemble des contraintes et f 1-homogène (la cas de la variance sera traité à part). Nous allons tout d'abord chercher à résoudre ce programme sous des hypothèses assez faibles.

2.3.1 Étude du programme général

Nous faisons pour f l'hypothèse **H0** i.e. f est semi-continue inférieurement. On note $K_0 = f^{-1}(\{0\})$, comme f est 1-homogène c'est un cône fermé. Soit $I = K_0 \cap \mathcal{S}$, s'il est non vide alors l'ensemble des solutions est I . Nous supposons donc dorénavant que I est vide.

Nous supposons de plus que \mathcal{S} est borné. En effet, si \mathcal{S} est non borné, comme

$$f(x) \rightarrow +\infty, \|x\| \rightarrow +\infty$$

pour $x \in \mathcal{S}$, il existe x_0 et $r \geq x_0$ tels que :

$$\forall x \in \mathcal{S}, \|x\| > r, f(x) > f(x_0).$$

Donc minimiser f sur \mathcal{S} revient à minimiser f sur $\mathcal{S} \cap \bar{B}(0, r)$.

Cas où f est s.c.i.

PROPOSITION 2.3

Sous les hypothèses précédentes, le programme d'optimisation a au moins une solution.

DÉMONSTRATION 2.1

Par hypothèse, \mathcal{S} est fermé borné dans \mathbb{R}^{n+1} donc compact.

On note $a = \inf\{f(x) | x \in \mathcal{S}\}$, $a \geq 0$. Soit (x_k) une suite d'élément de \mathcal{S} minimisant f , i.e $\lim f(x_k)_{x_k \rightarrow +\infty} = a$.

Par compacité de \mathcal{S} , il existe une sous-suite convergente (x_k^e) dans \mathcal{S} vers x^* . Et comme f est semi-continue inférieurement,

$$\liminf_{x_k^e \rightarrow x^*} f(x_k^e) \geq f(x^*).$$

Donc $a \geq f(x^*)$ et par définition de a , $f(x^*) \geq a$.

PROPOSITION 2.4

L'ensemble de ces solutions est situé sur la frontière de \mathcal{S} .

DÉMONSTRATION 2.2

En effet, si x est dans l'intérieur de \mathcal{S} , alors il existe $0 < \eta < \|x^*\|$ tel que $B(x, \eta)$ soit contenue dans \mathcal{S} . Donc pour $\alpha = 1 - \frac{\eta}{\|x^*\|}$, $0 < \alpha < 1$, on pose $y = \alpha x^*$ et par 1-homogénéité de f , $f(x^*) > f(y)$. Absurde.

En fait, on prouve directement que $x^* \in F$ avec

$$F = \{\lambda_x x \in \mathcal{S} | \lambda_x = \min\{\lambda | x \in \mathcal{S}, \lambda x \in \mathcal{S}\}\}.$$

Cas où f et \mathcal{S} sont convexes**DÉFINITION 2.3 (DIRECTION TANGENTE)**

Soit $x \in \mathcal{S}$ (fermé, convexe), $d \in \mathbb{R}^{n+1}$ est une direction tangente à \mathcal{S} en x lorsqu'il existe une suite (x_k) d'éléments de \mathcal{S} convergeant vers x et une suite (t_k) de réels strictement positifs tels que $\frac{x_k - x}{t_k} \rightarrow d$.

L'ensemble de ces directions est appelé cône tangent à \mathcal{S} en x , on le note $T(\mathcal{S}, x)$. Comme \mathcal{S} est un convexe fermé, ce cône est exactement l'adhérence du cône engendré par $\mathcal{S} - x$.

DÉFINITION 2.4 (DIRECTION NORMALE)

Une direction s est normale à \mathcal{S} en x lorsque :

$$\forall c \in \mathcal{S}, \langle s, c - x \rangle \leq 0.$$

L'ensemble de ces directions est le cône normal à \mathcal{S} en x , on le note $N(\mathcal{S}, x)$. Puisque \mathcal{S} est un convexe fermé, c'est exactement le cône polaire de $T(\mathcal{S}, x)$.

PROPOSITION 2.5

Les minima de f sur \mathcal{S} sont exactement les x pour lesquels $-\partial f(x)$ est contenu dans $N(\mathcal{S}, x)$.

DÉMONSTRATION 2.3

Si $s \in \partial f(x)$, alors en particulier :

$$\forall y \in \mathcal{S}, f(y) \geq f(x) - \langle -s, y - x \rangle.$$

L'équivalence est alors évidente.

Cas où f est strictement convexe

Il est alors possible d'obtenir le résultat :

PROPOSITION 2.6

L'ensemble des solutions du programme est un singleton.

DÉMONSTRATION 2.4

En effet, s'il existe x_1 et x_2 différents, minimums de f sur \mathcal{S} , la valeur de f au minimum étant m , alors par stricte convexité de f , pour $\lambda \in]0, 1[$, $y = \lambda x_1 + (1 - \lambda)x_2$ et :

$$f(y) < \lambda f(x_1) + (1 - \lambda)f(x_2) = m$$

soit $m < m$, absurde.

Application

Si on minimise le risque sous contrainte de richesse, l'ensemble \mathcal{S} est alors un hyperplan affine de \mathbb{R}^{n+1} , on le note \mathcal{S}_0 . Dès que f vérifie **H0**, le programme a au moins une solution (unique si f est strictement convexe). À cette solution, correspond une espérance de rendement r_0 et un niveau de risque k_0 minimum. On note par la suite $k_0 = f^*(r_0)$.

2.3.2 Minimisation du risque sous contrainte de rendement

En plus de la contrainte de richesse, il est intéressant pour le gestionnaire de portefeuille d'imposer une contrainte de rendement pour le portefeuille au niveau l strictement positif. L'ensemble des allocations réalisables est donc :

$$\mathcal{S} = \{a \in \mathbb{R}^{n+1} \mid \sum_{i=0}^{n+1} a_i = 1, g(a) \geq l\}$$

On suppose que \mathcal{S}_0 et l'hyperplan affine défini par $g^{-1}(0)$ ne sont pas parallèles. Ainsi, pour tout l , \mathcal{S} est non vide et de plus, convexe fermé (c'est un demi-espace affine de dimension n). Dès que f vérifie **H0**, ce programme a une solution. Au niveau de rendement l , on associe alors le risque optimal k , on note $k = f^*(l)$.

DÉFINITION 2.5 (FRONTIÈRE D'EFFICIENCE)

La courbe représentant la fonction f^* dans le plan « espérance de rendement - risque » s'appelle la frontière d'efficience du programme.

D'après les résultats de la partie précédente (l'ensemble contraint étant plus petit),

$$\forall l > 0, f^*(l) > f^*(r_0).$$

Il existe donc un risque minimum strictement positif pour les portefeuilles optimaux et seuls les portefeuilles de rendement supérieur à r_0 devraient être envisagés.

Cas où f est différentiable

Les contraintes étant affines, il est possible d'appliquer le théorème Karush–Kuhn–Tucker (cf. par exemple [20]) et donc d'écrire les conditions du premier ordre pour le programme. Soit L le lagrangien associé:

$$L(a, \lambda, \mu) = f(a) - \lambda \left(\sum_{i=0}^n a_i r_i - l \right) - \mu \left(\sum_{i=0}^n n a_i - 1 \right)$$

avec $r_i = \mathbb{E}[r_i^T]$, on a donc à l'optimum :

$$\forall i \in [0, n], \partial_i f(a^*) - \lambda r_i - \mu = 0$$

et en utilisant la propriété de 1-homogénéité de f :

$$f^*(l) - \lambda l - \mu = 0.$$

Et donc lorsque ces quantités sont bien définies pour $i, j \in [0, n]$:

$$\frac{\partial_i f(a^*) - f^*(l)}{\partial_j f(a^*) - f^*(l)} = \frac{r_i - l}{r_j - l}.$$

Ce qui lie les contributions au risque et les contributions au rendement des actifs à l'optimum.

2.3.3 En présence d'actif sans risque

Nous considérons maintenant que S_0 est un actif sans risque c'est à dire que son rendement est non aléatoire et donc connu en 0. Nous allons alors légèrement modifier le cadre de l'étude défini précédemment. Nous nous intéressons désormais aux écarts de rentabilité avec l'actif sans risque.

Écart de rentabilité avec l'actif sans risque

Le portefeuille sans risque est tel que tout le *cash* est investi dans l'actif sans risque, sa rentabilité est celle de l'actif sans risque i.e. Z_0 . « Sans risque » signifie ici « connu », i.e. non aléatoire. Attention : nous pouvons aussi considérer un portefeuille de valeur X tel que $\rho[X] \leq 0$, ce portefeuille sera aussi sans risque... au sens de ρ ! En notant $R = Z_0$ la rentabilité de l'actif sans risque, nous pouvons considérer l'écart de rentabilité avec l'actif sans risque :

$$Z_P = \sum_{i=1}^n a_i (Z_i - R).$$

Cette grandeur est intéressante car elle compare le portefeuille risqué qui a été choisi à celui sans risque.

Déterminons l'écart de rentabilité à l'actif sans risque μ lorsque l'investisseur décide de multiplier par un coefficient $\lambda > 0$ sa proportion de richesse placée dans les actifs risqués. Nous avons :

$$\Delta = \sum_{i=0}^n a_i W^0 Z_i$$

après rebalancement, nous obtenons :

$$a'_0 = 1 - \sum_{i=1}^n \lambda a_i$$

et

$$\Delta' = W^0 R + \lambda \sum_{i=1}^n a_i W^0 Z_i$$

soit

$$Z'_P = \lambda Z_P. \quad (2.3)$$

Ainsi lorsqu'il y a un actif non risqué et qu'on impose la contrainte de richesse, *seules les allocations en actifs risqués sont à prendre en compte dans les programmes d'optimisation.*

Fonction de risque et programme à résoudre

La fonction de risque est :

$$f(a) = \rho[-Z_P(a)] \mathbf{1}_{\{\rho[-Z_P(a)] \geq 0\}}.$$

Cette fonction d'après le point précédent est 1-homogène. La fonction objectif g considérée est l'espérance de l'écart de rentabilité du portefeuille avec celle de l'actif sans risque.

