

Master 2 MMMEF

Rapport de stage

CONTRIBUTION À L'AMÉLIORATION DE L'ENVIRONNEMENT
D'OPTIMISATION DE PORTEFEUILLE, ET IMPLÉMENTATION DE
STRATÉGIES QUANTITATIVES

Étudiant : Bergeret Hugo

Entreprise d'accueil : Candriam

Tuteur d'entreprise : Clapt Olivier

Tuteur de l'Université : Bich Philippe

Table des matières

1	Introduction	2
1.1	Présentation de l'entreprise et de l'équipe	2
1.2	Présentation des missions	2
2	Rappel : Programmation quadratique et portefeuille	
	Moyenne-Variance de Markowitz	3
2.1	Programmation quadratique	3
2.2	Portefeuille de Moyenne-Variance de Markowitz	5
3	Réplication de l'environnement initial en Python	7
3.1	Optimisation stratégique	7
3.2	Optimisation tactique	12
3.3	Optimisation avec resampling	15
3.4	Optimisation avec intégration de convictions	21
4	Amélioration de la robustesse de l'environnement	26
4.1	Structure du projet	26
4.2	Création d'un système de test efficace	27
4.3	Introduction du logging	27
5	Enrichissement des fonctionnalités de l'environnement d'optimisation	28
5.1	Entropy pooling	28
5.2	Optimisation avec coûts de transaction	38
6	Conclusion	44
A	Impacts de l'entropy pooling sur les probabilités	47

1 Introduction

1.1 Présentation de l'entreprise et de l'équipe

CANDRIAM (pour Conviction AND Responsibility In Asset Management) est un gestionnaire d'actifs européen. Pionnier et leader des investissements durables (ESG) depuis 1996, CANDRIAM gère environ 150 milliards d'euros d'actifs avec une équipe de plus de 600 salariés. Il dispose de bureaux à Luxembourg, Bruxelles, Paris et Londres, et de représentants de clients dans plus de 20 pays en Europe continentale, au Royaume-Uni, aux États-Unis et au Moyen-Orient. CANDRIAM propose des solutions d'investissement dans plusieurs domaines clés : obligations, actions, stratégies de performance absolue et allocation d'actifs, avec une gamme étendue et innovante de stratégies ESG couvrant toutes les classes d'actifs.

Mon stage a eu lieu au Front-Office de Candriam au sein de l'équipe de recherche quantitative en charge du développement de stratégies et d'outils quantitatifs pour la modélisation de portefeuille, pour les équipes de gestion Multi-Asset de Candriam (CTA, Global Macro, Asset Allocation).

Le stage a eu lieu en présentiel dans les bureaux parisiens de Candriam, situés dans le 8ème arrondissement de Paris, sous la tutelle d'Olivier Clapt : responsable de l'équipe de recherche quantitative. Il s'est étendu de début Avril à fin Septembre.

1.2 Présentation des missions

L'objectif de ce stage était double. Premièrement, il s'agissait de répliquer des outils d'optimisation de portefeuille préexistants au sein de l'entreprise (initialement codés en C/C++) dans un environnement en Python. Dans un second temps, il s'agissait de proposer des améliorations à cet environnement afin de le compléter. Dans un cadre général, l'optimisation de portefeuille a pour but de déterminer les poids à allouer aux différents actifs de l'univers d'investissement afin de maximiser le rendement du portefeuille ainsi obtenu, tout en minimisant une métrique de risque déterminée.

L'environnement d'optimisation de portefeuille initial comportait les éléments suivants :

- Un module d'optimisation stratégique
- Un module d'optimisation tactique
- Un module d'optimisation avec resampling
- Un module d'optimisation avec l'intégration de convictions qualitatives

En outre, il comportait également un module de pricing de dérivés simples (call/put européens et options barrières européennes et américaines).

Ainsi, la première partie du stage a été dédiée à la reproduction de ces modules, tandis que la seconde a eu pour objectif d'aller plus loin en proposant de nouvelles fonctionnalités.

2 Rappel : Programmation quadratique et portefeuille Moyenne-Variance de Markowitz

2.1 Programmation quadratique

La contribution de Harry Markowitz à l'économie est considérable. En effet, son optimisation moyenne-variance marque le début de l'allocation de portefeuille en finance. Ce concept d'optimisation est la pièce maîtresse de la théorie moderne du portefeuille, pour laquelle Markowitz a d'ailleurs reçu le prix Nobel en 1990. En plus de l'article fondateur de 1952, Harry Markowitz a proposé un algorithme pour résoudre les problèmes de programmation quadratique dès 1956. À cette époque, très peu de personnes connaissaient ce type d'optimisation. Ainsi Harry Markowitz est considéré comme le "père de la programmation quadratique". Historiquement, ce n'était pas la première fois que des économistes participaient au développement des mathématiques, mais c'est certainement la première fois que des mathématiciens ont exploré un domaine de recherche dont l'application principale pendant les premières années de recherche était exclusivement un problème économique. Le succès de l'optimisation moyenne-variance (OMV) est dû aux propriétés attrayantes de la fonction d'utilité quadratique, mais il doit également être évalué à la lumière du succès de la programmation quadratique (PQ). Étant donné qu'il est facile de résoudre des problèmes de PQ et que les problèmes de PQ sont disponibles dans les logiciels mathématiques, la résolution des problèmes de OMV ne nécessite que peu de compétences spécifiques. C'est pourquoi l'optimisation moyenne-variance est une méthode universelle qui est utilisée par tous les gestionnaires de portefeuille. Cependant, cette approche a été largement critiquée par les universitaires et les professionnels. En effet, l'optimisation moyenne-variance est très sensible aux paramètres d'entrée et produit des solutions en coin. Cette problématique rend la pratique de l'optimisation moyenne-variance moins attractive que la théorie (Michaud, 1989) [1]. En effet, la résolution des problèmes d'allocation OMV nécessite de spécifier les bonnes contraintes de poids afin d'obtenir des solutions acceptables. Il s'ensuit que la conception des contraintes est la composante la plus importante de l'optimisation. Dans ce cas, l'OMV semble être un processus itératif fait de corrections, et non une solution systématique. Le succès de l'OMV en finance s'explique également par le fait qu'il existe très peu de modèles d'allocation de portefeuille concurrents pouvant être mis en œuvre de manière industrielle. Il y a généralement deux raisons à cela. La première est que certains

modèles utilisent des paramètres d'entrée difficiles à estimer ou même à comprendre, ce qui rend ces modèles évidemment inutilisables. La seconde raison est que d'autres modèles utilisent une fonction d'évaluation plus complexe que la simple fonction d'utilité quadratique. Dans ce cas, la complexité de calcul rend ces modèles moins attrayants que le modèle OMV standard. Parmi ces modèles, certains sont basés sur la fonction d'évaluation moyenne-variance, mais introduisent des fonctions de pénalité de régularisation afin d'améliorer la robustesse de l'allocation de portefeuille (régularisation L1 et L2, ou encore Elastic Net). Là encore, ces modèles ont peu de chance d'être utilisés s'ils ne peuvent être convertis en un problème PQ.

Un problème de programmation quadratique (PQ) est un problème d'optimisation avec une fonction d'évaluation quadratique et des contraintes d'inégalité linéaires prenant la forme suivante :

$$\begin{aligned} x^* &= \arg \min_x \frac{1}{2} x^\top Q x - x^\top R \\ \text{t.q.} \quad & Sx \leq T \end{aligned} \tag{1}$$

où x est un vecteur $n \times 1$, Q est une matrice $n \times n$ et R est un vecteur $n \times 1$. On peut observer que le système de contraintes $Sx \leq T$ permet de spécifier des contraintes linéaires d'égalité $Ax = B$ ou des contraintes de bornes $x^- \leq x \leq x^+$. Dans la plupart des logiciels d'optimisation numérique, c'est la version suivante du problème qui est considérée :

$$\begin{aligned} x^* &= \arg \min_x \frac{1}{2} x^\top Q x - x^\top R \\ \text{t.q.} \quad & \begin{cases} Ax = B \\ Cx \leq D \\ x^- \leq x \leq x^+ \end{cases} \end{aligned} \tag{2}$$

car le problème (2) est équivalent au problème canonique (1) avec le système suivant d'inégalités linéaires :

$$\begin{bmatrix} -A \\ A \\ C \\ -I_n \\ I_n \end{bmatrix} x \leq \begin{bmatrix} -B \\ B \\ D \\ -x^- \\ x^+ \end{bmatrix}$$

Si l'espace Ω défini par $Sx \leq T$ est non vide et si Q est une matrice semi-définie positive symétrique, la solution existe car la fonction $f(x) = \frac{1}{2} x^\top Q x - x^\top R$ est convexe. Dans le cas général où Q est, seulement, une matrice carrée, la solution peut ne pas exister.

Il existe de plus une formulation duale du problème précédant. Le Lagrangien de problème étant :

$$\mathcal{L}(x; \lambda) = \frac{1}{2}x^\top Qx - x^\top R + \lambda^\top (Sx - T)$$

on peut déduire que le problème dual est de la forme :

$$\lambda^* = \arg \max_{\lambda} \left\{ \inf_x \mathcal{L}(x; \lambda) \right\} \text{ t.q. } \lambda \geq 0$$

En remarquant que $\partial_x \mathcal{L}(x; \lambda) = Qx - R + S^\top \lambda$. La solution à l'équation $\partial_x \mathcal{L}(x; \lambda) = 0$ est donnée par $x = Q^{-1} (R - S^\top \lambda)$. Nous obtenons finalement :

$$\begin{aligned} \inf_x \mathcal{L}(x; \lambda) &= \frac{1}{2} (R^\top - \lambda^\top S) Q^{-1} (R - S^\top \lambda) - (R^\top - \lambda^\top S) Q^{-1} R + \\ &\quad \lambda^\top (SQ^{-1} (R - S^\top \lambda) - T) \\ &= \frac{1}{2} R^\top Q^{-1} R - \lambda^\top S Q^{-1} R + \frac{1}{2} \lambda^\top S Q^{-1} S^\top \lambda - R^\top Q^{-1} R + \\ &\quad 2\lambda^\top S Q^{-1} R - \lambda^\top S Q^{-1} S^\top \lambda - \lambda^\top T \\ &= -\frac{1}{2} \lambda^\top S Q^{-1} S^\top \lambda + \lambda^\top (S Q^{-1} R - T) - \frac{1}{2} R^\top Q^{-1} R \end{aligned}$$

Ainsi le problème dual est un second problème de programmation quadratique :

$$\begin{aligned} \lambda^* = \arg \min_{\lambda} \frac{1}{2} \lambda^\top \bar{Q} \lambda - \lambda^\top \bar{R} \\ \text{t.q. } \lambda \geq 0 \end{aligned} \tag{3}$$

avec $\bar{Q} = S Q^{-1} S^\top$ et $\bar{R} = S Q^{-1} R - T$.

2.2 Portefeuille de Moyenne-Variance de Markowitz

Le concept d'allocation de portefeuille a une longue histoire et remonte à l'ouvrage fondateur de Markowitz (1952) [2]. Dans son article, Markowitz définit précisément le concept de sélection de portefeuille par Moyenne-Variance : "l'investisseur considère (ou devrait considérer) le rendement attendu comme une chose souhaitable et la variance du rendement comme une chose indésirable". En effet, Markowitz a montré qu'un portefeuille efficient est un portefeuille qui maximise le rendement attendu pour un niveau de risque donné (correspondant à la variance du rendement du portefeuille) ou un portefeuille qui minimise le risque pour un niveau de rendement attendu donné. Même si ce cadre a été étendu à d'autres problèmes d'allocation (échantillonnage d'indices, gestion du turnover, etc.), le modèle moyenne-variance reste

l'approche d'optimisation la plus utilisée en finance. Dans la suite de cette section, nous allons brièvement développer le framework de Markowitz afin de faire le lien avec ce qui a été détaillé dans la section précédente. Ensuite, nous verrons comment, en pratique, le stage a consisté à implémenter ces concepts.

Soit un univers de n actifs. Soit $x = (x_1, \dots, x_n)$ le vecteur des poids du portefeuille. Nous supposons que le portefeuille est entièrement investi, ce qui signifie que $\sum_{i=1}^n x_i = \mathbf{1}_n^\top x = 1$. Nous désignons par $R = (R_1, \dots, R_n)$ le vecteur des rendements des actifs où R_i est le rendement de l'actif i . Le rendement du portefeuille est alors égal à $R(x) = \sum_{i=1}^n x_i R_i = x^\top R$. Soit $\mu = E[R]$ et $\Sigma = E[(R - \mu)(R - \mu)^\top]$ le vecteur des rendements espérés et la matrice de covariance des rendements des actifs. Le rendement attendu du portefeuille est égal à : $\mu(x) = E[R(x)] = x^\top \mu$ tandis que sa variance est égale à :

$$\sigma^2(x) = E[(R(x) - \mu(x))(R(x) - \mu(x))^\top] = x^\top \Sigma x$$

Alors Markowitz a formulé le problème financier de l'investisseur comme suit :

1. Maximiser le rendement attendu du portefeuille sous une contrainte de volatilité (σ -problème) :

$$\max \mu(x) \quad \text{t.q.} \quad \sigma(x) \leq \sigma^* \quad (4)$$

2. Soit minimiser la volatilité du portefeuille sous une contrainte de rendement (μ -problème) :

$$\min \sigma(x) \quad \text{t.q.} \quad \mu(x) \geq \mu^* \quad (5)$$

La brillante idée de Markowitz fut de considérer une fonction d'utilité quadratique :

$$\mathcal{U}(x) = x^\top \mu - \frac{\phi}{2} x^\top \Sigma x$$

où ϕ est l'aversion au risque. Puisque la maximisation de $\mathcal{U}(x)$ est équivalente à la minimisation de $-\mathcal{U}(x)$, les problèmes de Markowitz (4) et (5) peuvent être transformés en un problème de PQ :

$$\begin{aligned} x^*(\gamma) = & \arg \min_x \frac{1}{2} x^\top \Sigma x - \gamma x^\top \mu \\ & \text{t.q.} \quad \mathbf{1}_n^\top x = 1 \end{aligned} \quad (6)$$

où $\gamma = \phi^{-1}$. Par conséquent, résoudre le μ -problème ou le σ -problème revient à trouver la valeur optimale de γ telle que $\mu(x^*(\gamma)) = \mu^*$ ou $\sigma(x^*(\gamma)) = \sigma^*$. On sait que les fonctions $\mu(x^*(\gamma))$ et $\sigma(x^*(\gamma))$ sont croissantes par rapport à γ et sont bornées. La valeur optimale de γ peut

alors être facilement calculée en utilisant la méthode de bisection. Il est évident qu'une grande partie du succès du framework de Markowitz repose sur l'astuce de la PQ. En effet, le problème (6) correspond au problème PQ (2) où $Q = \Sigma$, $R = \gamma\mu$, $A = 1_n^\top$ et $B = 1$. De plus, il est facile d'inclure des contraintes supplémentaires telles que des bornes concernant les poids des actifs, des inégalités entre les classes d'actifs, etc.

3 Réplication de l'environnement initial en Python

Dans le cadre des activités de Candriam, des fonctions préexistantes étaient codées en C/C++ et avaient pour but de calculer les portefeuilles optimaux selon le critère de Moyenne-Variance détaillé précédemment. D'un point de vue d'informatique, le changement d'un environnement C/C++ à un environnement en Python est motivé par le fait que la migration des ordinateurs vers un système 64 bits rend les bibliothèques en C/C++ inutilisables en l'état, notamment pour des problèmes d'interopérabilité des types de variables. Le choix de la migration en Python passe également par le fait de mettre les différentes fonctions développées à disposition des utilisateurs en passant par un système client/serveur, notamment utilisable au travers d'Excel. En effet, l'utilisateur n'a qu'à installer un complément Excel, qui permet d'utiliser des fonctions (VBA) qui appellent les fonctions développées en Python qui sont sur un serveur en réseau qui traite les requêtes de l'utilisateur et renvoie les résultats dans Excel. Malgré sa complexité, cette méthode a pour avantage de simplifier l'utilisation des fonctions :

1. Premièrement, les utilisateurs des fonctions (les gestionnaires de fonds par exemple) ne sont pas toujours à l'aise avec Python et ainsi, l'utilisation au travers d'Excel assure de ne pas perturber leurs habitudes de travail.
2. Deuxièmement, cela évite que chaque utilisateur ait à installer Python sur sa machine personnelle, et par la même cela évite les problèmes de compatibilité du code en fonction des versions de Python et des versions des packages.

3.1 Optimisation stratégique

Parmi les fonctions contenues dans l'ancienne bibliothèque C/C++, nous allons nous concentrer sur celles utilisées dans le cadre de l'optimisation stratégique. On entend par optimisation stratégique, le cas des portefeuilles de Moyenne-Variance classiques où le terme quadratique n'est autre que la variance du portefeuille (à l'inverse du cas d'optimisation tactique où le terme quadratique correspond à la tracking error par rapport à un benchmark : indice, fonds etc.).

