

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE.
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE.

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU
FACULTE DES SCIENCES
DEPARTEMENT MATHEMATIQUES



MÉMOIRE DE MASTER
en
MATHÉMATIQUES

Spécialité
Recherche Opérationnelle : Aide à la décision

Thème

Le modèle de Markowitz et détermination d'un portefeuille optimal

Présenté par :
Dehrib Lila

Devant le jury d'examen composé de :

<i>M^r.Chebbah</i>	M.assist A	UMMTO	Président
<i>M^{me}.Leslous</i>	M.assist A	UMMTO	Examinatrice
<i>M^r.Ouanes</i>	Professeur	UMMTO	Rapporteur

soutenu : Septembre 2018

Remerciements

Je remercie Dieu Tous Puissant de m'avoir accordé la force, la volonté et

La patience afin d'accomplir ce modeste travail de recherche.

Je tiens à adresser mes sincères remerciements à mon promoteur, Mr OUANES

Mohand pour son aide toute au long de notre travail.

Je remercie également Mme LESLOUS et Mr CHEBBAH d'avoir accepter de

Juger mon travail.

Mes remerciements vont également à toute personne qui a contribué à la

Réalisation de ce travail, Bahia, Fadila, Karima,...

Merci tous

Dédicaces

À ma mère et mon père

À mes sœurs et mes frères

À mes amis filles et garçons

À toute ma famille.

Table des matières

Table des matières	1
Introduction générale	2
1 La convexité et les problèmes d'optimisation convexe	5
1.1 Introduction	5
1.2 Quelques définitions :[5][8][14][15]	5
1.3 Ensemble convexe[10]	7
1.4 Cône convexe	8
1.5 Fonctions convexes[10][14]	9
1.6 Fonctions convexes différentiables (on a deux cas)[10][14]	10
1.6.1 Rappels : (Caractérisation des fonctions convexes dérivables)[9]	10
1.7 Fonctions quasi convexe :	12
1.7.1 Les problèmes d'optimisation non linéaire :[23]	12
1.7.2 Formulation générale des problèmes d'optimisation non linéaire : . . .	13
1.7.3 les problèmes d'optimisation sans contraintes :[6][14]	14
1.8 Condition nécessaires d'optimalité locale :[14]	15
1.9 Condition suffisantes d'optimalité locale :[6]	16
1.10 Cas des fonctions convexes :(Condition nécessaire et suffisante d'optimalité globale)	17
1.11 Cas des fonctions quelconques (difficulté du problème général) :	17
1.11.1 Les problèmes d'optimisation avec contraintes d'égalité, d'inégalité et mixtes :[5][6]	18
1.11.2 Les problèmes d'optimisation avec contraintes d'égalité :	18
1.11.3 Les problèmes d'optimisation avec contraintes d'inégalité :	19
1.11.4 Les problèmes d'optimisation avec contraintes mixtes :	21
1.12 La dualité de Lagrange :[21]	22
1.13 La Dualité :[14]	22
1.14 la programmation quadratique :[19]	24
1.14.1 Définition :	24
1.14.2 Problème quadratique avec contraintes d'égalités :[14][19]	24
1.14.3 Problème quadratique avec contraintes d'inégalités :[19]	26
1.14.4 Dualité en programmation quadratique convexe :[14]	28
1.14.5 Conclusion :	28

2	Le modèle de Markowitz	29
2.1	Introduction [1]	29
2.2	Les hypothèses du modèle de Markowitz [1][17]	29
2.2.1	Les hypothèses relatives aux actifs financiers	29
2.2.2	Les hypothèses relatives aux comportements des investisseurs	30
2.3	Présentation du modèle de Markowitz [7][17]	30
2.3.1	Introduction	30
2.3.2	Rentabilité et risque : [3]	31
2.3.3	L'approche espérance-variance :[17]	33
2.4	Portefeuille optimal :(optimisation d'un portefeuille)[4]	33
2.4.1	Comprendre la nécessité d'une gestion de portefeuille :	33
2.4.2	La diversification :[2][4]	34
2.4.3	Les actions, les portefeuilles d'actions :[1][8][16][20]	35
2.4.4	La frontière efficiente :[1][7]	37
3	Application	47
	Conclusion générale et perspectives	51
	Bibliographie	52

Introduction générale

L'optimisation [5] est un concept qui fait partie intégrante de la vie courante. Citons quelques exemples tout à fait banals, mais représentatifs :

- a. Quel est le meilleur itinéraire pour aller d'un point A à un point B en voiture ?
- b. Au tennis, comment maximiser l'effet, la vitesse d'une balle de service ?
- c. Peut-on gagner contre la banque à la roulette au casino ?
- d. A la bourse, comment maximiser les profits tout en minimisant les risques ?
- e. Pourquoi tel composant chimique réagit-il avec tel autre ? etc.

Une stratégie raisonnable est d'essayer de modéliser chacun de ces problèmes, c'est-à-dire de les reformuler sous une forme mathématique, puis de résoudre/optimiser les modèles mathématiques ainsi obtenus, et enfin de tester les résultats sur les situations pratiques. De fait, cette activité est du ressort du physicien, du chimiste, de l'économiste, du mathématicien ...

Parmi tous les problèmes rencontrés par le chercheur, l'ingénieur, le gestionnaire ou l'économiste, les problèmes d'optimisation occupent une place de premier ordre. Ils apparaissent dans divers domaines tels que la médecine, les télécommunications, le secteur économique. Par ailleurs, parmi ces problèmes d'optimisation discrets, un grand nombre, difficiles à résoudre, se formulent aisément comme des programmes mathématiques avec un objectif quadratique convexe.

L'ingénieur, à qui revient la charge de résoudre ces modèles, se doit de les bien connaître, notamment en ce qui concerne les hypothèses sous lesquelles le modèle est valide, avant d'envisager leur résolution.

Le domaine de l'économie offre un large éventail d'applications. En effet, la mesure de l'efficacité des systèmes étudiés s'exprime sous forme de rapport entre les critères techniques et/ou économiques, comme : Rendement/risque : Kallberg Ziemba (modèle concave), Rendement/risque : Harry Markowitz (modèle convexe quadratique), Rendement/investissement : Heinen, Mjelde (modèle concave quadratique avec contraintes linéaires) ; Coût/temps : Derman, Klein (modèle linéaire), Productivité : Gutenberg (modèle linéaire.).

Exemples :

- **Téléphonie mobile.**
- **Equilibre d'un fil pesant.**

- **Gestion d'énergie.**
- **Optimisation de portefeuille.**

Ce mémoire comprend trois chapitres, au premier chapitre nous présentons les bases théoriques et numériques de l'optimisation convexe et quadratique car elle est très utilisée dans les applications pratiques et théoriques. Au second chapitre, nous présentons le modèle de Markowitz pour la gestion de portefeuille. Une application est présentée au troisième chapitre, nous terminons ce mémoire par une conclusion et des perspectives.

La convexité et les problèmes d'optimisation convexe

1.1 Introduction

En général optimiser signifie le fait de chercher une configuration optimale d'un système, c'est à dire, chercher la meilleure configuration parmi tous les configurations possibles du système et ceci, par rapport à un critère donné.

Pour décrire (et éventuellement résoudre) un problème d'optimisation nous utilisons la modélisation mathématique, donc l'optimisation est une branche des mathématiques. Dans la pratique, on part d'un problème concret, on le modélise et on le résout mathématiquement (analytiquement : problème d'optimisation, numériquement : programme mathématique).

1.2 Quelques définitions :[5][8][14][15]

Définition1 :[14]

Dans les problèmes d'optimisation, avec ou sans contraintes, une notion va jouer un rôle très important : celle de convexité. En effet, pour la plupart des algorithmes que nous décrirons, la convergence vers un optimum global ne pourra être démontrée qu'avec des hypothèses de convexité (qu'il sera parfois possible d'affaiblir en ayant recours à la notion de quasi-convexité).

Définition2 :

- L'analyse convexe : est une théorie située entre l'algèbre linéaire et l'analyse non Linéaire, dans laquelle les objets étudiés, bien que non linéaires, sont contraints de vérifier une condition particulière qui leur confère des propriétés remarquables. Cette théorie et ses concepts interviennent en optimisation pour de nombreuses raisons. Par exemple, l'écriture des conditions d'optimalité passe par le linéarisé de l'ensemble admissible, qui est un cône ; un objet qui n'appartient pas à l'algèbre linéaire mais à l'analyse convexe. Autre exemple : dans les problèmes d'optimisation convexe (critère et ensemble admissible convexes), tous les points stationnaires sont des minima globaux, ce qui simplifie singulièrement le problème.

Définition3 :[14]

- Programmes convexes : on dit qu'un problème de programmation mathématique est convexe

s'il consiste à minimiser une fonction convexe (resp : maximiser une fonction convexe) sur un domaine convexe fermé.

La propriété fondamentale des programmes convexes apparaît alors dans le résultat suivant :

Théorème3 :

Pour un programme convexe, tout optimum local est un optimum global. Pour plus de détails, on pourra se reporter à l'ouvrage de référence [14].

Définition4 :

• Fonctions convexes : soit E un espace vectoriel sur R . On note $R = R \cup \{-\infty, +\infty\}$ la droite achevée.

En analyse convexe et en optimisation, il est parfois intéressant de pouvoir considérer des fonctions pouvant prendre des valeurs infinies. En analyse convexe, plutôt que de définir une fonction convexe comme un couple forme d'un ensemble convexe U et d'une fonction f définie sur U ayant une propriété bien particulière (une approche que l'on rencontre encore parfois), il est préférable de dire que cette fonction prend des valeurs infinies en dehors de U . On évite ainsi de devoir décrire U avant de définir f , en particulier lorsque f est construite par une des opérations non triviales que nous verrons dans lesquels il est compliqué de déterminer l'ensemble U .

Ce travail se situe donc sur le terrain de l'optimisation non linéaire, et plus exactement de l'optimisation convexe, comme le suggère son titre.

Remarque :

Comme on vient donc de le voir, l'objet principal est celui de l'optimisation :

- A un seul critère (sous contraintes),
- déterministe,
- en variables continues,
- et dans un cadre convexe.

Ce dernier point signifie que les fonctions du problème (coût et contraintes) seront supposées convexes (et même affines pour les contraintes d'égalité). On parle alors d'optimisation ou de programmation convexe.

Ce sujet de recherche a connu un fort développement après le succès initial de la programmation linéaire dans les années soixante[14] et il s'est constitué en une théorie cohérente représentant la première incursion structurée dans le domaine de l'optimisation non linéaire. Bien que la convexité ait une interprétation économique claire (phénomène des coûts marginaux croissants), beaucoup de problèmes rencontrés dans la réalité ne sont pas convexes. Cependant, puisque l'on s'intéresse à l'optimisation de fonctions disons à la minimisation pour fixer les idées au voisinage immédiat de leur minimum, les fonctions ont nécessairement une forme convexe. L'hypothèse de convexité globale peut donc être considérée comme une idéalisation (nécessaire dans toute théorie mathématique) ou l'on pourra étudier sur tout l'espace (globalement) des phénomènes qui risqueraient de n'apparaître en réalité que localement dans la région qui nous intéresse. Tant que l'on n'utilise que des concepts locaux (variations infinitésimales autour d'un point), il y a peu de chances que l'aspect qualitatif des phénomènes change radicalement du fait de cette extension à tout l'espace de propriétés seulement locales. On s'est simplement placé dans un cadre mathématique plus agréable. Cependant, dès que l'on commence à tirer des conclusions globales d'hypothèses de convexité (et on rencontrera plusieurs circonstances de cette nature), alors il y a fort à parier que ces conclusions soient qualitativement différentes si on renonce aux hypothèses de convexité, avant d'aborder le sujet de l'optimisation proprement.

En résumé, l'optimisation est une branche des "sciences de la décision" : son but est de

spécifier la décision souhaitable parmi toutes les possibilités d'un espace de décisions, par le choix d'un "critère" ou "fonction coût" et par le choix de contraintes. Ces choix préalables sont "arbitraires", mais une fois ceux-ci arrêtés, une décision devient "meilleure" que toutes les autres. Il s'agit alors de :

- caractériser cette décision pour la reconnaître (conditions d'optimalité) car la définition initiale de l'optimalité n'est pas utilisable de façon opérationnelle.
- puis de la calculer (algorithmes).

La théorie mathématique de l'optimisation est traitée de ces deux questions mais seule la première qui est abordée.

La convexité est à la base une propriété géométrique, assez intuitive d'ailleurs, qui permet de caractériser certains objets. On voit assez bien ce qu'est un objet convexe dans un espace à deux ou trois dimensions. Nous allons maintenant montrer comment cette propriété peut aussi s'appliquer aux fonctions de R^n dans R . [5][6][8][14]

1.3 Ensemble convexe[10]

Définition1

Un ensemble $U \subset R^n$ est dit convexe si pour tout couple $(x, y) \subset U^2$ et $\forall \lambda \in [0, 1]$ on a

$$\lambda x + (1 - \lambda)y \in U$$

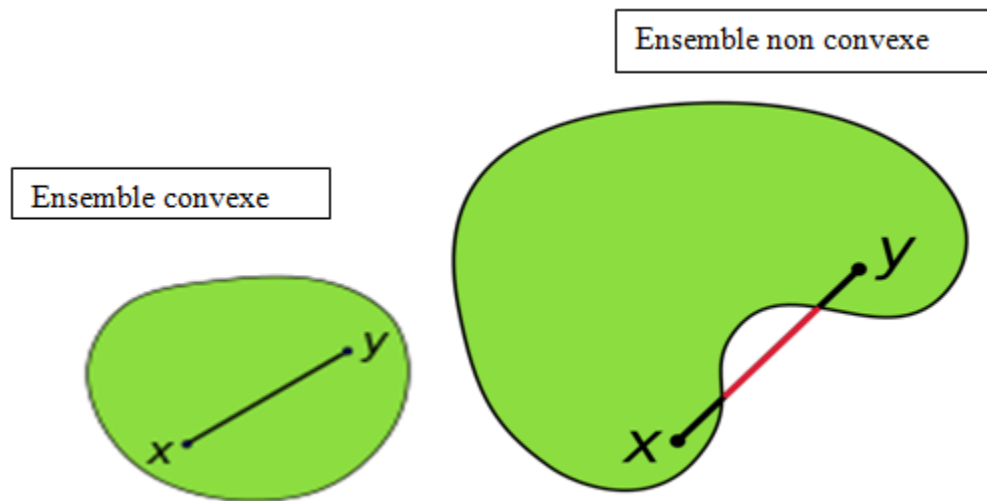


FIGURE 1.1 – La représentation d'un ensemble convexe et d'un ensemble non convexe.

D'une façon équivalente, on peut dire que U est convexe si et seulement si, pour deux points quelconque x et y pris dans U , le segment $[x, y]$ est tout entier contenu dans U . plus généralement, étant donné P points de R^n (x^1, x^2, \dots, x^p) on dit que $x \in R^n$ est combinaison convexe de ces P points s'il existe des coefficients $\lambda_1, \lambda_2, \dots, \lambda_p$ ($\lambda_i \geq 0, \forall i = 1, 2, \dots, p$) tels que :

$$x = \sum_{i=1}^p \lambda_i x^i$$

avec $\lambda_i \geq 0$ et $\sum_{i=1}^p \lambda_i = 1$

On vérifie aisément qu'un ensemble $U \subset \mathbb{R}^n$ est convexe si et seulement si tout point, combinaison convexe de points de U , est dans U . On peut citer quelques cas particuliers : \mathbb{R}^n tout entier est un ensemble convexe, de même qu'un singleton $\{a\}$.

Par convention l'ensemble vide $\{\phi\}$ est convexe.

