

HEC MONTRÉAL

Réplication des fonds de couverture

Par

Malek Ben Abdellatif

**Sciences de la gestion
(Ingénierie financière)**

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

Novembre 2009

© Malek Ben Abdellatif, 2009

Résumé

Dans ce travail de recherche, nous généralisons la méthode de répliation proposée par Papageorgiou *et al.* (2008) en répliquant les rendements d'un fonds de couverture synthétique conditionnellement à $(n - 1)$ portefeuilles et en utilisant un portefeuille de réserve. De plus, nous utiliserons l'algorithme de couverture élaboré par Rémillard & Rubenthaler (2009) qui garantira une répliation optimale. Les rendements sont modélisés à l'aide de mélanges gaussiens multivariés et la structure de dépendance du fonds avec les différents portefeuilles est prise en compte en utilisant les copules.

Remerciements

Je tiens à remercier mes directeurs Bruno Rémillard et Nicolas Papageorgiou pour leur soutien moral et financier tout au long de la réalisation de ce travail.

Sur le plan personnel, je dédie ce mémoire à mes chers parents *Mohamed* et *Wided* ainsi que mon mari *Brahim*. Je les remercie pour leurs encouragements et leurs sacrifices. Tout ce que j'ai accompli à ce jour est grâce à leur amour et leur soutien précieux. Je leur dois tout. Merci également à mon frère et ma petite soeur ainsi que ma belle famille pour leurs encouragements incessants.

Table des matières

Résumé	i
Remerciements	ii
Table des matières	iii
Table des figures	v
Liste des tableaux	vi
1 Introduction	1
2 Revue de la littérature	3
2.1 Amin & Kat (2003)	4
2.2 Kat & Palaro (2005)	6
2.3 Papageorgiou <i>et al.</i> (2008)	7
3 Modélisation des rendements	10
3.1 Les mélanges Gaussiens	10
3.1.1 Définition	11
3.1.2 Estimation	11
3.1.3 Estimation	12
3.2 Les copules	14
3.2.1 Définition	14

TABLE DES MATIÈRES **iv**

3.2.2	Copules archimédiennes	15
3.2.3	Propriétés	17
4	Stratégie de réplication	19
4.1	Fonction de paiement	19
4.2	Portefeuille de réplication	21
4.2.1	La couverture optimale	21
4.2.2	Le cas markovien	23
4.2.3	Implémentation de la stratégie dynamique de réplication	24
5	Résultats empiriques	26
5.1	Choix des portefeuilles	26
5.2	Modélisation des rendements	28
5.3	Résultats de la réplication	32
6	Conclusion	37
	Bibliographie	39

Table des figures

5.1	Erreur de répliation pour la période 1 avec la copule de Clayton.	34
5.2	Erreur de répliation pour la période 1 avec la copule de Gumbel.	35
5.3	Erreur de répliation pour la période 1 avec la copule de Frank.	35
5.4	Erreur de répliation pour la période 1 dans le cas de l'indépendance.	36

Liste des tableaux

5.1	Composition des portefeuilles.	27
5.2	Statistiques descriptives pour les portefeuilles sur les 3 périodes.	28
5.3	Paramètres des mélanges Gaussiens utilisés pour la modélisation des rendements sur la période 1 (99-09).	29
5.4	Paramètres des mélanges Gaussiens utilisés pour la modélisation des rendements pour la période 2 (99-03).	30
5.5	Paramètres des mélanges Gaussiens utilisés pour la modélisation des rendements pour la période 3 (04-09).	31
5.6	Paramètres cibles et ceux obtenus par la fonction g_T pour 10000 simulations selon les paramètres de la période 1.	32
5.7	Paramètres cibles et ceux obtenus par la fonction g_T pour 10000 simulations selon les paramètres de la période 2.	32
5.8	Paramètres cibles et ceux obtenus par la fonction g_T pour 10000 simulations selon les paramètres de la période 3.	33
5.9	Résultats de la réplication.	33
5.10	Erreur de réplication pour la période 1	34
5.11	Erreur de réplication pour la période 2	34
5.12	Erreur de réplication pour la période 3	34

Chapitre 1

Introduction

Les fonds de couverture ont suscité l'intérêt des recherches en raison des rendements élevés qu'ils promettent. En effet, grâce à leur caractère privé et en l'absence de restrictions, ils peuvent offrir des gammes de stratégies d'investissement très variées. Ces rendements élevés ont donc été expliqués non seulement par l'exposition au risque systématique du marché mais aussi par les compétences des gestionnaires. L'idée de pouvoir répliquer ces rendements a été mise en oeuvre au départ dans le but de pouvoir évaluer les performances de ces fonds de couverture. Cependant le manque de transparence et les frais élevés exigés pour investir dans ces fonds ont donné lieu à l'idée de pouvoir générer les mêmes rendements que le fonds.

Amin & Kat (2003) ont été les premiers à avoir eu l'idée de reproduire des rendements ayant la même distribution que les rendements du fonds de couverture. Par la suite, Kat & Palaro (2005) ont introduit l'idée de considérer un fonds de couverture conditionnellement à un portefeuille de titres. Et enfin, Papageorgiou *et al.* (2008) ont perfectionné cette approche de répliation qui consistait à construire une stratégie d'investissement avec une fonction de paiement déterminée par les rendements historiques du fonds de couverture, qui permet de reproduire la distribution des rendements du fonds conditionnellement au portefeuille détenu par l'investisseur. Cette stratégie d'investissement consiste en la formation d'un portefeuille dynamique de répliation dont les

composantes sont le portefeuille de l'investisseur et un actif de réserve.

Dans ce travail, nous nous proposons de mettre en oeuvre cette approche de réplification en considérant le fonds de couverture conditionnellement à plusieurs portefeuilles de titres détenus par l'investisseur. Pour ce faire, il faut commencer par modéliser les rendements et estimer leurs lois et la structure de dépendance entre eux. Par la suite, il convient de déterminer la fonction de paiement et le portefeuille de réplification qui vont permettre de générer des rendements ayant la même distribution que le fonds conditionnellement aux portefeuilles de référence. Enfin, nous étudions l'efficacité de la méthode implantée en essayant de répliquer les rendements d'un fonds synthétique conditionnellement à des portefeuilles formés à partir de données de marché.

Chapitre 2

Revue de la littérature

Les études des rendements historiques des fonds de couverture ont révélé que ces derniers ne suivent pas une loi normale et que leur relation avec le marché est non linéaire. Les mesures de performance classiques telles que le ratio de Sharpe ou l'alpha de Jensen ne sont donc pas appropriées aux fonds de couverture, c'est pourquoi les chercheurs ont tenté de trouver des mesures de performance adéquates à la nature de ces rendements. Ainsi, certains ont eu recours aux modèles factoriels pour expliquer les rendements des fonds de couverture dont la performance a été attribuée à deux facteurs. Le premier est associé à l'exposition au risque systématique et génère « les rendements bêta » alors que le deuxième est lié à la compétence des gestionnaires et génère « les rendements alpha ». Une fois les facteurs explicatifs identifiés, il suffit de comparer le rendement du fonds avec celui que le modèle prédit pour voir s'il a réalisé une bonne performance.

Cette approche permet, par la même occasion, de former un portefeuille de réplique ayant les mêmes expositions au risque que le fonds de couverture et, en théorie, de répliquer ses rendements. En l'absence de transparence et en raison des frais élevés d'investissement dans les fonds de couverture, cette réplique s'avère utile. Hasanhodzic & Lo (2007) montrent que cette manière de faire fonctionne bien pour les fonds de couverture qui adoptent certaines stratégies. D'autres travaux comme ceux de Mitchell & Pulvino (2001) et Chen

& Liang (2006) proposent d'introduire une prime de risque pour capturer les sources de la non-linéarité.

Cependant, la capacité de ces modèles linéaires à reproduire les rendements des fonds de couverture est limitée en raison de la non-linéarité de ces rendements avec le marché. De plus, la réplication devient compliquée à cause de l'utilisation de facteurs non échangés sur les marchés. Afin de capturer cette non-linéarité et de résoudre le problème de réplication, certains auteurs tels que Glosten & Jagannathan (1994) et Agarwal & Naik (2005) utilisent des options d'achat et de vente. Toutefois, il reste difficile à déterminer quelles options introduire et combien. De plus, les options ordinaires étant limitées, cette solution offre des familles limitées de modèles.

Une autre approche proposée en premier lieu par Amin & Kat (2003) est de construire une stratégie pour chaque fonds de couverture afin de reproduire toutes les caractéristiques statistiques des rendements. Cette idée a été ensuite développée par Kat & Palaro (2005) puis améliorée par Papageorgiou *et al.* (2008). Il en résulte la reproduction de la distribution des rendements du fonds de couverture mais aussi de sa dépendance avec le portefeuille que détient l'investisseur. Cette approche permet de constituer un portefeuille de réplication en échangeant le portefeuille de l'investisseur et un actif de réserve d'une manière dynamique, ce qui permet d'évaluer ex-post la performance du fonds de couverture et aussi de répliquer les rendements futurs en investissant dans la stratégie de réplication au lieu d'investir dans le fonds. C'est cette approche qui nous intéresse et qui sera exposée dans ce qui suit en parcourant les différents travaux de chacun de ces auteurs.