Nous cherchons à résoudre :

$$\begin{aligned} \min f(a) \\ \text{s.c. } g(a) \geq l. \end{aligned} \quad (2.4)$$

D'après les résultats montrés plus haut, dès que f vérifie **H0**, le programme a au moins une solution.

Résolution du programme avec H2

Dans le cas où f est différentiable (hypothèse **H2**), le lagrangien s'écrit :

$$L(a, \lambda) = f(a) - \lambda \left(\sum_{i=1}^n a_i m_i - l \right) \quad (2.5)$$

avec $\mathbb{E}[Z_i] = m_i$.

Les conditions du premier ordre (CPO) sont écrites au point solution, les valeurs des sensibilités en ce point sont notées β^* .

$$\forall i \in [1, n], \partial_i f(a^*) - \lambda m_i = 0.$$

Nous obtenons une expression du type « égalité des TMS », classique en théorie microéconomique :

$$\frac{\partial_i f}{\partial_j f} = \frac{m_i}{m_j} = \lambda.$$

De plus (via les CPO et contrainte de rendement),

$$\langle \nabla f(a^*), a^* \rangle = \lambda l.$$

Soit via le théorème d'Euler ,

$$\lambda = \frac{f^*(l)}{l}.$$

Et en introduisant les sensibilités, on obtient l'expression fondamentale:

$$m_i = \beta_i^* l,$$

qui lie, à l'optimum, l'espérance de rendement de chaque actif à l'espérance de rendement du portefeuille.

Nous pouvons de plus donner la forme de la frontière d'efficience. Soit $a(l)$, une allocation optimale correspondant au niveau de rendement l . On vérifie facilement que $\theta a(l)$ est solution du programme associé au niveau de rendement θl et donc $f^*(\theta l) = \theta f^*(l)$. Ainsi,

THÉORÈME 2.2 (FRONTIÈRE D'EFFICIENCE)

La frontière d'efficience est une droite.

2.3.4 Quelques exemples

Cas de la variance

Le cas où la mesure de risque est la variance constitue la référence pour les théories de sélection de portefeuille. Ce cas a été étudié par Markowitz [23] et Sharpe [30], entre autres. [23] considère le cas avec contraintes sur la vente d'actifs, notre étude, elle, ne prend pas en compte ce genre de contraintes.

Cette étude suppose implicitement que toutes les variables sont de carré intégrable. Dans le cas de risques extrêmes, il est connu que cette condition n'est pas toujours satisfaite. Cette limitation peut être contournée si l'on travaille avec d'autres mesures de risque (cf. section suivante).

Pour la variance, la frontière d'efficience est une parabole car la variance est 2-homogène. Nous avons :

$$\begin{aligned} \mathbb{V}[Z_p] &= \mathbb{V}[\alpha_1 Z_1 + \alpha_2 Z_2] \\ \mathbb{V}[Z_p] &= \alpha_1^2 \sigma_1^2 + 2\alpha_1 \alpha_2 \sigma_1 \sigma_2 \rho + \alpha_2^2 \sigma_2^2 \end{aligned}$$

avec $\sigma_i^2 = \mathbb{V}[Z_i]$ et $\text{Cov}[Z_1, Z_2] = \sigma_1 \sigma_2 \rho$.

Comme $\alpha_2 = (r - \alpha_1 m_1)/m_2$ en raison de la contrainte de rendement, nous obtenons un trinôme en α_1 pour la variance :

$$\mathbb{V}[Z_p] = \alpha_1^2 \sigma_1^2 (1 - 2\rho\theta + \theta^2) + 2\alpha_1 r \frac{\sigma_1^2 \theta}{m_1} (\rho - \theta) + r^2 \frac{\sigma_1^2}{m_1^2} \theta^2 \quad (2.6)$$

avec

$$\theta = \frac{m_1 \sigma_2}{m_2 \sigma_1}. \quad (2.7)$$

La condition du premier ordre $\partial_{\alpha_1} \mathbb{V}[Z_p] = 0$ nous donne :

$$\alpha_1^* = \frac{r}{m_1} \frac{\theta^2 - \rho\theta}{1 - \rho\theta + \theta^2} \quad (2.8)$$

puis

$$\alpha_2^* = \frac{r}{m_2} \frac{1 - \rho\theta}{1 - \rho\theta + \theta^2}. \quad (2.9)$$

Les expressions ci-dessus sont bien définies, car le dénominateur ne s'annule jamais pour $\rho \in [-1, 1]$

En dérivant l'expression $\mathbb{V}[Z_p]$, on obtient l'expression des sensibilités et la formule :

$$\mathbb{E}[Z_i] = \mathbb{E}[Z_p] \frac{\text{Cov}[Z_p, Z_i]}{\mathbb{V}[Z_p]}$$

Dans le cadre gaussien multivarié, il est possible d'interpréter les sensibilités comme la régression linéaire de la rentabilité des actifs sur celle du portefeuille optimal. Ces résultats sont exposés par exemple dans [18]. Dans ce même ouvrage, on retrouvera aussi les procédures d'estimation des paramètres du modèle ainsi que du ratio de Sharpe.

Ce modèle, étendu dans le cadre de la théorie de l'utilité, permet d'obtenir des modèles généraux de marché, par exemple, le *Capital Asset Pricing Model*, et est donc très intéressant. Pourtant, comme tout modèle, il a quelques limites. La plus grande critique est liée au caractère réducteur de la variance, qui, entre autre, ne permet pas de prendre en compte l'asymétrie des distributions et la particularité des queues épaisses. Dans certains cas, ce modèle est même non applicable. Ce modèle est évidemment parfait dans le cadre elliptique (voir page 45), or, il est connu que les rendements des actifs financiers ne respectent pas cette hypothèse.

Cas de la VaR

La VaR à un niveau α permet justement de prendre en compte les situations oubliées par la variance. Elle permet de distinguer le cas des distributions asymétriques et est toujours définie. L'étude de la sélection de portefeuille dans ce cadre, cf. [4], [19], a suscité un grand intérêt, d'autant plus que cette mesure devenait populaire dans les milieux financiers.

Pour autant, la VaR, en plus d'être critiquable comme mesure de risque, pose quelques problèmes techniques (pour une étude fine, cf. [11]). Par exemple, si les distributions des variables aléatoires ne sont pas continues, sa définition nécessite l'utilisation de l'inverse généralisée. De plus, la VaR_α – et alors la

fonction f – ne possède pas toujours la propriété de convexité. Cela entraîne des difficultés dans la résolution des programmes d'optimisation.

EXEMPLE 2.1

Soit X une variable aléatoire, on considère les deux pertes suivantes: $A = \mathbf{1}_{\{X>U\}}$ et $B = \mathbf{1}_{\{X<L\}}$. On choisit L et U tel que $\mathbb{P}[X > U] = \mathbb{P}[X < L] = 0.008$. Nous avons alors que $f_{A,B}$ n'est pas continue, en $(1,0)$ par exemple.

Cet exemple montre qu'il faut être très prudent lorsque l'on travaille avec la VaR et que les hypothèses intéressantes pour la résolution des programmes ne sont pas toujours vérifiées. Cela justifie aussi le fait que l'on tente d'estimer les dérivées première et seconde de la VaR_α du portefeuille en fonction de la quantité d'actif le composant. [17] et [31] obtiennent comme expression pour la sensibilité :

$$\beta_i = -\frac{1}{VaR_\alpha[Z_p]} \mathbb{E}[X_i | Z_p = -VaR_\alpha[Z_p]]. \quad (2.10)$$

Comme explicité dans l'introduction générale, la VaR étant positivement homogène, il vient que la frontière d'efficacité pour cette mesure de risque est une droite.

Au contraire de la variance, il est rarement possible (cf. point suivant) d'obtenir une expression analytique de l'allocation optimale. L'estimation de la VaR , obtenue par l'estimation des quantiles de la distribution des valeurs portefeuilles, et le calcul de l'allocation optimale nécessite des méthodes numériques. Le programme d'optimisation peut être résolu de deux manières, soit par une discrétisation assez fine de l'ensemble des portefeuilles possibles et une simple recherche du minimum de cet ensemble, soit par des méthodes de descente de gradient dès que les dérivées première et seconde le permettent. Ce dernier cas est discuté dans [17].

Mesure de risque de Wang

Lorsque nous avons défini ces mesures de risque en première partie, nous avons aussi montré que ces mesures sont sous-additives. Dans le cadre de notre étude, cela implique que la fonction f a la « bonne » propriété d'être convexe. Le programme de minimisation du risque sous contrainte de rendement a donc une solution, qui est unique dès que f est strictement convexe. De plus, ces mesures étant positivement homogène, nous obtenons que la frontière d'efficacité dans le cas de mesure de risque de Wang est une droite.

Le cas de la TVaR

La TVaR est une mesure de Wang couramment utilisée. [27] et [10] se sont intéressés à la résolution du programme de minimisation du risque à rendement fixé pour cette mesure.

Là encore, l'obtention de l'allocation optimale n'est pas aisée dans le cas général. Il faut avoir recours à des méthodes numériques. La première, de base, est de discrétiser l'espace des portefeuilles et de chercher le minimum sur cet ensemble par des méthodes de tri. Il est aussi possible d'utiliser des méthodes de gradient lorsque la dérivée première et la dérivée seconde de la mesure de risque par rapport aux quantités d'actifs sont bien définies. Une dernière méthode, proposée dans [3], est d'utiliser la régression sur les quantiles.

2.3.5 Maximisation de l'espérance de rendement

Il est possible de considérer le programme suivant, sous contrainte de richesse :

$$\max g(a) \quad (2.11)$$

$$s.c. f(a) \leq k \quad (2.12)$$

g étant la fonction objectif choisie précédemment, qui est affine.

Dès que f est semi-continue inférieurement, $\mathcal{S} = f^{-1}([0, k])$ est fermé. Il est alors possible d'appliquer les résultats précédents et de montrer que le programme a une solution.

Lorsque f est convexe, la théorie de l'optimisation et analyse convexe (cf. [20]) permet de montrer que la résolution des programmes de minimisation du risque ou de maximisation du rendement est équivalente.