Dans ce cadre, le but de l'optimisation est de déterminer soit :

1. un portefeuille optimal correspondant à un niveau de rendement donné (c'est-à-dire, le portefeuille avec la variance minimale étant donné un niveau de rendement),
2. un portefeuille optimal correspondant à un niveau de variance/volatilité donné (c'est-à-dire, le portefeuille avec le rendement maximal étant donné un niveau de variance/volatilité),
3. ou la frontière efficiente c'est à dire l'ensemble des portefeuilles optimaux pour des niveaux de rendement donnés (frontière étant découpée entre le rendement du portefeuille de variance/volatilité minimale et le rendement du portefeuille de rendement maximal).

3.1.1 Description du problème

Le portefeuille optimal pour un niveau de rendement donné s'exprime mathématiquement comme suit :

$$\begin{aligned}
 x^* &= \arg \min_x \frac{1}{2} x^\top \Sigma x \\
 \text{t.q. } &Ax \leq B
 \end{aligned} \tag{7}$$

où pour un objectif de rendement noté r_{tgt} on a :

$$\begin{bmatrix} \mathbf{1}_n^\top \\ -\mathbf{1}_n^\top \\ R^\top \\ -R^\top \\ I_n \\ -I_n \end{bmatrix} = A \quad \text{et} \quad B = \begin{bmatrix} 1 \\ -1 \\ r_{tgt} \\ -r_{tgt} \\ x^+ \\ -x^- \end{bmatrix}$$

ce qui assure que le portefeuille est totalement investi, que le rendement cible est atteint, et que les bornes des poids sont respectées. On peut enrichir ce problème en ajoutant des contraintes supplémentaires comme des contraintes de diversification du portefeuille. Par exemple si le portefeuille se compose de 3 actifs risqués, il est possible de spécifier dans le problème que l'on cherche une solution qui respecte la contrainte supplémentaire que la somme des poids des deux premiers actifs soit inférieure à 30%. Ainsi on pourra compléter la matrice A et le vecteur B avec cette contrainte supplémentaire. Dans le cadre d'une contrainte d'égalité, on ajoutera deux contraintes d'inégalité (comme réalisé ci-dessus).

Si l'on se situe dans le cadre d'un objectif de variance/volatilité, nous cherchons alors à résoudre le problème à l'aide de la méthode de la bisection comme évoqué à la fin de la section 2.2.

Finalement, pour résoudre le problème de la frontière efficiente, il suffit d'obtenir une discrétisation des rendements entre celui du portefeuille de variance/volatilité minimale et celui du portefeuille de rendement maximal. Pour obtenir le premier, on résout le problème de l'équation (7) sans cible de rendement, et pour le second on résout le problème de l'équation (1) avec la matrice de variance-covariance nulle (car le minimum de l'opposé d'une fonction est son maximum). Une fois la discrétisation obtenue, il suffit de trouver les portefeuilles optimaux pour chaque cible de rendement et l'on obtient ainsi la frontière efficiente.

3.1.2 Objectif et utilité de l'approche

L'idée des portefeuilles optimaux en termes de Moyenne-Variance est pertinente pour Candriam car les fonds sont définis par des prospectus qui les engagent, en définissant des profils de risque qui correspondent en réalité à des objectifs de rendement ou de volatilité. Ainsi l'approche de portefeuille de Moyenne-Variance paraît toute naturelle dans ce contexte. C'est pourquoi utiliser cette approche lors de la construction des portefeuilles paraît encore une fois naturelle. La flexibilité de cette approche permet en outre à chaque utilisateur des fonctions de spécifier les contraintes qui s'imposent à lui en fonction de la typologie du fonds.

En outre les matrices de variance-covariance étant de fait semi-définies positives et symétriques, nous pouvons utiliser les méthodes présentées précédemment. Toutefois, l'estimation numérique de ces quantités peut conduire à une instabilité pouvant résulter en des cas de matrice non semi-définie positive, chose qui est simplement corrigable numériquement en tiltant légèrement la matrice pour la rendre à nouveau semi-définie positive.

3.1.3 Méthodologie adoptée

Afin de résoudre les trois problèmes évoqués ci-dessus, nous avons fait le choix d'utiliser le solveur de programmation quadratique quadprog, initialement codé en C mais qui est doté d'un wrapper en Python qui permet de l'utiliser dans cet environnement. Le choix de ce solveur a été guidé par le fait qu'il soit gratuit, libre de droit et plus performant que les solveurs disponibles ailleurs (notamment celui du package Scipy). En effet, les performances de quadprog sont très bonnes comme l'atteste le tableau suivant qui compare différents solveurs de programmation quadratique disponibles en Python :

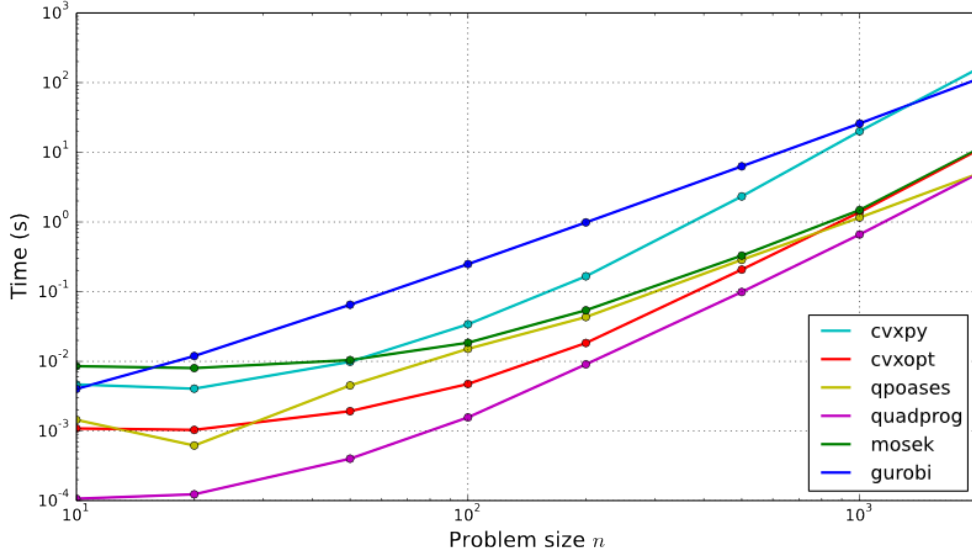


FIGURE 1 – Performance comparative des solveurs sur des problèmes denses (d’après la documentation de qpsolvers)

Le solveur quadprog s’appuie sur l’algorithme de D. Goldfarb et A. Idnani (1983) [3] qui utilise une méthode duale d’ensemble actif. L’idée de l’algorithme est de résoudre les problèmes de la forme suivante :

$$\min_x f(x) = a^\top x + \frac{1}{2}x^\top Gx, \quad \text{t.q.} \quad s(x) = C^\top x - b \geq 0 \quad (8)$$

où x et a sont des vecteurs $n \times 1$, G est une matrice semi-définie positive symétrique $n \times n$, C est une matrice $n \times m$, b est un vecteur $m \times 1$. Le vecteur de variables x peut également être soumis à des contraintes d’égalité. Par ensemble actif, on désigne un sous-ensemble des m contraintes dans (8) qui sont satisfaites en tant qu’égalités par l’estimation actuelle x de la solution du problème de PQ. Nous utiliserons K pour désigner l’ensemble $\{1, 2, \dots, m\}$ des indices des contraintes et $A \subseteq K$ pour désigner les indices de l’ensemble actif. Nous définissons un sous-problème $P(J)$ comme étant le problème de PQ avec la fonction d’évaluation (8) soumis uniquement au sous-ensemble des contraintes indexé par $J \subset K$. Par exemple, $P(\emptyset)$, où \emptyset désigne l’ensemble vide, est le problème de la recherche du minimum sans contrainte. Si la solution x d’un sous-problème $P(J)$ se trouve dans un ensemble actif de contraintes linéairement indépendantes indexées par $A \subseteq J$, nous appelons (x, A) une S-paire (S). Il est clair que si (x, A) est une S-paire pour le sous-problème $P(J)$, elle est également une S-paire pour le sous-problème $P(J)$. Par indépendance linéaire d’un ensemble de contraintes, nous entendons que les normales correspondant à ces contraintes sont linéairement indépendantes. Nous désignerons le vecteur normal de la i -ème contrainte, c’est-à-dire la i -ème colonne de C , par n_i . Nous pouvons

maintenant exposer l'approche des deux auteurs pour résoudre le problème de PQ :

1. Supposons qu'une S -paire (x, A) soit donnée.
2. Répéter jusqu'à ce que toutes les contraintes soient satisfaites :
 - (a) Choisir une contrainte violée $p \in K \setminus A$.
 - (b) Si $P(A \cup \{p\})$ est infaisable, on arrête car le problème est infaisable.
 - (c) Sinon, obtenir une nouvelle S -paire $(\bar{x}, \bar{A} \cup \{p\})$ où, $\bar{A} \subseteq A$ et $f(\bar{x}) > f(x)$ et définir $(x, A) \leftarrow (\bar{x}, \bar{A} \cup \{p\})$.
3. On arrête car x est la solution optimale du problème de PQ. Puisque le minimum non contraint, $x^0 = -G^{-1}a$ est facilement obtenu, on peut toujours commencer la procédure ci-dessus avec la paire (x^0, \emptyset) .

Ainsi les nouvelles fonctions en Python que j'ai implémenté utilisent cet algorithme, au travers du package quadprog. L'utilisation se fait soit directement au travers de Python (l'ensemble des fonctions développées sont réunies au sein d'un package), soit par l'intermédiaire d'Excel. Ci-dessous, nous présentons les résultats de l'optimisation stratégique avec 8 indices. Les paramètres de rendements attendus, de volatilité et de corrélation sont les suivants :

Rendements et vol annuels attendus:		
	Rendements (%)	Volatilités (%)
JPM UK Traded	-1.90	8.77
JPM EMU Aggregate in EUR	0.51	3.57
JPM US Traded	2.94	9.20
JPM Japan Traded	1.93	12.95
MSCI UK Net Div	6.08	14.14
MSCI Europe ex. UK in EURO	7.28	16.99
MSCI US Net Div	8.56	18.15
MSCI Japan Net Div	7.26	22.70

FIGURE 2 – Rendements et volatilités annuels attendus

Matrice de corrélation:								
	JPM UK Traded	JPM EMU Aggregate in EUR	JPM US Traded	JPM Japan Traded	MSCI UK Net Div	MSCI Europe ex. UK in EURO	MSCI US Net Div	MSCI Japan Net Div
JPM UK Traded	1.00	0.31	0.67	0.08	0.50	0.23	0.44	0.19
JPM EMU Aggregate in EUR	0.31	1.00	0.25	-0.01	0.16	0.06	0.17	0.03
JPM US Traded	0.67	0.25	1.00	0.12	0.48	0.38	0.60	0.21
JPM Japan Traded	0.08	-0.01	0.12	1.00	0.31	0.12	0.25	0.49
MSCI UK Net Div	0.50	0.16	0.48	0.31	1.00	0.73	0.78	0.54
MSCI Europe ex. UK in EURO	0.23	0.06	0.38	0.12	0.73	1.00	0.74	0.50
MSCI US Net Div	0.44	0.17	0.60	0.25	0.78	0.74	1.00	0.55
MSCI Japan Net Div	0.19	0.03	0.21	0.49	0.54	0.50	0.55	1.00

FIGURE 3 – Matrice de corrélation des indices

Les figures ci-dessous présentent les résultats obtenus sans et avec contraintes :

Output: sans contraintes:

	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84
Rendement cible:	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84
Volatilité du portefeuille	8.68	9.68	10.70	11.73	12.77	13.81	14.87	15.97
Rendement du portefeuille	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84
Poids	JPM UK Traded	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	JPM EMU Aggregate in EUR	34.84%	28.07%	21.29%	14.50%	7.45%	0.21%	0.00%
	JPM US Traded	13.75%	14.42%	15.10%	15.79%	16.39%	16.97%	8.50%
	JPM Japan Traded	2.48%	1.75%	1.01%	0.28%	0.00%	0.00%	0.00%
	MSCI UK Net Div	7.93%	9.21%	10.47%	11.72%	12.85%	13.90%	13.11%
	MSCI Europe ex. UK in EURO	16.95%	18.25%	19.55%	20.86%	22.27%	23.74%	24.75%
	MSCI US Net Div	19.12%	22.74%	26.36%	29.97%	33.66%	37.39%	45.94%
	MSCI Japan Net Div	4.93%	5.57%	6.22%	6.88%	7.38%	7.80%	7.70%

FIGURE 4 – Résultats de l'optimisation stratégique sans contraintes

Output avec contraintes:

	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84
Rendement cible:	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84
Volatilité du portefeuille	8.86	9.84	10.93	12.11	13.41	15.02	15.02	15.02
Rendement du portefeuille	4.34	4.84	5.34	5.84	6.34	6.73	6.73	6.73
Poids	JPM UK Traded	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	JPM EMU Aggregate in EUR	16.00%	15.45%	10.00%	10.00%	10.00%	10.00%	10.00%
	JPM US Traded	25.00%	23.46%	21.73%	11.81%	0.00%	0.00%	0.00%
	JPM Japan Traded	13.34%	4.46%	2.33%	0.00%	0.00%	0.00%	0.00%
	MSCI UK Net Div	13.90%	19.77%	24.96%	31.33%	32.64%	0.00%	0.00%
	MSCI Europe ex. UK in EURO	21.77%	26.87%	30.98%	36.86%	47.36%	80.00%	80.00%
	MSCI US Net Div	8.33%	7.69%	8.06%	9.73%	10.00%	10.00%	10.00%
	MSCI Japan Net Div	1.67%	2.31%	1.94%	0.27%	0.00%	0.00%	0.00%

FIGURE 5 – Résultats de l'optimisation stratégique avec contraintes

Les contraintes passées sont les suivantes :

- La somme des deux premiers poids doit être entre 10 et 16%,
- Le troisième poids doit être inférieur à 25%,
- La somme des deux derniers poids doit être entre 2 et 10%

On observe que l'ajout de contraintes aboutit à des portefeuilles moins performants, c'est-à-dire que pour un même rendement cible, la volatilité du portefeuille obtenu est plus grande avec les contraintes que sans. Nous voyons également que dans le cas avec contraintes, les trois derniers portefeuilles sont égaux. Cela s'explique par le fait que le rendement cible n'est pas atteignable et nous renvoyons donc le portefeuille de rendement maximal atteignable.

3.2 Optimisation tactique

On appelle optimisation tactique, le fait d'optimiser les poids des actifs par rapport à un benchmark. Dans ce cas, le terme quadratique $x^\top \Sigma x$ est remplacé par $(x - b)^\top \Sigma (x - b)$ ce qui revient à utiliser la tracking error comme fonction d'évaluation du risque au lieu de la variance/volatilité. Concernant le terme linéaire, ce n'est plus le rendement que l'on cherche à maximiser mais plutôt le rendement en excès $(x - b)^\top R$ au lieu de simplement $x^\top R$.

En définitive cela ne change rien au problème à l'exception des contraintes qui doivent être adaptées si elles sont définies de manière absolue ou bien si elles sont exprimées en termes de déviation par rapport au benchmark. Finalement, c'est le problème suivant que l'on cherche ici à résoudre :

$$\begin{aligned} x^* &= \arg \min_x \frac{1}{2}(x - b)^\top \Sigma (x - b) \\ \text{t.q. } &A(x - b) \leq B \end{aligned} \tag{9}$$

Encore une fois, trois fonctions ont été codées pour déterminer soit :

1. un portefeuille optimal correspondant à un niveau de rendement en excès donné (c'est-à-dire, le portefeuille avec la tracking error minimale étant donné un niveau de rendement en excès),
2. un portefeuille optimal correspondant à un niveau de tracking error donné (c'est-à-dire, le portefeuille avec le rendement en excès maximal étant donné un niveau de tracking error),
3. ou la frontière efficiente c'est à dire l'ensemble des portefeuilles optimaux pour des niveaux de rendement en excès donnés (frontière étant découpée entre le rendement en excès du portefeuille de tracking error minimale et le rendement en excès du portefeuille de rendement en excès maximal).

3.2.1 Objectif et utilité de l'approche

Le but de cette approche est de définir les meilleurs portefeuilles en fonction d'un benchmark, ce qui correspond aux besoins des fonds dont les prospectus spécifient des objectifs de rendement en excès et de tracking error. L'idée de tels fonds est de préciser aux investisseurs à quel point l'investissement prendra des libertés par rapport au benchmark. Intuitivement, pour un objectif de tracking error nul, le portefeuille devrait être une réplique parfaite du benchmark. Plus la tracking error augmente et plus le fonds aura de liberté dans le choix de ses allocations. Ainsi ces engagements étant contractuels, il apparaît encore une fois logique d'adopter cette approche pour lesdits fonds.