2.1 Amin & Kat (2003)

Dans le but de chercher une mesure de performance appropriée à la nature des fonds de couverture, Amin & Kat (2003) proposent de construire des contrats d'options spécifiques pour chaque fonds de couverture avec une fonction de paiement déterminée par les rendements historiques du fonds. En effet, dans le cas de rendements non normaux, il convient de considérer toute leur distri-

bution. Il faut donc construire un portefeuille de réplique ayant les mêmes caractéristiques statistiques que le fonds de couverture. Cette approche de réplique est basée sur un modèle d'évaluation proposé par Dybvig (1988) : *payoff distribution pricing model*. Ce modèle suppose que les préférences des agents dépendent de la distribution d'un seul actif, qu'ils préfèrent le plus au moins et que les marchés sont parfaits. Il en découle que les investisseurs qui maximisent leurs utilités espérées voudraient que le bien-être en fin de période soit une fonction monotone non décroissante de la densité du prix de l'actif en question. La construction dynamique de ce portefeuille résout le problème de non-normalité des rendements.

Afin d'évaluer la performance du fonds de couverture, il faut trouver la stratégie la moins coûteuse qui offre la fonction de paiement désirée et comparer son coût avec celui de l'investissement dans le fonds. La première étape est de déterminer la fonction de paiement de fin du mois du fonds de couverture en utilisant des données mensuelles. Aucune hypothèse n'est émise quant à la nature de la distribution de ces rendements. Cependant, l'indice de référence utilisé pour la réplique évolue dans un contexte Black et Scholes et suit une loi normale ; le S&P 500 est choisi dans un souci de diversification. Les auteurs supposent que la fonction de paiement doit être une fonction non décroissante de l'indice de référence et indépendante du parcours de celui-ci. Cette hypothèse découle des travaux de Cox & Leland (2000) qui démontrent que, dans un cadre Black et Scholes, les stratégies dépendantes du parcours sont inefficaces parce qu'il existe des stratégies indépendantes moins coûteuses et qui permettent d'obtenir la même fonction de paiement.

Après avoir trouvé la stratégie qui va permettre d'avoir la fonction de paiement souhaitée, la deuxième étape consiste à la tarifier. Les auteurs utilisent la simulation Monte Carlo pour générer les paiements de fin du mois et procèdent ensuite à une tarification en monde neutre au risque. Enfin, la dernière étape est de comparer les coûts d'investissement dans le fonds avec les coûts de cette stratégie et déterminer si le fonds a réalisé une bonne performance ou s'il aurait pu faire mieux.

La limite principale de ce travail est que le fonds de couverture est considéré

indépendamment du marché alors que la dépendance avec ce dernier doit être prise en compte surtout que la corrélation n'est pas suffisante.

2.2 Kat & Palaro (2005)

Kat & Palaro (2005) se sont intéressés à la distribution jointe du fonds de couverture et du portefeuille de l'investisseur. Ils proposent une version en temps continu du modèle de Dybvig (1988) et postulent de plus que les préférences d'un investisseur dépendent de la distribution jointe de son investissement et du portefeuille qu'il détient déjà, en fin de la période d'investissement. Il est donc possible de reproduire la distribution du fonds de couverture ainsi que sa dépendance avec le portefeuille de l'investisseur en formant une stratégie d'investissement faisant intervenir uniquement le portefeuille de l'investisseur et un actif de réserve.

La première étape est de collecter les données historiques du fonds de couverture, de l'actif de réserve et du portefeuille de l'investisseur afin de déterminer la distribution jointe des rendements du fonds et du portefeuille de l'investisseur et celle de ce dernier avec l'actif de réserve. Puisqu'il s'agit de reproduire la distribution du fonds à la fin du mois, ce sont les rendements mensuels qui sont utilisés. Kat & Palaro (2005) utilisent les copules pour reproduire la dépendance entre les rendements mensuels.

L'étape suivante est de calculer une fonction du portefeuille de l'investisseur et de l'actif de réserve donnant des rendements ayant la même distribution que celle du fonds conditionnellement au portefeuille de l'investisseur.

Ainsi la fonction de paiement est une fonction g telle que

$$P(S^P \leq x, g(S^P, S^R) \leq y) = P(S^P \leq x, S^I \leq y), \forall x, y \in \mathbb{R},$$

avec S^P le paiement de fin du mois du portefeuille de l'investisseur, S^R celui de l'actif de réserve et S^I celui du fonds de couverture.

A la fin du mois, le flux généré par la stratégie de réplcation avec un investis-

sement initial de 100 est égal à :

$$g^*(S^P, S^R) = 100 \exp \left\{ g \left(\log \left(\frac{S^P}{100} \right), \log \left(\frac{S^R}{100} \right) \right) \right\}.$$

Pour tarifier la fonction de paiement, Kat & Palaro (2005) calculent une espérance des paiements sous la probabilité neutre au risque. Le coût de cette stratégie permet de calculer son rendement et de décider s'il est intéressant de la mettre en oeuvre.

Une fois la fonction de paiement déterminée, la troisième étape est de construire la stratégie dynamique de réplication. Kat & Palaro (2005) font du *delta hedging* ce qui ne garantit pas d'obtenir de bons résultats puisqu'on sait que le contexte Black et Scholes n'est pas adapté aux fonds de couverture.

Kat & Palaro (2005) composent le portefeuille de l'investisseur du S&P 500 et de bons de trésor américains à long terme à égales proportions. Quant à l'actif de réserve, ils prennent des Eurodollar futures.

2.3 Papageorgiou *et al.* (2008)

Papageorgiou *et al.* (2008) présentent plusieurs critiques de l'article de Kat & Palaro (2005) et proposent une amélioration de leur méthode de réplication. Premièrement, Kat & Palaro (2005) utilisent les copules pour modéliser les dépendances entre les différents rendements mensuels. Cependant en passant à des rendements journaliers pour les besoins de la réplication, on n'est pas certain de pouvoir conserver cette structure de dépendance car il n'est pas encore prouvé que les propriétés des copules soient conservées quand on change d'horizon. C'est pour cette raison que Papageorgiou *et al.* (2008) utilisent des mélanges gaussiens pour modéliser la distribution jointe bivariée du portefeuille de l'investisseur et l'actif de réserve. Ce choix est justifié par la capacité des mélanges de lois à capturer l'asymétrie et parce qu'ils sont infiniment divisibles. Ainsi en passant de données journalières à des données mensuelles afin de construire la stratégie de réplication, la structure de dépendance est conservée.

La méthode d'estimation est basée sur l'algorithme de Dempster *et al.* (1977) et des tests d'ajustement sont effectués pour déterminer le nombre optimal de régimes suivant Genest *et al.* (2009).

Quant aux rendements du fonds de couverture, Papageorgiou *et al.* (2008) modélisent d'abord la distribution marginale du fond, puis la copule qui décrit sa dépendance avec le portefeuille de l'investisseur. L'utilisation des copules ne pose plus de problème dans ce cas car seuls les rendements mensuels du fonds nous intéressent. Bien entendu, il s'agit d'estimation si on souhaite répliquer un fonds existant et de choix s'il s'agit d'un fonds synthétique. La fonction g est calculée de la même manière que Kat & Palaro (2005).

Papageorgiou *et al.* (2008) déterminent la stratégie optimale de réplication en minimisant l'espérance du carré de l'erreur de réplication en développant les résultats de Schweizer (1995) et en utilisant, contrairement à Kat & Palaro (2005), une évaluation sous la probabilité historique. La stratégie de réplication étant autofinancée, les auteurs parviennent à déterminer les proportions journalières à investir dans l'actif sans risque, l'actif de réserve et le portefeuille de l'investisseur de manière récursive. Dans le cas où les actifs sont des processus Markoviens, la méthodologie de Del Moral *et al.* (2006) est utilisée. Papageorgiou *et al.* (2008) montrent, en utilisant des fonds synthétiques, que cette stratégie de réplication optimale est plus efficace que celle du *delta hedging* dans un cadre Black et Scholes. En effet, des données simulées permettent de comparer le carré de l'erreur de réplication de l'une et de l'autre méthode, il s'avère que l'erreur est remarquablement plus faible en utilisant le *hedging* optimal.

Papageorgiou *et al.* (2008) limitent les actifs utilisés pour former le portefeuille de l'investisseur ainsi que l'actif de réserve aux contrats futures et aux ETF (*Exchange Traded Funds*) en raison de leur liquidité. De plus, puisqu'ils font l'objet d'échanges quotidiens, leurs faibles coûts de transactions constituent un avantage. Les auteurs prennent le même portefeuille de l'investisseur que Kat & Palaro (2005) mais prennent différents actifs de réserve. Ils réussissent à répliquer les indices HFRI et EDHEC sur la période allant de 1997 à 2006. Il en ressort que le modèle réplique correctement les caractéristiques statistiques

des fonds de couverture et le surperforme. En particulier, la volatilité et la dépendance mesurée par le tau de Kendall sont reproduites avec précision. Par ailleurs, il s'avère que le choix de l'actif de réserve affecte le rendement moyen de la stratégie de réplication mais n'influence presque pas les autres moments de la distribution.