2.3.6 Application à la réassurance proportionnelle

Cadre du modèle : Soient n risques (X_i) auxquels on associe les primes π_i , on considère $Y_i = \pi_i + X_i$. On suppose que $\mathbb{E}[Y_i] > 0$, c'est à dire que la prime calculée contient un chargement de sécurité. Pour chaque risque, l'assureur a la possibilité de se réassurer proportionnellement et d'en garder une proportion α . La valeur du portefeuille de risque après réassurance s'écrit :

$$V_P(\alpha) = \sum_{i=1}^n \alpha_i Y_i.$$

L'ensemble des allocations possibles est :

$$\mathcal{S} = \{\alpha \in R^n | \forall i \in [1, n], \alpha_i \in [0, 1]\}.$$

Cette ensemble est convexe fermé.

Fonction de risque Comme précédemment, on utilise :

$$f(a) = \rho[-V_P(\alpha)] \mathbb{1}_{\{\rho[-V_P(\alpha)] > 0\}}.$$

Programme d'optimisation L'assureur cherche à optimiser sa réassurance, pour cela il résoud le programme suivant :

$$\max \mathbb{E}[V_p(\alpha)] \quad (2.13)$$

$$\text{s.c. } f(\alpha) \leq k. \quad (2.14)$$

Dès que f est semi-continue inférieurement, ce programme a une solution.

Il est donc possible pour l'assureur de se réassurer de manière optimale.

2.4 Sélection de portefeuille et allocation de fonds propres

Lors de la sélection optimale de portefeuille, un niveau de rendement ou un niveau de risque doit être fixé. Le gestionnaire de portefeuille peut alors déterminer son allocation en fonction de la mesure de risque utilisée. Le gestionnaire de risque lui doit généralement répondre à des contraintes légales et notamment constituer un montant minimum de fonds propres. Comme il a été fait en partie introductive, nous considérons une richesse initiale W_0 strictement positive. Une partie Q de cette richesse est assignée à l'investissement dans un portefeuille d'actifs, l'autre ($K > 0$) est destiné à couvrir cette investissement. On a donc :

$$W^0 = Q + K.$$

et la richesse finale est :

$$W^T = \Pi(Q, k) + (1 + \theta)K(Q, k).$$

où k représente un niveau de risque, calculé sur le rendement du portefeuille et donné pour la résolution du programme d'allocation (en présence d'actif sans risque) ; $\Pi(Q, k)$ représente cette solution et θ le niveau de rémunération des fonds propres.

La contrainte du gestionnaire se traduit elle par : $\rho[-W^T] = 0$. On doit donc avoir $K(Q, k) = \rho[-\Pi(Q, k)]$ soit :

$$K(Q, k) = Q \frac{k - (1 + R)}{1 + \theta}.$$

et donc la richesse à allouer pour l'investissement est :

$$Q = W_0 \frac{1}{1 + \frac{k - (1 + R)}{1 + \theta}}.$$

Dans la section suivante, nous présentons une méthode d'allocation de capital économique, qui est fondée sur la notion de champ de vecteurs adapté à la mesure de performance. La référence principale concernant cette méthode est l'article de Tasche [32], ou encore [31]. D'autres approches existent, telles celle de Denault [7] qui s'appuie sur la théorie des jeux coopératifs ou celle de Fischer [13].

2.5 Allocation de capital économique

Ce problème est un cas particulier de celui que nous avons examiné précédemment. Cependant, nous ne raisonnerons pas en termes de retour sur investissements, mais en termes de valeur des actifs.

Notre problème consiste à déterminer la part du risque total qui est apportée par chacun des actifs. En effet, les mesures de risque que nous avons considérées jusqu'ici permettent de déterminer le risque d'un portefeuille. Mais pour des raisons diverses – comme l'évaluation de la rentabilité des investissements passés ou les décisions d'investissement futurs – il est important de connaître la contribution au risque de chacun des actifs dans le portefeuille.

Une idée naïve serait de prendre la mesure de risque de l'actif seul : cette approche est vouée à l'échec, les corrélations entre actifs ayant un impact important sur le risque total. L'approche consistant à enlever l'actif considéré et à regarder la différence avec le risque total est elle aussi peu cohérente, le risque total dépendant du portefeuille tout entier. Nous allons donc utiliser une autre approche, prenant en compte tout le portefeuille à la fois. Les preuves des théorèmes peuvent être trouvées dans [32].

2.5.1 Définitions

Nous avons d actifs, de valeur positive ou négative. Le flux financier C_i d'un actif i se décompose comme la différence de :

- l'espérance de ce cash-flow m_i
- l'opposé de la variation par rapport à la moyenne : X_i , appelé fluctuation.

On peut donc écrire, X_i étant négatif si l'on a enregistré un gain par rapport à la moyenne et positif sinon :

$$C_i = m_i - X_i. \quad (2.15)$$

Pour un portefeuille, avec des quantités détenues u_i en chacun des actifs, cela donne :

$$C(u) = \sum_{i=1}^d m_i u_i - \sum_{i=1}^d u_i X_i. \quad (2.16)$$

La quantité qui va nous intéresser dans la suite est la fluctuation de la valeur du portefeuille

$$Z(u) = \sum_{i=1}^d u_i X_i. \quad (2.17)$$

2.5.2 Contributions au risque et différentiabilité des mesures de risque

Nous allons définir ce qu'est un champ de vecteurs *adapté à la mesure de performance*, ce qui nous donnera ensuite directement une expression des contributions au risque. Pour cela, nous avons besoin de la notion de fonction de retour sur investissement, analogue aux RORAC, RARORAC et autres indicateurs de rendement corrigés du risque :

DÉFINITION 2.6

Soit r une fonction d'un ouvert non vide U de \mathbb{R}^d dans \mathbb{R} . Soit $m \in \mathbb{R}^d$. Nous appelons alors fonction de retour la fonction $g = g_{r,m}$:

$$\{u \in U : r(u) \neq m'u\} \rightarrow \mathbb{R} \quad (2.18)$$

$$u \mapsto \frac{m'u}{r(u) - m'u}. \quad (2.19)$$

Un champ de vecteurs est qualifié d'adapté à la mesure de performance s'il vérifie les deux conditions suivantes :

1. Pour tout $m \in \mathbb{R}^d, u \in U$ tels que $r(u) \neq m'u$ et $i \in N_d$, l'inégalité :

$$m_i r(u) > a_i(u) m'u \quad (2.20)$$

implique qu'il existe un $\epsilon > 0$ tel que pour tout $t \in (0, \epsilon)$ on a :

$$g_{r,m}(u - te_i) < g_{r,m}(u) < g_{r,m}(u + te_i). \quad (2.21)$$

2. Pour tout $m \in \mathbb{R}^d, u \in U$ tels que $r(u) \neq m'u$ et $i \in N_d$, l'inégalité :

$$m_i r(u) < a_i(u) m'u \quad (2.22)$$

implique qu'il existe un $\epsilon > 0$ tel que pour tout $t \in (0, \epsilon)$ on a :

$$g_{r,m}(u - te_i) > g_{r,m}(u) > g_{r,m}(u + te_i). \quad (2.23)$$

REMARQUE 2.1

En d'autres termes, dans le cas où les quantités sont positives :

- si le rapport rendement/risque de l'actif est plus important que celui du portefeuille, alors on peut améliorer la fonction de retour du portefeuille en lui ajoutant une quantité infinitésimale d'actif i ;
- à l'opposé, si le rapport rendement/risque de l'actif est moins intéressant que celui du portefeuille, alors on peut améliorer la fonction de retour du portefeuille en lui soustrayant une quantité infinitésimale d'actif i .

REMARQUE 2.2

Il est important de comprendre que ce que nous appelons « rapport rendement/risque » dans la remarque ci-dessus est défini comme le rapport entre la *contribution au risque normalisée* $a_i(u)$ de l'actif i dans le portefeuille u et le rendement de l'actif. Cette quantité varie donc selon le portefeuille, elle ne dépend pas que de l'actif i . Le problème de l'allocation optimale est en effet de comparer des performances d'actifs *au sein d'un portefeuille donné*, et pas de comparer des performances individuelles absolues. Nous pouvons faire une analogie avec la gestion de portefeuille : pour composer un portefeuille on ne cherche pas à faire une collection d'actifs aux performances élevées, mais à choisir ceux qui pris collectivement donneront l'investissement le plus intéressant.

En fait, c'est le gradient de la mesure de risque qui est le seul champ de vecteurs adapté à la mesure de performance, pourvu que la mesure de risque soit assez régulière. C'est ce que montre le théorème ci-dessous :

THÉORÈME 2.3

Soit U un ouvert non vide de \mathbb{R}^d et $r : U \rightarrow \mathbb{R}$ une fonction partiellement différentiable sur U de dérivées partielles continues. Soit $a = (a_1, \dots, a_d) : U \rightarrow \mathbb{R}^d$ un champ de vecteurs continu. Alors a est adapté à la mesure de performance avec r si et seulement si

$$a_i(u) = \frac{\partial r}{\partial u_i}(u), \quad i = 1, \dots, d, u \in U. \quad (2.24)$$

Désormais, nous poserons lorsque l'expression a un sens : $a_i(u) = \frac{\partial r}{\partial u_i}(u)$ et nous dirons que $a_i(u)$ est la contribution au risque de l'actif i dans le portefeuille u .

Comparons cette notion de contribution au risque adaptée à la mesure de performance avec des notions proches existant dans la littérature financière ou utilisées par les praticiens.

Par exemple, une notion couramment utilisée, notamment par les spécialistes du risque de crédit, est celle du risque marginal d'un actif. Ce risque est défini comme la différence entre le risque du portefeuille considéré et celui

du portefeuille auquel on a ôté l'actif considéré :

$$r_i^m(u) = r(u) - r(u - u_i e_i), \quad u \in \mathbb{R}^d, \quad (2.25)$$

et nous définissons alors les *contributions marginales* :

$$c_i^m(u) = \frac{r_i^m(u)}{u_i}, \quad u \in \mathbb{R}^d, u_i \neq 0. \quad (2.26)$$

Cette notion est proche d'une dérivée discrète, et si r est différentiable les contributions obtenues seront en général différentes du gradient de r .