3.2.2 Méthodologie adoptée

La méthodologie adoptée pour cette série de fonction est exactement la même que celle présentée dans le cadre de l'optimisation stratégique. On utilise également le solveur quadprog pour les mêmes raisons.

L'unique difficulté de cette approche est la spécification des contraintes. En effet, le problème précédant peut se ramener au problème suivant :

$$x^* = \arg \min_y \frac{1}{2} y^\top \Sigma y \quad (10)$$

t.q. $A'y \leq B'$

en posant $y = (x - b)$, il faut donc adapter les contraintes.

Ainsi pour la résolution du problème avec cible de tracking error, nous nous ramenons au problème de cible de variance/volatilité au travers du changement de variable. On peut alors appliquer la méthode de bisection comme pour l'optimisation stratégique.

Les résultats de l'optimisation avec et sans contraintes sont présentés ci-après. Les paramètres utilisés sont ceux de la figure 2 et 3 et les contraintes celles de la section précédente. Le benchmark choisi est le suivant :

	Poids
JPM UK Traded	2.50%
JPM EMU Aggregate in EUR	32.50%
JPM US Traded	12.50%
JPM Japan Traded	2.50%
MSCI UK Net Div	7.41%
MSCI Europe ex. UK in EURO	17.59%
MSCI US Net Div	20.00%
MSCI Japan Net Div	5.00%

FIGURE 6 – Benchmark retenu

Avec ces paramètres nous obtenons les résultats suivants :

Output: sans contraintes:								
Rendement cible:	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84
Tracking error du portefeuille	0.00	0.87	1.94	3.01	4.09	5.17	6.28	7.46
Rendement du portefeuille	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84
Poids	JPM UK Traded	2.50%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	JPM EMU Aggregate in EUR	32.50%	28.07%	21.30%	14.53%	7.44%	0.19%	0.00%
	JPM US Traded	12.50%	14.42%	15.09%	15.76%	16.39%	16.99%	8.50%
	JPM Japan Traded	2.50%	1.75%	1.01%	0.28%	0.00%	0.00%	0.00%
	MSCI UK Net Div	7.41%	9.21%	10.46%	11.72%	12.85%	13.92%	13.11%
	MSCI Europe ex. UK in EURO	17.59%	18.25%	19.55%	20.86%	22.27%	23.75%	24.75%
	MSCI US Net Div	20.00%	22.74%	26.37%	29.99%	33.66%	37.35%	45.94%
	MSCI Japan Net Div	5.00%	5.57%	6.22%	6.87%	7.38%	7.81%	7.70%

FIGURE 7 – Résultats de l'optimisation tactique sans contraintes

Output avec contraintes:									
Rendement cible:	4.34	4.84	5.34	5.84	6.34	6.84	7.34	7.84	
Tracking error du portefeuille	1.79	1.97	2.95	4.26	5.79	8.08	8.08	8.08	8.08
Rendement du portefeuille	4.34	4.84	5.34	5.84	6.34	6.73	6.73	6.73	6.73
Poids	JPM UK Traded	2.27%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	JPM EMU Aggregate in EUR	13.73%	15.45%	10.00%	10.00%	10.00%	10.00%	10.00%	10.00%
	JPM US Traded	25.00%	23.46%	21.73%	11.81%	0.00%	0.00%	0.00%	0.00%
	JPM Japan Traded	12.28%	4.45%	2.33%	0.00%	0.00%	0.00%	0.00%	0.00%
	MSCI UK Net Div	13.85%	19.77%	24.96%	31.33%	32.64%	0.00%	0.00%	0.00%
	MSCI Europe ex. UK in EURO	22.87%	26.87%	30.98%	36.86%	47.36%	80.00%	80.00%	80.00%
	MSCI US Net Div	8.16%	7.69%	8.06%	9.73%	10.00%	10.00%	10.00%	10.00%
	MSCI Japan Net Div	1.84%	2.31%	1.94%	0.27%	0.00%	0.00%	0.00%	0.00%

FIGURE 8 – Résultats de l’optimisation tactique avec contraintes

On observe que sans contraintes le premier portefeuille obtenu n’est autre que le benchmark. En effet la cible de rendement est celle du benchmark, et comme il n’y a aucune contraintes, le benchmark convient car il donne évidemment une tracking error nulle. Avec l’ajout de contraintes, le benchmark ne respectant pas les contraintes spécifiées, on observe que pour atteindre un même rendement, le portefeuille optimal doit s’éloigner du benchmark d’où la tracking error plus élevée.

3.3 Optimisation avec resampling

L’optimisation Moyenne-Variance (MV) de Markowitz est la norme en matière de construction de portefeuille depuis plus de cinquante ans. Presque tous les optimiseurs de portefeuille commerciaux pour l’allocation d’actifs et la gestion de portefeuille sont basés sur une variante de la méthode de Markowitz. Bien que théoriquement importante pour la finance moderne, la sensibilité de l’optimisation MV à l’incertitude des estimations risque-rendement se traduit généralement par un cadre de gestion d’actifs instable et une optimalité de portefeuille ambiguë. Les tests démontrent que les portefeuilles optimisés MV non bornés sont dominés par des portefeuilles équiponderés et n’ont en substance aucune valeur dans l’industrie. Dans la pratique, l’optimisation MV est principalement utilisée comme une manière pour imposer des contraintes ad-hoc et fournir un vernis scientifique à des fins de marketing (Michaud 1989) [1]. Les limites de l’optimisation MV résultent principalement de la manière dont les informations sur les investissements sont utilisées. L’optimisation et le resampling, proposés pour la première fois dans (Michaud 2007) [4], utilisent des techniques de Monte-Carlo pour définir des portefeuilles optimisés par Markowitz et des règles de rééchantillonnage plus efficaces en termes d’investissement.

3.3.1 Description du problème

Le problème qui limite la qualité des portefeuilles optimisés MV n'est pas la théorie de Markowitz. Markowitz fournit une excellente façon d'investir sachant que les paramètres d'entrée (moyenne et variance-covariance) sont connus précisément. Le principal problème avec cette méthodologie est l'erreur d'estimation, ou l'incertitude des paramètres d'optimisation. Les estimations risque-rendements sont très incertaines dans la pratique et la sensibilité aux changements des paramètres d'entrée de l'optimisation conduit à une ambiguïté quant à l'optimalité du portefeuille. Le problème des estimations incertaines est aggravé par la manière dont les informations d'investissement sont représentées numériquement. Une estimation de rendement de 10% est stockée dans l'ordinateur avec 15 décimales. Dans la pratique, il est absurde de considérer que des informations importantes existent à la 16e ou même à la 5e décimale. Cependant, cette vision extrêmement irréaliste de l'information sur les investissements est implicite dans le fonctionnement de nombreux algorithmes d'optimisation. Par conséquent, l'optimisation MV donne bien souvent lieu à des portefeuilles irréalistes qui conduisent à une sous-performance probable.

Cette sensibilité aux paramètres d'entrée est illustrée ci-dessous, nous y comparons deux optimisations de Moyenne-Variance avec une faible variation des rendements et observons l'énorme impact sur les portefeuilles optimaux obtenus.

Rendements annuels et rendements choqués:		
	Rendements (%)	Rendements choqués (%)
JPM UK Traded	-1.90	-1.73
JPM EMU Aggregate in EUR	0.51	0.39
JPM US Traded	2.94	3.16
JPM Japan Traded	1.93	2.23
MSCI UK Net Div	6.08	5.78
MSCI Europe ex. UK in EURO	7.28	7.43
MSCI US Net Div	8.56	9.56
MSCI Japan Net Div	7.26	7.46

FIGURE 9 – Rendements et rendements choqués

Impact de la perturbation des rendements:		
	Poids du portefeuille standard	Poids du portefeuille avec rendements choqués
JPM UK Traded	0.00%	0.00%
JPM EMU Aggregate in EUR	34.84%	39.21%
JPM US Traded	13.75%	12.12%
JPM Japan Traded	2.48%	7.43%
MSCI UK Net Div	7.93%	0.00%
MSCI Europe ex. UK in EURO	16.95%	12.40%
MSCI US Net Div	19.12%	26.94%
MSCI Japan Net Div	4.93%	1.90%

FIGURE 10 – Impact de la perturbation des rendements sur la composition des portefeuilles

En outre, un second problème émerge avec l'optimisation MV, celle de la faible diversification des portefeuilles optimaux. Ci-dessous on observe que le portefeuille de rendement maximal est extrêmement concentré, ce qui va à l'encontre des pratiques des professionnels de l'investissement qui aiment diversifier les sources de risque et de rendement. Sur le graphique les portefeuilles sont en abscisse allant du portefeuille 1 de variance minimale à celui de rendement maximal 10, et les poids des différents actifs sont représentés par les surfaces colorées.

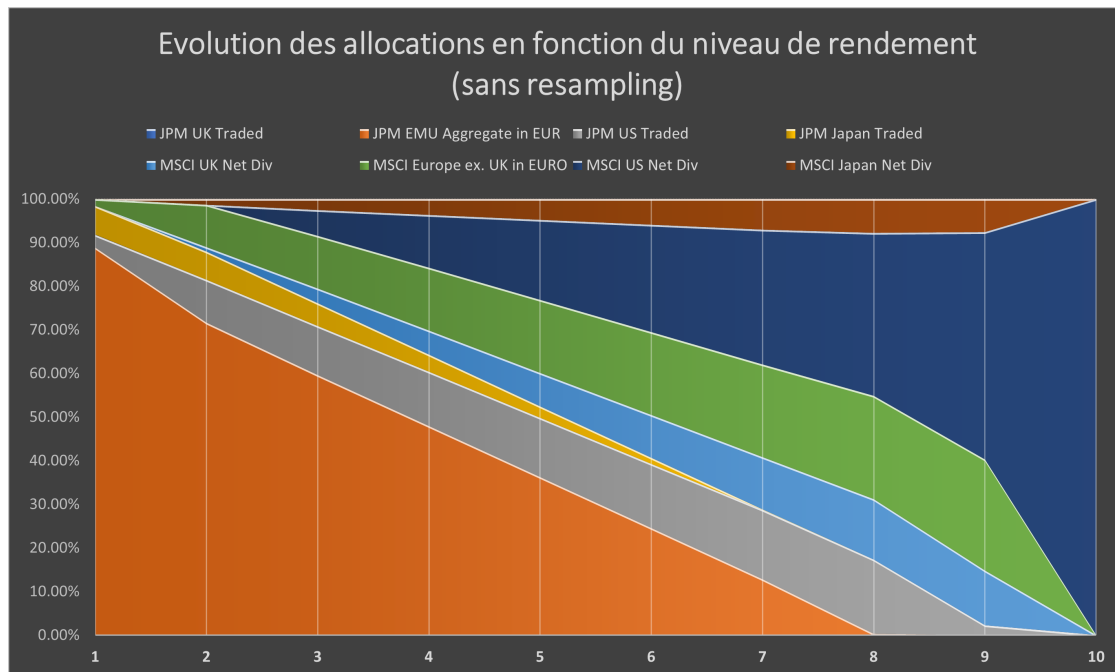


FIGURE 11 – Évolution des poids des portefeuilles non resamplés en fonction du niveau de rendement désiré

Ainsi l'approche MV est difficilement utilisable telle quelle.

3.3.2 Objectif et utilité de l'approche

Le resampling introduit les méthodes de rééchantillonnage par Monte-Carlo dans l'optimisation MV pour refléter de manière plus réaliste l'incertitude des informations d'investissement. Cela fournit généralement des portefeuilles optimisés MV plus stables, plus réalistes et plus performants en termes d'investissement. L'interprétation correcte de l'optimalité des portefeuilles resamplés face à l'approche MV classique est simple à expliquer. Si l'investisseur est 100% certain de ses estimations risque-rendement (avec une précision de 16 décimales ou plus), alors la frontière efficiente de Markowitz est la frontière efficiente pour lui. Sinon, du fait de son incertitude, il sera prêt à attendre moins de rendement et sera moins disposé à prendre des risques.

Ainsi, le portefeuille resamplé permet d'introduire cette sensibilité à l'incertitude de l'information. L'optimisation resamplée est le paradigme de choix pour la prise de décision rationnelle dans des conditions d'incertitude de l'information.

Ci-dessous nous pouvons observer que la frontière efficiente resamplée a tendance à se resserrer comparativement à la frontière efficiente MV, ce qui rejoint le fait qu'en cas d'incertitude l'investisseur sera plus averse au risque.

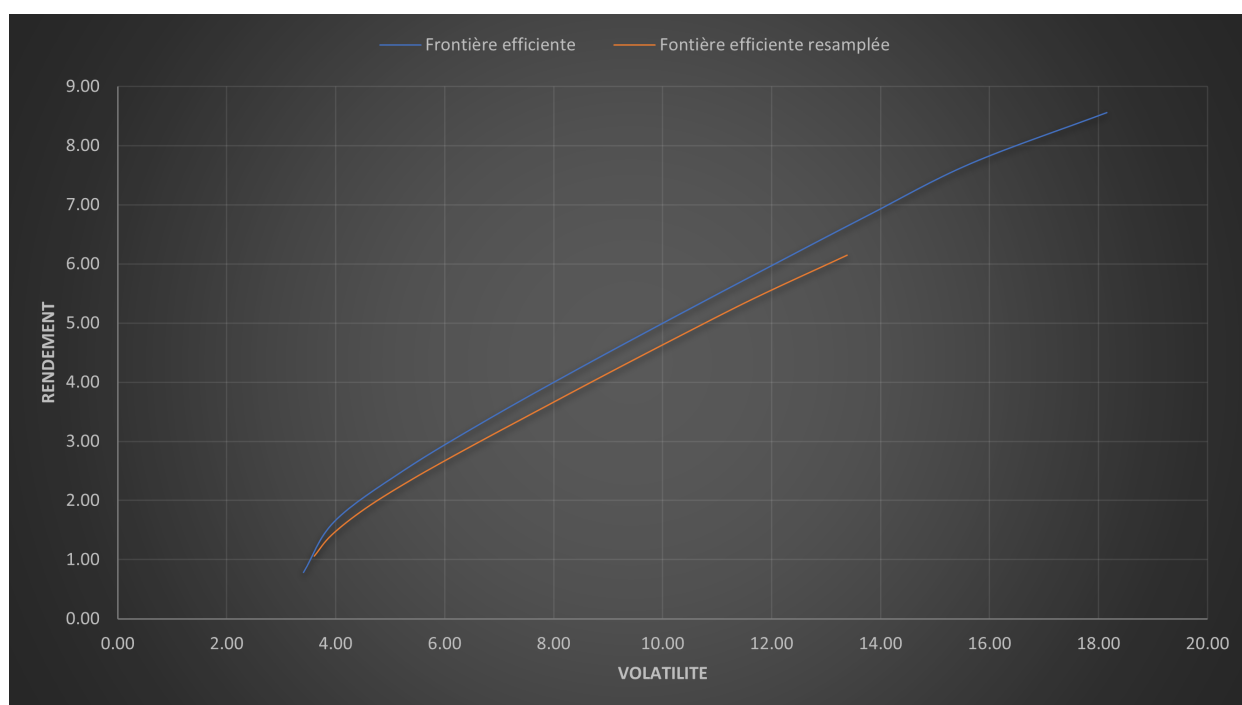


FIGURE 12 – Frontière efficiente et frontière efficiente resamplée

De plus, le rééchantillonnage permet de forcer une meilleure diversification des portefeuilles optimaux, ce qui répond au problème de la section précédente.