Propriété 1 :

Soit une famille $\{U_i\}_{i=1 \dots n}$ d'ensembles convexes et $S = \cap_{i=1}^n U_i$ Alors S est convexe.

1.4 Cône convexe

Dans la formulation des contraintes d'inégalités, les cônes joueront ultérieurement un rôle essentiel.

Définition 2

Un sous-ensemble U est un cône si :

$$\forall x \in U, \forall \lambda \geq 0 \Rightarrow \lambda x \in U$$

Un cône est donc une union de demi-droites fermées issues de l'origine.

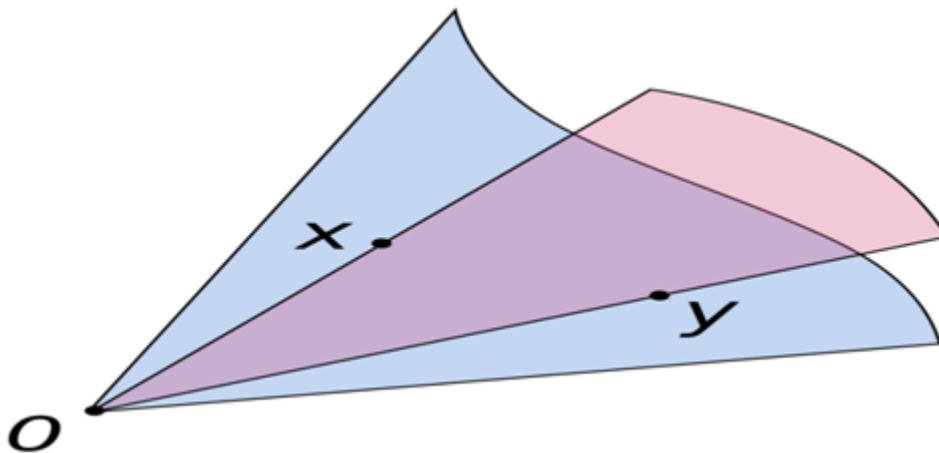


FIGURE 1.2 – La représentation d'un cône convexe.

1.5 Fonctions convexes[10][14]

Définition3

- Soit $U, U \neq \emptyset \subset \mathbb{R}^n$, convexe.

La fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est dite convexe dans U si, $\forall \lambda \in [0, 1], \forall x, y \in U$,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

- Fonctions affines : une fonction $f : U \rightarrow \mathbb{R}$ est dite affine si elle vérifie pour tout $x, y \in U$ et tout $\lambda \in \mathbb{R} : f((1 - \lambda)x + \lambda y) = (1 - \lambda)f(x) + \lambda f(y)$ (★)

Remarque

Avec la convention $(+\infty) + (+\infty) = +\infty$, $(+\infty) + a = +\infty$ pour tout $a \in \mathbb{R}$, et $(+\infty)a = (a \neq 0) \infty$ pour tout $a \in \mathbb{R}$ différent de 0.

- Soit $U, U \neq \emptyset \subset \mathbb{R}^n$, convexe.

la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est dite concave dans U si $(-f)$ est convexe dans U .

- Par convention, la fonction identiquement égale à $(-\infty)$ ou $(+\infty)$ est à la fois convexe et concave.

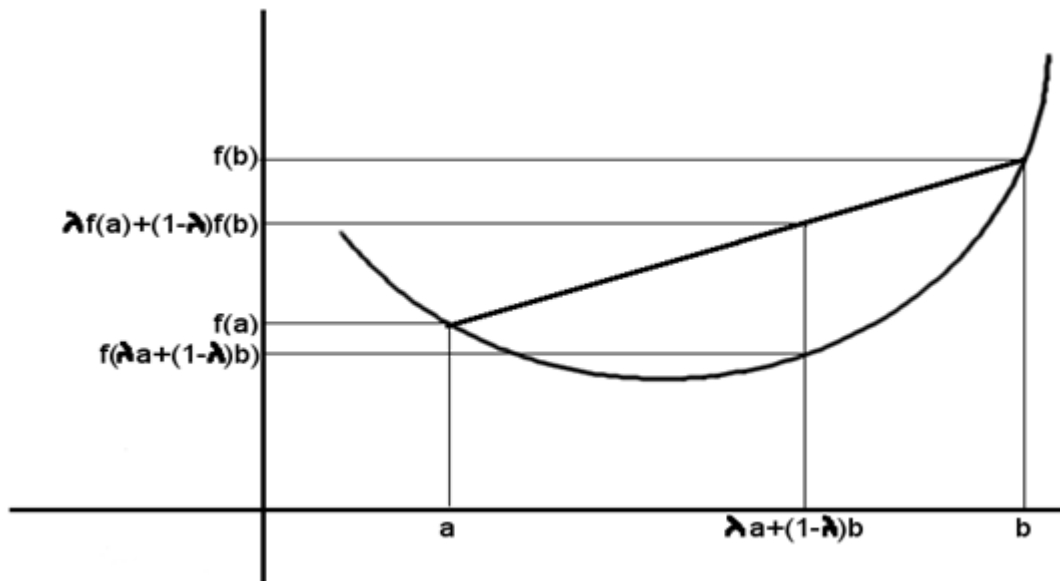


FIGURE 1.3 – la représentation d'une fonction convexe.

Soit $U, U \neq \emptyset \subset \mathbb{R}^n$,

- On appelle domaine de f noté par $dom(f)$ le sous-ensemble de U où f prend des valeurs finies :

$$dom(f) = \{x \in U \mid f(x) < +\infty\}$$

- L'épigraphe d'une fonction f , noté $epi(f)$, est l'ensemble des vecteurs à $(n + 1)$ composantes

(s, x) tels que $f(x) \leq s$:

$$\{(s, x) / f(x) \leq s, x \in R^n, s \in R\} \subset R^{n+1}$$

Propriété(épi(f)) :

Une fonction f est convexe si et seulement si $\text{épi}(f)$ est un ensemble convexe.

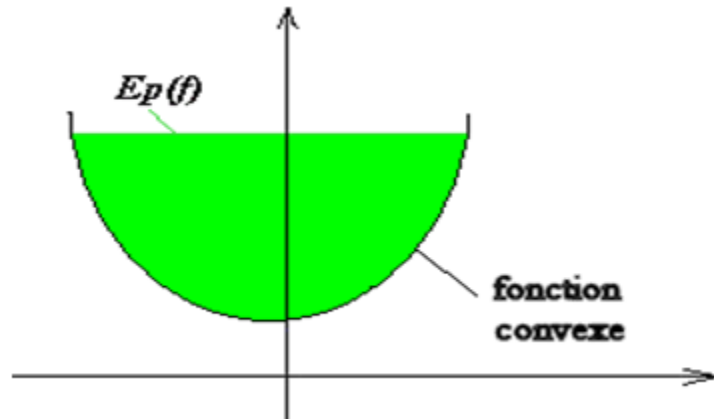


FIGURE 1.4 – La représentation de l'épigraph d'une fonction convexe.

On introduit maintenant les raffinements suivants.

- Une fonction f est strictement convexe si elle est convexe et si l'inégalité stricte a lieu dans (★) dès que $x \neq y$ et $\lambda \in]0, 1[$.

Autrement dit, le graphe de la fonction n'a pas de partie affine.

- Une fonction f est fortement convexe (de module b) si, $\exists b > 0 : \forall x, y \in U, \forall \lambda \in [0, 1] :$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - b\lambda(1 - \lambda) \|x - y\|^2 / 2$$

Une fonction fortement convexe est strictement convexe mais la réciproque n'est pas vraie.

- Enveloppe convexe :

Soit P une partie de U . L'intersection de convexes étant convexe, on peut parler du plus petit convexe contenant P , qui est donc l'intersection de tous les convexes contenant P . C'est ce que l'on appelle l'enveloppe convexe de P .

1.6 Fonctions convexes différentiables (on a deux cas)[10][14]

1.6.1 Rappels : (Caractérisation des fonctions convexes dérivables)[9]

Soit K un convexe dans Ω ouvert de U . Soit $f : \Omega \rightarrow R$ dérivable sur $\Omega \subset R$.

1. f est convexe sur K si et seulement si : $f(x) \geq f(x_0) + f'(x_0)(x - x_0), \forall x, x_0 \in K$.
2. f est strictement convexe sur K si et seulement si : $f(x) > f(x_0) + f'(x_0)(x - x_0), \forall x, x_0 \in$

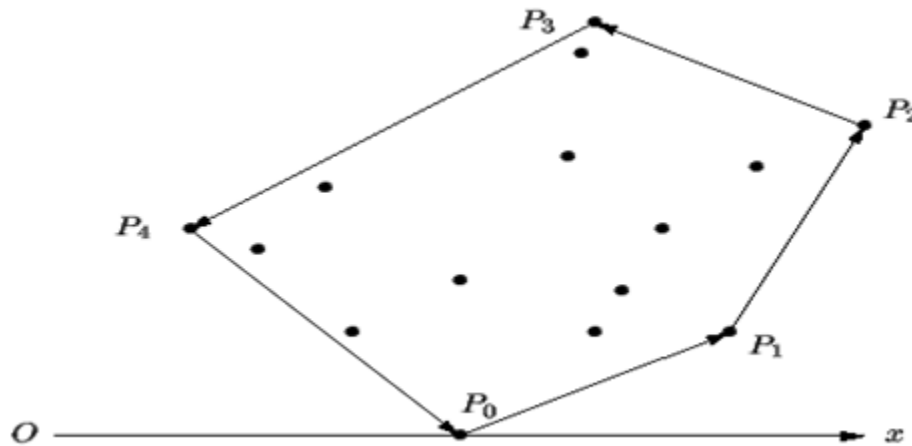


FIGURE 1.5 – La représentation d'une enveloppe convexe.

$K, x \neq x_0$.

3. On suppose que f est deux fois dérivable dans Ω . Alors f est convexe sur K si et seulement si : $f''(x) \geq 0; \forall x, \in K$.

4. On suppose que f est deux fois dérivable dans Ω . Si $f'' > 0 \forall x \in K$, Alors f est strictement convexe sur K .

Définition1 :

Convexité et Gradient : f fonction différentiable de $U \subset R^n$ dans R , U ensemble convexe ouvert :

- f convexe $\iff \forall x, y \in U, f(y) - f(x) \geq (y - x)^T g(x)$.
- f strictement convexe $\iff \forall x, y \in U, f(y) - f(x) > (y - x)^T g(x)$.

Interprétation géométrique : tangente au-dessous de la courbe.

Définition2 :

Convexité et Hessien : f fonction deux fois différentiable de $U \subset R^n$ dans R , U ensemble convexe ouvert :

- f convexe $\iff \forall x \in R^n, H(x)$ semi-définie positive.
- f strictement convexe $\iff \forall x, \in R^n, H(x)$ définie positive.

Définition3 :

► Définition de la matrice Hessienne :

En général, une fonction de deux variables admet quant à elle deux dérivées partielles, l'une par rapport à la première variable et l'autre par rapport à la seconde. Il n'existe donc pas un nombre dérivé comme avec une fonction d'une seule variable mais un couple de variables, c'est-à-dire un vecteur que l'on nomme gradient.

Une fonction de deux variables admet elle aussi des dérivées secondes, mais cette fois-ci il y en a quatre :

Soit une fonction $f(x, y)$ défini sur un ouvert :

Supposons que f peut être dérivée deux fois par x , deux fois par y , par x puis par y et enfin par y puis par x .

La matrice hessienne est la matrice $n \times n$ de ces quatre dérivées partielles en un point. Elle est symétrique puisque nous venons de voir que les deux dérivées croisées sont identiques.

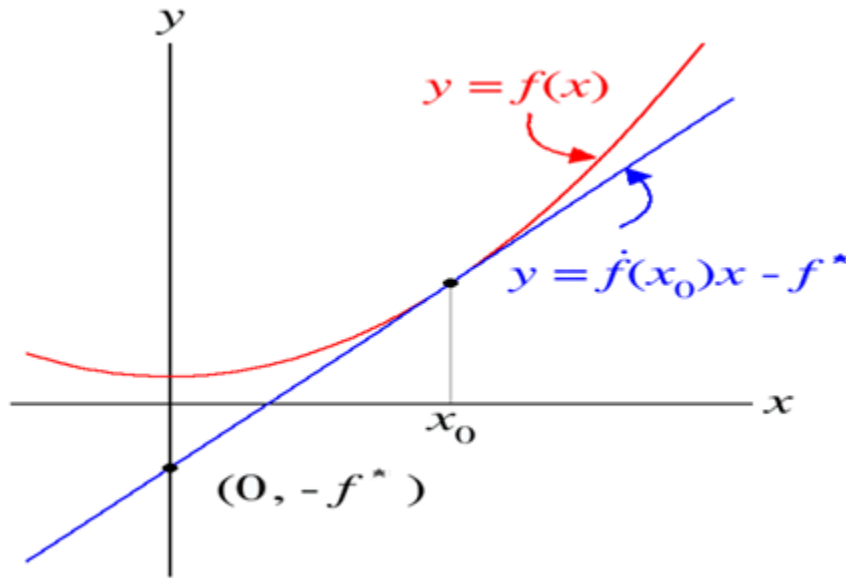


FIGURE 1.6 – La représentation d'une fonction différentiable convexe

Si $f : R^n \rightarrow R$ est de classe C^2 alors on définit la matrice Hessienne de f par :

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

1.7 Fonctions quasi convexe :

Une fonction f définie sur un ensemble convexe U dans R^n est dite quasi convexe si : pour chaque $x, y \in U$ on a : $f(x) \leq f(y)$. On ait :

$$\forall \lambda \in [0, 1], f(x) \leq f(y) \Rightarrow f[\lambda x + (1 - \lambda)y] \leq f(y)$$

Ceci nous amène à la notion moins restrictive de quasi-convexité.

Définition :[14]

La fonction $f : R^n \rightarrow R$ est quasi convexe dans un convexe $U \subset R^n$:

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}$$

1.7.1 Les problèmes d'optimisation non linéaire :[23]

Définition :

La dichotomie “linéaire/non linéaire” est assez classique en Mathématiques même si elle

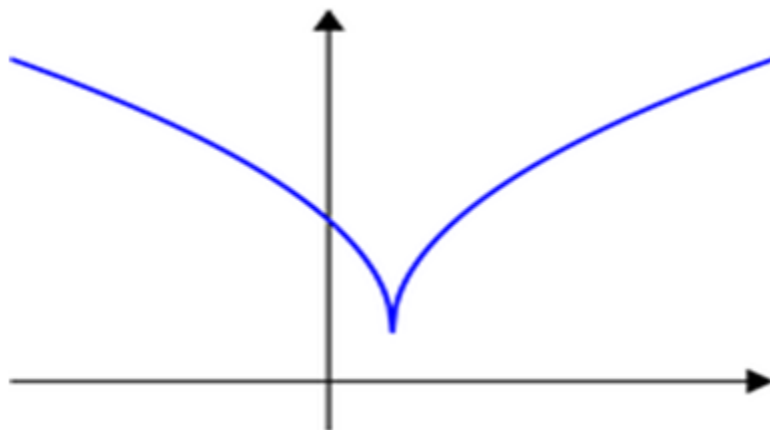


FIGURE 1.7 – La représentation d’une fonction quasi-convexe

représente une partition faussement symétrique (dans le monde du “non linéaire” il y a beaucoup de sous-catégories). En optimisation, on parle, du côté linéaire, de programmation linéaire. La programmation linéaire, qui pourrait n’être considérée que comme un cas particulier de la programmation convexe, elle-même sous-catégorie de la programmation non linéaire, est bien plus que cela en fait. Historiquement, elle a occupé le devant de la scène en optimisation dans les années soixante sous l’impulsion de George B. Dantzig[14], et on peut dire que dans beaucoup de cercles, “optimisation” a été pour un temps synonyme de “programmation linéaire”. Techniquement, bien que cas particulier de la programmation convexe, le contexte de la programmation linéaire permet des développements qui lui sont spécifiques. Ce fut en particulier le cas sur le plan algorithmique avec le fameux algorithme du simplexe[7] qui exploite un certain aspect combinatoire de la programmation linéaire même en variables continues. Les développements algorithmiques plus récents par les méthodes de points intérieurs[14][15] ont sérieusement réduit l’écart entre programmation linéaire et non linéaire. Toujours est-il que la programmation linéaire, par ses spécificités, constitue à elle seule un objet de cours. Elle ne sera pas abordée explicitement ici, même si les techniques de dualité qui y seront développées s’appliquent aussi bien à son cas[5].

