

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE  
Université Mohamed Khider-Biskra



FACULTE DES SCIENCES EXACTES ET SCIENCES DE LA NATURE ET DE LA VIE

## THESE

Présentée pour l'obtention du diplôme

## DOCTORAT EN SCIENCES

En : Mathématiques

Par :

**Nacer RAHMANI**

Thème

---

# Application des algorithmes évolutionnaires au calcul numérique et au finance

---

Devant le jury composé de :

**Mokhtari Zouhir**, *Prof, Université de Biskra*,..... **Président**

**Khelil Naceur**, *Prof, Université de Biskra*,..... **Rapporteur**

**Merad Ahcene**, *MCA, Université de Oum El-Bouaghi*,..... **Examineur**

**Rezzoug Imad**, *MCA, Université de Oum El-Bouaghi*,..... **Examineur**

**2020/2021**

# Dédicace

Je rends hommage a ma mère.

A la mémoire de mon père.

A mes frères HOCINE, Hafid, Adam et mes soeurs Hayette, Iness, Rahma.

A tous ceux et celles que j'aime et qui me sont chers.

A tous ceux qui ont participé de près ou de loin à la réalisation de ce travail.

A tous ceux que je ne nomme pas, mais qui se reconnaîtront.

A ma femme IMENE Ben messai et ma fille NADINE, les plus grandes sources de mon bonheur pour leur amour inestimable, leurs sacrifices, leur confiance, leur soutien et toutes les valeurs qu'ils ont su m'inculquer.

*Nacer Rahmani*

# Remerciements

C'est une vieille tradition que de remercier au début d'un tel travail tous ceux qui, de loin ou de près, directement ou indirectement, ont contribué à le rendre possible.

C'est avec mon enthousiasme le plus vif et le plus sincère que je voudrais rendre mérite à tous ceux qui à leur manière m'ont aidé à mener à bien cette thèse.

Je tiens tout d'abord à remercier mon directeur de thèse Monsieur *KHELIL Naceur*, Professeur d'enseignement supérieur à Université Mohamed Khider Biskra, qui a su me communiquer toutes ces années son enthousiasme et sa motivation.

Ses commentaires et suggestions tout au long de ma thèse ont considérablement amélioré à la fois le contenu et la présentation de cette dernière.

Avant de rentrer dans le vif du sujet, nous tenons à remercier Monsieur le Professeur *Ahmed BOUTERFAIA*, pour ses conseils et ses commentaires précieux, qui nous ont permis de surmonter nos difficultés et de progresser pour accomplir notre mission comme il le faut.

Je remercie également Monsieur *MOKHTARI Zouhir*, Professeur d'enseignement supérieur à l'Université Mohamed khider- BISKRA de l'honneur qui'il m'a fait d'avoir accepté de présider mon jury de thèse.

Je remercie aussi Monsieur *MERAD Ahcene* et Monsieur *REZZOUG Imad*, Professeurs à l'Université de Oum El Bouaghi.d'avoir accepté d'être membres de jury.

Je remercie tous mes collègues et mes amis. Je remercie particulièrement *Imad eddine LAKHDARI* et *Mokhtar HAFAYED* , leur support moral et intellectuel tout au long de notre travail, je remercie tous ceux qui m'ont aidé de près ou de loin et tous ceux qui m'ont motive même inconsciemment.

Sur le plan personnel, je tiens à exprimer mes remerciements et ma gratitude à tous les membres de la famille qui m'ont soutenue et encouragée sur tous les plans. Que tous ceux que je n'ai pas cités, et ils sont certainement nombreux, se sentent associés à ces remerciements. Je tiens à remercier mes parents de leurs encouragements, ma femme pour son aide précieuse et ses conseils lors de la préparation de la soutenance.

# Table des matières

<b>Table des matières</b>	<b>ii</b>
<b>Introduction</b>	<b>3</b>
<b>1 Théorie Moderne de Portefeuille : Cadre Conceptuel et Etat de l'Art</b>	<b>6</b>
1.1 Classes d'actifs . . . . .	6
1.1.1 Les actions . . . . .	6
1.1.2 Les obligations . . . . .	7
1.1.3 Les fonds monétaires . . . . .	7
1.1.4 L'immobilier . . . . .	7
1.2 Rendement . . . . .	7
1.2.1 Rendements simples . . . . .	8
1.2.2 Rendements logarithmiques . . . . .	10
1.3 Le risque . . . . .	10
1.3.1 Le concept de risque : . . . . .	11
1.3.2 Les types de risques . . . . .	11
1.3.3 Les mesures classiques de risque . . . . .	12
1.3.4 Caractéristiques d'une mesure de risque . . . . .	15
1.4 Portefeuille optimal : Optimisation du portefeuille . . . . .	19
1.4.1 Définition d'un portefeuille . . . . .	19
1.4.2 La diversification . . . . .	20
1.4.3 Portefeuille composé de deux actions . . . . .	20
1.4.4 Portefeuille composé de "n" actions . . . . .	21
1.4.5 Définition dans le cadre d'un investissement . . . . .	21
1.4.6 Le Rendement d'un portefeuille . . . . .	22
1.4.7 Espérance et variance d'un portefeuille . . . . .	22
1.4.8 Le Portefeuille efficient . . . . .	25
1.4.9 Le portefeuille de marché . . . . .	28
1.4.10 Le Portefeuille Tangent . . . . .	29
<b>2 Optimisation de portefeuille en finance : Modèles et méthodes</b>	<b>30</b>
2.1 Le Modèle Moyenne-Variance de Markowitz (1952) . . . . .	30
2.1.1 Hypothèses . . . . .	31

2.1.2	Le Modèle Mean–Variance de Markowitz . . . . .	31
2.2	Le modèle de marché de Sharpe(1963-1964) . . . . .	33
2.2.1	Les modèles à facteurs . . . . .	33
2.2.2	Le modèle à un facteur ou modèle de marché . . . . .	34
2.3	Le modèle simplifié de Sharpe . . . . .	38
2.3.1	Présentation mathématique du modèle . . . . .	38
2.4	Le modèle d'équilibre des actifs financiers (MEDAF) . . . . .	39
2.4.1	Définition du MEDAF . . . . .	39
2.4.2	Les principales hypothèses du MEDAF . . . . .	40
2.4.3	Le rôle de Bêta dans le MEDAF . . . . .	40
2.5	Le modèle d'évaluation par arbitrage (APT) . . . . .	41
2.6	Le modèle de Markowitz et Perold (1981) . . . . .	42
2.7	Le modèle de Lai (1991) . . . . .	43
2.8	Le modèle de Konno et Yamazaki (1991) . . . . .	45
2.9	Le modèle de Speranza (1993) . . . . .	47
2.10	Le modèle Moyenne–Semi Variances de Hamza et Janssen (1995) . . . . .	48
2.11	Le modèle de YOUNG (1998) . . . . .	49
2.12	Le modèle Basée sur La valeur à risque (2000) . . . . .	51
2.13	Le modèle Basée sur La VaR Conditionnelle . . . . .	53
<b>3</b>	<b>Métaheuristiques d'optimisation : Etat de l'art</b> . . . . .	<b>55</b>
3.1	Notions de Base en Optimisation . . . . .	55
3.2	Problème D'optimisation . . . . .	58
3.3	Les méthodes de résolution de problèmes d'optimisation . . . . .	59
3.3.1	Les méthodes exactes . . . . .	59
3.3.2	Les méthodes approchées . . . . .	61
3.4	<b>Optimisation par les Algorithmes Génétiques</b> . . . . .	67
3.5	Définition . . . . .	67
3.6	Description de l'algorithme . . . . .	68
3.6.1	Principes généraux . . . . .	69
3.6.2	La différence entre les Algorithmes génétiques et algorithmes clas- siques d'optimisation . . . . .	70
3.6.3	Fonctionnement des algorithmes génétiques . . . . .	71
3.7	Les étapes des algorithmes génétiques . . . . .	71
3.7.1	Population initiale . . . . .	72
3.7.2	Le codage . . . . .	72
3.7.3	Evaluation des individus . . . . .	74
3.7.4	Opérateurs de sélection . . . . .	74
3.7.5	Le Croisement . . . . .	77
3.7.6	La Mutations . . . . .	79
3.7.7	Critère d'arrêt . . . . .	81
3.7.8	Les avantages . . . . .	81
3.7.9	Les désavantages . . . . .	81
3.8	<b>Algorithme d'optimisation par essaim de particules (PSO)</b> . . . . .	81
3.8.1	L'origine de L'idée de L'optimisation par Essaim Particulaire . . . . .	81

3.8.2	Présentation de la Méthode . . . . .	83
3.8.3	Les Paramètres de L'algorithmes PSO . . . . .	86
<b>4</b>	<b>Algorithme d'optimisation de portefeuille d'actions à l'aide d'un AG</b>	<b>90</b>
4.1	Introduction . . . . .	90
4.2	Optimisation du portefeuille par la méthode moyenne-variance . . . . .	91
4.2.1	Le rendement espéré . . . . .	91
4.2.2	La variance d'un portefeuille . . . . .	91
4.2.3	Le modèle Moyenne-Variance . . . . .	92
4.3	Le modèle d'écart absolu moyen (MAD) . . . . .	93
4.4	Algorithme génétique . . . . .	96
4.5	Analyse en Composantes Principales (ACP) . . . . .	97
4.5.1	Théorème de l'analyse en composantes principales . . . . .	97
4.5.2	L'utilisation de l'analyse en composantes principales . . . . .	98
4.6	Minimisation du MAD à l'aide de PCA et d'algorithmes génétiques . . . . .	98
4.6.1	Problématique . . . . .	98
4.6.2	La proposition . . . . .	98
4.6.3	Pourquoi l'écart absolu moyen (MAD) . . . . .	99
4.6.4	Pourquoi l'analyse en composantes principales . . . . .	99
4.6.5	Le modèle mathématique : . . . . .	99
4.6.6	Algorithme d'optimisation MAD-AG . . . . .	101
4.6.7	Procedure d'optimisation (MAD-AG) : . . . . .	102
4.7	Optimisation de portefeuille d'actions à l'aide de l'algorithmes PSO-MAD .	108
4.7.1	Formulation du problème . . . . .	108
4.7.2	Configuration de la méthode MAD-PSO . . . . .	109
4.7.3	Algorithme d'optimisation MAD-PSO . . . . .	112
4.8	Partie pratique . . . . .	113
4.8.1	<b>Exemple 01</b> . . . . .	113
4.8.2	Exemple 02 . . . . .	117
4.8.3	Exemple 03 . . . . .	119
4.8.4	Exemple 04 . . . . .	120
4.8.5	Exemple 05 . . . . .	122
4.8.6	Exemple 06 . . . . .	124
	<b>Conclusion</b>	<b>126</b>
	<b>Bibliographie</b>	<b>128</b>

# Table des figures

1.1	Schéma des conventions de temps. . . . .	8
1.2	Frontière Efficiente. . . . .	26
3.1	Courbe représentant les optimums locaux et les optimums globaux . . . . .	57
3.2	Classification de méthodes de résolution de problèmes d'optimisation. . . . .	59
3.3	Les types des algorithmes évolutionnaires . . . . .	66
3.4	Organigramme d'un algorithme évolutionnaire . . . . .	67
3.5	Schéma d'une roulette . . . . .	76
3.6	Le croisement à un point. . . . .	78
3.7	Croisement multipoints ( $k = 2$ ) . . . . .	78
3.8	Opération de mutation . . . . .	80
3.9	Déplacement d'une particule. . . . .	84
3.10	Graphe d'influence d'un essaim de particules : (à gauche) Graphe complètement connecté, (à droite) Graphe d'information circulaire. . . . .	86
4.1	Organigramme général d'un algorithme MAD-AG. . . . .	102
4.2	Structure d'un chromosome . . . . .	102
4.3	Structure de la population initiale (exemple) . . . . .	103
4.4	Analyse en composantes principales (PCA) Classification diagram. . . . .	103
4.5	La fonction d'évaluation . . . . .	104
4.6	Opération de croisement . . . . .	106
4.7	Opérateur de mutation . . . . .	107
4.8	Exemple : Opérateur de mutation . . . . .	107
4.9	Organigramme général d'un algorithme 2 :MAD-AG. . . . .	108
4.10	Organigramme générale d'un algorithme PSO . . . . .	109
4.11	Résultats avec une population initiale de 50 avec 200 générations. . . . .	115
4.12	Résultats avec une population initiale de 50 avec 500 générations . . . . .	115
4.13	La convergence de la valeur objective. . . . .	117
4.14	Comment le risque affecte-t-il la rentabilité espéré d'un investissement. . . . .	118
4.15	Frontière efficiente pour différents capital investi. . . . .	118
4.16	La frontière efficiente pour MAD-AG et MAD standard. . . . .	119
4.17	La frontière efficiente pour MAD-AG et Markowitz. . . . .	120

4.18 Représentation graphique de la valeur du portefeuille en utilisant l'algorithme MAD-AG et les AG-Markowitz . . . . .	123
4.19 Représentation graphique du risque de portefeuille en utilisant l'algorithme et les AG-Markowitz . . . . .	123
4.20 Analyse en composantes principales (PCA) pour 70 actions . . . . .	125
4.21 Le rendement espéré de MAD-PAG et MAD-PSO . . . . .	125

# Introduction

L'une des directions les plus importantes de la finance est la théorie de la sélection de portefeuille, l'allocation des actifs d'une manière optimale dans un portefeuille est l'objectif de chaque investisseur ou société financière, alors l'optimisation de portefeuille ou le choix de portefeuille optimal d'actifs financiers est un sujet d'intérêt particulier dans la recherche en mathématiques financières.

L'optimisation du portefeuille d'investissement est le processus d'optimisation de la proportion de capital des actifs détenus pour s'adapter à diverses contraintes ; il donne le rendement le plus élevé avec le moins de risque [54], alors l'optimisation du portefeuille consiste à choisir le meilleur parmi l'ensemble d'opportunités de titres pour équilibrer l'objectif de maximiser le rendement espéré avec la contrainte de minimiser le risque.

Markowitz a été le premier à introduire la méthodologie d'optimisation de portefeuille de base en 1952. Le modèle de Markowitz ou le modèle de moyenne-variance (MV) a formulé le problème d'optimisation comme un problème de programmation quadratique dans lequel la fonction de risque était mesurée par la variance des rendements de portefeuille observés autour de leur moyenne, on a supposé que les rendements du portefeuille sont normalement distribués. Les solutions du problème de Markowitz pour différents niveaux de rendement forment la frontière efficiente qui représente le compromis optimal entre le risque et le rendement [71].

Pour le choix optimal de ce dernier, le modèle de Markowitz consiste à minimiser l'écart-type ou la variance pour un rendement donné ou de maximiser le rendement du portefeuille pour un risque donné.

Le modèle de moyenne-variance (MV) contrairement à sa réputation théorique, plusieurs critiques ont été adressées à ce modèle comme le choix de la variance en tant que mesure de risque et l'hypothèse sur le caractère quadratique de la fonction objectif et quelques difficultés lorsqu'il s'agit de traiter un grand nombre d'actifs ou de résoudre le problème de programmation quadratique et de calculer la matrice de variance-covariance (la charge de calcul).

Pour simplifier les difficultés associées au modèle de Markowitz, plusieurs auteurs ont tenté d'atténuer cette difficulté en utilisant divers schémas d'approximation [81]. De plus, plusieurs modèles alternatifs ont été développés comme alternative au modèle classique de Markowitz. Konno et Yamazaki [43] ont proposé l'écart absolu, tandis que Speranza [84] a proposé l'écart semi-absolu comme mesures du risque au lieu de la variance [34], plusieurs

modèles ont été proposés pour réduire la charge de calcul et linéariser le problème de choix optimal de portefeuille comme Sharpe, Stone, Konno et Yamazaki, Hamza et Janssen.

L'approche de l'écart absolu moyen (MAD) est proposée par Konno et Yamazaki [43] comme un modèle de programmation linéaire (PL), définissant l'écart absolu des variables aléatoires comme une mesure du risque au lieu de la variance. Le but du modèle MAD est de résoudre le problème de sélection de portefeuille, il conduit à une programmation linéaire au lieu d'une programmation quadratique. Konno et Yamazaki déclarent que le modèle MAD est comparable au modèle de Markowitz et montrent que si les rendements sont normaux multivariés, l'écart moyen absolu et l'écart type des mesures du risque sont essentiellement les mêmes [43]. Il y a très peu de travail récemment comparé les performances des modèles MAD et Markowitz.

En 2003, Mansini et al. a comparé les modèles Markowitz et MAD et a constaté que le modèle MAD fonctionne légèrement mieux que le modèle de Markowitz [56], Hoe et al. effectué une comparaison empirique pour les deux modèles MAD et Markowitz, ils constaté que les portefeuilles générés par les modèles MAD et Markowitz utilisent des actions très similaires mais ont des proportions de portefeuille différentes.

En général, les méthodes classiques de sélection du portefeuille optimal ne sont pas suffisamment efficaces. Les algorithmes heuristiques sont un moyen typique pour résoudre le problème d'efficacité. Nous nous concentrons dans ce travail sur l'algorithme génétique (AG) et l'algorithme d'optimisation par essaim de particules (PSO) à l'aide de l'analyse en Composantes Principales (ACP) pour Construire les meilleurs portefeuilles (portefeuille optimal)[69]. Nous l'avons développée dans ce travail deux modèles maximisant le rendement et minimisant le risque en même temps à travers des algorithmes dynamiques appelé MAD-AG et un autre MAD-PSO [69]. Ces modèles s'appliquent sur le portefeuille afin de déterminer les meilleurs, à travers l'optimisation, les proportions du capital investies qui rendent optimal ce dernier.

Notre travail est organisé comme suit :

Dans le premier chapitre, nous avons traité la Théorie Moderne de Portefeuille : Cadre Conceptuel et Etat de l'Art : la théorie et le concept puis nous avons présenté les classes d'actifs et la définitions des rendement et risques, en expliquant les diverses mesures classiques de risque. Nous avons défini le portefeuille optimal (Le portefeuille efficient, le portefeuille de marché et le portefeuille Tangent).

Dans le deuxième chapitre, nous avons présenté les différentes approches d'optimisation de portefeuille d'actifs financiers existants dans la littérature ainsi que quelques avantages et inconvénients de certaines d'entre elles.

Au niveau du troisième, nous avons défini les méthodes de résolution de problèmes d'optimisation (les méthodes exactes et les méthodes approchées). Aussi nous avons défini les algorithmes génétiques et l'algorithme d'optimisation par essaim de particules (PSO).

Dans le dernier chapitre, nous avons défini deux algorithmes d'optimisations des portefeuilles. Le premier MAD-AG, un algorithme de minimisation du MAD à l'aide de PCA et d'algorithmes génétiques, et le deuxième MAD-PSO un algorithme d'optimisation de portefeuille d'actions à l'aide d'un algorithme d'optimisation par essaim de particules (PSO) et l'écart absolu moyen (MAD).

Enfin, nous présentons nos conclusions ainsi que des perspectives.

# Théorie Moderne de Portefeuille : Cadre Conceptuel et Etat de l'Art

Les investisseurs ont pour objectif de générer des profits. Ceux qui opèrent sur le marché des actions se sont donc intéressés en premier lieu au cours des actions ou à leurs rendements. Les professionnels et les chercheurs ont analysé les séries des prix ou celles des rendements selon deux axes : l'un concerne les facteurs influençant le cours des actions et donc les rendements, et l'autre les caractéristiques des séries des rendements[17].

L'objectif de ce premier chapitre était de mettre en lumière la gestion de portefeuille, plus particulièrement aux différentes mesures de risque utilisées dans la littérature financière. La gestion de portefeuille consiste à gérer les capitaux confiés dans le respect des contraintes réglementaires et contractuelles et appliquant les politiques d'investissements définies en interne, pour en tirer le meilleur rendement possible en fonction du risque choisi, nous allons d'abord définir la rentabilité et le risque. Ensuite, la théorie d'utilité, la notion de dominance stochastique et la notion de cohérence seront présentées. Ces notions seront utilisées pour comparer des différentes mesures de risque que nous allons présenter par la suite [64].

## 1.1 Classes d'actifs

Il existe plusieurs catégories d'actifs, certains plus risqués d'autres moins, et il est nécessaire de connaître le couple rendement/risque que propose chaque catégorie. Nous retrouvons ci-après les quatre classes traditionnelles [14].

### 1.1.1 Les actions

Une action est un titre de propriété sur une fraction du capital qu'une entreprise décide de vendre aux investisseurs. L'action est l'actif le plus négocié sur les marchés financiers [14].

Une action est un titre de participation dans le capital social de son émetteur (société de capitaux) . Dans sa forme traditionnelle elle donne droit à la gestion de la société (une

action = une voix dans les votes en assemblée générale), et aux bénéficiaires (sous forme de dividendes) [8].

### 1.1.2 Les obligations

A l'inverse de l'action qui est un titre de propriété, une obligation est un titre d'endettement alors est un titre de créance. En effet, ce titre représente la part d'un emprunt long terme. L'émetteur de l'obligation s'engage à rembourser à terme son porteur et à lui verser des intérêts à des dates spécifiées. Elle peuvent être émises par l'état, les collectivités locales, les établissements publics, les intermédiaires financiers et les sociétés, ou encore par des entreprises privées, en fixant à l'avance la durée du prêt et ses modalités de rémunération. L'obligation ne donne aucun droit spécifique sur les bénéfices de l'entreprise mais son remboursement est prioritaire en cas de faillite. Les obligations sont généralement considérées comme un investissement plus sûr que les actions dans la mesure où le coupon et le taux d'intérêts sont connus à l'avance [8, 14].

### 1.1.3 Les fonds monétaires

Les fonds monétaires procurent aux investisseurs un revenu régulier et sont générés de manière à conserver une valeur stable.

En effet, ils sont considérés comme des investissements prudents. Ils offrent une performance certes minimale, mais restent très liquides et protègent le capital. C'est sans doute les placements les plus sécurisés et les moins risqués vue la disponibilité immédiate des liquidités en cas de besoin [14].

### 1.1.4 L'immobilier

L'immobilier est aussi un produit pris en compte dans la gestion d'actifs. Il reste quand même une classe d'actifs différente des classes d'actifs traditionnelles (actions, obligations) dans le sens qu'il est éminemment attaché à l'économie réelle. L'investissement immobilier permet de profiter d'un revenu sous forme de loyer. En effet, il consiste en l'achat de logements ou de bâtiments en vue d'augmenter les profits grâce aux loyers ou aux plus-values réalisées lors des reventes [14].

## 1.2 Rendement

La valeur d'un actif est modélisé par une variable aléatoire. Toute fonction de variables aléatoires est aussi une variables aléatoire. Le rendement d'un actif et celui d'un portefeuille d'actifs sont donc aussi des variables aléatoires [10].

La rentabilité est une notion fondamentale en finance et elle apparait dans l'expression de la plupart des modèles de gestion de portefeuille [64].

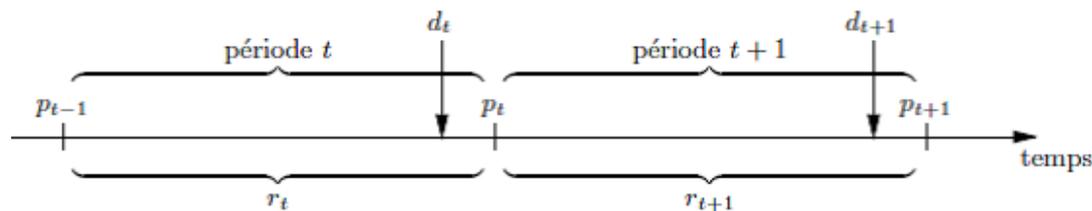


FIGURE 1.1: Schéma des conventions de temps.

### 1.2.1 Rendements simples

Nous ne considérons que le scénario en temps discret, dans lequel s'écoule une période (par exemple, une semaine ou un mois) entre les temps  $t$  et  $t + 1$ ,  $t \geq 0$  et entier. Par convention, la période  $t$  est celle écoulée entre les temps  $t - 1$  et  $t$ .

Soit  $\{P_t\}$ ,  $P_t \geq 0$ , le processus stochastique des prix d'un actif, pour un  $t$ .

Si  $P_t$  est le prix d'un actif à la période  $t$  et que  $P_{t-1}$  est le prix de cet actif à la période  $t - 1$ , son rendement simple entre  $t - 1$  et  $t$  est

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

**Remarque 1.1** *l'indice  $t$  identifie la période où le rendement est connu*

- Les rendements simples sont utiles pour les données en coupe transversale puisque le rendement d'un portefeuille est la moyenne pondérée des rendements des actifs qui le constituent
- Plusieurs actifs financiers génèrent des flux monétaires. Les actions, par exemple, paient parfois un dividende à une fréquence plus ou moins régulière. Si un actif paie un dividende  $D_t$  à la période  $t$ , son rendement simple est

$$r_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}}$$

- Certains analysts négligent les dividendes. Ils sous-estiment donc le rendement des titres qui paient des dividendes. Dans des analyses comparatives, les titres qui paient de gros dividendes sont désavantagés. Certaines sources de données (Yahoo! Finance, par exemple) ajustent les prix des actions historiques. Lorsqu'un dividende est payé, les prix historiques sont réduits du montant du dividende.

$$P_{t-1}^{ajust} = P_{t-1} - D_{t-1}$$

- D'autres événements de marché (fusions, acquisition, divisions, dividendes en actions, etc.) peuvent affecter l'interprétation du prix d'un actif financier.
- On utilise donc les rendements simple lorsqu'on considère plusieurs actifs et une seule période.

De nombreux modèles financiers utilisent des séries historiques de cours boursiers pour estimer les propriétés stochastiques des rentabilités correspondantes comme leur rentabilité moyenne. Plusieurs méthodes sont disponibles. Celle que nous avons adoptée consiste en le calcul des rentabilités de sous-périodes, ensuite nous utilisons la formule suivante

$$\bar{r} = \left(\frac{1}{T}\right) \sum_{t=1}^T r_t.$$

Nous considérons, en général, que les rentabilités des cours boursiers possèdent des densités de probabilités normales et identiquement distribuées. Soit 'a' un actif avec la rentabilité aléatoire  $r$  qui possède une telle distribution avec la moyenne  $\mu$  et l' écart-type  $\sigma$  c'est-à-dire  $r \sim N(\mu; \sigma)$ . Dans ce cas, la densité de probabilité de la variable aléatoire  $r$  s'écrit

$$f(r) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{r - \mu}{\sigma} \right)^2 \right].$$

L'hypothèse selon laquelle les rentabilités des actifs financiers sont normalement distribuées est souvent faite dans la littérature financière. Mais ceci ne correspond pas à la réalité. En général, les rentabilités sur les marchés ne suivent pas de loi gaussienne. Dans ce cas deux autres propriétés stochastiques entrent en jeu, ce sont la skewness et la kurtosis.

La skewness (ou coefficient d'asymétrie) d'une variable aléatoire se définit par :

$$Sk = \frac{E[r - E(r)]^3}{\sigma(r)^3}$$

Pour calculer la coefficient d'asymétrie d'une série de  $m$  observations, on utilise la formule :

$$Sk = \frac{\left(\frac{1}{m-1}\right) \sum_{t=1}^m [r_t - \mu]^3}{\hat{\sigma}(r)^3}$$

Pour une distribution normale,  $Sk$  est égale à zéro

La kurtosis (ou indice d'aplatissement) d'une variable aléatoire est :

$$K = \frac{E[r - E(r)]^4}{\sigma(r)^4}$$

et de la même manière que la coefficient d'asymétrie, l'indice d'aplatissement d'une série de  $m$  observations se calcule en utilisant la relation [64]

$$K = \frac{\left(\frac{1}{m-1}\right) \sum_{t=1}^m [r_t - \mu]^4}{\hat{\sigma}(r)^4}$$

Pour une distribution normale  $K = 3$ .

## 1.2.2 Rendements logarithmiques

Lorsqu'on utilise des séries temporelles, on préfère souvent calculer des rendements continus. Le rendement continu entre  $t - 1$  et  $t$  est défini par :

$$\begin{aligned} P_t &= P_{t-1} \exp(r_t) \\ r_t &= \ln\left(\frac{P_t}{P_{t-1}}\right) \\ &= P_t - P_{t-1} \end{aligned}$$

où  $P_t = \ln(P_t)$ . Ici encore, on prendra soin d'utiliser des prix ajustés pour les paiements de dividendes[10].

Les rendements continus présentent un avantage important : la composition des rendements est additive :

$$\begin{aligned} r^a &= P_t - P_{t-12} \\ &= (P_t - P_{t-1}) + (P_{t-1} - P_{t-2}) + \dots + (P_{t-11} - P_{t-12}) \\ &= r_t + r_{t-1} + \dots + r_{t-11}. \end{aligned}$$

On utilise donc les rendements simple lorsqu'on considère un seul actif et plusieurs périodes.

## 1.3 Le risque

Il est très difficile de définir de façon générale la notion de risque. Le risque est lié à la survenance d'un événement que l'on ne peut prévoir qui a des conséquences importantes sur le bilan de la banque. Il faut donc distinguer le caractère aléatoire et imprévisible (qui est l'origine du risque) de l'enjeu (conséquence finale) [75].

Le risque est défini de différentes manières. L'une d'elles que nous allons adopter est celle de ATHEARN et CROWE dont le risque est défini comme étant lié à l'incertitude et d'autre part, causé par les écarts non attendus des résultats par rapport à l'objectif attendu. C'est-à-dire que nous fixons l'objectif et nous comparons les réalisations des rentabilités des actifs par rapport à cet objectif. La définition des réalisations qu'elles soient attendues ou non attendues dépend de l'investisseur et/ou du modèle. D'après cette définition, le risque porte sur deux aspects importants. D'abord, le risque est de nature incertaine et de plus, il n'est pas attendu. Cette définition distingue les réalisations qui portent sur des profits et celles qui indiquent les pertes. Parfois les réalisations qui sont les plus proches de l'objectif fixé en amont sont considérées comme les réalisations attendues et parfois nous souhaitons seulement les réalisations qui sont supérieures à un objectif précis [64].

Dés que nous parlons de risque, nous pouvons distinguer deux types d'actifs financiers. Les actifs sans risques et les actifs risqués. Un actif sans risque est celui qui atteint le résultat,  $x$ , avec certitude, c'est-à-dire avec la probabilité de 1, ou  $p(x) = 1$  où  $p$  est la fonction de probabilité. A l'opposé, nous avons l'actif risqué pour un tel actif, il y a

plusieurs résultats possibles, comme  $x_1, x_2, \dots, x_n$ , parmi les quels il y a au moins un  $x_i$  avec  $0 < p(x_i) < 1$ .

l'investisseur ne peut pas négliger l'influence des décisions des autres investisseurs sur les siennes. En général, un investisseur peut avoir trois attitudes différentes :

1. **risque-averse (riscophobe)** : l'investisseur qui évite le risque,
2. **risque-neutre** : l'investisseur qui est indifférent à la prise de risque,
3. **riscophile** : l'investisseur qui n'hésite pas à prendre le risque.

Alors chaque investisseur éprouve une certaine aversion au risque.

### 1.3.1 Le concept de risque :

La théorie du risque a été développée par Frank Knight qui s'est intéressé à démontrer qu'il existe deux types de risques. Le premier type de risque est probabiliste et il peut être assuré tandis que le deuxième type de risque est celui de risque d'entreprise qui est non assuré car lié à la politique de cette dernière et non pas aux aléas.

- Le risque est un événement préjudiciable et aléatoire qui ne répond à aucun facteur déterminé. Il correspond au hasard et non à l'incertitude.
- Le risque est défini par l'imprécision au niveau de sa survenance, sa réalisation, la date de sa réalisation et son montant.
- La mesure de risque se base sur l'analyse de la probabilité, de survenance d'un événement et de son estimation.
- La gestion des risques au sein des entreprises a suscité ces dernières années un intérêt croissant ce qui a conduit à des investissements importants pour le développement de systèmes efficaces et la mise en oeuvre une série d'outils de gestion de risque.
- La gestion des risques joue un rôle très important dans la stabilité financière des entreprises. En effet, plusieurs entreprises ont connu des pertes financières importantes, ou même des faillites à cause d'une mauvaise maîtrise des risques.
- La gestion de risques joue un rôle important dans la stabilité financière des entreprises. Plusieurs entreprises ont connu des pertes financières importantes, ou même des faillites à cause d'une mauvaise maîtrise des risques.
- La mesure de risque se base sur l'estimation probabiliste (une approche probabiliste).

### 1.3.2 Les types de risques

#### 1.3.2.1 Les risques non quantifiables :

Sont des risques non mesurables peuvent engendrer des pertes financières importantes (le risque opérationnel, le risque médiatique, le risque légal) sont parmi les principaux risques appartenant à cette famille.

### 1.3.2.2 Les risques quantifiables :

Sont des risques mesurables comme le risque de crédit et le risque de marché. Le risque de marché résulte des variations de plusieurs facteurs du marché comme le risque de change et le risque de taux d'intérêt. Le risque de crédit (ou risque de défaut) survient lorsqu'une contrepartie ne peut ou ne veut remplir ses obligations contractuelles.

Voici une liste non exhaustive des différents risques que peut rencontrer un établissement financier :

- **Risque de crédit**
  - Risque de défaillance (default risk)
  - Risque de dégradation de la valeur de la créance (downgrading risk)
- **Risque de marché**
  - Risque de taux d'intérêt
  - Risque de change
  - Risque de modèle
- **Risque opérationnel**
  - Risque de désastre
  - Risque de fraude
  - Risque de traitement
  - Risque technologique
  - Risque juridique
- **Risque de liquidité**
- Risque stratégique

### 1.3.3 Les mesures classiques de risque

Dans cette section nous allons aborder la quantification de risque. Pour cela nous allons exposer certaines mesures de risque.

Nous rappelons ci-dessous quelques mesures de risque classiques.

On s'intéresse d'abord aux mesures de risque qui permettent de comparer des actifs ayant le même rendement espéré [10].

#### 1.3.3.1 La variance

Une mesure classique de risque est la variance,  $Var$ , et sa racine carrée dite écart-type,  $\sigma$ . Il est bien connu que Markowitz est la première personne qui a utilisé la variance comme la mesure de risque [57, 59].

Markowitz a bien marqué le début de la théorie moderne de portefeuille ou pour la première fois, le problème de choix de portefeuille a été clairement mis au point et résolu [64].

**Définition 1.1** Soit  $f$  la fonction de densité d'une variable aléatoire  $r$ , nous définissons la variance de  $r$  par

$$Var(r) = \sigma^2(r) := \int_{-\infty}^{+\infty} (r - E(r))^2 f(r) dr$$

Où  $E$  est l'opérateur d'espérance mathématique et  $r$  est une variable continue. Pour le cas où  $r$  est une variable discrète, la variance de  $r$  se définit par

$$\text{Var}(r) = \sigma^2(r) := \sum_r (r - E(r))^2 f(r)$$

Où la somme est effectuée sur toutes les valeurs possibles de  $r$ .

Si  $r$  est le rendement d'un portefeuille, alors la variance de rendement sera le carré de l'écart type du rendement par rapport à l'espérance mathématique du rendement [64].

### 1.3.3.2 écart-type

La variance satisfait cette condition. Dans le cas où un actif est normalement distribué, elle décrit complètement son risque. Pour obtenir une mesure dans les mêmes unités que l'actifs, on considère plutôt l'écart-type,

$$\sigma(r) = \sqrt{\frac{1}{T} \sum_{t=1}^T ((r_t - \bar{r}))^2}$$

Remarquons qu'on peut calculer l'écart-type de n'importe quel actif, ce qui permet d'ordonner tous les actifs selon leur risque [10].

### 1.3.3.3 écart absolu moyen

La variance échantillonnale est la moyenne du carré des écarts à la moyenne. Une alternative consiste à considérer la moyenne de la valeur absolue des écarts à la moyenne,

$$eam = \frac{1}{T} \sum_{t=1}^T |r_t - \bar{r}|$$

Intuitivement, le risque est une propriété qu'un investisseur veut éviter, c'est quelque chose de "mauvais". Si deux actifs ont le même rendement espéré, on voudrait celui qui est le moins risqué. Différents investisseurs pourraient donc avoir des définitions différentes du risque[10].

### 1.3.3.4 Semi-écart-type

L'écart-type est parfois critiqué comme étant une mesure symétrique, qui accorde autant d'importance aux écarts positifs qu'aux écarts négatifs. Certains investisseurs définissent le risque par le semi-écart-type

$$s^-(r) = \sqrt{\frac{1}{T^-} \sum_{t=1}^{T^-} 1_{(r_t < \bar{r})} (r_t - \bar{r})^2}$$

où  $T^- = \sum_{t=1}^{T^-} 1_{(r < \bar{r})}$ , soit le nombre de rendements inférieurs à la moyenne. Ainsi, la semi-variance est un estimateur de l'espérance conditionnelle

$$\sigma^{2-}(r) = E[(r - \bar{r})^2 / (r < \bar{r})].$$

Remarquons que l'utilisation de  $T^-$  n'est pas uniforme dans la littérature. On utilise parfois  $T$ , surtout dans les ouvrages plus techniques. C'est ce que nous ferons dès maintenant [10].

### 1.3.3.5 Semi-écart-type cible

Pour certains investisseurs, le rendement moyenne n'est pas nécessairement le rendement de référence. On peut définir le semi-écart-type cible par :

$$s_{\tau}^{-}(r) = \sqrt{\frac{1}{T} \sum_{t=1}^T 1_{(r_t < \tau)} (r_t - \tau)^2}$$

où  $\tau$  est une performance minimale cible (le rendement d'un indice de référence ou le rendement sans risque, par exemple). [10].

### 1.3.3.6 Moments partiels inférieurs

La notion de semi-écart-type cible (la racine carrée de la semi-variance cible) peut être généralisée. Un moment partiel inférieur (lower partial moments) prend la forme

$$\begin{aligned} LPM_{\alpha, \tau}(R) &= \int_{-\infty}^{\tau} (R - \tau)^{\alpha} f(R) dR \\ &= E[(R_t - \tau)^{\alpha} / (R < \tau)] \Pr(R < \tau). \end{aligned}$$

qu'on estime par :

$$\widehat{LPM}_{\alpha, \tau}(R) = \frac{1}{T} \sum_{t=1}^T 1_{(R_t < \tau)} (R - \tau)^{\alpha}.$$

Cette notion est aussi appelée **downside risk** dans la littérature. Le semi-écart-type cible peut donc s'écrire [1].

$$\bar{S}_{\tau} = \sqrt{\widehat{LPM}_{2, \tau}}.$$

### 1.3.3.7 Valeur à Risque

La valeur à Risque est une mesure populaire :

$$VaR_{\alpha} = \inf_y \{F_Y(y) \geq \alpha\}$$

Où  $y = x_0 - x$ , et  $x_0$  est la valeur initiale de l'actif :  $y$  est donc une perte. Elle mesure la perte associée à un actif lorsque que le scénario correspondant au quantile  $\alpha$  se réalise. Si la variable est continue, on peut aussi l'exprimer comme la solution de [10]:

$$\Pr(Y \leq VaR_{\alpha}) = \alpha.$$

La VaR est généralement utilisée de deux façons distinctes :

- en premier lieu, elle peut servir à calculer ( à posteriori) le risque auquel a été exposé un portefeuille dans le passé; cette mesure peut servir, par exemple, à comparer la performance de différents placements.
- En second lieu, la VaR peut servir à prévoir le risque auquel sera exposé un portefeuille dans le futur. Cette prévision peut permettre de choisir lequel, d'entre deux placements, offrira le rendement espéré le plus élevé pour un niveau de risque fixé.

Ces deux utilisations de la VaR sont complémentaires. La première est pertinente à l'évaluation des performances réalisées, alors que la deuxième sert dans la constitution des stratégies de placement. Nous faisons usage de ces deux points de vue dans ce memoire [13].

### 1.3.3.8 Méthode Morningstar

Les mesures précédentes ne permettent que de comparer des actifs d'espérance de rendement égales. La société Morningstar a adopté une méthode qu'elle a appelée Morningstar Risk-Adjusted Return pour classer les actifs qui repose sur l'estimation de [1, 10] :

$$MRAR = (E [(1 + R_t)^2])^{-\frac{1}{2}} - 1.$$

### 1.3.4 Caractéristiques d'une mesure de risque

Avant d'entamer la section sur les outils de mesure de risque, il est nécessaire de savoir ce qu'est une mesure de risque et quelles sont les propriétés qu'elle doit vérifier.

Les mesures de risque ont pour objectif de quantifier les pertes engendrées par le risque selon son type. Une mesure de risque est une fonction qui fait correspondre à un risque aléatoire  $X$  un nombre positif  $\rho(X)$ . [14].

En pratique,  $X$  peut représenter une perte financière de montant  $X$  et  $\rho(X)$  le montant du capital nécessaire pour faire face à la perte  $X$ . En d'autres termes,  $\rho(X)$  représente le niveau de danger inhérent à  $X$  [14].

Dans cette section, nous allons présenter quelques notions basées sur des axiomes. Ces notions vont nous aider à bien comparer les différentes mesures de risque. Chacune d'entre elles est un ensemble de critères qu'un portefeuille doit satisfaire pour qu'il soit classifié comme celui que l'investisseur préfère [64].

Pour un investisseur, l'objectif principal est la rentabilité et la profitabilité. Alors nous allons d'abord étudier une notion qui est basée sur l'hypothèse selon laquelle l'investisseur connaît bien la loi de distribution des rendements [64].

#### 1.3.4.1 Le critère de Maximum d'Espérance d'Utilité

La théorie d'utilité fournit une façon d'exprimer la sensibilité de l'individu au risque. Pour cela, la théorie d'utilité étudie les préférences des individus et leurs représentations numériques au sein des fonctions d'utilité.

En 1947, von Neumann et Morgenstern ont mis au point un ensemble d'axiomes concernant les préférences des individus. Dans leur étude, von Neumann et Morgenstern ont

adopté une convention. Ils supposent que l'individu connaît la loi de distribution des revenus aléatoires. Ici, il est supposé que l'individu est rationnel, c'est-à-dire que son comportement est défini au moyen de cet ensemble d'axiomes.

Pour chaque individu (rationnel) on peut définir une fonction d'utilité dont l'argument est la richesse et l'individu peut utiliser l'espérance mathématique afin de classifier tous les investissements risqués. Plus précisément, une fonction d'utilité est une règle par laquelle on associe un indice numérique à chacun des investissements de sorte que les préférences de l'individu se manifestent par le fait que cet indice d'utilité est d'autant plus élevé que la préférence est grande.

Soit  $L = \{p_1 A_1, p_2 A_2, \dots, p_n A_n\}$  un investissement risqué simple ou plus simplement un investissement (risqué), où les  $A_i$  sont des  $n$  résultats possibles avec les probabilités  $p_i$ .

Nous pouvons prouver qu'en acceptant ces axiomes, le critère de (MEU) est la meilleure méthode de prise de décision et les investissements doivent être classifiés selon leur espérance d'utilité.

Qu'en acceptant ces axiomes, le critère de (MEU) est la meilleure méthode de prise de décision et les investissements doivent être classifiés selon leur espérance d'utilité.

**Définition 1.2** : *(le critère d'espérance d'utilité). Les préférences d'un individu satisfont au critère d'espérance d'utilité s'il existe une fonction croissante  $U$  appelée fonction d'utilité telle que l'individu préfère le rendement aléatoire  $w_1$  au rendement aléatoire  $w_2$  si et seulement si l'espérance d'utilité de  $w_1$  est supérieure à celle de  $w_2$  :*

$$w_1 \text{ préféré à } w_2 \iff E[U(w_1)] \geq E[U(w_2)]$$

La croissance de la fonction d'utilité exprime simplement que l'individu aime la richesse.