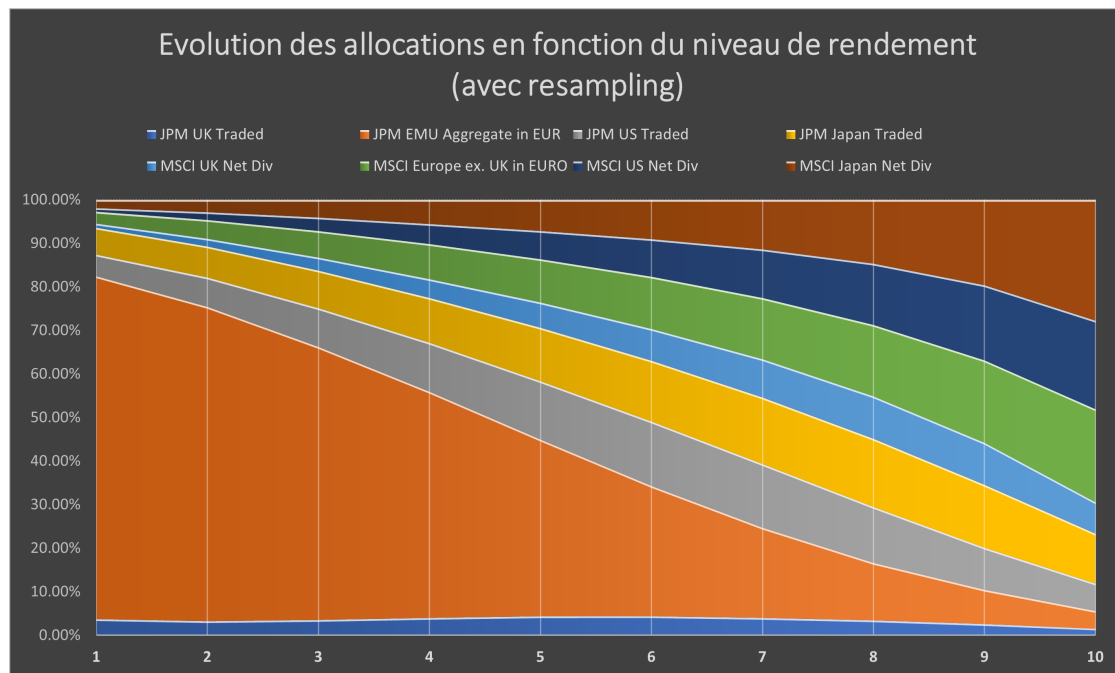


FIGURE 13 – Evolution des poids des portefeuilles resamplés en fonction du niveau de rendement désiré

3.3.3 Méthodologie adoptée

Les portefeuilles optimaux resamplés sont construits comme suit :

1. On échantillonne un vecteur de rendements moyens et une matrice de covariance des rendements centrées sur les valeurs originales (estimation ponctuelle) normalement utilisées dans l'optimisation MV,
2. On calcule une frontière efficiente MV basée sur ces estimations de risque et de rendement échantillonnés,
3. On répète les étapes 1 et 2 (jusqu'à ce que suffisamment d'observations soient disponibles pour la convergence),
4. On moyenne les poids des portefeuille de l'étape 2 pour former les portefeuilles optimaux resamplés.

La méthode de Monte Carlo fournit une méthode directe pour effectuer des simulations et intégrer. Elle est également robuste, puisque sa précision ne dépend que de la mesure la plus grossière de la complexité du problème. Par exemple, l'intégration de Monte-Carlo converge à un taux $O(N^{-1/2})$ qui est indépendant de la dimension de l'intégrale. Pour cette raison, c'est une des seules méthodes viable pour un large éventail de problèmes à haute dimension, allant de la physique atomique à la finance. Le prix à payer pour cette robustesse est que Monte-Carlo peut être extrêmement lent. Le taux de convergence est décélérant, puisqu'un facteur supplémentaire de 4 dans l'effort de calcul ne fournit qu'un facteur supplémentaire de 2 dans l'amélioration de

la précision. Le résultat de cette combinaison de facilité d'utilisation, de large gamme d'application et de convergence lente est qu'une énorme quantité de temps informatique est consacrée aux calculs de Monte-Carlo. Ainsi, même des améliorations modestes de la méthode peuvent avoir un impact considérable sur l'efficacité et le champ d'application des méthodes de Monte-Carlo. En effet, une grande partie de l'effort de développement de Monte-Carlo a porté sur la construction de méthodes de réduction de la variance qui accélèrent le calcul. Les méthodes de réduction de la variance accélèrent le taux de convergence en réduisant la constante devant $O(N^{-1/2})$ pour les méthodes utilisant des séquences aléatoires ou pseudo-aléatoires. Une autre approche d'accélération consiste à modifier le choix de la séquence. Les méthodes de quasi-Monte-Carlo utilisent des séquences quasi-aléatoires (également connues sous le nom de séquences à faible discrédence) au lieu de séquences aléatoires ou pseudo-aléatoires. Contrairement aux séquences pseudo-aléatoires, les séquences quasi-aléatoires ne tentent pas d'imiter le comportement des séquences aléatoires, elles sont déterministes et essaient de "remplir l'espace de manière régulière".

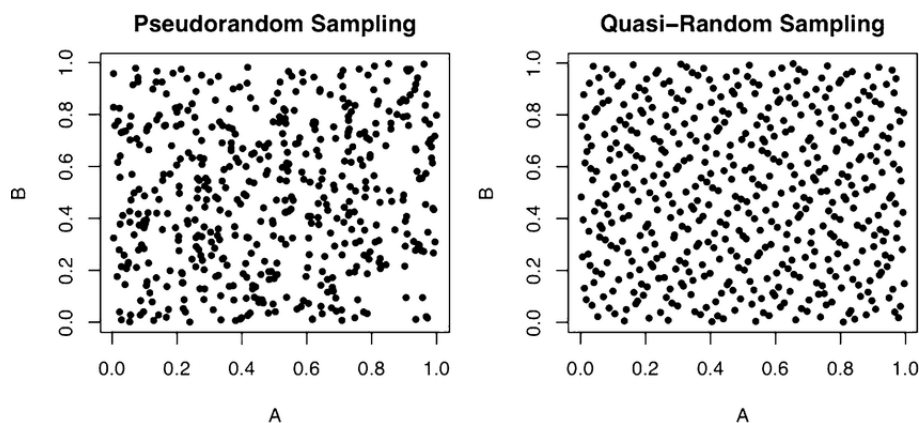


FIGURE 14 – Comparaison des méthodes pseudo-aléatoire et quasi-aléatoire

Le rééchantillonnage tire profit de la méthode de Monte-Carlo, mais il est néanmoins possible d'obtenir une convergence accélérée pour Monte-Carlo en utilisant des séquences quasi-aléatoires (également appelées séquences à faible discrédence), qui sont une alternative déterministe aux séquences aléatoires ou pseudo-aléatoires. Les points d'une séquence quasi-aléatoire sont corrélés pour assurer une plus grande uniformité. La méthode de quadrature qui en résulte, appelée quasi-Monte Carlo, a un taux de convergence d'environ $O((\log N)^k N^{-1})$.

Les méthodes de quasi Monte-Carlo s'appuient ainsi sur des séquences de faible discrédence. Notre choix s'est porté sur la séquence de Faure [5], un dérivé de la bien connue séquence de Halton. La séquence de Faure est comme la séquence de Halton, mais elle n'utilise qu'une seule base pour toutes les dimensions et elle utilise une permutation des éléments du vecteur pour chaque dimension.

1. La base d'une séquence de Faure est le plus petit nombre premier qui est supérieur ou égal au nombre de dimensions noté d .
2. Si $d = 50$, la dernière séquence de Halton (en dimension 50) utilise le 50ème nombre premier qui est 229, alors que la séquence de Faure utilise le premier nombre premier après 50, c'est-à-dire une base 53, qui est beaucoup plus petite que 229, ainsi, le "remplissage des trous" dans les hautes dimensions est plus rapide avec la séquence de Faure qu'avec celle de Halton.
3. En réordonnant la séquence dans chaque dimension, la séquence de Faure évite certains problèmes de corrélation pour les hautes dimensions qui se produisaient avec la séquence de Halton. Ainsi elle établit un lien entre la théorie des séquences à faible discrèpence et la théorie combinatoire pour le ré-ordonnement des vecteurs.

Une fois la séquence de Faure obtenue, nous exploitons la méthode de la fonction de répartition inverse afin d'obtenir des vecteurs gaussiens à partir des tirages uniformes de la séquence de Faure. Nous avons également décidé de coupler l'utilisation de la séquence de Faure, à l'utilisation de variables antithétiques. Les variables antithétiques ou l'échantillonnage antithétique est un ajustement simple pour générer les $\phi_n (1 \leq n \leq N)$. Au lieu de faire N tirages indépendants, on réalise des tirages par paires ; si la i ème variable normalement distribuée est ϕ_i , on prend la $i + 1$ ème variable ϕ_{i+1} pour être $-\phi_i$, puis on tire à nouveau pour ϕ_{i+2} . On notera ici que $-\phi_i$ est également distribuée normalement et, surtout, que la moyenne des deux tirages est nulle, ce qui garantit que la moyenne des trajectoires de l'échantillon sera correcte et que les distributions des tirages seront symétriques. Finalement à l'aide de la décomposition de Cholesky nous sommes capables d'obtenir des variables gaussiennes correctement corrélés, que l'on utilise comme input dans la fonction de la frontière MV classique.

3.4 Optimisation avec intégration de convictions

3.4.1 Objectifs et utilité de l'approche

La gestion active de portefeuille consiste à tirer parti des prévisions des gérants quant au comportement du marché. Comme moyen de prévision, les gestionnaires de portefeuille (GP) ou les analystes collectent des informations, génèrent des points de vue et cherchent à convertir ces points de vue en positions optimales de portefeuille. Ces opinions ne sont pas nécessairement des prévisions explicites du rendement des titres, mais peuvent être des opinions sur la performance relative ou les stratégies de portefeuille. D'autre part, les optimiseurs de portefeuille n'admettent pas directement les opinions comme entrées, mais attendent plutôt une prévision explicite du rendement pour chaque titre. Afin d'alimenter un optimiseur, les gérants doivent traduire leurs

opinions en prévisions de rendement attendus pour les titres concernés, et sont obligés de trouver un chiffre pour représenter "aucune opinion" (zéro par exemple). Cette pratique soulève immédiatement deux questions :

1. Quelle est la manière appropriée de traduire les opinions des gérants en prévisions de rendement explicites ?
2. Est-il légitime d'utiliser un rendement nul pour représenter une "absence d'opinion" ?

En ce qui concerne la deuxième question, les prévisions de rendement nul seront traitées par l'optimiseur comme des vues. Une réponse typique de l'optimiseur sera d'utiliser ce titre pour en exploiter d'autres sur lesquels l'analyste exprime son optimisme. Cela donne facilement lieu à des comportements inattendus (c'est-à-dire des solutions instables, contre-intuitives ou en coin). Des comportements erratiques similaires se produisent également en réponse à des erreurs d'estimation dans le modèle de risque. Pourtant, dans ces situations, l'utilisation courante de contraintes d'optimisation ne résout pas le problème sous-jacent, mais nuit définitivement à l'efficacité de l'approche Moyenne-Variance. À la première question, le modèle d'optimisation de portefeuille global de Black-Litterman (BL) (Black et Litterman, 1992) [6] apporte une réponse élégante. Le modèle établit la prévision dans un cadre analytique bayésien. Dans ce cadre, le GP ne doit produire qu'un nombre "flexible" de vues et le modèle les traduit en douceur en prévisions explicites de rendement des titres, accompagnées d'une matrice de covariance actualisée - exactement ce qu'attend un optimiseur de portefeuille conventionnel à moyenne variance. Si les vues arrivent sous une forme acceptable, c'est-à-dire des vues linéaires, ce modèle peut les consommer entièrement. De plus, le modèle traite la deuxième question avec facilité : sans vue, il existe des justifications théoriques pour prendre les rendements d'équilibre du marché comme prévisions par défaut. Une caractéristique remarquable de cette approche est sa robustesse. Comme les vues postérieures sont une combinaison des vues du marché et des vues de l'analyste, elles partagent une caractéristique commune, la vue du marché, comme point de départ. Sans vues, la meilleure stratégie consiste à s'en tenir à la vue du marché. Avec certaines vues, le portefeuille doit être déformé pour refléter ces vues combinées. Comme le point de vue du marché est toujours pris en compte, il est moins probable de se heurter à des solutions instables ou de coin. Dans le cas où le GP a des opinions fortes qui dominent la vision du marché, le modèle permet également d'ajuster les résultats de manière significative en fonction de ces opinions. Plutôt que d'être erratique, ce phénomène doit être considéré comme attendu et intuitif.

3.4.2 Méthodologie adoptée

Le modèle d'allocation d'actifs Black-Litterman utilise l'approche bayésienne pour déduire les rendements attendus des actifs. Avec l'approche bayésienne, les rendements attendus sont eux-mêmes des variables aléatoires. Ils ne sont pas observables. On peut seulement inférer leur distribution de probabilité. L'inférence commence avec une croyance préalable. Des informations supplémentaires sont utilisées avec la croyance préalable pour déduire la distribution postérieure. Dans le modèle Black-Litterman, la distribution d'équilibre du MEDAF (modèle d'évaluation des actifs financiers) est la croyance préalable, les opinions de l'investisseur sont les informations supplémentaires. L'approche bayésienne est utilisée pour déduire la distribution de probabilité des rendements attendus en utilisant à la fois la distribution préalable du MEDAF et les opinions supplémentaires. Supposons qu'il y ait N actifs sur le marché, qui peuvent comprendre des actions, des obligations, des devises et d'autres actifs. Le modèle suppose que les rendements de ces actifs ont une distribution normale, μ étant le rendement attendu et Σ la matrice de covariance. C'est-à-dire :

$$r \sim \mathcal{N}(\mu, \Sigma)$$

où r est le vecteur des rendements des actifs. À l'équilibre, l'ensemble des investisseurs détient le portefeuille de marché w_{eq} . Les primes de risque d'équilibre, Π , sont telles que si tous les investisseurs détiennent le même point de vue, la demande de ces actifs est exactement égale à l'offre en cours. En supposant que la tolérance moyenne au risque est représentée par le paramètre d'aversion au risque δ , les primes de risque d'équilibre sont données par :

$$\Pi = \delta \Sigma w_{eq}$$

L'a priori bayésien est que les rendements attendus, μ , sont centrés sur les valeurs d'équilibre, c'est-à-dire qu'ils sont normalement distribués avec la moyenne de Π ,

$$\mu = \Pi + \epsilon^{(e)}$$

où $\epsilon^{(e)}$ est un vecteur aléatoire normalement distribué avec une moyenne nulle et une matrice de covariance $\tau \Sigma$, où τ est un scalaire indiquant l'incertitude de la vue du MEDAF.

En plus de l'antériorité du MEDAF, l'investisseur a également un certain nombre d'opinions sur les rendements du marché. Une opinion est exprimée comme une déclaration selon laquelle le rendement attendu d'un portefeuille p a une distribution normale avec une moyenne égale à q et un écart type donné par w . Soit K le nombre total d'opinions, P une matrice $K \times N$

dont les lignes sont les poids de ces portefeuilles et Q un vecteur des rendements attendus de ces portefeuilles. C'est-à-dire :

$$\begin{aligned} P^\top &= \begin{pmatrix} p_1 & p_2 & \cdots & p_K \end{pmatrix} \\ Q^\top &= \begin{pmatrix} q_1 & q_2 & \cdots & q_K \end{pmatrix}. \end{aligned}$$

Alors la vue de l'investisseur peut être exprimée comme :

$$P\mu = Q + \epsilon^{(v)}$$

où $\epsilon^{(v)}$ est une variable aléatoire normalement distribuée non observable avec moyenne nulle et une matrice de covariance diagonale Ω . Nous supposons de plus que $\epsilon^{(e)}$ et $\epsilon^{(v)}$ sont indépendants :

$$\begin{pmatrix} \epsilon^{(e)} \\ \epsilon^{(v)} \end{pmatrix} \sim N \left(0, \begin{bmatrix} \tau\Sigma & 0 \\ 0 & \Omega \end{bmatrix} \right)$$

Ces vues sont ensuite combinées avec le MEDAF antérieur d'une manière Bayésienne. Ainsi, les rendements attendus sont distribués en $\mathcal{N}(\bar{\mu}, \bar{M}^{-1})$, avec $\bar{\mu}$ qui correspond à :

$$\bar{\mu} = [(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P]^{-1} [(\tau\Sigma)^{-1}\Pi + P^\top\Omega^{-1}Q] \quad (11)$$

et la matrice de covariance \bar{M}^{-1} donnée par :

$$\bar{M}^{-1} = [(\tau\Sigma)^{-1} + P^\top\Omega^{-1}P]^{-1}.$$

En pratique, il est assez simple d'appliquer l'équation (11) pour calculer la moyenne des rendements attendus. Cependant, il est souvent difficile de trouver les intuitions économiques originales des vues en observant ces nombres. Le problème lorsqu'on essaie de donner un sens aux rendements attendus des différentes classes d'actifs est qu'ils sont tous liés les uns aux autres par leurs volatilités et corrélations relatives. Par exemple, il est facile de conjecturer que le marché allemand des actions va surperformer le reste des marchés européens, mais il est beaucoup moins facile d'imaginer l'implication de cette conjecture sur les rendements attendus relatifs du marché allemand et des autres marchés. Augmenter le rendement attendu des actions allemandes et maintenir tous les autres rendements attendus fixes ne suggère pas, dans les mathématiques d'un optimiseur, une surpondération des actions allemandes - cela suggère plutôt à l'optimiseur qu'en utilisant les volatilités et corrélations relatives des différents marchés, il peut créer un portefeuille beaucoup plus compliqué avec un rendement attendu plus élevé et un risque plus faible que celui qui serait disponible simplement en créant une surpondération

du marché allemand. Dans l'approche Black Litterman, l'opinion selon laquelle les actions allemandes surperformeront le reste des actions européennes est exprimée sous la forme d'une attente de rendement positif pour un portefeuille composé d'une position importante sur les actions allemandes et de positions faibles pondérées en fonction de la capitalisation boursière sur le reste des marchés européens. Ce point de vue est ensuite traduit par l'équation (11), qui tient compte de manière appropriée des volatilités et des corrélations, comme ajustements des rendements attendus sur tous les marchés. Ces types de transformations complexes, des vues sur les portefeuilles au vecteur de rendement attendu, et du vecteur de rendement attendu au portefeuille optimal, sont généralement difficiles à comprendre. Ainsi, au lieu d'étudier directement les rendements attendus, nous examinons les pondérations du portefeuille optimal de Black-Litterman. Nous commençons par le cas d'un investisseur sans contrainte avec un paramètre d'aversion au risque égal à δ . Ce portefeuille nous donne des indications très intéressantes sur le modèle d'allocation d'actifs de Black-Litterman.