1.7.2 Formulation générale des problèmes d’optimisation non linéaire :

La forme générale d’un problème d’optimisation est la suivante :

$$(PC) \begin{cases} \min f(x) \\ sc \\ g(x) \leq 0 \\ h(x) = 0 \end{cases}$$

Avec g désigne ce que nous appelleront des contraintes d’inégalité et h des contraintes d’égalité.

l’objectif est la présentation de techniques permettant de résoudre le problème (PC) , ainsi

que des problèmes où soit un seul des deux types de contraintes est présent, soit des problèmes où il n'y a pas de contraintes du tout. Nous noterons ces types de problèmes ainsi :

- a. *(PM)* problème général, avec contraintes d'inégalité et d'égalité (mixtes),
- b. *(PCE)* problème avec contraintes d'égalité,
- c. *(PCI)* problème avec contraintes d'inégalité,
- d. *(P)* problème sans contraintes.

Il va de soi que la plupart des problèmes réels ou industriels ne sont pas initialement sous une des formes proposées. C'est pourquoi un des premiers travaux consiste en général à mettre le problème initial sous une forme *(PC)*.

1.7.3 les problèmes d'optimisation sans contraintes :[6][14]

Définitions dans le cas général :

On veut résoudre $\min_{x \in R^n} f(x)$ ou $\max_{x \in R^n} f(x)$ i.e. on cherche v valeur optimale et x_0 tel que $f(x_0) = v$.

Soit $x \in V(x_0)$.

- 1. On dit que x_0 est un minimum local si : $\exists V(x_0)$ tel que $\forall x \in V(x_0), f(x) \geq f(x_0)$,
- 2. On dit que x_0 est un minimum local strict si : $\exists V(x_0)$ tel que $\forall x \in V(x_0), f(x) > f(x_0)$,
- 3. On dit que x_0 est un minimum global si : $\forall x \in R^n, f(x) \geq f(x_0)$,
- 4. On dit que x_0 est un minimum global strict si : $\forall x \in R^n, f(x) > f(x_0)$,

On définit de la même façon un maximum local, maximum local strict, maximum global, maximum global strict, en renversant les inégalités.

Le problème que l'on étudie ici est celui de la recherche du minimum (ou maximum) d'une fonction f de n variables x_1, x_2, \dots, x_n avec $f \in C^2$ non continue.

Soit $f : R^n \rightarrow R$ qui à tout $x \in R^n, (x_1, x_2, \dots, x_n)^T$, associe la valeur réelle :

$$f(x) = (x_1, x_2, \dots, x_n)$$

On cherche à résoudre :

$$\begin{cases} \min f(x) \\ x \in R^n \end{cases}$$

Il s'agit donc de déterminer un point x^* de R^n tel que :

$$\forall x \in R^n : f(x^*) \leq f(x) \quad (I)$$

C'est-à-dire un minimum global de f sur R^n .

Pour beaucoup de problèmes d'optimisation sans contrainte, en l'absence de propriétés particulières (telles que la convexité et la quasi-convexité), les principales méthodes de résolution connues ne permettent pas la détermination d'un minimum global : il faut alors se contenter d'optimums locaux, c'est-à-dire de points qui vérifient (I) seulement dans un voisinage de x^* .

Définition1 :

► Rappelons que la condition de semi-définie positivité s'exprime par :

$$\forall y \in R^n, y^T \nabla^2 f(x^*) y \geq 0$$

► Et la condition de définie positivité s'exprime par :

$$\forall y \in R^n, y \neq 0, y^T \nabla^2 f(x^*) y > 0$$

► Sous-gradients :

Définition2 :

On appelle sous-gradient de f au point x^* tout vecteur $y = (y_1, y_2, \dots, y_n)^T \in R^n$ vérifiant : $f(x) \geq f(x^*) + y^T(x - x^*), \forall x \in R^n$.

On distingue deux cas :

1. Pour une fonction convexe f différentiable, le gradient $\nabla f(x^*)$ en tout point x^* vérifie l'inégalité fondamentale : $f(x) \geq f(x^*) + \nabla f^T(x^*)(x - x^*)$.
2. Et pour une fonction non partout différentiable, la notion de sous-gradient généralise celle de gradient.

► Direction de descente :

Définition3 :

- Gradient de f en $x \in R^n$: $g(x) = \nabla f(x)$,
- Dérivées directionnelle de f en x suivant $d \in R^n$: $f_d(x) = g(x)^T d$, avec d est une direction de descente en x si : $f_d(x) = g(x)^T d < 0$.
- La direction de la plus forte montée d^+ est la direction du gradient : $d^+ = g(x)$.
- La direction de la plus forte descente d^- est opposée au gradient : $d^- = -g(x)$.

$$\forall d \in R^n \parallel d \parallel = \parallel d^- \parallel, g(x)^T d^+ \geq g(x)^T d^- = - \parallel g(x) \parallel^2$$

1.8 Condition nécessaires d'optimalité locale : [14]

On suppose ici que f est continu et a des dérivées partielles premières $\partial f / \partial x_i$ et secondes $\partial^2 f / \partial x_i \partial x_j$ continues pour tout $x \in R^n$, alors :

Théorème1 :

Une condition nécessaire pour que x^* soit un minimum (local ou global) de f est :

- a). $\nabla f(x^*) = 0$ (stationnarité).
- b). le hessien $\nabla^2 f(x^*) = [\partial^2 f / \partial x_i \partial x_j(x^*)]$ est une matrice semi-définie positive.

Démonstration :

Soit x^* un minimum local (ou global) de f , comme f est deux fois continûment différentiable, le développement de Taylor à l'ordre 2 au voisinage de x^* donne :

$$f(x) = f(x^*) + \nabla f^T(x^*)(x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*) + \parallel x - x^* \parallel^2 \varepsilon(x - x^*)$$

Avec, $\varepsilon(x - x^*) \longrightarrow 0$ quand $x \longrightarrow x^*$.

- Si $\nabla f(x^*) \neq 0$ alors en choisissant $x = x^* - \theta \nabla f(x^*)$ on aurait, pour $\theta > 0$ suffisamment petit : $f(x) < f(x^*)$ ce qui contredirait le fait que x^* est un minimum local. Donc la condition a) est bien nécessaire, et l'on peut écrire :

$$f(x) = f(x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*) + \|x - x^*\|^2 \varepsilon(x - x^*)$$

- Si la matrice $\nabla^2 f(x^*)$ n'est pas semi-définie positive, c'est qu'il existe un vecteur $d \in R^n$ ($d \neq 0$) tel que :

$$d^T \nabla^2 f(x^*) d < 0$$

En choisissant alors $x = x^* + \theta d$, pour $\theta > 0$ suffisamment petit on aurait $f(x) < f(x^*)$ ce qui contredirait encore l'optimalité locale de x^* .

La condition b) est donc bien également nécessaire.

Un point x^* qui vérifie la condition a) c'est-à-dire : $\partial f / \partial x_i(x^*) = 0$; ($i = 1, \dots, n$) est appelé un point stationnaire.

La figure ci-dessous illustre le fait que les conditions a) et b) du théorème 1 ne sont pas suffisantes pour garantir un minimum local ou global. A fortiori la stationnarité seule, bien que nécessaire, n'est pas une condition suffisante d'optimalité locale.

Exemple : $f(x) = x^3$, avec : $f'(0) = 0$.

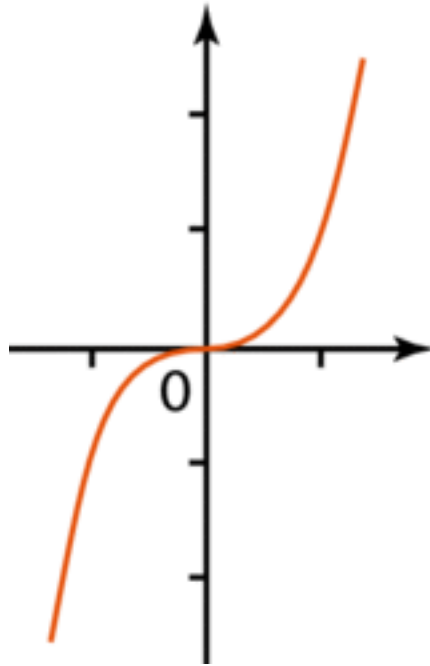


FIGURE 1.8 – Un point d'inflexion, mais ce n'est pas un minimum local.

1.9 Condition suffisantes d'optimalité locale : [6]

Théorème 2 :

Sous les mêmes hypothèses qu'au (1.8), une condition suffisante pour que x^* soit un optimum local de f sur R^n est :

a). $\nabla f(x^*) = 0$ (stationnarité).

b). le hessien $\nabla^2 f(x^*) = [\partial^2 f / \partial x_i \partial x_j (x^*)]$ est une matrice définie positive.

Démonstration :

Considérons un point x^* , satisfaisant les deux conditions a) et b) du théorème.

le développement de Taylor à l'ordre 2 au voisinage de x^* s'écrit alors :

$$f(x) = f(x^*) + \nabla f^T(x^*)(x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*) + \|x - x^*\|^2 \varepsilon(x - x^*)$$

Avec, $\varepsilon(x - x^*) \rightarrow 0$ quand $x \rightarrow x^*$.

Pour toute direction de déplacement $d \in R^n$, ($\|d\| = 1$) on a alors :

$$f(x^* + \theta d) = f(x^*) + \theta^2 \frac{1}{2} d^T \nabla^2 f(x^*) d + \theta^2 \varepsilon(\theta)$$

avec, $\varepsilon(\theta) \rightarrow 0$ quand $\theta \rightarrow 0$.

En vertu de la condition b) on a : $d^T \nabla^2 f(x^*) d > 0$ et par suite, pour θ suffisamment petit, on aura : $f(x^* + \theta d) > f(x^*)$.

Ce qui montre que x^* est bien un minimum local de f . Remarquons que la condition b) du théorème 2 revient à supposer que f est strictement convexe dans un voisinage de x^* .

La figure précédente montre un cas où les conditions suffisantes d'optimalités ne sont pas satisfaites puisque, f admettant un point d'inflexion, le hessien n'est pas défini positif mais seulement semi-défini positif.

1.10 Cas des fonctions convexes : (Condition nécessaire et suffisante d'optimalité globale)

Dans le cas d'une fonction convexe propre f définie sur R^n . Une condition nécessaire et suffisante pour que x^* soit un minimum global de f est que 0 soit un sous-gradient de f en x^* . Pour une fonction continûment différentiable, on obtient donc :

Théorème 3 : [14]

Si f est une fonction convexe continûment différentiable, une condition nécessaire et suffisante pour que x^* soit un optimum global de f sur R^n est que : $\nabla f(x^*) = 0$ (stationnarité). Autrement dit, dans le cas convexe, la stationnarité à elle seule constitue une condition nécessaire et suffisante d'optimalité globale.

1.11 Cas des fonctions quelconques (difficulté du problème général) :

Les conditions du 1.8), 1.9) sont pratiquement les conditions d'optimalité les plus générales que l'on connaisse pour des fonctions continûment différentiables sur R^n .

Lorsque la fonction à minimiser est convexe, on arrive bien à se passer de la différentiabilité en tout point en utilisant la notion de sous-gradient.

1.11.1 Les problèmes d'optimisation avec contraintes d'égalité, d'inégalité et mixtes :[5][6]

Remarque3 :

Les conditions de Lagrange ont été étudiées par Lagrange dans le cas de fonctions différentiables et pour des contraintes d'égalité uniquement. Les généralisations aux autres situations sont partiellement dues aux autres auteurs tels que Karush-Kuhn et Tucker. En fait, ces conditions constituent des conditions nécessaires dans le cas non convexe (sous des hypothèses de qualification des contraintes qui prennent alors des formes diverses), et elles deviennent des conditions suffisantes dans le cas convexe. En tout cas, elles peuvent être considérées comme des conditions “locales” en ce sens qu'elles font intervenir des notions de différentiabilité ou de sous-différentiabilité. Évidemment, dans le cas convexe, on a vu que l'on passe assez facilement des notions locales aux notions globales. Cependant, cette partie suivante abordera un autre point de vue spécifiquement global (c'est-à-dire où on ne parle pas de dérivées ou de sous-gradients mais seulement d'inégalités) sur les multiplicateurs et la dualité, Signalons l'appellation “variables duales” pour ce que nous avons appelé “multiplicateurs” (de Lagrange ou de Kuhn et Tucker).

1.11.2 Les problèmes d'optimisation avec contraintes d'égalité :

Définition1 :[9]

► Définition de La qualification des contraintes :

La qualification des contraintes sera réalisée en tout point $x \in U$ si les fonctions $g_i(x)$ sont soit linéaires, soit non linéaires convexe, et s'il existe $\bar{x} \in U$ vérifiant $g_i(\bar{x}) < 0$ (pour tout g_i non linéaire convexe).

► Régularité : x^* est régulier pour h si :

1. x^* est admissible.
2. les $\nabla h_j(x^*)$ sont linéairement indépendants.

► Définition de La matrice Jacobienne :

La différentielle de f en x est une application linéaire, elle peut être représentée par sa matrice dans la base canonique. On appelle cette matrice la Jacobienne de f en x .

Soit $f : R^n \longrightarrow R^p$ telle que :

$$f(x_1, x_2, \dots, x_n) = \{f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_p(x_1, x_2, \dots, x_n)\}$$

La matrice Jacobienne de f en x s'écrit alors :

Avec $1 \leq i \leq p$ et $1 \leq j \leq n$

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(x) & \frac{\partial f_p}{\partial x_2}(x) & \dots & \frac{\partial f_p}{\partial x_n}(x) \end{pmatrix}$$

Remarque

Le déterminant et le rang de la Jacobienne : on les définit par $\det J_f$ et $rg J_f$.

Si $m = n$ (c'est-à-dire le nombre de lignes égale aux nombre de colonnes) alors $J_f(x)$ est une matrice carré.

Si $\det J_f \neq 0$ donc le, $rg J_f = m$ (où le nombre de vecteurs qui sont linéairement indépendant).

Si $m \neq n$, la matrice n'est pas carrée, donc la Jacobienne est dite de rang plein c'est-à-dire : $rg J_f = m \leq n$.

Définition 2 :

Un problème d'optimisation avec contraintes d'égalité est défini par :

$$(PCE) \begin{cases} \text{minimiser } f(x) \\ \text{sc} \\ h_j(x) = 0, j = 1, \dots, p \\ x \in R^n \end{cases}$$

Conditions d'optimalité avec contraintes d'égalités (multiplicateurs de Lagrange) : [11][15]

Définition 3 :

Considérons le problème de programmation mathématique suivant :

$$(P) \begin{cases} \text{minimiser } f(x) \\ \text{sc} \\ h_j(x) = 0, j = 1, \dots, p \\ x \in R^n \end{cases} \quad (II)$$

Le lagrangien associé au problème (II) est obtenu comme suit, en associant un multiplicateur de Lagrange noté γ_j à chaque fonction de contrainte h_j :

$$L(\gamma, x) = f(x) + \sum_{j=1}^p \gamma_j h_j(x).$$

Remarque :

Théorème de Lagrange (voir Minoux) [14]

Exemple numérique :

On a, $\min h(x, y) = x^2 + 2y^2$.