En utilisant la fonction d'utilité, les comportements des individus à l'égard du risque se traduisent comme suit :

1. **risque-averse (riscophobe)** : Si l'investisseur est risque-averse, alors  $U(w) > E[U(w)]$ , tel que  $\dot{U}(w) < 0$ ,
2. **risque-neutre** : Si l'investisseur est risque-neutre, alors  $U(w) = E[U(w)]$ , tel que  $\dot{U}(w) = 0$ ,
3. **riscophile** : Si l'investisseur est riscophile, alors  $U(w) < E[U(w)]$ , tel que  $\dot{U}(w) > 0$

### 1.3.4.2 Dominance stochastique

La Dominance Stochastique était la généralisation des travaux en théorie de majorisation comme [36] et [51, 76]. En suite, il a été largement utilisé en économie et en finance voir [6, 51, 68] et les références incluses .

D'après [10], Les mesures classiques de risque sont quelque peu arbitraires. Peut-on définir une mesure du risque qui soit valide pour tous les investisseurs averses au risque et qui tiennent compte de l'espérance de rendement ? Oui, mais elles ne permettent généralement pas d'ordonner tous les actifs [10].

En gestion de portefeuille, la notion de DS est utilisée pour mettre en ordre les portefeuilles et nous supposons que nous connaissons la distribution des rendement aléatoires des actifs. A la base de cette convention sur les distributions, DS met en ordre les rendements des actifs tels que les rendements dominants ont une espérance d'utilité supérieure à celles des rendements dominés. Ce raisonnement s'applique sans avoir besoin d'information sur la fonction d'utilité des investisseurs.

### **Dominance stochastique du premier ordre :**

Une manière naturelle de comparer deux variables aléatoires est de comparer leur fonction de répartition. Ce critère correspond en économie à la notion de dominance stochastique à l'ordre 1, notée DSI.

On dit qu'un actif  $x$  domine un actif  $y$  au sens de la dominance stochastique du premier ordre si

$$F_X(z) \leq F_Y(z), \text{ pour tout } z.$$

Une variables aléatoires  $x$  domine un actif  $y$  au sens de la dominance stochastique du premier ordre si il existe une variable aléatoire  $\delta$  non positive telle que

$$y = x + \delta$$

Le critère de la dominance stochastique à l'ordre 1 est un critère très fort. Car cette définition du risque permet d'ordonner peu d'actifs financiers. Tout investisseur pour qui plus de rendement est préférable sera d'accord avec cet ordre. On peut utiliser des critères de dominance stochastique à l'ordre  $k$ ,  $k > 2$ .

### **Dominance stochastique du second ordre :**

On dit qu'un actif  $x$  domine un actif  $y$  au sens de la dominance stochastique du second ordre si

$$\int_{-\infty}^z F_X(s) - F_Y(s) ds \leq 0, \text{ pour tout } z.$$

La fonction  $F(x)$  est continue, convexe, positive et croissante. Elle correspond à l'aire sous la fonction de répartition. En termes de variables aléatoires,  $x$  domine un actif  $y$  au sens de la dominance stochastique du deuxième ordre.

si il existe une variable aléatoire  $\delta$  non positive et une variable aléatoire  $\theta$  d'espérance nulle telles que

$$y = x + \delta + \theta.$$

Cette définition du risque permet d'ordonner plus d'actifs financiers que la dominance stochastique du premier ordre, mais pas tous. Tout investisseur averse au risque sera d'accord avec cet ordre [10]

### **Dominance stochastique à l'ordre k.**

Plus généralement, on peut définir les fonctions  $F_X^k$ ,  $k \geq 2$  pour des variables aléatoires appartenant à  $L^{k-1}(\Omega, F, P)$  de la manière suivante :

$$\forall \eta \in R, F_X^k(\eta) = \int_{-\infty}^{\eta} F_X^{k-1}(u) du,$$

auxquelles on associe la dominance stochastique à l'ordre  $k$ .

**Définition 1.3** On dit qu'une variable aléatoire  $X \in L^{k-1}(\Omega, \mathcal{F}, P)$  domine stochastiquement une variable aléatoire  $Y \in L^{k-1}(\Omega, \mathcal{F}, P)$  à l'ordre  $K$  si :

$$\forall \eta \in R, F_X^k(\eta) \leq F_Y^k(\eta).$$

Remarquons que la dominance stochastique à un ordre  $k$  entraîne la dominance stochastique à tous les ordres supérieurs.

### 1.3.4.3 Mesures de risque cohérentes

Dans cette partie, nous rappelons les propriétés fondamentales des mesures de risque à travers la notion de mesure de risque cohérente.[80], les mesures de risque les plus utilisées en pratique ne reflètent pas de façon correcte les préférences des investisseurs. Afin de surmonter ce problème, des mesures de risque cohérentes introduite par Artzner et al. dans [5] ont défini quatre propriétés qu'une mesure de risque doit satisfaire pour qu'elle soit cohérente.

#### Mesures de risque monétaires :

Soit  $X$  un ensemble de variables aléatoires à valeurs réelles tel que :

$$\forall G \in \chi, \forall m \in \mathbb{R}, G + m \in \chi$$

**Définition 1.4** Une mesure de risque  $R$  est dite monétaire si elle vérifie les deux propriétés suivantes :

**P1-. Monotonicité** (dominance stochastique du premier ordre) :  $\forall G_1, G_2 \in \chi, G_1 \geq G_2 \implies R(G_1) \leq R(G_2)$

**P2- Equivariance par translation** :  $\forall m \in \mathbb{R}, R(G + m) = R(G) - m$ .

— La Monotonicité exprime que si le risque d'un portefeuille est supérieur à celui d'un autre, le capital requis pour le premier portefeuille est supérieur à celui requis pour le deuxième. Donc un portefeuille rapporte toujours plus qu'un autre alors il est moins risqué.

— La propriété d'invariance par translation signifie que si l'addition (ou la soustraction) d'un montant initial sûr  $m$  au portefeuille initial dans l'actif de référence, la mesure de risque  $R$  accroît (ou décroît) la mesure du risque par  $m$ . On constate que l'addition d'un montant initial égal au  $R(G)$  réduit le risque à 0 soit  $R(G + R(G)) = R(G) - R(G) = 0$ .

#### Mesures de risque cohérentes :

Le risque auquel est soumis l'actif financier pour une période de temps est décrit par la variation de sa valeur ou son rendement dans cette période.

Dans ce cas, la mesure de risque est une fonction  $R$  faisant correspondre à un risque  $X$ , un nombre positif noté  $R(X)$  qui permet de quantifier le niveau de danger inhérent à ce risque.

Pour une mesure de risque  $R$  donnée, considérons les propriétés suivantes :

**P3. Sous-additivité** :  $\forall G_1, G_2 \in \chi, R(G_1 + G_2) \leq R(G_1) + R(G_2)$ .

La condition de sous-additivité requière que la mesure de risque tienne compte des effets de diversification. Pour qu'elle soit satisfaite, le risque d'un portefeuille doit être inférieure ou égal à la somme des risques de ses composantes.

Si nous fusionnons deux portefeuilles  $A$  et  $B$ , la mesure de risque du portefeuille résultant ne doit pas être supérieure à la somme des mesures de risque initiales de  $A$  et  $B$ .

Donc la propriété de sous-additivité garantit la diminution du risque par diversification du portefeuille.

cette propriété reflète le gain de diversification.

**P4. Positive homogénéité :** Si  $k > 0$  et Si  $G \in \chi$  alors  $R(kG) = kR(G)$

La propriété d'homogénéité positive peut être vue comme un cas limite de la sous-additivité qui représente l'absence de diversification.

$$R(G_1 + G_2 + \dots + G_k) = kR(G)$$

La multiplication de chaque risque d'un portefeuille par un scalaire positif augmente la mesure de risque par le même scalaire.

**Définition 1.5** Une mesure de risque est dite **cohérente** si elle vérifie les propriétés **PI**, **P2**, **P3** et **P4**.

**Remarques :**

1. La Value-at-Risk, bien que très largement utilisée en finance, n'est pas une mesure de risque cohérente car elle n'est pas sous-additive, de même pour la variance. Par contre la CVaR est une mesure de risque cohérente.
2. La sous-additivité et la positivement homogénéité garantissent la convexité de la mesure de risque, ce qui est un avantage en gestion de portefeuille.
3. Les mesures de risques classiques sont définies arbitrairement, mais peuvent néanmoins être utiles pour fins de communication ou pour simplifier certains calculs.
4. Finalement, il est possible de retenir le choix d'une mesure de risque selon des critères de cohérence. Ces critères nous incitent à retenir la VaR conditionnelle.

## 1.4 Portefeuille optimal : Optimisation du portefeuille

### 1.4.1 Définition d'un portefeuille

Le portefeuille boursier est la représentation de l'ensemble des actifs sur lesquels un agent économique a investi sur le marché financier. Ces actifs peuvent provenir de différentes classes : actions, obligations, produits dérivés, matières premières, fonds, cash, etc. Ce portefeuille peut être géré par classe d'actif ou par type de gestion (active, passive, etc.). Que l'investisseur soit un professionnel ou un particulier, le portefeuille est généralement constitué en référence à un couple rendement/risque. Plus le rendement du portefeuille est élevé et plus les actifs sont risqués. Pour limiter les risques. La valeur des actifs financiers va évoluer notamment en fonction de multiples paramètres qui sont le capital disponible, le temps, le niveau des taux d'intérêt réels et enfin, la croissance économique et l'inflation.

L'investisseur, pour diminuer son risque, procède souvent à une diversification de ses actifs. Le principe de la diversification consiste à trouver un bon équilibre entre les différentes catégories d'actifs financiers composant un portefeuille. Cette répartition dépend bien évidemment des souhaits et des objectifs spécifiques à chaque investisseur.

Ces derniers possèdent chacun une volatilité qui leur est propre et sont plus ou moins corrélés entre eux. La détention de plusieurs actifs différents tend donc, généralement, à diminuer la volatilité globale du portefeuille. Le degré de risque (nature des titres détenus, volatilité historique, diversification), le dynamisme de la stratégie (fréquence de réajustement de l'allocation) et le rendement obtenu sont des caractéristiques importantes, servant à comparer les portefeuilles boursiers entre eux.

Pour optimiser la valeur future d'un portefeuille, la stratégie d'investissement à mettre en œuvre reposera sur une bonne analyse fondamentale. Celle-ci servira de cadre de référence pour l'allocation stratégique des actifs. La combinaison de stratégies d'investissement ayant une prépondérance des facteurs macroéconomiques dans le processus d'investissement) et une prépondérance des facteurs microéconomiques apporte de la valeur ajoutée dans la gestion de portefeuille.

### 1.4.2 La diversification

La maximisation de rentabilité exige une gestion minutieuse des risques et cette dernière se fait à base de la diversification. La diversification stipule le mixage d'un portefeuille d'actifs entre ceux risqués ou bien les combiner avec d'autres sans risque. C'est l'investissement dans différentes classes d'actifs ou dans différents secteurs, cette diversification ne signifie pas seulement détenir beaucoup d'actifs.

Par exemple, si vous détenez 50 titres liés au secteur Informatique, votre portefeuille n'est pas diversifié. Par contre, si vous détenez 50 titres qui sont éparpillés parmi 20 différentes industries, vous êtes en possession d'un portefeuille diversifié.

Le concept de diversification est à la base de la théorie. En effet, Markowitz pense que les différents titres composant un portefeuille ne peuvent être sélectionnés individuellement et doivent au contraire être choisis selon la corrélation de leurs variations à celles du reste des actifs du portefeuille

### 1.4.3 Portefeuille composé de deux actions

On suppose qu'un actionnaire dispose de deux actions  $A$  et  $B$  dont les caractéristiques sont :

- Une rentabilité mesurée par l'espérance (de  $A$  et  $B$ ),
- Cet actionnaire investit en plaçant  $\alpha$  actions  $A$  et  $1 - \alpha$  actions  $B$ , ce qui donne le portefeuille suivant :

$$P = \alpha A + (1 - \alpha) B$$

#### 1.4.4 Portefeuille composé de “n” actions

En général, pour un portefeuille comportant  $n$  actifs

$$P = \sum_{i=1}^n A_i x_i. \quad \text{où} \quad \sum_{i=1}^n x_i = 0$$

Soit  $P$  un portefeuille d'actions  $(A_1, \dots, A_n)$  représenté par un vecteur  $x = (x_1, \dots, x_n)$  où  $x_i$  désigne la proportion du capital  $C$  investie dans l'action  $i$  a caractérisée par son rendement incertain  $r_i = (1, \dots, n)$ .

Le rendement de ce portefeuille est défini comme suit :

$$R(x) = \sum_{i=1}^n r_i x_i.$$

#### 1.4.5 Définition dans le cadre d'un investissement

Nous introduisons ici les notations dans le cadre d'un investissement statique sur une unique période de temps, où le portefeuille est constitué à l'instant initial et conserve sa composition jusqu'à l'instant final. Le gérant décide de la composition du portefeuille au début d'une période prédéfinie selon sa fréquence d'investissement (par exemple une semaine, un mois, un an, etc.) d'après ses estimations sur les caractéristiques des actifs disponibles, et ne modifie pas la quantité détenue dans chaque actif jusqu'à la fin de la période. De façon externe à la gestion du portefeuille, la valeur de chaque actif varie dans le temps. Les variations de la valeur du portefeuille entre les instants initial et final sont donc entièrement déterminées par les actifs le constituant. Une grande partie de la littérature démarre de ce cadre élémentaire. Certains travaux traitent de périodes multiples, ou encore de la gestion dite “active” où, de façon réaliste, le gérant peut décider de modifier la composition selon les circonstances au cours de la période [89].

On considère  $K$  actifs de marché ayant chacun un prix dans le temps  $S_t^1, \dots, S_t^K$ . à l'instant initial  $t = 0$ , on compose avec le capital  $C$  un portefeuille  $P$  en investissant sur les  $K$  actifs avec des poids  $w = (w^1, \dots, w^K)$ , c'est-à-dire qu'on a

$$C = \sum_{k=1}^K w^k S_0^k.,$$

et que la valeur  $P_t$  du portefeuille dans le temps est :

$$P_t = \sum_{k=1}^K w^k S_t^k,$$

On utilisera de manière plus pratique le rendement  $R_T^k$  des actifs entre l'instant initial et l'instant final  $T$ , défini par :

$$R_T^k = \frac{S_T^k - S_0^k}{S_0^k},$$

- étant sur une unique période, on notera simplement  $R = (R_T^1, \dots, R_T^K)$  le vecteur des rendements d'actifs, sans l'indice  $T$ .
- Le capital est unitaire ( $C = 1$ ), et les poids représentent des fractions de capital avec  $\sum_{k=1}^K w^k = 1$ .

### 1.4.6 Le Rendement d'un portefeuille

Soit un ensemble  $A$  de  $N$  actifs, dont les rendements simples à la période  $t$  sont donnés par  $R_t = (R_{1t}, R_{2t}, \dots, R_{Nt})^t$  (où  $t$  dénote la transposée d'une matrice ou d'un vecteur).

**Définition 1.6** *un portefeuille  $x_t$  définie par rapport l'ensemble d'actifs  $A$  est le vecteur des montants  $x_{it}$  investis dans chaque actif à un temps  $t$  donné :*

$$x_t = (x_{1t}, x_{2t}, \dots, x_{Nt})^t, \quad x_{it} \in \mathbb{R} \quad \text{et} \quad -\infty < x_{it} < +\infty$$

Le **rendement total** du portefeuille  $x_{t-1}$  pour la période  $t$  est

$$R_t(x_{t-1}) = \sum_{k=1}^K x_{k(t-1)} R_{kt} = \dot{x}_{t-1} R_t$$

Ce rendement est un rendement "brut" (exprimé par exemple en dollars)

### 1.4.7 Espérance et variance d'un portefeuille

#### 1.4.7.1 Le cas de deux actifs :

On se place dans un marché où il y a deux actifs risqués 1 et 2, on essaie de construire le portefeuille qui intéresse les investisseurs autrement dit celui qui a la rentabilité anticipée la plus élevée [20].

Un tel portefeuille peut s'écrire de cette façon :  $P_x = \{(1, x), (2, 1 - x)\}$ .

- $x$  et  $1 - x$  sont les proportions à investir du capitale de détenteur de portefeuille respectivement sur les deux actifs 1 et 2. Elles peuvent varier entre  $-1$  et  $1$ .
- Les rentabilités aléatoires notées  $R_1$  et  $R_2$ , les rentabilités espérées de chacun des deux actifs  $E(R_1)$  et  $E(R_2)$ , tel que  $E(R_1) \neq E(R_2)$ .
- Les variances de ces rentabilités sont notées  $\sigma_1^2$  et  $\sigma_2^2$ . Et  $\sigma_{12}$  la covariance entre la rentabilité de ces deux actifs.

$$\sigma_{12} = \text{cov}(R_1, R_2) = \rho_{12} \sigma_1 \sigma_2$$

Où  $\rho_{12}$  ( $\rho_{12} \in [-1, 1]$ ) désigne le coefficient de corrélation entre les rentabilités des deux actif.

Notre portefeuille a une rentabilité aléatoire  $R_p$  :

$$R_p = xR_1 + (1 - x)R_2$$

On peut même écrire l'égalité au dessus en utilisant les rentabilités esprées :

$$E[R_p] = xE[R_1] + (1 - x)E[R_2]$$

Ce qui donne l'expression suivante :

$$x = \frac{E(R_p) - E(R_2)}{E(R_1) - E(R_2)}$$

La variance du portefeuille est donnée ainsi :

$$\begin{aligned} \sigma_p^2 &= x^2 \sigma_1^2 + (1-x)^2 \sigma_2^2 + 2x(1-x) \text{cov}(R_1, R_2). \\ &= x^2 \sigma_1^2 + (1-x)^2 \sigma_2^2 + 2x(1-x) \rho_{12} \sigma_1 \sigma_2 \\ &= x^2 \sigma_1^2 + (1-x)^2 \sigma_2^2 + 2x(1-x) \sigma_{12} \end{aligned} \quad (1)$$

Cette équation indique dans quelle mesure le risque du portefeuille  $P$  dépend des écarts types des titres détenus et de leur covariance.

Le risque d'un portefeuille de deux actifs 1 et 2 est :

$$\sigma_p = \sqrt{x^2 \sigma_1^2 + (1-x)^2 \sigma_2^2 + 2x(1-x) \text{cov}(R_1, R_2)}$$

où  $\sigma_1$  et  $\sigma_2$  désignent respectivement l'écart type des titres 1 et 2.

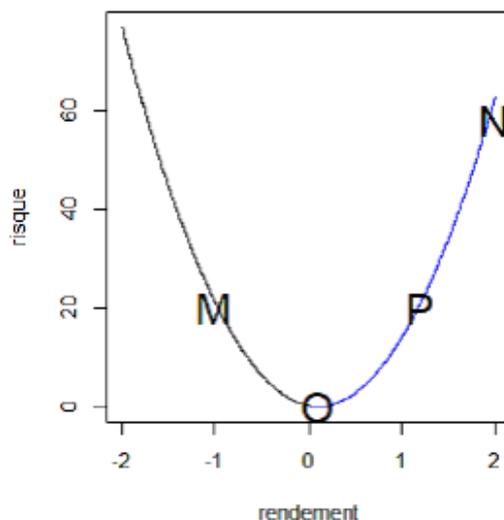
En remplaçant  $x$  par  $\frac{E(R_p) - E(R_2)}{E(R_1) - E(R_2)}$ , dans la formule 1 On obtient une relation entre  $\sigma_p^2$  et  $E[R_p]$  :

$$\begin{aligned} f(E(R_p)) &= \left( \frac{E(R_p) - E(R_2)}{E(R_1) - E(R_2)} \right)^2 x^2 + \left( 1 - \frac{E(R_p) - E(R_2)}{E(R_1) - E(R_2)} \right)^2 (1-x)^2 \\ &+ 2 \left( \frac{E(R_p) - E(R_2)}{E(R_1) - E(R_2)} \right) \left( 1 - \frac{E(R_p) - E(R_2)}{E(R_1) - E(R_2)} \right) \sigma_{12} \\ &= \sigma_p^2. \end{aligned}$$

En développant les calculs, on trouve bien une fonction quadratique en  $E[R_p]$  :

$$\begin{aligned} f(E(R_p)) &= \frac{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}{(E(R_1) - E(R_2))^2} E(R_p)^2 \\ &+ \left( \frac{4E(R_2)\sigma_{12}}{(E(R_1) - E(R_2))^2} + \frac{2\sigma_{12}}{E(R_1) - E(R_2)} \right) E[R_p] \\ &- \left( \frac{2E(R_1)\sigma_2^2}{(E(R_1) - E(R_2))^2} + \frac{2E(R_2)\sigma_1^2}{(E(R_1) - E(R_2))^2} \right) E[R_p] \\ &+ \frac{E(R_2)^2 \sigma_1^2}{(E(R_1) - E(R_2))^2} + \frac{E(R_1)^2 \sigma_2^2}{(E(R_1) - E(R_2))^2} \\ &- \frac{2E(R_2)\sigma_{12}}{E(R_1) - E(R_2)} - \frac{2E(R_2)^2 \sigma_{12}^2}{(E(R_1) - E(R_2))^2}. \end{aligned}$$

Donc on a une fonction qui s'écrit sous forme :  $f(E(R_p)) = aE(R_1)^2 + bE(R_1) + c$   
On peut visualiser notre situation, en traçant la courbe de  $\sigma_p^2$  en fonction de  $E(R_p)$  :



La partie en bleu s'appelle la frontière efficiente ou bien de Markowitz, elle représente l'ensemble de tout les points qui intéressent les détenteurs du portefeuille.

Le point **M** a le même risque que le point **P** mais **M** ne se situe pas sur la frontière efficiente, on remarque clairement cela en projetant sur l'axe des abscisses, on voit que le rendement du point **P** est supérieur à celui du point **M**.

#### 1.4.7.2 Le cas de “n” actifs :

Soit  $R_p$  le rendement du portefeuille composé de  $n$  actifs caractérisés par leur rendement respectif  $R_1, R_2, \dots, R_n$ . On suppose, en outre, que chaque actif  $i$  entre pour une proportion  $x_i$  dans la composition du portefeuille  $P$ .

$$R_p = \sum_{i=1}^n x_i R_i$$

En d'autres termes :

$$E(R_p) = E\left(\sum_{i=1}^n x_i R_i\right) = \sum_{i=1}^n x_i E(R_i)$$

$$V(R_p) = \sum_{i=1}^n \sum_{j=1}^n R_i R_j Cov(x_i, x_j)$$

Sélectionner un portefeuille revient à choisir celui qui :

- Maximise  $E(R_p)$ .
- Minimise  $V(R_p)$ .
- Sous la contrainte que  $\sum_{i=1}^n x_i = 1$ .

Considérons dans un premier temps le cas d'un marché financier composé uniquement d'actifs risqués, l'univers d'investissement contient  $N$  titres financiers risqués indexés par ( $i = 1, \dots, N$ ).

Nous adaptons les notations suivantes :

- $x \in \mathbb{R}^{n \times 1}$  un vecteur représentant les poids d'un portefeuille  $P$ .
- $R \in \mathbb{R}^{n \times 1}$  un vecteur représentant les rentabilités aléatoires des actifs financiers de l'univers d'investissement.  $E[R]$  désigne son espérance.
- $e \in \mathbb{R}^{n \times 1}$  dont toutes les composantes sont égales à 1.
- $V \in M_n(\mathbb{R})$  la matrice des variances-covariances des rentabilités des actifs financiers. On suppose que cette matrice est inversible
- $\sigma_{ii} = \sigma_i^2 = \sigma^2(R_i)$  : variance de la rentabilité du  $i^{\text{ème}}$  actif financier.
- $\sigma_{ij} = \text{cov}(R_i, R_j)$  : covariance entre le taux de rendement du  $i^{\text{ème}}$  actif financier et le taux de rendement du  $j^{\text{ème}}$  actif financier.

Soit en notation matricielle :

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad E(R) = \begin{pmatrix} E(R_1) \\ \vdots \\ E(R_n) \end{pmatrix}, \quad e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad V = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \dots & \sigma_{nn} \end{pmatrix}$$

Donc le risque dédié au portefeuille est défini par :

$$\sigma_p^2 = x^t V x$$

La rentabilité espérée d'un portefeuille  $P$  est égale à :

$$E_0 = \sum_{i=1}^n x_i E(R_i) = \hat{x} E(R)$$

L'espérance mathématique du rendement du portefeuille de  $n$  actifs est la moyenne pondérée des différents rendements des titres qui le composent.

$$\bar{R} = \sum_{i=1}^n x_i R_i$$

Markowitz utilise comme mesure du risque la variance du portefeuille :

$$\sigma_p^2 = \sum_{i=1}^n x_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n x_i x_j \rho_{ij} \sigma_i \sigma_j$$

### 1.4.8 Le Portefeuille efficient

Développée par l'économiste américain Harry Markowitz (l'un des Prix Nobel d'économie 1990), la théorie du portefeuille efficient définit le mécanisme de sélection de titres constituant un portefeuille boursier idéal. Les travaux de Markowitz ont constitué la première tentative de théorisation de la gestion financière de portefeuilles et son modèle suggère une procédure de sélection de plusieurs titres boursiers, à partir de critères statistiques, afin d'obtenir des portefeuilles optimaux. Plus précisément, Markowitz a montré que l'investisseur cherche à optimiser ses choix en tenant compte non seulement de la rentabilité attendue de ses placements, mais aussi du risque de son portefeuille qu'il définit mathématiquement par la variance de sa rentabilité. Ainsi, le "portefeuille efficient" est le

portefeuille le plus rentable pour un niveau de risque donné. Il est déterminé au mieux par application de méthodes de programmation quadratique ou sinon de manière heuristique

H. Markowitz distingue deux portefeuilles : un portefeuille efficient purement risqué et un portefeuille efficient non risqué.

Dans les deux cas, il s'agit de trouver la combinaison optimale de  $X$  titres risqués (ou non risqués) grâce auxquels un investisseur obtiendra le niveau maximal de rendement espéré pour un niveau de variance (risque) donné.

Les portefeuilles efficientes sont des portefeuilles d'actifs dont la composition permet à un investisseur d'optimiser le couple rendement /risque. Les portefeuilles efficientes doivent donc avoir la plus forte rentabilité possible pour un même risque. Ils impliquent également, pour une même rentabilité, de se voir associer le niveau de risque le plus faible.

Le modèle de Markowitz ne définit pas un unique portefeuille optimal mais génère une frontière efficiente comprenant l'ensemble des portefeuilles optimaux. C'est à l'investisseur de choisir son portefeuille optimal

En finance, on part du principe que l'investisseur a une aversion au risque. Plus les possibilités de pertes sont élevées, moins l'investisseur acceptera un investissement dit équitable. Un portefeuille efficient est en général diversifié autour d'une base composée d'instruments sûrs, comme des bons du Trésor, des obligations convertibles, ou des placements à terme à taux d'intérêt fixe. Le nombre de titres à très haut rendement mais à très haut risque est donc marginal dans un portefeuille efficient.

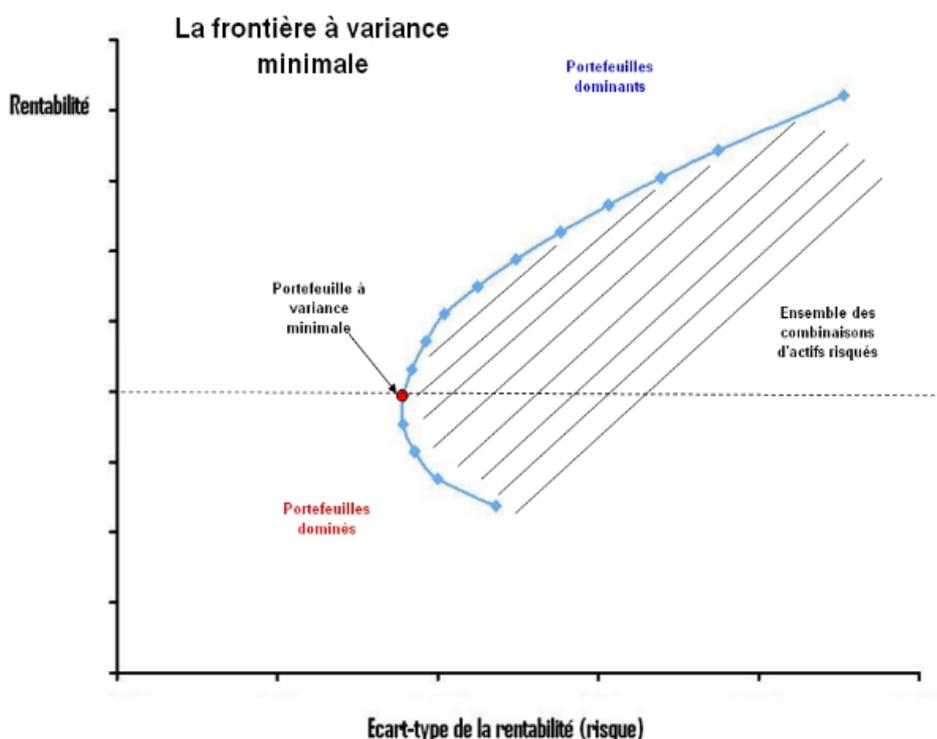


FIGURE 1.2: *Frontière Efficiente.*

Trois propriétés de la frontières efficiente sont parfois utiles :

1. Une combinaison de portefeuilles efficients est un portefeuille efficient.
2. On peut construire toute la frontière avec 2 portefeuilles efficients.
3. Pour chaque portefeuille efficient, il existe un autre portefeuille efficient avec lequel il n'est pas corrélé .

Un portefeuille efficient est un portefeuille dont la rentabilité moyenne est maximale pour un niveau de risque donné, ou dont le risque est minimal pour une rentabilité donnée.

Les portefeuilles efficients sont sur la «*frontière efficiente*» de l'ensemble des portefeuilles dans le plan  $(\sigma_p^2; E_0)$ .

Notre objectif est de pouvoir déterminer cette frontière efficiente ou du moins exprimer une fonction qui permet de déterminer le portefeuille pour un niveau de rendement cible  $E_0$ . Ce problème peut se formuler comme ci-dessous :

$$\left( \begin{array}{l} \min (x^t V x) \\ \text{sc : } \left\{ \begin{array}{l} x^t E (R) = E_0 \\ x^t e = 1 \end{array} \right. \end{array} \right)$$

La frontière efficace est l'ensemble de portefeuilles efficaces (efficients).

Dans le cas où ,  $x_j = 1, \dots, n$  est quelconque, c'est-à-dire les ventes à découverte sont autorisées et si la contrainte sur la rendement minimum attendu par l'investisseur est égal à  $\rho$ , on utilise la technique des multiplicateurs de Lagrange pour calculer la solution optimale. La fonction Lagrangienne utilisée est donnée par [22] :

$$L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2) = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} + \lambda_1 \left( \sum_{j=1}^n x_j \bar{R}_j - \rho \right) + \lambda_2 \left( \sum_{j=1}^{n+1} x_j - 1 \right)$$

Il s'ensuit que

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= 2 \sum_{j=1}^n x_j \sigma_{ij} + \lambda_1 \bar{R}_i + \lambda_2 = 0 \quad i = 1, \dots, n \\ \frac{\partial L}{\partial x_{n+1}} &= \lambda_1 r_o + \lambda_2 = 0 \\ \frac{\partial L}{\partial \lambda_1} &= \sum_{j=1}^n x_j \bar{R}_j - \rho = 0 \\ \frac{\partial L}{\partial \lambda_2} &= \sum_{j=1}^n x_j - 1 = 0 \end{aligned}$$

**Définition 1.7** Soit  $A_1$  et  $A_2$  deux portefeuilles de  $\mathbb{R}^d$ . On dit que  $A_1$  est préférable à  $A_2$  au sens du critère de moyenne variance, et on note  $A_1 \geq_{mv} A_2$  ssi :

$$E [R_{A_1}] \geq E [R_{A_2}] \quad \text{et} \quad Var [R_{A_1}] \leq Var [R_{A_2}]$$

La relation  $\geq_{mv}$  définit un préordre sur l'ensemble  $A = \mathbb{R}^d$ . Ce préordre est partiel.

**Définition 1.8** Une portefeuille est dit **efficient** s'il n'est dominé par aucun autre portefeuille eu sens du préordre  $\geq_{mv}$ . On appelle frontière efficiente l'ensemble des couples

$$\{(E[R_A], Var[R_A]), \text{ avec } A \text{ portefeuille efficient}\},$$

Ou encore l'ensemble des couples

$$\left\{ \left( E[R_A], \sqrt{Var[R_A]} \right), \text{ avec } A \text{ portefeuille efficient} \right\}.$$

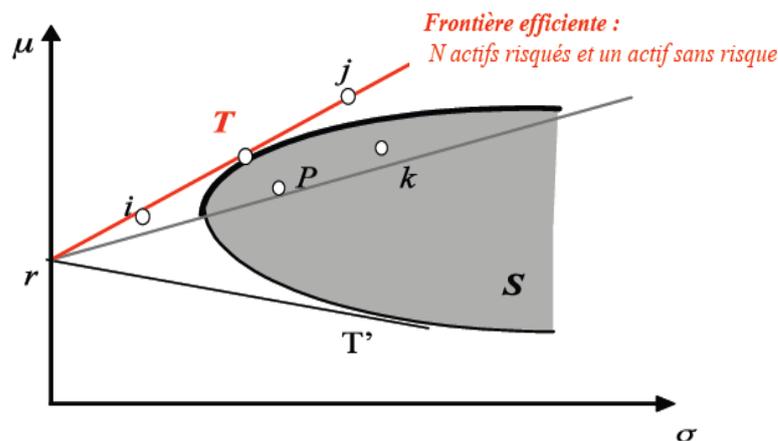
**Remarque 1.2** Détermination de la frontière efficiente. Ce qui revient à la résolution du problème (Pr) ou d'une manière équivalente, du problème (Pr) :

$$(Pr)_\sigma : \begin{cases} \sup_{A \in \mathbb{R}^d} E[R_A] \\ \text{s.c. } Var[R_A] = \sigma^2 \end{cases} \quad (Pr)_m : \begin{cases} \inf_{A \in \mathbb{R}^d} Var[R_A] \\ \text{s.c. } E[R_A] = m \end{cases}$$

### 1.4.9 Le portefeuille de marché

Dans la détermination du portefeuille efficient, nous incluons l'actif hors risque, nous allons obtenir une droite qui représentera la frontière efficiente. Il a été démontré que cette droite est la tangente à la frontière efficiente lorsque celle-ci est calculée sans l'actif hors risque (courbe) au point qui représente le portefeuille de marché.

En présence d'un actif sans risque, l'investisseur peut prêter (et, parfois, emprunter) au taux d'intérêt du marché, noté  $r$  et supposé positif. Un portefeuille investi intégralement dans cet actif sans risque est représenté sur la figure suivante par le point d'abscisse nulle et d'ordonnée  $r$ .



On peut démontrer, dans ce cas, les résultats très importants suivants :

a) La frontière efficiente devient la demi-droite (en traits épais sur la figure) tangente à l'hyperbole de Markowitz passant par le point  $r$ .  $T$  est le portefeuille efficient tangent et

ne comprend que des actifs risqués. L'équation de cette droite est la suivante :

$$\mu_p = r + \left( \frac{\mu_T - r}{\sigma_T} \right) \sigma_P$$

b) Pour les portefeuilles efficients, c'est-à-dire bien diversifiés, la relation entre rentabilité espérée et risque est linéaire, résultat justifiant par sa simplicité l'adoption de l'écart-type plutôt que de la variance comme mesure de risque.

c) Tous les portefeuilles efficients sont des combinaisons de l'actif sans risque et du portefeuille tangent est appelé « théorème de séparation en deux fonds (ou portefeuilles) : bien que  $N + 1$  titres différents ( $N$  risqués et un sans risque) soient présents sur le marché, tous les portefeuilles efficients se construisent à partir des mêmes deux portefeuilles ( $r$  et  $T$ ).

d) La frontière efficiente domine en tous points sauf le point de tangence  $T$  qui située au nord-ouest de l'hyperbole dans le plan  $(\sigma, \mu)$ .

**Remarque 1.3** *Le portefeuille de marché est le portefeuille qui contient tous les actifs risqués (toutes les actions) dans les proportions de leurs capitalisations boursières.*

### 1.4.10 Le Portefeuille Tangent

Il existe un actif supplémentaire sans risque, de rendement constant  $\mu_0$ . L'actif sans risque est un concept théorique qui n'existe pas en pratique. On l'associe souvent au rendement des bons du trésor de pays développés ayant un marché financier stable, car on considère peu probable qu'ils fassent défaut. Cependant, comme la crise souveraine récente l'a rappelé en Europe, la dette souveraine est loin d'être un actif sans risque [89].

L'investisseur répartit le capital entre les actifs risqués et cet actif sans risque en attribuant à celui-ci un poids  $x_0$  dans le portefeuille, dont le rendement devient d'espérance  $\mu_P = x_0\mu_0 + x^T\mu$  mais toujours de variance  $\sigma_p^2 = x^T\Sigma x$ . Le problème d'optimisation sur les actifs risqués devient :

$$x^T = \arg \min x^T \Sigma x \quad \text{s.c.} \quad x_0\mu_0 + x^T\mu = r, \quad x_0 + x^T\mathbf{1} = 1$$

On peut montrer que les poids des actifs risqués s'écrivent alors :

$$x^T = \frac{\Sigma^{-1}(\mu + \mu_0\mathbf{1})}{\mathbf{1}^T \Sigma^{-1}(\mu + \mu_0\mathbf{1})}$$

Ces poids ne dépendent alors plus de la contrainte de rendement  $r$ , mais du rendement  $\mu_0$  de l'actif sans risque.

On définit le portefeuille tangent dans le plan (écart-type, moyenne) comme l'intersection entre la frontière efficiente (actifs risqués) et sa tangente passant par le point  $(0, \mu_0)$  représentant l'actif sans risque. Tout portefeuille efficient est donc une combinaison linéaire entre l'actif sans risque et le portefeuille tangent, et on appelle la droite passant par ces deux points la droite de marché (*Capital Market Line*). Tous les investisseurs ayant les mêmes estimations de  $\mu$  et  $\Sigma$  auront donc des portefeuilles appartenant à cette même droite, répartissant leur capital entre le portefeuille tangent et l'actif sans risque.

## Optimisation de portefeuille en finance : Modèles et méthodes

### 2.1 Le Modèle Moyenne-Variance de Markowitz (1952)

La théorie moderne du portefeuille et les modèles de choix de portefeuille est née en 1952 avec la publication de l'article fondateur de Harry Markowitz. et dans le cadre de la théorie de l'espérance d'utilité, aussi d'après l'axiomatisation des préférences individuelles de von Neumann et Morgenstern. et les travaux de Harry Markowitz [57] pour la sélection de titres pour créer le portefeuille le plus efficient possible Ce mode de sélection permet de minimiser le risque pour un niveau de rendement choisi. c'est à dire qui possède la rentabilité maximum pour un niveau de risque minimum.

Le modèle proposé par Markowitz est la recherche d'un portefeuille efficace, il a l'espérance de rentabilité la plus forte parmi les portefeuilles qui ont la même variance de rentabilité que lui. L'ensemble de tous les portefeuilles efficaces constitue la frontière efficiente de Markowitz (appelée à frontière efficace).

L'investisseur est alors présumé prendre ses décisions en fonction seulement de deux paramètres : l'espérance de sa richesse qu'il souhaite la plus grande possible, et sa variance qu'il désire la plus faible possible. Cette approche a offert le premier traitement systématique d'un dilemme auquel les investisseurs font face [79].

L'approche de Markowitz est plus connue sous le nom de l'approche Moyenne-Variance (MV). Cette approche connaît aussi certains inconvénients. D'abord, elle suppose que,

- Soit les rendements suivent une distribution normale (ou multi-normale),
- Soit la fonction d'utilité est quadratique [59].
- Manque de sensibilité du modèle quant aux différences entre les gains et les pertes.
- L'approche MV n'est pas, en général, compatible avec la dominance stochastique.
- MV n'est pas cohérente car elle ne respecte pas les axiomes de monotonie et d'invariance par translation [64].

## 2.1.1 Hypothèses

### 2.1.1.1 Les hypothèses relatives aux actifs financiers

- **Hypothèse 01** : Tout investissement est une décision prise dans une situation de risque : le return d'un actif financier est une variable aléatoire, qu'elle est distribuée selon une loi normale, c'est à dire une distribution symétrique stable entièrement définie par deux paramètres, c'est-à-dire une distribution symétrique stable entièrement définie par deux paramètres, l'espérance mathématique de return et l'écart-type de la distribution de probabilité de return.
- **Hypothèse 02** : Les rendements des différents actifs financiers sont corrélés ( $Cov(R_i, R_j) = 0$ ) ne fluctuent pas indépendamment les uns des autres
- **Hypothèse 03** : Les marchés sont parfaits : toutes les conditions pour que les prix correspondent à la réalité du moment sont réunies. L'entrée et la sortie sont libres et sans coût. L'information circule de manière totalement transparente. La concurrence est parfaite entre les acteurs composant le marché, il y a absence totale d'impôt [79].

### 2.1.1.2 Les hypothèses relatives aux comportements des investisseurs

- **Hypothèse 01** : Le comportement de tous les investisseurs est caractérisé par un degré plus ou moins prononcé d'aversion vis-à-vis du risque. Ce dernier est mesuré par l'écart-type de la distribution de probabilité du rendement.
- **Hypothèse 02** : Les investisseurs sont rationnels : bien que leur fonction de préférence soit purement subjective, leur comportement est caractérisé par un degré plus ou moins prononcé d'aversion au risque.
- **Hypothèse 03** : Tous les investisseurs ont la même période de l'investissement, ont prennent leurs décisions en même temps. Cette simplification, qui peut paraître exagérée, permet de mettre en œuvre un modèle de décision qui tient compte du caractère hautement combinatoire du portefeuille.

## 2.1.2 Le Modèle Mean–Variance de Markowitz

Le rendement du portefeuille est une variable aléatoire dont l'espérance sera donnée par [22] :

$$E[(R(x))] = E\left[\sum_{j=1}^n r_j x_j\right] = \sum_{j=1}^n x_j E[r_j] = \sum_{j=1}^n x_j \bar{r}_j$$

La variance du rendement de portefeuille est :

$$\begin{aligned}
 \sigma^2 [R(x)] &= E [R(x) - E(R(x))]^2 \\
 &= E \left[ \sum_{i=1}^n (r_i - \bar{r}_i) x_i \right]^2 \\
 &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \text{Cov}(r_i, r_j) \\
 &= \sum_{i=1}^n x_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} \\
 \sigma^2 [R(x)] &= x^t Q x
 \end{aligned}$$

Nous pouvons résoudre le problème quadratique convexe suivant qui minimise le risque  $x^t Q x$ , la rendement minimum attendu par l'investisseur est égal à  $\rho$ .