Comme les rendements attendus sont eux-mêmes des variables aléatoires dans le modèle de Black-Litterman, la distribution des rendements n'est plus simplement $\mathcal{N}(\bar{\mu}, \Sigma)$. Selon les équations précédentes, la distribution des rendements est la suivante :

$$r \sim \mathcal{N}(\bar{\mu}, \bar{\Sigma}),$$

où $\bar{\Sigma} = \Sigma + \bar{M}^{-1}$. Étant donné la moyenne $\bar{\mu}$ et la matrice de covariance $\bar{\Sigma}$, le portefeuille optimal peut être construit en utilisant la méthode standard d'optimisation moyenne-variance. Pour un investisseur dont le paramètre d'aversion au risque est δ , le problème de maximisation peut être écrit comme suit

$$\max w^\top \bar{\mu} - \frac{\delta}{2} w^\top \bar{\Sigma} w.$$

La condition du premier ordre donne que

$$\bar{\mu} = \delta \bar{\Sigma} w^*,$$

ou de manière équivalente,

$$w^* = \frac{1}{\delta} \bar{\Sigma}^{-1} \bar{\mu},$$

où w^* est le vecteur des poids optimaux du portefeuille. En utilisant l'équation (11), les poids optimaux du portefeuille peuvent être écrits comme suit

$$w^* = \frac{1}{\delta} \bar{\Sigma}^{-1} \bar{M}^{-1} [(\tau \Sigma)^{-1} \Pi + P^\top \Omega^{-1} Q].$$

On notera le fait que,

$$\bar{\Sigma}^{-1} = (\Sigma + \bar{M}^{-1})^{-1} = \bar{M} - \bar{M} (\bar{M} + \Sigma^{-1})^{-1} \bar{M}$$

le terme $\bar{\Sigma}^{-1} \bar{M}^{-1}$ peut être simplifié comme suit

$$\bar{\Sigma}^{-1} \bar{M}^{-1} = \frac{\tau}{1 + \tau} \left(I - P^\top A^{-1} P \frac{\Sigma}{1 + \tau} \right)$$

où la matrice $A = \Omega/\tau + P\Sigma/(1 + \tau)P^\top$. Les poids optimaux du portefeuille, $w^* = \bar{\Sigma}^{-1} \bar{\mu}/\delta$ peuvent maintenant être écrits comme suit

$$w^* = \frac{1}{1 + \tau} (w_{eq} + P^\top \times \Lambda)$$

où $w_{eq} = (\delta\Sigma)^{-1}\Pi$ est le portefeuille d'équilibre du marché et Λ est un vecteur défini comme suit

$$\Lambda = \tau\Omega^{-1}Q/\delta - A^{-1}P \frac{\Sigma}{1 + \tau} w_{eq} - A^{-1}P \frac{\Sigma}{1 + \tau} P^\top \tau\Omega^{-1}Q/\delta$$

Puisque chaque colonne de la matrice P^\top est un portefeuille de vue, le portefeuille optimal de l'investisseur est le portefeuille d'équilibre du marché w_{eq} , plus une somme pondérée des portefeuilles formant les vues, puis mis à l'échelle par un facteur de $1/(1 + \tau)$. Le poids de chaque portefeuille est donné par l'élément correspondant dans le vecteur Λ .

4 Amélioration de la robustesse de l'environnement

Au-delà de la réplication de l'environnement d'optimisation de portefeuille, le stage a été l'occasion de travailler sur la structure du projet. En effet, le passage à une version en Python a été l'occasion de développer un écosystème de suivi et de test.

4.1 Structure du projet

Le projet est scindé en deux entités distinctes, la première correspond au coeur de l'environnement et contient l'ensemble des fonctions. Elle prend la forme d'un package Python, avec une documentation. Ce package peut être directement utilisé par des utilisateurs de Python, sous condition d'avoir l'ensemble des dépendances installées. Les fonctions ont été construites sur le principe de la programmation orientée objet, si bien que l'utilisation en est grandement facilitée.

Parallèlement, la seconde partie de l'environnement est le serveur qui permet de faire appel aux

différentes fonctions du coeur depuis Excel. Lorsque l'utilisateur utilise Excel, un complément doit être installé en local sur sa machine, il peut ensuite utiliser les fonctions du coeur directement depuis Excel. Une couche de VBA permet de traiter les paramètres d'entrée, puis réalise une requête au serveur (sur un port et à une adresse IP donnée) qui reçoit la requête, la traite puis renvoie le résultat au code VBA qui la traite et l'affiche in-fine dans Excel.

4.2 Création d'un système de test efficace

Auparavant, des fichiers Excel permettaient de tester les fonctions codées en C/C++, à l'aide de tests de non-régression. L'idée était de pouvoir comparer les résultats des fonctions lorsque des modifications étaient apportées au code. Dans le nouvel environnement, nous avons décidé de migrer intégralement les tests en Python. Grâce au package de base de Python, `unittest`, j'ai créé une interface simple qui permet de comparer les résultats des fonctions à des résultats pré-enregistrés dans des fichiers de texte. Par ailleurs, l'interface permet de générer des nouveaux fichiers de référence à la volée, en ne changeant qu'un des paramètres dans le code.

De manière schématique, cette interface génère des tests très rapidement avec pour seule contrainte la définition des plages dans lesquelles se situent les paramètres de la fonction à tester. En outre, tout y est dynamique, si bien qu'il sera aisé de rajouter des tests au fil de l'enrichissement de l'environnement.

4.3 Introduction du logging

Le principe du logging est de garder une trace de l'utilisation d'un service par l'utilisateur pour s'assurer que le service fonctionne comme il se doit, mais aussi pour des raisons de sécurité. Étant donné que le projet est composé de deux parties distinctes, nous avons décidé d'y intégrer deux systèmes de logging. L'un pour le coeur qui stocke les erreurs, lors de l'utilisation, localement sur l'ordinateur de l'utilisateur, et l'autre sur le serveur, qui garde une trace de toutes les connexions entrantes et des éventuelles erreurs. Les deux modules de logging sont réglables directement en ne changeant qu'un paramètre binaire et le niveau de suivi (avoir la trace des erreurs critiques, des avertissements, ou bien même des messages d'information).

5 Enrichissement des fonctionnalités de l'environnement d'optimisation

Après la réplication de l'environnement d'optimisation de portefeuille, le but a été de proposer des améliorations à celui-ci. Trois types d'améliorations ont été réalisées :

1. la possibilité de spécifier des vues non-linéaires quant à la distribution des rendements des actifs utilisés dans les optimisations,
2. l'introduction des coûts de transaction à l'approche de MV traditionnelle afin d'obtenir un modèle d'optimisation de portefeuille plus proche des contraintes réelles d'investissement,
3. la possibilité d'optimiser par rapport à des variables discrètes et non plus continues (comme le poids des actifs),

5.1 Entropy pooling

5.1.1 Description du problème

L'analyse de scénarios permet au gérant d'explorer les implications sur un portefeuille donné d'un ensemble d'opinions subjectives sur les réalisations possibles du marché. L'approche pionnière de Black et Litterman (BL) généralise l'analyse de scénarios, en ajoutant une incertitude sur les opinions et sur le modèle de risque de référence. Pour aller plus loin, Meucci a étendu le modèle ci-dessus pour agir sur les facteurs de risque au lieu des rendements, et couvrir ainsi les marchés dérivés hautement non linéaires et les vues sur les facteurs externes qui n'influencent le $P&L$ que statistiquement. Dans méthode de BL, la distribution de référence des facteurs de risque est normale et les vues non linéaires ne sont pas autorisées.

Le problème ici est donc que les techniques habituelles formulent des hypothèses quant aux distributions des rendements des actifs. Des hypothèses qui sont souvent fallacieuses. La linéarité des vues contraint aussi énormément l'utilisation de la méthode de BL, car donner une estimation du rendement futur d'un actif (exemple de vue linéaire) est beaucoup plus difficile que d'estimer la volatilité futur d'un actif (sa variance). C'est pour cette raison que nous présentons ici l'approche par 'entropy pooling' (EP dans la suite) qui généralise BL. Les paramètres d'entrée de cette méthode sont un modèle de marché arbitraire, appelé "antérieur", et des vues sur ce marché concernant son évolution futur. Le résultat est une distribution, appelée "postérieure", qui incorpore toutes les vues et peut être utilisée pour la gestion du risque et l'optimisation du portefeuille.

Entre autres, l'EP gère les marchés non normaux, les vues sur les combinaisons non linéaires

de facteurs de risque qui ont un impact sur le $P&L$ directement ou seulement statistiquement par le biais des corrélations, les vues sur les rendements attendus, mais aussi sur les quantiles, notamment utiles en gestion de risque mais encore les vues sur les volatilités, les corrélations, etc.

5.1.2 Objectif et utilité de l'approche

La méthode de l'entropy pooling a pour but de spécifier des vues quant à la distribution des rendements des actifs ; elle a l'avantage d'être une approche non-paramétrique dans le sens où il n'est pas nécessaire d'associer une loi de probabilité aux rendements, à l'inverse des méthodes traditionnelles, souvent paramétriques comme Black-Litterman. De plus, l'approche de Meucci, permet de spécifier des vues bien plus riches que BL, comme des vues sur le kurtosis ou le skew de la distribution. En quelque sorte cette méthode permet d'étendre l'idée de BL à des contraintes plus larges, tout en préservant au maximum la distribution initiale des rendements grâce à la minimisation de l'entropie.

Une telle flexibilité est parfaitement adaptée aux besoins concrets des professionnels de l'allocation de portefeuille, dans la mesure où il est souvent difficile de formuler des vues sur les premiers moments de la distribution des rendements. Ainsi un gestionnaire peut avoir une opinion sur la volatilité future d'un actif et intégrer cette opinion à sa décision d'allocation en générant une nouvelle distribution de rendements qui respecte cette opinion en étant le moins éloignée de la distribution initiale que possible. C'est heuristiquement l'idée fondatrice de cette approche. Par ailleurs, il est même possible de fournir un niveau de confiance pour l'opinion qui peut être utilisé pour mélanger les distribution initiale et postérieure à l'instar de ce qui est fait dans BL.

5.1.3 Méthodologie adoptée

Considérons un portefeuille piloté par un vecteur à N dimensions de facteurs de risque \mathbf{X} . En d'autres termes, en désignant par f le temps actuel, par \mathcal{L}_t l'information actuellement disponible, et par π le temps jusqu'à l'horizon d'investissement, il existe une fonction déterministe P qui fait correspondre les réalisations de \mathbf{X} et de l'information \mathcal{I}_t au prix $P_{t+\pi}$ de chaque titre du portefeuille à l'horizon :

$$P_{t+\pi} \equiv P(X, \mathcal{I}_t).$$

Cette description est très simplement généralisable. Par exemple, dans un portefeuille optionnel, \mathbf{X} peut représenter les changements de tous les sous-jacents et les volatilités implicites : dans ce cas, l'équation précédente est approximée par une expansion de Taylor du second ordre dont les

coefficients sont les deltas, vegas, gammas, vanna, volgas, etc. De même, \mathbf{X} peut représenter un ensemble de facteurs de risque dont dépend une fonction de pricing Monte-Carlo coûteuse en calcul etc. En outre, \mathbf{X} peut être augmenté d'un ensemble de facteurs de risque externes qui n'alimentent pas directement la fonction de pricing, mais qui influencent tout de même le $P\&L$ statistiquement, par corrélation par exemple.

MODÈLE DE RÉFÉRENCE

Le modèle de Meucci suppose l'existence d'un modèle de risque, c'est-à-dire un modèle qui régie la distribution conjointe des facteurs de risque, telle que représentée par sa fonction de densité de probabilité ($f_{\mathbf{X}}$)

$$\mathbf{X} \sim f_{\mathbf{X}}.$$

Avec BL, ce qui est utilisé est la distribution "antérieure" des facteurs. Plus généralement, il s'agit d'un modèle que les gestionnaires de risque utilisent pour effectuer des analyses de risque, telles que le calcul de la volatilité, de la tracking error, de la VaR, de l'expected shortfall d'un portefeuille, ainsi que les contributions des différentes sources de risque à ces mesures. Les gestionnaires de portefeuille et les traders, quant à eux, utilisent ce modèle pour optimiser leurs positions. Ils spécifient un indice subjectif de satisfaction \mathcal{S} , tel que le compromis VaR moyen. Dans notre cas, la satisfaction dépend à la fois de la distribution du marché $f_{\mathbf{X}}$ à travers les prix, et des positions du portefeuille, représentées par un vecteur \mathbf{w} . Alors le portefeuille optimal \mathbf{w}^* est défini comme suit

$$\mathbf{w}^* \equiv \operatorname{argmax}_{\mathbf{w} \in \mathcal{C}} \{ \mathcal{S}(\mathbf{w}; f_{\mathbf{X}}) \}$$

où \mathcal{C} est un ensemble donné de contraintes d'investissement. Ce modèle de référence peut être estimé à partir d'une analyse historique, ou calibré en fonction des observables actuelles du marché. Dans notre cas c'est l'approche historique qui est utilisée et fera donc foi.

Dans le cas le plus général, l'utilisateur exprime des opinions sur des fonctions génériques du marché $g_1(\mathbf{X}), \dots, g_K(\mathbf{X})$. Ces fonctions constituent une variable aléatoire à K dimensions dont la distribution conjointe est contenue dans le modèle de référence :

$$\mathbf{V} \equiv \mathbf{g}(\mathbf{X}) \sim f_{\mathbf{V}}$$

Ici contrairement à BL, dans EP nous n'émettons pas l'hypothèse que les fonctions g_k sont

linéaires. Toutefois, les vues sont des déclarations sur les variables qui peuvent entrer en conflit avec le modèle de référence. Par conséquent, la meilleure spécification de vue est celle qui est la plus détaillée possible, une distribution conjointe complète et subjective, évitant de fait les conflits. On fournit donc les vues comme un distribution conjointe :

$$\mathbf{V} \sim \tilde{f}_{\mathbf{V}} \neq f_{\mathbf{V}}.$$

Détaillons ci-dessous le type de vues qu'il est possible de fournir :

- Les vues classiques sont des opinions sur $\tilde{\mathbb{E}}\{V_k\}$, les valeurs attendues de chacun des V_k selon la nouvelle distribution $\tilde{f}_{\mathbf{V}}$. De plus, Meucci considère les vues sur une mesure de localisation plus générale $\tilde{m}\{V_k\}$, qui peut être, par exemple, l'espérance ou la médiane. Les vues sont alors définies comme suit :

$$\tilde{m}\{V_k\} \underset{\geq}{\lesseqgtr} m_k, \quad k = 1, \dots, K,$$

Souvent on définira les valeurs m_k de manière exogène. Ainsi si l'utilisateur n'a que des vues qualitatives, il est pratique de définir comme dans Meucci (2010) [7] :

$$m_k \equiv m\{V_k\} + \varkappa\sigma\{V_k\}.$$

Dans cette expression, σ est une mesure de la volatilité du modèle de référence, telle que l'écart type ou, sur les marchés à queue épaisses, l'écart interquantile ; et \varkappa est un multiplicateur ad hoc, tel que $-2, -1, 1,$ et 2 pour "très baissier", "baissier", "haussier" et "très haussier" respectivement.

- En outre, EP peut traiter des vues exprimées sous forme d'inégalités. En particulier, EP peut traiter des relations d'ordre, fréquentes dans la gestion des actions et des obligations :

$$\tilde{m}\{V_1\} \geq \tilde{m}\{V_2\} \geq \dots \geq \tilde{m}\{V_K\}.$$

De cela on déduit que l'on peut par exemple fournir un vue comme : "le rendement du premier actif sera plus faible que celui du deuxième qui sera lui même plus faible que celui du troisième actif".

- Les vues peuvent également concerner les volatilités. Une formulation commode est la suivante :

$$\tilde{\sigma}\{V_k\} \underset{\geq}{\lesseqgtr} \varkappa\sigma\{V_k\}, \quad k = 1, \dots, K.$$

en reprenant les idées des points précédents.