Sc : $h_1(x, y) = x + y - b = 0$.

$L(\gamma, x, y) = (x^2 + 2y^2) + \gamma(x + y - b)$.

L convexe en (x) . Donc le minimum atteint lorsque $\nabla_{x,y} L(\gamma, x, y) = 0$, d'où :

$$\nabla_{x,y} L(\gamma, x, y) = (2x + \gamma, 4y + \gamma) = (0, 0) \Leftrightarrow (\gamma = -2x, \gamma = -4y) \Leftrightarrow x = 2y.$$

$$x + y - b = 2y + y - b = 3y - b = 0 \Leftrightarrow y = b/3 \Rightarrow x = 2b/3.$$

1.11.3 Les problèmes d'optimisation avec contraintes d'inégalité :

Définition 1 :

► Contraintes actives et contraintes inactives (cas générale) :

Soient $g : R^n \rightarrow R$ et $h : R^n \rightarrow R$.

Une contrainte d'inégalité $g(x) \leq 0$ est dite active en x^* si : $g(x^*) = 0$, et inactive en x^* si : $g(x^*) < 0$.

L'ensemble des indices des contraintes actives en x^* sera généralement noté $A(x^*)$.

Définition 2 :

Un problème d'optimisation avec contraintes d'inégalité est défini par :

$$(PCI) \begin{cases} \text{minimiser } f(x) \\ \text{sc} \\ g_i(x) \leq 0, i = 1, \dots, m \\ x \in R^n \end{cases}$$

Remarque1 :

On notera S l'ensemble des solutions de (PCI) c'est-à-dire :

$$S = \{x \in R^n / g_i(x) \leq 0, \forall i = 1, \dots, m\}$$

Toutes fonctions f et $g_i(x)$ sont supposées continues et différentiables.

Remarque2 :

on dit que la qualification des contraintes (QC) est vérifiée en tout point $x \in S$, si l'une des conditions **a)** ou **b)** soit réalisée :

a) Toutes les fonctions g_i sont linéaire ou affines (Karlin 1959).

b) Toutes les fonctions g_i sont convexes différentiables et il existe $\bar{x} \in R^n$ vérifiant $(g_i(\bar{x}) < 0, i = 1 \dots m)$ (Slater 1950).

Aussi, pour (QC) soit vérifiée en un point $x^* \in S$, il suffit que l'on ait :

c) les gradients $\nabla g_i(x^*)$ des contraintes saturées en x^* sont linéairement indépendants (Fiacco et McCormick 1968).

les conditions nécessaires et suffisantes de Karush-Kuhn-Tucker :[11]

Pour obtenir des conditions plus facilement vérifiables, il faut poser des hypothèses sur S et sur les fonctions $f, g_i(x)$.

Si S est convexe, si f et $g_i(x)$ sont différentiables et convexes dans le problème (PCI) suivant :

$$(PCI) \begin{cases} \text{minimiser } f(x) \\ \text{sc} \\ g_i(x) \leq 0, i = 1, \dots, m \\ x \in S \end{cases}$$

Alors le lagrangien $L(\gamma, x) = f(x) + \sum_{i=1}^m \gamma_i g_i(x)$ est aussi une fonction convexe sur S si $\gamma \geq 0$.
Puisque $\gamma \geq 0$ et $g_i(x)$ convexe $\Rightarrow \gamma_i g_i(x)$ convexe.

$f(x) + \sum_{i=1}^m \gamma_i g_i(x)$ Somme de fonctions convexes, d'où si f et $g_i(x)$ sont différentiables et convexes, alors le lagrangien est une fonction différentiable et convexe en x , et il s'ensuit qu'il possède un minimum global en x^* sur S lorsque $\gamma = \gamma^*$ si :

$$\nabla_x L(\gamma^*, x) = \nabla f(x) + \sum_{i=1}^m \gamma_i^* g_i(x) = 0$$

D'après le **théorème** de *KKT*[14].

A. Conditions suffisantes de *KKT* :

Le théorème précédant (*KKT*[14]), nous démontrer que les conditions de *KKT* sont suffisantes sous les hypothèses supplémentaires que S est convexe et que f et $g_i(x)$ sont convexes sur S .

Nous pouvons démontrer ce résultat en utilisant plutôt l'inégalité du gradient, on obtient le théorème1 suivant :

Théorème1 :

Supposons que S est convexe et que les fonctions f et $g_i(x)$ sont différentiables et convexes. Si les conditions de *KKT* sont vérifiées à x^* , alors x^* est un minimum global du problème (*PCI*).

Preuve : Le lagrangien étant convexe si $\lambda^* \geq 0$ (démontrer précédemment[14]), alors par l'inégalité du gradient il s'ensuit que pour tout $x \in S$:

$$f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) \geq f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) + [\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x)]^T (x - x^*)$$

Avec $\sum_{i=1}^m \lambda_i^* g_i(x^*) = 0$.

Et $[\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x)]^T = 0$.

Alors, pour tout $x \in S$: $f(x^*) - f(x) \leq \sum_{i=1}^m \lambda_i^* g_i(x^*) \leq 0$, et $f(x^*) \leq f(x)$.

B. Conditions nécessaires de *KKT* :

Le théorème suivant est fondamental et donne, sous l'hypothèse de (*QC*), une condition nécessaire d'optimalité locale pour un problème d'optimisation avec contraintes d'inégalité.

Théorème2 : (Karush 1939, Kuhn et Tucker 1951).[14]

Interprétation géométrique des conditions de *KKT* :

Les conditions de *KKT* peuvent s'interprétés géométriquement (figure9). Sous l'hypothèse de (*QC*), l'ensemble des directions admissible y forme un cône convexe fermé C_{ad} intersection des $|F|$ demi-espaces d'équation : $\nabla g_i^T(x^*) \cdot d \leq 0 (\forall i \in F)$.

Remarque2 : (C_{ad} est encore appelé cône tangent en x^* à S).

1.11.4 Les problèmes d'optimisation avec contraintes mixtes :

La définition d'un problème d'optimisation avec contraintes mixtes est :

$$(PM) \begin{cases} \min f(x) \\ sc \\ g_i(x) \leq 0, i = 1, \dots, m \\ h_j(x) = 0, j = 1, \dots, p \\ x \in R^n \end{cases}$$

Remarque : les contraintes de *KKT* s'étendent sans difficulté à des problèmes comportant des contraintes mixtes.

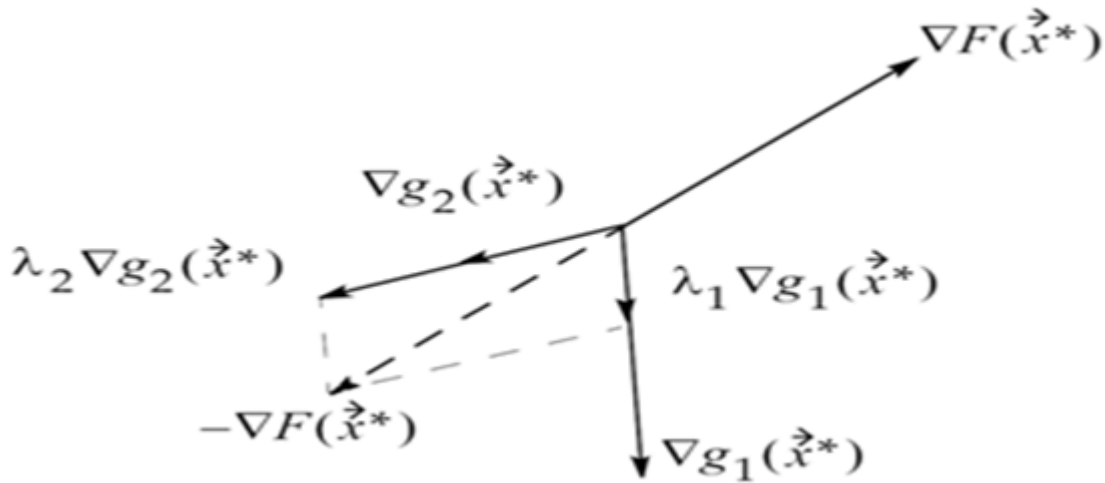


FIGURE 1.9 – Illustration des conditions de KKT sur un exemple à deux dimensions.

Théorème 3[14] : on suppose que les fonctions $f, g_i(x)$ et $h_j(x)$ sont continûment différentiable. Et que (QC) est vérifiée en x^* , solution de (PM) . Alors, une condition nécessaire pour que x^* soit un optimum local de (PM) est qu'il existe des nombres $(\lambda_i \geq 0)$ et $(\mu_j \text{ avec } : \mu_j \text{ non contraints au signe})$ tels que :

$$(P) \begin{cases} \nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla g_i(x^*) + \sum_{j=1}^p \mu_j \nabla h_j(x^*) = 0 \\ \text{et} \\ \lambda_i g_i(x^*) = 0, (\forall i = 1, \dots, m) \end{cases}$$

1.12 La dualité de Lagrange :[21]

Introduction :

forme générale d'un problème d'optimisation (P) est défini par :

$$(PM) \begin{cases} \text{minimiser } f(x) \\ \text{sc} \\ g_i(x) \leq 0, i = 1, \dots, m \\ h_j(x) = 0, j = 1, \dots, p \\ x \in R^n \end{cases}$$

Avec C l'ensemble des contraintes ou l'ensemble des points admissibles ou réalisables tel que :

$$C = \{x \in R^n / h_j(x) = 0, j = 1, \dots, p \text{ et } g_i(x) \leq 0, i = 1, \dots, m\}.$$

Remarque (cas convexe) : si le problème (PM) est convexe, les conditions de KKT sont nécessaires et suffisantes en un point x^* régulier.

1.13 La Dualité :[14]

- Lagrangien : le Lagrangien du problème (PM) est la fonction :

$$L(\gamma, \mu, x) = f(x) + \sum_{i=1}^m \gamma_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x)$$

Quand f , h et g sont différentiables, les conditions de KKT s'expriment par :

$$\nabla_x L(\gamma^*, \mu^*, x^*) = 0$$

Les γ_i et les μ_j sont les multiplicateurs de Lagrange (ou les variables duales) associés aux contraintes.

- On définit la fonction duale de Lagrange par : $D(\gamma, \mu) = \inf_x L(\gamma, \mu, x)$.

Avec D est toujours concave (La concavité vient de la définition en tant qu'infimum de fonctions affines).

Pour $\mu \geq 0$, $D(\gamma, \mu) \leq \inf_{x \in C} f(x)$. D'où le problème dual (W) associé au problème primal (PM) est :

$$(W) \begin{cases} \text{maximiser}_{\gamma, \mu} D(\gamma, \mu) \\ \text{sc} \\ \mu_j \geq 0, \forall j = 1, \dots, p \end{cases}$$

- On définit le saut de dualité par : $\inf_{x \in C} f(x) - \max_{\mu \geq 0} D(\gamma, \mu)$

Remarque : on appelle saut de dualité le fait que la valeur $\inf_{x \in C} f(x) - \max_{\mu \geq 0} D(\gamma, \mu)$ soit non nul. C'est le cas lorsqu'un PM ne possède pas de point-selle pour sa fonction duale associée. C'est le cas par exemple de la plupart des PMD .

- La définition d'un point selle :

Symétrisation du problème : (PM) est équivalent à $(\inf_x \sup_{\gamma, \mu \geq 0} L(\gamma, \mu, x))$.

Le problème dual (W) est : $(\sup_{\gamma, \mu \geq 0} \inf_x L(\gamma, \mu, x))$.

D'où le point selle est minimal par rapport à une variable, maximal par rapport à l'autre, (γ^*, μ^*, x^*) est un point selle du Lagrangien si pour tout $(\gamma, \mu, x) (\mu^* \geq 0, \mu \geq 0)$

$$L(\gamma, \mu, x^*) \leq L(\gamma^*, \mu^*, x^*) \leq L(\gamma^*, \mu^*, x)$$

Propriété1. (La dualité faible)

Théorème1 : (Théorème faible de la dualité, Wolfe 1961)[14][22]

Soit $F(x)$ et $L(k, y)$ les fonctions objectifs respectivement du problème primal et son dual.

On a alors :

$$L(k, y) \leq F(x), \forall (k, y) \in V, \forall x \in S.$$

Théorème2 : (Théorème fort de la dualité, Wolfe 1961)[14]

Soit x^* une solution optimale du problème primal (PM). Alors il existe $y^* \in R^m, y^* \geq 0$, tel que le couple (x^*, y^*) et une solution optimale du dual (PMD) et on a :

$$F(x^*) = L(x^*, y^*).$$

Remarque : Pour le premier théorème de dualité faible, nous n'avons pas besoin d'hypothèses particulières sur S ni sur les fonctions f et $g_i, (i = 1, \dots, m)$.

Propriété 2. (Théorème de dualité) [14] : (γ^*, μ^*, x^*) est un point selle avec $\mu^* \geq 0$ si et seulement si x^* est une solution de (PM) , (γ^*, μ^*) est une solution de (W) et le saut de dualité est nul, pour résoudre le problème, on peut donc chercher un point selle du Lagrangien.

Remarque : un point selle du Lagrangien vérifie les conditions de KKT (sans hypothèse autre que f, h et g différentiable). [12]

Si le problème est convexe : point selle $\Leftrightarrow KKT$.

a. Si le problème (PM) admet un point selle $:(\gamma^*, \mu^*, x^*)$ alors, on a :

$$Max(W) = D(\gamma^*, \mu^*) = f(x^*) = Min(PM).$$

Autrement dit, la valeur optimale du problème (PM) est égale à la valeur optimale du problème dual (W) .

b. Réciproquement, s'il existe x^* solution de (PM) et $\mu^* \geq 0$ tels que : $D(\mu^*) = f(x^*)$ alors (PM) admet un point selle.

1.14 la programmation quadratique : [19]

1.14.1 Définition :

a) Historiquement, Barankin et Dorfman furent les premiers à remarquer qu'en combinant les conditions d'optimalité de Lagrange avec celles du système original, la solution optimale d'un problème quadratique était une solution de base d'un système élargi ayant la propriété que seuls certains couples de variables figuraient dans la base. De son côté Markowitz montra qu'il est possible de modifier le système élargi et d'engendrer paramétriquement une classe de solutions de base ayant la propriété particulière ci-dessus et convergeant vers l'optimum en un nombre fini d'itérations. Enfin, Wolf montra qu'en modifiant légèrement la méthode de simplexe d'une façon à ne pas autoriser l'introduction d'une variable dans la base si sa variable complémentaire s'y trouvait déjà, on parvenait aisément à l'optimum recherché.

b) Pour la grande plupart des problèmes dans ce sous-ensemble, il faut que les fonctions qui définissent les contraintes et l'objective soient convexes et h_j linéaires (affines).

1.14.2 Problème quadratique avec contraintes d'égalités : [14][19]

Définition :

Un cas important où les conditions sont suffisantes :

- Q quadratique convexe.

$$(PQE) \left\{ \begin{array}{l} \min Q(x) = \frac{1}{2}x^T Cx + P^T x \\ \text{sc} \\ Ax = b \\ x \geq 0 \end{array} \right.$$

Avec C symétrique semi-défini positive.

Rappel :

- Matrice C carrée $n \times n$,
- C est symétrique : $C = C^T$, ie : $C_{ij} = C_{ji} \forall i, j$.
- $x^T C x \geq 0, \forall x$.