L'algorithme d'optimisation de Markowitz s'écrit comme suit :

$$\min \left[ \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} \right]$$

Sous les contraintes :

$$\begin{cases} \sum_{j=1}^n x_j E[r_j] \geq \rho \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 \quad j = 1, \dots, n \end{cases}$$

Nous pouvons résoudre le problème quadratique convexe suivant qui minimise le risque  $(x^t Q x)$  pour un gain fixé  $\rho$  :

$$\min [x^t Q x]$$

Sous les contraintes :

$$\begin{cases} \sum_{j=1}^n x_j E[r_j] \geq p \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 \quad j = 1, \dots, n \end{cases}$$

En pratique, le modèle suivant est très utilis

$$\min \left[ \lambda x^t Q x - (1 - \lambda) \sum_{j=1}^n r_j x_j \right]$$

Sous les contraintes :

$$\begin{cases} \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 \quad j = 1, \dots, n \end{cases}$$

où  $\lambda$  est le paramètre d'aversion au risque et  $0 \leq \lambda \leq 1$  correspond à l'individu purement risque-averse et  $\lambda = 0$  correspond à celui qui n'a aucune peur de prendre le risque.

De la même manière, si nous faisons varier  $\lambda$  entre 0 et 1, nous pourrions tracer la courbe des portefeuilles efficients [64].

Pour  $x_i = 1, \dots, n$  est quelconque, c'est-à-dire les ventes à découverte sont autorisées et si la contrainte sur la rendement minimum attendu par l'investisseur est égal à  $\rho$ , on utilise la technique des multiplicateurs de Lagrange pour calculer la solution optimale. La fonction Lagrangienne utilisée est donnée par [22] :

$$L(x_1, \dots, x_n, \lambda_1, \lambda_2) = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} + \lambda_1 \left( \sum_{j=1}^n x_j \bar{R}_j - \rho \right) + \lambda_2 \left( \sum_{j=1}^n x_j - 1 \right)$$

Il s'ensuit que

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= 2 \sum_{j=1}^n x_j \sigma_{ij} + \lambda_1 \bar{R}_i + \lambda_2 = 0 \quad i = 1, \dots, n \\ \frac{\partial L}{\partial \lambda_1} &= \sum_{j=1}^n x_j \bar{R}_j - \rho = 0 \\ \frac{\partial L}{\partial \lambda_2} &= \sum_{j=1}^n x_j - 1 = 0 \end{aligned}$$

L'utilisation de la variance comme mesure de risque a des avantages et des inconvénients. L'avantage de cette mesure de risque est sa simplicité. De plus, en utilisant la matrice de variance-covariance, nous pouvons construire un modèle de programmation quadratique convexe, afin de calculer les portefeuilles efficaces. Actuellement, un problème quadratique convexe se résout facilement même pour des dimensions très grandes comme 10000. Mais cela ne suffit pas car certaines contraintes du monde réel ne sont pas prises en compte dans le modèle de base [59].

L'approche de Markowitz est plus connue sous le nom de l'approche Moyenne-Variance (MV). Cette approche connaît aussi certains inconvénients. D'abord, elle suppose que,

- soit les rendements suivent une distribution normale (ou multi-normale),
- soit la fonction d'utilité est quadratique [58].

Ces deux cas sont malheureusement assez peu réalistes. La deuxième critique est due au manque de sensibilité du modèle quant aux différences entre les gains et les pertes car l'approche MV pénalise de la même manière les gains ou les pertes s'éloignant trop de la moyenne [64]. L'approche MV n'est pas, en général, compatible avec la dominance stochastique [59]. Finalement, l'approche MV n'est pas cohérente car elle ne respecte pas les axiomes de monotonie et d'invariance par translation.

Après le travail de Markowitz, les modèles alternatifs ont été développés pour pallier aux défauts du modèle MV.

## 2.2 Le modèle de marché de Sharpe(1963-1964)

### 2.2.1 Les modèles à facteurs

On a vu précédemment que la théorie de base de Markowitz portait essentiellement sur la maximisation de l'utilité de richesse finale de l'investisseur, une théorie que certes

a proposé une solution au problème de choix de portefeuille d'un investisseur averse au risque où les portefeuilles optimaux sont définies comme ceux de plus faible risque pour une rentabilité donnée, mais aussi a présenté un problème majeur concernant le temps nécessaire pour calculer et inverser la matrice des corrélations. Suite à cette difficulté plusieurs algorithmes ont été proposés afin de résoudre ce problème on peut citer à titre d'exemple la méthode de Wolf (1959) ainsi que d'autres méthodes qui s'enracinent de la méthode simplifiée de sélection de portefeuille qui s'inscrit dans le cadre du modèle de marché de Sharpe [81] qu'on va détailler en ce qui suit.

L'idée d'un modèle factoriel de portefeuille est d'expliquer les revenus procurés par les titres financiers par un petit nombre de variables faciles à observer (les facteurs). Ces facteurs peuvent être eux-mêmes assimilables à des titres (par exemples des indices boursiers).

L'objectif principal de cette classe de modèles est la réduction massive du nombre de paramètres, ensuite l'estimation de la matrice de covariance à partir de n'importe quel modèle linéaire. Le calcul de la matrice de covariance par un modèle factoriel est donc très contraignant dans la mesure où la matrice n'incorpore pas de modélisation temporelle. Ce qui implique que : «La classe des modèles factoriels consiste à exprimer l'évolution d'une variable en fonction d'une fonction linéaire choisie »

Les modèles à facteurs, contrairement au MEDAF, sont de pures constructions statistiques sans fondement théorique. Ils permettent de simplifier grandement l'estimation des covariances des rentabilités des différents titres. Nous commençons par le modèle à un facteur de Sharpe, avant de présenter les modèles multi-facteurs.

## 2.2.2 Le modèle à un facteur ou modèle de marché

### 2.2.2.1 Contribution de ce modèle :

« *Management Science* » l'article intitulé *A Simplified Model for Portfolio Analysis*. Cet article complète l'analyse de Markowitz (la détermination d'un portefeuille optimal) : l'analyse de portefeuille.

En 1963, William Sharpe [81] présente un modèle simplifié, en montrant comment cette simplification facilite la résolution du modèle du choix des titres, par l'adjonction de quelques hypothèses. Mais quelles hypothèses retenir ?

Sharpe reprend donc le partitionnement de Markowitz entre la cause commune et les causes propres de la rentabilité des titres. Il postule sans démonstration que le rendement de tout titre  $i$  est lié linéairement à un indice de marché noté  $R_M$ .

### 2.2.2.2 Hypothèses du modèle :

Pour simplifier les procédures d'estimation dans ce modèle, deux hypothèses ont été relevées :

**H** : les rendements des différentes valeurs sont exclusivement liés entre eux par une relation avec un facteur de base sous-jacent, autrement dit que chaque valeur doit être estimée par un seul facteur exogène, il s'agit du modèle unifactoriel ou modèle des indices. Le terme "unifactoriel" vient donc du fait qu'à la base le but du modèle de Sharpe est de définir le rendement d'un placement financier en fonction de son risque non diversifiable,

assimile au seul risque de marche (ou risque systématique) donne par un nombre appelé "coefficient beta". Les investisseurs et gestionnaires distinguent trois sortes de risques :

1. Le "risque spécifique" relatif (implicite) au titre lui-même (sa variance).
2. Le "risque systématique/non diversifiable" relatif à l'économie/marché au sens le plus large (variance du portefeuille de référence du marché).
3. Le "risque global" qui est en quelque sorte la somme des deux (c'est un peu plus subtil qu'une simple somme).

Le principe de W.Sharpe (1963) est basé sur le développement de la gestion du portefeuille moyenne variance simplifiée, au lieu du modèle avancé par Markowitz pour le calcul de la frontière d'efficience qu'est caractérisée par sa lourdeur. Ce modèle est dit "**Modèle De Marché**" il établit une relation linéaire entre le rendement d'une action  $i$  et le rendement du marché, ou modèle a un seul indice, il ne contient pas de notion d'équilibre et ne fait pas d'hypothèses particulières sur le marché et les investisseurs il propose simplement une vision simplifier.

### 2.2.2.3 Présentation mathématique du modèle :

Malgré ses airs de parenté avec le MEDAF, le modèle à un facteur de Sharpe, dit également modèle de marché, est fondé sur une décomposition statistique *ad hoc* et n'est absolument pas nécessaire à la validité dudit CAPM. Il permet cependant de donner une explication très intuitive de la théorie du portefeuille. Le modèle de marché repose sur l'hypothèse selon laquelle l'aléa sur chaque titre est dû :

- pour partie, à l'aléa entachant la rentabilité globale  $R_M$  du marché ;
- pour partie, à un aléa spécifique au titre considéré, indépendant du précédent.

En considérant la même hypothèse que dans le modèle de Markowitz, à savoir un portefeuille dont le rendement  $R_p$  est défini par :

$$R_p = \sum_{i=1}^n x_i R_i$$

Et selon l'hypothèse posé par Sharpe : La rentabilité  $R_i$  d'un titre  $i$  quelconque s'écrit ainsi comme la somme des composantes :

$$R_i = \alpha_i + \beta_i R_M + \varepsilon_i \quad i = 1, \dots, n$$

- **La première composante**  $\alpha_i$  valeur espérée lorsque  $I$  est nul est une grandeur non aléatoire spécifique à  $i$  ;
- **La deuxième**,  $\beta_i R_M$ , exprime la réaction de la rentabilité  $R_i$  aux fluctuations du marché et reflète donc l'intensité de la co-variation de  $R_i$  et de  $R_M$  ; il s'agit de la pente de la droite de régression de  $R_i$  sur  $R_M$ ,
- **La troisième composante**  $\varepsilon_i$  représente le risque spécifique, ou diversifiable, du titre  $i$  ; l'aléa  $\varepsilon_i$  est d'espérance nulle, indépendant de la rentabilité du marché  $R_M$  et de tous les  $\varepsilon_i$  caractérisant les autres titres (c'est donc un bruit blanc).
- $I$  est donc le rendement d'un indice économique donné (indice boursier, indice du produit national brut, indice des prix ou voir même rendement le rendement du

portefeuille du marché lui-même... ) au temps  $t$  et est la variable expliquée de la régression (selon la terminologie utilisée dans le chapitre de Méthodes Numériques) considérée comme une variable aléatoire.

- $\alpha_i, \beta_i$  sont des estimateurs non biaisés des paramètres propres à cette valeur. Le premier terme appelée en finance "coefficient alpha" est simplement l'ordonnée à l'origine de la régression (le rendement de l'actif lorsque le rendement de l'indice de référence est nul soit lorsque le marché à un rendement nul) et le deuxième paramètre est pour rappel simplement le bêta du portefeuille risqué  $i$ . Les paramètres  $\alpha_i$  et  $\beta_i$  sont obtenus par la régression simple comme suit :

$$\beta_i = \frac{\text{cov}(r_i, R_M)}{\text{var}(R_M)} \quad (2.1)$$

et

$$\alpha_i = \quad (2.2)$$

La rentabilité du portefeuille  $P$  s'écrit également comme la somme de trois composantes. En notant  $x_i$  le poids du titre  $i$  dans le portefeuille  $P$ , l'on a :  $R_p = \sum_{i=1}^n x_i R_i$  et il vient :

$$\begin{aligned} R_p &= (x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n) + (x_1\beta_1 + x_2\beta_2 + \dots + x_n\beta_n) R_M \\ &\quad + (x_1\varepsilon_1 + x_2\varepsilon_2 + \dots + x_n\varepsilon_n) \\ R_p &= I_1 + I_2 + I_3, \end{aligned}$$

telle que :

$$\begin{aligned} I_1 &= (x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n) \\ &= \alpha_P, \\ I_2 &= (x_1\beta_1 + x_2\beta_2 + \dots + x_n\beta_n) R_M \\ &= \beta_P R_M, \end{aligned}$$

et

$$\begin{aligned} I_3 &= (x_1\varepsilon_1 + x_2\varepsilon_2 + \dots + x_n\varepsilon_n) \\ &= \varepsilon_P. \end{aligned}$$

La composante  $\varepsilon_P$ , est la somme pondérée des risques indépendants, spécifiques aux titres individuels. Or, la loi des grands nombres permet d'annuler pratiquement ce dernier risque pour vu que le portefeuille  $P$  soit bien diversifié. En effet, quand certains  $\varepsilon_i$  sont positifs, d'autres sont négatifs et leur somme pondérée est toujours pratiquement nulle, si  $n$  est grand et aucun  $x_i$  trop important. C'est pourquoi le composant  $\varepsilon_i$  de chaque rentabilité  $R_i$  est appelé risque diversifiable car il est (presque) éliminé dans un portefeuille bien diversifié.

Puisque le bêta est une covariance (normée par une constante, la variance du marché), c'est un opérateur linéaire et le bêta d'un portefeuille,  $\beta_P$ , est égal à la moyenne pondérée des bêtas des titres qui le composent. Par conséquent, la deuxième composante,  $\beta_P R_M$ ,

ne peut pas être éliminée par diversification et constitue de ce fait le seul véritable aléa entachant la rentabilité d'un portefeuille bien diversifié.

La rentabilité aléatoire d'un portefeuille, non nécessairement bien diversifié, s'écrit :

$$R_p = \alpha_P + \beta_P R_M + \varepsilon_P$$

Dès lors, sa variance totale (Le risque de portefeuille),  $\sigma_p^2$ , est égale à :

$$\begin{aligned} \sigma_p^2(R_p) &= \text{Var}(R_p) \\ &= \text{Var}(\alpha_P + \beta_P R_M + \varepsilon_P) \\ &= \beta_P^2 \sigma^2(R_M) + \sigma^2(\varepsilon_P) \\ &= \beta_P^2 \sigma^2(R_M) + \sum_{i=1}^n x_i^2 \sigma_{\varepsilon_i}^2. \end{aligned}$$

Comme dans le cas d'un seul titre, le risque de portefeuille peut être décomposé en deux parties :

-  $\beta_P^2 \sigma_M^2$  : le risque systématique de portefeuille  
 -  $\sum_{i=1}^n x_i^2 \sigma_{\varepsilon_i}^2$  : le risque non systématique qui peut être éliminé en diversifiant le capital investi entre l'ensemble des actions.

Cette décomposition de risque est appelée l'effet de portefeuille ou l'effet de construction de portefeuille sur le risque, qui se décompose en deux éléments :

- L'effet Markowitz (ou effet des corrélations négatives)
- L'effet de diversification (ou effet des non-corrélations)

Donc l'interprétation du bêta peut être résumée comme suit :

- -Un investissement ne présentant aucun risque afficherait donc un bêta nul.
- -Un bêta inférieur à 1 indique que si le marché est à la baisse, le titre sera susceptible de baisser moins que le marché.
- -Un bêta supérieur à 1 indiquera que si le marché est à la hausse, le titre sera susceptible de suivre moins rapidement la tendance à la hausse.

Le concept de bêta ayant été introduit, passons maintenant à la théorie du modèle qui a pour objectif donc de simplifier celui de Markowitz en utilisant ce fameux coefficient. Par définition, le bêta global d'un portefeuille est déterminé à partir des bêta pondérés respectifs de chacun des titres ou bêta sous-jacents qui le composent tel que :

$$\beta_P = x_1 \beta_1 + x_2 \beta_2 + \dots + x_n \beta_n = \sum_{i=1}^n x_i \beta_i$$

Avec  $i$  étant le bêta du portefeuille global,  $X_i$  la proportion du titre  $i$  dans le portefeuille  $P$ ,  $\beta_i$  le bêta du titre  $i$  et  $n$  le nombre d'actifs financiers présents dans le portefeuille.

On peut déduire que le risque d'un portefeuille bien diversifié est constitué donc uniquement du risque de marché.

## 2.3 Le modèle simplifié de Sharpe

### 2.3.1 Présentation mathématique du modèle

Sharpe [81] a été le premier qui a tenté de simplifier le modèle de Markowitz en développant les modèles à indice qui se base sur la simplification de la matrice de variances-covariances afin de réduire la charge de calcul. Sharpe a proposé une diagonalisation de cette matrice en se basant sur le modèle à un seul indice en supposant que les fluctuations des rendements des actions peuvent être exprimés à l'aide d'une régression simple. Autrement dit,

$$r_i = a_i + bR_I + \varepsilon_i \quad \text{pour } i = 1, \dots, n$$

telle que :

- $R_I$  : est le rendement de l'indice I.
- $\varepsilon_i$  est une variable aléatoire appelée bruit blanc qui vérifié :  $E(\varepsilon_i) = 0$ , et  $\sigma_{\varepsilon_i}^2 \neq 0$  pour  $i = 1, \dots, n$

Le rendement de portefeuille devient :

$$R(x) = \sum_{i=1}^n x_i r_i = \sum_{i=1}^n x_i a_i + R_I \left( \sum_{i=1}^n x_i b_i \right) + \sum_{i=1}^n x_i \varepsilon_i$$

Soit

$$x_{n+1} = \sum_{i=1}^n x_i b_i$$

il en résulte :

$$R(x) = \sum_{i=1}^n x_i a_i + x_{n+1} R_I + \sum_{i=1}^n x_i \varepsilon_i$$

Le rendement espéré est donné par :

$$E[R(x)] = \sum_{i=1}^n x_i a_i + x_{n+1} E[R_I]$$

La variance du rendement est :

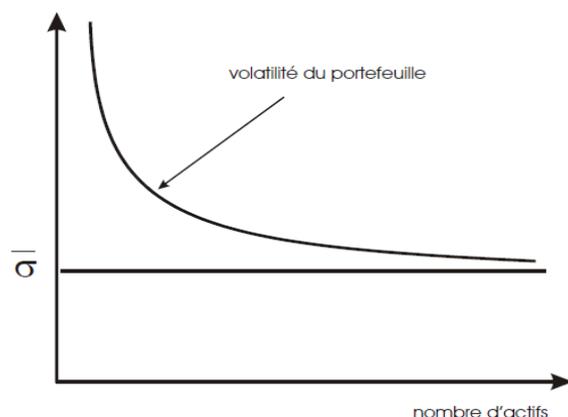
$$\sigma_p^2 = \sum_{i=1}^n x_i^2 \sigma_{\varepsilon_i}^2 + x_{n+1}^2 \sigma_I^2$$

Donc on a besoin que de  $(n+1)$  termes à estimer au lieu de  $n \cdot (n+1)/2$  variances et covariances pour l'approche de Markowitz.

- Le concept de l'approche de Markowitz est basé sur la diversification qui permet de réduire davantage le risque du portefeuille. Malheureusement on ne peut réduire complètement le risque en augmentant indéfiniment la taille de portefeuille.
- Sharpe a montré que le risque d'un portefeuille quelconque peut être décomposé en deux parties : le risque diversifiable ou risque non systématique et le risque non diversifiable ou risque de marché.

#### Remarque :

Il faut souligner que l'augmentation du nombre de titres au-delà d'un certain seuil ne permet pas de réduire le risque. La figure suivante montre bien cette situation :



## 2.4 Le modèle d'équilibre des actifs financiers (MEDAF)

### 2.4.1 Définition du MEDAF

En 1963. A la suite des travaux de Markowitz, puis Sharpe 1962, le modèle linéaire vient de pénétrer par effraction dans la finance moderne, on liant l'indice de marché à celui de portefeuille.

Grâce aux travaux de Jack Treynor (1961), William Sharpe [81], John Lintner (1965) et Jan Mossin (1966). basé sur la théorie de choix de portefeuille de Markowitz, Sharpe un an plus tard, qui va rapidement s'apercevoir que l'on peut dire beaucoup mieux, et beaucoup plus, que le modèle de marché, et forger le modèle d'équilibre des actifs financiers, ou MEDAF. Ce modèle d'évaluation des actifs financiers (MEDAF) appelé aussi « Capital Asset Pricing Model (CAPM), en anglais » a été développé par Sharpe (1964), Lintner (1965), et Mossin (1966).

(MEDAF) est un modèle dont le but est de déterminer les rentabilités appropriées des différents actifs qui forment le marché financier. C'est un modèle d'équilibre, c'est à dire que les rentabilités sont expliquées par l'adéquation entre l'offre et la demande en titres des différents agents qui interviennent sur le marché. Le message essentiel du MEDAF est que le risque de chaque titre peut être décomposé en une composante systématique commune à tous les actifs du marché et une composante spécifique diversifiable, le marché ne rémunérant que la composante systématique.

La prise en compte de la pluralité des investisseurs conduit sous certaines hypothèses à un modèle exprimant les rentabilités espérées d'équilibre, le Modèle d'Equilibre des Actifs Financiers (MEDAF), ou Capital Asset Pricing Model (CAPM). Son message central est que, pour tout actif financier pris isolément, la relation entre son risque et sa rentabilité espérée est linéaire, à condition de mesurer ce risque par sa covariance avec le marché pris dans son ensemble, et non sa variance ou son écart-type.

La façon la plus simple (il y en a d'autres) de déterminer la relation entre rentabilité espérée et risque d'un titre financier à l'équilibre du marché est de supposer que tous les investisseurs obéissent au critère  $E - V$  et qu'il existe un actif sans risque. De plus,

on suppose que les individus ont le même horizon d'investissement et des anticipations identiques, c'est-à-dire qu'ils utilisent tous le même vecteur des rentabilités espérées et la même matrice de variance covariance.

## 2.4.2 Les principales hypothèses du MEDAF

Le modèle d'évaluation des actifs financiers (MEDAF), est un corollaire direct de la droite de marché. On se place dans les hypothèses suivantes :

### 2.4.2.1 Les Hypothèses Relatives aux Comportements des Investisseurs

1. les investisseurs exigent une rentabilité d'autant plus forte que le risque est élevé : il existe donc une relation croissante entre rendement et risque.
2. les investisseurs raisonnent tous dans le cadre de Markowitz, avec le même objectif d'optimisation statique,
3. les investisseurs disposent de la même information et des mêmes estimations pour les rendements.
4. les investisseurs ont la même période de l'investissement.
5. les investisseurs prennent leurs décisions en même temps.
6. les investisseurs détiennent leurs actifs pendant la même période.

### 2.4.2.2 Les Hypothèses Relatives aux Actifs Financiers

1. le marché est efficient.
2. il n'y a pas de coûts de transaction.
3. Un actif sans risque est disponible.
4. Les dividendes et les gains de capitaux ne sont pas taxés.
5. Pas d'influence sur les prix par les acheteurs et les vendeurs qui interviennent sur le marché.
6. le taux d'emprunt est égal au taux de prêt.

## 2.4.3 Le rôle de Bêta dans le MEDAF

Le modèle développé par Sharpe (MEDAF) consiste à mesurer le degré de sensibilité du rendement d'un actif par rapport à celui du marché. Ce modèle postule que la prime de risque attendue pour un actif est proportionnelle à son bêta. Cela implique que chaque titre doit se trouver sur la droite du MEDAF, reliant l'actif sans titre au portefeuille de marché.

Etant donné un portefeuille constitué de  $n$  actions de rendements  $r_1, r_2, \dots$  et  $r_n$ ; et un actif sans risque de rendement  $r_0$ .

Le rendement espéré de ce portefeuille est donné par :

$$\bar{r} = \sum_{i=1}^n x_i r_i$$

Sous ces conditions, un modèle linéaire exprime l'excès de rendement des actifs risqués (par rapport au rendement sans risque) en fonction du rendement  $R_M$  d'un portefeuille de marché :

pour tout  $i \in \{1, \dots, n\}$ ,

$$r_i - r_0 = \beta_i (R_M - r_0) + \varepsilon_i$$

avec  $\varepsilon_i$  un bruit indépendant de  $R_M$ . En espérance, on écrit aussi, avec  $\bar{R}_M = E(R_M)$ , où  $x_i$  représente la proportion investie dans l'action  $A_i$  pour  $i = 1, \dots, n$

La relation qui caractérise le modèle d'équilibre des actifs financiers MEDAF est donnée par :

$$\bar{r}_i = r_0 + \beta_i (\bar{R}_M - r_0)$$

où :

- $\bar{r}_i$  : le rendement espéré de l'action  $A_i$ .
- $\bar{R}_M$  : le rendement de portefeuille de marché.
- $\sigma_M^2$  : le risque de portefeuille de marché.
- $\beta_i = \frac{\sigma_{iM}}{\sigma_M^2}$ .
- $\sigma_{iM} = \sum_{j=1}^n x_j \sigma_{ij}$

## 2.5 Le modèle d'évaluation par arbitrage (APT)

Créé par Stephen Ross en 1976, le modèle APT (Arbitrage pricing theory) est l'un des plus célèbres modèles d'évaluation d'actifs financiers. C'est en quelque sorte le principal concurrent du modèle CAPM ou MEDAF et beaucoup plus réaliste par ses hypothèses et sa façon de traiter l'association rendement/risque d'un actif. Se fondant en particulier sur l'hypothèse selon laquelle les rentabilités des actifs financiers suivent une loi normale (ou que les investisseurs ont une fonction d'utilité quadratique) et sur le concept d'équilibre Ross (1976) a, en conséquence, tenté de mettre en place un modèle d'évaluation des actifs financiers dans lequel le portefeuille de marché n'intervient pas, et de façon plus générale, qui n'est pas basé sur l'équilibre des actifs financiers [44].

Cette théorie se base sur une absence d'opportunité d'arbitrage. Par conséquent, il devient impossible de réaliser des profits excessifs sur une longue période avec des actifs de faible risque. Cette nouvelle règle se vérifie sans cesse sur les marchés financiers. En effet, un actif A aussi risqué qu'un actif B, mais plus rentable, verrait sa demande augmenter rapidement, jusqu'à ce que sa rentabilité redevienne égale à celle de l'actif B, annulant ainsi toute opportunité d'arbitrage.

L'autre hypothèse de base de l'APT repose dans le fait que l'on peut modéliser la rentabilité espérée d'une action par une fonction linéaire de différents facteurs macro-économiques ou propres au secteur de la valeur, pondérés selon leur impact sur l'action par un coefficient beta spécifique. ROSS résulte de cette hypothèse une relation linéaire entre les rentabilités espérées des différents actifs et leurs covariances avec les facteurs [44].

Ainsi, pour un actif  $j$  on aura le modèle repose sur un modèle multifactoriel :

$$\begin{aligned} E(r_j) &= r_f + b_{j1}F_1 + b_{j2}F_2 + \dots + b_{jn}F_k \\ &= r_f + \epsilon_j + \sum_{i=1}^k b_{ji}F_i. \end{aligned}$$

avec :

- $E(r_j)$  : le rendement espéré de l'actif  $j$
- $r_f$  : le rendement de l'actif sans risque.
- $F_k$  : est la variation non anticipée du facteur  $k$ , et est donc d'espérance nulle, les  $F_k$  sont non corrélés entre eux
- $b_{jk}$  : désigne la sensibilité de la rentabilité du titre  $i$  au facteur commun  $F_k$ .
- $\epsilon_j$  : est le risque spécifique du titre  $i$ .

On considère un marché formé par  $n$  actifs financiers sur une période de temps  $[0, T]$ . On note  $R_j$  la rentabilité de l'actif  $j$  sur la période  $[0, T]$ ,  $j = 1, \dots, n$ .

Le modèle factoriel stipule que les rentabilités de tous les titres sont expliquées par les mêmes  $K$  *facteurs* représentés par des variables aléatoires  $F_1, \dots, F_K$ . Plus précisément, le modèle factoriel linéaire suppose

$$R_j = r_f + \epsilon_j + \sum_{i=1}^k b_{ji}F_i.$$

avec :

pour  $i \in \{1, \dots, n\}$ ,  $k, l \in \{1, \dots, k\}$ ,  $k \neq l$

$$\begin{aligned} E(\epsilon_j) &= 0 \\ E(F_k) &= 0 \\ cov(\epsilon_j, F_k) &= 0 \\ cov(\epsilon_i, \epsilon_j) &= 0 \end{aligned}$$

**Remarque** : Critiques du modèle d'évaluation par arbitrage :

- Les facteurs qui influencent l'actif ne sont pas déterminés dans le modèle. L'investisseur doit donc les déterminer lui même, ce qui demande technique et patience.
- L'influence d'un facteur sur un actif évolue dans le temps.
- Demande de lourds calculs, qui de plus ne sont pas à la portée de tous.

## 2.6 Le modèle de Markowitz et Perold (1981)

Le modèle à indice simple de Sharp peut être généralisé à un modèle à plusieurs indices qui offre à l'investisseur la possibilité d'investir sur un marché international où figure plusieurs indices boursiers notamment le marché européen, marché américain ou le marché asiatique [22].

Dans ce cadre, Markowitz et Perold [60] ont développé un modèle multi-indiciels qui suppose qu'il y a une relation entre les titres sous la forme suivante :

$$R_i = \alpha_i + \epsilon_i + \sum_{j=1}^k \beta_{ij} F_j, \quad i = 1, \dots, n$$

où :

- $F_k$  : représente le  $k^{\text{ème}}$  facteur aléatoire ;
- $\alpha_i, \beta_{ij}$  : sont des constants.
- $\epsilon_i$  : est un bruit aléatoire de moyenne 0 et non corrélé avec  $F_k$  (pour tout  $k = 1, \dots, K$ ).

Si  $\sigma_i^2 = E[\epsilon_i^2]$  et  $\gamma_{rs} = \text{cov}[F_k, F_s]$  alors on obtient la relation suivante :

$$\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j = \sum_{j=1}^n \sigma_j^2 x_j^2 + \sum_{i=1}^n \sum_{j=1}^n \sum_{r=1}^k \sum_{s=1}^k \gamma_{rs} \beta_{ir} \beta_{js} x_i x_j$$

Le programme d'optimisation est :

$$\min \left[ \sum_{i=1}^n \sigma_i^2 x_i^2 + \sum_{r=1}^k \sum_{s=1}^k \gamma_{rs} y_r y_s \right]$$

Sous les contraintes :

$$\begin{cases} \sum_{j=1}^n x_j E[r_j] \geq p_0 \\ \sum_{j=1}^n \beta_{jk} x_j - y_k = 0 \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 \quad j = 1, \dots, n \end{cases} \quad k = 1, \dots, K$$

## 2.7 Le modèle de Lai (1991)

Afin d'améliorer la qualité d'optimisation, nous expliquons l'approche multicritère de sélection des portefeuilles optimaux **MVS** proposée par Lai (1991). Son idée principale dans le développement d'un modèle MVS (mean-variance-skewness) réalisable en supposant le Skewness (l'asymétrie) et un actif sans risque pour rechercher un portefeuille maximisant à la fois le rendement et l'asymétrie, pour un niveau de variance donné [50].

Soit le portefeuille  $x = (x_1, \dots, x_n)$  où  $x_i$  désigne la proportion du capital  $C_0$  à investir par un investisseur dans les différents titres  $i = 1, \dots, n$  et soit  $x_{sr}$  un actif sans risque. Donc le portefeuille  $(x_1, \dots, x_n, x_{sr}) \in \mathbb{R}^{n+1}$  est un vecteur de proportions dans chacun de ces  $n$  actifs financiers et l'actif sans risque avec le rendement et le rendement espéré de ce portefeuille sont donnés respectivement par :

$$\sum_{i=1}^n x_i + x_{sr} = 1$$

Si un actif sans risque est absent ou ne peut pas être sélectionné

L'ensemble des portefeuilles admissibles est donné par :

$$\mathfrak{S} = \left\{ (x, x_{sr}) \in \mathbb{R}^{n+1}; \sum_{i=1}^n x_i + x_{sr} = 1 \right\}$$

Les actifs parmi lesquels l'investisseur fait un choix sont caractérisés par leurs rendements  $R_i$  pour  $i \in \{1, \dots, n\}$ , le rendement espéré  $E[R_i]$ , pour  $i \in \{1, \dots, n\}$ , peut être dérivée, ainsi que la matrice de covariance  $\Omega$  avec

$$\Omega_{ij} = Cov[R_i, R_j] = E[(R_i - E[R_i])(R_j - E[R_j])],$$

Pour  $i, j \in \{1, \dots, n\}$  Le Coskewness est le troisième moment central croisé standardisé, lié à l'asymétrie car la covariance est liée à la variance est donné comme suit :

$$\Lambda_{ijk} = E[(R_i - E[R_i])(R_j - E[R_j])(R_k - E[R_k])]$$

Le rendement du portefeuille  $(x, x_{sr})$  est défini par  $R(x, x_{sr}) = \sum_{i=1}^n x_i R_i + x_{sr} R_{sr}$ . Le rendement espéré du portefeuille  $(x, x_{sr})$

$$Z_1 = E[R(x, x_{sr})] = \sum_{i=1}^n x_i E[R_i] + x_{sr} R_{sr}$$

Les moments d'ordres deux et trois du portefeuille sont donnés respectivement par :

$$Var[R(x, x_{sr})] = E[(R(x, x_{sr}) - E[R(x, x_{sr})])^2] = \sum_{i,j=1}^n x_i x_j \Omega_{ij}$$

$$Z_3 = SK[R(x, x_{sr})] = E[(R(x, x_{sr}) - E[R(x, x_{sr})])^3] = \sum_{i,j,k=1}^n x_i x_j x_k \Lambda_{ijk}$$

Lai a proposé l'espace variance- unité  $\{x/V[R(x)] = 1\}$  pour l'optimisation de portefeuille. Selon cette hypothèse, l'algorithme d'optimisation multicritère à résoudre s'écrit comme suit :

$$\begin{aligned} & \max \{E[R(x, x_{sr})]\}, \\ & \max \{SK[R(x, x_{sr})]\} \end{aligned}$$

Sous les contraintes suivantes :

$$\begin{aligned} Var[R(x, x_{sr})] &= 1 \\ \sum_{i=1}^n x_i + x_{sr} &= 1 \\ x_i &\geq 0 \quad j = 1, \dots, n \end{aligned}$$

Pour faire l'optimisation du portefeuille, Lai utilise les techniques de programmation polynomiale **PGP** (Polynomial goal programming) :

**Définition 2.1** Pour des valeurs de paramètres données  $\alpha, \beta \in \mathbb{R}_+$  le modèle PGP est défini par :

$$PGP(\alpha, \beta) = \left\{ d_1^\alpha + d_2^\beta; d_1 = Z_1^* - Z_1, d_3 = Z_3^* - Z_3, Var[R(x, x_{sr})] = 1 \right\}$$

avec  $Z_1^*$  : le rendement espéré,  $Z_3^*$  valeur d'asymétrie espéré.

$$Z_1^* = \max_{(x, x_{sr}) \in \mathfrak{S}} \{Z_1; Var[R(x, x_{sr})] = 1\}$$

et

$$Z_3^* = \max_{(x, x_{sr}) \in \mathfrak{S}} \{Z_3; Var[R(x, x_{sr})] = 1\}$$

Lai d'après l'approche multicritère et propose l'algorithme suivant :

$$\min \left( d_1^\alpha + d_2^\beta \right)$$

Sous les contraintes suivantes :

$$\begin{aligned} Z_1 + d_1 &= Z_1^* \\ Z_3 + d_2 &= Z_3^* \\ Var[R(x, x_{sr})] &= 1 \\ \sum_{i=1}^n x_i + x_{sr} &= 1 \\ x_i &\geq 0 \quad j = 1, \dots, n \\ d_1, d_2 &\geq 0 \end{aligned}$$

où :

- $Z_1^*$  : est la valeur optimale de l'objective  $R(x)$  dans l'algorithme uni-critère,
- $Z_3^*$  : est la valeur optimale de l'objective  $\gamma(R(x))$  dans l'algorithme uni-critère,
- $d_1$  : est la variable positive représente l'écart entre  $R(x)$  et  $Z_1^*$ .
- $d_2$  : est la variable positive représente l'écart entre  $sk(R(x))$  et  $Z_3^*$
- $\alpha, \beta$  : paramètre subjectif positif qui mesure le degré de  $N$  préférence de l'investisseur entre les objectives.

## 2.8 Le modèle de Konno et Yamazaki (1991)

En raison de la difficulté de calcul associée au modèle de Markowitz, Koono et Yamazaki introduit une mesure alternative du risque, qui permettrait de résoudre facilement des problèmes à grande échelle [44].

Malgré son importance dans la théorie du portefeuille, le modèle d'optimisation de portefeuille de Markowitz a fait l'objet de critiques pour plusieurs raisons. En 1991, Hiroshi Konno et Hiroaki Yamazaki ont souligné trois défauts dans le modèle d'optimisation de portefeuille de Markowitz [43].

**Charge de calcul :** Le modèle de Markowitz a une fonction de coût quadratique qui peut être très longue à résoudre en fonction du nombre de stocks considérés. Konno et Yamazaki ont fait valoir que ce coût de calcul peut être peu pratique pour les investisseurs qui négocient en temps réel.

**Perception de l'investisseur :** L'utilisation de la variance comme mesure du risque semble pessimiste car des études sur le marché boursier de Tokyo ont montré que la plupart des rendements ne sont pas normalement distribués, de sorte que la moyenne et la variance peuvent ne pas être une représentation précise de la distribution des rendements d'une action.

**Coût de transaction / gestion et effet de coupure :** en raison du montant important des coûts de transaction associés à un très large portefeuille d'investissements. La gestion d'un très grand portefeuille peut également prendre beaucoup de temps. Pour essayer d'éviter ce problème, un investisseur peut choisir d'éliminer les actions du portefeuille optimal qui ont des pondérations plus petites ; cependant, cela fausse la solution optimale et peut se traduire par un portefeuille avec un risque plus élevé.

Konno a introduit l'approche de l'EAM en 1988. Au lieu d'utiliser les variances comme dans le cas de l'approche MV [38]. ici la mesure de la dispersion est basée sur les écarts absolus par rapport à la moyenne. Konno et Yamazaki ont introduit le risque  $L_1$  ou écart absolu moyen par rapport à la moyenne.

Cette fonction de risque est exprimé comme suit :

$$MAD(R) = E \left[ \left| \left( \sum_{j=1}^n x_j (R_j - E[R_j]) \right) \right| \right]$$

**Définition 2.2** *L"écart absolu moyen MAD pour un portefeuille de rendement R est*

$$MAD(R) = E(|R - E(R)|)$$

*Notez que MAD est une fonction de risque  $L_1$ . Le MAD est en général une fonction non convexe, non différentiable. Le MAD n'est pas une mesure de risque cohérente.*

**Théorème 2.1** *Si les rendements du portefeuille sont distribués normalement à plusieurs variables, alors*

$$MAD(R) = \sqrt{\frac{2Var(R)}{\pi}}$$

*indique que la minimisation de la variance, qui est une fonction de risque  $L_2$ , équivaut à minimiser MAD.*

Konno et Yamazaki ont illustré que sous l'hypothèse de la normalité des rendements des actifs, la mesure de risque de Konno est équivalent à celui de Markowitz. En utilisant cette mesure de risque linéaire,

Le modèle de Mean-Absolute Deviation (MAD) si les rendements sont distribués norma-

lement est comme suit :

$$\max \left[ \sum_{j=1}^n x_j R_j - \mu \frac{1}{T} \sum_{t=1}^T [y_t] \right]$$

$$S.C : \begin{cases} y_t - \sum_{j=1}^n x_j (R_j(t) - R_j) \geq 0; & t = 1, \dots, T \\ y_t + \sum_{j=1}^n x_j (R_j(t) - R_j) \geq 0; & t = 1, \dots, T \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 & j = 1, \dots, n \end{cases}$$

## 2.9 Le modèle de Speranza (1993)

Dans le modèle original de Markowitz [57], le risque de marché ou bien d'autre est mesuré par l'écart type ou la variance. Markowitz dans ce modèle classique utilise la variance comme mesure du risque, ce qui donne un problème d'optimisation quadratique. Suite aux travaux de Sharpe sur l'approximation linéaire du modèle de moyenne-variance, de nombreuses travaux ont été faites pour linéariser le problème d'optimisation du portefeuille. Il a été introduit plusieurs mesures de risque alternatives qui sont intéressantes sur le plan du calcul car (pour les variables aléatoires discrètes) elles permettent de résoudre des problèmes de programmation linéaire (LP). La solvabilité LP est très importante pour les applications aux décisions financières réelles. Alors que le modèle original de Markowitz est un problème de programmation quadratique, à la suite de Sharpe [81], de nombreuses tentatives ont été faites pour linéariser le problème d'optimisation de portefeuille.

En 1993 Speranza [84] a proposé une mesure de risque sous forme d'une combinaison linéaire entre l'écart absolu en-dessous de la moyenne et l'écart absolu au-dessus de la moyenne, le but de Speranza est d'améliorer le modèle de Konno et Yamazaki [22].

Cette mesure de risque est définie comme suit :

$$S(x) = \alpha E \left[ - \min \left\{ 0, \sum_{j=1}^n R_j x_j - E \left[ \sum_{j=1}^n R_j x_j \right] \right\} \right] + \beta E \left[ \max \left( 0, \sum_{j=1}^n R_j x_j - E \left[ \sum_{j=1}^n R_j x_j \right] \right) \right],$$

où  $\alpha$  et  $\beta$  sont les deux paramètres qui représentent les poids attribués à la fonction de risque permettant de mesurer le degré d'aversion au risque de l'investisseur.

Ce modèle peut prendre plusieurs formes selon les valeurs des paramètres  $\alpha$  et  $\beta$ .

Speranza a choisi  $\alpha = 1$  et  $\beta = 0$  afin de rendre son modèle plus efficace, dans ce cas la mesure de risque devient :

$$S(x) = E \left[ - \min \left\{ 0, \sum_{j=1}^n R_j x_j - E \left[ \sum_{j=1}^n R_j x_j \right] \right\} \right]$$

Alors l'algorithme d'optimisation est comme suit :

$$\frac{1}{T-1} \sum_{t=1}^T \left\{ -\text{Min} \left\{ 0, \sum_{j=1}^n (r_{jt} - \hat{r}_j) x_j \right\} \right\}$$

$$S.C. \begin{cases} \sum_{j=1}^n \hat{r}_j x_j \geq p \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 \quad j = 1, \dots, n \end{cases}$$

Cet algorithme est équivalent à l'algorithme suivant :

$$\text{Min} \left\{ \frac{1}{T-1} \sum_{t=1}^T \mu_t \right\}$$

$$S.C. \begin{cases} \mu_t - \sum_{j=1}^n (r_{jt} - \hat{r}_j) x_j \geq 0 \\ \sum_{j=1}^n \hat{r}_j x_j = \rho \\ \sum_{j=1}^n x_j = 1 \quad j = 1, \dots, n \\ \mu_t \geq 0 \quad t = 1, \dots, T \end{cases}$$

## 2.10 Le modèle Moyenne–Semi Variances de Hamza et Janssen (1995)

Lorsqu' il y a une asymétrie de l'information dans les données utilisées, les mesures de risque précédent (variance, écart absolu) sont symétriques, elles ne permettent pas de prendre en considération l'asymétrie de données. Pour remédier à ce problème, Hamza et Janssen ont proposé une mesure de risque définie par une combinaison convexe des deux semi-variances de rendement de portefeuille par rapport à sa rentabilité espérée. Ces deux semi-variances de rendement du portefeuille sont définies comme suit :

La première est celle qui mesure la variance en dessous de la moyenne :

$$E \{ \min(0, R(x) - E[R(x)]) \}^2$$

La deuxième est celle qui mesure la variance au-dessus de la moyenne :

$$E \{ \max(0, R(x) - E[R(x)]) \}^2$$

Afin de simplifier cet algorithme d'optimisation, Hamza et Janssen ont proposé une représentation plus simple en introduisant les variables auxiliaires suivantes :

$$u_t = - \min \left( 0, \sum_{j=1}^n (r_{jt} - \hat{r}_j) x_j \right) \quad \text{pour tout } t = 1, \dots, T$$

et

$$v_t = -\max \left( 0, \sum_{j=1}^n (r_{jt} - \hat{r}_j) x_j \right) \quad \text{pour tout } t = 1, \dots, T$$

L'algorithme d'optimisation précédent devient :

$$\text{Min} \left\{ \frac{1}{T-1} \sum_{t=1}^T (\alpha u_t^2 + \beta v_t^2) \right\}$$

Sous les contraintes suivantes :

$$\left\{ \begin{array}{l} \sum_{j=1}^n \hat{r}_j x_j \geq p \\ \sum_{j=1}^n x_j = 1 \\ \sum_{j=1}^n (r_{jt} - \hat{r}_j) x_j = v_t - u_t ; \quad t = 1, \dots, T \\ u_t v_t = 0 \quad t = 1, \dots, T \\ x_j \geq 0 \quad j = 1, \dots, n \\ u_t \geq 0 \text{ et } v_t \geq 0 \quad t = 1, \dots, T \end{array} \right.$$

## 2.11 Le modèle de YOUNG (1998)

Sharpe (1971) [81] a remarquer que «si le problème de l'analyse de portefeuille pouvait être correctement saisie sous une forme convenant aux méthodes de programmation linéaire, les perspectives d'application pratique s'en trouveraient grandement améliorées». Dans cet article de YOUNG [88].

le portefeuille optimal est défini comme celui qui minimiserait la perte maximale sur toutes les périodes historiques passées.