Cette approche de réplication basée sur le modèle de Dybvig (1988) a permis à Kat & Palaro (2005) de trouver une mesure de performance adaptée aux fonds de couverture caractérisés par leurs rendements élevés, non-normaux et ayant une relation non-linéaire avec le marché. Ensuite, ce sont Papageorgiou *et al.* (2008) qui ont pu déterminer, en développant les travaux de Schweizer (1995), la stratégie optimale de réplication. Ainsi, cette méthode a fait ses preuves et permet de répliquer correctement les propriétés statistiques des fonds de couverture. Maintenant, il serait intéressant de généraliser la méthode de Papageorgiou *et al.* (2008) et de pouvoir répliquer la distribution d'un fonds de couverture conditionnellement à plusieurs portefeuilles détenus par l'investisseur ce qui va permettre une plus grande flexibilité de la méthode.

Chapitre 3

Modélisation des rendements

L'idée est de reproduire les rendements du fonds de couverture en considérant $n - 1$ portefeuilles détenus par l'investisseur et ce, en investissant dans ces derniers ainsi que dans un actif de réserve. La première étape est de modéliser les rendements des différents actifs. Nous utiliserons pour cette fin des mélanges Gaussiens qui permettent de capturer plusieurs formes de distributions. Quant au fonds de couverture, on choisit sa distribution marginale mensuelle et on utilise les copules pour capturer sa dépendance avec les différents portefeuilles de référence. On aura ainsi une copule à n dimensions.

Dans ce chapitre, nous exposons en premier lieu, les propriétés des mélanges Gaussiens et la méthode d'estimation utilisée. Ensuite, nous parcourons les différentes définitions et propriétés des copules qui vont être utilisées dans le cadre de ce travail.

3.1 Les mélanges Gaussiens

Modéliser la distribution jointe des portefeuilles de référence et de l'actif de réserve en utilisant des mélanges Gaussiens permet de capturer différents niveaux d'asymétrie et de reproduire différentes formes de distribution. De plus, ils permettent de conserver la bonne distribution en passant de l'horizon journalier à l'horizon mensuel.

3.1.1 Définition

Un mélange Gaussien est caractérisé par le nombre de régimes m et les paramètres $(\pi_k)_{k=1}^m$, $(\mu_k)_{k=1}^m$ et $(A_k)_{k=1}^m$. Sa densité est donnée par

$$f_X(x) = \sum_{k=1}^m \pi_k \phi_n(x, \mu_k, A_k),$$

où $\phi_n(x, \mu_k, A_k) = \frac{e^{-\frac{1}{2}(x-\mu_k)^\top A_k^{-1}(x-\mu_k)}}{(2\pi)^{n/2}|A_k|^{1/2}}$ est la densité d'un vecteur Gaussien multivarié de dimension n avec moyenne $\mu_k = (\mu_{k1}, \dots, \mu_{kn})^\top$ et matrice de variance-covariance A_k , et $|A_k|$ dénote le déterminant de la matrice A_k .

Sa fonction de répartition est donnée par $F_X(x) = \sum_{k=1}^m \pi_k \Phi_n(x, \mu_k, A_k)$ avec $\Phi_n(x, \mu, A)$ est la fonction de répartition de la loi normale multivariée de dimension n , de paramètres μ et A .

3.1.2 Propriétés ¹

Propriété 1 *La somme de mélanges Gaussiens indépendants reste un mélange Gaussien. Ainsi, si X_1, \dots, X_N sont indépendants et sont identiquement distribués selon un mélange Gaussien alors $X = X_1 + \dots + X_N$ est aussi un mélange Gaussien. Afin d'en déterminer les paramètres, posons*

$$\mathcal{A} = \{\alpha = (\alpha_1, \dots, \alpha_m); \alpha_j \geq 0 \text{ et } \alpha_1 + \dots + \alpha_m = N\},$$

où $\alpha_1, \dots, \alpha_m$ sont des entiers. Alors $\text{card}(\mathcal{A}) = \binom{N+m-1}{m-1}$, il existe $\binom{N+m-1}{m-1}$ régimes et les paramètres $(\pi_\alpha)_{\alpha \in \mathcal{A}}$, $(\mu_\alpha)_{\alpha \in \mathcal{A}}$ et $(A_\alpha)_{\alpha \in \mathcal{A}}$ sont calculés de la manière suivante :

$$\pi_\alpha = \pi_{(\alpha_1, \dots, \alpha_m)} = \frac{n!}{\alpha_1! \dots \alpha_m!} \prod_{k=1}^m \pi_k^{\alpha_k},$$

$$\mu_\alpha = \sum_{k=1}^m \alpha_k \mu_k,$$

1. Cette sous-section est tirée de l'article Papageorgiou *et al.* (2008)

et

$$A_\alpha = \sum_{k=1}^m \alpha_k A_k.$$

Cette propriété des mélanges Gaussiens permet, en théorie, de trouver la distribution jointe des rendements mensuels après avoir estimé celle des rendements journaliers. Cependant, si N est assez grand alors m^N est très grand et il serait impossible de calculer les nouveaux paramètres. Dans ce cas, c'est possible de ré-estimer la loi jointe par un mélange Gaussien en utilisant les rendements mensuels simulés à partir de la loi des rendements journaliers.

Propriété 2 *Étant donné un vecteur $X = (X_1, \dots, X_n)^\top$ modélisé par un mélange Gaussien de paramètres m , $(\pi_k)_{k=1}^m$, $(\mu_k)_{k=1}^m$ et $(A_k)_{k=1}^m$, la distribution conditionnelle $F_{X_n|X_1, \dots, X_{n-1}}(x_n|x_1, \dots, x_{n-1})$ est aussi un mélange Gaussien de paramètres $\{\tilde{\pi}_k(x')\}_{k=1}^m$, $\{\tilde{\mu}_k(x')\}_{k=1}^m$ et $\{\tilde{\sigma}_k^2\}_{k=1}^m$, avec*

$$\tilde{\pi}_k(x') = \frac{\pi_k \phi_{n-1}(x', \mu'_k, \alpha_k)}{\sum_{j=1}^m \pi_j \phi_{n-1}(x', \mu'_j, \alpha_j)},$$

$$\tilde{\mu}_k(x') = \mu_k + B_k^\top (x' - \mu'_k)$$

et

$$\tilde{\sigma}_k^2 = \sigma_{kn}^2 - \beta_k^\top \alpha_k^{-1} \beta_k$$

où

$$x' = (x_1, \dots, x_{n-1})^\top, \mu'_k = (\mu_{k1}, \dots, \mu_{kn-1})^\top,$$

α_k est la matrice de covariance obtenue de A_k en enlevant la ligne n et la colonne n , $\beta_k = (A_{k,1,n}, \dots, A_{k,n-1,n})^\top$, et $B_k = \alpha_k^{-1} \beta_k$.

3.1.3 Estimation

L'estimation d'un mélange Gaussien à m régimes est basée sur l'algorithme EM de Dempster *et al.* (1977). Il s'agit d'un algorithme itératif qui alterne entre deux étapes. La première (E) est le calcul de l'espérance d'une log-vraisemblance en cohérence avec l'estimation actuelle de la distribution et la

deuxième (M) calcule les paramètres qui maximisent l'espérance de la log-vraisemblance trouvée à l'étape (E). Ces paramètres sont utilisés pour déterminer la distribution dans la prochaine itération (E).

Algorithme EM pour l'estimation d'un mélange Gaussien multivarié

Soit x_1, x_2, \dots, x_N un échantillon d'un mélange Gaussien multivarié de dimension n avec les paramètres $(\pi_k)_{k=1}^m$, $(\mu_k)_{k=1}^m$ et $(A_k)_{k=1}^m$. L'estimateur de départ de $\theta = (\pi, \mu, A)$ est dénoté par $\theta^{(0)} = (\pi^{(0)}, \mu^{(0)}, A^{(0)})$.

Étant donné un estimateur $\theta^{(\ell)} = (\pi^{(\ell)}, \mu^{(\ell)}, A^{(\ell)})$ à l'étape ℓ , on pose

$$\pi_k(x_i, \theta^{(\ell)}) = \frac{\pi_k^{(\ell)} \phi_n(x_i, \mu_k^{(\ell)}, A_k^{(\ell)})}{\sum_{j=1}^m \pi_j^{(\ell)} \phi_n(x_i, \mu_j^{(\ell)}, A_j^{(\ell)})}, \quad i = 1, \dots, N.$$

On calcule ensuite le nouvel estimateur $\theta^{(\ell+1)} = (\pi^{(\ell+1)}, \mu^{(\ell+1)}, A^{(\ell+1)})$ comme suit :

$$\pi_k^{(\ell+1)} = \frac{1}{N} \sum_{i=1}^N \pi_k(x_i, \theta^{(\ell)}),$$

$$\mu_k^{(\ell+1)} = \frac{1}{N} \sum_{i=1}^N x_i \pi_k(x_i, \theta^{(\ell)}) / \pi_k^{(\ell+1)},$$

et

$$A_k^{(\ell+1)} = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_k^{(\ell+1)})(x_i - \mu_k^{(\ell+1)})^\top \pi_k(x_i, \theta^{(\ell)}) / \pi_k^{(\ell+1)},$$

pour $k = 1, \dots, m$.