Pourquoi préférer le gradient ? En particulier pour la raison suivante : si r est positivement homogène et différentiable, alors d'après le théorème d'Euler :

$$r(u) = \sum_{i=1}^d u_i a_i(u), \quad (2.27)$$

alors que généralement :

$$r(u) \geq u_i c_i^m(u). \quad (2.28)$$

Si les contributions au risque calculées avec le gradient permettent de prendre en compte le risque total du portefeuille, les contributions marginales calculées en ôtant les actifs *sous-estiment* ce risque total, et sont donc dangereuses du point de vue du gestionnaire de risque.

2.5.3 Exemples de contributions au risque

Nous cherchons maintenant à exprimer explicitement les contributions au risque dans le cas de certaines mesures de risque.

Écart-type

Dans le cas de l'écart-type, il faut se restreindre aux portefeuilles de carré intégrable et de variance non nulle. On a alors le champ vectoriel adapté à la mesure de performance suivant :

$$a_i(u) = \frac{\text{Cov}[X_i, Z(u)]}{\sqrt{\mathbb{V}[u]}}. \quad (2.29)$$

VaR et Tail-Var

Tout d'abord, remarquons que dans le cas d'un vecteur gaussien, l'allocation ne sera pas grandement changée par rapport à celle obtenue par la variance, mais simplement multipliée (tout comme la mesure de risque) par une constante. Ceci sera prouvé dans la section 3.1.

Dans le cas général, le quantile n'est pas toujours différentiable en u et nous avons besoin d'hypothèses techniques supplémentaires, voir [32]. Si ces conditions sont remplies, alors on peut définir des contributions au risque par différentiation de la mesure de risque considérée, qui seront adaptées à la mesure de performance et vérifieront en particulier :

$$r(u) = \sum_{i=1}^d u_i a_i(u), \quad u \in U. \quad (2.30)$$

Conclusion

Dans cette partie, nous avons vu que les principaux résultats théoriques obtenus dans le cadre de l'approche moyenne-variance du CAPM pouvaient être étendus au cas d'autres mesures de risque. Sous des hypothèses faibles, il est possible de déterminer la frontière d'efficacité d'un tel programme. Avec l'hypothèse plus forte de différentiabilité, il est possible de définir et d'interpréter les sensibilités. Nous avons également défini des contributions au risque adaptées à la mesure de performance et donné les valeurs de ces contributions dans le cas de certaines mesures de risque.

Néanmoins, comme bien souvent quand on traite d'un cas général, il est rarement aisé d'obtenir une formule analytique pour l'allocation optimale de ces programmes. Cela est seulement possible dans le cas de la variance. Dans la partie suivante, nous nous intéressons à l'obtention, de manière numérique, de ces allocations optimales.

Chapitre 3

Impact de la structure de dépendance

Dans le chapitre précédent, nous avons supposé que nous connaissions parfaitement la loi suivie par les actifs, c'est à dire dans le cas de deux actifs, la loi jointe du couple (X, Y) : non seulement les lois marginales mais aussi toute la structure de dépendance entre X et Y .

Or, pour des lois marginales données, il existe de nombreuses formes de dépendance, qui conduisent à des résultats bien différents lorsqu'on agrège les risques. Prenons l'exemple d'une variable aléatoire symétrique X fixée, et considérons les deux cas $Y_1 = X$ et $Y_2 = -X$. Dans le premier cas, $X + Y_1 = 2X$ alors que dans le second cas, $X + Y_2 = 0$. Pourtant, on a bien $(F_X, F_{Y_1}) = (F_X, F_{Y_2})$. Cet exemple simple montre que faire varier la structure de dépendance peut avoir un impact non négligeable sur la valeur des portefeuilles, donc sur l'allocation optimale.

Nous allons tout d'abord présenter un cadre simple de lois multivariées : celui des lois elliptiques. Puis, après avoir procédé à un rappel succinct sur la notion de copules, nous allons étudier comment faire l'allocation dans le cas où la structure de dépendance est inconnue. Enfin, le problème dans le cas général posant des difficultés, nous allons nous intéresser à l'influence de la structure de dépendance dans le cas de copules paramétriques.

3.1 Les distributions elliptiques

3.1.1 Définitions et propriétés

Dans le cadre gaussien, le vecteur aléatoire \mathbf{X} , de loi gaussienne multivariée $N(\mathbf{m}, \Sigma)$ a pour fonction caractéristique :

$$\forall \mathbf{t} \in \mathbb{R}^n, \mathbb{E}[\exp(i\mathbf{t}'\mathbf{X})] = \exp(i\mathbf{t}'\mathbf{m}) \exp\left(-\frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}\right). \quad (3.1)$$

Les distributions elliptiques multivariées sont une extension des lois gaussiennes dans le sens suivant :

DÉFINITION 3.1

Le vecteur aléatoire \mathbf{X} a une distribution elliptique de paramètres \mathbf{m} , Σ et ϕ , avec Σ définie positive et ϕ fonction scalaire (characteristic generator), si sa fonction caractéristique a pour expression :

$$\forall t \in \mathbb{R}^n, \mathbb{E}[\exp(it'\mathbf{X})] = \exp(it'\mathbf{m})\phi(-\frac{1}{2}\mathbf{t}\Sigma\mathbf{t}). \quad (3.2)$$

On notera alors $X \rightsquigarrow E_n(\mathbf{m}, \Sigma, \phi)$.

Toutes les fonctions ϕ ne conviennent pas [12] notamment il faut $\phi(0) = 1$.

PROPOSITION 3.1

Si X a une densité alors celle-ci est de la forme suivante :

$$f_{\mathbf{X}}(x) = \frac{c_n}{\sqrt{|\Sigma|}} g_n((x - \mathbf{m})'\Sigma^{-1}(x - \mathbf{m})).$$

g_n est le *density generator* et $c_n = \frac{\Gamma(n/2)}{(2\pi)^{n/2}} [\int_0^\infty x^{n/2-1} g_n(x) dx]^{-1}$.

PROPOSITION 3.2

Si la moyenne est bien définie, alors $\mathbb{E}[X] = \mathbf{m}$.

PROPOSITION 3.3

Si la covariance est bien définie, alors $\text{Cov}[X] = -\phi'(0)\Sigma$.

PROPOSITION 3.4

Soit A une matrice $p \times n$ de rang $p \leq n$ et $\mathbf{b} \in \mathbb{R}^p$ alors :

$$A\mathbf{X} + b \rightsquigarrow E_p(A\mathbf{m} + \mathbf{b}, A\Sigma A', g_p).$$

En particulier, si $S = \sum_{i=1}^n \alpha_i X_i = \alpha'\mathbf{X}$, alors

$$S \rightsquigarrow E_1(\alpha'\mathbf{m}, \alpha'\Sigma\alpha, g_1).$$

3.1.2 Expression des mesures de risque

Dans le cadre de notre étude, nous considérons $\mu \rightsquigarrow E_1(m, \sigma, g)$ et $f_\mu = \frac{c}{\sigma} g(\frac{x-m}{\sigma})$. Alors, en notant $\theta_{g,c}(x) = \int_{-\infty}^x g(y^2) dy$, après changement de variable, on obtient pour le quantile :

$$Q_q[-\mu] = m + \sigma_\mu \theta_{g,c}^{-1}(q). \quad (3.3)$$

Il apparaît alors clairement qu'optimiser sur la VaR est équivalent à optimiser sur la variance et donc les résultats de la section précédente consacrée à la variance s'appliquent ici.

Nous avons vu, en première partie, que la VaR ne possédait pas toujours les caractéristiques idoines pour une mesure de risque. Ceci entraîne même des problèmes lors de la résolution des programmes d'optimisation. C'est pourquoi nous étudions maintenant le cas des mesures dites de Wang, présentées en première partie.

Le cas des distributions elliptiques multivariées :

Une mesure de Wang peut s'exprimer ainsi :

$$\rho_g[-\mu] = \int_0^1 Q_q[-\mu] dg(q)$$

avec g fonction de distorsion associée. D'après la section précédente, nous obtenons donc que :

$$\rho_g[-\mu] = m + \sigma \int_0^1 \theta_{g,c}^{-1}(q) dg(q).$$

Dans ce cas là aussi, il apparaît clairement qu'optimiser sur une mesure de risque de Wang est équivalent à optimiser sur la variance et donc les résultats de la section consacrée à la cette mesure de risque s'appliquent.

3.2 Copule et dépendances

3.2.1 Motivation

Nombre de résultats, en particulier dans le domaine de la finance et de la gestion de portefeuille, sont fondés sur l'hypothèse de rendement gaussien, un cas particulier des distributions multivariées présentées ci-dessus. Nous allons nous intéresser maintenant à l'impact sur l'allocation optimale de richesse lorsque l'on sort du cadre gaussien. Les mesures de risque ne se ramènent alors pas toutes à la variance. Pour motiver cette approche, il suffit de prendre deux des indices les plus importants du marché et de les comparer au cadre gaussien.

La figure 3.1 représente, à gauche, la loi du couple (X, Y) (où X est le CAC40 et Y le Dow Jones), dont les lois marginales ont été normalisées. En fait, on a représenté non pas (X, Y) mais $(\Phi^{-1}(F_X(X)), \Phi^{-1}(F_Y(Y)))$, où Φ est la fonction de répartition d'une gaussienne centrée réduite. À droite, on a simulé un vecteur gaussien de marginales centrées réduites et dont la corrélation linéaire est celle du couple (CAC, DJ).

La différence est frappante en particulier dans les extrêmes (qui sont plus corrélés que dans le cas gaussien), ce qui signifie qu'il est d'autant plus important de s'écarter du cas gaussien, que l'on utilise des mesures de risque autre que la variance. Nous allons donc chercher à introduire d'autres types de distribution, un outil très efficace pour ce faire est la théorie des copules.