Le critère "Minimax" dans le problème d'optimiser un portefeuille d'actions transformer le problème en un problème d'optimisation qui maximise le rendement minimal du portefeuille ou minimise le risque maximal du portefeuille.

YOUNG[88] a introduit une fonction alternative de risque minimax dans l'optimisation du portefeuille. Cette fonction de risque est définie comme la moyenne des risques individuels maximums sur un certain nombre de périodes passées. La signification pratique de cette fonction de risque est de satisfaire l'objectif d'un investisseur par la minimisation de la moyenne des risques individuels maximum parmi les actifs à investir. Le problème d'optimisation de portefeuille correspondant est formulé comme un problème de programmation linéaire par morceaux à deux critères.

Dans cette partie, nous présentons notre mesure du risque et formulons le problème d'optimisation de portefeuille correspondant à cette mesure. Supposons qu'un investisseur dispose d'une capital initial  $M_0$ , qui doit être investie dans  $n$  actifs possibles  $S_j$ ,  $j = 1, \dots, n$ . Soit  $R_j$  une variable aléatoire représentant le taux de rendement de l'actif  $S_j$ , Soit  $x_j \geq 0$  l'allocation de  $M_0$  pour l'investissement à  $S_j$ , ( $x_j \geq 0$ , nous sommes concernés par la situation où la vente à découvert n'est pas autorisée). Ainsi, la région envisageable pour le problème d'optimisation du portefeuille est

$$\mathcal{F} = \left\{ x = \{x_1, \dots, x_n\} : \sum_{j=1}^n x_j = M_0, x_j \geq 0, j = 1, \dots, n \right\}$$

Soit  $E(R)$  l'espérance mathématique d'une variable aléatoire  $R$

$$r_j = E(R_j) \quad \text{and} \quad q_j = E(|R_j - r_j|)$$

$r_j$  et  $q_j$  désignent respectivement le taux de rendement attendu de l'actif  $S_j$  et l'écart absolu attendu de  $R_j$  par rapport à sa moyenne.

Le rendement attendu d'un portefeuille  $x = \{x_1, \dots, x_n\}$  est donné par :

$$x = \{x_1, \dots, x_n\} = E \left[ \sum_{j=1}^n R_j x_j \right] = \sum_{j=1}^n E(R_j) x_j = \sum_{j=1}^n r_j x_j$$

**Définition 2.3** la mesure de risque minimax, avec leur fonction de risque  $I_\infty$  est définie comme :

$$\omega_\infty(x) = \max_{1 \leq j \leq n} E(|R_j x_j - r_j x_j|) = \max_{1 \leq j \leq n} q_j x_j.$$

soit  $x \in \mathcal{F}$ . Alors

$$\omega_\infty(x) = \max_{1 \leq j \leq n} E(|R_j - E(R_j)|) x_j = \max_{1 \leq j \leq n} q_j x_j.$$

Dans l'hypothèse où un investisseur souhaite maximiser le rendement attendu et d'autre part minimiser le niveau de risque, ce problème d'optimisation vise deux critères en conflit, à savoir qu'un rendement plus élevé s'accompagne toujours d'un niveau de risque plus élevé. Pour cette raison, le problème d'optimisation de portefeuille peut être formulé comme un programme linéaire par morceaux à deux critères comme suit, qui est noté **POL** $_\infty$  (le problème d'optimisation de portefeuille avec la mesure de risque  $I_\infty$ ).

**Définition 2.4** Le programme d'optimisation d'un portefeuille linéaire par morceaux à deux critères **POL** $_\infty$  d'après la mesure de risque  $I_\infty$

$$\begin{aligned} \min \quad & \left( \max_{1 \leq j \leq n} q_j x_j, - \sum_{j=1}^n r_j x_j \right) \\ \text{S.c :} \quad & x \in \mathcal{F} \end{aligned}$$

On peut transformer **POL** $_\infty$  en un problème équivalent de programmation linéaire bi-critères (**BLP**). Cet algorithme d'optimisation est équivalent au programme linéaire suivant :

$$\begin{aligned} \min \quad & \left( y, - \sum_{j=1}^n r_j x_j \right) \\ \text{S.c :} \quad & \begin{cases} q_j x_j \leq y; & j = 1, \dots, n \\ x \in \mathcal{F}. \end{cases} \end{aligned}$$

Nous convertissons maintenant le problème de programmation linéaire à deux critères **BLP** en un problème d'optimisation paramétrique avec un seul critère. Pour un  $\lambda$  fixe,

où  $0 < \lambda < 1$ , le problème d'optimisation paramétrique de *BLP*, noté **PO** ( $\lambda$ ), et se présente comme suit :

$$\min \left[ \lambda x + (1 - \lambda) \left( - \sum_{j=1}^n r_j x_j \right) \right]$$

$$S.c : \begin{cases} q_j x_j \leq y; & j = 1, \dots, n \\ x \in \mathcal{F}. \end{cases}$$

**Proposition 2.1** *Considérez les problèmes **BLP** et **PO** ( $\lambda$ ). Le couple  $(x, y)$  est une solution efficace de **BLP** si et seulement s'il existe un  $\lambda \in (0, 1)$  tel que  $(x, y)$  soit une solution optimale de **PO**( $\lambda$ ). avec  $r_1 \leq r_2 \leq \dots \leq r_n$  et  $q_j > 0$ ;  $j = 1, \dots, n$*

On peut considérer  $\lambda$  comme le paramètre de tolérance au risque d'un investisseur, plus  $\lambda$  est grand, plus l'investisseur est risqué.

## 2.12 Le modèle Basée sur La valeur à risque (2000)

- Tous les modèles ne permettent pas de calculer de manière explicite la perte que pourra subir un investisseur ou une institution financière (établissement de crédit, banque, compagnie de l'assurance société cotée en bourse, . . .), d'où la naissance d'une nouvelle mesure de risque appelée valeur à risque (Value at Risk).

- La valeur à risque noté (VaR) ont utilisée pour la première fois dans les années 1980 par la **banque Bankers Trust** sur les marchés financiers américains, la notion de Value-At-Risk (ou VAR) a principalement été démocratisée par la **banque JP Morgan** dans les années 1990 grâce à son système de RiskMetrics.

- Jusqu'alors, les méthodes utilisées pour détecter et gérer les risques de marché ne permettaient pas de comparer les mesures de risque entre les différentes activités de marché. Le VaR calculer la perte probable de manière explicite.

- L'accroissement de la volatilité des marchés financiers, le développement des produits dérivés et surtout une série de faillites et de krachs boursiers ont poussé les institutions financières à mettre en place un indicateur commun et synthétique des risques financiers.

- Il faudra attendre 1995 et les accords de **Bâle**, pour voir la concrétisation et l'émergence de ce nouvel indicateur considéré très rapidement comme un standard dans l'évaluation des risques financiers.

**Définition 2.5** *La VaR est un cadre d'analyse qui permet d'agréger l'exposition du risque de marché à travers différentes classes de titres. En d'autres termes, la VaR est la prédiction de la perte maximale de la valeur d'un portefeuille.*

Le calcul de la VaR est basé sur trois paramètres :

► La distribution des pertes et profits du portefeuille en fin de période : le modèle peut retenir une distribution empirique ou une distribution connue (la plus usitée est la loi normale).

► La période de détention : qui correspond à la période jugée nécessaire pour vendre les titres et retourner sa position. Elle est donc liée à la liquidité des actifs.

► Le seuil de confiance  $\alpha$  : qui signifie que la probabilité d'occurrence d'une perte supérieure à la VaR est de  $(1 - \alpha)$ .

La Value-at-Risk correspond au montant des pertes qui ne devraient pas être dépassées pour un niveau de confiance donné sur un horizon temporel donné. le niveau de confiance choisi est en général de 95 ou 99%. Considérons :

- Un seuil de risque de  $\alpha\%$ , équivalent à un seuil de confiance de  $(1 - \alpha)\%$ .
- Un horizon temporel  $N$ .

Il s'agit donc de la perte maximale potentielle qui ne devrait être atteinte avec une probabilité donnée (ou un risque donnée) sur un horizon temporel donnée.

Mathématiquement, la notion de la Value-at-Risk se traduit ainsi :

$$\Pr(\Delta V < VaR) + 1 - c$$

Avec

- $\Delta V$  : La variation de la valeur  $V$  du portefeuille sur la période de détention.
- $c$  : Le niveau de confiance.

Pour un point de vue probabiliste, on considère la variable aléatoire  $X$  représentant la valeur de portefeuille. Alors la fonction de répartition  $F(X)$  de  $X$  est :

$$F_X(m) = \Pr(X \leq m)$$

C'est la probabilité que la valeur de portefeuille soit inférieur à un certain montant  $m$ .

La VaR correspond une probabilité supérieur  $(1 - \alpha)$

$$VaR_\alpha(X) = -\inf \{x / F_X(x) \geq 1 - \alpha\}$$

D'après la définition précédente la VaR représente la parte maximale qui pourrait se présenter dans un portefeuille (ou un action) sur une période  $[0; t]$  pour un risque  $\alpha$  comme suite :

d'après La méthode d'estimation classique Variance-Covariances. pour la VaR on obtient l'estimateur :

$$VaR_\alpha = V_0 \left\{ [-r_0 - \lambda(R_m(T) - 2r_0)] + Z_\alpha \sqrt{(\lambda\sigma_m)^2 + \sum_{i=1}^n \left[ \frac{V_i(0) x_i}{V(0)} \right]} \right\}.$$

Cette approche consiste à minimiser le risque de ce portefeuille donné par :

$$S.C : \begin{cases} \min [VaR_\alpha] \\ \sum_{j=1}^n x_j V_j \geq V_0 \\ \sum_{j=1}^n x_j = C_0 \\ x_j \geq 0 \quad j = 1, \dots, n \end{cases}$$

Où  $V_0$  est la valeur du portefeuille attendue par l'investisseur.

## 2.13 Le modèle Basée sur La VaR Conditionnelle

La VaR ne constitue pas une mesure de risque adéquate dans de nombreux cas à cause de la propriété de sous-additivité qu'elle ne parvient pas à respecter. Récemment, une mesure de risque est de plus en plus utilisée dans les établissements financiers. Cette mesure de risque est nommée La CVaR, (Valeur à risque Conditionnelle) ou encore **ES** (Expected Shortfall).

La CVaR possède en effet la caractéristique de sous-additivité et l'ensemble des autres caractéristiques (homogénéité, monotonie et invariance transitionnelle) qui furent mentionnées ; ce qui fait d'elle une mesure de risque valable dans tous les cas. CVaR qui n'est pas aussi populaire que VaR. Pourtant, CVaR possède certaines propriétés qui la rendent plus logique que VaR .

La CVaR est considérée comme le quantile correspondant à la perte potentielle que peut subir un titre ou un portefeuille suite à des mouvements défavorables des prix de marché avec un certain seuil de confiance sachant que cette perte dépasse la VaR. La CVaR est donc une mesure de risque plus pertinente et aussi puissante que la VaR, puisqu'elle tient compte des cas extrêmes et des événements de grands dommages encourus.

Si on considère deux portefeuilles  $P_1$  et  $P_2$ , la sous-additivité de la mesure de risque est définie par  $\rho(P_1 + P_2) \leq \rho(P_1) + \rho(P_2)$ . La VaR n'est pas sous-additive d'après [S. Basak and A. Shapiro. Value-at-risk-based risk management : optimal policies and asset prices et des portefeuilles peuvent afficher une VaR plus élevée alors qu'ils sont mieux diversifiés. Afin de compenser ce défaut, on utilise la CVaR définie comme :

$$CVaR(P, \alpha) = E [R_p / R_p < VaR(P, \alpha)]$$

Cette mesure à notamment été étudiée par *Artzner et al.* [5],

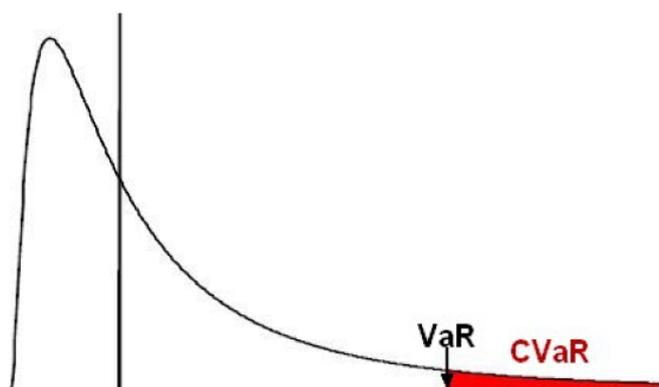


FIGURE 3.4 – La Valeur à Risque conditionnelle

Ainsi, nous pouvons dire que la VaR mesure la valeur qui separe les  $(1 - \alpha)\%$  de la distribution,

On peut exprimer La CVaR comme suit :

$$CVaR_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_\phi(X) d\phi$$

En se basant sur le critère « CVaR », l'algorithme d'optimisation de portefeuille est donné par :

- Considérons un portefeuille  $(x_1, x_2, \dots, x_n)$  où  $x_i$  désigne la proportion du capital  $M$  investie dans un actif de valeur  $V_i$  ( $i = 1, \dots, n$ )
- cette approche consiste à minimiser le risque de ce portefeuille donné par :

$$\min(CVaR_\alpha(X))$$

- Sous les contraintes :

$$\begin{aligned} \sum_{j=1}^n x_j V_j &\geq V_0 \\ \sum_{j=1}^n x_j &= M \\ x_j &\geq 0 \quad j = 1, \dots, n \end{aligned}$$

# Métaheuristiques d'optimisation : Etat de l'art

## 3.1 Notions de Base en Optimisation

Il y a Deux types de problèmes d'optimisation sont distingués : des problèmes de minimisation et des problèmes de maximisation. on définit un problème d'optimisation (de minimisation ou de maximisation) par un ensemble de données et un ensemble de contraintes.  $S$  est un ensemble de solutions associé au problème d'optimisation.  $X \subseteq S$  un sous-ensemble représente des solutions réalisables respectant les contraintes  $C$  du problème, à chaque solution  $s$  est associée une valeur  $f(s)$  qui représente sa qualité. La résolution du problème d'optimisation consiste à trouver une solution  $s^* \in X$  qui minimise ou maximise la valeur  $f(s)$ .

Dans ce qui suit nous présentons quelques définitions tirées de la littérature liées aux problèmes d'optimisation en général

**Définition 3.1** Une instance d'un problème de minimisation (maximisation) est un couple  $(X, f)$ . Où  $X \subseteq S$  est un ensemble fini de solutions potentielles admissibles et  $f$  est une fonction du coût (fonction objectif) à minimiser (à maximiser),  $f : X \rightarrow \mathbb{R}$ . L'objectif est de trouver  $s^* \in X$  tel que  $f(s^*) \leq f(s)$  (Au cas de maximisation :  $f(s^*) \geq f(s)$ ) pour n'importe quelle solution  $s \in X$

**Définition 3.2** Un problème d'optimisation combinatoire est défini par l'ensemble  $S$  des solutions possibles d'un problème.  $X \subseteq S$  est l'ensemble des solutions réalisables.  $f : X \rightarrow \mathbb{R}$ , une fonction que l'on nomme fonction «objectif». La résolution du problème consiste à minimiser (maximiser) la valeur  $f(s)$ , où  $s \in X$ .

**Définition 3.3** Un problème mono-objectif est décrit par la formule 1.1

$$\text{Tel que } \begin{cases} \text{Mode}_x f(s) \\ C_i(s) \Delta 0, i = 1 \dots m \\ H_i(s) = 0, j = 1 \dots p \\ s \in S \subset \mathbb{R}^n, \end{cases} \quad (1.1)$$

où

$$Mode = \begin{cases} \text{Min au cas de problème de minimisation} \\ \text{Max au cas de problème de maximisation} \end{cases}$$

$$\Delta \text{ est remplacé par } \begin{cases} \leq \text{ au cas de problème de minimisation} \\ \geq \text{ au cas de problème de maximisation} \end{cases}$$

$f$  est la fonction du coût (la fonction «objectif») à minimiser (ou à maximiser).  $C_i$  sont des contraintes d'inégalité et  $H_i$  sont des contraintes d'égalité.  $S$  est l'ensemble de solutions possibles (l'espace de recherche).  $s$  est une solution admissible, elle appartient à  $S$  et respecte les contraintes du problème.

**Définition 3.4** *Un problème multi-objectif est décrit par la formule (1.2)*

$$\text{Tel que } \begin{cases} Mode_x \vec{f}(s) \\ \vec{C}_i(s) \Delta 0, i = 1 \dots m \\ \vec{H}_i(s) = 0, j = 1 \dots p \\ s \in S \subset \mathbb{R}^n, \end{cases} \quad (1.2)$$

où

$$Mode = \begin{cases} \text{Min au cas de problème de minimisation} \\ \text{Max au cas de problème de maximisation} \end{cases}$$

$$\Delta \text{ est remplacé par } \begin{cases} \leq \text{ au cas de problème de minimisation} \\ \geq \text{ au cas de problème de maximisation} \end{cases}$$

Dans le cas de problèmes d'optimisation multi-objectifs, la fonction «objectif» est représentée par un vecteur  $\vec{f}$  regroupant  $N$  fonctions «objectif». Aussi les contraintes sont représentées par des vecteurs regroupant les contraintes de chaque fonction «objectif».

**Définition 3.5** *Une solution d'un problème d'optimisation est un ensemble de quantités souvent numériques pour lesquelles des valeurs sont à choisir. L'ensemble de ces valeurs est généralement regroupé dans un vecteur représentant une solution. Supposant un problème de taille  $n$ , le vecteur représentant la solution  $s$  est représenté par :*

$$\vec{s} = [s_1, s_2, s_3, \dots, s_n] \quad (1.3)$$

Les différentes valeurs prises par les variables  $(s_1, s_2, s_3, \dots, s_n)$  constituent l'ensemble des solutions envisageables. Selon le type des variables  $(s_1, s_2, s_3, \dots, s_n)$ , le type du problème d'optimisation peut être reconnu. Par conséquent, on distingue deux classes de problèmes d'optimisation : des problèmes continus et des problèmes discrets. Dans la première classe (*i.e.* la classe des problèmes continus), les variables composant une solution donnée sont de type réel. Tandis que, dans les problèmes de la deuxième classe (*i.e.* la classe des problèmes discrets), les variables composant une solution donnée peuvent être de type entier, naturel ou binaire.

**Définition 3.6** Une contrainte d'un problème est une restriction imposée par la nature et les caractéristiques du problème sur les solutions proposées.

**Définition 3.7** L'espace de recherche  $S$  d'un problème est composé de l'ensemble de valeurs pouvant être prises par les variables  $(s_1, s_2, s_3, \dots, s_n)$  qui construisent la solution  $s$ .

**Définition 3.8** Le voisinage  $V(s)$  d'une solution  $s$  est un sous ensemble de  $S$ , dont les membres sont des solutions proches (voisines) de la solution  $s$ . En effet, on dit qu'une solution «  $s'$  » est une voisine de  $s$ , si elle peut être obtenue en modifiant légèrement la solution  $s$ .

**Définition 3.9** L'optimum local est la meilleure solution appartenant à un voisinage  $V(s)$  de la solution  $s$ . Une solution  $s'$  (appartenant à  $S$ ) est un optimum local de la structure du voisinage  $V(s)$  de la solution  $s$  si elle vérifie la condition suivante (1.4) :

$$f(s') \leq f(s) \quad \forall s \in V(s) \quad (1.4)$$

Pour un problème de maximisation l'inégalité est inversée, c'est-à-dire la condition (1.4) sera remplacée par la condition (1.5) :

$$f(s') \geq f(s) \quad \forall s \in V(s) \quad (1.5)$$

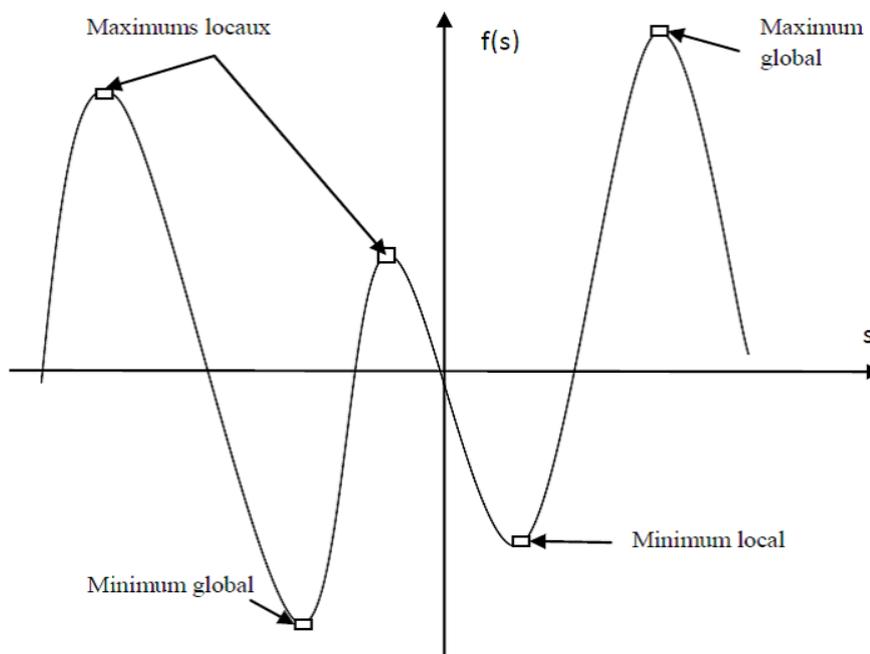


FIGURE 3.1: Courbe représentant les optimums locaux et les optimums globaux

**Définition 3.10** Une solution optimale (L'optimum global) : Une solution  $s^*$  est dite optimale (ou optimum global) si les deux contraintes suivantes sont vérifiées :

- Elle est réalisable : tirée de l'ensemble de solutions possibles (l'espace de recherche  $S$ ) et respecte toutes les contraintes du problème posé.

$$\begin{cases} f(s^*) = \max_{s \in X} \{f(s)\} & \text{En cas de problème de maximisation} \\ f(s^*) = \min_{s \in X} \{f(s)\} & \text{En cas de problème de minimisation} \end{cases}$$

Où  $X$  est l'ensemble de solutions réalisables.

Autrement dit, l'optimum global est le meilleur optimum local. Ainsi, une solution est dite optimum global si :

$$\begin{cases} f(s^*) \leq f(s) \forall s \in E(S) & \text{au cas de problème de maximisation} \\ f(s^*) \geq f(s) \forall s \in E(S) & \text{au cas de problème de minimisation} \end{cases}$$

Où  $E(S)$  est l'ensemble d'optimums locaux. Il est à noter que l'optimum global est aussi nommé maximum global au cas de problème de maximisation et minimum global au cas de problème de minimisation.

La figure (1.1) représente une courbe représentant les optimums locaux et les optimums globaux d'une fonction d'évaluation.

## 3.2 Problème D'optimisation

1) Un problème d'optimisation est défini par :

- un espace de recherche (de décision) : ensemble de solutions ou de configurations constitué des différentes valeurs prises par les variables de décision.
- une ou plusieurs fonction(s) dite objectif(s), à optimiser (minimiser ou maximiser).
- un ensemble de contraintes à respecter.

2) Un problème d'optimisation est caractérisé par :

- le domaine des variables de décision : soit Continu et on parle alors de problème continu, soit discret et on parle donc de problème combinatoire.
- la nature de la fonction objectif à optimiser : soit linéaire et on parle alors de problème linéaire, soit non linéaire et on parle donc de problème non linéaire.
- le nombre de fonctions objectifs à optimiser : soit une fonction scalaire et on parle alors de problème mono-objectif, soit une fonction vectorielle et on parle donc de problème multiobjectif.
- la présence ou non des contraintes : on parle de problème sans contrainte ou avec contrainte.
- sa taille : problème de petite ou de grande taille.
- L'environnement : problème dynamique (la fonction objectif change dans le temps).

3) Face à un problème d'optimisation :

- Elaborer un modèle (mathématiques) : l'expression de l'objectif à optimiser et les contraintes à respecter.
- Développer un algorithme de résolution.
- Evaluer la qualité des solutions produites.

### 3.3 Les méthodes de résolution de problèmes d'optimisation

La résolution de différents types de problèmes dans notre vie quotidienne a poussé la recherche pour proposer des méthodes de résolution fiable et à réaliser de grands efforts par les chercheurs pour améliorer leurs performances en termes de temps de calcul nécessaire et/ou de la qualité de la solution proposée. Au fil des années, de nombreuses méthodes de résolution de problèmes de différentes complexités ont été proposées. Ainsi, une grande variété et des différences remarquables au niveau du principe, de la stratégie et des performances ont été discernées. Cette variété et ces différences ont permis de regrouper les différentes méthodes de résolution de différents problèmes en deux classes principales : la classe de méthodes exactes et la classe de méthodes approchées. L'hybridation des méthodes de ces deux classes a donné naissance à une pseudo classe qui englobe des méthodes dites hybrides .

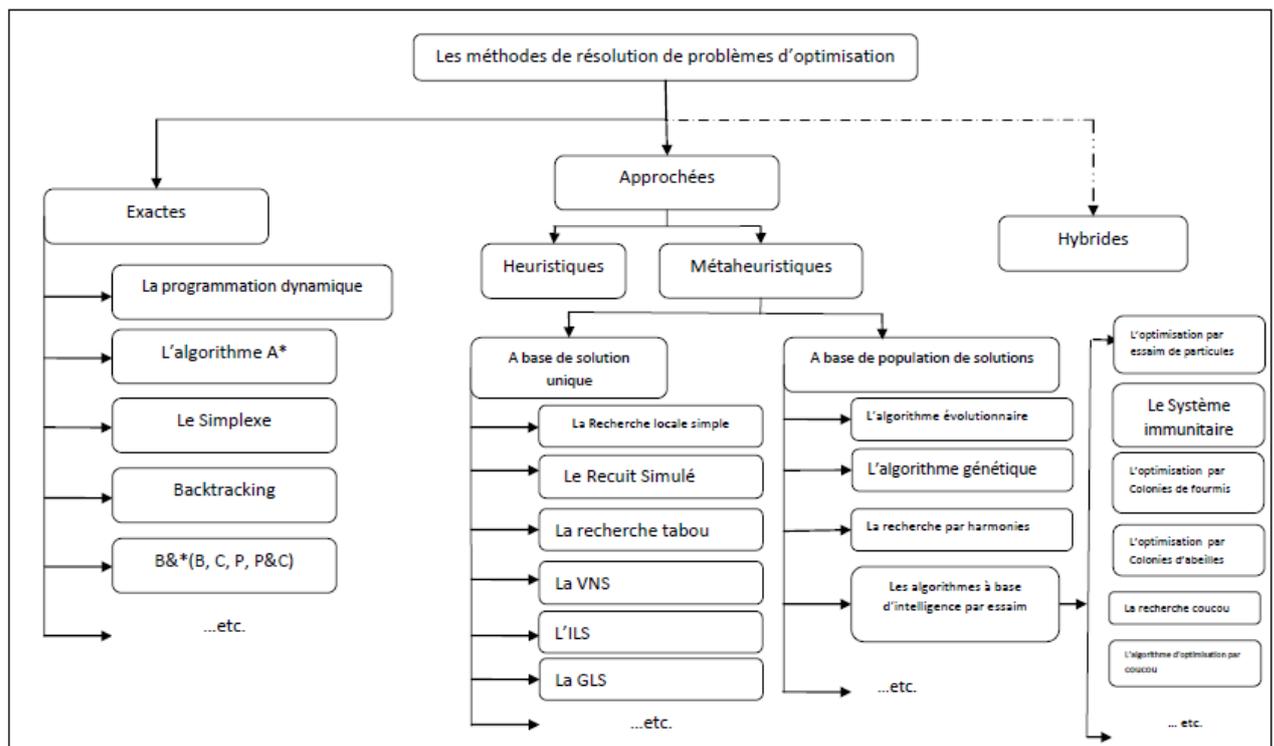


FIGURE 3.2: Classification de méthodes de résolution de problèmes d'optimisation.

#### 3.3.1 Les méthodes exactes

Les méthodes exactes cherchent à trouver de manière certaine la solution optimale en examinant de manière explicite ou implicite la totalité de l'espace de recherche. elles permettent de parcourir la totalité de l'ensemble de l'espace de recherche de manière à assurer l'obtention de toutes les solutions ayant le potentiel d'être meilleures que la solution opti-

male trouvée au cours de la recherche. Les méthodes exactes sont connues aussi par le fait qu'elles garantissent l'optimalité de la solution mais elles sont très gourmandes en termes de temps de calcul et de l'espace mémoire nécessaire. Cependant, sont très connues par le fait qu'elles nécessitent un coût de recherche souvent prohibitif en termes de ressources requises. C'est la raison pour laquelle, elles sont beaucoup plus utilisées pour la résolution de problèmes faciles.

Il existe de nombreuses algorithmes exacts y compris : les algorithmes de séparation et évaluation (Branch and Bound-Branch and Cut-Branch and Price - Branch and Cut and Price), les algorithmes de retour arrière (Backtracking) ; l'algorithme du simplexe, l'algorithme de Johnson [Johnson, 1954] pour la résolution de problèmes d'ordonnement. Notre vocation de citer quelques unes dans ce qui suit.

### 3.3.1.1 L'algorithme de retour arrière (Backtracking)

Appelé aussi backtracking en anglais. Est un algorithme d'énumération le plus classique et général pour trouver toutes (ou certaines) solutions à certains problèmes de calcul. Notamment les problèmes de satisfaction de contraintes. Son principe à revenir légèrement en arrière sur des décisions prises afin de sortir d'un blocage. autrement dit ; traverser les valeurs des variables par instanciation en remettant à cause la dernière affectation si une contrainte du problème est violée. Afin d'aboutir à une configuration consistante, une nouvelle valeur est affectée à la variable  $x_i$  en respectant son domaine de définition. Le processus est répété à chaque itération jusqu'à l'émergence d'une solution complète. Si toutes les possibilités ont été essayées sans donner de bonnes solutions, dans ce cas là on dit que le problème est irréalisable. Grâce à son principe qui permet une exploitation complète de l'espace de recherche, cet algorithme garantit l'aboutissement à la solution optimale. Néanmoins, il est très lent : si on se trouve devant un problème de  $n$  variables ayant chacune un domaine de définition de  $k$  valeurs, il y a  $(kn)$  combinaisons possibles et donc une complexité et un temps de recherche exponentiels et peut être irraisonnable (dans le cas de problèmes de grande taille).

### 3.3.1.2 La méthode Branch and Bound (B&B)

La méthode par séparation et évaluation (nommée Branch and Bound en anglais) est notée B&B Comme son nom l'indique, l'algorithme possède deux temps :

- la séparation : séparer un ensemble de solutions en sous-ensembles ;
- l'évaluation : évaluer les solutions d'un sous-ensemble en majorant la valeur de la meilleure solution de ce sous-ensemble.

C'est une méthodes qui permettent la résolution exacte de problèmes d'optimisation, notamment les problèmes d'optimisation combinatoires dont on cherche à minimiser le coût de la recherche. B&B propose un mécanisme de recherche très intelligent, grâce à lequel elle permet une bonne exploitation de l'espace de recherche et l'aboutissement à la solution optimale plus rapidement que d'autres méthodes exactes.

L'algorithme de Branch-and-Bound (B&B) consiste à construire nœud par nœud l'arbre d'énumération de l'ensemble des solutions du modèle  $P$  d'une manière intelligente, tout en faisant usage des bornes inférieure et supérieure afin d'éviter la génération de tous les

nœuds de l'arbre. Il s'agit principalement de diviser récursivement le problème initial en sous-problèmes plus petits et plus faciles à résoudre : c'est la phase de séparation (branching). Quand à la phase d'évaluation (bound), elle consiste à identifier les sous-ensembles qui peuvent contenir la solution optimale et à supprimer ceux qui ne la contiennent pas.

L'utilisation de la méthode B&B nécessite :

- Une solution initiale permettant d'entamer la recherche.
- Une stratégie permettant la division du problème  $P$  en sous problèmes  $P_i$ .
- Une fonction permettant le calcul des différentes bornes.
- Une stratégie de parcours de l'arbre : parcourir en profondeur, en largeur... etc.

L'efficacité de la méthode B&B a attiré l'attention de nombreux chercheurs. Par conséquent, plusieurs améliorations de l'algorithme B&B ont été proposées, y compris les algorithmes : Branch and Cut (noté B&C) [Padberg et Rinaldi, 1991], Branch and Price (noté B&P) [Barnhart et al, 2005], Branch and Cut and Price (B&C&P) Le branch and bound est parfois comparé à une autre technique de recherche de solution, l'algorithme A\*, très souvent utilisé en intelligence artificielle, alors que le branch and bound est plutôt destiné aux problèmes de recherche opérationnelle [30].

### 3.3.2 Les méthodes approchées

Méthodes souvent inspirées de mécanismes d'optimisation rencontrés dans la nature comme en biologie de l'évolution (cas des algorithmes génétiques). Elles sont utilisées pour les problèmes où on ne connaît pas d'algorithmes de résolution en temps polynomial et pour lesquels on espère trouver une solution approchée de l'optimum global. Elles cherchent à produire une solution de meilleure qualité possible dictée par des heuristiques avec un temps de calcul raisonnable en examinant seulement une partie de l'espace de recherche.

Les méthodes approchées (i.e. Heuristiques) permettent de trouver de manière rapide une solution réalisable à un problème donné. Cependant cette solution n'est pas forcément la solution optimale. L'objectif d'une heuristique est donc de trouver une solution la plus proche possible de celle d'une méthode exacte tout en étant plus rapide. La qualité d'une méthode approchée va donc se calculer par rapport à l'écart obtenu entre sa solution et l'optimale. Plus ce résultat est proche de la solution optimale, meilleure est l'heuristique.

#### 3.3.2.1 Les méthodes Heuristiques :

Sont des méthodes spécifiques à un problème particulier. Elles nécessitent des connaissances du domaine du problème traité. En fait, ce sont des règles empiriques qui se basent sur l'expérience et les résultats acquis afin d'améliorer les recherches futures. Une heuristique est un algorithme qui fournit rapidement (en un temps polynomial) une solution approchée et réalisable, pas nécessairement optimale, pour un problème d'optimisation difficile. Cette méthode approximative est le contraire d'un algorithme exact qui donne une solution optimale pour un problème donné.

Un raisonnement ou une méthode heuristique (ou une heuristique) est une méthode de résolution de problème qui ne s'appuie pas sur une analyse détaillée ou exhaustive du problème. Elle consiste à fonctionner par approches successives en s'appuyant, par

exemple, sur des similitudes avec des problèmes déjà traités afin d'éliminer progressivement les alternatives et ne conserver qu'une série limitée de solutions pour tendre vers celle qui est optimale.

Les heuristiques peuvent être classées en deux catégories :

- **Méthodes constructives** : Essayez d'accéder à la solution complète pour ajouter des éléments progressivement à partir de la solution primaire
- **Méthodes de fouilles locales** : Travailler pour améliorer la solution complète primaire en explorant son voisinage par de manière répétitive.

Plusieurs définitions des heuristiques ont été proposées par plusieurs chercheurs dans la littérature, parmi lesquelles.

**Définition 3.11** « Une heuristique (règle heuristique, méthode heuristique) est une règle d'estimation, une stratégie, une astuce, une simplification, ou tout autre type de dispositif qui limite considérablement la recherche de solutions dans des espaces problématiques importants. Les heuristiques ne garantissent pas des solutions optimales. En fait, elles ne garantissent pas une solution du tout. Tout ce qui peut être dit d'une heuristique utile, c'est qu'elle propose des solutions qui sont assez bonnes la plupart du temps ».

**Définition 3.12** « Une méthode heuristique (ou simplement une heuristique) est une méthode qui aide à découvrir la solution d'un problème en faisant des conjectures plausibles mais faillible de ce qui est la meilleure chose à faire.».

**Définition 3.13** Une heuristique est une règle d'estimation, une stratégie, une méthode ou astuce utilisée pour améliorer l'efficacité d'un système qui tente de découvrir les solutions des problèmes complexes ».

**Définition 3.14** « Les heuristiques sont des règles empiriques et des morceaux de connaissances, utiles (mais non garanties) pour effectuer des sélections différentes et des évaluations ».

**Définition 3.15** « Les heuristiques sont des ensembles de règles empiriques ou des stratégies qui fonctionnent, en effet, comme des règles d'estimation ».

**Définition 3.16** « Les heuristiques sont des critères, des méthodes ou des principes pour décider qui, parmi plusieurs d'autres plans d'action promet d'être le plus efficace pour atteindre un certain but ».

**3.3.2.1.1 Algorithme du recuit simulé** : L'origine de la méthode du recuit simulé provient de la métallurgie, a été proposé par Kirkpatrick, Gelatt et Vecchi [Kirkpatrick et al, 1983] pour atteindre les états de basse énergie d'un solide, on chauffe celui-ci jusqu'à des températures élevées, avant de le laisser refroidir lentement. Ce processus est appelé le recuit.

Le recuit simulé repose sur l'algorithme de Metropolis [Metropolis et al, 1953; Hastings, 1970] . Cette procédure permet de sortir des minima locaux avec une probabilité élevée

si la température  $T$  est élevée et, quand l'algorithme atteint de très basses températures, de conserver les états les plus probables .

Ces derniers chauffe à blanc le métal, puis ils laissent l'alliage se refroidir très lentement afin d'aboutir à un alliage sans défauts. En fait, les thermodynamiciens ont remarqué qu'une baisse brutale de la température d'un liquide entraîne une reproduction d'un optimum local, i.e. une structure amorphe. Alors qu'une baisse progressive de la température du liquide permet d'aboutir à un optimum global, i.e. une structure bien construite. C'est l'idée prise en considération par les métallurgistes qui savent que si le métal refroidit trop vite, il contiendra beaucoup de défauts microscopiques et s'il refroidit lentement ils obtiendront une structure bien ordonnée. La métaheuristique du recuit simulé s'inspire de l'algorithme de Métropolis [Métropolis et al, 1953], dont le principe (pour un problème de maximisation) peut être résumé comme suit :

- Entamer la recherche avec une solution initiale  $s$  ;
- Affecter une valeur initiale à la température  $T$  ;
- Calculer la fitness  $f(s)$  de la solution initiale  $s$ .
- Générer une solution  $\acute{s}$  voisine de  $s$  ;
- Calculer la fitness  $f(\acute{s})$  de  $\acute{s}$ .

Un schéma général de l'algorithme du recuit simulé (RS) est présenté comme suit :

Algorithme :	Le <b>recuit simulé</b>
<b>Debut</b> : Construire une solution initiale $S$ ;	
Calculer la fitness $f(S)$ de $S$ ;	
Initialiser une valeur de la température $T$ ;	
$S_{best} = s$ ;	
<b>Tant que</b> la condition d'arrêt n'est pas satisfaite faire	
Générer une solution $S^*$ voisine de $s$ ;	
Calculer $f(S^*)$ ;	
Calculer $\Delta(f) = f(S^*) - f(S)$ ;	
<b>Si</b> $\Delta(f) \geq 0$ <b>alors</b> // cas de maximisation	
$S_{best} = S^*$ ;	
$S = S^*$ ;	
<b>Sinon Si</b> $r < \exp\left(\frac{\Delta(f)}{T}\right)$ <b>alors</b>	
$S = S^*$ ;	
<b>Fin Si</b>	
Décroître la température $T$ ;	
<b>Fin Tant que</b>	
Retourner $S_{best}$ ;	
<b>Fin</b>	

**Les inconvénients du recuit simulé** résident dans :

1. Les « réglages », car l'algorithme dispose d'un nombre élevé de paramètres (température initiale, règle de décroissance de la température, durées des paliers de température, etc.) qui rendent les réglages de l'algorithme assez empiriques ;

2. Les « temps de calcul », qui peuvent devenir très importants. est un algorithme lent surtout avec les problèmes de grande taille.

il existe des études qui s'attachent à régler de manière optimale les paramètres de l'algorithme [Cour 94]. Par ailleurs, pour surmonter le problème de temps de calcul, plusieurs méthodes de parallélisation des calculs ont été introduites. la méthode du recuit simulé a l'avantage d'être souple vis-à-vis des évolutions du problème et facile à implémenter.

**3.3.2.1.2 Algorithme de recherche tabou :** L'algorithme de Recherche Tabou a été introduit par Glover en 1986. Est une métaheuristique itérative à base d'une solution unique. Le but de cette méthode est d'inculquer aux méthodes de recherche locale un surcroît- un ensemble de règles et de mécanismes généraux pour guider la recherche de manière intelligente L'idée de la recherche tabou consiste, à partir d'une position donnée, à en explorer le voisinage et à choisir le voisin qui minimise la fonction objectif est de pouvoir poursuivre la recherche de solutions même lorsqu'un optimum local est rencontré et ce, en permettant des déplacements qui n'améliorent pas la solution, et en utilisant le principe de la mémoire pour éviter les retours en arrière (mouvements cycliques). Comme l'algorithme de recuit simulé, la méthode de recherche tabou fonctionne avec une seule configuration courante, qui est actualisée au cours des itérations successives.

La nouveauté ici est que, pour éviter le risque de retour à une configuration déjà visitée, on tient à jour un liste de mouvements interdits, appelée "liste tabou" . Cette liste contient  $m$  mouvements ( $t \rightarrow s$ ) qui sont les inverses des  $m$  derniers mouvements ( $s \rightarrow t$ ) effectués. L'algorithme modélise ainsi une forme primaire de mémoire à court terme. L'algorithme de recherche tabou peut être résumé par l'Algorithme :

Algorithme :	<b>La recherche Tabou</b>
<b>Déterminer</b> une configuration aléatoire $s$	
<b>Initialiser</b> une liste tabou vide	
<b>Tant que</b> le critère d'arrêt n'est pas atteint <b>faire</b>	
<b>Perturbation</b> de $s$ suivant $N$ mouvements non tabous	
<b>évaluation</b> des $N$ voisins	
<b>Sélection</b> du meilleur voisin $t$	
<b>Actualisation</b> de la meilleure position connue $s^*$	
<b>Insertion</b> du mouvement $t \rightarrow s$ dans la liste tabou	
$s = t$	
<b>Fin Tant que.</b>	

La recherche tabou a été largement utilisée. C'est une méthode facile à mettre en oeuvre. Dans sa forme de base, l'algorithme de recherche tabou présente l'avantage de comporter moins de paramètres que l'algorithme de recuit simulé de manière à ne pas être très exigeante en temps de parcours et d'espace mémoire requis. Cependant, l'algorithme n'étant pas toujours performant, il est souvent approprié de lui ajouter des processus d'intensification et/ou de diversification, qui introduisent de nouveaux paramètres de contrôle [Glov1997].