Quand ℓ augmente, les probabilités $\{\pi_k(y_i, \theta^{(\ell)}); k = 1, \dots, m; i = 1, \dots, N\}$ se stabilisent et les estimateurs convergent.

Choix des valeurs de départ

Le choix des valeurs de départ pour l'initialisation de l'algorithme qui estime un mélange Gaussien à m régimes se fait comme suit : les probabilités $(\pi_k)_{k=1}^m$ sont fixées uniformément égales à $\frac{1}{m}$; l'échantillon est ensuite divisé en m sous-

échantillons, $m - 1$ sous-échantillons de taille $\frac{N}{m}$ et un $m^{\text{ème}}$ sous-échantillon de taille $N - \frac{N(m-1)}{m}$.

Pour chaque sous-échantillon k , $k = 1, \dots, m$, on calcule μ_k , la moyenne de l'échantillon et A_k sa matrice de variance-covariance.

Choix du nombre de régimes

Après avoir estimé les paramètres, des tests d'ajustement sont nécessaires pour s'assurer d'avoir pris le nombre de régimes optimal. Le test proposé est celui basé sur les travaux de Genest *et al.* (2009) où la transformée de Rosenblatt est utilisée. Le nombre de régimes est déterminé par le plus petit entier m pour lequel le test d'ajustement basé sur m régimes donne une probabilité critique supérieure à 5%. Voir Papageorgiou *et al.* (2008) pour plus de détails.

3.2 Les copules

Les copules vont permettre de tenir compte de la structure de dépendance du fonds de couverture avec les différents portefeuilles détenus par l'investisseur. On aura ainsi une copule de dimension n . Nous présentons dans cette section les définitions et propriétés des copules qui seront utilisées par la suite.

3.2.1 Définition

Une copule C d'un vecteur aléatoire $X = (X_1, \dots, X_n)$ est la distribution jointe du vecteur aléatoire $U = (U_1, \dots, U_n)$, avec $U_1 = F_1(X_1), \dots, U_n = F_n(X_n)$, où F_1, \dots, F_n sont les fonctions de répartition respectives de X_1, \dots, X_n , supposées continues. On a alors

$$F_X(x) = C(F_1(x_1), \dots, F_n(x_n)), \quad \forall x_1, \dots, x_n \in \mathbb{R}. \quad (3.1)$$

où F_X est la fonction de répartition jointe du vecteur X .

Théorème 1 (Théorème de Sklar) *Si F_X est une fonction de répartition jointe du vecteur $X = (X_1, \dots, X_n)$ avec les fonctions de répartition marginales F_1, \dots, F_n , alors il existe une copule C telle que*

$$F_X(x) = C(F_1(x_1), \dots, F_n(x_n)), \forall (x_1, \dots, x_n) \in \mathbb{R}^2.$$

Si F_1, \dots, F_n sont continues alors C est unique.

Inversement, si C est une copule et F_1, \dots, F_n sont des fonctions de répartition alors la fonction F_X définie en (3.1) est une fonction de répartition jointe ayant pour marges F_1, \dots, F_n .

Ce théorème permet de caractériser une distribution jointe par les distributions marginales et une copule. Ainsi, il est possible de capturer la structure de dépendance du fonds de couverture avec les portefeuilles de référence sans avoir à poser des restrictions sur la distribution marginale du fonds.

L'identité suivante sera aussi utilisée : Soit $(U, V) \sim C(u, v)$, où C est une copule de dimension n , $U \sim \text{Unif}(0, 1)$ et $V \sim C(1, v)$, $v = (v_1, \dots, v_{n-1})$. Alors on a

$$C_{U|V}(u; v) = \frac{\partial_{v_1} \dots \partial_{v_{n-1}} C(u, v)}{\partial_{v_1} \dots \partial_{v_{n-1}} C(1, v)}. \quad (3.2)$$

3.2.2 Copules archimédiennes

Dans ce travail, nous nous limiterons à la famille des copules archimédiennes telles que la copule de Clayton, de Gumbel et de Frank.

Les copules archimédiennes sont caractérisées par leur générateur ϕ . On pose $D_\phi = \{u = (u_1, \dots, u_n) \in (0, 1]^n; \phi(u_1) + \dots + \phi(u_n) < \phi(0)\}$ et on définit la copule de générateur ϕ par

$$C(u) = \begin{cases} \phi^{-1}\{\phi(u_1) + \dots + \phi(u_n)\} & \text{si } u \in D_\phi \\ 0 & \text{sinon} \end{cases}$$

Pour la copule de Clayton, on a :

$$\phi_\theta(t) = \frac{t^{-\theta} - 1}{\theta}, \quad \theta \geq -\frac{1}{n-1},$$

Pour $-\frac{1}{n-1} \leq \theta < 0$

$$C_\theta(u) = \left(\left[\sum_{i=1}^n u_i^{-\theta} - n + 1 \right]^+ \right)^{\frac{1}{\theta}},$$

et pour $\theta > 0$

$$C_\theta(u) = \left(\sum_{i=1}^n u_i^{-\theta} - n + 1 \right)^{\frac{1}{\theta}}.$$

Pour la copule de Gumbel, on a :

$$\phi_\theta(t) = (-\ln t)^{\frac{1}{\theta}}, \quad 0 < \theta \leq 1,$$

et

$$C_\theta(u) = \exp \left[- \left\{ \sum_{i=1}^n (-\ln u_i)^{\frac{1}{\theta}} \right\}^\theta \right].$$

Pour la copule de Frank,

$$\phi_\theta(t) = -\ln \left(\frac{1 - \theta^t}{1 - \theta} \right), \quad \theta > 0.$$

Pour $\theta > 1$

$$C_\theta(u) = \ln \left\{ 1 + (\theta - 1) \prod_{i=1}^n \left(\frac{\theta^{u_i} - 1}{\theta - 1} \right) \right\} / \ln(\theta),$$

et pour $0 < \theta < 1$

$$C_\theta(u) = \ln \left\{ 1 - (1 - \theta) \prod_{i=1}^n \left(\frac{1 - \theta^{u_i}}{1 - \theta} \right) \right\} / \ln(\theta).$$

Le tau de Kendall est considéré comme mesure de la dépendance. Sa valeur, pour les copules précédentes est :

Pour la copule de Clayton,

$$\tau = \frac{\theta}{\theta + 2},$$

tandis que pour la copule de Gumbel,

$$\tau = 1 - \theta,$$

et finalement pour la copule de Frank, on a

$$\tau = \frac{\ln(\theta)^2 + 4 \ln(\theta) + 4 \operatorname{dilog}(\theta)}{\ln(\theta)^2}$$

où $\operatorname{dilog}(x) = \int_1^x \frac{\ln t}{1-t} dt$.

3.2.3 Propriétés

Dans le cas d'une copule archimédienne de générateur ϕ , (3.2) donne

$$C_{U|V}(u|v) = \frac{(\phi^{-1})^{(n-1)} \circ \phi \circ C(u, v)}{(\phi^{-1})^{(n-1)} \circ \phi \circ C(1, v)} = \frac{f_{n-1}\{C(u, v)\}}{f_{n-1}\{C(1, v)\}},$$

avec $f_i(t) = \left. \frac{d^{i+1}}{ds^{i+1}} \phi^{-1} \right|_{s=\phi(t)}$.

Pour la copule de Clayton,

$$f_i(t) = (-1)^{i+1} t^{1+(i+1)\theta} \prod_{j=0}^i (1 + j\theta).$$

Pour la copule de Gumbel,

$$f_i(t) = (-1)^{i+1} \theta t p_i(-\ln t) (-\ln t)^{1-(i+1)/\theta},$$

avec $p_0(x) = 1$ et pour $i \geq 0$,

$$p_i(x) = \theta x \{p_{i-1}(x) - p'_{i-1}(x)\} + (i - \theta) p_{i-1}(x).$$

Pour la copule de Frank,

$$f_i(t) = q_i(\theta^{-t}) / \ln(\theta),$$

avec $q_0(x) = x - 1$ et pour $i \geq 0$,

$$q_i(x) = x(x - 1) q'_{i-1}(x).$$

La dépendance entre les rendements du fonds de couverture et ceux des $(n - 1)$ portefeuilles va être choisie en choisissant le copule de paramètre θ . Par la suite, le choix de la distribution des rendements mensuels du fonds de couverture sera sans contrainte. En effet, dans ce travail, nous n'auront pas à répliquer les rendements d'un fonds de couverture existant, nous nous limiterons sans perte de généralité à un fonds synthétique.

Les formules précédentes pour les fonction f_i proviennent de Barbe *et al.* (1996).

Chapitre 4

Stratégie de répliation

Une fois les distributions estimées, l'étape suivante est de déterminer la fonction de paiement qui permet de générer des rendements ayant la même distribution que le fonds de couverture et ayant la même structure de dépendance avec les portefeuilles de référence (déterminés par l'investisseur), et ce en investissant dans un portefeuille de répliation ayant pour composantes les portefeuilles de référence ainsi qu'un actif de réserve. Il faut donc trouver la fonction de paiement de fin du mois permettant de répliquer les rendements du fonds et ensuite déterminer la stratégie dynamique de répliation qui va permettre d'avoir le paiement désiré en fin du mois.