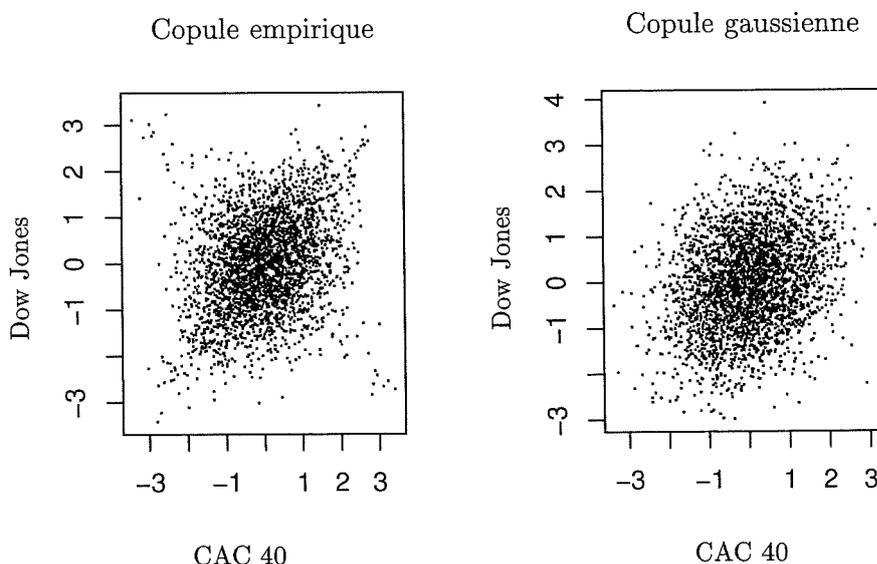


FIG. 3.1 – Comparaison des rendements croisés observés au cas gaussien idéal.

3.2.2 Définition

Nous supposons que les lois marginales sont connues. En notant ces lois marginales F_1 et F_2 , cela revient à se placer dans la classe de Fréchet $\mathcal{F}(F_1, F_2)$ définie comme suit :

DÉFINITION 3.2 (CLASSE DE FRÉCHET)

Soit F_1, \dots, F_n , n distributions sur \mathbb{R} . On définit $\mathcal{F}(F_1, \dots, F_n)$ l'ensemble de toutes les distributions sur \mathbb{R}^n ayant comme marginales unidimensionnelles F_1, \dots, F_n . Cet ensemble prend le nom de l'ensemble de Fréchet.

Une fois choisie la classe de Fréchet, nous cherchons à caractériser de manière simple la structure de dépendance, et à la définir précisément par la même occasion. Comment le faire en évitant une redite du comportement marginal ? Une des solutions est d'utiliser la notion de copule, qui permet de se ramener à des marginales simples (celles d'une loi uniforme sur $[0, 1]$) pour mieux exprimer la dépendance.

Définissons tout d'abord les notions de n -croissance ou n -positivité.

DÉFINITION 3.3 (n -CROISSANCE)

Soit $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$. Soit (e_i) la base canonique de \mathbb{R}^n . On définit l'opérateur de différences finies

$$\Delta_i^{\varepsilon} \varphi(x) = \varphi(x + \varepsilon e_i) - \varphi(x).$$

φ est dite n -positive, ou n -croissante si pour tout $\varepsilon_1, \dots, \varepsilon_n \geq 0$, et pour tout $x \in \mathbb{R}$ on ait

$$\Delta_1^{\varepsilon_1} \circ \dots \circ \Delta_n^{\varepsilon_n} \varphi(x) \geq 0.$$

Parmi les exemples de fonctions n -croissantes, nous pouvons citer:

1. $x \rightarrow -\max(x^i)$;
2. Les fonctions de répartition;
3. $(x_1, x_2) \rightarrow -|x_1 - x_2|^p$.

DÉFINITION 3.4 (COPULES)

Soit $n \geq 2$.

Une copule n -dimensionnelle est une fonction C qui vérifie les conditions suivantes:

1. Le domaine de définition de C est I^n avec $I = [0, 1]$;
2. C est n -croissante, et elle s'annule dès que l'un de ses arguments s'annule;
3. Pour chaque entier $m \leq n$ les marginales m -dimensionnelles (qui sont les fonctions admettant $n - m$ arguments égales à la borne sup de leur domaine qui est 1 dans notre cas) sont l'identité sur \mathbb{R}^n .

En remarquant que la copule est invariante par toute transformation croissante des marginales, nous pouvons donc définir la structure de dépendance d'un vecteur aléatoire comme l'ensemble des caractéristiques invariantes par toute transformation croissante des marginales. L'intérêt des copules résulte du théorème suivant :

THÉORÈME 3.1

Soit H une n -distribution et F_1, \dots, F_n ses marginales (unidimensionnelles) alors il existe une n -copule telle que

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

pour tout x_1, \dots, x_n dans \mathbb{R} . Une de ces copules est donnée par

$$C(u_1, \dots, u_n) = H(F_1^{-1}(x_1), \dots, F_n^{-1}(x_n))$$

Insistons sur l'importance de l'hypothèse de continuité de F_1 et F_2 : si l'une de ces deux lois n'est pas continue, alors l'unicité n'est plus assurée.

Pour revenir au graphique 3.1, nous voyons que celui-ci ne représente pas réellement la copule entre CAC et DJ, mais une copule « gaussianisée » avec des marginales gaussiennes au lieu des marginales uniformes de la théorie.

3.2.3 Calculs des bornes pour la VaR

Nous donnons ici les programmes d'optimisation pour calculer les bornes pour la VaR de la somme de deux variables aléatoires, lorsque la structure de dépendance est inconnue. L'étude fine est faite dans [11].

Considérons deux variables aléatoires réelles X et Y de distributions connues F et G respectivement. Leur loi jointe est inconnue. On se propose de déterminer un encadrement de la fonction de répartition de $X + Y$. Le maximum et le minimum sur cette distribution n'ont pas de copules correspondantes, et ils sont limites point par point de distributions bornant celle de la somme et sont les meilleures bornes possibles.

Ces bornes peuvent être exprimées en tout point comme le résultat d'un programme d'optimisation réelle sous contraintes. On peut aussi exprimer le résultat pour les inverses des distributions, soit les VaR. Et on a :

$$\begin{aligned} \underline{\text{VaR}}_+(F,G) &= \sup_{\min(u+v,1)=x} F^{-1}(u) + G^{-1}(v) \\ &= \begin{cases} F^{-1}(1) + G^{-1}(1), & \text{si } x = 1; \\ \sup_{0 \leq u \leq x} F^{-1}(u) + G^{-1}(x - u), & \text{sinon.} \end{cases} \end{aligned} \quad (3.4)$$

et de même

$$\begin{aligned} \overline{\text{VaR}}_+(F,G) &= \inf_{\max(u+v-1,0)} F^{-1}(u) + G^{-1}(v) \\ &= \begin{cases} \inf_{u+v-1 < 0} F^{-1}(u) + G^{-1}(v), & \text{si } x = 0; \\ \inf_{x \leq u \leq 1} F^{-1}(u) + G^{-1}(x - u + 1), & \text{sinon.} \end{cases} \end{aligned} \quad (3.5)$$

3.3 Structure de dépendance inconnue

Nous nous intéressons maintenant au cas où la structure de dépendance est inconnue. Il se trouve alors que l'agent ne peut plus résoudre les programmes explicités dans la partie précédente. En effet, quelle copule doit-il choisir pour minimiser son risque? Comment déterminer l'allocation optimale? Nous avons essayé, à l'instar de [22], de répondre à cette question, non par l'estimation d'une copule empirique mais par la résolution d'un programme d'optimisation général.

3.3.1 Présentation du programme

Le programme général pour déterminer l'allocation optimale se décompose en deux programmes. Nous travaillons avec une mesure de risque ρ , de type mesure de risque monétaire, par exemple la VaR ou la TVaR.

L'idée est simple et raisonnable: ne connaissant pas la structure de dépendance, nous allons déterminer le cas le plus défavorable. Pour obtenir l'allocation optimale, nous choisissons donc de minimiser le pire cas, et de résoudre :

$$\boxed{\arg \min_{\alpha_1} \max_{(Z_1, Z_2) \in \mathcal{F}(F_1, F_2)} \rho[\alpha_1 Z_1 + \alpha_2 Z_2]}.$$

Nous supposons la contrainte de rendement saturée, i.e. $\alpha_1 m_1 + \alpha_2 m_2 = r$. Les notations sont celles de la partie précédente.

3.3.2 Programme de maximisation sur l'ensemble des copules

Nous nous intéressons tout d'abord au programme suivant :

$$\max_{(F_1, F_2) \in \mathcal{F}(Z_1, Z_2)} \rho[\alpha_1 Z_1 + \alpha_2 Z_2].$$

Il y a quatre cas à considérer pour (α_1, α_2) :

1. $\alpha_1 \geq 0, \alpha_2 \geq 0$;
2. $\alpha_1 \geq 0, \alpha_2 \leq 0$;
3. $\alpha_1 \leq 0, \alpha_2 \geq 0$;
4. $\alpha_1 \leq 0, \alpha_2 \leq 0$.

Comme la contrainte est linéaire, (α_1, α_2) peut appartenir au plus à trois quadrants du plan sur quatre. Nous considérerons donc sans perte de généralité les cas 1., 2. et 3. (comme dans nos exemples).

Dans le cas de la VaR et des mesures de risque de Wang, d'après [9], nous avons l'additivité pour des risques comonotones i.e.

$$\rho[\alpha_1 Z_1^c + \alpha_2 Z_2^c] = \alpha_1 \rho[Z_1] + \alpha_2 \rho[Z_2]$$

pour $\alpha_1 \geq 0, \alpha_2 \geq 0$ et (Z_1^c, Z_2^c) version comonotone du vecteur (Z_1, Z_2) .

Cas des mesures de risques de Wang

Nous obtenons les trois inégalités suivantes:

1. pour $\alpha_1 \geq 0, \alpha_2 \geq 0, \rho[\alpha_1 Z_1 + \alpha_2 Z_2] \leq \alpha_1 \rho[Z_1] + \alpha_2 \rho[Z_2]$;
2. pour $\alpha_1 \geq 0, \alpha_2 \leq 0, \rho[\alpha_1 Z_1 + \alpha_2 Z_2] \leq \alpha_1 \rho[Z_1] + |\alpha_2| \rho[-Z_2]$;
3. pour $\alpha_1 \leq 0, \alpha_2 \geq 0, \rho[\alpha_1 Z_1 + \alpha_2 Z_2] \leq |\alpha_1| \rho[-Z_1] + \alpha_2 \rho[Z_2]$.