### 3.3.2.2 Les méthodes dites métaheuristiques

Les métaheuristiques sont une famille d'algorithmes stochastiques. Sont des méthodes générales, des heuristiques polyvalentes applicables sur une grande gamme de problèmes. Elles peuvent construire une alternative aux méthodes heuristiques lorsqu'on ne connaît pas l'heuristique spécifique à un problème donné. Selon la définition proposée en 1996 par Osman et Laportes.

Dans la vie pratique, on se retrouve souvent confronté à des problèmes de différentes complexités, pour lesquelles on cherche des solutions qui satisfont deux notions antagonistes : la rapidité et la qualité.

Les métaheuristiques forment une famille d'algorithmes d'optimisation visant à résoudre des problèmes d'optimisation difficile, pour lesquels nous ne connaissons pas de méthodes classiques plus efficaces. Elles sont généralement utilisées comme des méthodes génériques pouvant optimiser une large gamme de problèmes différents, d'où le qualificatif méta.

L'efficacité d'une métaheuristique dans la résolution d'un problème d'optimisation est liée à sa capacité d'établir un certain équilibre entre l'exploitation des expériences acquises (i.e. les solutions trouvées) au cours de la recherche et l'exploration de l'espace de recherche pour identifier d'autres solutions de très bonne qualité. Il est à noter que les termes « exploitation » et « exploration » peuvent être remplacés par les termes « intensification » et « diversification » respectivement. Selon Glover et Laguna, la différence principale entre l'intensification et la diversification réside dans le fait qu'une phase d'intensification se concentre sur le test du voisinage d'une solution élue. Tandis qu'une phase de diversification encourage le processus de recherche à examiner des régions non visitées et à découvrir des solutions différentes des solutions rencontrées dans des points divers [33].

La particularité qui différencie les méthodes métaheuristiques des méthodes heuristiques c'est que les métaheuristiques sont applicables sur de nombreux problèmes. Tandis que, les heuristiques sont spécifiques à un problème donné. De nombreuses métaheuristiques ont été proposées dans la littérature afin de faire face aux différents problèmes en minimisant le coût de la recherche. Les métaheuristiques peuvent être classées en deux catégories : les méthodes à base de solution unique et les méthodes à base de population de solutions.

La majorité des métaheuristiques sont inspirées des systèmes naturels, nous pouvons citer à titre d'exemple : le recuit simulé qui est inspiré d'un processus métallurgique, les algorithmes évolutionnaires et les algorithmes génétiques qui sont inspirés des principes de l'évolution Darwinienne et de la biologie, la recherche tabou qui s'inspire de la mémoire des êtres humains, les algorithmes basés sur l'intelligence d'essaim comme l'algorithme d'optimisation par essaim de particules, l'algorithme de colonies de fourmis, l'algorithme de colonies d'abeilles, la recherche coucou et l'algorithme d'optimisation par coucou qui s'inspirent du comportement social de certaines espèces évoluant en groupe.

**Définition 3.17** « Une métaheuristique est un processus itératif qui subordonne et guide une heuristique, en combinant intelligemment plusieurs concepts pour explorer et exploiter tout l'espace de recherche. Des stratégies d'apprentissage sont utilisées pour structurer l'information afin de trouver efficacement des solutions optimales, ou presque-optimales. » [Osman et Laporte, 1996]

### 3.3.2.3 Les algorithmes évolutionnaires

Les algorithmes évolutionnaires, élaborés au cours des années 1950, Fraser forment une famille d'algorithmes de recherche inspirés de l'évolution biologique des espèces. L'idée ici est de s'inspirer de la théorie darwiniste de sélection naturelle pour résoudre des problèmes d'optimisation. Le principe de cette dernière stipule que les individus bien adaptés à l'environnement ont plus de chances de se reproduire et de survivre que les autres individus. En fait, la combinaison des caractéristiques des individus peut former au fil des générations de nouveaux individus de plus en plus mieux adaptés à leur environnement et qui peuvent avoir plus de chances de survivre que leurs parents. L'idée général est que, les individus qui ont hérité des caractères bien adaptés à leur milieu ont tendance à vivre assez longtemps pour se reproduire, alors que les plus faibles ont tendance à disparaître.

Au cours des années 1970, avec l'avènement des calculateurs de forte puissance, de nombreuses sous classes d'algorithmes évolutionnaires ont été réalisées pour modéliser cette approche. quatre approches se sont détachées :

- Les stratégies d'évolution [73] qui ont été conçues pour résoudre des problèmes d'optimisation à variables continues. Elles sont axées sur la modélisation des paramètres Métaheuristiques d'optimisation. ;
- La programmation évolutionnaire [28, 29], qui vise à faire évoluer les structures d'automates à états finis par des successions de croisements et de mutations ;
- Les algorithmes génétiques [70], Il sera bien détaillé dans la section suivante. Les algorithmes génétiques été conçus pour résoudre des problèmes d'optimisation à variables discrètes, en modélisant l'évolution génétique.
- La programmation génétique , basée sur les algorithmes génétiques, mais où les individus (ou chromosomes) sont des programmes informatiques, représentés en utilisant une structure d'arbre ; Cette dernière utilise des structures arborescentes pour représenter les individus de la population.

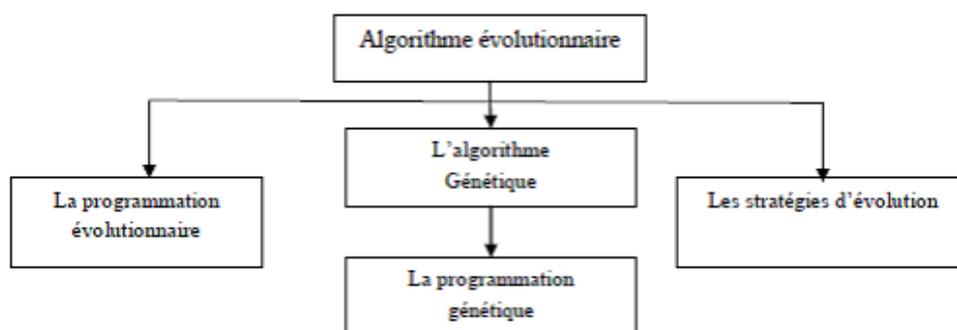


FIGURE 3.3: Les types des algorithmes évolutionnaires

Le processus d'évolution d'un algorithme évolutionnaire est basé sur trois opérations principales : la sélection, la reproduction et l'évaluation :

- **La Sélection** : Cette opération s'intervient dans deux phases. Elle est appliquée au premier lieu pour choisir les meilleurs individus parents qui vont se reproduire pour construire de nouveaux individus enfants. Ensuite, elle est appliquée à la fin de chaque itération pour opter pour les individus qui vont survivre et construire la

nouvelle population.

- **La reproduction** : Cette opération est en général composée de deux autres opérations : le croisement et la mutation. Elle permet la génération de nouveaux individus en combinant (phase de croisement) les caractéristiques des individus sélectionnés puis en appliquant quelques modifications de certains individus (phase de mutation) pour améliorer leurs qualités.
- **L'évaluation** : Cette opération consiste à mesurer la qualité de chaque individu (calculer la fitness des individus).

l'Organigramme suivant représente les étapes d'un algorithme évolutionnaire :

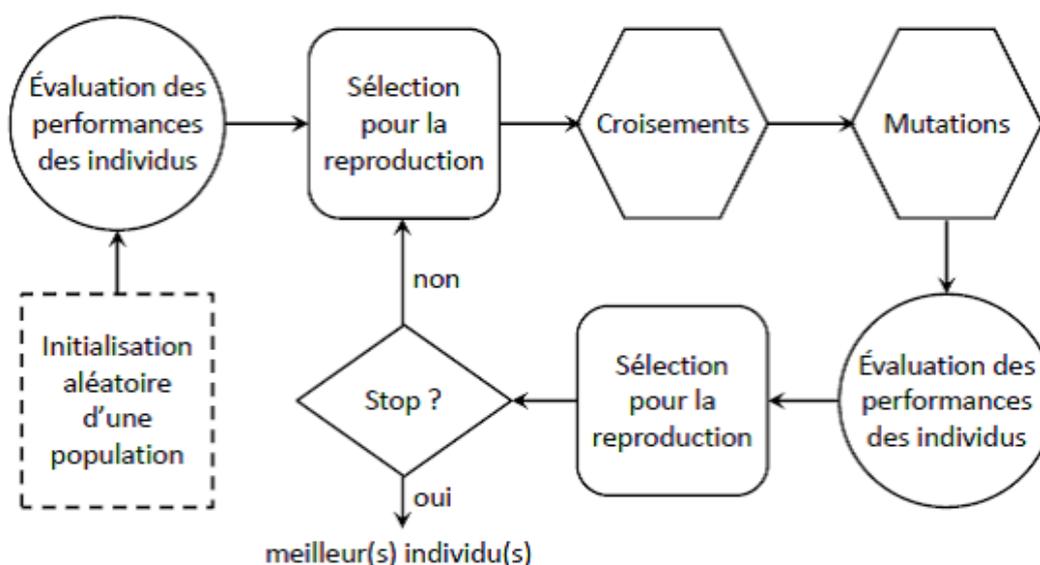


FIGURE 3.4: Organigramme d'un algorithme évolutionnaire

## 3.4 Optimisation par les Algorithmes Génétiques

### 3.5 Définition

Les algorithmes génétiques (AGs) sont des algorithmes d'optimisation stochastique inspirés des mécanismes de la sélection naturelle et de la génétique. Les (GAs) sont inspirés de la théorie de l'évolution proposée par Charles Darwin et des processus biologiques qui permettent à des organismes de s'adapter à leur environnement.

- Dans chaque environnement, seules les espèces les mieux adaptées perdurent au cours du temps ; les autres étant condamnées à disparaître.
- Au sein de chaque espèce, le renouvellement des populations est essentiellement dû aux meilleurs individus de l'espèce.

Charles Darwin en 1860 publie son livre intitulé L'origine des espèces au moyen de la sélection naturelle ou la lutte pour l'existence dans la nature. Dans ce livre, Darwin rejete

l'existence «de systèmes naturels figés», déjà adaptés pour toujours à toutes les conditions extérieures, et expose sa théorie de l'évolution des espèces : sous l'influence des contraintes extérieures, les êtres vivants se sont graduellement adaptés à leur milieu naturel au travers de processus de reproductions.

En 1866, Mendel ( « le moine des pois » ) publie l'article retraçant dix années d'expériences d'hybridation chez les végétaux (recombinaison des gènes) et l'adresse aux sociétés scientifiques des quatre coins du monde. Les réactions sont mitigées, voire inexistantes. Le monde scientifique n'est pas prêt à reconnaître la qualité de ses résultats. Ce n'est seulement en 1900, que la publication de trois nouveaux articles signés Hugo de Vries, Carl Correns et Erich von Tschermak révèle des résultats similaires à ceux de Mendel, et feront que ces premiers seront reconnus.

Dans les années 1960, à l'Université du Michigan les algorithmes génétiques ont été développés par John Holland, il introduit le premier modèle formel des algorithmes génétiques d'après l'étude des systèmes évolutifs (the canonical genetic algorithm AGC) dans son livre *Adaptation in Natural and Artificial Systems*. Il expliqua comment ajouter de l'intelligence dans un programme informatique avec les croisements (échangeant le matériel génétique) et la mutation (source de la diversité génétique).

C'est au livre de Goldberg (1989) "Genetic Algorithms in Search, Optimization, and Machine Learning" que nous devons leur popularisation. Leurs champs d'application sont très vastes. Outre l'économie, ils sont utilisés pour l'optimisation de fonctions (De Jong (1980)), en finance (Pereira (2000)), en théorie du contrôle optimal (Krishnakumar et Goldberg (1992), Michalewicz, Janikow et Krawczyk (1992) et Marco et al. (1996)), ou encore en théorie des jeux répétés (Axelrod (1987)) et différentiels (zyildirim (1996, 1997) et zyildirim et Alemdar (1998)). La raison de ce grand nombre d'application est claire : simplicité et efficacité. Bien sûr d'autres techniques d'exploration stochastique existent, la plus connue étant le recuit simulé (simulated annealing – voir Davis (1987) pour une association des deux méthodes).

### 3.6 Description de l'algorithme

Les algorithmes génétiques sont des algorithmes d'optimisation s'appuyant sur des techniques dérivées de la génétique et de l'évolution naturelle<sup>1</sup> : croisements, mutations, sélection, etc. Les algorithmes génétiques ont déjà une histoire relativement ancienne, puisque les premiers travaux de John Holland sur les systèmes adaptatifs remontent à 1962. L'ouvrage de David Goldberg a largement contribué à les vulgariser

Selon Lerman et Ngouenet (1995) un algorithme génétique est défini par :

- **Individu/chromosome/séquence** : une solution potentielle du problème ;
- **Population** : un ensemble de chromosomes ou de points de l'espace de recherche ;
- **Environnement** : l'espace de recherche ;
- **Fonction de fitness** : la fonction positive que nous cherchons à maximiser.

Terme	signification algorithmique
Chromosome	Chaîne
Gène	Trait, caractéristique (une partie du génotype)
Allèle	la valeur d'un gène
Phénotype	la solution actuelle (Ensemble de paramètres, structure décodée)
Génotype	une représentation de la solution sur la quelle sont appliqués les opérateurs de croisement et de mutation
Locus	Position dans la chaîne
Mutation	un opérateur aléatoire, qui reçoit en entrée une solution, et donne en sortie une nouvelle solution
Croisement	un opérateur qui croise les informations de deux individus, pour donner en sortie un ou deux nouveaux individus
Individu	une solution à un problème donné.
Parent	l'individu utilisé pour produire de nouveaux individus (après une mutation et/ou un croisement)
Enfant	l'individu engendré par la mutation et/ou le croisement
Population	un ensemble de solutions
Génération	une itération de l'algorithme génétique

### 3.6.1 Principes généraux

Un algorithme génétique recherche le ou les extrema d'une fonction définie sur un espace de données. Pour l'utiliser, on doit disposer des cinq éléments suivants :

1. Un principe de codage de l'élément de population. Cette étape associe à chacun des points de l'espace d'état une structure de données. Elle se place généralement après une phase de modélisation mathématique du problème traité. Le choix du codage des données conditionne le succès des algorithmes génétiques. Les codages binaires ont été très employés à l'origine. Les codages réels sont désormais largement utilisés, notamment dans les domaines applicatifs, pour l'optimisation de problèmes à variables continues.
2. Un mécanisme de génération de la population initiale. Ce mécanisme doit être capable de produire une population d'individus non homogène qui servira de base pour les générations futures. Le choix de la population initiale est important car il peut rendre plus ou moins rapide la convergence vers l'optimum global. Dans le cas où l'on ne connaît rien du problème à résoudre, il est essentiel que la population initiale soit répartie sur tout le domaine de recherche.
3. Une fonction à optimiser. Celle-ci prend ses valeurs dans  $\mathbb{R}^+$  et est appelée fitness ou fonction d'évaluation de l'individu. Celle-ci est utilisée pour sélectionner et reproduire les meilleurs individus de la population.
4. Des opérateurs permettant de diversifier la population au cours des générations et d'explorer l'espace d'état. L'opérateur de croisement recompose les gènes d'individus existant dans la population, l'opérateur de mutation a pour but de garantir l'exploration de l'espace d'état.

5. Des paramètres de dimensionnement : taille de la population, nombre total de générations ou critère d'arrêt, probabilités d'application des opérateurs de croisement et de mutation.

Un algorithme génétique fait évoluer une population dans le but d'en améliorer les caractéristiques des individus. Autrement dit, l'information obtenue lors de l'évaluation des points déjà testés à l'étape précédente est exploitée.

Les AGs sont alors basés sur les phases suivantes

- **1. Initialisation** : La première étape d'un algorithme génétique est la création du design initial. Une population initiale de N chromosomes est tirée aléatoirement
- **2. évaluation** : L'évaluation des individus/points choisis. Chaque chromosome est décodé, puis évalué.
- **3. Sélection** : La sélection des individus/points qui vont participer à la reproduction. c-à-d la création d'une nouvelle population de N chromosomes par l'utilisation d'une méthode de sélection appropriée.
- **4. Reproduction** : La sélection des points que l'on garde pour la suite de l'algorithme et la possibilité de croisement et mutation au sein de la nouvelle population.
- **5. Retour à la phase d'évaluation jusqu'à l'arrêt de l'algorithme** : La répétition du processus à partir de l'étape numéro 2.
- **6. Critère d'arrêt** : le processus évolutif d'un AG se termine lorsque la condition d'arrêt du problème est satisfaite.

### 3.6.2 La différence entre les Algorithmes génétiques et algorithmes classiques d'optimisation

Les algorithmes génétiques diffèrent des algorithmes classiques d'optimisation et de recherche essentiellement en quatre points fondamentaux :

1. Ils utilisent un codage des éléments de l'espace de recherche et non pas les éléments eux-mêmes ;
2. Ils recherchent une solution à partir d'une population de points et non pas à partir d'un seul point ;
3. Ils n'imposent aucune régularité sur la fonction étudiée (continuité, dérivabilité, convexité, etc.). C'est un des gros atouts des algorithmes génétiques.
4. Ils ne sont pas déterministes ; ils utilisent des règles de transition probabilistes.

Les algorithmes génétiques peuvent être particulièrement utiles dans les domaines suivants :

- Optimisation : optimisation de fonctions, planification, etc ...
- Apprentissage : classification, prédiction, robotique, etc ...
- Programmation automatique : programmes LISP, automates cellulaires, etc ...
- Etude du vivant, du monde réel : marchés économiques, comportements sociaux, systèmes immunitaires, etc ...

### 3.6.3 Fonctionnement des algorithmes génétiques

L'algorithme génétique commence par générer une population d'individus, de façon aléatoire. Puis, on sélectionne un certain nombre d'individus dans la population, afin de générer une population intermédiaire, appelée aussi *mating pool*. Deux parents sont ensuite sélectionnés ( $P_1$  et  $P_2$ ) en fonction de leurs adaptations. On applique, aléatoirement, un opérateur de croisement avec une probabilité  $p_c$  qui génère deux enfants  $C_1$  et  $C_2$ . Si la probabilité de croisement  $P_c$  vaut 0.6, par exemple, on appliquera l'opérateur de croisement dans 60% des cas sur les individus  $P_1$  et  $P_2$ . On modifie ensuite certains gènes de  $C_1$  et  $C_2$  en appliquant un opérateur de mutation avec une probabilité  $p_m$ , ce qui produit deux nouveaux individus  $\acute{C}_1$  et  $\acute{C}_2$  pour lesquels on évalue le niveau d'adaptation avant de les insérer dans la nouvelle population. Contrairement à la reproduction et au croisement qui favorisent l'intensification, l'opérateur de mutation favorise la diversification des individus. On réitère les opérations de sélection, de croisement et de mutation afin de compléter la nouvelle population. Enfin, on réitère les opérations précédentes à partir de la sélection, jusqu'à ce qu'un critère d'arrêt soit satisfait. Différents critères d'arrêt de l'algorithme peuvent être choisis : nombre de générations fixé, limite de convergence de la population, population qui n'évolue plus suffisamment, etc. L'algorithme suivant résume le fonctionnement d'un algorithme génétique de base :

#### Algorithme

- 1 : Générer aléatoirement une population de  $n$  individus.
- 2 : Evaluer la fitness  $f(x)$  de chaque individu  $x$  de la population.
- 3 : **Tant que** : Condition d'arrêt n'est pas vérifiée **faire**
- 4 : Sélectionner deux individus parents  $P_1$  et  $P_2$  dans la population.
- 5 : Avec une probabilité  $p_c$ , appliquer l'opérateur de croisement sur les deux parents  $P_1$  et  $P_2$  pour former deux nouveaux enfants  $C_1$  et  $C_2$ .
- 6 : Avec une probabilité  $p_m$ , appliquer l'opérateur de mutation à chacun des deux enfants  $C_1$  et  $C_2$  pour former deux nouveaux enfants  $\acute{C}_1$  et  $\acute{C}_2$ .
- 7 : Placer les deux enfants  $\acute{C}_1$  et  $\acute{C}_2$
- 8 : **Fin**
- 9 : **Retourner** : Meilleure individu

## 3.7 Les étapes des algorithmes génétiques

Les algorithmes génétiques sont des algorithmes itératifs de recherche d'optimum. A chaque itération, appelée génération, est créée une nouvelle population avec le même nombre de chromosomes. Cette génération consiste en des chromosomes mieux "adaptés" à leur environnement tel qu'il est représenté par la fonction sélective. Au fur et à mesure des générations, les chromosomes vont tendre vers l'optimum de la fonction sélective. Voyons maintenant plus en détail les autres phases de l'algorithme génétique. Nous présentons les étapes des algorithmes génétiques et les différents opérateurs.

### 3.7.1 Population initiale

Le choix de la population initiale d'individus conditionne fortement la rapidité de l'algorithme. Si la position de l'optimum dans l'espace d'état est totalement inconnue, il est naturel d'engendrer aléatoirement des individus en faisant des tirages uniformes dans chacun des domaines associés aux composantes de l'espace d'état.

La population initiale sera créée de façon aléatoire à condition que chaque individu dans la population crée une solution du problème. La taille initiale de la population doit être logiquement grande en tenant compte à la fois de la qualité des solutions trouvées et du temps d'exécution de notre algorithme.

Le choix des solutions initiales de bonne qualité accélère la convergence vers l'optimum. On peut obtenir la population initiale par :

- Génération aléatoire, permettant d'avoir des solutions de diverses caractéristiques.
- Des heuristiques, donnant des individus 'forts' mais qui risquent de faire converger l'algorithme vers des optimums locaux .
- Un mélange de solutions heuristiques et aléatoires, permettant d'avoir à la fois des individus 'forts' et une diversité.
- Duplication et évolution d'une seule solution, lorsque le problème est fortement contraint et l'obtention de plusieurs solutions est difficile. On prend une solution et on la modifie par mutation et croisement.

### 3.7.2 Le codage

Historiquement, le codage utilisé par les algorithmes génétiques était représenté sous forme de chaînes de bits contenant toute l'information nécessaire à la description d'un point dans l'espace d'état. Ce type de codage a pour intérêt de permettre de créer des opérateurs de croisement et de mutation simples. C'est également en utilisant ce type de codage que les premiers résultats de convergence théorique ont été obtenus. Cependant, ce type de codage n'est pas toujours bon.

Le codage est une modélisation d'une solution d'un problème donné sous forme d'une séquence de caractères appelée chromosome (est une suite de gène) qui représente une variable ou une partie du problème. Alors le chromosome est une suite de gène, on peut par exemple choisir de regrouper les paramètres similaires dans un même chromosome (chromosome à un seul brin) et chaque gène sera repérable par sa position : son locus sur le chromosome en question. Chaque individu est représenté par un ensemble de chromosomes, et une population est un ensemble d'individus.

Il existe trois principaux type de codage : binaire, gray ou réel. Nous pouvons facilement passer d'un codage à l'autre relativement facilement (voir Michalewicz (1992)). Certains auteurs n'hésitent pas à faire le parallèle avec la biologie et parlent de génotype en ce qui concerne la représentation binaire d'un individu, et de phénotype pour ce qui est de sa valeur réelle correspondante dans l'espace de recherche.

Rappelons que la transformation la plus simple (fonction de décodage  $d$ ) d'une chaîne binaire  $A$  en nombre entier  $x$  s'opère par la règle suivante:

$$x = d(A) = \sum_{i=1}^l a_i 2^{l-i-1}$$

où  $l$  est un longueur totale du chromosome

Ainsi le chromosome  $A = \{1, 0, 1, 1\}$  vaut  $1 * 2^3 + 0 * 2^2 + 1 * 2^1 + 1 * 2^0 = 8 + 2 + 1 = 11$ .  
évidemment, la fonction  $d$  sera modifiée selon le problème

Une autre façon de faire est de choisir  $d$  telle que :

$$x = d(A) = \sum_{i=1}^l \frac{a_i 2^{l-i-1}}{2^{l+1} - 1}$$

### 3.7.2.1 Le codage binaire :

Ce type de codage est le plus utilisé. Son principe est de coder la solution selon une chaîne de bits comme suit :

- Chaque gène dispose du même alphabet binaire  $\{0, 1\}$ ,
- Un gène est alors représenté par un entier long (32 bits),
- les chromosomes sont représentés par des tableaux de gènes et les individus de notre espace de recherche sont représentés par des tableaux de chromosomes.
- Ce cas peut être généralisé à tout alphabet allélique  $n$ -aire permettant un codage plus intuitif, par exemple pour le problème du voyageur de commerce on peut préférer utiliser l'alphabet allélique  $\{c_1, c_2, c_3, \dots, c_n\}$  où  $c_i$  représente la ville de numéro  $i$ .
- Le codage binaire est également indépendant des opérateurs génétiques (croisement et mutation).

<b>chromosome A</b>	1	0	1	0	0	1	0	1	0	0	1	1	0
<b>chromosome B</b>	1	1	0	1	0	0	1	1	0	1	0	1	1

**Le codage binaire** (3.1)

### 3.7.2.2 Le codage réel :

Les codages réels sont désormais largement utilisés cela peut-être utile notamment dans le cas où l'on recherche le maximum d'une fonction réelle, pour l'optimisation de problèmes à variables réelles. Le codage réel permet d'augmenter l'efficacité de l'algorithme génétique et d'éviter des opérations de décodage supplémentaires (le surplus).

Ce type de codage est beaucoup plus efficace pour représenter des problèmes de type continu. Il représente les solutions par des suites de type réel

0.23	1.25	0.5	0.12	3.2	7.8	20.7	11.3	0.78	17.28
------	------	-----	------	-----	-----	------	------	------	-------

**Codage réel d'un chromosome.** (3.2)

On peut distinguer un autre type spécial du codage réel, c'est le codage entier ou discret. Il utilise des entiers au lieu de réels

2	6	3	7	6	4	10	8	1	9
---	---	---	---	---	---	----	---	---	---

**Codage entier d'un chromosome.** (3.3)

### 3.7.2.3 Le codage de Gray :

dans le cas d'un codage binaire on utilise souvent la "distance de Hamming" comme mesure de la dissimilarité entre deux éléments de population, cette mesure compte les différences de bits de même rang de ces deux séquences.

Le codage de Gray, également appelé code Gray ou code binaire réfléchi, est un type de codage binaire permettant de ne modifier qu'un seul bit à la fois quand un nombre est augmenté d'une unité. Cette propriété est importante pour plusieurs applications.

Le code de Gray est une fonction qui associe à chaque nombre une représentation binaire. Cette méthode est différente du codage binaire naturel.

### 3.7.3 Evaluation des individus

Après avoir créé la population initiale, nous attribuons une valeur d'adaptation ou une évaluation des individus selon leur performance par rapport au coût de distance totale. L'évaluation d'un individu ne dépend pas de celle des autres individus (on peut alors paralléliser cette phase), Il faudrait donc créer une fonction d'évaluation pour évaluer la qualité de chaque individu. Le résultat fourni par cette fonction va permettre de sélectionner ou de refuser un individu pour ne garder que les individus ayant le meilleur coût en fonction de la population courante. c'est le rôle de la fonction de fitness

L'évaluation consiste à mesurer la 'force' de chaque individu (fitness) de la population.

- Si le problème est de maximisation, la fonction objectif peut être utilisée comme mesure.
- En cas de minimisation, on utilise une fonction qui varie inversement avec la fonction objectif. c'est à dire :

soit  $H(x)$  la fonction-objectif ;

$$\begin{cases} \text{Si problème max, critère} \implies H(x) \\ \text{Si problème min, critère} \implies \frac{1}{H(x)} \end{cases}$$

Le choix de  $H$  n'est pas unique, mais cette transformation est la plus simple.

Goldberg [89] propose une fonction de force  $F(x)$  avec un critère  $G(x)$  pour optimiser l'individu  $x$  donnée par :

$$F(x) = \begin{cases} G_{\max} - G(x), & G(x) \geq 0 \\ 0, & G(x) \leq 0 \end{cases}$$

où  $G_{\max}$  est un coefficient qui désigne la plus grande valeur observée de  $G(x)$ .

### 3.7.4 Opérateurs de sélection

La sélection permet d'identifier statistiquement les meilleurs individus d'une population et d'éliminer les mauvais. Dans la littérature il existe un nombre important de méthodes de sélection plus ou moins adaptées aux problèmes qu'elles traitent, à chaque génération, des individus se reproduisent, survivent ou disparaissent de la population sous l'action de deux opérateurs de sélection :

1. **La sélection parentale** : ou sélection pour la reproduction, ou plus simplement sélection, qui détermine combien de fois un individu sera reproduit en une génération ; cette sélection est très simple à mettre en oeuvre, puisqu'elle ne fait que choisir les  $P_{size}$  meilleurs individus d'une population,  $n$  étant un paramètre que l'utilisateur doit fixer.
2. **La sélection environnementale** : ou sélection pour le remplacement ou encore plus simplement : le remplacement, qui détermine quels individus devront disparaître de la population à chaque génération de façon que, de génération en génération, la taille de population reste constant soit.

On trouve dans la littérature un nombre important de principes de sélection plus ou moins adaptés aux problèmes qu'ils traitent. Il existe plusieurs méthode de sélection, les principales utilisées sont :

- La méthode élitiste
- La sélection par roulette ou proportionnelle.
- La sélection par tournoi.
- La sélection par le rang.
- La sélection uniforme

#### 3.7.4.1 La méthode élitiste :

L'objectif de cette méthode d'éviter que les meilleurs individus soient perdus après les opérateurs de croisement et de mutation ou après les opérations d'hybridation. Elle consiste à sélectionner les  $n$  individus dont on a besoin pour la nouvelle génération  $\acute{P}$  en prenant les  $n$  meilleurs individus de la population  $P$  après l'avoir triée de manière décroissante selon la fitness de ses individus.

#### 3.7.4.2 La sélection par roulette :

La méthode la plus connue et utilisée est sans nul doute, la roue de loterie biaisée (roulette wheel) de Goldberg (1989). Cette méthode est la plus ancienne, où chaque individu, de la population de taille maximale  $P_{max}$ , occupe une section de la roue proportionnellement à sa fonction d'adaptation  $Fitness(P)$ , la probabilité de sélection d'un individu ( $P$ ) s'écrit :

$$Prob(p) = \frac{fitness(P)}{\sum_{P=1}^{P_{max}} fitness(P)}$$

à chaque fois qu'un individu doit être sélectionné, On effectue ensuite un tirage au sort homogène sur cette roue, les individus possédant une plus grande fonction d'adaptation ayant plus de chance d'être sélectionnés.

Le principe de Roulette wheel selection 2 consiste à associer à chaque individu un segment dont la longueur est proportionnelle à sa fitness. On reproduit ici le principe de tirage aléatoire utilisé dans les roulettes de casinos avec une structure linéaire. Ces segments sont ensuite concaténés sur un axe que l'on normalise entre 0 et 1. On tire alors un nombre aléatoire de distribution uniforme entre 0 et 1, puis on "regarde" quel est le segment sélectionné. Avec ce système, les grands segments, c'est-à-dire les bons individus, seront plus souvent choisis que les petits. Lorsque la dimension de la population est réduite, il

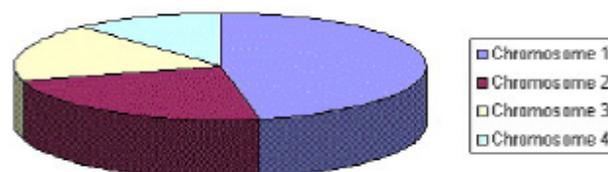


FIGURE 3.5: Schéma d'une roulette

est difficile d'obtenir en pratique l'espérance mathématique de sélection en raison du peu de tirages effectués. Un biais de sélection plus ou moins fort existe suivant la dimension de la population.

---

**Algorithme 2 : Sélection par roulette**

---

On calcule la somme  $S1$  de toutes les fonctions d'évaluation d'une population;  
 On génère un nombre  $r$  entre 0 et  $S1$ ;  
 On calcule ensuite une somme  $S2$  des évaluations en s'arrêtant dès que  $r$  est dépassé;  
 Le dernier individu dont la fonction d'évaluation vient d'être ajoutée est sélectionné;

---

### 3.7.4.3 La sélection par tournoi :

C'est la méthode la plus facile à mettre en oeuvre. Cette technique utilise la méthode de la roulette biaisée pour sélectionner deux individus. On récupère celui dont la valeur de la fonction d'adaptation est la plus grande. Cette méthode choisit toujours une valeur de la fonction d'adaptation plus élevée par rapport à la technique de la roulette biaisée.

Le principe de cette méthode est le suivant : on effectue un tirage avec remise de deux individus de  $P$ , et on les fait "combattre". Celui qui a la fitness la plus élevée l'emporte avec une probabilité  $P$  comprise entre 0.5 et 1. On répète ce processus  $n$  fois de manière à obtenir les  $n$  individus de  $P^*$  qui serviront de parents.

La variance de cette méthode est élevée et le fait d'augmenter ou de diminuer la valeur de  $P$  permet respectivement de diminuer ou d'augmenter la pression de la sélection.

### 3.7.4.4 La sélection par le rang :

La sélection par roulette présente des inconvénients lorsque la valeur d'évaluation des individus varie énormément. En effet, on risquerait d'atteindre une situation de stagnation de l'évolution. Imaginons le cas où 90% de la roulette est allouée à l'individu qui a la meilleure évaluation, alors les autres individus auront une probabilité très faible d'être sélectionnés.

La sélection par rang trie d'abord la population par évaluation. Ensuite, chaque individu se voit associé un rang en fonction de sa position. Ainsi le plus mauvais individu aura le rang 1, le suivant 2, et ainsi de suite jusqu'au meilleur individu qui aura le rang  $N$ , pour une population de  $N$  individus.

La sélection par rang d'un individu est identique à la sélection par roulette, mais les proportions sont en relation avec le rang plutôt qu'avec la valeur de l'évaluation. Avec cette méthode de sélection, tous les individus ont une chance d'être sélectionnés. Cependant, elle conduit à une convergence plus lente vers la bonne solution. Ceci est dû au fait que les meilleurs individus ne diffèrent pas énormément des plus mauvais.

### 3.7.4.5 La sélection uniforme

C'est une technique très simple qui consiste à sélectionner un individu aléatoirement de la population  $P$ . La probabilité  $P_i$  pour qu'un individu soit sélectionné est définie par :

$$P_i = \frac{1}{\text{taille pop}}$$

### 3.7.5 Le Croisement

Le phénomène de croisement est une propriété naturelle de l'ADN<sup>2</sup>. Il a pour objectif d'enrichir la diversité de la population en manipulant la structure des chromosomes des individus, c'est-à-dire les chromosomes. C'est par analogie avec la notion d'hybridation de Darwin qu'a été conçu cet opérateur.

Généralement, les croisements sont effectués entre deux parents pour générer deux enfants.

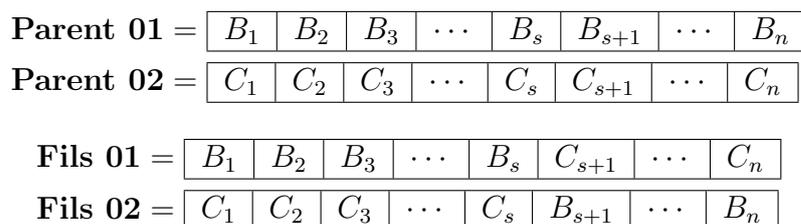
Bien que de nombreuses variantes de ces opérateurs existent, nous présentons ici celles qui sont, le plus couramment, utilisées.

Le croisement est effectué en tirant aléatoirement une position appelée site de croisement dans chacun des parents, puis on échange les deux sous-chaînes terminales de chacun des deux chromosomes, ce qui donne lieu à deux enfants. Ce type de croisement est appelé le croisement à un-point. Ce principe peut être étendu au croisement k-points, où  $k$  représente le nombre de sites de croisement en générant  $k+1$  sous chromosomes qui sont recombinaisonnés pour créer deux chromosomes fils. La figure suivante montre bien cette situation.

#### 3.7.5.1 Le croisement à un point :

On choisit au hasard un point de croisement, pour chaque couple. Notons que le croisement s'effectue directement au niveau binaire, et non pas au niveau des gènes. Un chromosome peut donc être coupé au milieu d'un gène, sauf si un bit représente un gène.

C'est le plus simple. Si  $n$  est la longueur du chromosome, on choisit aléatoirement une position  $s$  entre 1 et  $n - 1$ . On coupe les deux parents à cette position et on les recolle en les croisant.





peut être vu comme un croisement multipoint dont le nombre de coupures est indéterminé a priori. Pratiquement, on utilise un "masque de croisement", c'est-à-dire une chaîne de bits de même longueur que les chromosomes des parents sur lesquels il sera appliqué, qui est un mot binaire de même longueur que les individus. Un "0" à la  $n^{\text{ième}}$  position du masque laisse inchangés les symboles à la  $n^{\text{ième}}$  position des deux Chaînes, un "1" déclenche un échange des symboles correspondants. Le masque est engendré aléatoirement pour chaque couple d'individus. Les valeurs "0" ou "1" des éléments du masque sont généralement tirées avec une probabilité 0.5.

**Exemple :**

$$\begin{aligned} \text{Parent 01} &= \boxed{B_1} \boxed{B_2} \boxed{B_3} \boxed{B_4} \boxed{B_5} \\ \text{Parent 02} &= \boxed{C_1} \boxed{C_2} \boxed{C_3} \boxed{C_4} \boxed{C_5} \end{aligned}$$

On génère deux fois cinq valeurs aléatoires dans  $\{0,1\}$ , soient  $(0, 1, 1, 0, 1)$  et  $(1, 0, 0, 1, 1)$ . On obtient :

$$\begin{aligned} \text{Fils 01} &= \boxed{B_1} \boxed{C_2} \boxed{C_3} \boxed{B_4} \boxed{C_5} \\ \text{Fils 02} &= \boxed{C_1} \boxed{B_2} \boxed{B_3} \boxed{C_4} \boxed{C_5} \end{aligned}$$

### 3.7.6 La Mutations

**Une mutation** est l'inversion d'un bit aléatoire dans un chromosome. Elle joue le rôle de bruit et assure en outre une recherche aussi bien globale que locale, selon le poids et le nombre des bits mutés. Plus un individu a été muté, plus la recherche s'éloignera du voisinage local de ce même individu non muté.

L'opérateur de mutation [55] est un opérateur qui permet à un algorithme génétique d'atteindre tous les points de l'espace d'état d'une manière susceptible, sans les parcourir tous dans le processus de résolution, ce qui permet la convergence des algorithmes génétiques vers l'optimum global. Cette opération consiste à remplacer aléatoirement un gène dans le chromosome par une valeur aléatoire. Celle-ci peut être choisie dans le voisinage de la valeur initiale [22].

En général, on décide de muter une solution avec une probabilité assez faible. Le but de la mutation est d'introduire un élément de diversification et d'innovation. Il existe de nombreuses méthodes simulant une mutation. Souvent la probabilité de mutation  $P_m$  par bit et par génération est fixée entre 0,001 et 0,01. On peut prendre également une valeur  $P_m = 1$ . La mutation peut créer des chromosomes difficiles à avoir avec le croisement. Elle correspond à une petite perturbation de la solution. On cite les trois variantes suivantes : **Modifier la valeur d'un gène** : Le gène de position 3 est modifié de  $B_3$  en  $B_5$ .

$$\text{Parent} = \boxed{B_1} \boxed{B_2} \boxed{B_3} \boxed{B_4} \boxed{B_5} \longrightarrow \text{Fils} = \boxed{B_1} \boxed{B_2} \boxed{B_5} \boxed{B_4} \boxed{B_5}$$

**Permuter deux gènes consécutifs** : Les gènes de positions 3 et 4 sont échangés.

$$\text{Parent} = \boxed{B_1} \boxed{B_2} \boxed{B_3} \boxed{B_4} \boxed{B_5} \longrightarrow \text{Fils} = \boxed{B_1} \boxed{B_2} \boxed{B_4} \boxed{B_3} \boxed{B_5}$$

**Permuter deux gènes quelconques** : Les gènes de positions 1 et 4 sont échangés.

$$\text{Parent} = \boxed{B_1} \boxed{B_2} \boxed{B_3} \boxed{B_4} \boxed{B_5} \longrightarrow \text{Fils} = \boxed{B_4} \boxed{B_2} \boxed{B_3} \boxed{B_1} \boxed{B_5}$$

**Inverser l'ordre d'une sous chaîne du chromosome** : La sous chaîne  $B_2 B_3 B_4 B_5$  est inversée.

$$\text{Parent} = \boxed{B_1} \boxed{B_2} \boxed{B_3} \boxed{B_4} \boxed{B_5} \longrightarrow \text{Fils} = \boxed{B_1} \boxed{B_5} \boxed{B_4} \boxed{B_3} \boxed{B_2}$$

**Remarque 3.1** :

1. Les méthodes de croisement et de mutation sont liées au codage mais leur principe est identique.
2. Dans un environnement stable, il est préférable d'avoir un taux de mutation faible, la survie d'une espèce dans un environnement subissant une évolution rapide nécessite un taux de mutation élevé, permettant une adaptation rapide. Les taux de mutation d'une espèce dépendent donc de leur environnement.

La figure ci-dessous illustre bien ce mécanisme.

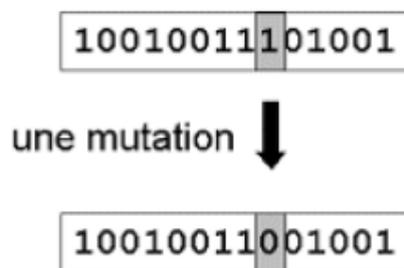


FIGURE 3.8: *Opération de mutation*

### 3.7.6.1 Mutation en codage binaire

Dans un algorithme génétique, la mutation en codage binaire est la modification aléatoire occasionnelle (de probabilité faible) de la valeur d'un caractère de la chaîne.

### 3.7.6.2 Mutation en codage réel

Pour le codage réel, les opérateurs de mutation les plus utilisés sont les suivants :

- ▶ L'opérateur d'inversion simple : consiste à choisir aléatoirement deux points de coupure et inverser les positions des bits situés au milieu.
- ▶ L'opérateur d'insertion : consiste à sélectionner au hasard un bit et une position dans le chromosome à muter, puis à insérer le bit en question dans la position choisie.
- ▶ L'opérateur d'échange réciproque : cet opérateur permet la sélection de deux bits et les inter changés.

### 3.7.7 Critère d'arrêt

Dans tout algorithme itératif, il faut définir un critère d'arrêt comme suit :

- ▶ Arrêt de l'algorithme lorsque le résultat atteint une solution satisfaisante.
- ▶ Arrêt si un certain nombre de générations est dépassé.

Le processus évolutif d'un AG se termine lorsqu'un critère d'arrêt est atteint. Le critère d'arrêt peut être une des conditions suivantes :

- Un nombre de générations maximal fixé a été atteint.
- Si l'on a trouvé la meilleure solution (si les conditions sur la solution recherchée sont connues).
- Si tous les individus de la population sont identiques.
- Un temps de calcul maximal est atteint.
- Un nombre donné d'évaluations de la fonction est atteint.
- Après une certaine taille de la population (s'il y a une évolution de la taille).

Une fois que la condition d'arrêt est satisfaite, les meilleurs individus de la population sont retenus comme les solutions au problème initial.