- De plus, grâce à l’approche de Meucci, il est possible de saisir des vues sur le comportement de la queue gauche (ou droite) de la distribution, représentées par exemple par $\tilde{Q}_V(u)$, le quantile de V_k selon la nouvelle distribution $\tilde{f}_{\mathbf{V}}$, où le niveau de la queue u est proche de zéro (un). Une bonne spécification est :

$$\tilde{Q}_V(u) \geq Q_V(u),$$

où Q_V est le quantile de référence induit par f_V , ou bien des quantiles de référence tels que ceux de la loi normale ou la loi de Student- t .

Ce qui précède est une liste très partielle de toutes les caractéristiques possibles sur lesquelles l’utilisateur peut souhaiter exprimer des points de vue, et qui peuvent être traitées par EP.

DISTRIBUTION POSTÉRIEURE

La distribution postérieure doit satisfaire les vues sans trop perturber la distribution antérieure, c’est l’idée centrale de Meucci. Le fait de travailler au niveau de la distribution (ce qui est très général) permet de rester le plus proche possible du modèle de référence. Ceci passe par la minimisation de l’entropie relative de la distribution postérieure, rapport à celle de la distribution antérieure. L’entropie relative entre une distribution générique $\tilde{f}_{\mathbf{X}}$ et une distribution de référence $f_{\mathbf{X}}$ est définie par

$$\mathcal{E}(\tilde{f}_{\mathbf{X}}, f_{\mathbf{X}}) \equiv \int \tilde{f}_{\mathbf{X}}(\mathbf{x}) \left[\ln \tilde{f}_{\mathbf{X}}(\mathbf{x}) - \ln f_{\mathbf{X}}(\mathbf{x}) \right] d\mathbf{x}$$

et c’est une mesure naturelle de la quantité d’information dans $\tilde{f}_{\mathbf{X}}$; de plus, elle mesure également la distorsion de $\tilde{f}_{\mathbf{X}}$ par rapport à $f_{\mathbf{X}}$. En effet, si les deux distributions coïncident, l’entropie relative est égale à zéro, et en imposant des contraintes sur $\tilde{f}_{\mathbf{X}}$ cette distribution s’écarte de $f_{\mathbf{X}}$ et l’entropie relative augmente. Par conséquent, Meucci définit la distribution postérieure du marché comme une minimisation de la fonction précédente pour un ensemble de contraintes :

$$\tilde{f}_{\mathbf{X}} \equiv \operatorname{argmin}_{f \in \mathbb{V}} \{ \mathcal{E}(f, f_{\mathbf{X}}) \}$$

où $f \in \mathbb{V}$ représente toutes les distributions compatibles avec les vues énoncées (contraintes). Finalement, si la minimisation de l’entropie est une méthode largement appliquée en physique et en statistique, elle est encore plus naturelle dans notre contexte, car elle généralise la technique de la mise à jour bayésienne [8].

NIVEAU DE CONFIANCE

Après la minimisation de l'entropie relative, une dernière étape est nécessaire. La postérieure $\tilde{f}_{\mathbf{X}}$ suit toutes les vues du praticien mais cela suppose qu'il a une confiance totale dans ses déclarations. Si ce n'est pas le cas, la distribution postérieure doit se rapprocher de la distribution des facteurs de référence. Ceci est facilement réalisable, de manière intuitive, en mettant en commun les opinions du modèle de référence et de la distribution postérieure :

$$\tilde{f}_{\mathbf{X}}^c \equiv (1 - c)f_{\mathbf{X}} + c\tilde{f}_{\mathbf{X}}$$

où $c \in [0, 1]$ représente le niveau de confiance dans les vues : dans le cas extrême où la confiance est totale, le postérieur à confiance totale est récupérée ; en revanche, en l'absence de confiance du gérant ($c = 0$), le modèle de risque de référence est récupéré. La mise en commun des vues est très utile dans un contexte de multigestion. En effet, considérons S utilisateurs qui saisissent leurs opinions distinctes sur (éventuellement, mais pas nécessairement) différentes fonctions du marché. Nous obtenons alors S distributions postérieures de confiance totale $\tilde{f}_{\mathbf{X}}^{(s)}$, $s = 1, \dots, S$. La distribution postérieure est alors naturellement la moyenne pondérée par les niveaux de confiance des distributions postérieures individuelles :

$$\tilde{f}_{\mathbf{X}}^c \equiv \sum_{s=1}^S c_s \tilde{f}_{\mathbf{X}}^{(s)}$$

Concrètement, ces niveaux de confiance peuvent être liés au track-record des managers (leurs performances respectives passées), c'est-à-dire que la s -ième confiance c_s peut être définie comme une fonction croissante du nombre de vues correctes passées, comme une sorte de mesure de la compétence des gérants. En étendant cette méthode, on peut facilement spécifier différents niveaux de confiance pour les différentes vues d'un même utilisateur et les intégrer dans un contexte multi-utilisateurs. Il s'avère que cela revient à spécifier une mesure de probabilité pour chaque niveau de confiance.

IMPLÉMENTATION

En général, la méthode EP ne peut pas être mise en œuvre de manière analytique. Cependant, l'implémentation numérique de l'EP dans le cas historique, le nôtre, est extrêmement

simple et efficace sur le plan informatique. Tout d’abord, nous représentons la distribution de référence du marché \mathbf{X} en termes d’un historique $J \times N$ \mathcal{X} de réalisations (les rendements quotidiens des actifs par exemple) : la j ème ligne générique de \mathcal{X} représente un scénario parmi un très grand nombre de scénarios conjoints pour les N variables \mathbf{X} , tandis que la n ème colonne générique de \mathcal{X} représente la distribution marginale du n ème facteur X_n . Aux scénarios nous associons le vecteur $J \times 1$ des probabilités respectives \mathbf{p} , dont chaque entrée est typiquement, mais pas nécessairement, égale à $1/J$. Le fait de prendre $1/J$ revient à considérer qu’il y a équiprobabilité de chaque réalisation.

Nous supposons que chacun des scénarios conjoints dans \mathcal{X} a été mis en correspondance avec les scénarios de prix conjoints respectifs pour les I actifs du marché considéré par l’utilisateur (dans le cas où les réalisations ne sont pas les rendements eux mêmes), au moyen d’une fonction, potentiellement coûteuse, à l’instar d’une fonction de pricing par Monte-Carlo, générant ainsi un panel de prix \mathcal{P} de taille $J \times I$. Le panel des prix des titres \mathcal{P} , ainsi que les probabilités respectives \mathbf{p} , est ensuite introduit dans un algorithme d’optimisation pour effectuer l’étape d’allocation d’actifs. Ensuite, l’utilisateur exprime des points de vue sur des fonctions (non linéaires) du marché. Alors, leur distribution, telle qu’impliquée par le modèle de référence, est facilement représentée par le panneau $J \times K$, \mathcal{V} défini en entrée comme suit :

$$\mathcal{V}_{j,k} \equiv g_k(\mathcal{X}_{j,1}, \dots, \mathcal{X}_{j,N}),$$

Pour représenter la distribution postérieure du marché qui inclut les vues, au lieu de générer de nouvelles réalisations, nous utilisons les mêmes réalisations avec des probabilités différentes $\tilde{\mathbf{p}}$. Ensuite, les vues peuvent être écrites comme un ensemble de contraintes linéaires sur les nouvelles probabilités, encore à déterminer :

$$\underline{\mathbf{a}} \leq \mathbf{A}\tilde{\mathbf{p}} \leq \bar{\mathbf{a}}$$

où \mathbf{A} , $\underline{\mathbf{a}}$ et $\bar{\mathbf{a}}$ sont des expressions simples du panel. Dans ce cadre, l’entropie relative devient sa contrepartie discrète et la distribution postérieure avec une confiance totale est définie comme suit :

$$\mathcal{E}(\tilde{\mathbf{p}}, \mathbf{p}) \equiv \sum_{j=1}^J \tilde{p}_j [\ln(\tilde{p}_j) - \ln(p_j)]$$

Ce problème d’optimisation peut être résolu très efficacement car la formulation duale est un simple problème convexe à contraintes linéaires dans un nombre de variables égal au nombre de vues, et non au nombre de réalisations de l’historique, qui peut être maintenu élevé. Cela permet notamment d’obtenir une excellente précision même dans le cas de vues extrêmes. On peut

finaleme nt calculer la distribution postérieure pondérée en fonction des niveaux de confiance et de la mise en commun des vues : elle est représentée par $(\mathcal{X}, \mathbf{p}_c)$, les mêmes réalisations que pour le modèle de référence, mais avec de nouvelles probabilités

$$\mathbf{p}_c \equiv (1 - c)\mathbf{p} + c\tilde{\mathbf{p}}$$

Puisque la distribution postérieure des facteurs est obtenue en modifiant les probabilités relatives des scénarios \mathcal{X} sans affecter les scénarios eux-mêmes, la distribution postérieure des prix du marché est représentée par $(\mathcal{P}, \mathbf{p}_c)$, le panel original des prix conjoints et les nouvelles probabilités. Par conséquent, aucun repricing n'est nécessaire pour traiter les vues et les stress-tests.

EXEMPLES

Dans cette section nous proposons d'étudier les trois premiers types de vues. Les données sur lesquelles nous avons travaillé sont les historiques de rendement quotidien de 9 indices de référence que nous considérons ici comme des actifs financiers investissables. L'historique est de 15 ans. Les principales caractéristiques des rendements quotidiens sont présentées ci-dessous, avec la matrice de corrélation (la moyenne et la volatilité sont exprimés en %, ainsi que la corrélation).

	Mean	Volatility	Skewness	Kurtosis	CVaR
Equity US	0.045	1.351	-0.368	12.186	-0.165
Equity Europe ex. EMU	0.021	1.172	-0.383	11.783	-0.145
Equity Japan	0.018	1.326	-0.021	8.301	-0.153
Govies EUR	0.010	0.170	0.408	12.632	-0.019
Govies US	0.009	0.308	-0.035	6.685	-0.034
Credit IG Europe	0.011	0.150	-1.234	14.612	-0.019
Credit IG US	0.014	0.360	-0.866	11.364	-0.043
Credit HY US	0.021	0.354	-1.563	31.551	-0.046
Credit HY Europe	0.021	0.326	-2.455	32.514	-0.044

FIGURE 15 – Principales caractéristiques de la distribution des rendements quotidiens

	0	1	2	3	4	5	6	7	8
(0, Equity US)	100.0	53.0	9.0	-6.0	-37.0	-2.0	-19.0	30.0	18.0
(1, Equity Europe ex. EMU)	53.0	100.0	17.0	-10.0	-36.0	-5.0	-10.0	56.0	42.0
(2, Equity Japan)	9.0	17.0	100.0	8.0	6.0	26.0	18.0	26.0	32.0
(3, Govies EUR)	-6.0	-10.0	8.0	100.0	38.0	67.0	35.0	3.0	15.0
(4, Govies US)	-37.0	-36.0	6.0	38.0	100.0	40.0	83.0	-16.0	-11.0
(5, Credit IG Europe)	-2.0	-5.0	26.0	67.0	40.0	100.0	54.0	27.0	43.0
(6, Credit IG US)	-19.0	-10.0	18.0	35.0	83.0	54.0	100.0	27.0	26.0
(7, Credit HY US)	30.0	56.0	26.0	3.0	-16.0	27.0	27.0	100.0	76.0
(8, Credit HY Europe)	18.0	42.0	32.0	15.0	-11.0	43.0	26.0	76.0	100.0

FIGURE 16 – Matrice de corrélation des rendements quotidiens

Les scénarios que nous avons sélectionnés sont exprimés de manière annuelle (rendements annuels et volatilités annuelles), en utilisant l'approximation de 252 jours ouvrés annuels ($r_y = r_d \times 252$ et $\sigma_y = \sigma_d \times \sqrt{252}$ où r_d et r_y sont respectivement les rendements quotidiens et annuels et σ_d et σ_y sont respectivement les volatilités quotidiennes et annuelles). Les scénarios sont les suivants :

1. La moyenne des rendements de l'indice Equity US est égale à 40% i.e. :

$$r_y^{Eq.US} = 40\%,$$

2. La moyenne des rendements annuels de l'indice Equity US est inférieure à celle de l'indice Equity Europe ex. EMU - 3% qui est elle même inférieure à celle de l'indice Equity Japan et la moyenne des rendements annuels de l'indice Govies US est inférieure à celle de l'indice Govies EUR moins 3% i.e. :

$$r_y^{Eq.US} \leq r_y^{Eq.EU} - 3\% \leq r_y^{Eq.JPN} \text{ et } r_y^{Gov.US} \leq r_y^{Gov.EU},$$

3. La volatilité annuelle des rendements de l'indice Equity US est supérieure à 32% i.e. :

$$\sigma_y^{Eq.US} \geq 32\%.$$

Après minimisation de l'entropie dans les différents cas nous obtenons :

- Scénario 1 :

	Mean	Volatility	Skewness	Kurtosis	CVaR
Equity US	0.16	1.35	0.39	12.09	-0.13
Equity Europe ex. EMU	0.07	1.16	-0.03	11.57	-0.13
Equity Japan	0.03	1.33	-0.05	8.25	-0.15
Govies EUR	0.01	0.17	0.47	13.13	-0.02
Govies US	-0.00	0.31	-0.19	6.88	-0.04
Credit IG Europe	0.01	0.15	-1.09	12.88	-0.02
Credit IG US	0.01	0.36	-0.87	11.10	-0.04
Credit HY US	0.03	0.35	-1.16	31.65	-0.04
Credit HY Europe	0.02	0.32	-2.05	29.70	-0.04

ENS = 99.64%.
RE = 0.36%.

FIGURE 17 – Caractéristiques de la distribution des rendements post EP

où $r_d = 0.16\%$ et donc $r_y = 0.16\% \times 252 \approx 40\%$, ce qui correspond bien à nos attentes. L'entropie relative est assez faible, 0.36%, ce qui signifie que la déformation des probabilités initiales est faible. Cela peut d'ailleurs s'observer à la lumière de la figure en annexe qui compare les probabilités initiale et après EP, sur laquelle on observe une perturbation restreinte.

— Scénario 2 :

	Mean	Volatility	Skewness	Kurtosis	CVaR
Equity US	0.03	1.35	-0.46	11.82	-0.17
Equity Europe ex. EMU	0.04	1.17	-0.22	11.81	-0.14
Equity Japan	0.03	1.33	0.11	8.80	-0.15
Govies EUR	0.01	0.17	0.65	14.32	-0.02
Govies US	-0.00	0.31	-0.15	6.65	-0.04
Credit IG Europe	0.01	0.15	-1.34	16.28	-0.02
Credit IG US	0.01	0.36	-0.97	12.25	-0.05
Credit HY US	0.02	0.35	-1.33	30.42	-0.04
Credit HY Europe	0.03	0.32	-2.25	31.10	-0.04

ENS = 99.87%.
RE = 0.13%.

FIGURE 18 – Caractéristiques de la distribution des rendements post EP

où $r_d^{Eq.US} = 0.03\%$ et donc $r_y^{Eq.US} \approx 7.56\%$, $r_d^{Eq.EU} = 0.043\%$ et donc $r_y^{Eq.EU} \approx 10.56\%$, $r_d^{Eq.JPN} = 0.03\%$ et donc $r_y^{Eq.JPN} \approx 7.56\%$, $r_d^{Gov.US} = 0\%$ et donc $r_y^{Gov.US} \approx 0\%$ et finalement $r_d^{Gov.EU} = 0.0119\%$ et donc $r_y^{Gov.EU} \approx 3\%$. Ainsi nous avons bien que : $r_y^{Eq.US} \leq r_y^{Eq.EU} - 3\% \leq r_y^{Eq.JPN}$ et $r_y^{Gov.US} \leq r_y^{Gov.EU}$. En outre, l'entropie relative est de 0.13%, ce qui signifie que la déformation des probabilités est encore plus faible que dans le scénario précédent. Une chose intéressante à remarquer est que les contraintes d'inégalités sont satisfaites à l'égalité, en effet les probabilités ont été déformées afin que

les contraintes soient respectées 'tout pile' (les inégalités étaient inversées initialement), ainsi c'est bien l'illustration que la méthode EP déforme le moins possible les probabilités.

— Scénario 3 :

	Mean	Volatility	Skewness	Kurtosis	CVaR
Equity US	-0.04	2.01	-1.46	17.53	-0.33
Equity Europe ex. EMU	-0.03	1.38	-0.70	12.97	-0.22
Equity Japan	-0.00	1.39	-0.15	8.03	-0.18
Govies EUR	0.00	0.19	-0.13	12.47	-0.03
Govies US	0.02	0.35	0.55	8.40	-0.04
Credit IG Europe	-0.00	0.19	-2.50	18.04	-0.03
Credit IG US	0.01	0.39	-1.28	12.94	-0.05
Credit HY US	-0.02	0.47	-2.88	24.38	-0.08
Credit HY Europe	-0.02	0.48	-4.26	32.98	-0.08

ENS = 97.02%.
RE = 3.03%.