Ces inégalités sont des égalités pour (Z_1^c, Z_2^c) , $(Z_1^c, (-Z_2)^c)$ et $((-Z_1)^c, Z_2^c)$. La solution du programme de maximisation sur l'ensemble des copules est donc affine par morceaux et est donnée par :

1. pour $\alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_1 \rho[Z_1] + \alpha_2 \rho[Z_2]$;
2. pour $\alpha_1 \geq 0, \alpha_2 \leq 0, \alpha_1 \rho[Z_1] + |\alpha_2| \rho[-Z_2]$;
3. pour $\alpha_1 \leq 0, \alpha_2 \geq 0, |\alpha_1| \rho[-Z_1] + \alpha_2 \rho[Z_2]$.

Cas de la VaR

Le cas de la VaR est légèrement différent puisque le maximum n'est pas atteint pour la copule comonotone, cela résulte de la non sous-additivité de la VaR pour certaines copules. Sur le graphe 3.2, nous avons tracé la courbe affine par morceaux pour la copule comonotone et la courbe maximale dans le cas de la VaR. Cette courbe a été obtenue grâce au programme présenté en 3.2.3.

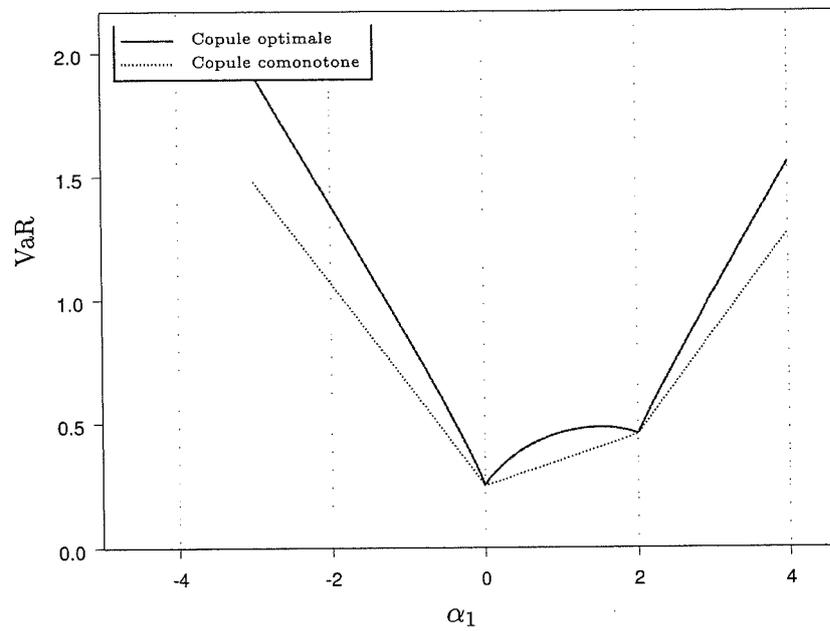


FIG. 3.2 – Évolution de la VaR pour le cas comonotone et le cas le plus défavorable.

Résolution du programme d'allocation

La résolution de la deuxième partie du programme, i.e. le programme d'allocation en lui-même, est immédiate.

Dans le cas des mesures de risque de Wang, le minimum est atteint lorsque le portefeuille est constitué uniquement de l'actif le moins risqué au sens de ρ . Dans le cas de la VaR, la courbe maximum passe par ce point car c'est un cas dégénéré où il n'y a qu'un seul actif. On obtient donc le même résultat.

Ainsi, la résolution de ce programme montre que l'allocation optimale est telle que tout l'investissement se fait dans l'actif le moins risqué. Et il n'y a donc pas de phénomène de diversification. Ce résultat est bien sûr trop fruste et ne peut pas être appliqué en pratique.

L'échec du programme résulte de la maximisation du risque sur l'ensemble des copules. Pour que la solution de ce problème soit non triviale, il faut se restreindre à un ensemble plus petit, par exemple, à un ensemble de copules paramétriques.

3.4 Structure de dépendance paramétrée

Il n'est pas intéressant de résoudre un programme général, comme celui présenté ci-dessus. Pour la recherche de l'allocation optimale, il faut avoir une « idée » de la structure de dépendance. Dans cette partie, nous allons nous intéresser particulièrement à l'influence de la structure de dépendance sur l'allocation optimale.

Nous allons utiliser pour notre étude trois mesures de risque : la variance, la VaR(90%) et la TVaR(90%), car ce sont les plus usitées et rechercher les allocations optimales pour ces trois mesures dans le cas de deux structures de dépendance. Nous considérons uniquement le cas de deux rentabilités d'actif par rapport à l'actif sans risque. Z_1 et Z_2 suivent des lois normales de paramètres $(m_1 = 0,03, \sigma_1 = 0,2)$ et $(m_2 = 0,1, \sigma_2 = 0,4)$. Pour construire notre portefeuille nous envisageons une contrainte de rendement de 6%.

L'étude porte alors sur une copule paramétrique, dans chacun des deux cas suivants nous faisons varier le paramètre dans une plage naturelle et nous étudions l'impact sur l'allocation optimale et la valeur de la mesure de risque pour cette allocation.

3.4.1 Structure gaussienne

Notre première application se fait dans le cas de lois gaussiennes multivariées. Le paramètre influent est donc le coefficient de corrélation ρ entre

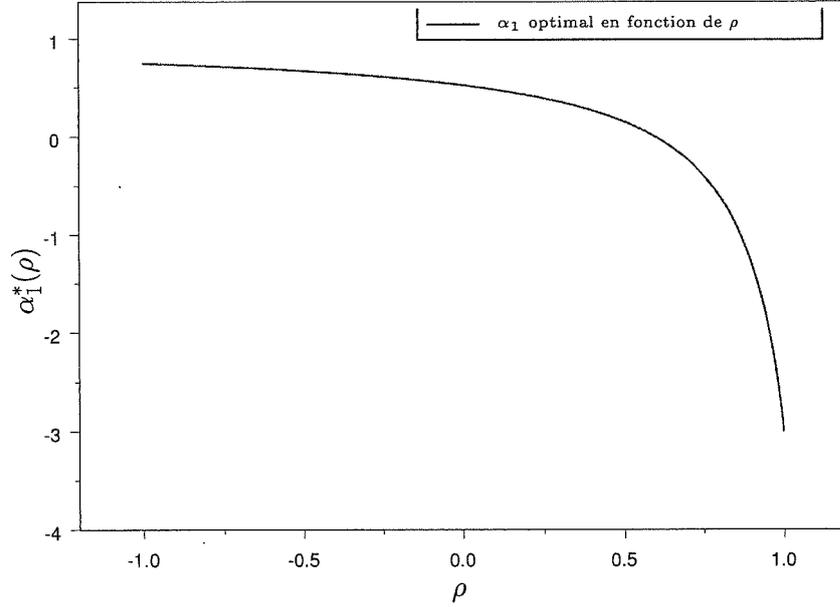


FIG. 3.3 – Évolution de l'allocation optimale en fonction de la corrélation.

Z_1 et Z_2 .

Rappelons que pour une loi normale X de moyenne m et σ , et si N suit une loi normale centrée réduite –cas particulier des lois elliptiques– nous avons :

$$\text{VaR}_p[X] = m + \sigma \text{VaR}_p[N] \quad (3.6)$$

et

$$\text{TVaR}_p[X] = m + \sigma \text{TVaR}_p[N]. \quad (3.7)$$

Nous obtenons alors que pour la VaR et la TVaR, l'allocation optimale est la même que dans le cas de la variance. Cette allocation (cf. fig 3.3), obtenue dans la partie précédente, est :

$$\alpha_1^* = \frac{r}{m_1} \frac{\theta^2 - \rho\theta}{1 - \rho\theta + \theta^2}$$

qui fournit comme variance optimale :

$$\text{Var}[\mu^*] = \frac{r^2 \theta^2 \sigma_1^2}{(1 - \rho\theta + \theta^2)^2 m_1^2} ((\theta - \rho)^2 + 2\rho(\theta - \rho)(1 - \theta\rho) + (1 - \theta\rho)^2). \quad (3.8)$$

Sur le graphe 3.4, nous constatons bien que les formes des courbes de la VaR, TVaR et écart-type optimaux sont homothétiques.

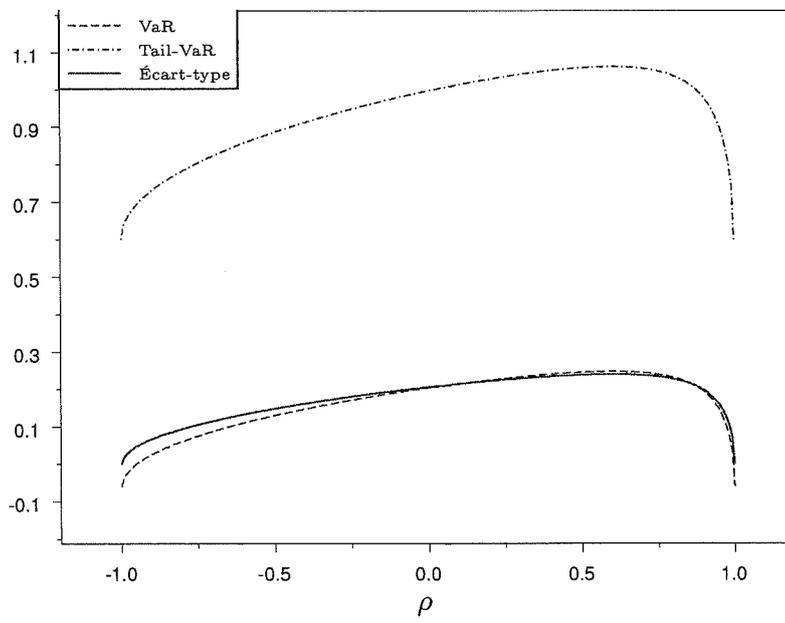


FIG. 3.4 – VaR, TVaR et écart-type à l'optimum.