C symétrique semi-définie positive $\implies Q$ convexe.

Q convexe \implies optimum local = optimum global.

Conditions de Kuhn et Tucker pour le problème quadratique (PQE) :

$$(KT) \begin{cases} Ax = b \\ Cx - v + A^T u = -p \\ x \geq 0, v \geq 0 \\ x^T v = 0 \end{cases}$$

Remarque :

à part $x^T v = 0$, les autres équations sont linéaires.

Définition de la méthode de Wolfe : [14]

La méthode de Wolfe s'applique lorsque le point de coordonnées nulles ne vérifie pas les contraintes. Dans ce cas, on ne peut pas appliquer la méthode de Dantzig car celle-ci prend l'origine comme point de départ. Il faut donc trouver un point de départ qui appartienne au domaine réalisable. La méthode de Wolfe est très similaire à la méthode des valeurs ajoutées qui traite le même problème de point de départ dans le cas de la méthode du simplexe. Deux formes pour la méthode de Wolfe ; la courte et la longue.

Forme courte : Suppose que $p = 0$ ou C définie positive,

Forme longue : Pas de contrainte sur p ou C .

l'étude de la forme courte, revient à l'appliquer deux fois dans notre cas.

$$(PQE) \begin{cases} \min Q(x) = \frac{1}{2} x^T C x + P^T x \\ sc \\ Ax = b \\ x \geq 0 \end{cases} \longrightarrow (KT) \begin{cases} Ax = b \\ Cx - v + A^T u = -p \\ x \geq 0, v \geq 0 \\ x^T v = 0 \end{cases}$$

Idée de la méthode de Wolfe : [13][14]

Ajout de variables artificielles \implies solution réalisable des conditions de Kuhn et Tucker ;

Suppression de ces variables via l'algorithme du simplexe \implies similaire à une Phase1 du simplexe excepté qu'on rajoute une règle pour assurer que $x^T v = 0$.

Kuhn et Tucker vers une solution réalisable : (introduire des variables artificielles)

$$w^T = (w_1, \dots, w_m), z^{1T} = (z_1^1, \dots, z_n^1), z^{2T} = (z_1^2, \dots, z_n^2),$$

\implies nouveau système élargi :

$$(KT) \begin{cases} Ax + w = b \\ Cx - v + A^T u + z^1 - z^2 = -p \\ x \geq 0, v \geq 0 \\ z^1 \geq 0, z^2 \geq 0, w \geq 0 \\ x^T v = 0. \end{cases}$$

Solution évidente du système :

$$(KT) \begin{cases} x = 0, v = 0, u = 0 \\ w = b \\ z_i^1 = -p_i, \text{ si } p_i \text{ négatif} \\ z_i^2 = p_i, \text{ si } p_i \text{ positif.} \end{cases}$$

La base réalisable : les w_i , les z_i^1 ou z_i^2 non nuls.

La méthode de Wolfe consiste à éliminer des z_i^j et des w_i de la base.

1.14.3 Problème quadratique avec contraintes d'inégalités :[19]

Nous introduisons les différentes approches développées pour la résolution des problèmes de programmation quadratique convexe avec contraintes d'inégalités.

Dans la littérature, plusieurs méthodes ont été développées pour la réalisation des problèmes de programmation quadratique. Elles sont généralement des extensions de celles développées pour la programmation linéaire (PL).

Remarque :

On considère, pour simplifier, le problème avec contraintes d'inégalités seulement, celles pour lesquelles les algorithmes de points intérieurs ont été conçus. On l'écrit sous la forme :

$$(PQI) \begin{cases} \min Q(x) \\ sc \\ Ax \leq b \\ x \geq 0 \end{cases}$$

Technique de la méthode de points intérieurs :[13][14]

En introduisant le vecteur des variables d'écart $s = b - Ax$, on peut réécrire les conditions d'optimalité qui sont à la fois nécessaires et suffisantes pour l'optimalité de couple $(x, \lambda) = (x^*, \lambda^*)$ comme suit :

$$\begin{cases} Dx + B\lambda + c = 0 \\ Ax + s - b = 0 \\ s_i \lambda_i = 0, \forall_i \\ (\lambda, s) \geq 0 \end{cases} \quad (*)$$

Les méthodes de points intérieurs de type primal-dual calculent les solutions (x, λ, s) du système $(*)$ en appliquant une variante de la méthode de Newton aux trois premières équations du système, en modifiant les directions de recherche et le pas de telle sorte que l'inégalité de la quatrième condition du système $(*)$ restent satisfaites à chaque itération.

La méthode de Newton :[13]

Définition :

La méthode de Newton consiste à approcher une solution x^* de l'équation $f(x) = 0$ pour laquelle $Df(x^*)$ est inversible en la regardant comme solution de $f(x^*) = x^*$ avec

$$F(x) = x - D^{-1}f(x) \circ f(x)$$

Idée de la méthode de Newton :

Soit $f : R^n \rightarrow R$ une fonction deux fois différentiable. Le modèle quadratique de f en x^* est une fonction $m_{x^*}(x) : R^n \rightarrow R$ définie par :

$$m_{x^*}(x) = f(x^*) + (x - x^*)^T \nabla f(x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*)$$

Où $\nabla f(x^*)$ est le gradient de f en x^* et $\nabla^2 f(x^*)$ est la matrice hessienne de f en x^* . En posant $d = x - x^*$, on obtient la formulation équivalente :

$$m_{x^*}(x^* + d)(x) = f(x^*) + d^T \nabla f(x^*) + \frac{1}{2}d^T \nabla^2 f(x^*)d$$

$$\min_x m_{x^*}(x) = f(x^*) + (x - x^*)^T \nabla f(x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^*)(x - x^*)$$

- Condition suffisante d'optimalité (premier ordre) :

$$m_{x^*}(x^* + d)(x) = f(x^*) + \nabla f(x^*) + \nabla^2 f(x^*)d = 0$$

C'est-à-dire :

$$d = -\nabla^2 f(x^*)^{-1} \nabla f(x^*)$$

Ou encore :

$$x = x^* - \nabla^2 f(x^*)^{-1} \nabla f(x^*)$$

- Condition suffisante d'optimalité (second ordre) :

$\nabla^2 f(x^*)$ définie positive.

La méthode de Newton est intéressante car sa convergence est quadratique au voisinage de la solution, c'est-à-dire lorsque la matrice hessienne de la fonction est définie positive en x_k , une itération de la méthode de Newton revient à minimiser le modèle quadratique de la fonction en x_k , et ainsi définit :

$$x_{k+1} = \operatorname{argmin}_{x \in R^n} m_{x_k}(x)$$

Algorithme :

Trouver une approximation de la solution du système $\nabla f(x) = 0$.

Initialisation

k=0.

Itérations

1. Construire le modèle quadratique :

$$m_{x^*}(x^* + d)(x) = f(x^*) + d^T \nabla f(x^*) + \frac{1}{2} d^T \nabla^2 f(x^*) d$$

2. Calculer : $d_{k+1} = \operatorname{argmin}_d m_{x^*}(x^* + d)$.

3. $x_{k+1} = x_k + d_{k+1}$.

4. $k = k + 1$.

Critère d'arrêt :

Si, $\|\nabla f(x_k)\| \leq \varepsilon$, Alors, $x^* = x_k$.

1.14.4 Dualité en programmation quadratique convexe : [14]

Le concept de dualité joue un rôle très important en programmation mathématique. Le but est de trouver une formulation alternative équivalente du problème de programmation mathématique, qui convient le plus au calcul ou qui a une signification théorique importante. Le problème original est appelé problème primal et le problème transformé est le problème dual. Souvent, les variables du dual peuvent être interprétées comme des multiplicateurs de Lagrange pour le cas linéaire et prennent la valeur γ^* comme solution duale, quand γ^* est le multiplicateur associé à la solution optimale primale x^* . Cependant, dans le cas non linéaire il existe toujours une fonction objective, souvent reliée à la fonction de Lagrange, qui doit être optimisée. Ici, on traitera de la dualité associée à un problème de programmation quadratique convexe comme problème primal. Il est important de remarquer que si le problème primal n'est pas convexe, alors le problème dual peut bien ne pas avoir de solution à partir de laquelle la solution primale peut être déduite. Donc ce n'est pas toujours possible d'appliquer la dualité comme technique générale dans le but de chercher une solution.

1.14.5 Conclusion :

Dans cette partie, on a abordé les problèmes d'optimisation convexe sous contraintes. Les contraintes ont d'abord été décrites implicitement ou géométriquement par la donnée d'un ensemble "admissible", puis par une collection finie de contraintes explicites scalaires sous forme d'égalités et d'inégalités à respecter.

Et nous avons présenté les résultats fondamentaux pour l'optimisation non linéaire, les différentes méthodes de minimisation numérique essentiellement les méthodes de gradient et la méthode de newton et enfin nous avons vu des notions sur la dualité.

Le modèle de Markowitz

2.1 Introduction [1]

Le terme "MARCHÉ" résume parfaitement la nature des marchés financiers dans leur perspective historique.

Depuis des centaines d'années, les hommes ont choisi de se retrouver en des lieux donnés et à des heures convenues pour échanger les produits les plus divers. Ces marchés furent d'abord basés sur le troc puis avec l'apparition des monnaies, ils ont rempli le rôle de centralisation de l'offre pour la demande.

Ce qui est vrai pour les produits de la vie courante, l'est aussi pour le secteur financier. Depuis le milieu du siècle dernier, la gestion de portefeuille a subi une profonde mutation. Il est loin, en effet, le temps où le gestionnaire pouvait se contenter d'appliquer quelques règles de bon sens et de bien connaître les sociétés cotées. Ce sont les travaux de Markowitz qui, au cours des années 1950, ont marqué le point de départ des développements théoriques modernes relatifs à la gestion des investissements en actifs financiers et au fonctionnement des marchés financiers.

A partir de là on peut entamer notre modèle tel que : **Le modèle de Markowitz.**

2.2 Les hypothèses du modèle de Markowitz [1][17]

Le modèle de Markowitz repose cependant sur des hypothèses fortes, relatives aux préférences des agents ou à la distribution de probabilité des rentabilités (Markowitz, 1952).

Markowitz lui-même avait déjà mentionné que cette mesure à savoir la variance n'était peut-être pas la meilleure en suggérant une alternative, la semi-variance, qui tient uniquement compte des rentabilités inférieures à la moyenne.

2.2.1 Les hypothèses relatives aux actifs financiers

H1 : «tout investissement est une décision prise dans une situation de risque; le return (rendement) d'un actif financier pour toute période future est par conséquent une variable aléatoire, dont on fait l'hypothèse qu'elle est distribuée selon une loi normale ».

C'est-à-dire une distribution symétrique stable entièrement définie par les deux paramètres :

$E(R_i) = \mu$: espérance mathématique du rendement.

$\sigma(R_i) = \sigma$: écart-type de la distribution de probabilité du rendement.

Tel que le (ROI : ou Return On Investment) c'est l'accroissement de la fortune initiale que l'investisseur cherche à Maximiser.

D'où :

$$r_t = (P_t - P_{(t-1)}) + C_t$$

r_t = rendement de l'actif financier pour la période (se terminant au temps) t .

P_t = prix de marche au temps t de l'actif financier.

C_t = revenu liquide attache à la détention de l'actif financier durant la période (se terminant au temps) t .

Remarque (Roi : en français Retour Sur Investissement).

H2 : La rentabilité des différents actifs financiers ne fluctuent pas indépendamment les uns des autres : ils sont corrélés ou, ce qui revient au même, ont des covariances non nulles. Chaque actif peut être acquis en quantité illimitée.

2.2.2 Les hypothèses relatives aux comportements des investisseurs

H3 : 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 return.

H4 : Les investisseurs sont rationnels

H5 : «Tous les investisseurs ont le même horizon de décision, qui comporte une seule période».

A partir des 5 hypothèses, Markowitz propose un modèle de décision qui tient compte du caractère hautement combinatoire du portefeuille.

2.3 Présentation du modèle de Markowitz [7][17]

2.3.1 Introduction

La théorie du choix de portefeuille a été développée initialement par Harry Markowitz dans un article publié en 1952. Cet article a posé les fondements de la finance moderne.

Son auteur s'est d'ailleurs vu attribué le prix Nobel d'économie en 1990. (Il a reçu le Prix de théorie John Von Neumann en 1989, et est lauréat du Prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel en 1990).

Une dizaine d'année plus tard, Bill Sharpe et John Lintner publiaient (indépendamment l'un de l'autre) leurs articles présentant le MEDAF.

Bill Sharpe a également reçu le prix Nobel pour sa découverte, la même année qu'Harry

Markowitz. En dépit de sa simplicité, ce modèle résiste remarquablement à l'usure du temps. Les vérifications empiriques du modèle font toujours l'objet de recherches et de discussion dans les milieux académiques [1] pour une synthèse de la littérature. Une enquête récente [1] a cependant révélé que le MEDAF reste le modèle le plus couramment utilisé en pratique pour déterminer le coût du capital d'une entreprise.

Plusieurs modèles d'allocation d'actifs ont été étudiés pour faire de l'optimisation de portefeuille à travers les années. Les premiers développements sont dus à Markowitz par ses travaux sur la quantification et la diversification du risque. Il introduit la notion de frontière efficiente déduite à partir des portefeuilles à variance minimale pour une espérance de rendement donnée, qui représente la combinaison optimale de rendement et de risque.

L'optimisation se fait en définissant une fonction d'utilité représentant les préférences des investisseurs en tenant compte de leur aversion au risque et en maximisant celle-ci étant donné la contrainte représentée par la frontière efficiente. Markowitz a cependant beaucoup travaillé son modèle à travers les années, ce qui fait que sa référence aujourd'hui est surtout son modèle de moyenne-variance (1952). (Plusieurs mesures de risque peuvent être considérées par les modèles d'optimisation de portefeuille. Le cadre classique (Markowitz) considère en particulier la variance. Cette mesure n'est cependant valable que si les rendements sont distribués suivant une loi normale ou si l'investisseur ne se préoccupe que des deux premiers moments de la distribution des rendements de son portefeuille). La figure suivante donne un aperçu visuel de la théorie de portefeuille. Le cadre de Markowitz utilise une fonction d'utilité quadratique convexe.

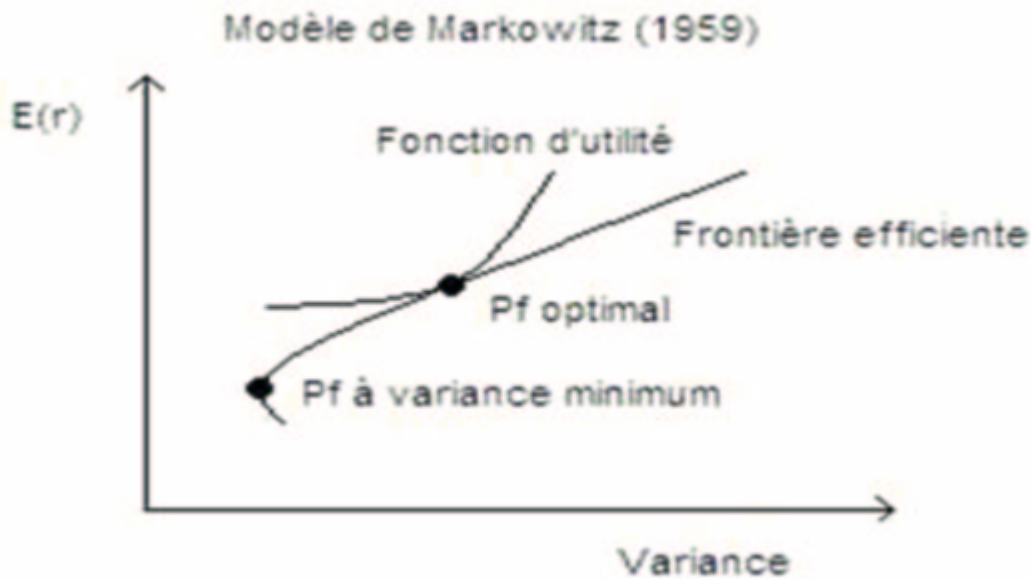


FIGURE 2.1 – Représentation du modèle de Markowitz.