FIGURE 19 – Caractéristiques de la distribution des rendements post EP

où $\sigma_d^{Eq.US} = 2.01\%$ et donc $\sigma_y^{Eq.US} \approx 32\%$ avec une entropie relative est de 3.03%, bien plus que précédemment (voir annexe pour les probabilités).

Typiquement, après l'obtention des probabilités postérieures pour chaque vue, la suite du processus serait de fournir les résultats comme entrées d'un algorithme d'optimisation (MVO par exemple) et d'obtenir les portefeuilles optimaux qui intègrent ainsi les convictions. On pourrait également les combiner, ou bien ajouter un niveau de confiance pour les mitiger par rapport à la probabilité uniforme de départ.

5.2 Optimisation avec coûts de transaction

5.2.1 Description du problème

Si l'approche de MV a fait ses preuves, c'est qu'elle fournit un outil mathématique d'allocation. Bien qu'imparfaite, à l'instar de son énorme sensibilité aux paramètres d'entrée, nous avons vu qu'il est possible de remédier en partie à ce problème par l'approche avec rééchantillonnage. Néanmoins un autre problème surgit dans la pratique. Alors que les gestionnaires de portefeuille cherchent une allocation performante tout au long d'une fenêtre temporelle, l'approche MV fournit une allocation ponctuelle. Se pose alors la question du rebalancement, c'est-à-dire la fréquence à laquelle le gestionnaire doit rebalancer son portefeuille. En effet si les paramètres d'entrée de l'approche de MV sont à un horizon de 1 an (par exemple des rendements attendus à 1 an) : le gestionnaire doit-il calculer le portefeuille optimal, investir selon

celui-ci et attendre un an avant de refaire la même procédure? Cette question montre bien le décalage entre la pratique de la gestion en continue et la méthode MV ponctuelle. De manière intuitive, le gestionnaire pourrait décider de réaliser des prévisions à 1 mois (pour les entrées du MV) et de réajuster chaque mois. Ici encore, l'approche classique MV fournirait des allocations optimales totalement indépendantes ce qui signifie que le gérant devrait réajuster, peut être de manière brutale, ses positions. Évidemment, en pratique ces réajustements ont un coût et il s'agira pour le gérant de réaliser un arbitrage entre le gain espéré suite à la correction de sa position, et la perte due au coût du rebalancement. Ainsi, le problème est double, il nous faut trouver une procédure d'allocation multi-période qui permette de prendre en compte la constitution de l'ancien portefeuille, les potentiels gains dus à un rebalancement et finalement les coûts du rebalancement. L'approche que nous avons retenue, est celle de l'introduction des coûts de transaction dans le problème de PQ.

Reprenons ici le problème de minimisation de l'équation 6 et introduisons maintenant les coûts de transaction. On note \tilde{w} le portefeuille actuel et $\mathcal{C}(w | \tilde{w})$ le coût de rééquilibrage du portefeuille \tilde{w} vers le portefeuille w . On en déduit que le rendement net est égal au rendement brut moins les coûts de transaction :

$$R(w | \tilde{w}) = R(w) - \mathcal{C}(w | \tilde{w})$$

Il s'ensuit que :

$$\begin{aligned}\mu(w | \tilde{w}) &= \mathbb{E}[R(w | \tilde{w})] \\ N &= \mu(w) - \mu_{\mathcal{C}}(w | \tilde{w})\end{aligned}$$

et :

$$\begin{aligned}\sigma^2(w | \tilde{w}) &= \mathbb{E}[(R(w | \tilde{w}) - \mu(w | \tilde{w}))^2] \\ &= \mathbb{E}[(R(w) - \mu(w) + \mu_{\mathcal{C}}(w | \tilde{w}) - \mathcal{C}(w | \tilde{w}))^2] \\ &= \sigma^2(w) + \sigma_{\mathcal{C}}^2(w | \tilde{w}) - 2\rho_{\mathcal{C}}(w | \tilde{w})\sigma(w)\sigma_{\mathcal{C}}(w | \tilde{w})\end{aligned}$$

où $\mu_{\mathcal{C}}(w | \tilde{w}) = \mathbb{E}[\mathcal{C}(w | \tilde{w})]$ est le coût attendu du rééquilibrage et $\sigma_{\mathcal{C}}(w | \tilde{w})$ est l'écart-type de $\mathcal{C}(w | \tilde{w})$. La fonction $\rho_{\mathcal{C}}(w | \tilde{w})$ est la corrélation entre le rendement brut $R(w)$ et le coût de transaction $\mathcal{C}(w | \tilde{w})$. En général, nous prendrons que $\rho_{\mathcal{C}}(w | \tilde{w}) = 0$. Il est intéressant de remarquer que les coûts de transaction ont un impact à la fois sur le rendement attendu et sur la volatilité du portefeuille. Cependant, ce n'est pas le seul effet. En effet, il faut également financer le processus de rééquilibrage puisque la richesse avant et après n'est pas la même. Par conséquent, la contrainte budgétaire devient :

$$\mathbf{1}_n^\top w + \mathcal{C}(w | \tilde{w}) = 1$$

Ici, nous sommes confrontés à un problème car la contrainte budgétaire est stochastique. C'est pourquoi les gestionnaires de portefeuille supposent que les coûts de transaction sont connus et non aléatoires. Dans ce cas, le problème d'optimisation devient :

$$\begin{aligned} w^* &= \arg \min \frac{1}{2} w^\top \Sigma w - \gamma (w^\top \mu - \mathcal{C}(w | \tilde{w})) \\ \text{s.t.} & \begin{cases} \mathbf{1}_n^\top w + \mathcal{C}(w | \tilde{w}) = 1 \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{cases} \end{aligned} \quad (12)$$

Une première idée est de considérer des coûts de transaction constants. Dans ce cas, nous avons :

$$\mathcal{C}(w | \tilde{w}) = \sum_{i=1}^n c_i \cdot |w_i - \tilde{w}_i|$$

où c_i est le coût unitaire associé à l'actif i . Une meilleure formulation consiste à distinguer les prix acheteur et vendeur. En suivant Scherer (2007) [9], nous avons :

$$\begin{aligned} \mathcal{C}(w | \tilde{w}) &= \mathcal{C}^-(w | \tilde{w}) + \mathcal{C}^+(w | \tilde{w}) \\ &= \sum_{i=1}^n c_i^- \cdot \max(\tilde{w}_i - w_i, 0) + \sum_{i=1}^n c_i^+ \cdot \max(w_i - \tilde{w}_i, 0) \end{aligned}$$

où c_i^- et c_i^+ sont les coûts unitaires de transaction à l'achat et à la vente. On en déduit que le coût de transaction de l'actif i satisfait à :

$$\mathcal{C}_i(w | \tilde{w}) = \begin{cases} c_i^- \cdot (\tilde{w}_i - w_i) & \text{if } w_i < \tilde{w}_i \\ 0 & \text{if } w_i = \tilde{w}_i \\ c_i^+ \cdot (w_i - \tilde{w}_i) & \text{if } w_i > \tilde{w}_i \end{cases}$$

Dans cette approche, le coût unitaire de transaction est fixe et ne dépend pas du poids de rééquilibrage :

$$c_i(w | \tilde{w}) = \frac{\mathcal{C}_i(w | \tilde{w})}{|w_i - \tilde{w}_i|} = \begin{cases} c_i^- & \text{if } w_i < \tilde{w}_i \\ 0 & \text{if } w_i = \tilde{w}_i \\ c_i^+ & \text{if } w_i > \tilde{w}_i \end{cases}$$

Deux possibilités s'offrent à nous, considérer que le coût unitaire de transaction est une fonction linéaire du poids de rééquilibrage, ou bien un fonction quadratique. Voyons tout d'abord le cas

linéaire :

$$c_i(w | \tilde{w}) = \begin{cases} c_i^- + \delta_i^- \cdot (\tilde{w}_i - w_i) & \text{if } w_i < \tilde{w}_i \\ 0 & \text{if } w_i = \tilde{w}_i \\ c_i^+ + \delta_i^+ \cdot (w_i - \tilde{w}_i) & \text{if } w_i > \tilde{w}_i \end{cases}$$

Il s'ensuit que pour le cas quadratique :

$$\mathcal{C}_i(w | \tilde{w}) = \begin{cases} c_i^- \cdot (\tilde{w}_i - w_i) + \delta_i^- \cdot (\tilde{w}_i - w_i)^2 & \text{if } w_i < \tilde{w}_i \\ 0 & \text{if } w_i = \tilde{w}_i \\ c_i^+ \cdot (w_i - \tilde{w}_i) + \delta_i^+ \cdot (w_i - \tilde{w}_i)^2 & \text{if } w_i > \tilde{w}_i \end{cases}$$

Par construction, le choix d'une spécification, ou de l'autre, aura un impact sur l'optimisation du portefeuille, surtout si le rééquilibrage est important. Dans la section suivante nous allons développer la méthodologie suivie pour implémenter le problème avec coûts de transaction linéaires.

5.2.2 Méthodologie adoptée

Comme $\mathcal{C}(w | \tilde{w})$ est une fonction non linéaire de w , le problème (12) n'est pas un problème de PQ standard. C'est pourquoi Scherer (2007) [9] a suggéré de réécrire les coûts de transaction comme suit :

$$\mathcal{C}(w | \tilde{w}) = c_i^- \cdot \Delta w_i^- + c_i^+ \cdot \Delta w_i^+$$

où $\Delta w_i^- = \max(\tilde{w}_i - w_i, 0)$ et $\Delta w_i^+ = \max(w_i - \tilde{w}_i, 0)$ représentent la vente et l'achat de l'actif i . Par définition, nous avons $\Delta w_i^- \cdot \Delta w_i^+ = 0$ et :

$$w_i = \tilde{w}_i + \Delta w_i^+ - \Delta w_i^-$$

Alors le problème (12) devient :

$$w^* = \arg \min \frac{1}{2} w^\top \Sigma w - \gamma \left(\sum_{i=1}^n w_i \cdot \mu_i - \sum_{i=1}^n \Delta w_i^- \cdot c_i^- - \sum_{i=1}^n \Delta w_i^+ \cdot c_i^+ \right)$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^n w_i + \sum_{i=1}^n \Delta w_i^- \cdot c_i^- + \sum_{i=1}^n \Delta w_i^+ \cdot c_i^+ = 1 \\ w_i + \Delta w_i^- - \Delta w_i^+ = \tilde{w}_i \\ \mathbf{0}_n \leq w \leq \mathbf{1}_n \end{cases}$$

On a maintenant un problème QP standard par rapport aux variables $x = (w, \Delta w^-, \Delta w^+)$. En effet, nous avons :

$$x^* = \arg \min \frac{1}{2} x^\top Q x - x^\top R$$

$$\text{s.t.} \begin{cases} Ax = B \\ x^- \leq x \leq x^+ \end{cases}$$

où :

$$Q = \begin{pmatrix} \Sigma & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} \\ \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} \\ \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} \end{pmatrix}$$

et :

$$R = \gamma \begin{pmatrix} \mu \\ -c^- \\ -c^+ \end{pmatrix}$$

Pour la contrainte d'égalité, on obtient :

$$A = \begin{pmatrix} \mathbf{1}_n^\top & (c^-)^\top & (c^+)^\top \\ I_n & I_n & -I_n \end{pmatrix}$$

et :

$$B = \begin{pmatrix} 1 \\ \tilde{w} \end{pmatrix}$$

Et pour les bornes, on constate que :

$$\begin{aligned} 0 \leq w_i \leq 1 &\Leftrightarrow 0 \leq \tilde{w}_i + \Delta w_i^+ - \Delta w_i^- \leq 1 \\ &\Leftrightarrow -\tilde{w}_i \leq \Delta w_i^+ - \Delta w_i^- \leq 1 - \tilde{w}_i \\ &\Leftrightarrow \begin{cases} -\tilde{w}_i \leq \Delta w_i^+ \leq 1 - \tilde{w}_i & \text{if } \Delta w_i^- = 0 \\ \tilde{w}_i - 1 \leq \Delta w_i^- \leq \tilde{w}_i & \text{if } \Delta w_i^+ = 0 \end{cases} \end{aligned}$$

Cependant, l'on sait que $\Delta w_i^- \geq 0$ et $\Delta w_i^+ \geq 0$. D'où l'on déduit que $x^- = 0_{3n}$ et :

$$x^+ = \begin{pmatrix} \mathbf{1}_n \\ \tilde{w} \\ \mathbf{1}_n - \tilde{w} \end{pmatrix}$$

Le problème obtenu est appelé un problème de PQ augmenté (Roncalli, 2013) [10], car c'est une augmentation du nombre de variables afin de trouver la solution optimale w^* qui est donnée par la relation suivante :

$$w^* = \begin{pmatrix} I_n & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} \end{pmatrix} x^* \quad (13)$$

Finalement, sous cette formulation il est possible de résoudre le problème à l'aide du solveur quadprog présenté précédemment. C'est ce que nous présentons ci-dessous. Le cas présenté ci-dessous correspond à un cas avec 7 actifs dont les rendements attendus et les volatilités sont :

i	1	2	3	4	5	6	7
μ_i	1.00	2.00	3.00	4.00	5.00	7.50	10.00
σ_i	1.00	2.00	3.00	4.00	5.00	7.50	10.00

FIGURE 20 – Rendements attendus et volatilités des actifs en %

La corrélation entre les actifs est fixée à 25%, et nous prenons un portefeuille initial équilibré (1/7 pour chaque actif). Le coût de transaction à l'achat est de 10bps, et de 20bps à la vente pour l'ensemble des actifs. Avec ces paramètres nous avons calculé la frontière efficiente traditionnelle (sans coûts de transaction) et celle qui les inclut :

Comparaison des frontières efficientes avec et sans coûts de transaction: $c^- = 0.002$ and $c^+ = 0.001$

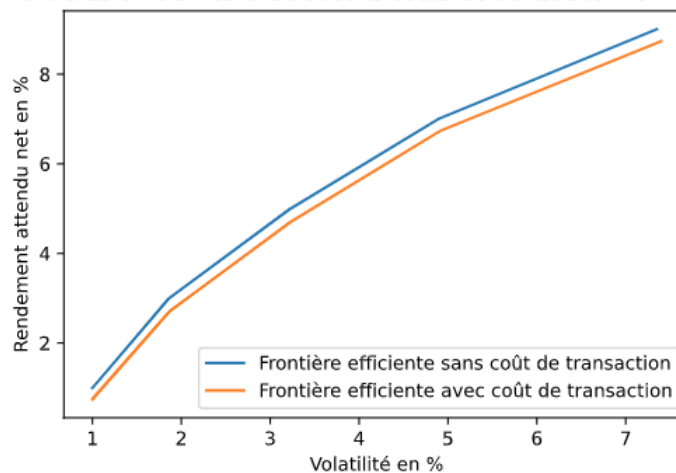


FIGURE 21 – Comparaison de la frontière efficiente avec et sans coûts de transaction

où les coûts de transaction en fonction du niveau de volatilité sont les suivants :

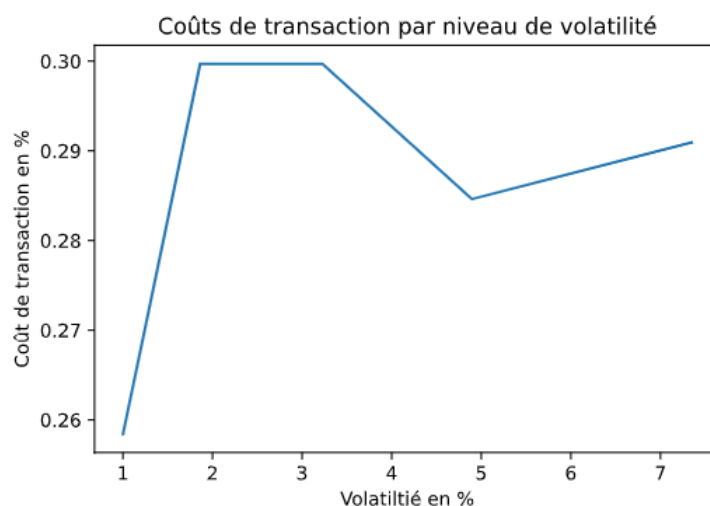


FIGURE 22 – Coûts de transaction en fonction du niveau de volatilité du portefeuille

Finalement, nous obtenons les pondérations suivantes :

	Cible rdt: 1.0%	Cible rdt: 3.0%	Cible rdt: 5.0%	Cible rdt: 7.0%	Cible rdt: 9.0%
Actif 1	99.48	24.13	11.92	-0.00	0.00
Actif 2	0.26	21.42	12.60	0.00	0.00
Actif 3	-0.00	18.70	13.26	6.64	0.00
Actif 4	-0.00	15.96	13.90	13.32	0.00
Actif 5	-0.00	13.21	14.53	19.42	0.00
Actif 6	-0.00	6.27	16.04	29.46	38.84
Actif 7	-0.00	0.00	17.46	30.88	60.87
CT	0.26	0.30	0.30	0.28	0.29

FIGURE 23 – Poids des portefeuilles de la frontière avec coûts de transaction

	Cible rdt: 1.0%	Cible rdt: 3.0%	Cible rdt: 5.0%	Cible rdt: 7.0%	Cible rdt: 9.0%
Actif 1	100.0	24.33	12.11	0.00	0.0
Actif 2	0.0	21.55	12.72	0.00	0.0
Actif 3	0.0	18.77	13.32	6.96	0.0
Actif 4	0.0	15.98	13.92	13.52	0.0
Actif 5	0.0	13.19	14.52	19.34	0.0
Actif 6	0.0	6.19	15.99	29.37	40.0
Actif 7	0.0	0.00	17.43	30.81	60.0

FIGURE 24 – Poids des portefeuilles de la frontière sans coûts de transaction

6 Conclusion

En définitive, ce stage chez Candriam m'a beaucoup appris, notamment sur l'optimisation quadratique : un domaine que je connaissais peu. L'aspect théorique des méthodes de résolution de ce type de problème est très intéressant, bien qu'assez technique pour quelqu'un qui débute

dans le domaine de l'optimisation. C'est en réalité l'aspect pratique qui m'a demandé le plus de temps. Tout d'abord il a fallu comparer les différents solveurs disponibles, leurs résultats, leurs robustesses et leurs temps d'exécution. Après cette comparaison, le choix s'est porté sur le solveur quadprog, pour lequel il a fallu adapter les différentes fonctions. La confrontation de cette méthode avec les exigences de la gestion d'actifs a également été riche en découvertes, comme la nécessité d'employer des méthodes de rééchantillonnage afin de stabiliser les résultats d'optimisation. Concernant le rééchantillonnage, ça a été la première fois que j'ai utilisé les méthodes de quasi Monte-Carlo. Des méthodes dont je connaissais l'existence, mais dont je n'avais pas exploré l'aspect théorique latent. L'aspect financier, avec l'intégration de vues qualitatives aux outils quantitatifs a été également très enrichissant. En effet, les méthodes bayésiennes comme Black-Litterman et l'entropy pooling, paraissent 'logiques' d'un point de vue financier, mais la théorie sous-jacente n'est pas si évidente en termes de mathématiques. Leur implémentation a aussi été complexe, surtout pour le passage des vues en entrée.