### 3.7.8 Les avantages

1. large espace de solution.
2. étudier de nombreuses solutions simultanément.
3. facilité d'application
4. pas de besoin de connaître les règles du problème.
5. une bonne performance sur les problèmes d'optimisation à grande échelle

### 3.7.9 Les désavantages

1. représentation du problème.
2. problème choisir différents critères tels que la taille de la population, taux de mutation, le taux de crossover, méthode de sélection des individus et leur force.
3. L'arrêt de l'algorithme n'est pas efficace

## 3.8 Algorithme d'optimisation par essaim de particules (PSO)

### 3.8.1 L'origine de L'idée de L'optimisation par Essaim Particulaire

L'Optimisation par Essaim Particulaire (OEP) a été proposée par James Kennedy (ingénieur en électricité) et Russel Eberhart (socio-psychologue) en 1995, voir [46]. Cette méthode est inspirée du comportement social des animaux évoluant en essaim. L'exemple le plus souvent utilisé est le comportement des bancs de poissons. En effet, on peut observer chez

ces animaux des dynamiques de déplacement relativement complexes, alors qu'individuellement chaque individu a une intelligence limitée et une connaissance seulement locale de sa situation dans l'essaim. Un individu de l'essaim n'a pour connaissance que la position et la vitesse de ses plus proches voisins. Chaque individu utilise donc, non seulement, sa propre mémoire, mais aussi l'information locale sur ses plus proches voisins pour décider de son propre déplacement. Des règles simples, telles que "aller à la même vitesse que les autres", "se déplacer dans la même direction" ou encore "rester proche de ses voisins" sont des exemples de comportements qui suffisent à maintenir la cohésion de l'essaim, et qui permettent la mise en oeuvre de comportements collectifs complexes et adaptatifs. L'intelligence globale de l'essaim est donc la conséquence directe des interactions locales entre les différentes particules de l'essaim. La performance du système entier est supérieure à la somme des performances de ses parties.



L'idée de s'inspirer des systèmes naturels pour proposer des méthodes de résolution des problèmes d'optimisation a donné naissance à une sous classe des métaheuristiques, ce sont des métaheuristiques basées sur l'intelligence par essaim (en anglais : Swarm Intelligence). La métaheuristique d'optimisation par essaim de particules (PSO : Particle Swarm Optimization) est la métaheuristique principale dans la classe des métaheuristiques basées sur l'intelligence par essaim. C'est une des métaheuristiques à base de population de solutions inspirées par une analogie avec l'éthologie.

L'optimisation par essaim de particules est une métaheuristique populaire basée sur l'intelligence par essaim. Elle a été proposée en 1995 par Kennedy et Eberhart [46]. L'optimisation par essaim de particules s'inspire du comportement social des oiseaux évoluant en groupe et des bancs de poissons. L'algorithme d'optimisation par essaim de particule lance la recherche avec une population de solutions, où chacune est appelée « particule ». Cette dernière est caractérisée par une vitesse de déplacement et une position dans l'espace de recherche. Au cours du processus de la recherche, chaque particule se déplace pour modifier sa position dans l'espace de recherche en fonction de sa vitesse actuelle, sa position actuelle, sa meilleure position trouvée au cours des itérations passées et la meilleure position trouvée par l'essaim. Son déplacement lui permet de mettre à jour sa position et sa vitesse de déplacement à chaque itération.

Il porte quelque ressemblance à computation évolutionnaire.

- L'objectif de PSO est trouver l'optimum global de quelque multidimensionnel (habituellement non linéaire) fonction. L'algorithme a prouvé efficace dans résoudre

beaucoup de problèmes.

- Dans PSO, la recherche à travers l'espace du problème peut être pensée de comme le vol d'un essaim de particules (points dans l'espace).
- L'objectif est avoir les particules converger sur l'optimum de la fonction, beaucoup comme un troupeau d'oiseaux converge sur quelque destination.
- Les particules sont distribuées initialement aléatoirement à travers le problème espacez et donné une vitesse initiale.
- Chaque particule se tient au courant de son emplacement et aptitude (la valeur de la fonction qui est optimisée), aussi bien que la meilleure place (et aptitude correspondante) il a rencontré si loin dans son vol.
- Avec le temps, la vitesse de chaque particule est ajustée afin qu'il déplace stochastique vers sa propre meilleure place et la meilleure place a trouvé par une autre particule dans son voisinage.
- Le voisinage d'une particule est le sous-ensemble de particules dans l'essaim avec lequel il a la communication directe. Ce réseau de rapports entre toutes les particules est connu comme la sociométrie, ou topologie de l'essaim.

## 3.8.2 Présentation de la Méthode

### 3.8.2.1 Principe général :

L'Optimisation par Essaim Particulaire (OEP, ou PSO en anglais) a été proposée par (Kennedy et Eberhart 1995) [46]. Cette méthode est inspirée du comportement social des animaux évoluant en essaim. Au départ, ils cherchaient à simuler la capacité des oiseaux à voler de façon synchrone et leur aptitude à changer brusquement de direction tout en restant en une formation optimale. Les particules sont les individus et elles se déplacent dans l'hyperm espace de recherche en se basant sur des informations limitées :

1. Chaque particule est dotée d'une mémoire qui lui permet de mémoriser le meilleur point par lequel elle est déjà passée et elle a tendance à retourner vers ce point.
2. Chaque particule est informée du meilleur point connu au sein de son voisinage et elle va tendre à aller vers ce point.

L'OEP est un algorithme à population. Il commence par une initialisation aléatoire de l'essaim dans l'espace de recherche. A chaque itération de l'algorithme, chaque particule est déplacée Une fois le déplacement des particules effectué, les nouvelles positions sont évaluées.

Kennedy et Eberhart se sont inspirés de ces comportements socio-psychologiques pour créer l'OEP. Un essaim de particules, qui sont des solutions potentielles au problème d'optimisation, survole l'espace de recherche, en quête de l'optimum global. Le déplacement d'une particule est influencé par les trois composantes suivantes :

- Une composante physique : la particule tend à suivre sa direction courante de déplacement ;
- Une composante cognitive : la particule fait tout pour se diriger vers la meilleure position rencontrée jusqu'à présent.

- Une composante sociale : la particule tend à se fier à l'expérience de ses congénères et, ainsi, à se diriger vers la meilleure position rencontrée par ses voisins.

Dans le cas d'un problème d'optimisation, la qualité d'un site de l'espace de recherche est déterminée par la valeur de la fonction objectif en ce point. La figure suivante illustre la stratégie de déplacement d'une particule.

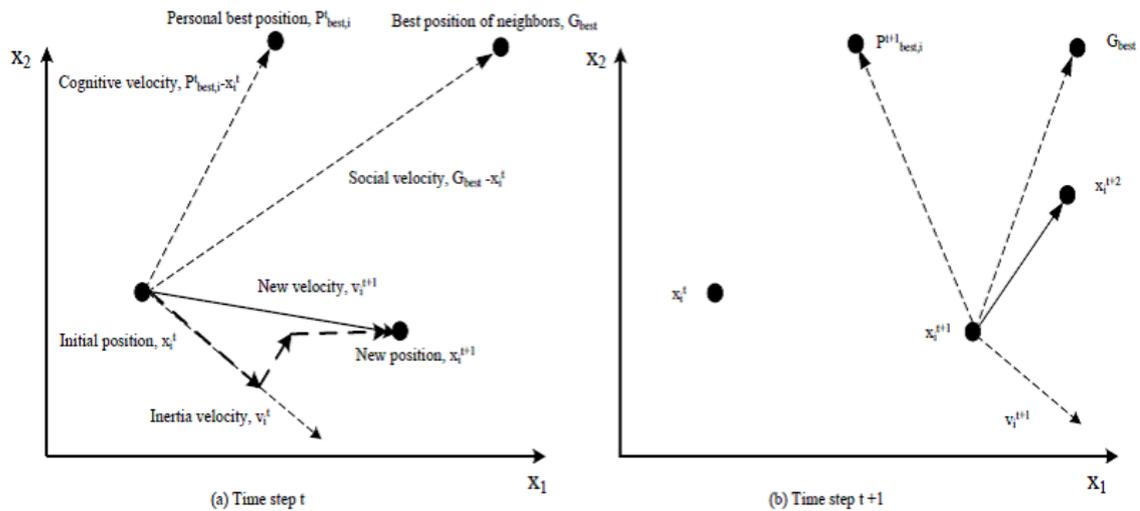
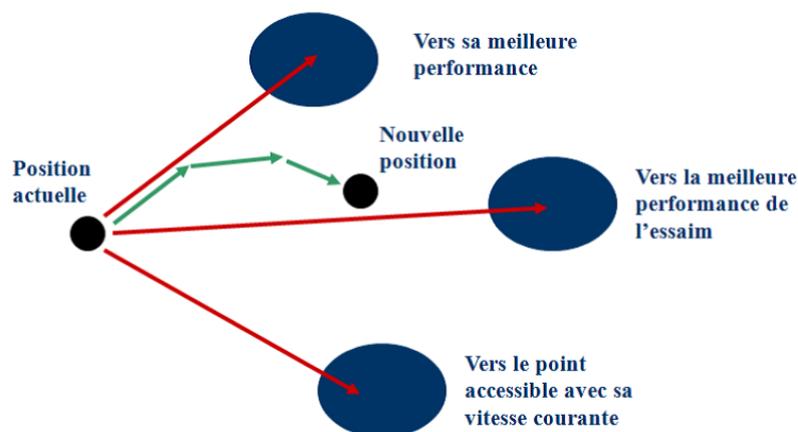


FIGURE 3.9: Déplacement d'une particule.

**Information utilisée par une particule pour déterminer sa prochaine position :**

- Sa vitesse actuelle.
- Sa meilleure performance.
- La meilleure performance de ses voisins.
- D'où trois types de comportement ...
- **égoïste** : suivre sa propre voie.
- **conservateur** : revenir en arrière.
- **panurgien** : suivre aveuglement le meilleur de tous.



### 3.8.2.2 Algorithme de base :

Un essaim de particules survole l'espace de recherche, en quête de l'optimum global. Dans  $\mathbb{R}^n$ , la particule  $i$  (solution potentielle) de l'essaim est modélisée par son vecteur position  $x_i = (x_{i,1}, x_{i,2}, x_{i,3}, \dots, x_{i,n})$  et par son vecteur vitesse  $v_i = (v_{i,1}, v_{i,2}, v_{i,3}, \dots, v_{i,n})$ . Cette particule garde en mémoire la meilleure position par laquelle elle est déjà passée, que l'on note  $p_i = (p_{i,1}, p_{i,2}, p_{i,3}, \dots, p_{i,n})$ . La meilleure position atteinte par toutes les particules de l'essaim est notée  $p_g = (p_{g,1}, p_{g,2}, p_{g,3}, \dots, p_{g,n})$ . Au temps  $t + 1$ , le vecteur vitesse est calculé à partir de (1.6)

$$v_{ij}(t+1) = v_{ij}(t) + c_1 r_1 (p_{ij}(t) - x_{ij}(t)) + c_2 r_2 (p_{gj}(t) - x_{ij}(t)) \quad (1.6)$$

$c_1$  et  $c_2$  sont deux constantes, appelées coefficients d'accélération;  $r_1$  et  $r_2$  sont deux nombres aléatoires tirés uniformément dans  $[0,1]$ .

- $v_{ij}(t)$  correspond à la composante physique du déplacement.
- $c_1 r_1 (p_i(t) - x_i(t))$  correspond à la composante cognitive du déplacement.  $c_1$  pondère les tendances de la particule à vouloir suivre son instinct de conservation et à aller vers sa meilleure position connue.
- $c_2 r_2 (p_g(t) - x_i(t))$  correspond à la composante sociale du déplacement.  $c_2$  contrôle l'aptitude sociale de la particule en se rapprochant plus de la meilleure position de ses informatrices

La position de la particule  $i$  est alors définie par :

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (1.7)$$

Il est à noter que le terme "vitesse" est ici abusif car les vecteurs  $v_i$  ne sont pas homogènes à une vitesse. Il serait plus approprié de parler de "direction de déplacement".

Le pseudo code pour la version la plus générale de l'algorithme - la version continue - est présenté par:

Algorithme : Optimisation par Essaim Particulaire (OEP)
<p><b>Initialiser</b> aléatoirement <math>p_s</math> particules : position et vitesse ;  <b>Evaluer</b> les positions des particules ;  <b>Tant que</b> le critère d'arrêt n'est pas atteint faire              <b>Pour</b> <math>i = 1, \dots, P_s</math> <b>faire</b>                  <b>Déplacer</b> les particules selon (1.6) et (1.7) ;                  <b>Si</b> <math>f(x_i) &lt; f(p_i)</math>                      <math>p_i = x_i</math> ;                      <b>Si</b> <math>f(x_i) &lt; f(p_g)</math>                          <math>p_g = x_i</math> ;                      <b>Fin Si</b>                  <b>Fin Si</b>              <b>Fin Pour</b>  <b>Fin Tant que</b></p>

### 3.8.3 Les Paramètres de L'algorithmes PSO

#### 3.8.3.1 Nombre de particules :

La quantité de particules  $n$  allouées à la résolution du problème dépend essentiellement de deux paramètres : la taille de l'espace de recherche et le rapport entre les capacités de calcul de la machine et le temps maximum de recherche. Il n'y a pas de règle pour déterminer ce paramètre, faire de nombreux essais permet de se doter de l'expérience nécessaire à l'appréhension de ce paramètre. Un grand nombre de particules peut réduire le nombre d'itérations ont besoin d'obtenir un bon résultat d'optimisation [41].

#### 3.8.3.2 Topologie du voisinage :

La topologie du voisinage défini avec qui chacune des particules va pouvoir communiquer. Si l'on se réfère à la version basique de l'*OEP* résumée par l'Algorithme 1, la meilleure particule  $p_g$  est choisie à partir de la population entière, c'est une topologie entièrement connectée. Cette version de *PSO* est appelée version "globale" (*Gbest*).

De même, on trouve aussi une version dite "locale" de l'*OEP*. Cette version utilise un graphe d'information circulaire. Les particules de l'essaim sont virtuellement disposées en cercle et numérotées séquentiellement à partir de 1 en parcourant le cercle. La particule n'est donc plus informée par toutes les particules. Dans les versions locales, le terme *Gbest* est remplacé par les termes *Lbest<sub>i</sub>*, où, pour chaque particule  $i$ , on définit un voisinage (*i.e.* l'information qui doit être partagée est la meilleure solution trouvée dans le voisinage de chaque particule (*Lbest<sub>i</sub>*)). Bien qu'elle converge moins rapidement que la version "globale", la version "locale" de l'*OEP* donne de meilleurs résultats, car elle est moins sujette à l'attraction par des minimas locaux.

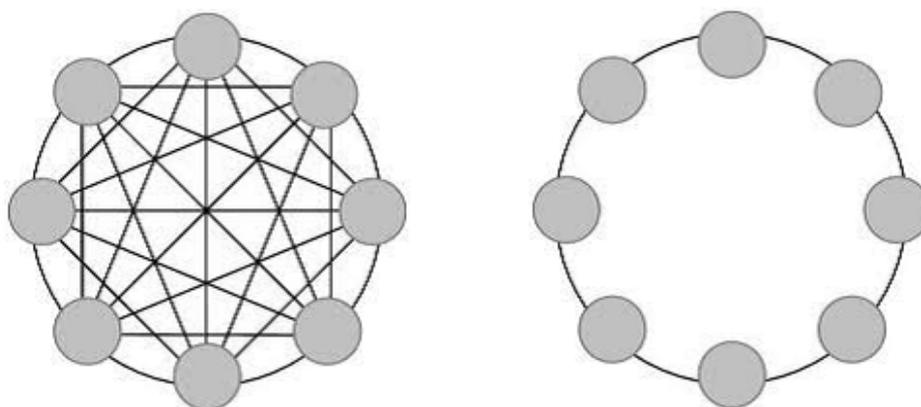


FIGURE 3.10: Graphe d'influence d'un essaim de particules : (à gauche) Graphe complètement connecté, (à droite) Graphe d'information circulaire.

De nombreuses autres topologies ont été testées. Parmi celles-ci, on peut retenir celles inspirées de modèles socio-psychologiques ou celles incluant des stratégies "petit-monde". Des résultats d'*OEP* utilisant différentes topologies sont disponibles. L'information importante qui ressort de ces résultats est qu'aucune topologie n'est meilleure que les autres,

dans le sens où il n'y a aucune topologie qui donne de meilleurs résultats que toutes les autres sur un large panel de problèmes .

Le voisinage géographique auquel nous sommes amenés à penser en premier lieu n'est pas nécessairement pertinent car, d'une part, il s'agirait d'un voisinage trop local, et d'autre part car la sociabilisations des particules tend à rendre tout voisinage social en voisinage géographique.

### 3.8.3.3 Coefficients de confiance :

Les variables de confiance pondèrent les tendances de la particule à vouloir suivre son instinct de conservation ou son panurgisme. Les variables aléatoires  $r_1$  et  $r_2$  peuvent être définis de la façon suivante

$$\begin{cases} \rho_1 = r_1 c_1 \\ \rho_2 = r_2 c_2 \end{cases}$$

Où  $r_1$  et  $r_2$  suivent une loi uniforme sur  $[0;1]$  et  $c_1$  et  $c_2$  sont des constantes positives déterminées de façon empirique et suivant la relation  $c_1 + c_2 \leq 4$ .

### 3.8.3.4 Vitesse maximale :

Pour éviter l'explosion du système, i.e. que les particules se déplacent trop rapidement d'une région à une autre dans l'espace de recherche, un mécanisme de confinement d'intervalle est souvent mis en place. La pratique la plus courante est de définir une vitesse maximale  $V_{max}$  à ne pas dépasser. Ainsi, on pose

$$v_{ij} \in [-V_{max}, V_{max}] \quad (1.8)$$

### 3.8.3.5 Coefficient d'inertie :

Pour contrôler l'influence de la direction de la particule sur le déplacement futur, un coefficient d'inertie  $\omega$  a été introduit par [82]. Le but de l'introduction de ce paramètre est de réaliser un équilibre entre la recherche locale (exploitation) et la recherche globale (exploration).

Le coefficient d'inertie  $\omega$  varie linéairement avec le temps selon la formule suivante :

$$\omega = \omega_{\min} + (\omega_{\max} - \omega_{\min}) \cdot \left( \frac{iter}{\max_{iter}} \right)$$

où  $iter$  est l'itération courante et  $\max_{iter}$  est le nombre maximal d'itérations.

$\omega_{\max}$  et  $\omega_{\min}$  : désignent respectivement les valeurs maximum et minimum du coefficient  $\omega$  (généralement,  $\omega_{\max}$  et  $\omega_{\min} \in [0, 1]$ ).

Le calcul de la vitesse est alors défini par :

$$v_{ij}(t+1) = \omega (v_{ij}(t) + c_1 r_1 (p_{ij}(t) - x_{ij}(t)) + c_2 r_2 (p_{gj}(t) - x_{ij}(t))) \quad (1.9)$$

Une grande valeur de  $\omega$  est synonyme d'une grande amplitude de mouvement et donc d'exploration globale de l'espace de recherche. Les études menées par Shi et Eberhart [82] indiquent une meilleure convergence pour  $\omega$  entre 0.8 et 1.2. La détermination de la meilleure

valeur de ce paramètre pour chaque algorithme se fait à travers des expérimentations numériques. Eberhart et Shi ont proposé une autre variante, dans laquelle le coefficient d'inertie est choisi au hasard, selon une distribution uniforme, dans l'intervalle  $[0.5, 1]$ . Cet intervalle a été inspirée du facteur de constriction proposé par Clerc et Kennedy (la valeur attendue du coefficient d'inertie, dans ce cas, est égale à  $0,75 \approx 0,729$ ).

### 3.8.3.6 Facteur de constriction :

L'étude de la dynamique des particules au sein de l'essaim a conduit à la recherche de solutions pour éviter la divergence de l'algorithme.

Dans la même idée, une version met en place un facteur de constriction qui permet d'améliorer la convergence de l'algorithme sans avoir besoin de confiner la vitesse des particules. Dans [46], il a été démontré qu'une bonne convergence peut être assurée en rendant dépendants les paramètres  $\omega$ ,  $c_1$  et  $c_2$ . En utilisant le coefficient de constriction, l'équation 1.9 devient :

$$v_{ij}(t+1) = \chi(v_{ij}(t) + \phi_1 r_1(p_{ij}(t) - x_{ij}(t)) + \phi_2 r_2(p_{gj}(t) - x_{ij}(t))) \quad (1.10)$$

où

$$\chi = \frac{2}{\left| \phi - 2 + \sqrt{\phi^2 - 4\phi} \right|}, \phi = \phi_1 + \phi_2 > 4. \quad (1.11)$$

Les valeurs optimales de  $\phi_1$  et  $\phi_2$  ont été déterminées dans [46], en effectuant de nombreux tests. En général, on utilise  $\phi = 4.1$  et  $\phi_1 = \phi_2$ , ce qui donne un coefficient  $\chi = 0.7298844$ . On remarque que l'*OEP* avec coefficient de constriction est équivalente à l'*OEP* avec coefficient d'inertie :

$$\chi \leftrightarrow \omega, c_1 \leftrightarrow \chi\phi_1 \text{ et } c_2 \leftrightarrow \chi\phi_2$$

La définition du paramètre permet de contrôler le système, afin qu'il présente les caractéristiques suivantes :

- Le système converge vers un état d'équilibre ;
- Plusieurs régions de l'espace de recherche sont explorées, tout en évitant une convergence prématurée.

### 3.8.3.7 Initialisation de l'essaim :

La position des particules ainsi que leur vitesse initiale doivent être initialisés aléatoirement selon une loi uniforme sur  $[0; 1]$ . Cependant, en ce qui concerne la position des particules, il est préférable d'utiliser un générateur de séquence de *SOBOL* qui est plus pertinent dans la disposition homogène des particules dans un espace de dimension  $n$ .

### 3.8.3.8 Critères d'arrêt :

Le critère d'arrêt peut être défini suivant le problème posé. Si l'optimum global est connu a priori, on peut définir une "erreur acceptable"  $\epsilon$  comme critère d'arrêt. Sinon,

il est commun de fixer un nombre maximum d'évaluations de la fonction objectif ou un nombre maximum d'itérations.

la convergence vers la solution optimale globale n'est pas garantie dans tous les cas de figure même si les expériences dénotent la grande performance de la méthode. De ce fait, il est fortement conseillé de doter l'algorithme d'une porte de sortie en définissant un nombre maximum d'itération (que nous noterons `nbItermax`).

Critères d'arrêt est utilisé pour terminer le processus de recherche itérative. certains arrêt critères sont discutés ci-dessous :

- le nombre maximum d'itération a été atteint ;
- la variation de la vitesse est proche de 0 ;
- le *fitness* de la solution est suffisant.

# Algorithme d'optimisation de portefeuille d'actions à l'aide d'un AG

## 4.1 Introduction

L'optimisation de portefeuille est l'un des principaux processus dans les marchés financiers pour la sélection du meilleur portefeuille (répartition des actifs), parmi l'ensemble de tous les portefeuilles considérés, nous présentons dans ce chapitre une nouvelle approche pour obtenir un portefeuille optimal, qui minimise le risque pour un profit requis ou maximise le profit d'un risque donné. Pour résoudre le problème, nous introduisons d'abord un concept de risque d'écart absolu moyen (MAD). La fonction de risque MAD ( $L_1$ ) peut éliminer la plupart des difficultés associées au modèle de Markowitz. Nous utilisons un algorithme évolutif heuristique pour trouver le portefeuille optimal.

Nous avons proposé deux approches pour trouver un portefeuille d'actions réalisable ou un portefeuille optimal investi dans un marché financier basé sur le risque MAD en utilisant L'analyse en composantes principales ACP :

- Pour la première approche en utilise l'algorithme génétique (GA). Cette approche est organisée en deux étapes : la première consiste à utiliser la méthode de classification ACP pour classer les actions en classes. Dans un deuxième temps, nous utilisons un algorithme d'optimisation appelé MAD-PAG basé sur l'algorithme génétique et l'écart absolu moyen pour minimiser le risque mesuré par le MAD et maximiser la valeur du portefeuille.
- Pour la deuxième approche en utilise métaheuristique d'optimisation, L'optimisation par essais particuliers (OEP ou PSO en anglais). Cette approche est organisée en deux étapes : la première consiste à utiliser la méthode de classification ACP pour classer les actions en classes pour déterminer la meilleur population initial . Dans la deuxième étapes : nous utilisons un algorithme d'optimisation appelé MAD-PSO pour minimiser le risque mesuré par le MAD et maximiser la valeur du portefeuille.

## 4.2 Optimisation du portefeuille par la méthode moyenne-variance

Quel portefeuille est le meilleur ? Cette question est probablement aussi ancienne que la bourse elle-même. Cependant, lorsque Markowitz a publié son article sur la sélection des portefeuilles en 1952, il a jeté les bases de la théorie moderne du portefeuille en tant que problème mathématique .

Supposons qu'il y ait  $n$  actifs risqués, dont les taux de rendement au temps  $t$  sont donnés par les variables aléatoires  $R_1, \dots, R_n$ , où  $V_t$  la valeur totale de l'actif  $j$  au temps  $t$  :

$$R_j = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1$$

Soit  $X = (x_1, \dots, x_n)^T$ ,  $x_i$  est défini comme la proportion de la richesse investie dans l'actif  $i$ , avec  $\sum_{j=1}^n x_j = 1$ .

Le taux de rendement du portefeuille est :

$$R_p = \sum_{j=1}^n x_j R_j = \sum_{t=1}^{\infty} \sum_{j=1}^n x_j d_{jt} R_{jt} = \sum_{j=1}^n x_j \left( \sum_{t=1}^{\infty} d_{jt} R_{jt} \right)$$

$R_j = \sum_{t=1}^{\infty} d_{jt} R_{jt}$ , est le rendement actualisé du jème actif .

Où  $R_j$  est indépendant de  $X_j$ . Car  $X_j \geq 0$  pour tous  $j$ .

Soit  $X = (x_1, \dots, x_n)^T$  et  $R = (R_1, \dots, R_n)^T$ , le rendement du portefeuille est :

$$R_p = X^T R$$

### 4.2.1 Le rendement espéré

Soit  $r(x_1, x_2, \dots, x_n)$  le rendement espéré (par période) du portefeuille est donné par [58] :

$$r(x_1, x_2, \dots, x_n) = E [R_p] = E \left[ \sum_{j=1}^n x_j R_j \right] = \sum_{j=1}^n E [R_j] x_j = \sum_{j=1}^n x_j r_j = X^T \mu$$

Où :  $E [R_j] = r_j$  Le rendement espéré pour l'actif numéro  $j$  et  $\mu = (r_1, r_2, \dots, r_n)$ .

Le rendement espéré  $E (R_p)$  représente la valeur espéré de la variable aléatoire  $R_p$

### 4.2.2 La variance d'un portefeuille

Par définition, la variance d'une variable aléatoire  $R$  est donnée par Harry Markowitz (1959) [59] comme l'écart type du rendement :

$$\sigma^2 (x_1, \dots, x_n) = \sigma_p^2 = E \left[ \left\{ \sum_{j=1}^n x_j R_j - E \left[ \sum_{j=1}^n x_j R_j \right] \right\}^2 \right] = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j$$

Où  $\sigma_{ij}$  est la covariance entre le  $i^{\text{ème}}$  et  $j^{\text{ème}}$  actifs, le rendement des actifs est donné par :

$$\sigma_{ij} = Cov(R_i, R_j) = E \{ [R_i - E(R_i)] [R_j - E(R_j)] \} = E \{ [R_i - r_i] [R_j - r_j] \}$$

Soit  $\Omega$  la matrice de covariance de sorte que

$$\sigma^2(x_1, \dots, x_n) = \sigma_p^2 = X^T \Omega X$$

et  $\Omega$  est la matrice semi-définie positive symétrique ( $n \times n$ ) avec ( $\Omega \geq 0$ ).

Dans l'analyse de Markowitz (moyenne-variance). Le risque de rendement du portefeuille est quantifié par  $\sigma_p^2$ , seuls les deux premiers moments sont considérés dans le modèle de portefeuille, on suppose que  $\mu_i, \sigma_i$  et  $\sigma_{ij}$  sont tous connus.

### 4.2.3 Le modèle Moyenne-Variance

Le modèle de moyenne-variance de la sélection de portefeuille avec des rendements incertains est présenté par Liu [53] comme une extension du modèle de moyenne-variance de Markowitz [57]. Dans cette section, nous définissons le travail de Lin [52] et donnons les variations du modèle moyenne-variance.

Pour un portefeuille composé de  $n$  actifs (par exemple, des actions), soit  $x_i$  le poids de la valeur du portefeuille investie dans l'actif  $i$  alors  $X = (x_1, \dots, x_n)^T \in \{X \in \mathbb{R}^n : X^T \mathbf{1}_n = 1\}$  désigne la part du portefeuille à investir dans les  $n$  actifs risqués.

Il existe une variété de combinaisons de portefeuilles, chacune ayant son propre rendement de portefeuille. La question qui se pose est : parmi toutes les combinaisons de portefeuilles possibles, laquelle a la distance minimale, au sens des moindres carrés, entre le rendement du portefeuille et le rendement attendu du portefeuille ? C'est un problème d'optimisation

$$\min_x \{f(R_p, E[R_p])\} \implies \min_x X^T \Omega X$$

Le rendement du portefeuille a une moyenne  $E[R_p]$  et variance  $\sigma_p^2$ , où  $\Omega$  est la matrice de covariance des rendements des actifs [50].

Le modèle MV classique est défini comme suit :

$$\left( \begin{array}{l} \min \left\{ E \left( \sum_{j=1}^n x_j R_j - E \left[ \sum_{j=1}^n x_j R_j \right] \right)^2 \right\} \\ S.C \left\{ \begin{array}{l} E \left[ \sum_{j=1}^n x_j R_j \right] \geq \rho \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right) \iff \left( \begin{array}{l} \min X^T \Omega X \\ S.C \left\{ \begin{array}{l} X^T \mu \geq \rho \\ X^T \mathbf{1}_n = 1 \\ X \geq 0 \\ (i.e., x_i \geq 0 \text{ for all } i) \end{array} \right. \end{array} \right)$$

Où  $\rho$  le rendement minimum espéré et  $\mathbf{1}_n$  désigne un vecteur de tous les uns

Mathématiquement, minimiser la variance soumise aux contraintes d'objectif et de budget

conduit à un problème de programmation quadratique convexe ; peut s'écrire :

$$\left( \begin{array}{l} \min [V [x_1 R_1 + x_2 R_2 + \dots + x_n R_n]] \\ S.C \left\{ \begin{array}{l} E [x_1 R_1 + x_2 R_2 + \dots + x_n R_n] \geq \rho \geq \rho \\ x_1 + x_2 + \dots + x_n = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right) \\ \iff \\ \left( \begin{array}{l} \min \left[ \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \right] \\ S.C \left\{ \begin{array}{l} \sum_{j=1}^n x_j r_j \geq \rho \\ \sum_{j=1}^n x_j = 1 \\ x_i \geq 0; \quad j = 1, \dots, n \end{array} \right. \end{array} \right)$$

Si nous ne savons pas comment définir le niveau de rendement et / ou de risque conduit au problème de programmation bi-objectif suivant :

$$\left( \begin{array}{l} \max E \left[ \sum_{j=1}^n x_j R_j \right] \\ \min V \left[ \sum_{j=1}^n x_j R_j \right] \\ S.C \left\{ \begin{array}{l} \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right) \iff \left( \begin{array}{l} \max \left[ \sum_{j=1}^n x_j r_j \right] \\ \min \left[ \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \right] \\ S.C \left\{ \begin{array}{l} \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right)$$

Ce modèle peut être formulé comme une programmation à objectif unique comme suit

$$\left( \begin{array}{l} \max \left( \sum_{j=1}^n x_j r_j - \lambda \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j \right) \\ S.C \left\{ \begin{array}{l} \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right) \iff \left( \begin{array}{l} \max (X^T \mu - \lambda X^T \Omega X) \\ S.C \left\{ \begin{array}{l} X^T \mathbf{1}_n = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right)$$

Où  $\lambda > 0$  est considéré comme un paramètre d'aversion au risque (représente le degré d'aversion au risque des investisseurs).

Avec l'émergence d'une variété de logiciels pour résoudre les problèmes de programmation quadratique, le MV peut facilement être résolu par n'importe lequel des algorithmes standard si  $n$  est petit.

Lorsque  $n$  est supérieur à mille, la plupart de ces algorithmes nécessiteraient une quantité de calcul unreasonable [44].

### 4.3 Le modèle d'écart absolu moyen (MAD)

Nous sommes confrontés à certaines difficultés liées au modèle de Markowitz lorsqu'il s'agit d'un grand nombre d'actifs. La résolution d'un problème de programmation quadratique à variables multiples avec une matrice de covariance dense est une tâche de calcul très longue. Afin de simplifier la complexité du calcul et l'exigence d'information de la covariance, Konno et Yamazaki (Konno, H et Yamazaki, H, 1991) ont montré que la mesure de l'écart absolu moyen (MAD) du risque est équivalente à la mesure du modèle de variance et a avantages substantiels par rapport au modèle classique [43], Konno et Yamazaki ont introduit le modèle MAD pour transformer la fonction de risque de covariance en

MAD, il n'est pas nécessaire de calculer la matrice de covariance pour mettre en place le modèle, ce qui permettrait de résoudre des problèmes à grande échelle résolu facilement. Le modèle résultant est un problème de programmation linéaire (LP) et nécessite donc moins de temps de calcul [7].

**Définition 4.1** *L'écart absolu moyen (MAD) des rendements du portefeuille peut être introduit comme fonction de risque  $L_1$*

$$w(x) = E \left[ \left| \sum_{j=1}^n x_j R_j - E \left[ \sum_{j=1}^n x_j R_j \right] \right| \right]$$

Le MAD est en général une fonction non convexe, non différentiable.

**Theorem :**

Si les rendements du portefeuille sont distribués normalement, alors

- $(R_1, \dots, R_n)$  est une distribution normal multivarié est  $W(x) = \sqrt{\frac{2}{\pi}} \sigma(x)$ .
- $(\mu_1, \dots, \mu_n)$  est les moyennes de  $(R_1, \dots, R_n)$ .
- $\sigma_{ij}$  est la matrice de covariance de  $(R_1, \dots, R_n)$ .
- Alors  $\sum_{j=1}^n R_j x_j$  est normalement distribué.

$$\sigma(x) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j}$$

donc

$$w(x) = \frac{1}{\sqrt{2\pi}\sigma(x)} \int_{-\infty}^{+\infty} |r| \exp\left(-\frac{r^2}{2\sigma^2(x)}\right) dr = \sqrt{\frac{2}{\pi}} \sigma(x)$$

Konno and Yamazaki proposed the portfolio optimization model based on the piecewise linear risk function, they introduced the  $L_1$  risk function instead of  $L_2$  risk (standard deviation) function. If  $(R_1, \dots, R_n)$  are multivariate normally distributed then  $L_1$  and  $L_2$  are the same.[42]

The proposed MAD model led to the minimization problem :

$$\left( \begin{array}{l} \min E \left[ \left| \sum_{j=1}^n x_j R_j - E \left[ \sum_{j=1}^n x_j R_j \right] \right| \right] \\ S.C \left\{ \begin{array}{l} \left[ \sum_{j=1}^n x_j R_j \right] \geq \rho \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right) \iff \left( \begin{array}{l} \min E |X^T R - X^T \mu| \\ S.C \left\{ \begin{array}{l} X^T \mu \geq \rho \\ X^T \mathbf{1}_n = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right)$$

Dans Konno et Yamazaki [43] [42] et ont mis au point un modèle de choix de portefeuille exprimable sous forme d'un programme linéaire.

La déviation absolue  $W(x)$  du portefeuille  $x = (x_1, \dots, x_n)$  se définit par :

$$\begin{aligned} w(x) &= E \left[ \left| \sum_{j=1}^n x_j R_j - E \left[ \sum_{j=1}^n x_j R_j \right] \right| \right] \\ &= \frac{1}{T} \sum_{t=1}^T \left| \sum_{j=1}^n (r_{jt} - r_j) x_j \right| \end{aligned}$$

Soit

$$a_{jt} = r_{jt} - r_j, \quad j = 1, \dots, n \quad t = 1, \dots, T$$

Alors

$$w(x) = \frac{1}{T} \sum_{t=1}^T \left| \sum_{j=1}^n (r_{jt} - r_j) x_j \right| = \frac{1}{T} \sum_{t=1}^T \left| \sum_{j=1}^n a_{jt} x_j \right| \quad ((01))$$

D'après (01) le modèle Mean-Absolute Deviation (*MAD*) est :

$$\left( \begin{array}{l} \min \quad \frac{1}{T} \sum_{t=1}^T \left| \sum_{j=1}^n a_{jt} x_j \right| \\ S.C \quad \left\{ \begin{array}{l} \sum_{j=1}^n r_j x_j \geq \rho \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right) \quad ((02))$$

Nous pouvons mettre (02) sous forme d'un modèle de programmation linéaire par une technique standard [53] , pour cela nous introduisons les variables supplémentaires :  $y_t$  ,  $a_{jt}$  tel que :

$$\left| \begin{array}{l} y_t = \sum_{j=1}^n (r_{jt} - r_j) x_j \\ a_{jt} = (r_{jt} - r_j) \end{array} \right|$$

Par conséquent, le modèle *MAD* devient :

$$\left( \begin{array}{l} \min \quad \left[ \frac{1}{T} \sum_{t=1}^T y_t \right] \\ S.C \quad \left\{ \begin{array}{l} y_t + \sum_{j=1}^n a_{jt} x_j \geq 0 \\ y_t - \sum_{j=1}^n a_{jt} x_j \geq 0 \\ \sum_{j=1}^n r_j x_j \geq \rho \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right) \quad ((03))$$

Certains des avantages du modèle *MAD* (03) sont dans [53].

- (a) Cela, la résolution d'un programme linéaire (03) est plus facile et beaucoup plus rapide que la résolution de (01) et (02)
- (b) Une solution optimale  $x^* = (x_1^*, \dots, x_n^*)$  de (4) contient au plus  $2T + 2$  actifs avec un poids positif.
- (c) Peut utiliser  $T$  comme variable de contrôle lorsque nous limitons le nombre d'actifs dans le portefeuille [43].

$W(x)$  une fonction de risque peut être estimé de la manière suivante :

$$\widehat{W}(x) = \frac{1}{T-1} \sum_{t=1}^T \left| \sum_{j=1}^n (r_{jt} - \widehat{r}_j) x_j \right|,$$

où :

$$\widehat{r}_j = \frac{1}{T} \sum_{t=1}^T r_{jt} \quad j = 1, \dots, n$$

Alors l'algorithme d'optimisation exprimé dans l'équation (02) devient :

$$\left( \begin{array}{l} \min \left\{ \widehat{W}(x) = \frac{1}{T-1} \sum_{t=1}^T \left| \sum_{j=1}^n (r_{jt} - \widehat{r}_j) x_j \right| \right\} \\ S.C \left\{ \begin{array}{l} \sum_{j=1}^n \widehat{r}_j x_j \geq p \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0, \quad j = 1, \dots, n \end{array} \right. \end{array} \right)$$

Selon le théorème de Chvátal :

$$\min |x| \iff \min y$$

Sous les contraintes suivantes :

$$\left\{ \begin{array}{l} y \geq x \\ y \geq -x \end{array} \right.$$

Donc l'algorithme d'optimisation de Konno et Yamazaki est donné par :

$$\min \left\{ \frac{1}{T-1} \sum_{t=1}^T y_t \right\}$$

## 4.4 Algorithme génétique

L'algorithme génétique est une méthode heuristique inventée par Holland (1975) en tant qu'idées évolutives de sélection naturelle. L'algorithme génétique est une technique d'optimisation stochastique utilisée pour trouver des solutions aux problèmes d'optimisation. De nombreux articles de recherche ont prouvé la validité de l'approche GA dans les problèmes d'optimisation des fonctions [52]. GA est un algorithme itératif pour trouver l'optimum. Ainsi, il existe un ensemble d'individus souvent appelés population, chaque individu de la population présente une solution satisfaisante du problème d'optimisation. Les individus sont généralement appelés chromosomes [71]. Chaque chromosome représente le codage d'une solution potentielle au problème à résoudre [63] ; chaque chromosome représente un vecteur de décision constitué d'un ensemble d'éléments appelés gène.

L'AG lance le processus de recherche en générant aléatoirement une population initiale de solutions possibles, chaque solution est évaluée à l'aide d'une fonction de fitness, qui est une mesure de la performance de la solution [2]. Ensuite, les étapes de base des algorithmes génétiques sont [63], [85].

- **étape 01** : Générée une population Initiale aléatoirement.
- **étape 02** : Evaluer les individus de la population.
- **étape 03** : Appliquez une sélection élitiste : portez les meilleurs individus à la génération suivante de la reproduction, du croisement et de la mutation.
- **étape 04** : Remplacez la population actuelle par la nouvelle population.
- **étape 05** : Si la condition de terminaison est satisfaite, arrêtez, sinon passez à l'étape 2.

## 4.5 Analyse en Composantes Principales (ACP)

L'analyse en composantes principales (ACP) est un algorithme mathématique qui réduit la dimensionnalité des données tout en conservant l'essentiel de la variation dans l'ensemble de données [37].

Les premières descriptions de la technique connue sous le nom de PCA ont été données par Pearson (1901) et Hotelling (1933). L'article d'Hotelling est en deux parties. La première partie, la plus importante, avec l'article de Pearson, la collection d'articles édités par Brayant et Atchely (1975).

PCA est une transformation linéaire qui choisit un nouveau système de coordonnées appelé le composant principal où les données avec la plus grande variance viennent se trouver sur le premier axe de la composante principale et la deuxième plus grande variance sur le deuxième axe de la composante principale.

### 4.5.1 Théorème de l'analyse en composantes principales

Pour compresser un ensemble de points donné, nous utilisons l'ACP en recherchant la corrélation entre les points.

Soit  $X$  la matrice  $M_{(N,n)}$  à colonnes  $x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}$

$$X = [x_1 - \bar{x} \quad x_2 - \bar{x} \quad x_3 - \bar{x} \quad \dots \quad x_n - \bar{x}]$$

$x_1, x_2, \dots, x_n$  un ensemble de  $(N, 1)$  vecteurs et  $\bar{x}$  leur moyenne :

$$x_1 = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{bmatrix} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{bmatrix}$$

La matrice de covariance  $Q$  est :

$$Q = XX^t$$

$$= [x_1 - \bar{x} \quad x_2 - \bar{x} \quad \dots x_n - \bar{x}] \begin{bmatrix} (x_1 - \bar{x})^t \\ (x_2 - \bar{x})^t \\ \vdots \\ (x_n - \bar{x})^t \end{bmatrix}$$

$x_j$  peut s'écrire :

$$x_j = \bar{x} + \sum_{i=1}^n g_{ji} e_i \quad \text{and} \quad g_{ji} = (x_j - \bar{x}) e_i$$

Où  $e_i$  sont les  $n$  vecteurs propres.

Si les points sont fortement corrélés, de nombreuses coordonnées de  $x$  seront nulles ou fermées à zéro.

Puis leur valeur propre  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . donné  $\lambda_i \simeq 0$  si  $i > k$

Puis

$$x_j \simeq \bar{x} + \sum_{i=1}^{i=k} g_{ji} e_i$$

## 4.5.2 L'utilisation de l'analyse en composantes principales

Il existe de nombreuses façons d'utiliser l'ACP. Certains sont assez formels et mathématiques, et certains sont assez graphiques. Pour résoudre notre problème, nous utilisons une stratégie qui a fonctionné en 5 étapes :

- **Etape 1** : Calculez la matrice de covariance.
- **Etape 2** : Calculez les vecteurs propres et les valeurs propres de la matrice de covariance.
- **Etape 3** : Tracé les données normalisées.
- **Etape 4** : Calculez le vecteur fonctionnel.
- **Etape 5** : Dérivation de nouvelles coordonnées de données.