3.4.2 La famille des copules de Clayton

Nous pouvons également choisir une autre famille de copules :

DÉFINITION 3.5

La famille des copules de Clayton, paramétrée par $\gamma > 0$, est celle des copules s'écrivant :

$$C_\gamma(u_1, u_2) = (u_1^{-\gamma} + u_2^{-\gamma} - 1)^{-1/\gamma}. \quad (3.9)$$

Sur la figure 3.5, il est possible de comparer l'effet de la structure de dépendance dans le cas de la copule de Clayton avec celui de la copule gaussienne. Les valeurs négatives, c'est à dire les pertes, paraissent beaucoup plus corrélées dans le cas de la copule de Clayton. Il est donc judicieux de choisir cette copule pour prendre en compte des scénarios de crise.

Il n'est pas possible d'obtenir une expression analytique de l'allocation optimale en fonction du paramètre γ pour la VaR et la TVaR. Nous avons donc dû calculer numériquement ces grandeurs. Pour ce faire nous avons utilisé le logiciel R [26]. Notre méthode est fondée sur des simulations de Monte-Carlo. Pour chaque valeur du paramètre γ , nous simulons un jeu de valeurs d'actif. Dans un second temps, après avoir discrétisé l'ensemble des allocations, nous calculons la VaR(90%) et la TVaR(90%) pour chaque portefeuille. Enfin, l'allocation optimale est obtenue par recherche du minimum

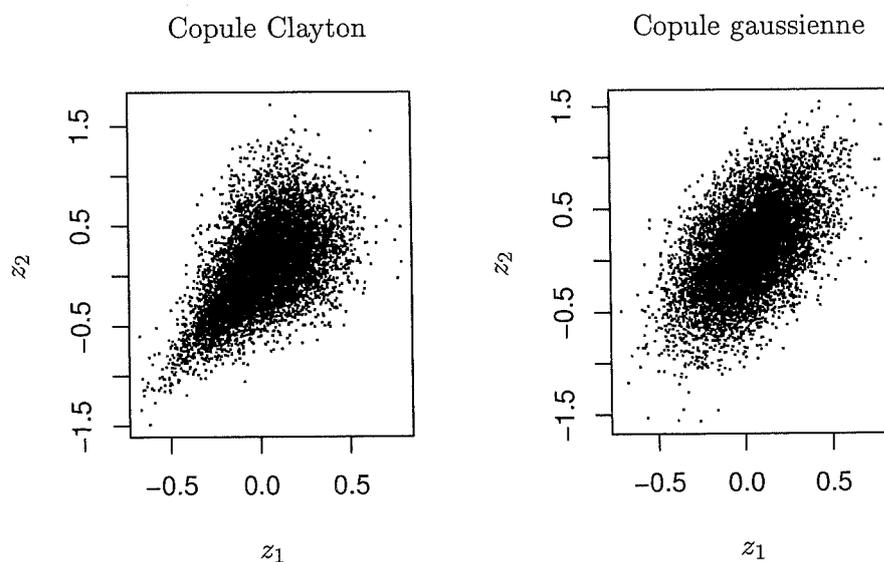
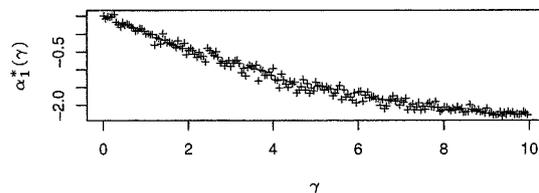


FIG. 3.5 – Comparaison des deux distributions pour un même coefficient de corrélation de 0,5.

Évolution de l'allocation optimale en fonction de γ



Évolution de la VaR à 90% à l'optimum en fonction de γ

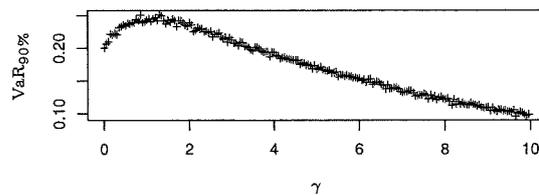
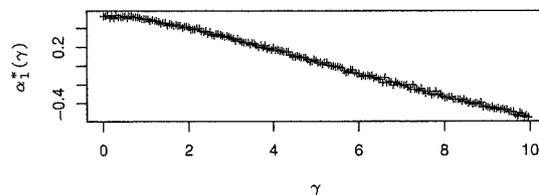


FIG. 3.6 – Influence sur le programme d'optimisation dans le cas de la VaR.

Évolution de l'allocation optimale en fonction de γ



Évolution de la TVaR à 90% à l'optimum en fonction de γ

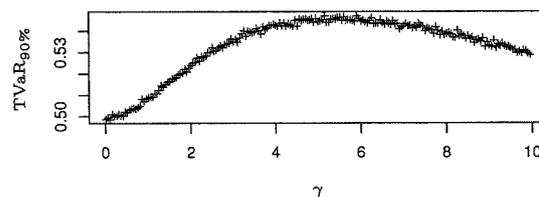


FIG. 3.7 – Influence sur le programme d'optimisation dans le cas de la TVaR.

sur cette ensemble d'allocation. Les résultats sont visibles sur les figures 3.6 et 3.7. Nous remarquons que les courbes de la VaR et de la TVaR ne sont pas semblables.

3.4.3 Analyse de l'étude

Les résultats présentés dans cette partie montre que dès que l'on s'écarte du cas gaussien – ou plus généralement du cas elliptique – le choix effectif de l'allocation optimale se complique. Tout d'abord, l'influence de la mesure de risque est importante puisque les allocations optimales varient grandement d'une mesure à une autre. Il est donc essentiel pour l'agent économique de choisir la mesure de risque la plus pertinente dans le cadre de son travail.

D'autres difficultés sont liées à l'estimation des paramètres. En effet, nous voyons sur les graphes précédents que l'allocation optimale varie aussi en fonction du coefficient de corrélation. Si, pour la copule de Clayton l'effet s'atténue lorsque l'on considère la TVaR, il reste important pour la VaR et la variance figure 3.8 . Il est à noter de plus que dans le cas de la copule Clayton, la VaR optimale et la TVaR optimale n'ont pas des formes similaires.

Enfin, terminons cette étude par la présentation d'une situation de « conflit ». Imaginons que, dans une banque, le gestionnaire de portefeuille utilise la variance pour allouer sa richesse, et que le gestionnaire des risques lui s'intéresse à la VaR, à cause entre autres de contraintes légales. Le graphique 3.9 montre à quel point leurs intérêts divergent : la VaR de l'allocation obtenue par minimisation de la variance est supérieure à celle de l'allocation obtenue par minimisation de la VaR, ce qui est normal. L'écart devient important lorsque γ augmente.

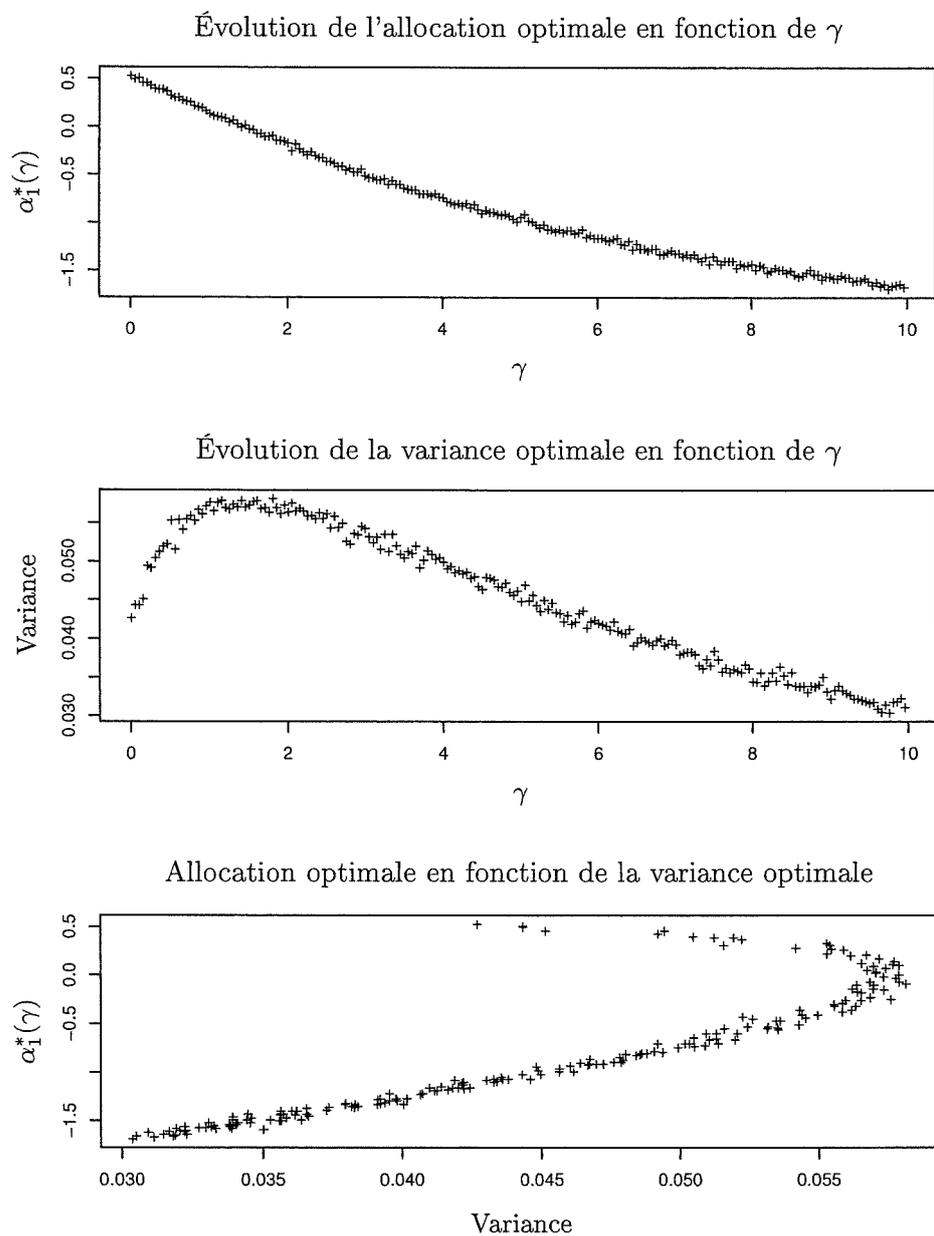


FIG. 3.8 – Influence sur le programme d'optimisation dans le cas de la variance pour la copule Clayton.