Finalement, c'est sur cet aspect, plus financier, que la majorité des pistes de développement futurs se portent. Face à l'intuition qu'un gérant de portefeuille n'investit pas 0.001% de son portefeuille dans un actif, mais qu'il doit faire face à des tailles minimales d'investissement, nous avons pensé à enrichir l'environnement avec un module d'optimisation avec un seuil d'investissement minimal ('buy-in-treshold' dans la littérature). Concrètement, cela consiste à dire que le gérant investit soit rien du tout, soit une valeur supérieure au seuil. Dans le même esprit, le gérant pourrait n'investir que par incréments c'est à dire par exemple par pas de 1% (1, 2, 3% etc.). Mathématiquement, ces deux problèmes sont dans le champs de la programmation mixte en nombre entiers ('mixed integer programming' en anglais), car ils font intervenir des variables continues et entières ou binaires. Plus précisément, ce sont des problèmes de programmation mixte quadratique en nombre entiers, que nous pourrions essayer de résoudre avec une méthode de séparation et évaluation (branch-and-bound en anglais)[11].

Références

- [1] Richard O Michaud. The markowitz optimization enigma : Is ‘optimized’optimal? *Financial analysts journal*, 45(1) :31–42, 1989.
- [2] Henry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1) :77–91, 1952.
- [3] Donald Goldfarb and Ashok Idnani. A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical programming*, 27(1) :1–33, 1983.
- [4] Richard O Michaud and Robert Michaud. Estimation error and portfolio optimization : a resampling solution. *Available at SSRN 2658657*, 2007.
- [5] Henri Faure. Discrépances de suites associées à un système de numération (en dimension un). *Bulletin de la Société Mathématique de France*, 109 :143–182, 1981.
- [6] Fischer Black and Robert Litterman. Global portfolio optimization. *Financial analysts journal*, 48(5) :28–43, 1992.
- [7] Attilio Meucci. The black-litterman approach : Original model and extensions. *Shorter version in, THE ENCYCLOPEDIA OF QUANTITATIVE FINANCE*, Wiley, 2010.
- [8] Ariel Caticha and Adom Giffin. Updating probabilities. In *AIP Conference Proceedings*, volume 872, pages 31–42. American Institute of Physics, 2006.
- [9] Bernd Scherer. Portfolio construction and risk budgeting. 2007.
- [10] Thierry Roncalli. *Introduction to risk parity and budgeting*. CRC Press, 2013.
- [11] Pierre Bonami, Lorenz T Biegler, Andrew R Conn, Gérard Cornuéjols, Ignacio E Grossmann, Carl D Laird, Jon Lee, Andrea Lodi, François Margot, Nicolas Sawaya, et al. An algorithmic framework for convex mixed integer nonlinear programs. *Discrete optimization*, 5(2) :186–204, 2008.

A Impacts de l'entropy pooling sur les probabilités

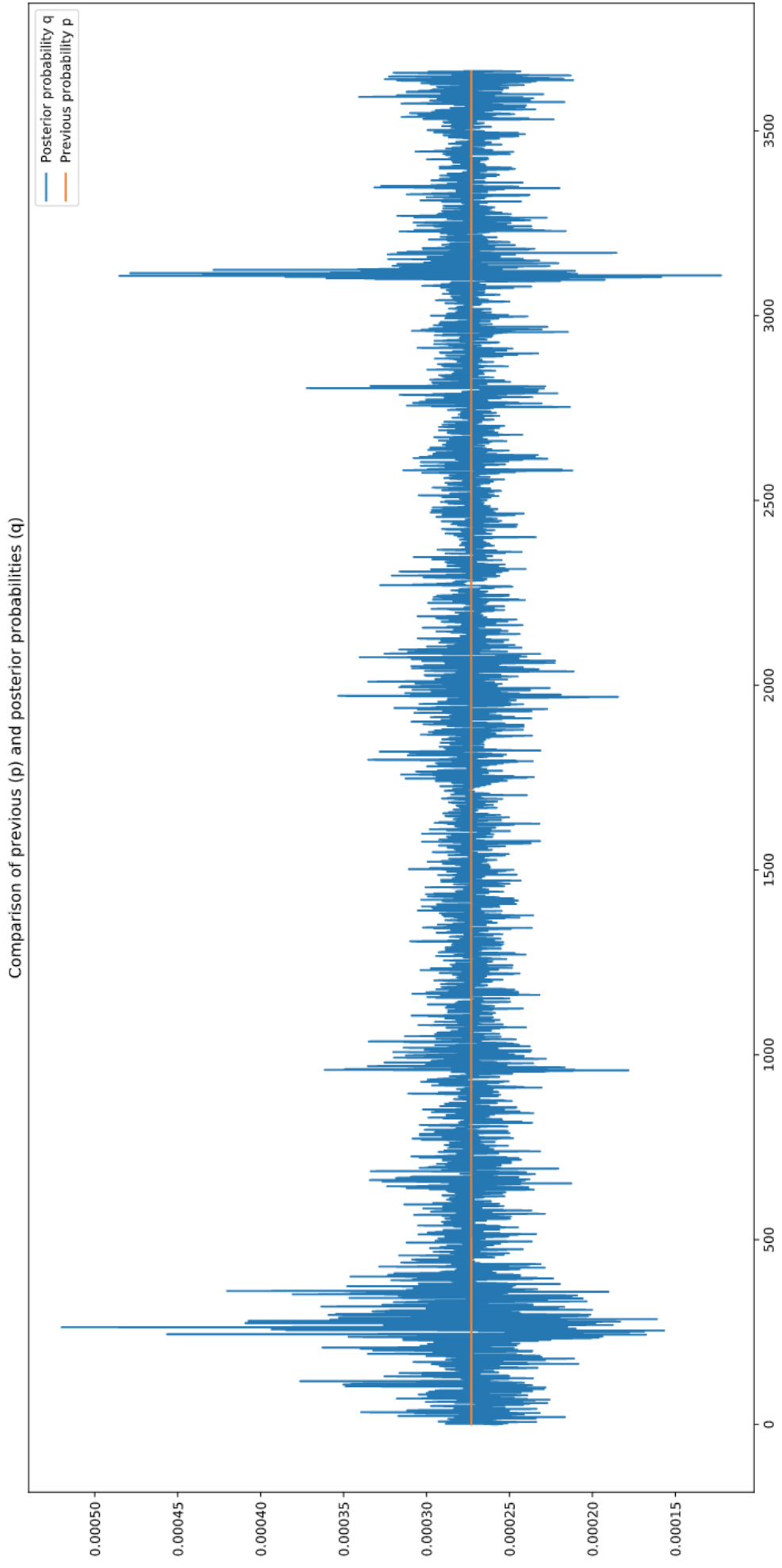


FIGURE 25 – Impact de l'EP sur les probabilités - Scénario 1

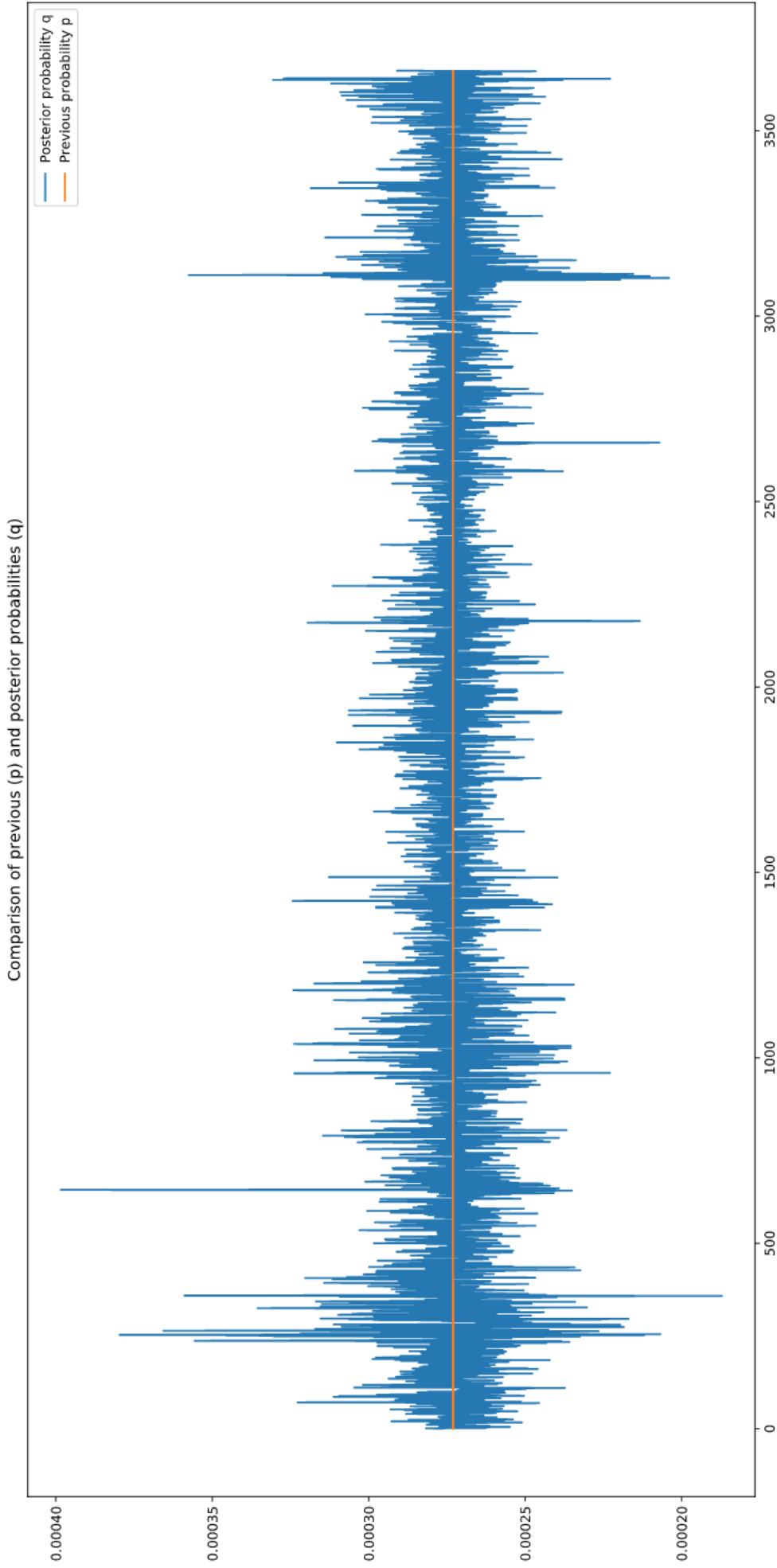


FIGURE 26 – Impact de l'EP sur les probabilités - Scénario 2

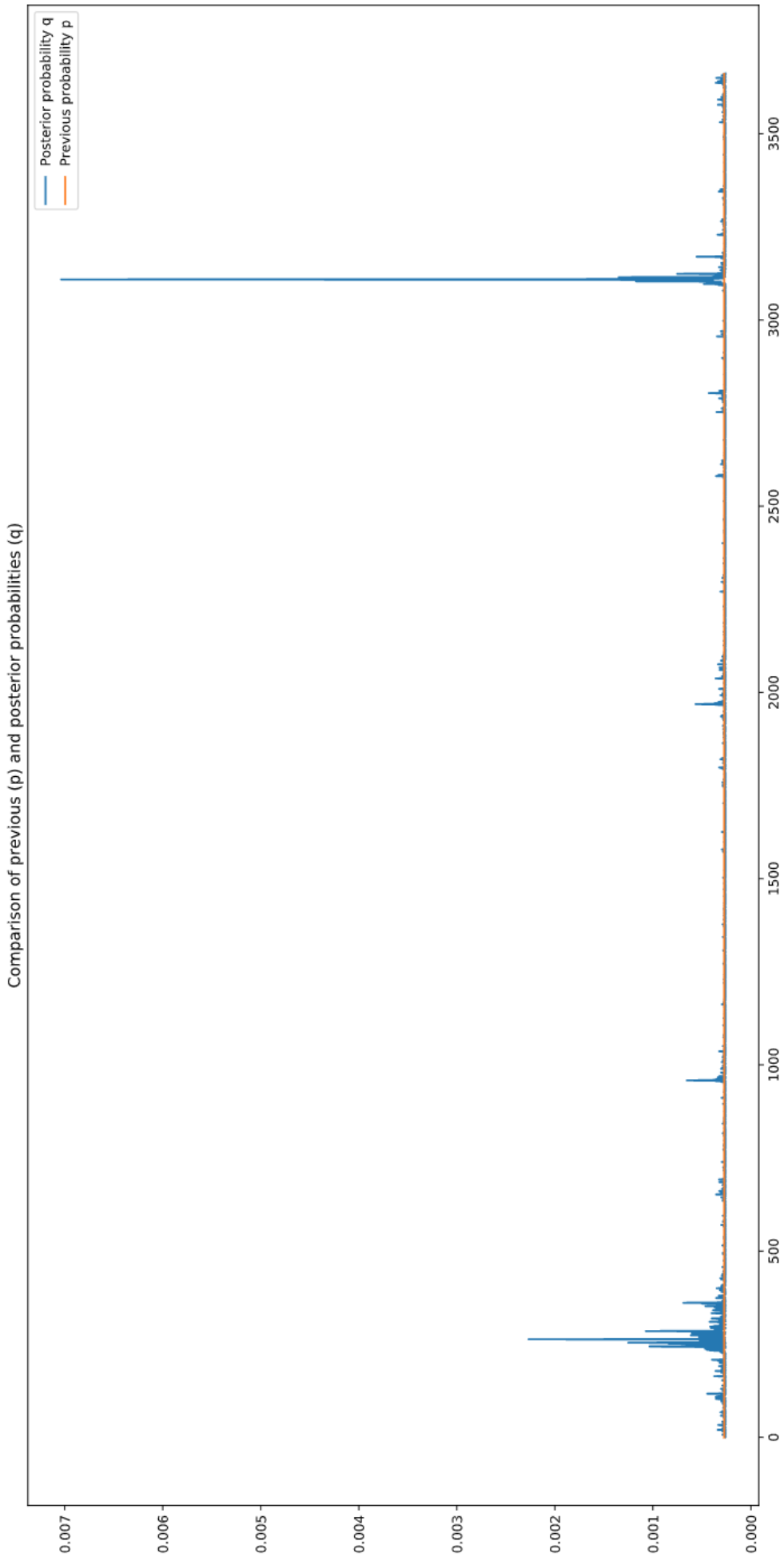


FIGURE 27 – Impact de l'EP sur les probabilités - Scénario 3