4.1 Fonction de paiement

Nous commençons par la fonction qui permet de générer en fin du mois un rendement ayant la même distribution du fonds de couverture conditionnellement aux rendements des portefeuilles de référence. On note par $Y = (Y_1, \dots, Y_{n-1})$ les rendements en fin du mois des portefeuilles de référence, par X celui de l'actif de réserve et I celui du fonds de couverture qu'on cherche à répliquer. Ainsi la fonction de paiement est la fonction g telle que pour tout $(x, y) \in \mathbb{R}^n$, on a

$$P(I \leq x | Y = y) = P(g(X, y) \leq x | Y = y). \quad (4.0)$$

Il est facile de constater que la solution est similaire à celle trouvée par Kat & Palaro (2005) dans le cas $n = 2$. Plus précisément,

$$g(x, y) = F_{I|Y}^{-1}[F_{X|Y}(x; y); y], \quad x \in \mathbb{R}, y \in \mathbb{R}^{n-1},$$

où $F_{X|Y}$ est la fonction de répartition de X conditionnellement à Y , $F_{I|Y}$ est la fonction de répartition rendement I du fonds à répliquer conditionnellement à Y , et $F_{I|Y}^{-1}$ est l'inverse de la fonction de répartition $F_{I|Y}$. La fonction g étant une composition de deux fonctions non décroissantes en x , elle est non décroissante en x .

La preuve que la fonction g a les propriétés prescrites est assez simple. D'après le théorème dans , on sait que $U = F_{X|Y}(X; Y)$ est distribuée selon une loi uniforme sur $(0, 1)$ et est indépendante de P . Donc $F_{I|Y}^{-1}[U; Y]$ a la même loi que I conditionnellement à Y .

Si G représente la fonction de répartition marginale du rendement mensuel du fonds de couverture et C est la copule associée à (I, Y) , traduisant la dépendance entre le fonds et les portefeuilles de référence, on a ¹

$$F_{I|Y}(x; y_1, \dots, y_{n-1}) = C_{U|V}[G(x); v_1 = F_1(y_1), \dots, v_{n-1} = F_{n-1}(y_{n-1})],$$

où $C_{U|V}$ est déterminée par (3.2).

La fonction g s'écrit donc

$$g(x, y) = G^{-1} \left\{ C_{U|V}^{-1} \left[F_{R|P}(x; y_1, \dots, y_{n-1}); F_1(y_1), \dots, F_{n-1}(y_{n-1}) \right] \right\}.$$

Il faut noter que seule la fonction de répartition conditionnelle $C_{U|V}$ est importante et non la copule C , ce qui permet beaucoup plus de flexibilité car en spécifiant C , on spécifie en outre la dépendance entre les portefeuilles de référence, ce qui n'est pas le cas avec la fonction de densité jointe.

La valeur du paiement de fin de mois à répliquer à partir d'un montant initial

1. Tiré des notes de cours de Pr. Bruno Rémillard : Méthodes statistiques en ingénierie financière

de 100 et donnant le rendement $g(x, y)$ est donc égal à

$$100 \exp \{g(x, y)\}.$$

4.2 Portefeuille de répliation

Après avoir trouvé la fonction de paiement, il faut déterminer la stratégie d'investissement optimale. Papageorgiou *et al.* (2008) minimisent l'espérance du carré de l'erreur de répliation suivante :

$$E\{\beta_T^2[V_T(V_0, \varphi) - C_T]^2\},$$

où β_T est le facteur d'actualisation à la fin du mois, $V_T(V_0, \varphi)$ est la valeur à la fin du mois du portefeuille de répliation ayant pour valeur initiale V_0 et pour pondérations le vecteur φ , et C_T est le paiement de fin de mois du fonds de couverture. Notons que suivant Papageorgiou *et al.* (2008), tous les calculs sont faits sous les probabilités historiques. Nous suivrons la méthodologie de la couverture optimale présentée par Rémillard & Rubenthaler (2009).

4.2.1 La couverture optimale

Soit le processus de prix multidimensionnel $S_t = (S_t^{(1)}, \dots, S_t^{(n)})$ défini dans un espace probabilisé (Ω, P, \mathcal{F}) avec une filtration $\mathbb{F} = \{\mathcal{F}_0, \dots, \mathcal{F}_T\}$. β_t dénote le facteur d'actualisation et est un processus prévisible, c'est à dire que β_t est \mathcal{F}_{t-1} mesurable pour tout $t = 1, \dots, T$.

La stratégie de répliation dynamique est définie par une valeur initiale V_0 et une séquence de vecteur de poids aléatoire $\varphi = (\varphi)_{t=0}^T$, où pour chaque $i = 1, \dots, n$, $\varphi_t^{(i)}$ est le nombre de parts du titre $S^{(i)}$ investis durant la période $(t-1, t]$. φ est aussi un processus prévisible.

Initialement, $\varphi_0 = \varphi_1$ et le portefeuille de départ à une valeur égale à V_0 . Il en

résulte que le montant investi dans l'actif sans risque est égal à

$$V_0 - \sum_{i=1}^n \varphi_1^{(i)} S_0^{(j)} = V_0 - \varphi_1^\top S_0.$$

Le portefeuille de répliation doit être autofinancé c'est à dire que pour tout $t = 1, \dots, T$ on a :

$$V_t(V_0, \varphi) - V_{t-1}(V_0, \varphi) = \varphi_t^\top (\beta_t S_t - \beta_{t-1} S_{t-1}),$$

où V_t représente la valeur actualisée du portefeuille à la période t .

Cette condition permet d'établir que

$$V_T = V_T(V_0, \varphi) = V_0 + \sum_{t=1}^T \varphi_t^\top (\beta_t S_t - \beta_{t-1} S_{t-1}).$$

Le problème de répliation pour un certain paiement C est donc équivalent à trouver la stratégie (V_0, φ) de telle sorte que

$$G(V_0, \varphi) = \beta_T C - V_T(V_0, \varphi)$$

soit le plus faible possible.

Suivant Papageorgiou *et al.* (2008), l'espérance du carré de l'erreur de couverture est utilisée comme mesure de la qualité de la répliation. Pour cela, Schweizer (1995) suppose que les prix $S_t^{(i)}$ ont des seconds moments finis. De plus, la stratégie de couverture φ satisfait une condition similaire c'est à dire que pour tout $t = 1, \dots, T$, $\varphi_t^\top (\beta_t S_t - \beta_{t-1} S_{t-1})$ a des seconds moments finis.

On pose $\Delta_t = \beta_t S_t - \beta_{t-1} S_{t-1}$, $t = 1, \dots, T$.

L'objectif est de trouver le montant de l'investissement initial V_0 et la stratégie d'investissement $\varphi = (\varphi_t)_{t=1}^n$ qui minimise l'espérance du carré de l'erreur de répliation $E \left[\{G(V_0, \varphi)\}^2 \right]$, avec

$$G = G(V_0, \varphi) = \beta_T C_T - V_T,$$

et

$$V_t = V_0 + \sum_{j=1}^t \varphi_j^\top \Delta_j, t = 0, \dots, T.$$

On pose $P_{T+1} = 1$, et pour tout $t = T, \dots, 1$, on calcule

$$\begin{aligned} A_t &= E(\Delta_t \Delta_t^\top P_{t+1} | \mathcal{F}_{t-1}), \\ b_t &= A_t^{-1} E(\Delta_t P_{t+1} | \mathcal{F}_{t-1}), \\ \alpha_t &= A_t^{-1} E(\beta_t C_T \Delta_t P_{t+1} | \mathcal{F}_{t-1}), \\ P_t &= \prod_{j=t}^T (1 - b_j^\top \Delta_j). \end{aligned}$$

Théorème 1 *En supposant que $E(P_t | \mathcal{F}_{t-1}) \neq 0$, pour $t = 1, \dots, T$, alors la solution (V_0, φ) du problème de minimisation est donnée par*

$$V_0 = E(\beta_T C_T P_1) / E(P_1),$$

et

$$\varphi_t = \alpha_t - V_{t-1} b_t, \quad t = 1, \dots, T.$$

La preuve de ce théorème est fournie dans Rémillard & Rubenthaler (2009).

4.2.2 Le cas markovien

Si les processus de prix S sont markoviens c'est à dire que la loi de S_t sachant \mathcal{F}_{t-1} dépend seulement de S_{t-1} , et si le paiement terminal C_T dépend uniquement des prix terminaux, c'est à dire que $C_T = C_T(S_T)$, alors Rémillard & Rubenthaler (2009) montrent que $C_t = C_t(S_t)$, $\alpha_t = \alpha_t(S_{t-1})$ et $b_t = b_t(S_{t-1})$.