## 4.6 Minimisation du MAD à l'aide de PCA et d'algorithmes génétiques

L'objectif de notre algorithme est de déterminer de manière dynamique les proportions du portefeuille.

Nous minimisons donc le risque d'un scénario MAD (écart absolu moyen) et maximisons la valeur du portefeuille optimal. La procédure de l'algorithme se résume comme suit[43] :

---

Soit  $n$   $\Leftarrow$  Le nombre d'actifs (actions)

Meilleurs actifs  $\Leftarrow$  Générer une population initiale (en utilisant  $k$  meilleurs actifs).

Tant que : Générations  $<$  max [return] and min [MAD]

— Population Sélectionnez  $\Leftarrow$  Sélectionnez (meilleure population).

— Population Croiser  $\Leftarrow$  Croisement (meilleure population)

— Mutations d'une Population  $\Leftarrow$  Mutations (meilleure population)

— Population optimale  $\Leftarrow$  Meilleure (Population Sélectionnez  $\cup$  Population Croiser  $\cup$  Mutations d'une Population)

---

### 4.6.1 Problématique

Comment on forme un portefeuille optimal autant que possible à partir du nombres des actions

### 4.6.2 La proposition

Pour résoudre ce problème ; on recours à l'utilisation du méthode de Konno et Yamazaki à raison de leur caractéristique. Et pour éviter les points faibles de cette méthode, on va la amélioré par l'AG et PSO pour trouver la meilleure solution dans le plus court du temps possible (la solution optimale).

Pour obtenir un portefeuille optimal, nous utilisons le modèle MAD (l'écart absolu moyen de Konno Yamazaki) et l'algorithme génétique [42].

les algorithmes génétiques (AG) sont des algorithmes itératifs de recherche d'optimum. A chaque itération (génération) créer une nouvelle population avec le même nombre de chromosomes[44].

Nous avons  $n$  actions caractérisées par :

- L'écart absolu moyen  $L_1$ .
- Le rendement espéré d'une action  $r_i$ .

Dans notre approche, nous classons les actions en utilisant ACP (en anglais Principal Component Analysis) pour choisir les meilleures actions utilisées dans la population initiale pour l'algorithme AG.

### 4.6.3 Pourquoi l'écart absolu moyen (MAD)

Konno et Yamazaki déclarent que le modèle  $MAD$  est plus crédible que  $MV$  [43]. Leurs arguments sont les suivants :

1. Dans la formulation de  $MAD$  nous n'avons pas besoin de la matrice de variance covariance.
2.  $MAD$  est un modèle de programme linéaire que  $MV$  est un modèle de programme quadratique. (Les programmes linéaires sont moins coûteux que ceux de quadratique sur tout pour les problèmes de grandes tailles (le cas d'optimisation de portefeuille).
3. Les portefeuilles efficients fournis par  $MAD$  ont moins d'actifs que ceux de  $MV$  . Cette propriété est un avantage quand il y a des coûts de transactions.
4. Le modèle  $MAD$  est compatible avec le domaine stochastique quelle que soit la distribution des rendements mais nous avons que, en général, le modèle  $MV$  n'est pas compatible avec le dominance stochastique.

**Remarque 4.1** *Konno et Yamazaki ont illustré que sous l'hypothèse de la normalité des rendements des actifs, la mesure de risque MAD est équivalent à celui de Markowitz*

### 4.6.4 Pourquoi l'analyse en composantes principales

L'inconvénient des algorithmes de clustering (k-means, k-nearest voisins (KNN), ...) est que le nombre de clusters doit être prédéterminé et le deuxième problème que nous devons classer les  $n$  actions avec plus de deux caractères. Pour résoudre ces problèmes, nous avons utilisé l'analyse en composantes principales (ACP) [69, 41].

### 4.6.5 Le modèle mathématique :

Dans cette section, nous abordons le modèle MAD. Premièrement ; nous définissons les notions nécessaires [43] :

#### 4.6.5.1 Les indices :

- $j$  : indice de sécurité ( $j = 1, \dots, n$ ).
- $t$  : indice de période "de temps" ( $t = 1, \dots, T$ ).

#### 4.6.5.2 Les paramètres :

- $r_{jt}$  : taux de rendement observé de l'action  $j$  à la période  $t$ .
- $r_j$  : taux de rendement moyen *i.e*  $r_j = \frac{1}{T} \sum_{t=1}^T r_{jt}$ .
- $C_L, C_U$  : les limites inférieures et supérieures de l'argent à investir.
- $p_j$  : prix d'achat unitaire de l'action  $j$ .
- $d_j$  : taux de frais de transaction de l'action  $j$ .
- $L_j$  : unités minimales des actions qui doivent être incluse dans le portefeuille.
- $U_j$  : unités maximales des actions qui doivent être incluse dans le portefeuille.
- $\rho$  : taux de rendement requis (exiger)

#### 4.6.5.3 Les variables de décisions :

- $x_j$  : les unités des actions  $j$  qui devient être incluser dans le portefeuille.
- $y_t$  : écart inférieur au taux de rendement moyen à la période  $t$ .

Donc le modèle MAD (Mean absolute deviation) est le suivant :

#### 4.6.5.4 Le problème initial (Original Problem OP)

$$\left( \begin{array}{l} \min \left\{ \frac{1}{T} \sum_{t=1}^T y_t \right\} \quad (1) \\ S.C : \left\{ \begin{array}{l} -\sum_{j=1}^n (r_{jt} - r_j) p_j x_j \leq y_t \quad (2) \\ \sum_{j=1}^n (1 + d_j) p_j x_j \geq C_L \quad (3) \\ \sum_{j=1}^n (1 + d_j) p_j x_j \leq C_U \quad (4) \\ \sum_{j=1}^n (r_j - d_j - p) p_j x_j \geq 0 \quad (5) \\ L_j \leq x_j \leq U_j \quad (j = 1, \dots, n) \quad (6) \\ y_t \geq 0 \quad (t = 1, \dots, T) \quad (7) \end{array} \right. \end{array} \right)$$

Dans le cas simple, le modèle de MAD sera :

$$\left( \begin{array}{l} \min \left\{ \frac{1}{T} \sum_{t=1}^T y_t \right\} \\ S.C : \left\{ \begin{array}{l} y_t + \sum_{j=1}^n a_{jt} x_j \geq 0 \\ y_t - \sum_{j=1}^n a_{jt} x_j \geq 0 \\ \sum_{j=1}^n r_j x_j \geq \rho \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 \quad j = 1, \dots, n \end{array} \right. \end{array} \right)$$

## 4.6.6 Algorithme d'optimisation MAD-AG

Pour obtenir un portefeuille optimal, nous utilisons le modèle MAD (l'écart moyen absolu de Konno Yamazaki) et l'algorithme génétique comme suit :

les algorithmes génétiques (AG) sont des algorithmes itératifs de recherche d'optimum. A chaque itération (génération) créer une nouvelle population avec le même nombre de chromosomes.

### 4.6.6.1 Les étapes d'un AG de base

- **étape 1** : l'algorithme commencer par la génération d'une population d'individus de façon aléatoire. Pour passer d'une génération  $k$  à la génération  $k + 1$ , trois opérations appelées opérations génétique sont répétées pour tous les éléments (la sélection, le croisement et la mutation) [42].
- **étape 2** : la sélection des meilleurs chromosomes est la première opération. Au cours de cette opération l'algorithme sélectionne les éléments pertinents qui optimisent mieux la fonction objective.
- **étape 3** : pour la croisement, il permet de générer deux chromosomes nouveaux "enfants" à partir de deux chromosomes sélectionnés "parents".
- **étape 4** : la mutation, elle réalise l'inversion d'un ou plusieurs gènes d'un chromosome.

Les étapes d'un AG de base
La génération de la population initiale
Evaluation des individus
Calcul de la fonction sélective
<b>Répéter</b>
Sélection
Croisement
Mutation
Calcul de la fonction sélective
<b>jusqu'à</b> satisfaction du critère d'arrêt

### 4.6.6.2 Algorithme (MAD-AG)

- Nous proposons une approche pour l'optimisation de portefeuille d'action, cette approche consiste à minimiser le risque de portefeuille en utilisant le modèle MAD.
- La procédure d'optimisation consiste à minimisé l'écart moyen absolu.

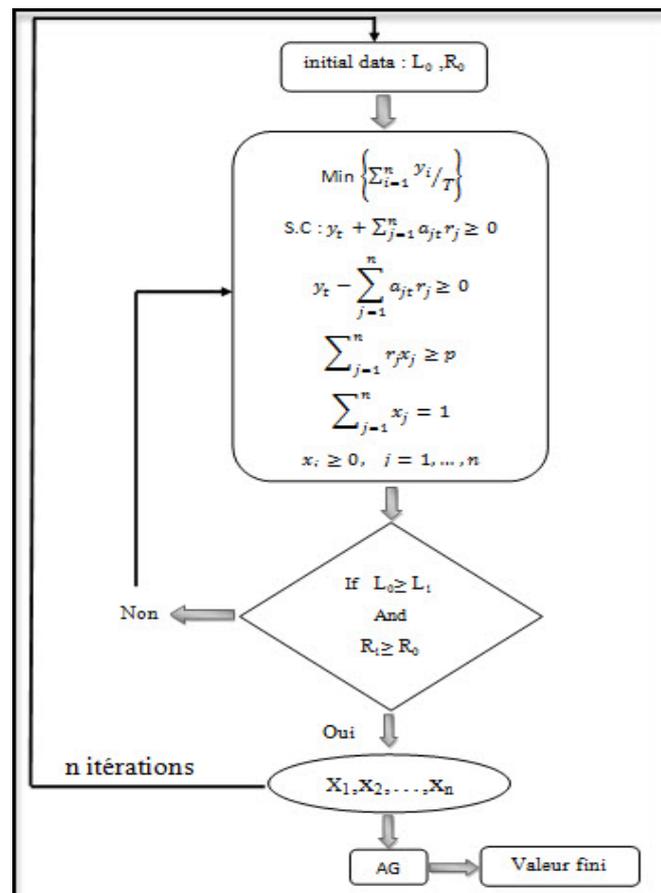


FIGURE 4.1: Organigramme général d'un algorithme MAD-AG.

#### 4.6.7 Procédure d'optimisation (MAD-AG) :

##### a- Initialisation :

La population est un ensemble de chromosomes, représente une solution unique au problème étant donné un ensemble de revenus, et chaque chromosome est composé de  $k$  gènes, chaque gène  $x_i$  représente une proportion de l'investissement dans les actifs (action). La population considérée ici est l'ensemble de chromosomes qui sont composés des gènes. Les gènes représentant les proportions  $x_i$  ( $i = 1, \dots, n$ ) de la richesse investie dans les actions.

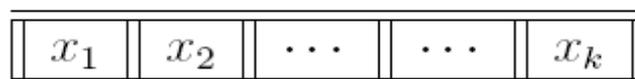


FIGURE 4.2: Structure d'un chromosome

Chaque individu de l'espace de points à associer une structure de données sous forme des chaînes de bits contenant toutes les informations nécessaires pour la description de ce point (s'appelle codage).

La population initiale est une ensemble de chromosomes ou des éléments de recherche (les proportions).

La génération d'une population est l'ensemble d'individus (chromosome) ayant été créés à la même data.

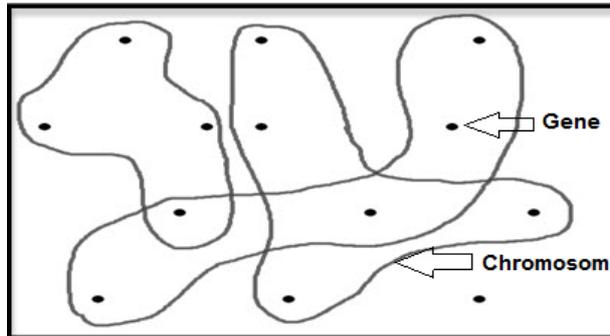


FIGURE 4.3: Structure de la population initiale (exemple)

**b- Choosing the right financial assets :**

Avant de créer la population, nous devons choisir les meilleurs actifs financiers (actions, actions) en utilisant la classification PCA (Principal Component Analysis) pour éliminer les actions faibles. L'analyse en composantes principales permet de restituer l'importance de chaque actifs dans la prédiction et ainsi d'identifier les chromosomes sur lesquelles le gérant de portefeuille doit se concentrer. Ces actifs sont celles qui permettent de distinguer au mieux les baisses des hausses de l'actif cible.

L'ACP en conjonction avec k-means est une méthode puissante pour visualiser des données de grande dimension.

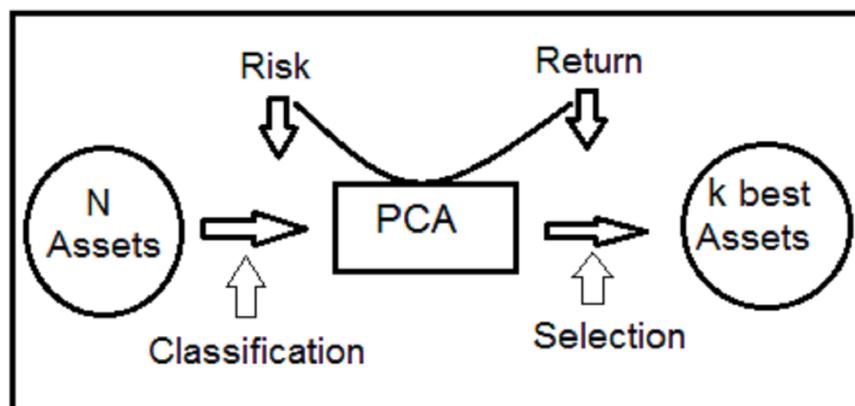


FIGURE 4.4: Analyse en composantes principales (PCA) Classification diagram.

**c- la fonction d'évaluation :**

L'opération suivante consiste à calculer la fonction d'évaluation (la fonction objective) de la population initiale . tel que chaque chromosome générée par les opérations précédentes.

Pour l'évaluation des individus (chromosomes) on s'intresse à calculer la force de chaque chromosome ce qui permet de retenir les individus les plus forts lors de l'étape suivante (la sélection)

La fonction objective utilisée dans ce travail est la fonction de minimisation[44] :

$$f(x) = MAD(\alpha, t),$$

où

$$MAD(\alpha, t) = L(x) = \frac{1}{T} \sum_{t=1}^T y_t = \frac{1}{T} \sum_{t=1}^T \left| \sum_{j=1}^n (r_{jt} - r_j) x_j \right|.$$

Représente la mesure de performance d'un individu au sein d'une population.

L'évaluation des individus(chromosomes) sera comme suit :

- Dans le cas de la minimisation :  $f(x) = L(x)$  .
- Dans le cas de la maximisation :  $g(x) = \Delta V(x) = V_p - V_0$ .

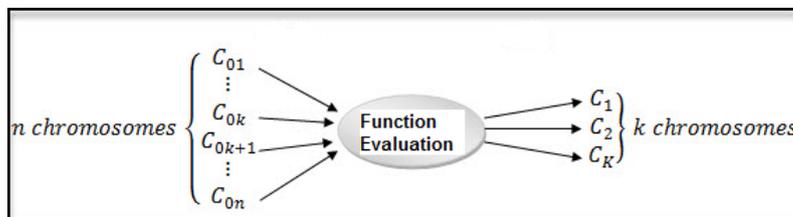


FIGURE 4.5: La fonction d'évaluation

**d- Opérateur de sélection :**

La sélection permet d'identifier statistiquement les meilleurs individus d'une population et d'éliminer les mauvais, pendant le passage d'une génération à une autre, ce processus est basé sur la performance de l'individu. L'opérateur de sélection doit être conçu pour donner également une chance aux mauvais éléments, car ces éléments pouvant par croisement ou mutation, engendrer une descendance pertinente par rapport au critère d'optimisation.

Le but est de sélectionner les meilleurs chromosomes de la population actuelle, les meilleurs chromosomes ont la meilleure probabilité de sélection.

Il existe différents techniques de solution, on propose

- La sélection uniforme.
- La sélection par tournoi.
- L'élitisme.
- La sélection par roulette.

La méthode adaptée à notre problème est la sélection par roulette. Cette méthode consiste à affecter à chaque individu  $c_i$  une force relative appelée probabilité d'apparition donnée par :

$$P(c_i) = \frac{1}{N-1} \left[ 1 - \frac{f_i}{\sum_{i \in N} f_i} \right]$$

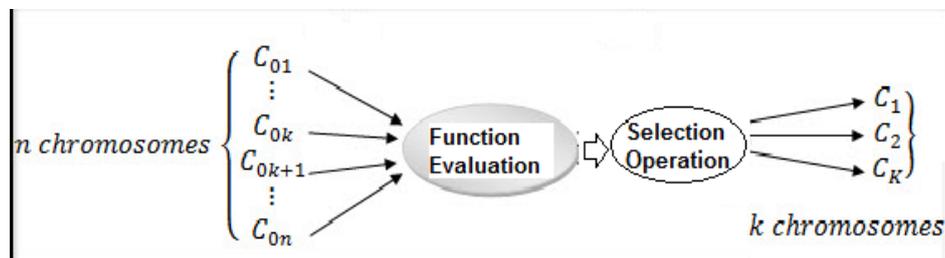
Où  $N$  est le nombre de chromosomes. dans la population.

la sélection d'un individu se déroule comme suit :

chaque chromosome est reproduit avec une probabilité, certains chromosomes seront "plus" reproduits et d'autres "mauvais" qui vont être éliminés.

Pour la maximisation de la valeur de portefeuille, l'opération de sélection s'appuie sur la probabilité :

$$P(c_i) = \frac{g_i}{\sum_{j \in Pop} g_j}$$



Les meilleurs chromosomes sont produits plus que les autres (éliminés).

#### e- Opérateur de croisement :

L'opérateur de croisement joue un rôle central dans l'algorithme génétique. L'opérateur de croisement favorise l'exploration de l'espace de recherche et enrichit la diversité de la population en manipulant la structure de chromosomes.

Après avoir utilisé la méthode de sélection pour sélectionner les meilleurs chromosomes, nous appliquons l'opérateur de croisement pour deux chromosomes appelés parents, désignés par  $C_1$  et  $C_2$  (sélectionnés de manière aléatoire). Nous divisons chaque parent en deux parties à la même position, puis produisons (génère) deux enfants  $\hat{C}_1$  et  $\hat{C}_2$ .

En espérant qu'un des deux enfants au moins héritera de bons gènes des deux parents et sera adapté qu'eux .

$$\begin{array}{|c|} \hline \begin{array}{l} \hat{C}_1 = rC_1 + (1-r)C_2 \quad \hat{C}_2 = rC_2 + (1-r)C_1 \\ r \in [0, 1] \end{array} \\ \hline \end{array}$$

Les enfants remplacent leurs parents si  $\hat{C}_1$  et  $\hat{C}_2$  sont possibles, sinon répéter le processus de croisement jusqu'à ce que deux enfants soient réalisables.

il existe plusieurs méthodes de croisement pa exemple (le croisement en un point, ou en multiples points). dans notre travail, la méthode utilisée est la première [41].

le croisement est effectué en tirant aléatoirement une position appelée site de croisement dans chacun des parents, puis on échange les deux sous-chaines terminales de chacun des deux chromosomes.

Ce principe peut être étendu au croisement  $K$ -points où  $K$  représente le nombre de sites de croisement en générant  $(K + 1)$  sous chromosomes [41].

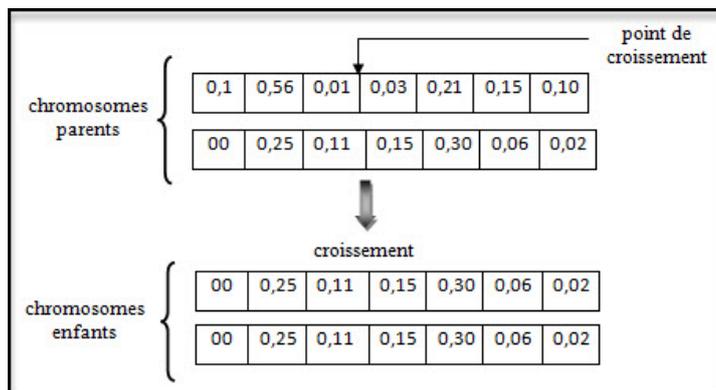


FIGURE 4.6: *Opération de croisement*

#### f- Opérateur de mutation :

Cette opérateur consiste à remplacer aléatoirement un gène dans une nouvelle génération le chromosome par une valeur aléatoire pour fournir la possibilité d'évolution [39]. En tirant un gène aléatoire dans le chromosome et en le remplaçant par une valeur aléatoire. Celle-ci peut être choisie dans le voisinage de la valeur initiale. Elle est utilisée généralement pour les problèmes discrets [69].

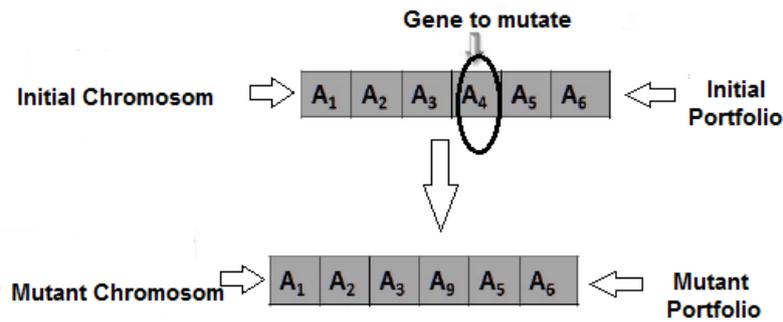


FIGURE 4.7: Opérateur de mutation

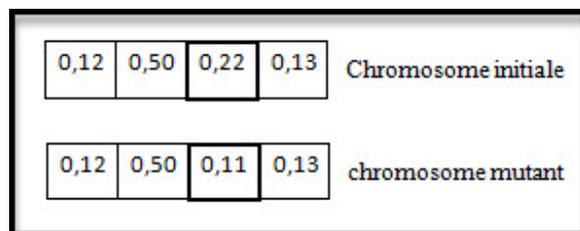


FIGURE 4.8: Exemple : Opérateur de mutation

**g- Conditions for convergence (critère d'arrêt) :**

La dernière génération est considérée. Si le chromosome optimal est obtenu, le résultat est favorable. Sinon les étapes d'évolution et de reproduction sont répétées jusqu'à un certain nombre de générations .

Pour l'arrêt de cet algorithme, un critère d'arrêt peut être défini arbitrairement comme le nombre maximum d'itérations, la détection d'un optimum ou le nombre de générations.

Pour les algorithmes évolutionnaires, trois types de critères d'arrêt de base sont connus[69] :

- L'arrêt après un nombre fixe d'itérations (ou d'évaluation),
- L'arrêt dès que la valeur de la performance du meilleur individu a atteint un certain seuil,
- L'arrêt après qu'un certain nombre d'itérations ont été faites sans améliorations.

L'organigramme général d'un algorithme MAD-AG peut être schématisé comme suit :

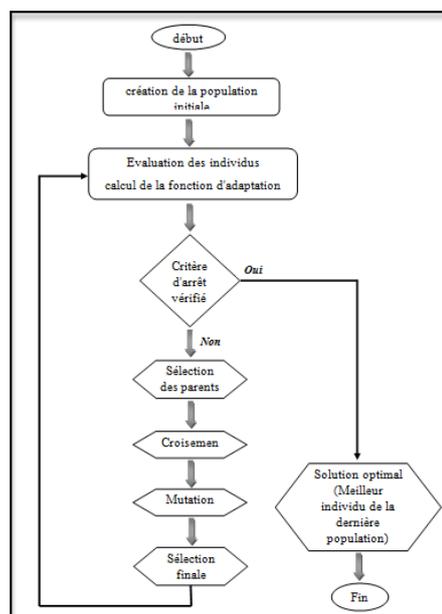


FIGURE 4.9: Organigramme général d'un algorithme 2 :MAD-AG.

## 4.7 Optimisation de portefeuille d'actions à l'aide de l'algorithmes PSO-MAD

### 4.7.1 Formulation du problème

Les essaims de particules forment une approche d'intelligence collective permettant de résoudre des problèmes d'optimisation. Cette approche ou méthode d'optimisation part d'une population de particules ou d'individus (les portefeuilles) se déplaçant dans l'espace de recherche. Le processus de recherche suit les étapes suivantes

- Chaque particule à une mémoire qui lui permet de mémoriser le meilleur point par lequel elle est déjà passée et elle a tendance à retourner vers ce point.
- Chaque particule appartient à un voisinage. Ce dernier a une influence sur son comportement. Cela signifie qu'une particule sera affectée par le meilleur point (particule) trouvé par n'importe quel membre de son voisinage topologique. Elle va tendre à aller vers ce point [69].

Pour résoudre le problème de portefeuille optimal par la PSO il faut déterminer un espace de recherche constitué de particules et une fonction objectif à optimiser. Le principe de l'algorithme est de déplacer ces particules afin qu'elles trouvent l'optimum .

Nous constatons que les paramètres qui interviennent dans cet algorithme sont :

- La dimension du problème,

- Le nombre  $N$  de particules,
- Les valeurs des coefficients  $P_1$  et  $P_2$ ,
- La taille du voisinage,
- La vitesse.

**Dans notre problème :**

les particules (essaim)  $\implies$  les portefeuilles

**Ces particules est notée :**

- Une position, c'est-à-dire ces coordonnées dans l'ensemble de définition
- Une vitesse, qui permet à la particule de se déplacer.
- Un voisinage : un ensemble de particules (portefeuilles).
- Sa meilleur position visitée.
- La position de meilleur voisin de l'essaim.
- La valeur qu'elle donne à la fonction objectif car à chaque itération il faut un comparaison entre la valeur du critère donnée par la particule courante et la valeur optimal.

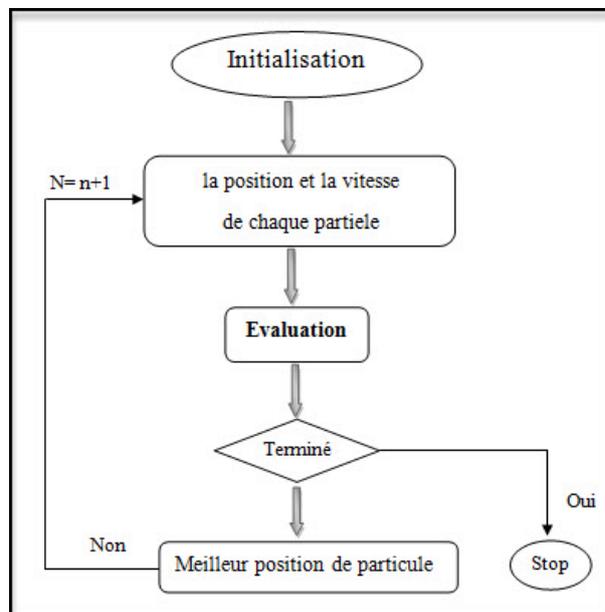


FIGURE 4.10: Organigramme générale d'un algorithme PSO

## 4.7.2 Configuration de la méthode MAD-PSO

### 4.7.2.1 Nombre de particules

La quantité de particules allouées à la résolution du problème dépend essentiellement :

- le nombre de portefeuille possible dans l'espace de recherche, il dépend le nombre des actions.
- les capacité de la calcul de la machine.

Un essaim de particules est une population de particules, où chaque particule est un objet en mouvement qui traverse l'espace de recherche et est attiré vers des endroits précédemment visités avec une grande aptitude. Contrairement aux individus dans le calcul évolutif (AG), les particules ne se reproduisent ni ne sont remplacées par d'autres particules .

Supposons que l'essaim de particules comprenne  $N$  particules,  $X_i = (x_{i,1}, \dots, x_{i,n})$  et  $V_i = (v_{i,1}, \dots, v_{i,n})$  désignent respectivement l'emplacement et la vitesse de la particule  $i$  dans l'espace de recherche à  $n$  dimensions,  $P_i = (p_{i,1}, \dots, p_{i,n})$  désignent l'emplacement optimal local de chaque particule (*lbest*),  $P_{gl} = (p_{g,1}, \dots, p_{g,n})$  désignent l'emplacement optimal global de toutes les particules (*gbest*) [69].

#### 4.7.2.2 Topologie du voisinage

Chacun des portefeuilles (particules) va pouvoir communiquer. dans notre problème on va utilisé la topologie en rayon, de sorte que chaque particule est reliée à  $n$  particules. (un portefeuille  $\rightarrow$  reliée à  $n$  portefeuille) [69].

#### 4.7.2.3 Coefficient de confiance

Ces variables aléatoires de confiance sont :

$$\begin{cases} P_1 = r_1 c_1 \\ P_2 = r_2 c_2 \end{cases}$$

où  $r_1$  et  $r_2$  sont deux nombres aléatoires indépendants variait de 0 à 1 suivent une loi uniforme sur  $[0, 1]$

$c_1$  et  $c_2$  sont des constantes positives déterminées de façon  $c_1 + c_2 \leq \mu$ . alors  $c_1$ ,  $c_2$  sont des chiffres de vitesse, régulant la longueur en volant vers la particule la plus optimiste de tout l'essaim et vers la particule individuelle la plus optimiste, voir [69].

#### 4.7.2.4 Vitesse maximale :

La vitesse dans le problème du portefeuille représente le taux de rendement maximal pour éviter que les particules ne se déplacent trop rapidement dans l'espace de recherche, passant éventuellement à coté de l'optimum

$$R_{\max} \rightarrow \vec{V}_{\max}$$

#### 4.7.2.5 La vitesse d'un particule :

- La vitesse d'un particule représente le rendement de portefeuille :  $\vec{V}_i(t) \rightarrow R_i(t)$
- $t$  est le nombre d'itérations
- $k$  est le poids d'inertie
- cependant, l'équation de la vitesse devient alors :

$$\vec{V}_i(t) = k(\vec{V}_i(t-1)) + P_1(\vec{x}_{pbest_i} - \vec{x}_i(t)) + P_2(\vec{x}_{vbest_i} - \vec{x}_i(t)) \quad (\text{pso1})$$

avec

$$K = 1 - \frac{1}{p} + \frac{\sqrt{|p^2 - 4p|}}{2}$$

aussi

$$p = p_1 + p_2 \succ 4$$

#### 4.7.2.6 Position d'un particule :

Après avoir mis à jour la vitesse des particules, utilisez la formule suivante pour mettre à jour l'emplacement de chaque particule [69] :

$$x_i(t+1) = x_i(t) + V_i(t) \quad (\text{pso2})$$

la position d'un particule  $i$  ( $x_i(t)$ ) représente la valeur de portefeuille à l'instant  $t$

$$x_i(t) \leftarrow V_{part_i}(t)$$

Ainsi :

- $\vec{x}_{pbest_i}$  : la meilleur valeur de portefeuille  $i$  a l'instant  $t$ .
- $\vec{x}_{vbest_i}$  : la meilleur valeur de portefeuille  $i$  en voisinage.
- $pbest_i$  : la meilleur valeur de la fonction objective pour le portefeuille  $i$ .
- $vbest_i$  : la meilleur valeur de la fonction objective comun du voisinage.

$$x_i(t) \leftarrow V_{part_i}(t)$$

Le comportement de la particule est un compromis entre les trois possibilités suivantes [69] :

- La particule suit son chemin personnel.
- La particule tend à retourner vers sa meilleure position.
- La particule tend à suivre la meilleure position trouvée par le voisinage.

#### 4.7.2.7 Critère d'arrêt

Ce processus se répète tant que le critère de convergence n'est pas atteint. Cela peut être :

- Un nombre fixe d'itérations ;
- En fonction de la fonction objectif ;
- Lorsque la variation de vitesse est proche de 0.

### 4.7.3 Algorithme d'optimisation MAD-PSO

Pour obtenir un portefeuille optimal, nous utilisons le modèle MAD (l'écart moyen absolu de Konno Yamazaki) et l'algorithme d'optimisation par essaim de particules (PSO).

Nous expliquons ici comment fonctionne l'algorithme PSO :

---

**Etape 1 :** Avant de créer l'espace de solutions réalisable (avant de générer des particules), nous devons choisir les meilleurs actifs financiers (actions, actions) en utilisant la classification PCA (Principal Component Analysis) pour éliminer les actions faibles.

l'analyse en composantes principales permet de restituer l'importance de chaque actifs dans la prédiction et ainsi d'identifier les portefeuilles sur lesquelles l'investisseur doit se concentrer.

**Etape 2 :** Réglez les paramètres (le nombre de portefeuille possible, Position d'un particule, La vitesse d'un particule :,...).

**Etape 3 :** Générer  $S$  particules selon la formulation des particules.

**Etape 3.1 :** définies Les positions initiales de la particule dans l'espace de recherche en initialisant chaque coordonnée à une valeur tirée de la distribution uniforme sur l'intervalle  $[x_{\max}, x_{\min}]$

**Etape 3.2 :** Générer des vitesses aléatoirement à partir de  $[-1.0, 1.0]$ .

**Etape 4 :** Répéter pour un nombre maximal d'itérations donné.

**Etape 4.1 :** évaluer l'aptitude de chaque particule en fonction des moments précalculés.

**Etape 4.2 :** Déterminer la meilleure position personnelle la plus visitée jusqu'à présent par chaque particule.

**Etape 4.3 :** Déterminer la meilleure position locale  $\mathbf{gbest}_i$  visitée jusqu'à présent par les particules locales voisines de chaque particule.

**Etape 5 :** Mise à jour les vitesses des particules en utilisant (pso2).

**Etape 6 :** Mise à jour les positions des particules en utilisant (pso1).

---

### 4.7.3.1 Algorithme PSO pour résoudre le problème des portefeuilles :

```

Répéter
  pour  $i = 1, \dots, n$  faire
    si  $f(\vec{x}_i) \succ pbest_i$  alors
       $pbest_i = f(\vec{x}_i)$ 
    fin si
    si  $f(\vec{x}_i) \succ vbest_i$  alors
       $vbest_i = F(\vec{x}_i)$ 
       $\vec{x}_{vbest_i} = \vec{x}_i$ 
    fin si
  fin pour
  pour  $i = 1$  to  $n$  faire
     $\vec{V}_i = K [\vec{v}_i + p_1 (\vec{x}_{pbest_i} - \vec{x}_i) + p_2 (\vec{x}_{vbest_i} - \vec{x}_i)]$ 
     $\vec{x}_i = \vec{x}_i + \vec{V}_i$ 
  fin pour
jusqu'à (un critères de convergence est

```

$$\left\{ \begin{array}{l} \min \quad \frac{1}{T} \sum_{i=1}^n y_i \\ \text{sous contraintes : } y_t + \sum_{j=1}^n a_{jt} x_j \geq 0 \quad t = 1, \dots, T \\ y_t - \sum_{j=1}^n a_{jt} x_j \geq 0 \quad t = 1, \dots, T \\ \sum_{j=1}^n r_j x_j \geq p \\ \sum_{j=1}^n x_j = 1 \\ x_j \geq 0 \quad j = 1, \dots, n \end{array} \right.$$

## 4.8 Partie pratique

### 4.8.1 Exemple 01

Le stade statique de l'AG nécessite certains paramètres pour obtenir une qualité de solution et maintenir une évaluation contrôlable du processus d'optimisation.

Taille de la population : popSize = 100 ;
Max de Génération : maxGen = 1000
Probabilité de croisement : pC = 0.70
probabilité de mutation : pM = 0.50 ;

Dans cet exemple numérique, 40 échantillons de données des entreprises, collectées du marché boursier de Tokyo, sont utilisés pour démontrer la méthode proposée. L'application de l'algorithme MAD-AG sur ce portefeuille donne les résultats suivantes.

ID	$r^{(3)}$	$r^{(12)}$												
1	0.07	0.10	9	0.03	0.04	17	0.03	0.05	25	0.01	0.03	33	0.11	0.10
2	0.14	0.12	10	0.04	0.03	18	0.05	0.06	26	0.01	0.04	34	0.05	0.07
3	0.10	0.03	11	0.01	0.04	19	0.06	0.10	27	0.21	0.22	35	0.04	0.08
4	0.31	0.23	12	0.19	0.21	20	0.13	0.17	28	0.12	0.14	36	0.04	0.02
5	0.01	0.00	13	0.13	0.14	21	0.11	0.11	29	0.11	0.07	37	0.02	0.01
6	0.05	0.01	14	0.09	0.10	22	0.08	0.10	30	0.06	0.07	38	0.03	0.01
7	0.01	0.05	15	0.06	0.05	23	0.13	0.18	31	0.07	0.08	39	0.17	0.12
8	0.12	0.10	16	0.09	0.17	24	0.22	0.19	32	0.03	0.01	40	0.14	0.13

tab : le revenu de 3 mois et 12 mois pour les 40 entreprises

Puis, nous choisissons les titres candidats selon la performance de 3 mois et 12 mois et réorganisons le numéro d'identification de sécurité comme suit :

R.ID	ID	$r^{(3)}$	$r^{(12)}$
1	04	0.31	0.23
2	24	0.22	0.19
3	27	0.21	0.22
4	12	0.19	0.21
5	39	0.17	0.12
6	21	0.15	0.11
7	02	0.14	0.12
8	40	0.14	0.13

tab : Réorganiser les données

La solution optimale déterminée par l'algorithme MAD-AG est :

Les proportions des actions dans le portefeuille optimal	(0.34, 0.22, 0.13, 0.09, 0.10, 0.02, 0.04, 0.06)
Le taux de rendement espéré	0.1943
Le risque	0.898

L'algorithme a été exécuté plusieurs fois en utilisant des valeurs variables pour la taille de la population initiale (50,100 et 500) et le nombre de générations (de 100 à 1000). Le nombre initial de mutations et de combinaisons a été fixé à 50% de la population initiale, puis a augmenté à mesure que la population augmentait.

Nous pouvons voir sur la figure 4.11 que nous avons une population raisonnablement diverse. La valeur de risque la plus basse de 3,7 correspond à un rendement de 86, alors que pour obtenir un rendement supérieur de 338, un risque considérablement plus élevé de 20,6. Bien qu'il existe un plus grand nombre de solutions à la figure 4.12, il y a peu de changement, sauf qu'il existe maintenant une augmentation de solutions avec une mesure à haut rendement et à risque élevé. Le rendement maximal est maintenant de 381 et le risque maximum 23,6.

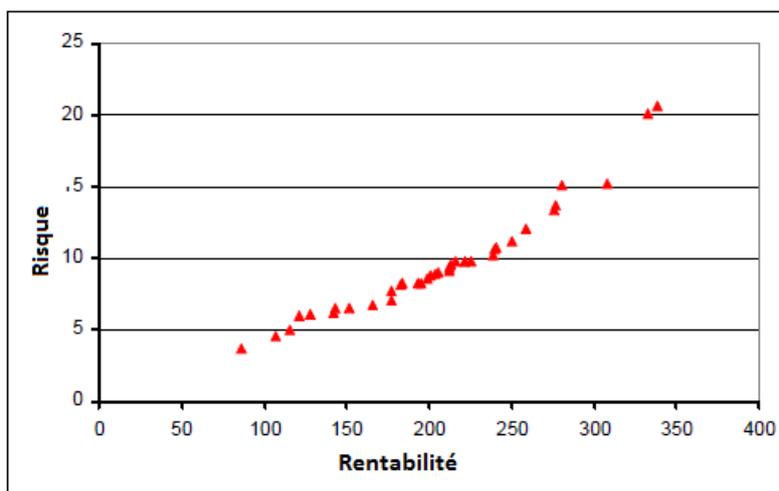


FIGURE 4.11: Résultats avec une population initiale de 50 avec 200 générations.

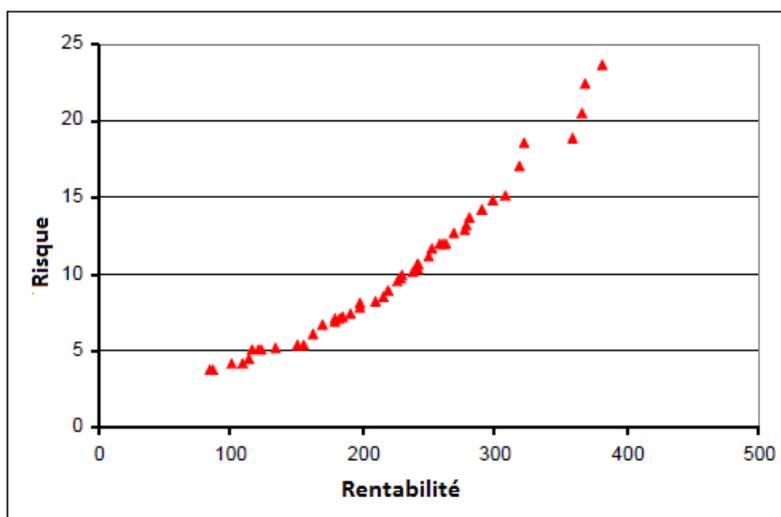
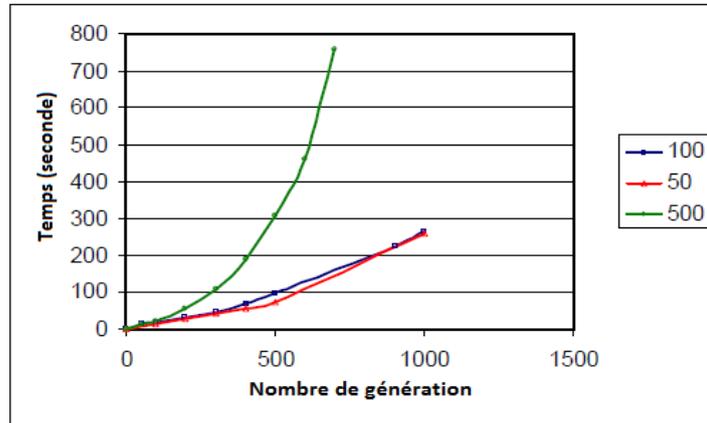


FIGURE 4.12: Résultats avec une population initiale de 50 avec 500 générations

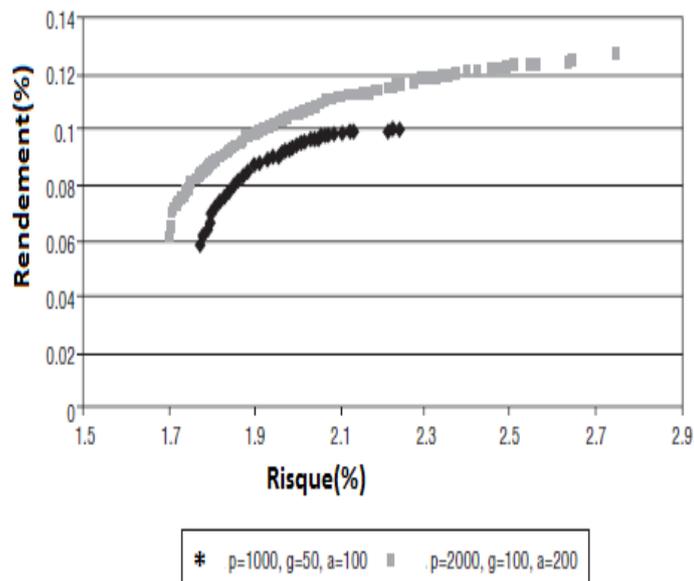
L'augmentation de la population initiale fait peu de différence avec les résultats, sauf pour le temps d'exécution de l'algorithme. Il n'y a pas beaucoup de différence de temps de calcul entre les populations initiales de 50 et 100 (seulement une question de secondes), mais cela s'allonge considérablement lorsqu'elles sont augmentées à 500. La taille de la population finale est également très différente comme on le voit dans le tableau suivant :

Population initiale	Population finale
50	3356
100	9063
500	42727

Le graphique suivant montrera comment le temps de calcul augmente par rapport au nombre de générations utilisées avec des tailles de population initiales de 50, 100 et 500.