2.3.2 Rentabilité et risque : [3]

- **La rentabilité financière :**

La rentabilité d'un titre financier ou d'un portefeuille est une mesure relative de la rémunération totale de son détenteur calculée à une date t et pour une période de détention données.

La rentabilité prend donc en compte :

- a) le gain ou la perte en capital tiré de la détention du titre ou du portefeuille (la plus- ou moins-value) et
- b) les revenus réels versés sur la période.

La rentabilité dépend également de la durée de détention.

Nous allons passer en revue les rentabilités disponibles et comprendre que les conventions de calcul induisent des représentations différentes des mouvements de cours. Nous étudierons les éventuelles distorsions qu'elles pourraient induire les unes par rapport aux autres. Les mesures de rentabilité que nous considérons n'intègrent pas les coûts de transaction que l'on supporte dans les opérations d'achat-vente.

• **Les rentabilités ex post ou ex ante :**

La différence entre les rentabilités ex post (Les rentabilités arithmétiques) et ex ante (Les rentabilités géométriques) porte sur la nature de la période de détention considérée. Elle est « passée » pour la rentabilité ex post et « à venir » pour celle ex ante. Par construction, les rentabilités ex post sont observées et déduites des cours présents ou passés. De leur côté, les rentabilités ex ante ne peuvent être connues avec certitude puisqu'elles impliquent des prix futurs.

Les rentabilités ex post ou ex ante vont aussi se distinguer par les méthodologies utilisées pour les analyser. Pour décrire, comprendre et modéliser les rentabilités ex post, on peut exploiter des outils statistiques et économétriques sur des données collectées en coupes transversales ou en séries temporelles. Pour décrire, comprendre et modéliser les rentabilités ex ante, on doit développer une théorie financière qui implique le plus souvent des concepts probabilistes. Parfois, le modèle obtenu permet d'extraire la rentabilité ex ante moyenne à partir des prix cotés. Mais la valeur estimée empiriquement dépend intimement du prix du produit financier considéré et des hypothèses du modèle.

Les rentabilités arithmétique et géométrique mobilisent des formules différentes.

• **Mesure de Rendement :**

Il existe plusieurs façons de calculer le rendement.

- Le rendement réalisé au cours d'une période t , noté r_t , est donné en fonction du prix du titre P_t au temps t et du dividende D_t versé entre les temps t et $t - 1$ par :

$$R_t = \frac{(p_t - p_{t-1}) + D_t}{p_{t-1}}$$

- Le rendement arithmétique moyen :

$$\bar{R}_t = \frac{\sum_{t=1}^n r_t}{n}$$

Où n représente le nombre de rendements périodiques.

- Le rendement géométrique :

$$\bar{R}_t = [\prod_{i=1}^n (1 + r_t)]^{\frac{1}{n}} - 1$$

Il constitue une meilleure mesure du rendement périodique moyen historique.

- **Le risque**

L'évolution historique d'un actif financier permet de déterminer un rendement moyen et un écart-type, lequel écart-type est considéré comme une mesure du risque. L'espoir d'un gain substantiel, lors d'un investissement, est généralement proportionnel à l'incertitude du résultat. Ce que les prudents appellent le risque, est qualifié d'espoir par d'autres. La mesure du risque, donc de la difficulté – incertitude – de prévoir le rendement d'un investissement ou d'un portefeuille est bien connue en statistique par : l'ÉCART-TYPE. Ainsi, Postuler une diversification parfaite pour un portefeuille revient à supprimer l'incertitude existant sur la variation relative du return (rendement) de ce portefeuille par rapport au return du marché

...

dans une telle hypothèse, les variations aléatoires – par rapport aux variations de return du marché – du return des éléments constituant ce portefeuille se compensent parfaitement, de sorte que l'on a le même rendement pour le portefeuille que le rendement du marché. Ceci ne veut pas dire que l'on a supprimé toute incertitude sur le return du portefeuille considéré mais que cette incertitude – **écart-type** – est identique à l'incertitude existant au niveau du return du marché.

2.3.3 L'approche espérance-variance :[17]

- **L'espérance mathématique** : mesure de la rentabilité espérée

Placé dans un univers incertain, l'investisseur ne peut pas calculer la rentabilité d'avance car la valeur du titre en fin de période est aléatoire, on utilise alors une rentabilité espérée qui est la moyenne des Rentabilité possible pondérées par leur probabilité de réalisation.

$$\text{Rentabilité espérée} = \sum r_i P(r)$$

Où :

r_i : taux de rentabilité de l'action.

$P(r)$: Probabilité de réalisation du taux de rentabilité.

- **La variance** : outil statistique d'analyse du risque

Intuitivement, on conçoit que, plus le risque d'un titre financier n'est élevé, plus son taux de rentabilité ne variera.

Si le détenteur d'obligations du trésor est assuré de toujours percevoir ses coupons, il est loin d'en être de même pour l'actionnaire d'une société qui intervient dans un -secteur pré potentiel : il pourra perdre ou gagner.

2.4 Portefeuille optimal :(optimisation d'un portefeuille)[4]

2.4.1 Comprendre la nécessité d'une gestion de portefeuille :

L'attention grandissante accordée à la productivité et à la rentabilité oblige les sociétés à concentrer leurs ressources limitées en développement de produits sur des projets assurés de donner des produits gagnants. Mais la gestion des compromis en matière d'investissement entre les nouvelles opportunités, les services et produits existants et la propriété intellectuelle présente sur le marché s'avère très malaisée, surtout lorsque votre équipe utilise des outils

manuels ou déconnectés. Lorsque vous gérez un portefeuille d'idées, de projets et de produits, il est difficile de reconnaître les produits réellement prometteurs de ceux qui sont à revoir ou à abandonner.

Les portefeuilles présentent par conséquent un risque disproportionné par rapport à la rentabilité associée. Les décisions d'investissement se basent souvent sur une intuition, conduisant à négliger des projets de grande valeur.

Dans ce scénario par trop connu, le processus de développement de produits est souvent bien trop lourd par rapport aux capacités. Des projets sont ajoutés sans aucune visibilité des capacités en ressources, ce qui a pour effet de surcharger des fonctions critiques ou des collaborateurs clés. Les projets de développement prennent du retard, manquent de prévisibilité ou pâtissent de problèmes de qualité dus à une implication fonctionnelle insuffisante. Les priorités évoluent dans un flux constant. Les modifications apportées en milieu de projet obligent à parer au plus pressé et à réorganiser massivement les ressources.

L'impact négatif de cette situation est très important. Lorsque les prévisions de revenus et de bénéfices se basent sur des calendriers de projet irréalistes et que le plan de projet n'est pas aligné sur les objectifs stratégiques, les affectations budgétaires qui en résultent ne correspondent pas aux objectifs commerciaux. Dans ces conditions, l'entreprise est incapable de mettre en œuvre sa stratégie commerciale générale.

- **Meilleure rentabilité du portefeuille :**

- Augmentation du pourcentage de produits ou projets qui atteignent les objectifs financiers.
- Équilibre des risques et des succès dans l'ensemble du portefeuille de projets, permettant de maximiser la valeur du portefeuille des nouveaux projets.
- Augmentation de la valeur nette actuelle totale des produits déjà sur le marché.

2.4.2 La diversification :[2][4]

L'approche de Markowitz permet de donner au concept de diversification une signification rigoureuse. au sens large, ce terme signifie : atténuation du risque par la combinaison au sein du portefeuille de plusieurs actifs financiers ; le concept d'efficience permet d'énoncer la proposition suivante : pour tout investisseur, le portefeuille d'utilité maximale, qu'il choisit s'il est rationnel, est un portefeuille optimalement diversifié.

Le cadre « espérance-variance » permet de formaliser une démarche somme toute assez intuitive : la diversification. La formalisation fournit néanmoins des résultats moins évidents qu'il n'y paraît.

Les résultats :

Les résultats de Markowitz en matière de diversification peuvent se résumer par trois propositions :

1. L'ajout d'un actif dans un portefeuille diminue le risque de ce dernier, dès lors que cet actif n'est pas parfaitement et positivement corrélé au portefeuille.
2. Il existe une limite à la diminution du risque par diversification d'un portefeuille. Le risque limite et résiduel peut être appelé **le risque incompressible par diversification** ou **risque non diversifiable**.
3. La volatilité du portefeuille le plus diversifié est la racine carrée de la covariance moyenne qui prévaut en son sein.

La proposition1, est conforme à ce que l'on pouvait attendre du concept de diversification. Plus formellement, elle nous apprend que la volatilité du portefeuille est inférieure à la moyenne pondérée des volatilités individuelles.

La proposition2, est un résultat important qui précise l'existence d'une limite et,

la proposition3, la caractérise. L'existence d'une limite (non nulle) suggère que l'efficacité de la diversification par un $n + 1 - ième$ actif décroît à mesure que le nombre n d'actifs déjà dans le portefeuille est grand. Le coût engendré par la diversification ouvre la question de l'optimisation de cette procédure. Combien d'actifs suffiront à diversifier de manière significative mon portefeuille ? Comment les choisir pour optimiser la diversification à un nombre d'actifs fixé ?

2.4.3 Les actions, les portefeuilles d'actions :[1][8][16][20]

Les actions : Une action est une part des capitaux propres d'une entreprise qui s'échange sur un marché financier.

Les portefeuilles d'actions : Un portefeuille d'actions contient des actions de différentes sociétés. Si l'on note $N_s(t)$ le nombre de sociétés présentes dans le portefeuille à la date t , $n_i(t)$ Le nombre d'actions détenues de la société i et P_t^i le cours de l'action de la $i^{ième}$ société, alors la valeur du portefeuille à la date t peut s'écrire :

$$P_t^p = \sum_{i=1}^{N_s(t)} n_i(t) p_t^i N_s(t)$$

Et $n_i(t)$ sont a priori des fonctions déterministes dont les valeurs sont des décisions de gestion.

$\sum_{i=1}^{N_s(t)} n_i(t)$: est le nombre total de titres détenus dans le portefeuille.

Portefeuille composé de deux actions

Considérons un portefeuille P constitué de deux actifs. Les poids investis dans chacun des actifs sont x_1 et x_2 avec $x_1 + x_2 = 1$. Une valeur positive de x_i représente une position longue.

« Long » : on a acheté l'actif i . Une valeur négative représente une position à découvert

« Short » : l'actif i a été emprunté.

La distribution de probabilité du portefeuille est une loi normale. Elle est identifiée par deux paramètres :

–la rentabilité attendue :

$$r_p = x_1 r_1 + x_2 r_2$$

–l'écart type qui mesure le risque du portefeuille :

$$\sigma_p = \sqrt{x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2(x_1 x_2 \sigma_1 \sigma_2 \rho_{12})}$$

Remarque :

Le coefficient de corrélation(p) : Notez l'apparition du coefficient de corrélation p dans la formule de l'écart type du portefeuille. Les rentabilités des deux actifs peuvent être liées et la corrélation est une mesure de leur relation. Elle peut prendre des valeurs entre -1 et +1. Une autre formulation de l'écart type du portefeuille fait apparaître la covariance entre les rentabilités, une autre manière de mesurer la variabilité conjointe des titres :

$$\sigma_p = \sqrt{x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + 2(x_1x_2\sigma_{12})}$$

Nous commençons par analyser trois cas particuliers avant d'aborder le cas général.

L'un des deux actifs est un actif sans risque :

Cela correspond à la situation où l'on peut placer ou emprunter au taux d'intérêt sans risque r_f . L'actif étant non risqué, sa variance est nulle ($\sigma_{r_f} = 0$). Désignons par A l'actif risqué (A pour actions) et notons x la fraction du portefeuille investie en actions. Les formules générales deviennent :

$$r_p = (1 - x)r_f + xr_A = r_f + (r_A - r_f)x$$

$$\sigma_p = x\sigma_A$$

Il en résulte une relation linéaire entre la rentabilité attendue du portefeuille et son risque :

$$r_p = r_f + \frac{r_A - r_f}{\sigma_A}\sigma_p$$

Les deux actifs sont risqués et parfaitement corrélés positivement ($p_{12} = +1$)

Dans ce cas, la formule de la variance est un carré parfait : le risque du portefeuille est égal à la moyenne pondérée des écarts types de ses titres, ce qui se traduit par : $\sigma_p = -x_1\sigma_1 + x_2\sigma_2$ et $\sigma_p = x_1\sigma_1 - x_2\sigma_2$.

Ces formules font apparaître la possibilité de créer un portefeuille sans risque avec des fractions investies dans chacun des actifs :

$$x_1 = \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

et

$$x_2 = \frac{\sigma_1}{\sigma_1 + \sigma_2}$$

La relation entre le risque et la rentabilité attendue du portefeuille est donnée par deux segments de droites d'équation :

$$r_p = r_1 + \frac{r_2 - r_1}{\sigma_1 + \sigma_2}(\sigma_p - \sigma_1) \text{ pour } x_1 > \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

$$r_p = r_2 + \frac{r_1 - r_2}{\sigma_1 + \sigma_2}(\sigma_p - \sigma_2) \text{ pour } x_1 < \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

Les deux actifs ne sont pas parfaitement corrélés ($-1 < p_{12} < +1$) :

Dans le cas général, la relation entre le risque du portefeuille et sa rentabilité attendue est non linéaire.

Généralisation pour un portefeuille composé de N actifs

Si le nombre de titres en portefeuille est N , la rentabilité attendue et le risque du Portefeuille s'écrivent :

$$r_p = \sum_{i=1}^N x_i r_i$$

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N x_i x_j \sigma_{ij} = \sum_i x_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} x_i x_j \sigma_{ij}$$

Les covariances entre les titres jouent un rôle prépondérant. Le calcul de la variance du portefeuille comprend N^2 termes dont N termes sont des variances et $N^2 - N$ termes des covariances.

Limites de la diversification [1]

Il existe deux limites à la diversification : une limite théorique et une limite pratique.

• Limite théorique :

Prenons des portefeuilles comportant des investissements égaux dans N actions. La proportion investie dans chaque action est donc de $\frac{1}{N}$. Dans chacune des cases de la matrice de variance, nous aurons $(\frac{1}{N})^2 * (Var)$ la variance, et dans chacune des cases de covariances, nous aurons $(\frac{1}{N})^2 * (Cov)$ la covariance. Il y a N cases de variances et $(N^2 - N)$ cases de covariance. Ainsi avec l'augmentation de N , la variance du portefeuille se rapproche de la covariance moyenne. Si la covariance moyenne était nulle, il serait possible d'éliminer totalement le risque en détenant suffisamment de titres. Malheureusement, les actions évoluent ensemble, et non de façon indépendante.

• **Limite pratique :** Au centre de la théorie de Markowitz, on trouve le postulat suivant : « chaque titre comporte un risque que l'on peut décomposer en deux catégories : le risque spécifique de chaque titre, et le risque systématique, lié aux mouvements du marché », donc dans la limite pratique on constate deux types de risque tels :

a. Le risque systématique : Risque qui affecte le système financier dans son ensemble, par opposition au risque associé à un ou plusieurs titres considérés individuellement. On l'appelle aussi le «risque non diversifiable».

b. Le risque spécifique (non systématique) : Risque associé à une valeur en particulier, est celui qui est uniquement dépendant du titre lui-même, donc de la manière et des conditions dans lesquelles est gérée la société. On peut éliminer le risque non systématique par la diversification du portefeuille.