En posant $c_t(S_{t-1}) = E(P_t | \mathcal{F}_{t-1})$ pour $t = 1, \dots, T + 1$ et en supposant que

$\beta_t = \beta_1 \beta_{t-1}$, on obtient pour $t = T, \dots, 1$,

$$\begin{aligned} A_t(s) &= \beta_{t-1}^2 E_s \{ (\beta_1 S_1 - s) (\beta_1 S_1 - s)^\top c_{t+1}(S_1) \} \\ b_t(s) &= \beta_{t-1} A_t^{-1}(s) E_{s,h} \{ (\beta_1 S_1 - s) c_{t+1}(S_1) \}, \\ c_t(s) &= E_s \{ c_{t+1}(S_1) \} - b_t^\top(s) A_t(s) b_t(s), \\ C_{t-1}(s) &= \frac{\beta_1}{c_t(s)} E_s \{ C_t(S_1) c_{t+1}(S_1) [1 - \beta_{t-1} b_t(s)^\top (\beta_1 S_1 - s)] \}, \\ \alpha_t(s) &= \beta_t \beta_{t-1} A_t^{-1}(s) E_s \{ C_t(S_1) c_{t-1}(S_1) (\beta_1 S_1 - s) \}. \end{aligned}$$

Ces équations se simplifient dans le cas du modèle étudié ici qui se nomme aussi marche aléatoire géométrique. En effet, $X_t^i = \beta_t S_t^i = \beta_{t-1} S_{t-1}^i (1 + \xi_t^i) = X_{t-1}^i (1 + \xi_t^i)$ pour $i = 1, \dots, n$, où les ξ_t sont indépendants, de moyenne μ et $B = E(\xi_t \xi_t^\top)$. Dans ce cas on a $c_t = c^{T-t+1}$, avec $c = 1 - \mu^\top B^{-1} \mu$, $A_t(s) = \beta_{t-1}^2 D(s) B D(s) c^{T-t}$, $b_t(s) = \frac{1}{\beta_{t-1}} D^{-1}(s) B^{-1} \mu$, et $D(s)$ est la matrice diagonale éléments diagonaux s_1, \dots, s_n .

4.2.3 Implémentation de la stratégie dynamique de répliation

Pour notre modèle, nous posons $X_t^i = \beta_t S_t^i = \beta_{t-1} S_{t-1}^i (1 + \xi_t^i) = X_{t-1}^i (1 + \xi_t^i)$ pour $i = 1, \dots, n$ et où $\xi_t = (\xi_t^1, \dots, \xi_t^n)^\top$ est un vecteur simulé à partir des paramètres du mélange Gaussien estimés précédemment, modélisant la distribution jointe des rendements des n portefeuilles.

Nous commençons par construire une grille en discrétisant l'espace des 3 portefeuilles. Soient $R_{max}^{(i)}$ et $R_{min}^{(i)}$ respectivement les rendements maximum et minimum du portefeuille i et notons par Δ l'intervalle de de discrétisation. Nous avons alors pour $i = 1, \dots, n$

$$\begin{aligned} R_{max}^{(i)} &= \max_k (\mu_{ki} + 4\sigma_{ki}) \\ R_{min}^{(i)} &= \min_k (\mu_{ki} - 4\sigma_{ki}), \quad k = 1, \dots, m \end{aligned}$$

et

$$\Delta^{(i)} = (R_{max}^{(i)} - R_{min}^{(i)})/p,$$

où p est le nombre de discrétisation. Nous auront ainsi p^n points sur la grille.

L'algorithme est basée sur des simulations Monte Carlo combinées à une séquence d'approximations. Il peut être décrit comme suit :

- Définir $\tilde{C}_T = C_T$;
- Pour chaque $t = T, \dots, 1$
 - générer $\xi_{1,t}, \dots, \xi_{N,t}$;
 - Pour tout point de la grille , calculer

$$\begin{aligned}\mu_t &= \frac{1}{N} \sum_{i=1}^N \xi_{i,t}, \\ B_t &= \frac{1}{N} \sum_{i=1}^N \xi_{i,t} \xi_{i,t}^\top, \\ c_t &= 1 - \mu_t^\top B_t \mu_t, \\ C_t &= \frac{1}{N} \sum_{i=1}^N \tilde{C}_{t+1} \frac{1 - B_t^{-1} \xi_{i,t} \mu_t}{c_t}, \\ \alpha_t &= \frac{1}{N} \sum_{i=1}^N \tilde{C}_{t+1} B_t^{-1} \mu_t.\end{aligned}$$

- Interpoler linéairement C_t et définir \tilde{C}_t à tous les points de la grille.

Cet algorithme calqué sur celui proposé par Papageorgiou *et al.* (2008) garantit la répliation optimale des rendements du fonds de couverture. Maintenant, il reste à choisir un fonds synthétique à réplier et à appliquer la méthode en utilisant des données de marché pour former les portefeuilles et l'actif de réserve, ce qui fera l'objet du prochain chapitre.

Chapitre 5

Résultats empiriques

Dans ce chapitre, nous mettrons en oeuvre la méthode de répliation présentée précédemment en prenant des données de marché afin de répliquer les rendements d'un fonds synthétique. Pour l'implémentation de notre méthode de répliation, nous choisissons, sans perte de généralité, de considérer un problème à trois dimensions, c'est à dire que nous prenons deux portefeuilles de référence auxquels s'ajoute un actif de réserve.

Nous commençons par former deux portefeuilles de référence et un portefeuille de réserve à partir de données de marché et nous procédons à l'estimation de la loi jointe des rendements de ces portefeuilles par un mélange Gaussien. Par la suite, nous choisirons la distribution marginale d'un fonds synthétique à répliquer en considérant différentes copules pour tenir compte de sa dépendance avec les deux portefeuilles de référence. Nous présenterons enfin les résultat de la répliation.

5.1 Choix des portefeuilles

Nous commençons par construire deux portefeuilles de référence notés $P^{(1)}$ et $P^{(2)}$ et un portefeuille de réserve noté $P^{(3)}$. Ces portefeuilles font l'objet d'échange quotidien, nous limitons donc le choix aux titres liquides ayant de faibles coûts de transaction. C'est pour cette raison que nous utilisons des

contrats futures. Le taux d'actualisation est le taux Libor BBA 1 mois. Les log-rendements des contrats futures sont calculés à partir du réinvestissement d'une stratégie de roulement (*rolling strategy*) sur le contrat le plus proche de la maturité sur l'échéancier Mars/juin/Septembre/Décembre, et au premier jour ouvrable du mois de maturité au prix de fermeture du jour précédent.

Le premier portefeuille de l'investisseur est composé d'actions et le deuxième d'obligations. Pour l'actif de réserve, on forme un portefeuille assez diversifié. Le tableau 5.1 fournit une description détaillée de la composition des trois portefeuilles.

$P^{(1)}$	60%	S&P/TSE 60 IX future
	40%	S&P500 EMINI future
$P^{(2)}$	100%	CAN 10YR BOND future
$P^{(3)}$	10%	E-mini NASDAQ-100 futures
	20%	Russell 2000 TR
	20%	MSCI Emerging Markets TR
	10%	GOLD 100 OZ future
	10%	WTI CRUDE future
	30%	US 2YR NOTE (CBT)

TABLE 5.1 – Composition des portefeuilles.

Nous utilisons un échantillon de rendements journaliers sur une période de 115 mois allant du 01/10/1999 jusqu'au 30/04/2009. En plus de la période totale, nous considérons deux sous-périodes, la première allant du 01/10/1999 au 31/12/2003 (51 mois) et la deuxième allant du 01/01/2004 au 30/04/2009 (64 mois). Le tableau 5.2 donne quelques statistiques descriptives des portefeuilles choisis en considérant la période totale mais aussi les deux sous-périodes.

Portefeuilles	Statistiques	Période 1 (99-09)	Période 2 (99-03)	Période 3 (04-09)
$P^{(1)}$	Moyenne	0.00007	0.00006	0.00008
	Écart-type	0.00836	0.00785	0.00874
	Asymétrie	-0.64470	-0.57495	-0.68121
	Aplatissement	11.54778	7.37982	13.50644
$P^{(2)}$	Moyenne	0.00008	0.000073	0.00009
	Écart-type	0.00372	0.003934	0.00355
	Asymétrie	-0.32614	-0.404236	-0.23816
	Aplatissement	5.05828	4.149706	6.02384
$P^{(3)}$	Moyenne	0.00014	0.000121	0.000161
	Écart-type	0.00779	0.007143	0.008280
	Asymétrie	-0.44181	-0.161823	-0.580953
	Aplatissement	8.14154	3.948375	9.751693

TABLE 5.2 – Statistiques descriptives pour les portefeuilles sur les 3 périodes.

5.2 Modélisation des rendements

La loi jointe des rendements des trois portefeuilles est modélisée par un mélange Gaussien multivarié avec m régimes noté MGM(m). Les tableaux 5.3, 5.4 et 5.5 fournissent les résultats des estimations de la loi jointe des rendements journaliers et mensuels sur toute la période ainsi que sur les deux sous-périodes en utilisant la méthode d'estimation présentée dans le chapitre 3. Les rendements journaliers sont modélisés avec un mélange Gaussien à cinq régimes pour la période totale (99-09) et pour la période 2 (99-03) alors que pour la période 3 (04-09), la loi des rendements journaliers est donnée par un mélange Gaussien à quatre régimes. Par contre, les rendements mensuels correspondants sont modélisés à l'aide d'un mélange Gaussien à deux régimes et ce, pour les trois périodes considérées. Il est prévisible que le nombre de régimes pour les rendements mensuels est moindre que celui des rendements journaliers car plus on augmente l'horizon d'investissement, plus les rendements tendent à être de loi normale.