Dans la figure suivante, nous montrons graphiquement les différences entre les portefeuilles efficaces en fonction des valeurs de paramètres utilisées dans l'MAD-AG puisque ces valeurs affectent de manière significative le rendement et le risque.



Les paramètres	MAD-AG(1)	MAD-AG(2)
Population initiale	1000	2000
Nombre de génération	50	100
Nombre d'action	100	200

Après 100 générations. Comme la figure précédente montre que les changements dans les valeurs des paramètres impliquaient une amélioration significative dans la frontière efficiente. La plus grande amélioration est faite par l'augmentation du nombre de générations et non à l'augmentation du nombre d'action (entre 50 et 100).

### 4.8.2 Exemple 02

Pour étudier l'efficacité de notre algorithme MAD-AG sur .10 entreprises (actions) investir dans le marché boursier de toronto.

L'application de l'algorithme MAD-AG sur ce portefeuille donne les résultats suivantes :

Action	Taille ( population)	Temps de simulation	Nb de génération	déviatiion absolu	erreur relative(%)
1	50	1500	600	0.8353	1.02
2	90	1300	400	0.8331	0.75
3	75	1000	300	0.8330	0.74
4	120	1500	400	0.8336	0.81
5	150	2500	200	0.8270	0.01
6	300	2000	220	0.8377	1.31
7	200	2200	250	0.8412	1.72
8	350	1500	250	0.8290	0.25
9	100	2500	500	0.8269	0
10	400	2500	100	0.8375	1.28

Le rendement maximal attendu correspondant est de 1,72. De plus, les proportions du portefeuille optimal sont :

Action i	1	2	3	4	5	6	7	8	9	10
Proportion (%)	0.00	0.43	2.26	6.40	9.06	58.75	2.83	1.93	0.00	18.34

L'évolution d'une valeur objective optimale à chaque génération est illustrée à la figure suivante. On constate que le résultat reste stable après 300 générations.

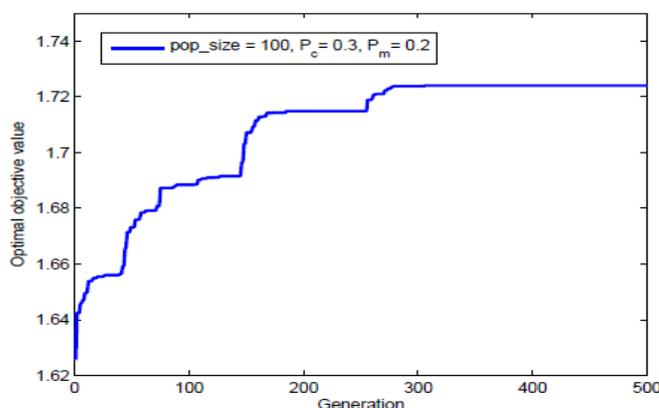


FIGURE 4.13: La convergence de la valeur objective.

La figure suivante est donnée pour préciser la relation entre le taux de risque de chaque portefeuille et la rentabilité ou bien le rendement :

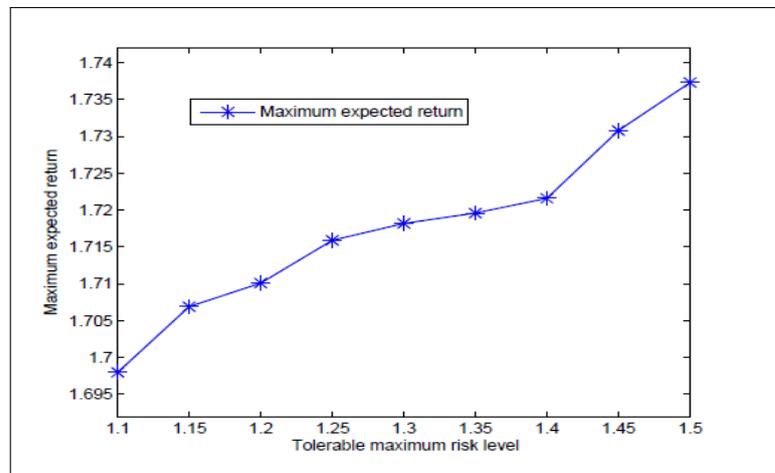


FIGURE 4.14: Comment le risque affecte-t-il la rentabilité espéré d'un investissement.

Pour le même exemple on va augmenter le capital investi ; Le but est de dessiner une frontière efficiente pour les différents capital investi. Voir la figure suivante :

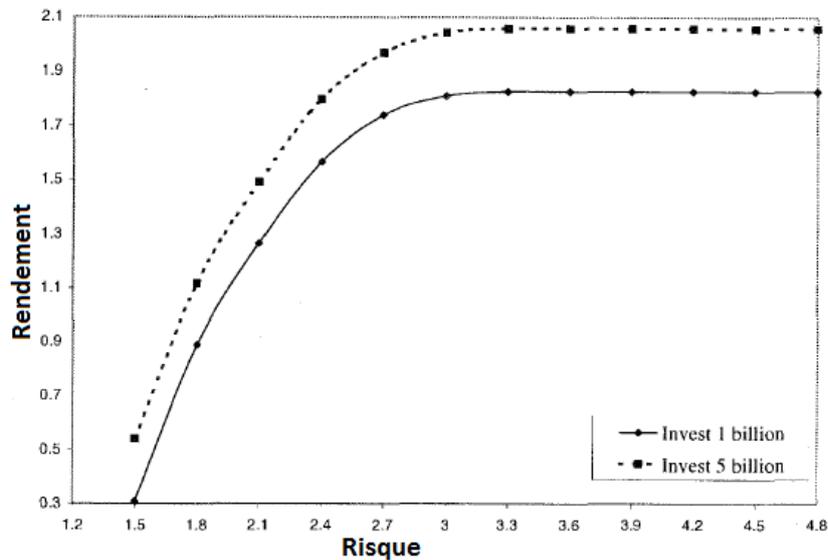


FIGURE 4.15: Frontière efficiente pour différents capital investi.

La figure suivante représente la frontière efficiente pour le taux de rendement réel calculé par le modèle MAD-AG sans coût de transaction, puis avec le coût de transaction ajusté par le coût de transaction pour l'achat du portefeuille. On va comparé avec le modèle MAD standard.

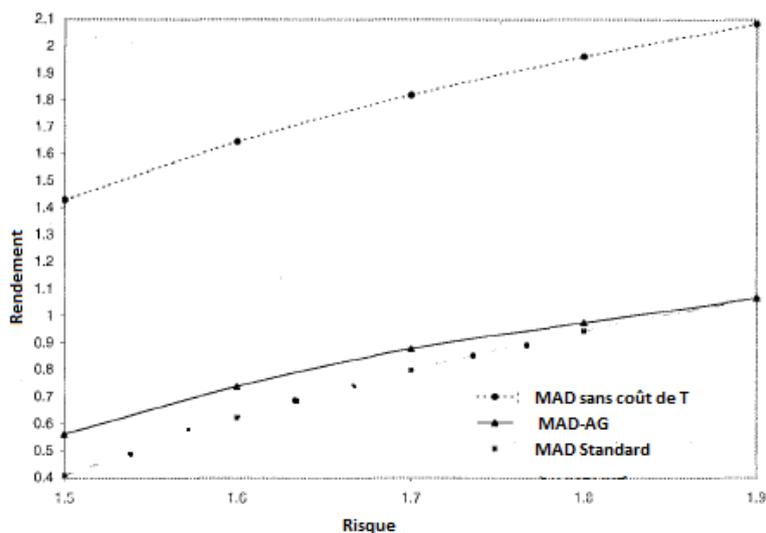


FIGURE 4.16: La frontière efficiente pour MAD-AG et MAD standard.

### 4.8.3 Exemple 03

Dans cet exemple, on va comparer les performances pour le modèle proposé MAD-AG et le modèle de Markowitz (comme modèle de référence)

Les données utilisées dans cet exemple ont été extraites de la base de données Wall Street (1995-2004). Il s'agit d'un ensemble composé de douze indices (action) composite des états-Unis, du Canada, du Japon, du Royaume-Uni, de la France, de l'Allemagne, de l'Espagne, des Pays-Bas et de la Suède. Nous employons des données hebdomadaires de ces indices. De Janvier 1995 à décembre 2004. Cela nous a fourni 939 observations par indice (action, portefeuille,...).

Pays	Indice	Moyenne	Déviations	Asymétrie
USA	DJ Industrial	0.2118	2.035	-0.878
	SP500	0.1853	2.075	-0.815
	Nasdaq	0.2328	3.805	-0.855
Canada	SPTSX	0.1504	2.321	-0.973
UK	Footsie100	0.1424	2.174	-0.233
France	CAC40	0.1922	2.911	-0.245
Germany	DAX	0.2315	2.902	-0.678
Spain	IBEX35	0.2306	3.020	-0.339
Holland	AEX	0.2453	2.652	-0.813
Sweden	OMX	0.3254	3.361	-0.195
Japan	Nikkei225	-0.1661	3.112	-0.202
Europe	Euro Stoxx 50	0.2781	2.598	-0.510

On obtient la frontière efficiente pour les deux différentes méthodes (MAD-AG et la méthode espérance-variance de Markowitz).

La méthode proposée MAD-AG ait donné globalement de meilleurs résultats que l'autre de Markowitz (pour un risque de 2.7 la rentabilité par le MAD-AG est de 0.27% par contre par la méthode de Markowitz la rentabilité 0.21% pour le même risque  $w=2.7$ ) ; voir la figure suivante :

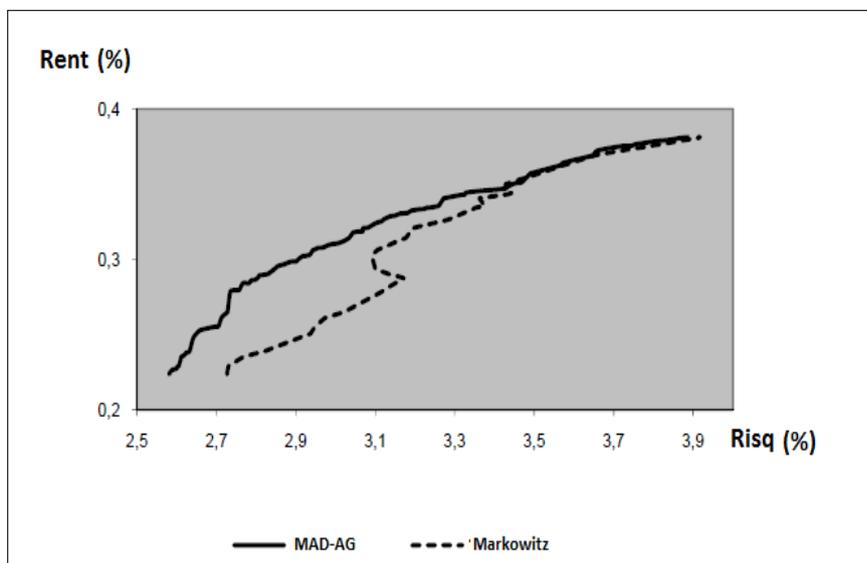


FIGURE 4.17: La frontière efficiente pour MAD-AG et Markowitz.

#### 4.8.4 Exemple 04

Dans cet exemple, nous allons déterminer les portefeuilles optimaux par le modèle proposé MAD-AG pour des différentes valeurs de risque ( $w=1$ ,  $w=0.5$ ,  $w=0$ ) les tableaux suivants sont des tableaux de proportionnalité (l'algorithme MAD-AG qui donne les proportions des actions pour construire un portefeuille optimal) En utilisant l'outil d'optimisation MAD-GA, pour un échantillon de 10 actions de la Bourse de Bucarest (BSE) entre janvier 2002 et décembre 2009, nous avons calculé les valeurs de la fonction objective et les proportions du portefeuille.

Action	Proportion pour un portefeuille d'actions
ALR	7.66%
ATB	0.00%
BRD	0.00%
OIL	64.29%
OLT	1.81%
SIF1	0.00%
SIF2	0.00%
SIF3	0.00%
SNP	0.00%
TBM	2.14%
TLV	24.10%
<b>Fonction objective</b>	-0.22

Tab 1. La proportion d'investissement avec un risque  $w=1$

La combinaison optimale consiste à investir 7,66% des fonds de portefeuille dans le secteur des «Matériaux», 64,9% dans le secteur «Services et équipements énergétiques», 1,81% dans le secteur «Chimie», 2,14% dans le secteur «Equipements» et 24,10% dans le secteur "Banques, assurances et services financiers". Pour cette situation, la valeur de la fonction objective est de -0.17%.

Action	Proportion pour un portefeuille d'actions
ALR	0.00%
ATB	0.00%
BRD	0.00%
OIL	70.53%
OLT	0.00%
SIF1	0.00%
SIF2	0.00%
SIF3	0.00%
SNP	0.00%
TBM	0.00%
TLV	29.47%
<b>Fonction objective</b>	0.07%

Tab 2. La proportion d'investissement avec un risque  $w=0.5$

En termes de risque ( $w = 0,5$ ) L'investisseur investit 70,60% dans le secteur «Chimie» et 29,40% dans le secteur «Banques, assurances et services financiers» et la valeur de la fonction objective est de 0,07%.

Action	Proportion pour un portefeuille d'actions
ALR	0.00%
ATB	0.00%
BRD	0.00%
OIL	100.00%
OLT	0.00%
SIF1	0.00%
SIF2	0.00%
SIF3	0.00%
SNP	0.00%
TBM	0.00%
TLV	0.00%
<b>Fonction objective</b>	<b>0.36%</b>

Tab 3. La proportion d'investissement avec un risque  $w=0$

L'amoureux des risques ( $w = 0$ ) investit tous ses fonds dans le secteur des entreprises «Banques, assurances et services financiers». Ce résultat montre que ce type d'investisseur préfère faire des investissements risqués et la valeur de la fonction objective est de 0,36%

#### 4.8.5 Exemple 05

Considérons un portefeuille constitué de 48 actions de la Bourse de Tokyo prises mensuelles du 01/01/2010 au 01/01/2012. La distribution de valeurs ou rendement de ces données est normal.

L'application de l'algorithme MAD-AG sur ce portefeuille donne lieu aux résultats illustrés sur les figures suivantes.

Selon la figures suivante, on remarque que les valeurs de portefeuille obtenues par notre algorithme sont supérieures à celles obtenues par les algorithmes génétiques de Markowitz.

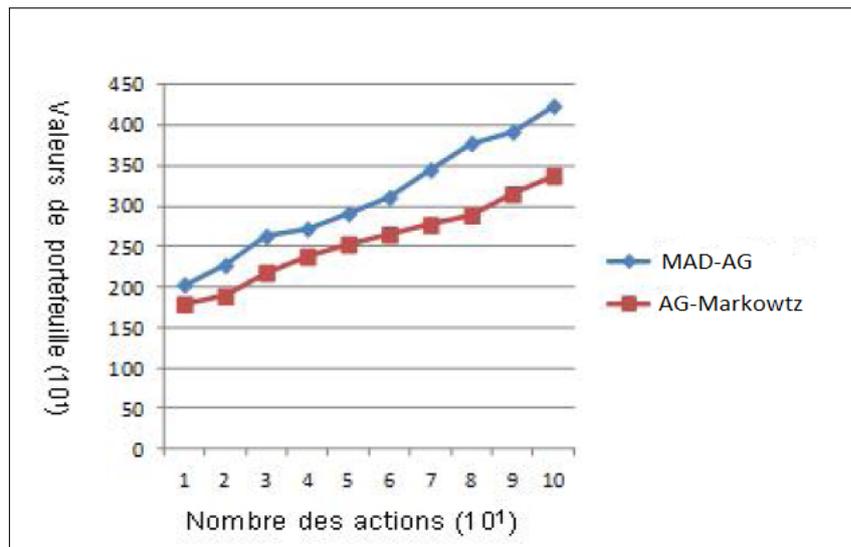


FIGURE 4.18: Représentation graphique de la valeur du portefeuille en utilisant l'algorithme MAD-AG et les AG-Markowitz.

En outre, d'après la figure suivante, les risques de portefeuille obtenues par notre algorithme sont inférieures à celles obtenues par les algorithmes génétiques de markowitz. Ces résultats de simulation sont performants et montrent la validité de notre approche proposée.

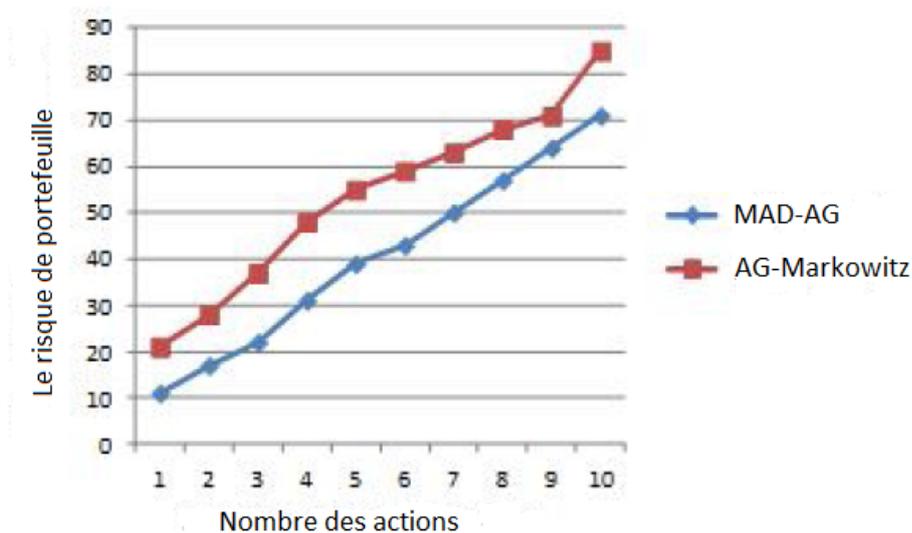


FIGURE 4.19: Représentation graphique du risque de portefeuille en utilisant l'algorithme et les AG-Markowitz

### 4.8.6 Exemple 06

Les exemples numériques illustrent le succès de l'algorithme proposé MAD-PAG et L'algorithme MAD-PSO . Les données sont d'Euronext (Axa, Air france, KLM, Michelin, ..., Nicox), nous considérons un portefeuille contenant 70 actions de la bourse Euronext paris pris mensuellement du 01-01-2015 au 01-01-2016. [69]

Asset Number	$L^{(1)}$	$L^{(2)}$	$L^{(3)}$	$L^{(4)}$	...	$L^{(12)}$	Mean return ( $\bar{r}_i$ )
<b>1</b>	0.19	0.17	0.18	0.20	...	0.19	0.23
<b>2</b>	0.03	0.10	0.07	0.08	...	0.06	0.89
<b>3</b>	0.37	0.39	0.27	0.29	...	0.32	1.12
<b>4</b>	0.11	0.13	0.19	0.16	...	0.17	0.03
<b>⋮</b>	<b>⋮</b>	<b>⋮</b>	<b>⋮</b>	<b>⋮</b>	<b>⋮</b>	<b>⋮</b>	<b>⋮</b>
<b>70</b>	0.22	0.19	0.21	0.20	...	0.18	0.65

Nous avons utilisé l'analyse en composantes principales (PCA) pour réorganiser les meilleures actions comme suit :

Reordering assets	Asset Number	Company	MAD ( $L_i$ )	Mean return ( $\bar{r}_i$ )
<b>1</b>	44	VINCI	0.15	2.96
<b>2</b>	16	ATOS	0.32	8.88
<b>3</b>	27	BOUYGUES	0.25	3.12
<b>4</b>	12	AIRBUS	0.46	4.16
<b>5</b>	56	MICHELIN N	0.27	6.11
<b>6</b>	23	PERNOD RICARD	0.29	1.55
<b>7</b>	67	KERING (Ex : PPR)	0.16	1.96
<b>8</b>	39	PETROLE BRENT	0.04	2.22
<b>9</b>	07	FR BENCHMARK 10A	0.13	3.41
<b>10</b>	32	ESSILOR INTL	0.06	0.96

A partir du tableau, nous avons utilisé les 10 actifs (actions) pour construire le portefeuille optimal par l'algorithme proposé **MAD-AG**. Le but de l'initialisation de l'algorithme AG est d'utiliser les 10 parts comme gène pour obtenir des chromosomes réalisables pour la première population. Nous pouvons utiliser les autres actifs (actions) dans le processus de mutation. Nous générons le vecteur de poids des chromosomes de manière aléatoire. Aussi en utilise les 10 actifs (actions) pour construire le portefeuille optimal par l'algorithme proposé **MAD-PSO**.

Après les expériences, la valeur optimale de chaque génération est montrée dans la figure 09, par contre, nous nous intéressons à l'analyse des performances de MAD-AG et à la comparaison avec MV-AG (Moyenne-Variance avec algorithme génétique), la figure montre que la moyenne obtenue avec l'algorithme MAD-AG est meilleure à 95% que les résultats de l'algorithme MV-AG. Les portefeuilles MAD-AG et MAD-PSO optimaux ont été sélectionnés pour chacune des 500 générations simulées (scénarios). La solution

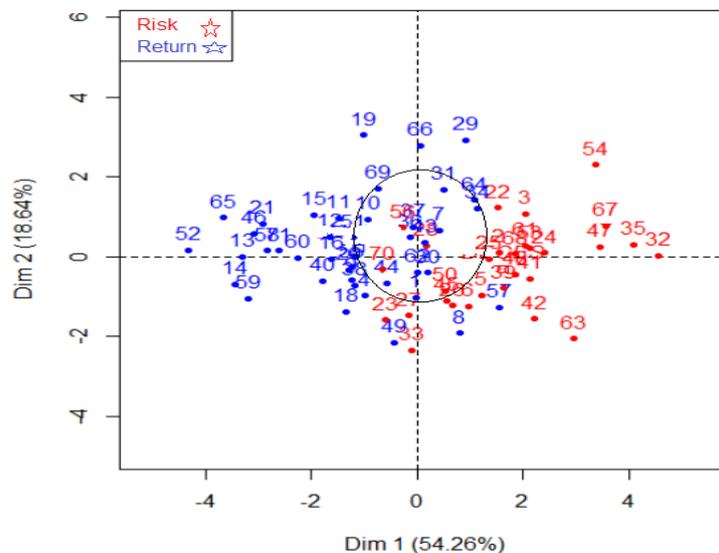


FIGURE 4.20: Analyse en composantes principales (PCA) pour 70 actions

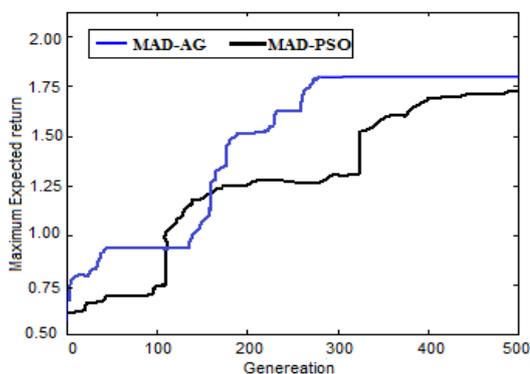


FIGURE 4.21: Le rendement espéré de MAD-PAG et MAD-PSO

optimale pour le problème de test indiqué dans le tableau :

Les actifs (action)	1	2	3	4	5	6	7	8	9	10
Proportion via MAD-AG	0.07	0.11	0.05	0.30	0.16	0.03	0.14	0.04	0.08	0.02
Proportion via MAD-PSO	0.06	0.13	0.05	0.31	0.14	0.02	0.15	0.03	0.10	0.01

La proportion du titre (actifs) utilisée pour représenter le poids pour la construction d'un portefeuille, pour l'investisseur qui accepte l'écart absolu moyen comme risque, la proportion représente l'allocation d'argent à chaque Actif (Actions) ; où  $\sum_{i=1}^n prop = 1$ .

## Conclusion

Dans ce travail nous avons présenté des approches visant l'optimisation de portefeuille d'actifs financiers en combinant l'intelligence artificielle et les méthodes statistiques.

Nous avons utilisé les méthodes dite métaheuristiques (algorithmes génétiques, algorithme essaim particulaire) pour réaliser deux algorithmes l'une appelé MAD-AG, qui permet de minimiser le risque mesuré par la méthode de konno yamazaki pour une valeur de portefeuille donnée dans une première étape, puis de la maximiser le rendement dans une seconde étape, et ce d'une manière dynamique, et l'autre appelé MAD-PSO, qui permet de minimiser le risque mesuré par la méthode de konno yamazaki pour une valeur de portefeuille donnée dans une première étape, puis de la maximiser la rentabilité dans une seconde étape, et ce d'une manière dynamique. Les deux algorithmes construire des portefeuilles efficaces.

Le résultat obtenu par ce processus donne lieu à une valeur de portefeuille supérieure à celle du portefeuille fixée à la première étape et à un risque inférieur à celui obtenu au niveau de la même étape.

D'après le MAD-AG nous avons montré que l'augmentation de la population initiale fait peu de différence avec les résultats, sauf pour le temps d'exécution de l'algorithme Il n'y a pas beaucoup de différence de temps de calcul entre les populations initiales de 50 et 100 (seulement une question de secondes), mais cela s'allonge considérablement lorsqu'elles sont augmentées à 500. La taille de la population finale est également très différente.

La méthode proposée MAD-AG ait donné globalement de meilleurs résultats que l'autre de Markowitz ; pour un risque  $\alpha$  la rentabilité par le MAD-AG est la meilleur par rapport la rentabilité de Markowitz.

Un algorithme génétique et algorithme essaim particulaire est désigné pour résoudre les problèmes d'optimisation correspondants, car ces problèmes de maximisation nonconcave sont avec une structure particulière et ne peuvent pas être résolus efficacement par les méthodes d'optimisation traditionnelles existantes.

Les données financières d'Euronext Stock Exchange ont été analysées par l'utilisation de l'écart absolu moyen et les algorithmes génétiques ou l'algorithme sessaim particulaire pour construire le portefeuille optimal. l'algorithmes génétiques ont été sélectionnés, alors que le portefeuille optimal est généralement difficile à obtenir. 10 actifs (actions) avec les meilleurs ratios ont été sélectionnés et comparés à plus de 70 actifs pour créer les meilleurs

portefeuilles (en MAD-PAG appelés meilleurs chromosomes). notre programme propose un nouveau modèle pour soutenir la prise de décision pour les portefeuilles d'investissement rapidement et efficacement. L'exemple numérique a été donné pour prouver et tester les algorithmes MAD-PAG et MAD-PSO et pour éviter que le nouveau modèle proposé une orientation théorique efficace et une prise de décision pour les investisseurs.

# Bibliographie

- [1] Aftalion, F., Poncet, P. Portait, R. : La théorie moderne du portefeuille. FeniXX, (1998).
- [2] Alfaro, C.E., Baixauli, J.S., Fernandez, M.O. : Minimising value-at-risk in a portfolio optimisation problem using a multi-objective genetic algorithm. *International Journal of Risk Assessment and Management* 15(5-6), 453-477 (2011).
- [3] Amédée, S., Francois-Gérard, R. : 'ALGORITHMES GENETIQUES'. TE de fin d'année (2004).
- [4] Arnaud, A.J. : Jean-louis le moigne, la modélisation des systèmes complexes, 1990. *Droit et Société* 19(1), 424-424 (1991).
- [5] Artzner, P., Delbaen, F., Eber, J.M., Heath, D. : Coherent measures of risk. *Mathematical finance* 9(3), 203-228 (1999).
- [6] Bawa, V.S. : Research bibliography—stochastic dominance : A research bibliography. *Management science* 28(6), 698-712 (1982).
- [7] Belovas, I., Kabasinskas, A., Sakalaukas, L. : MAD Approach to the Management of the Portfolio with Stable Assets : ASMDA. *Proceedings of the International Conference Applied Stochastic Models and Data Analysis 2009*, p. 491. Vilnius Gediminas Technical University, Department of Construction Economic.
- [8] Ben-Tahar, I. : *Gestion de Portefeuilles*, Université Paris Dauphine, Mathématiques de la Modélisation et de la Décision, 2011.
- [9] Beyer, H.G. : *The theory of evolution strategies*. Springer Science and Business Media, (2001).
- [10] Blais, S. : *Théories avancées de portefeuille*. Département des sciences administratives, UQO. 15 avril 2014.
- [11] Bernard, P. : *Le Modele d'Equilibre des Actifs Financiers*. PhD thesis, Université Paris-Dauphine (2007).
- [12] Brown, D.S. : *Learning and control techniques in portfolio optimization*. Brigham Young University (2011).
- [13] Chapados, N. : *Critères d'optimisation d'algorithmes d'apprentissage en gestion de portefeuille*. Université de Montréal (2000).

- [14] Chabbar, S., Lallouch, S. : Optimisation De Portefeuille Sous Des Contraintes De Risque, Université Mohamed V Agdal - Rabat Ecole Mohammadia D'ingénieurs, Soutenu le : 01/06/2015.
- [15] Clerc, M., Kennedy, J. : The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE transactions on Evolutionary Computation* 6(1), 58-73 (2002).
- [16] Cooren, Y. : Perfectionnement d'un algorithme adaptatif d'Optimisation par Essaim Particulaire : application en génie médical et en électronique. Université Paris-Est (2008).
- [17] Coulon, J. : Mémoire longue, volatilité et gestion de portefeuille. Université Claude Bernard-Lyon I (2009).
- [18] Durand, N. : Optimisation de trajectoires pour la résolution de conflits aériens en route. INPT (1996).
- [19] Cai, X., Teo, K.-L., Yang, X., Zhou, X.Y. : Portfolio optimization under a minimax rule. *Management science* 46(7), 957-972 (2000).
- [20] Doucoure, D. : Théorie du portefeuille. (2018).
- [21] El Dor, A. : Perfectionnement des algorithmes d'optimisation par essaim particulaire : applications en segmentation d'images et en électronique. Université Paris-Est (2012).
- [22] El Hachloufi, M. : Les Apports de l'Intelligence Artificielle aux Approches Probabilistes pour l'Optimisation de Portefeuille d'Actifs Financiers. (2013).
- [23] Fabozzi, F.J., Focardi, S.M., Kolm, P.N. : Quantitative equity investing : Techniques and strategies. John Wiley and Sons, (2010).
- [24] Fabozzi, F.J., Markowitz, H.M. : The theory and practice of investment management : Asset allocation, valuation, portfolio construction, and strategies, vol. 198. John Wiley and Sons, (2011).
- [25] Fatou D. B., Abdoulaye W. Le modèle de Markowitz et détermination d'un portefeuille optimal.. Unité de Formation et de Recherche Sciences Appliquées et Technologie. 2012.
- [26] Feigenbaum, E.A. : The simulation of verbal learning behavior. In : Papers presented at the May 9-11, 1961, western joint IRE-AIEE-ACM computer conference 1961, pp. 121-132.
- [27] Feigenbaum, E.A., Feldman, J. : Computers and thought. New York McGraw-Hill, (1963).
- [28] Fogel, L.J. : Autonomous automata. *Industrial research* 4, 14-19 (1962).
- [29] Fogel, L.J., Owens, A.J., Walsh, M.J. : Artificial intelligence through simulated evolution. (1966).
- [30] Gherboudj, A. : Méthodes de résolution de problèmes difficiles académiques. Université de Constantine2 (2013).
- [31] Goldberg, D.E., Deb, K. : A comparative analysis of selection schemes used in genetic algorithms. In : Foundations of genetic algorithms, vol. 1. pp. 69-93. Elsevier, (1991).
- [32] Grainia, S. : L'algorithme de Branch and Price and Cut pour le problème de conception de réseaux avec coûts fixes et sans capacité. (2015).

- [33] Glover, F., Laguna, M. : Tabu search. In : Handbook of combinatorial optimization. pp. 2093-2229. Springer, (1998).
- [34] Hamza, F., Janssen, J. : Linear approach for solving large-scale portfolio optimization problems in a lognormal market. In : IAA/AFIR Colloquium, Nürnberg, Germany 1996, pp. 1019-1039.
- [35] Hanoch, G., Levy, H. : The efficiency analysis of choices involving risk. *The Review of Economic Studies* 36(3), 335-346 (1969).
- [36] Hardy, G., Littlewood, J. : and G. Polya, *Inequalities*. In. Cambridge University Press, (1934)
- [37] Jolliffe, I.T. : Principal components in regression analysis. In : *Principal component analysis*. pp. 129-155. Springer, (1986).
- [38] Kariya, T. : Distribution of Stock Prices in the Stock Market of Japan. In. Toyo Keizai Publishing Co., Tokyo, (1989).
- [39] Ke, J., Yu, Y., Yan, B., Ren, Y. : Asset risk diversity and portfolio optimization with genetic algorithm. In : *Recent Advances on Applied Mathematics and Computational Methods in Engineering, Proceedings of the International Conference on Applied Mathematics and Computational Methods in Engineering (AMCME 2015)*.
- [40] Khelil, N., Rahmani, N., Djerou, L. : Improvement of Euler's Method Using Particle Swarm Optimization. *Journal of Mathematics and System Science* 2(9), 535 (2012).
- [41] Khelil, N., Djerou, L., Rahmani, N., Dakhia, H. : On the application of Particle Swarm Optimization to Differential Equations. *Conference on META'2014, Merrakech 2014*.
- [42] Konno, H. : Piecewise linear risk function and portfolio optimization. *Journal of the Operations Research Society of Japan* 33(2), 139-156 (1990).
- [43] Konno, H., Yamazaki, H. : Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management science* 37(5), 519-531 (1991).
- [44] Konno, H., Suzuki, K.-i. : A fast algorithm for solving large scale mean-variance models by compact factorization of covariance matrices. *Journal of the Operations Research Society of Japan* 35(1), 93-104 (1992).
- [45] Konfršt, Z. : *The mean-variance model*. In. Technical Computing Prague, (2013).
- [46] Kennedy, J., Eberhart, R. : Particle swarm optimization. In : *Proceedings of ICNN'95-international conference on neural networks 1995*, pp. 1942-1948. IEEE
- [47] Koza, J.R. : Hierarchical genetic algorithms operating on populations of computer programs. In : *IJCAI 1989*, pp. 768-774
- [48] Koza, J.R. : *Genetic programming : A paradigm for genetically breeding populations of computer programs to solve problems*, vol. 34. Stanford University, Department of Computer Science Stanford, CA, (1990)
- [49] Lai, T.Y. : Portfolio selection with skewness : a multiple-objective approach. *Review of Quantitative Finance and Accounting* 1(3), 293-305 (1991).
- [50] Lai, T.L., Xing, H. : *Statistical models and methods for financial markets*. Springer, (2008).
- [51] Levy, H. : Stochastic dominance and expected utility : Survey and analysis. *Management science* 38(4), 555-593 (1992).

- [52] Lin, C.M., Gen, M. : An effective decision-based genetic algorithm approach to multiobjective portfolio optimization problem. *Applied Mathematical Sciences* 1(5), 201-210 (2007).
- [53] Liu, B. : Some research problems in uncertainty theory. *Journal of Uncertain systems* 3(1), 3-10 (2009).
- [54] Lin, P.C. : Portfolio optimization and risk measurement based on non-dominated sorting genetic algorithm. *Journal of Industrial and Management optimization* 8(3), 549-564 (2012).
- [55] Lutton, E. : Algorithmes génétiques et Fractales. Habilitation à diriger des recherches, Université Paris XI Orsay, France (1999).
- [56] Mansini, R., Ogryczak, W., Speranza, M.G. : On LP solvable models for portfolio selection. *Informatica* 14(1), 37-62 (2003).
- [57] Markowitz, H. : Portfolio Selection, *Journal of Finance*. Markowitz HM-1952-N, 77-91 (1952).
- [58] MARKOWITZ, H. : The Nature and Applications of Process Analysis. In. RAND CORP SANTA MONICA CALIF, (1954).
- [59] Markowitz, H. : Portfolio selection, efficient diversification of investments. J. Wiley, (1959).
- [60] Markowitz, H. : Sparsity and piecewise linearity in large scale portfolio optimization problems. *Sparse matrices and their uses*, 89-108 (1981).
- [61] Mayambala, F. : Mean-Variance Portfolio Optimization : Eigendecomposition-Based Methods. Linköping University Electronic Press (2015).
- [62] Moullier, I. : Florin Aftalion, L'économie de la Révolution française. In : *Annales historiques de la Révolution française* 2008.
- [63] Mostafa, E.H., Mohammed, E.H., Abderrahim, E.A. : Minimization of value at risk of financial assets portfolio using genetic algorithms and neural networks. *Journal of Applied Finance. Banking* 6(2) (2016).
- [64] Moeini, M. : La programmation DC et DCA pour l'optimisation de portefeuille. Université Paul Verlaine-Metz (2008).
- [65] Peyrard, J. : Risque, rendement et évaluation des actifs financiers. *Journal de la société française de statistique* 114, 113-129 (1973).
- [66] Pearl, J. : Heuristics : intelligent search strategies for computer problem solving. (1984).
- [67] Osman, I.H., Laporte, G. : Metaheuristics : A bibliography. In. Springer, (1996).
- [68] Ogryczak, W., Ruszczyński, A. : From stochastic dominance to mean-risk models : Semideviations as risk measures. *European journal of operational research* 116(1), 33-50 (1999).
- [69] Rahmani, N., Khelil, N. : Portfolio optimization using pca classification and genetics algorithm. *Journal of Global Economics, Management and Business Research*, 129-141 (2019).
- [70] Ramkumar, S. : Algorithmes génétiques. Rapport, Université Libre de Bruxelles, Belgique 26 (2007).

- [71] Rankovic, V., Drenovak, M., Stojanovic, B., Kalinic, Z., Arsovski, Z. : The mean-value at risk static portfolio optimization using genetic algorithm. *Computer Science and Information Systems* 11(1), 89-109 (2014).
- [72] Rao, C.R., Statistiker, M. : *Linear statistical inference and its applications*, vol. 2. Wiley New York, (1973).
- [73] Rechenberg, I. : *Cybernetic solution path of an experimental problem*. Royal Aircraft Establishment Library Translation 1122 (1965).
- [74] : Rockafellar, R.T., Uryasev, S. : Conditional value-at-risk for general loss distributions. *Journal of banking & finance* 26(7), 1443-1471 (2002).
- [75] Roncalli, T. : *Introduction à la Gestion des Risques*. Groupe de Recherche Opérationnelle Crédit Lyonnais, p21 (2001).
- [76] Rothschild, M., Stiglitz, J.E. : Increasing risk : I. A definition. *Journal of Economic theory* 2(3), 225-243 (1970).
- [77] Ringnér, M. : What is principal component analysis? *Nature biotechnology* 26(3), 303-304 (2008).
- [78] Sadaf, A., Ghodrati, H. : An Improved Genetic Algorithm Method for Selection and Optimizing the Share Portfolio. *Int. J. of Computer Science and Mobile Computing* 4(1) (2015).
- [79] Salah, H.B. : *Gestion des actifs financiers : de l'approche Classique à la modélisation non paramétrique en estimation du DownSide Risk pour la constitution d'un portefeuille efficient*. Université Claude Bernard-Lyon I (2015).
- [80] Seck, B. : *Optimisation stochastique sous contrainte de risque et fonctions d'utilité*. Ecole des Ponts ParisTech (2008).
- [81] Sharpe, W.F. : A linear programming approximation for the general portfolio analysis problem. *Journal of Financial and Quantitative Analysis*, 1263-1275 (1971).
- [82] Shi, Y., Eberhart, R. : A modified particle swarm optimizer. In : 1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360) 1998, pp. 69-73. IEEE
- [83] Slagle, J.R., Koniver, D.A. : Finding resolution proofs and using duplicate goals in AND/OR trees. *Information Sciences* 3(4), 315-342 (1971).
- [84] Speranza, M.G. : *Linear programming models for portfolio optimization*. (1993).
- [85] Stancu, S., Predescu, M. : „Genetic Algorithm for the Portfolio Selection Problem on the Romanian Capital Market”. In : *Proceedings of International Conference on Engineering an Meta-Engineering 2010*, pp. 57-60.
- [86] Vallée, T., Basar, T. : Incentive stackelberg solutions and the genetic algorithm. In : *Proceedings of the 8th International Symposium on Dynamic Games and Applications 1998*, pp. 633-639.
- [87] Witt, S.F., Dobbins, R. : *The Markowitz contribution to portfolio theory*. Managerial Finance (1979).
- [88] Young, M.R. : A minimax portfolio selection rule with linear programming solution. *Management science* 44(5), 673-683 (1998).
- [89] Zhang, R. : *Apprentissage statistique en gestion de portefeuille*. Télécom ParisTech (2014).

# Résumé

L'optimisation de portefeuille est l'un des principaux investisseurs sur les marchés financiers, nous présentons dans ce travail une nouvelle approche pour obtenir un portefeuille optimal, qui minimise le risque pour un profit requis ou maximise le profit d'un risque donné. Pour résoudre le problème, nous introduisons d'abord un concept de risque d'écart absolu moyen (MAD). La fonction de risque MAD ( $L_1$ ) peut éliminer la plupart des difficultés associées au modèle de Markowitz. Nous utilisons un algorithme évolutif heuristique pour trouver le portefeuille optimal.

Nous avons proposé une approche pour construire un portefeuille d'actions réalisable investi sur le marché basé sur MAD en utilisant PCA (analyse en composantes principales) et algorithme génétique (GA) et optimisation des essaims de particules (PSO). Cette approche est organisée en deux étapes : la première consiste à utiliser la méthode de classification PCA pour classer les actions en classes. Dans un deuxième temps, nous utilisons un algorithme d'optimisation appelé MAD-AG basé sur l'algorithme génétique et l'écart absolu moyen pour minimiser le risque mesuré par le MAD et maximiser la valeur du portefeuille et un autre algorithme d'optimisation appelé MAD-PSO basé sur l'optimisation par l'essaim de particules (PSO) et écart absolu moyen pour minimiser le risque mesuré par le MAD.

**Mots clés :** Ecart absolu moyen, optimisation de portefeuille, Analyse en composantes principales, Actifs financiers, Risque, Algorithmes génétiques (AG), Optimisation par l'essaim de particules (PSO).

# Abstract

Portfolio optimization is one of the main investors in financial markets, we present in this work a new approach to obtain an optimal portfolio, which minimizes the risk for a required profit or maximizing the profit of a given risk. To solve the problem we first introduce a concept of mean absolute deviation risk (MAD). The MAD ( $L_1$ ) risk function can remove most difficulties associated with the Markowitz's model. We use heuristic evolutionary algorithm to find the optimal portfolio.

We have proposed an approach to find a feasible shares portfolio invested in market based on MAD by using PCA (principal component analysis) and genetic algorithm (GA) and particle swarm optimization (PSO). This approach is organized in two steps: the first one is to use the PCA classification method to classify the actions into classes. In second step we use an algorithm of optimization called MAD-AG based on genetic algorithm and mean absolute deviation to minimize the risk measured by the MAD and maximize the value of portfolio and another algorithm of optimization called MAD-PSO based on particle swarm optimization (PSO) and mean absolute deviation to minimize the risk measured by the MAD.

**Keywords:** Mean absolute deviation, Portfolio optimization, Principal Component Analysis, Genetic algorithm (AG), particle swarm optimization (PSO).