Cas non gaussien : résultats différents selon les programmes

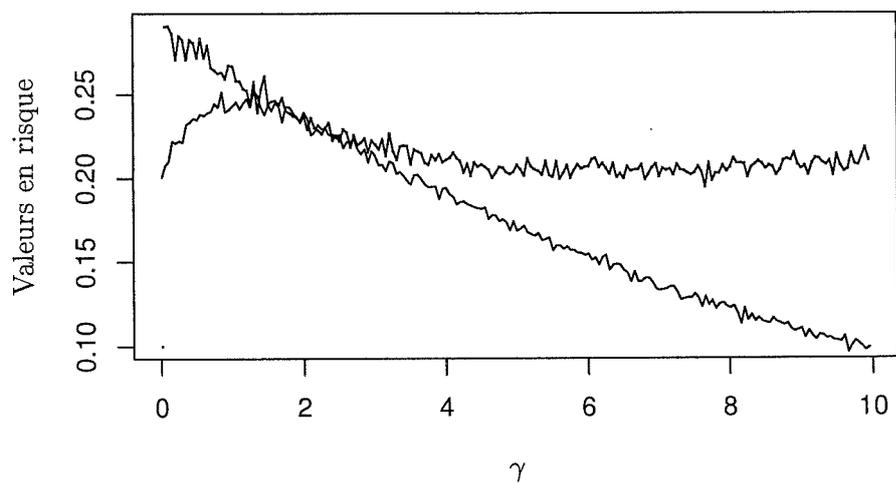


FIG. 3.9 – VaR de l'allocation obtenue en minimisant la variance (en haut) et VaR de l'allocation obtenue en minimisant la VaR (en bas).

Conclusion

La gestion du risque dans la banque a évolué ces dernières années, la VaR et la TVaR ont fait leur apparition aux côtés de la traditionnelle variance. Cette évolution est essentiellement due aux régulateurs, qui cherchent à obliger la constitution de réserve de fonds propres minimum pour assurer la solvabilité des banques et, au delà, à assurer ainsi la stabilité du système financier dans son ensemble.

D'un point de vue théorique, la conception de mesure de risque a elle aussi évoluée, notamment par la recherche d'une rationalisation de ses propriétés. De nouvelles classes de mesures de risques ont pu être déterminées dans un cadre rigoureux, mesures de risque cohérentes ou mesures de risque de Wang. Ces dernières sont fondées sur un mélange de VaR et englobent notamment le cas de la TVaR. Des liens ont été exhibés entre ces mesures et les nouvelles théories du choix dans l'incertain, résultant de l'affaiblissement de certains axiomes de la théorie traditionnelle de l'espérance d'utilité.

Dans ce nouveau cadre, nous avons cherché à résoudre le problème de l'allocation optimale de richesse pour l'agent, qui espère un certain rendement et veut limiter sa prise de risque. Il apparaît que sous des hypothèses assez faibles les résultats théoriques établis dans le cadre moyenne-variance peuvent se retrouver dans le cas d'autres mesures de risque, comme la VaR ou les mesures de risques de Wang.

Dans le cas général, le problème principal est la détermination effective de la mesure de risque. Il faut avoir recours à des techniques numériques. De plus, en comparant deux cas d'école : lois gaussiennes multivariées et copule de Clayton avec marginales gaussiennes, nous avons pu montrer l'importance de la structure de dépendance dans le problème de l'allocation optimale.

En ce qui concerne la sélection de portefeuille, il peut être intéressant d'étendre les résultats théoriques obtenus à une classe plus grande de mesure de risque, par exemple, les mesures de risque cohérentes, et d'établir dans ce cadre un modèle de marché.

Index

- adapté à la mesure de performance, 40
- additivité pour les risques comonotones, 21
- characteristic generator*, 46
- classe de Fréchet, 48
- contribution au risque, 41
- convexité, 14
- copule, 49
 - de Clayton, 56
- density generator*, 46
- direction normale, 30
- distribution elliptique, 46
- espérance, 10
- fluctuation, 39
- fonction de distorsion, 19
 - concave, 21
- fonction de retour, 40
- fonction de risque, 26
- fonction objectif, 24, 26
- frontière d'efficience, 31, 34
- invariance au cash, 14
- Karusch–Kuhn–Tucker, 31
- mesure de risque, 10
 - cohérente, 14
 - convexe, 14
 - de Wang, 19
 - dual-power, 20
 - monétaire, 13
 - représentation, 16
 - PH, 20
- mesure martingale, 17
- monotonie, 14
- n -croissance, 48
- n -homogénéité, 26
- neutralité, 14
- normalisation, 14
- prime pure, voir espérance
- programme d'optimisation, 25, 28
- quantile, 11
- risque, 9
- réassurance proportionnelle, 37
- région des portefeuilles acceptables, 15
- semi-continue inférieurement, 27
- sensibilité, 28
- sous-additive, 27
- sous-additivité, 15
- sous-différentiel, 27
- stress testing*, 17
- TVaR, 12
- théorème d'Euler, 28
- théorie duale, 18
- transformée de Wang, 21
- valeur en risque, voir VaR
- VaR, 11
- variance, 11

Bibliographie

- [1] Maurice Allais. Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine. *Econometrica*, 21(4):503–546, 1953.
- [2] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(9):203–228, 1999.
- [3] Gilbert W. Bassett Jr., Roger Koenker, and Gregory Kordas. Pessimistic portfolio allocation and Choquet expected utility. Preprint, 2004.
- [4] Rachel Campbell, Ronald Huisman, and Kees Koedijk. Optimal portfolio selection in a value-at-risk framework. *Journal of Banking & Finance*, 25:1789–1804, 2001.
- [5] Michèle Cohen and Jean-Marc Tallon. Décision dans le risque et l'incertain : l'apport des modèles non-additifs. *Revue d'économie politique*, pages 631–681, 2000.
- [6] Freddy Delbaen. Coherent risk measures on general probability spaces, 2000. Preprint.
- [7] Michel Denault. Coherent allocation of risk capital. *The journal of risk*, 4(1):1–34, 2001.
- [8] Michel Denuit and Arthur Charpentier. *Mathématiques de l'assurance non-vie*, volume 1, chapter 5. Economica, 2004.
- [9] Jan Dhaene, Steven Vanduffel, Qihe Tang, Marc J. Goovaerts, Rob Kaas, and David Vyncke. Risk measures and optimal portfolio selection, 2003. Preprint.
- [10] Annalisa Di Clemente and Claudio Romano. Beyond Markowitz: building optimal portfolio using non-elliptical asset return distributions, 2004. Working paper.
- [11] Khalil A. El Dairi. Bornes sur les transformées de couples de variables aléatoires. Rapport de stage, ENSAE, 2004.
- [12] K.T. Fang, S Kotz, and K.W. Ng. *Symmetric multivariate and related distribution*. Chapman Hall, London, 1987.
- [13] Tom Fischer. Risk capital allocation by coherent risk measures based on one-sided moments. *Insurance Mathematics and Economics*, 32:135–146, 2003.

- [14] Hans Föllmer. Incertitude financière, préférences et mesures de risque. In *mathématiques financières, conférence-débat à l'académie des sciences*, 1^{er} février 2005.
- [15] Hans Föllmer and Alexander Schied. *Stochastic finance: an introduction in discrete time*, volume 27 of *de Gruyter studies in mathematics*, chapter 4. Walter de Gruyter, Berlin, 2nd edition, 2004. [78 FOL 00 A].
- [16] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. *J. of math. economics*, 18:141–153, 1989.
- [17] Christian Gouriéroux, Jean-Pierre Laurent, and Olivier Scaillet. Sensitivity analysis of value at risk. *Journal of Empirical Finance*, 7:225–245, 2000.
- [18] Christian Gouriéroux, Olivier Scaillet, and Ariane Szafarz. *Econometrie de la finance : analyses historiques*. Economica, 1997. [78 GOU 00 A].
- [19] Wilfried G. Hallerbach. Capital allocation, portfolio enhancement and performance measurement : a unified approach, 2003. Working paper.
- [20] Jean-Baptiste Hiriart-Urruty. *Optimisation et analyse convexe*. PUF, 1998.
- [21] Elyes Jouini, Moncef Meddeb, and Nizar Touzi. Vector-valued coherent risk measures. *Finance and stochastics*, 8:531–552, 2002.
- [22] Roger Laeven and Marc Goovaerts. An optimization approach to the dynamic allocation of economic capital. Preprint, 2004.
- [23] Harry M. Markowitz. Portfolio selection. *Journal of finance*, 7(1):77–91, 1952.
- [24] Committee on the global financial system. Stress testing at major financial institutions: survey results and practice. Technical report, Bank for international settlements, janvier 2005.
- [25] John Quiggin. A theory of anticipated utility. *J. of econ. behavior and organization*, 3:323–343, 1982.
- [26] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2004. <http://www.R-project.org>.
- [27] R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. Working paper, 1999.
- [28] Alexander Schied. Risk measures and robust optimization. In *Cours du séminaire bachelier*, janvier 2005.
- [29] David Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.
- [30] William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.
- [31] Dirk Tasche. Risk contributions and performance measurement. Working paper, 2000.

- [32] Dirk Tasche. Allocating portfolio economic capital to sub-portfolios. Working paper, 2004.
- [33] Shaun S. Wang and Virginia R. Young. Ordering risks: Expected utility theory versus Yaari's dual theory of risk. *Insurance Mathematics and Economics*, 22:145–161, 1998.
- [34] Menahem E. Yaari. The dual theory of choice under risk. *Econometrica*, 55(1):95–115, 1987.