Rendements journaliers MGM(5)				
π_k	μ_k	A_k		
0.292993	-0.001893	0.000071	-0.000007	0.000037
	0.000713	-0.000007	0.000009	-0.000006
	-0.003163	0.000037	-0.000006	0.000047
0.069351	0.002841	0.000077	0.000007	0.000076
	-0.004523	0.000007	0.000019	0.000017
	0.005476	0.000076	0.000017	0.000097
0.076443	0.002720	0.000018	0.000008	-0.000003
	0.000385	0.000008	0.000017	-0.000002
	0.005375	-0.000003	-0.000002	0.000008
0.506822	0.000748	0.000020	-0.000002	0.000014
	0.000247	-0.000002	0.000008	-0.000001
	0.000887	0.000014	-0.000001	0.000021
0.054392	-0.002708	0.000542	-0.000066	0.000360
	0.000620	-0.000066	0.000051	-0.000055
	-0.003104	0.000360	-0.000055	0.000386
Rendements mensuels MGM(2)				
π_k	μ_k	A_k		
0.593347	0.003947	0.001132	-0.000101	0.000831
	0.000334	-0.000101	0.000329	-0.000166
	0.005347	0.000831	-0.000166	0.001169
0.406653	-0.000508	0.002322	-0.000233	0.001475
	0.004215	-0.000233	0.000246	-0.000101
	0.001884	0.001475	-0.000101	0.001698

TABLE 5.3 – Paramètres des mélanges Gaussiens utilisés pour la modélisation des rendements sur la période 1 (99-09).

Rendements journaliers MGM(5)				
π_k	μ_k	A_k		
0.148714	0.001950	0.000027	0.000001	0.000017
	-0.002880	0.000001	0.000008	-0.000010
	-0.000278	0.000017	-0.000010	0.000045
0.127828	-0.000129	0.000134	-0.000003	0.000040
	0.000374	-0.000003	0.000009	-0.000003
	-0.001093	0.000040	-0.000003	0.000076
0.183805	-0.000740	0.000128	-0.000019	0.000101
	-0.001099	-0.000019	0.000036	-0.000009
	0.000051	0.000101	-0.000009	0.000102
0.180678	-0.003857	0.000021	-0.000003	0.000016
	0.002058	-0.000003	0.000007	-0.000003
	-0.004473	0.000016	-0.000003	0.000020
0.358974	0.001739	0.000024	0.000001	0.000010
	0.000791	0.000001	0.000008	-0.000001
	0.003064	0.000010	-0.000001	0.000014
Rendements mensuels MGM(2)				
π_k	μ_k	A_k		
0.349753	-0.002927	0.001337	-0.000104	0.001039
	0.004444	-0.000104	0.000200	-0.000119
	-0.000854	0.001039	-0.000119	0.001434
0.650247	0.004562	0.001351	-0.000111	0.000679
	0.000140	-0.000111	0.000387	-0.000072
	0.005211	0.000679	-0.000072	0.000934

TABLE 5.4 – Paramètres des mélanges Gaussiens utilisés pour la modélisation des rendements pour la période 2 (99-03).

Rendements journaliers MGM(4)				
π_k	μ_k	A_k		
0.056112	0.003732	0.000085	0.000008	0.000083
	-0.004652	0.000008	0.000013	0.000019
	0.007432	0.000083	0.000019	0.000111
0.749031	0.001314	0.000020	-0.000000	0.000012
	0.000197	-0.000000	0.000008	-0.000000
	0.001373	0.000012	-0.000000	0.000020
0.117727	-0.006774	0.000044	-0.000004	0.000009
	0.001155	-0.000004	0.000007	-0.000004
	-0.007828	0.000009	-0.000004	0.000017
0.077131	-0.003897	0.000559	-0.000064	0.000381
	0.000899	-0.000064	0.000048	-0.000057
	-0.004668	0.000381	-0.000057	0.000418

Rendements mensuels MGM(2)				
π_k	μ_k	A_k		
0.418199	0.006294	0.000901	-0.000052	0.000610
	0.001151	-0.000052	0.000196	-0.000058
	0.008343	0.000610	-0.000058	0.000823
0.581801	-0.001268	0.002333	-0.000276	0.001684
	0.002449	-0.000276	0.000299	-0.000249
	-0.000114	0.001684	-0.000249	0.002060

TABLE 5.5 – Paramètres des mélanges Gaussiens utilisés pour la modélisation des rendements pour la période 3 (04-09).

5.3 Résultats de la répliation

Dans ce travail, nous tentons de réplier les rendements d'un fonds synthétique. Nous choisissons un fonds avec des rendements Gaussiens de moyenne mensuelle égale à 1% et de volatilité 1% en considérant différentes copules pour tenir compte de sa dépendance avec les portefeuilles de référence et en considérant aussi le cas de l'indépendance.

Nous commençons par tester la capacité de la fonction g à reproduire des flux ayant les mêmes caractéristiques statistiques que ceux du fonds de couverture à la fin du mois. Les tableaux 5.6, 5.7 et 5.8 rapportent les propriétés statistiques du fonds comparées à celles obtenues par la fonction g_T en se basant sur 10000 simulations pour les trois périodes respectivement. Nous considérons les dépendances induites par les copules de Clayton, de Gumbel et de Frank avec un tau de Kendall égal à 0.3 pour modéliser la dépendance des rendements marginaux à réplier avec les rendements des portefeuilles de référence ainsi que le cas de l'indépendance.

Statistiques	Cible	g_T			
		Clayton	Gumbel	Frank	Indépendance
Moyenne	0.1	0.01118	0.01094	0.00700	0.01017
Écart type	0.08	0.00953	0.00898	0.01180	0.01002
Asymétrie	0	0.03268	-0.10355	0.13810	0.05289
Aplatissement	3	3.06726	3.05829	2.74312	2.97004

TABLE 5.6 – Paramètres cibles et ceux obtenus par la fonction g_T pour 10000 simulations selon les paramètres de la période 1.

Statistiques	Cible	g_T			
		Clayton	Gumbel	Frank	Indépendance
Moyenne	0.1	0.01118	0.01099	0.00710	0.01014
Écart type	0.08	0.00956	0.00905	0.01189	0.01001
Asymétrie	0	0.01864	-0.09879	0.15382	0.06249
Aplatissement	3	3.03397	3.04850	2.72936	2.98936

TABLE 5.7 – Paramètres cibles et ceux obtenus par la fonction g_T pour 10000 simulations selon les paramètres de la période 2.

Statistiques	Cible	g_T			
		Clayton	Gumbel	Frank	Indépendance
Moyenne	0.1	0.01114	0.01087	0.00703	0.01013
Écart type	0.08	0.00951	0.00896	0.01170	0.01001
Asymétrie	0	0.03704	-0.09683	0.15093	0.05071
Aplatissement	3	3.02849	3.04176	2.75429	2.96485

TABLE 5.8 – Paramètres cibles et ceux obtenus par la fonction g_T pour 10000 simulations selon les paramètres de la période 3.

Considérons maintenant le résultat de la méthode de réplication avec un investissement initial égal à 100. La réplication est intéressante à mettre en oeuvre si la construction du portefeuille de réplication coûte moins cher. Le tableau 5.9 rapporte le résultat de réplication du fonds de couverture décrit précédemment en utilisant les trois copules et dans le cas de l'indépendance, et en considérant les trois périodes. On remarque que dans tous les cas, la valeur de l'investissement initial est supérieure à 100. Ceci peut être attribué au choix du portefeuille de réserve. En effet, Papageorgiou *et al.* (2008) expliquent que le choix de l'actif de réserve peut affecter les résultats de la réplication au niveau du rendement moyen.

	V_0			
	Clayton	Gumbel	Frank	Indépendance
Période 1	100.68832	100.61500	100.33382	100.54353
Période 2	100.70821	100.65338	100.31486	100.55541
Période 3	100.64903	100.61967	100.31204	100.54850

TABLE 5.9 – Résultats de la réplication.

Examinons maintenant l'erreur de réplication. Les tableaux 5.10, 5.11 et 5.12 rapportent les erreurs moyennes et les racines carrées de la moyenne des erreurs au carré (RMSE) pour les trois périodes respectivement. Nous pouvons constater que l'erreur est centrée autour de zéro dans tous les cas, comme prévu par la théorie puisque V_0 est choisi de façon à avoir une erreur moyenne nulle. De plus, les graphes 5.1, 5.2, 5.3 et 5.4 représente la distribution de l'erreur de réplication pour la période 1 pour les trois copules et dans le cas de l'indépendance. Les graphes de la distribution des erreurs pour les périodes 2 et 3 sont semblables.