Ce qui découle de cette représentation : Un portefeuille «bien diversifié», une douzaine de titres, a éliminé l'essentiel du risque non-systématique.

2.4.4 La frontière efficiente :[1][7]

La frontière efficiente regroupe les portefeuilles qui, à niveau de risque donné, maximisent la rentabilité espérée. Ce sont les seuls portefeuilles qu'il est rationnel de détenir sur le marché. On dispose de trois approches pour construire la frontière efficiente.

1. La première approche consiste à construire l'enveloppe ξ des portefeuilles de variance minimale à rentabilité espérée fixée, à repérer le portefeuille de variance minimale ($\sigma_{min}, E[R_{min}]$) et à ne retenir que les portefeuilles de l'enveloppe ξ dont la rentabilité espérée est supérieur à $E[R_{min}]$.

2. La deuxième approche consiste à résoudre le programme d'optimisation qui traduit fidèlement la définition d'un portefeuille efficient. On résout :

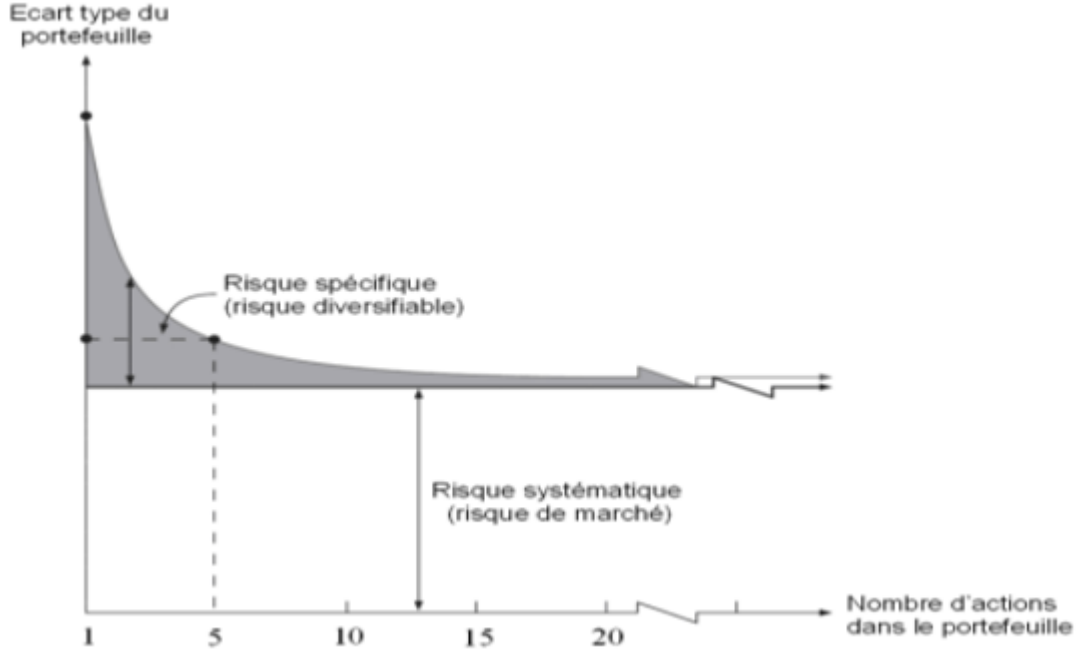


FIGURE 2.2 – représentation d'un risque spécifique et risque systématique

$$\left\{ \begin{array}{l} \max_{x_1, \dots, x_{N_s}} E[R_p] \\ \text{sc} \\ \sigma_p^2 = C_{\sigma^2} \\ \sum_{i=1}^{N_s} x_i = 1 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \max_x R'X \\ \text{sc} \\ X'VX = C_{\sigma^2} \\ U'X = 1 \end{array} \right.$$

Où C_{σ^2} est une constante.

Et "sc" signifie « sous contraintes ». Le système à droite rappelle que la rentabilité espérée $E[R]$ et la variance σ_p^2 sont bien deux fonctions du vecteur poids X .

Chaque portefeuille solution est alors décrit par un vecteur des poids $X(C_{\sigma^2})$ qui assurent la variance minimale (à rentabilité espérée fixée C_{σ^2}).

3. La troisième approche consiste à construire l'ensemble des portefeuilles optimaux de tous les investisseurs. Mais il convient au préalable de savoir caractériser les portefeuilles optimaux de chacun d'entre eux.

Le concept de choix de portefeuille optimal :[22]

L'article de Markowitz publié en 1952 dans le (Journal of Finance) constitue le début de la théorie moderne du choix de portefeuilles. Ce cadre d'analyse dit « espérance-variance » donne une forte cohérence aux décisions financières. Il aide encore aujourd'hui le gestionnaire à formaliser et rationaliser sa prise de décision.

Souvent, il s'impose comme outil d'analyse et de justification des décisions prises. La théorie du choix de portefeuille est essentiellement normative.

Les portefeuilles efficients développés par Markowitz en 1959, sont des portefeuilles d'actifs dont la composition permet à un investisseur d'optimiser le couple rendement /risque. Les

portefeuilles efficients 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.

Deux définitions sont envisageables pour caractériser le portefeuille optimal d'un investisseur donné. Les deux font appel à la notion d'utilité.

1. Est optimal pour un investisseur donné le portefeuille de la frontière efficiente qui lui procure l'utilité la plus grande.
2. Est optimal pour un investisseur donné le portefeuille réalisable qui maximise sa fonction d'utilité.

Ces définitions suggèrent deux approches différentes pour construire le portefeuille optimal. Au préalable, on rappelle très succinctement quelques concepts liés aux fonctions d'utilité. Admettons que l'utilité des investisseurs soit bien décrite par une fonction quadratique des coordonnées du portefeuille dans l'espace $(0, \sigma, E[R])$. On pose conventionnellement :

$$U(P) = U(\sigma_p, E[R_p]) = E[R_p] - \frac{\Phi}{2} \sigma_p^2 = R'X - \frac{\Phi}{2} X'VX$$

Avec Φ un nombre positif qui capture l'aversion au risque de l'investisseur. Les portefeuilles vérifiant $U(P) = c$ procurent à l'investisseur la même utilité. Leurs coordonnées vérifient :

$$E[R_p] = c + \frac{\Phi}{2} \sigma_p^2$$

Cette équation décrit les courbes d'indifférence ou d'iso-utilité.

La première définition suggère alors une procédure à trois étapes qui consiste à :

- a) construire la frontière efficiente,
- b) identifier les courbes d'indifférence (c.-à-d. les ensembles de portefeuilles procurant le même degré de satisfaction) et
- c) en déduire le portefeuille efficient le plus « utile ». Le portefeuille optimal se situe à la fois sur la frontière efficiente et sur la courbe d'indifférence de plus grande utilité. C'est le portefeuille qui est tangent à la frontière efficiente et qui appartient à une des courbes d'indifférence de l'investisseur.

Remarque : Si Φ est nul, alors l'investisseur ne privilégie que l'espérance de rentabilité qu'il cherchera à maximiser (sans considération aucune du niveau de risque pris). Si, au contraire, Φ est très grand, alors la variance (le risque) est son souci majeur, et maximiser son utilité est pour l'essentiel un exercice de minimisation de la variance.

La seconde suggère de maximiser directement la fonction d'utilité en conservant la contrainte $\sum_{i=1}^{N_s} x_i = 1$.

On cherche à résoudre ici :

$$\left\{ \begin{array}{l} \max_{x_1, \dots, x_{N_s}} (\sigma_p, E[R_p]) \\ sc \\ \sum_{i=1}^{N_s} x_i = 1 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \max_x R'X - \frac{\Phi}{2} X'VX \\ sc \\ U'V = 1 \end{array} \right.$$

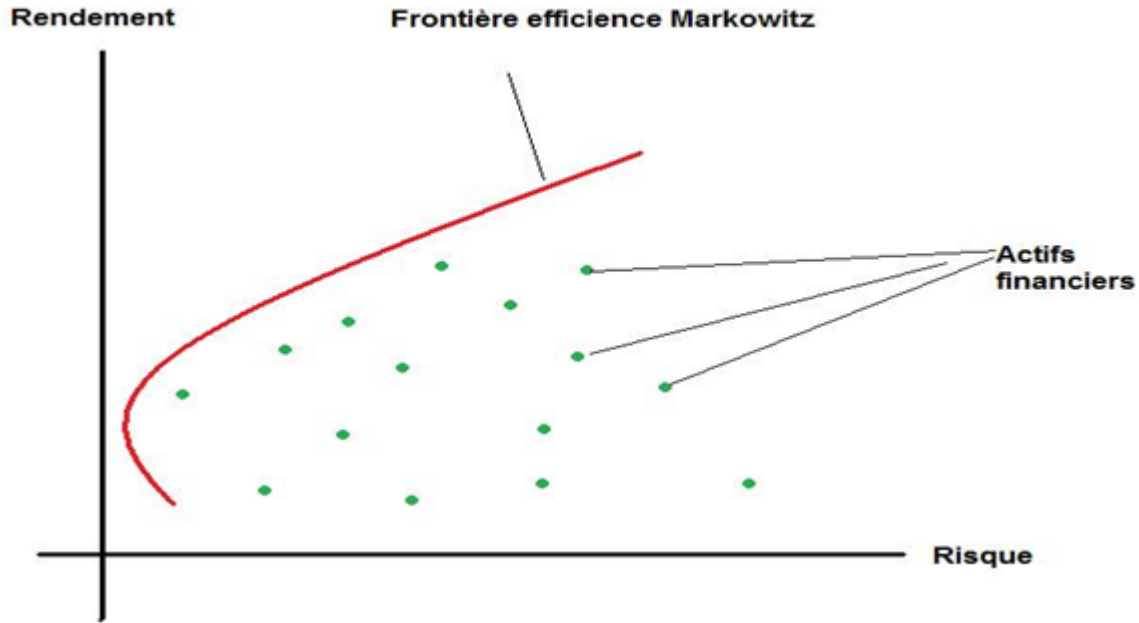


FIGURE 2.3 – La représentation de la frontière efficiente de Markowitz

La théorie du choix de portefeuille optimal :[1][22]

La gestion de portefeuille est un problème de décision en incertitude. Seule la fonction d'utilité quadratique permet une analyse algébrique et intuitive de la composition des portefeuilles optimaux sans qu'on ait à la spécifier complètement. Les autres fonctions ne permettent d'identifier un portefeuille optimal que si elles sont complètement spécifiées et reposent sur des méthodes d'optimisation numériques.

La plupart des idées vues précédemment proviennent de l'article "Portfolio Selection" d'Harry Markowitz (Journal of Finance, 1952). Celui-ci a montré comment un investisseur peut réduire le risque de son portefeuille à l'aide de la diversification. C'est la théorie moderne du portefeuille.

Pour aborder ce problème rigoureusement, on doit d'abord le formuler mathématiquement comme un problème d'optimisation.

- L'espérance de rentabilité d'un portefeuille est donné par :

$$x_i = \frac{\text{valeur du titre } i}{\text{valeur totale du portefeuille}}$$

Tel que : $\sum_{i=1}^N x_i = 1$.

- La rentabilité d'un portefeuille est égale à la moyenne pondérée des titres qui le composent donc :

$$R_p = x_1 * R_1 + \dots + x_N * R_N = \sum_{i=1}^N x_i R_i$$

$$E[R_p] = E[\sum_{i=1}^N x_i R_i] = \sum_{i=1}^N E[x_i R_i] = \sum_{i=1}^N x_i E[R_i]$$

. Avec l'espérance de R_i est simplement la moyenne arithmétique.

Calcul de la covariance et de la corrélation : nous supposons que les rendements des différents actifs financiers ne fluctuent pas indépendamment les uns des autres : ils sont corrélés, c'est-à-dire ont des covariances non nulles : $cov(R_i, R_j) \neq 0$.

Covariance des rentabilités R_i et R_j est :

$$cov(R_i, R_j) = E[(R_i - E[R_i])(R_j - E[R_j])]$$

- Calcul de la variance et de l'écart type d'un portefeuille :

- Variance d'un portefeuille composé de deux titres :

$$var[R_p] = cov(R_p, R_p) = cov(x_1 R_1 + x_2 R_2, x_1 R_1 + x_2 R_2);$$

$$= x_1 x_1 cov(R_1, R_1) + x_1 x_2 cov(R_1, R_2) + x_2 x_1 cov(R_2, R_1) + x_2 x_2 cov(R_2, R_2);$$

$$= x_1^2 var[R_1] + x_2^2 var[R_2] + 2x_1 x_2 cov(R_1, R_2);$$

$$= x_1^2 var[R_1] + x_2^2 var[R_2] + 2x_1 x_2 \sigma_{R_1} \sigma_{R_2} cov(R_1, R_2).$$

$$\text{Avec } \sigma_{R_p} = \sqrt{(var[R_p])}.$$

- La volatilité d'un portefeuille composé de N actions :

$$var[R_p] = cov(R_p, R_p) = cov(\sum_{i=1}^N x_i R_i, R_p) = \sum_{i=1}^N x_i cov(R_i, R_p).$$

$$var[R_p] = \sum_{i=1}^N x_i cov(R_i, R_p) = \sum_{i=1}^N x_i cov(R_i, \sum_{j=1}^N x_j R_j) = \sum_{i=1}^N x_i \sum_{j=1}^N x_j cov(R_i, R_j).$$

Donc :

$$var[R_p] = \sum_{i=1}^N \sum_{j=1}^N x_i x_j cov(R_i, R_j).$$

- Formulation matricielle du système :

Notons X le vecteur des parts d'actifs et X^t est son transposé donc :

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}; X^T = (X_1 X_2 \cdots X_n)$$

D'où la matrice des covariances est :

$$(C_{i,j}) = \begin{bmatrix} cov(R_1, R_1) & cov(R_1, R_2) & \dots & cov(R_1, R_n) \\ cov(R_2, R_1) & cov(R_2, R_2) & \dots & cov(R_2, R_n) \\ \vdots & \vdots & & \vdots \\ cov(R_n, R_1) & cov(R_n, R_2) & \dots & cov(R_n, R_n) \end{bmatrix}$$

Une matrice qui se simplifie directement en :

$$(C_{i,j}) = \begin{bmatrix} \sigma_1^2 & cov(R_1, R_2) & \dots & cov(R_1, R_n) \\ cov(R_2, R_1) & \sigma_2^2 & \dots & cov(R_2, R_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(R_n, R_1) & cov(R_n, R_2) & \dots & \sigma_n^2 \end{bmatrix}$$

Car $cov(R_i, R_i) = \sigma_i^2$.

En résumé, la relation de la variance sous forme matricielle est :

$$Var(R_p) = X^T \cdot (C_{i,j}) \cdot X.$$

Sélectionner un portefeuille revient donc à résoudre le problème de maximisation sous les contraintes suivantes :

$$\begin{cases} \min Var(R_p) \\ E(R_p) = C^{te} \\ \sum X_i = 1 \end{cases}$$

En utilisant la programmation quadratique.

Dans la pratique, on cherche non pas un, mais tous les portefeuilles qui pour une espérance donnée minimise la variance.