	Clayton	Gumbel	Frank	Indépendance
Erreur moyenne	0.00001362	0.00003193	0.00006658	0.00001144
RMSE	0.01130528	0.01114527	0.01171583	0.01410102

TABLE 5.10 – Erreur de réplication pour la période 1

	Clayton	Gumbel	Frank	Indépendance
Erreur moyenne	-0.00001626	0.00000920	0.00004257	-0.00002735
RMSE	0.01128122	0.01111760	0.01166290	0.01408088

TABLE 5.11 – Erreur de réplication pour la période 2

	Clayton	Gumbel	Frank	Indépendance
Erreur moyenne	-0.00000431	0.00001634	0.00005218	-0.00001301
RMSE	0.01130883	0.01116341	0.01173963	0.01409300

TABLE 5.12 – Erreur de réplication pour la période 3

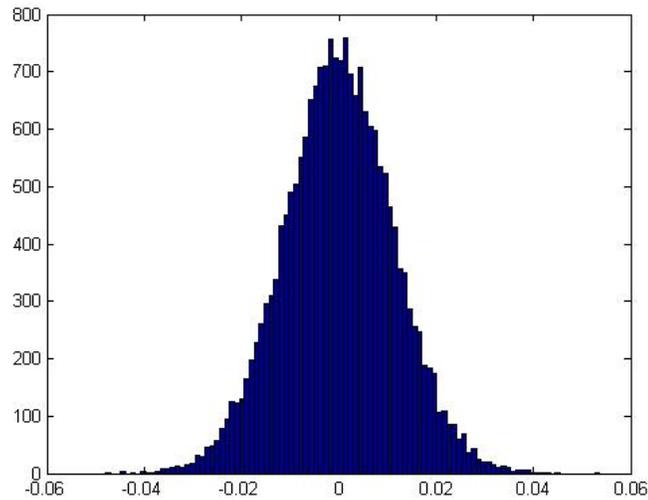


FIGURE 5.1 – Erreur de réplication pour la période 1 avec la copule de Clayton.

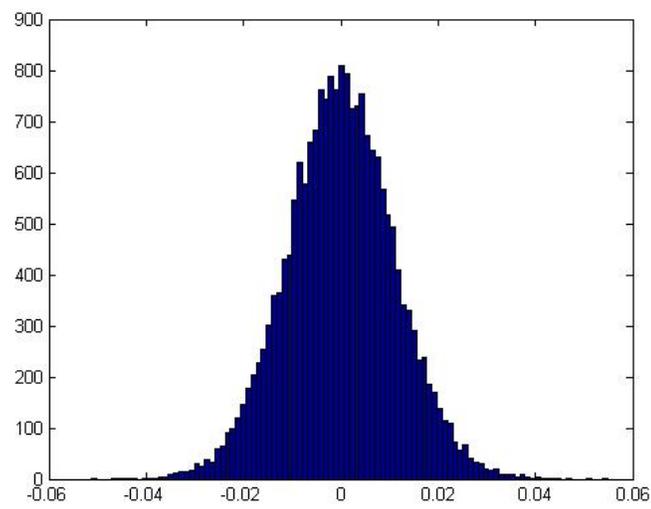


FIGURE 5.2 – Erreur de réplication pour la période 1 avec la copule de Gumbel.

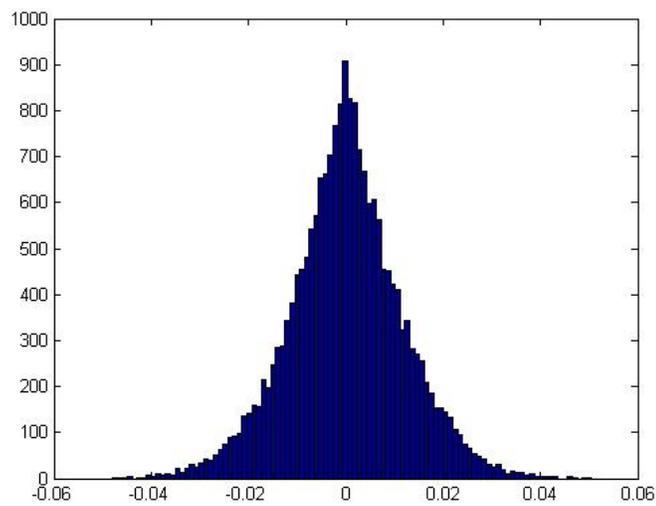


FIGURE 5.3 – Erreur de réplication pour la période 1 avec la copule de Frank.

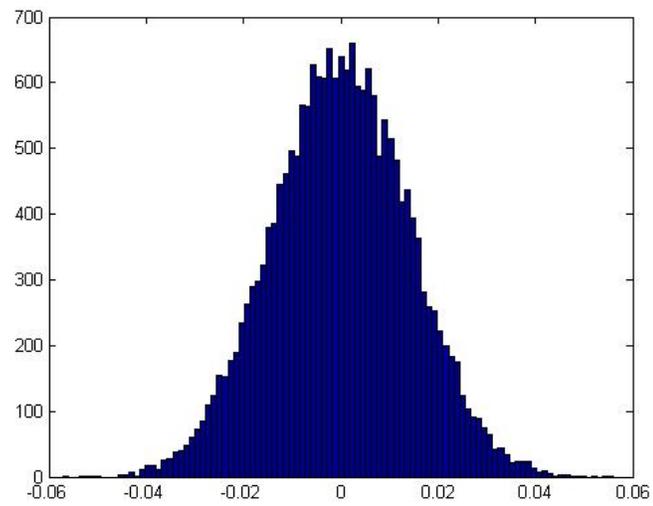


FIGURE 5.4 – Erreur de réplication pour la période 1 dans le cas de l'indépendance.

Chapitre 6

Conclusion

Dans ce travail, nous avons généralisé la méthode de réplication proposée par Papageorgiou *et al.* (2008) en y intégrant l'algorithme de couverture optimale élaboré par Rémillard & Rubenthaler (2009). On réplique ainsi les rendements d'un fonds de couverture conditionnellement à plusieurs portefeuilles de référence. Cette réplication se traduit par la formation d'un portefeuille en investissant dans les portefeuilles de référence et un portefeuille de réserve de telle sorte à avoir en fin de mois un rendement ayant la même distribution que la distribution conditionnelle du fonds de couverture à répliquer.

Les rendements ont été modélisés par des mélanges Gaussiens multivariés et la dépendance entre les portefeuilles a été prise en compte en utilisant les copules. Ainsi, après avoir estimé la loi jointe des portefeuilles, nous avons choisi la distribution marginale du fonds à répliquer et la copule qui modélise la dépendance de celui-ci avec les portefeuilles de référence. Par la suite, l'outil de la réplication, la fonction g généralisant celle établie par Kat & Palaro (2005) a permis de générer des flux ayant les mêmes caractéristiques statistiques que ceux du fonds en fin de mois. Enfin, l'algorithme établi par Rémillard & Rubenthaler (2009) a permis d'effectuer la réplication optimale afin de garantir les meilleurs résultats.

Les résultats empiriques de la réplication d'un fonds synthétique présentés au dernier chapitre ont permis de valider le bon fonctionnement de notre méthode

de répliation. Il serait à présent possible de répliquer les rendements de n'importe quel fonds de couverture existant. De plus, il serait possible d'examiner la sensibilité des résultats au choix du portefeuille de réserve.

Bibliographie

- Agarwal, V., & Naik, N. 2005. Risks and portfolio decisions involving hedge funds. *Review of Financial Studies*, **17**, 63–98.
- Amin, G., & Kat, H. 2003. Hedge fund performance 1990-2000 : Do the money machines really add value? *Journal of Financial and Quantitative Analysis*, **38**(2), 251–275.
- Barbe, Philippe, Genest, Christian, Ghoudi, Kilani, & Rémillard, Bruno. 1996. On Kendall's process. *Journal of Multivariate Analysis*, **58**(2), 197–229.
- Chen, Y., & Liang, B. 2006. Do market timing hedge funds time the market? *Technical report, Carroll School of Management, Boston College*.
- Cox, J., & Leland, H. 2000. On dynamic investment strategies. *Journal of Economic Dynamics and Control*, **24**, 1859–1880.
- Del Moral, P., Rémillard, B., & Rubenthaler, S. 2006. Monte carlo approximations of american options. *Technical report, GERAD*.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. roy. statist. soc. ser. b*, **39**, 1–38.
- Dybvig, P. 1988. Distributional analysis of portfolio choice. *The Journal of Business*, **61**(3), 369–393.
- Genest, C., Rémillard, B., & Beaudoin, D. 2009. Omnibus goodness-of-fit tests for copulas : A review and a power study. *Insurance : Mathematics and Economics*, **44**, 199–213.

- Glosten, L., & Jagannathan, R. 1994. A contingent claim approach to performance evaluation. *Journal of Empirical Finance*, **1**, 133–160.
- Hasanhodzic, J., & Lo, A. W. 2007. Can hedge-fund returns be replicated? : The linear case. *Journal of Investment Management*, **5**(2), 5–45.
- Kat, H., & Palaro, H. 2005. Who needs hedge funds? A copula-based approach to hedge fund return replication. *Technical report, Cass Business School, City University*.
- Mitchell, M., & Pulvino, T. 2001. Characteristics of risk and return in risk arbitrage. *Journal of Finance*, **56**, 2135–2175.
- Papageorgiou, N., Rémillard, B., & Hocquard, A. 2008. Replicating the properties of hedge funds returns. *Journal of Alternative Investments*, **11**, 8–38.
- Rémillard, B., & Rubenthaler, S. 2009. Optimal hedging in discrete time and its continuous limit. *Papier de recherche*.
- Rosenblatt, M. 1952. Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, **23**, 470–472.
- Schweizer, M. 1995. Variance-optimal hedging in discrete time. *Math. Oper. Res.*, **20**(1), 1–32.