Définition :

La frontière qui caractérise le polygone ou la courbe des contraintes s'appelle dans cette situation la "frontière efficiente (de Markowitz)" et dans le polygone/courbe se situent tous les portefeuilles à rejeter dits "portefeuilles dominés". Une autre manière de formuler ceci consiste à dire que les combinaisons (rendement, risque) de cette frontière forment un ensemble d'optima de Pareto, c'est-à-dire que si l'un des éléments augmente, l'autre doit augmenter aussi. Maintenant, formalisons l'optimisation comme cela était fait au moment où les gens devaient encore développer les algorithmes eux mêmes.

Soit Z la fonction économique définie par :

$$Z = \varphi E(R_p) - V(R_p)$$

Qui doit être maximisée sous la contrainte que $\sum_{i=1}^n X_i = 1$ et où φ est un paramètre qui représente le degré d'aversion au risque des investisseurs (histoire aussi d'homogénéiser la relation...).

Le problème de maximisation sous contrainte consiste à déterminer le maximum de la fonction économique Z définie par :

$$Z = \varphi \sum_{i=1}^n X_i E(R_i) - \sum_{i=1}^n \sum_{j=1}^n X_i X_j cov(R_i, R_j) + \lambda (1 - \sum_{i=1}^n X_i)$$

Cette fonction de $n+1$ variables $(X_1, X_2, \dots, X_n, \lambda)$ est maximisée si sa dérivée partielle par rapport à chacune de ces variables est nulle, ce qui revient à poser le système suivant :

$$\begin{cases} \frac{\partial Z}{\partial X_1} = \varphi E(R_1) - 2X_1 \text{cov}(R_1, R_1) - 2X_2 \text{cov}(R_1, R_2) - \dots - 2X_n \text{cov}(R_1, R_n) - \lambda = 0 \\ \frac{\partial Z}{\partial X_2} = \varphi E(R_2) - 2X_2 \text{cov}(R_2, R_1) - 2X_2 \text{cov}(R_2, R_2) - \dots - 2X_n \text{cov}(R_2, R_n) - \lambda = 0 \\ \vdots \\ \frac{\partial Z}{\partial X_n} = \varphi E(R_n) - 2X_n \text{cov}(R_n, R_1) - 2X_n \text{cov}(R_n, R_2) - \dots - 2X_n \text{cov}(R_n, R_n) - \lambda = 0 \\ \frac{\partial Z}{\partial \lambda} = 1 - X_1 - X_2 - \dots - X_n = 0 \end{cases}$$

Posont, $\text{cov}(R_i, R_j) = \sigma_{ij}$ d'où le système devient :

$$\begin{cases} 2X_1\sigma_{11} + 2X_2\sigma_{12} + \dots + 2X_n\sigma_{1n} + \lambda = \varphi E(R_1) \\ 2X_1\sigma_{12} + 2X_2\sigma_{22} + \dots + 2X_n\sigma_{2n} + \lambda = \varphi E(R_2) \\ \vdots \\ 2X_1\sigma_{n1} + 2X_2\sigma_{n2} + \dots + 2X_n\sigma_{nn} + \lambda = \varphi E(R_n) \\ X_1 + X_2 + \dots + X_n = 1 \end{cases}$$

Soit sous forme matricielle :

$$\begin{pmatrix} 2\sigma_{11} & 2\sigma_{12} & \dots & 2\sigma_{1n} & 1 \\ 2\sigma_{21} & 2\sigma_{22} & \dots & 2\sigma_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 2\sigma_{n1} & 2\sigma_{n2} & \dots & 2\sigma_{nn} & 1 \\ 1 & 1 & \dots & & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \\ \lambda \end{pmatrix} = \begin{pmatrix} \varphi E(R_1) \\ \varphi E(R_2) \\ \vdots \\ \varphi E(R_n) \\ 1 \end{pmatrix}$$

Par conséquent, $X = A^{-1}B$.

La détermination du poids de chacun des n actifs susceptibles d'entrer dans la composition d'un portefeuille passe donc par l'inversion d'une matrice carrée de $n + 1$ lignes et $n + 1$ colonnes comportant $(n^2 - n)/2$ covariances (la diagonale comportant des variances seulement et la matrice étant symétrique). Ce qui est relativement long à calculer pour de gros portefeuilles.

Cependant, même une fois la pondération des actifs terminée, le problème lui ne l'est pas complètement. Effectivement, nous pouvons donc connaître la frontière efficiente mais le client va lui imposer une contrainte bien logique au niveau du risque nul de son portefeuille et du rapport rendement/risque maximum.

Il est donc possible de constituer une infinité de portefeuilles en faisant varier les proportions investies dans chacun des titres. La prochaine étape consiste à sélectionner, parmi l'ensemble des portefeuilles disponibles, un portefeuille donné. Pour ce faire, on doit considérer les préférences individuelles de l'investisseur.

Un investisseur rationnel ne devrait donc considérer que les portefeuilles se trouvant sur la frontière efficiente pour ses choix d'investissement. Son portefeuille optimal se situera donc au point de tangence entre la frontière efficiente et sa courbe d'indifférence la plus haute qu'il serait capable d'atteindre. En procédant ainsi, chaque investisseur maximisera son utilité espérée. En présence d'une économie ne contenant que des actifs risqués, la composition du portefeuille d'actifs risqués varie d'un individu à un autre.

En pratique, les investisseurs ont également la possibilité d'investir dans des actifs financiers

sans risques. Nous allons donc chercher à déterminer la nouvelle frontière efficiente en tenant compte de cette nouvelle opportunité d'investissement.

Considérons alors un portefeuille qui est une combinaison de l'actif sans risque et d'un portefeuille de marché à risque.

Nous avons alors,

$$R_p = X_m R_m + (1 - X_m) R_f.$$

Où X_m est la fraction du portefeuille investie dans le portefeuille du marché (m), et R_f est le "taux de rendement certain".

Rappel : L'espérance d'une constante est égale à cette constante.

D'où : $E(R_p) = X_m E(R_m) + (1 - X_m) E(R_f) = X_m E(R_m) + (1 - X_m) R_f$.

Donc :

$$Var(R_p) = E((R_p - E(R_p))^2)$$

$$= E((X_m R_m + (1 - X_m) R_f - X_m E(R_m) - (1 - X_m) R_f)^2)$$

$$= E((X_m (R_m - E(R_m)))^2)$$

$$= X_m^2 Var(R_m)$$

Soit : $\sigma_p^2 = X_m^2 \sigma_m^2$.

La dérivée de du rendement espéré par rapport à X_m nous donne :

$$\frac{\partial E(R_p)}{\partial X_m} = E(R_m) - R_f$$

La dérivée de l'écart-type par rapport à X_m nous donne :

$$\frac{\partial \sigma(R_p)}{\partial X_m} = \sigma(R_m)$$

En mettant ces deux résultats ensemble on aura :

$$\frac{\partial E(R_p)}{\partial \sigma(R_p)} = \frac{E(R_m) - R_f}{\sigma(R_m)}$$

Cette équation nous donne la pente de la "capital market line" (C.M.L).

Définition de la C.M.L : [1]

La **C.M.L** est la droite formée par l'ensemble des portefeuilles composés d'actif sans risque et du portefeuille risqué (mesuré par l'écart type des rentabilités des titres).

Elle est constante (la pente), comme la C.M.L est une droite, L'ordonnée à l'origine est évidemment R_f .

Puisque :

$$E(R_p) = X_m E(R_m) + (1 - X_m) E(R_f) = X_m E(R_m) + (1 - X_m) R_f.$$

L'équation de la C.M.L. se réduit alors à :

$$E(R_p) = \frac{\sigma(R_p)}{\sigma(R_m)} (E(R_m) - R_f) + R_f$$

Et puisque dans la finance l'intérêt est de représenter graphiquement.

$$E(R_p) = f(\sigma(R_p))$$

$$\text{Donc, } E(R_p) = \frac{E(R_m) - R_f}{\sigma(R_m)} \sigma(R_p) + R_f.$$

Par construction, cette droite associe donc à chaque niveau de risque, la rentabilité espérée la plus élevée. Ainsi, étant donnée le rendement d'un actif sans risque il devient facile à partir de cette équation de déterminer le point de tangence avec la frontière d'efficience de Markowitz ou de Sharpe pour obtenir le portefeuille le plus efficace sur la base du rendement sans risque.

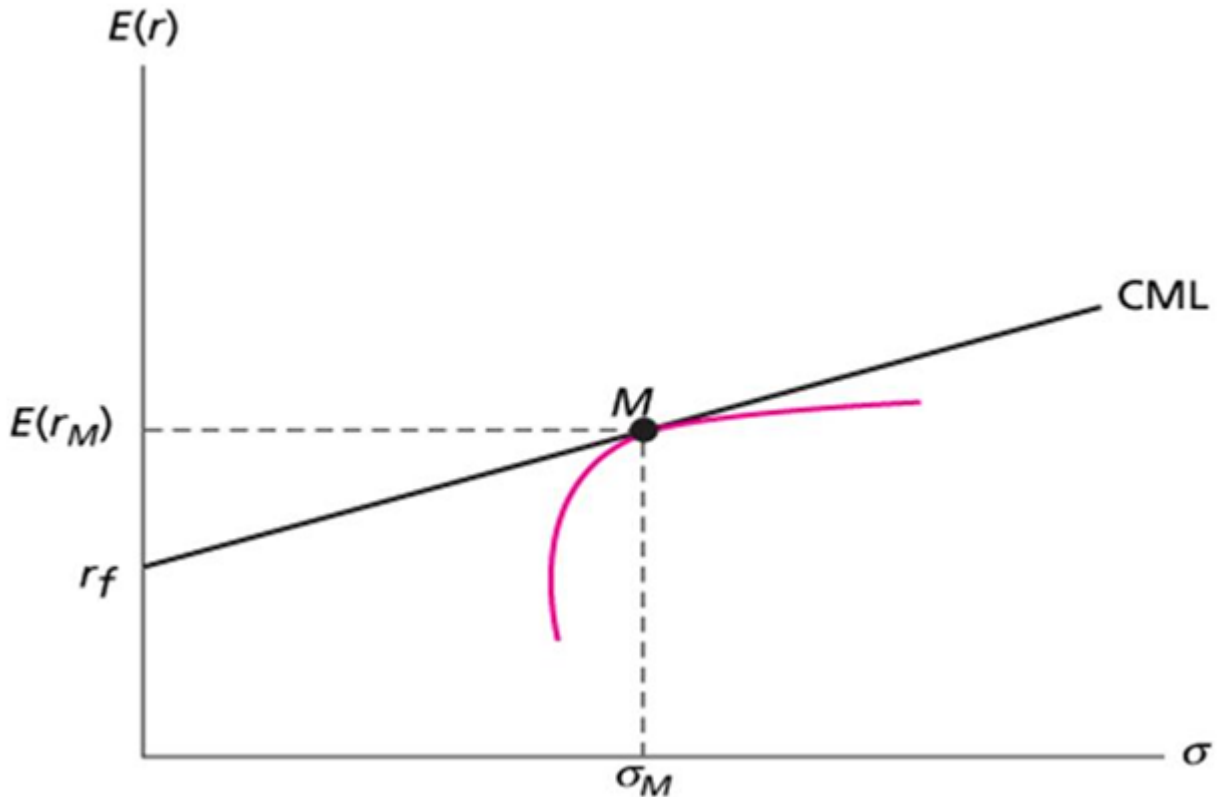


FIGURE 2.4 – La représentation de la C.M.L

Les limites du Modèle de Markowitz :[17]

Depuis son apparition, le modèle de Markowitz a pris une place très importante dans l'évolution de la finance moderne et il a réalisé beaucoup de succès avec son apport en matière de gestion de portefeuille.

Mais avec les ajustements récents, ce modèle s'est trouvé plusieurs limites soulevées par plusieurs praticiens de la théorie financière. Parmi ces limites, on note :

- Le modèle suppose la rationalité des investisseurs. Or, la réalité a prouvé qu'une croyance tout à fait irrationnelle peut être vue légitime par le seul fait qu'elle soit collectivement admise par un opérateur crédible ;
- Le modèle ne s'est pas intéressé à la décomposition du risque global du marché mais s'est

limité à l'analyse et à l'évaluation du risque individuel ou spécifique ; d'où l'apparition d'un nouveau modèle d'évaluation des actifs financiers (MEDAF) ;

- Le modèle suppose également la normalité de la distribution des rentabilités, chose qui n'est pas toujours vérifiable dans la réalité. Cette limite a été résolue par l'apparition du modèle de «Dominance stochastique» qui s'applique à tout type de distribution ;

- La variance a été considérée comme une mesure simplificatrice de la fonction de la rentabilité, tandis que la «Dominance stochastique» admet une comparaison de la distribution entière ;

- La variance étant une mesure non parfaite du risque, une nouvelle technique de mesure a été développée en 1993, appelé Value-At-Risk (VaR). Cette technique permet de déterminer la perte maximale probabilisée sur un portefeuille quelconque.

- Le temps est concentré entre deux instants. Les mêmes actifs financiers sont à la disposition de tous les investisseurs et il y a un actif sans risque.

- Tous les investisseurs utilisent le modèle de Markowitz pour constituer leurs portefeuilles.

- Tous les investisseurs ont le même horizon de temps et les mêmes anticipations.

- Les investisseurs ne peuvent pas influencer individuellement les prix (concurrence parfaite).

- Le marché est sans friction (pas de coût de transaction ou d'information, pas de taxe, les ventes à découvert sont autorisées, les titres sont parfaitement divisibles).

Application

Construction de portefeuille composé de deux actions :

Plan du travail :

- Espérance de rentabilité et écart type d'un portefeuille.
- Calcul de la covariance et de la corrélation.
- Choix du portefeuille efficient.
- Le portefeuille de marché.

Soient A, B, C trois entreprises et a, b, c leurs actions

1. L'espérance de rentabilité d'un portefeuille :

v_t = valeur du titre

v_{tp} = valeur totale du portefeuille.

$$x_i = \frac{v_t}{v_{tp}}$$

Avec : $\sum_i x_i = 1$.

La rentabilité d'un portefeuille est égale à la moyenne pondérée des titres qui le composent.

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

$$E[R_p] = \sum_{i=1}^N x_i E[R_i].$$

2. L'écart type d'un portefeuille composé de deux actions :

Soit le tableau suivant :

Année	A	B	C	1/2R _A +1/2R _B	1/2R _B +1/2R _C
2003	21%	9%	-2%	15%	4%
2004	30%	21%	-5%	26%	8%
2005	7%	7%	9%	7%	8%
2006	-5%	-2%	21%	-4%	10%
2007	-2%	-5%	30%	20%	13%
2008	9%	30%	7%	20%	19%
Rentabilités	10%	10%	10%	10%	10%
écart type	13%	13%	13%	12%	5%

3. Calcule de la variance et de la corrélation :

$$\text{cov}(R_i, R_j) = E[(R_i - E[R_i])(R_j - E[R_j])]$$

Estimation de la covariance à partir de rentabilités :

$$\text{cov}(R_i, R_j) = \frac{1}{T-1} \sum_{t=1}^T (R_{i,t} - \bar{R}_i)(R_{j,t} - \bar{R}_j)$$

D'où la corrélation est égale :

$$\text{cor}(R_i, R_j) = \frac{\text{cov}(R_i, R_j)}{\sigma_{R_i} * \sigma_{R_j}}$$

Application numérique :

$T = 6$.

$$\text{somme} = \sum_{t=1}^6 (R_{i,t} - \bar{R}_i)(R_{j,t} - \bar{R}_j).$$

Pour A et B somme = 0.0558.

Pour B et C , somme = -0.0642.

Covariance :

$$\text{cov}(R_i, R_j) = \frac{1}{T-1} * \text{somme} = \frac{1}{5} * \text{somme}.$$

Covariance pour A et B , $\text{cov} = 62\%$.

Pour B et C , $\text{cov} = -71.33\%$.

Voir le tableau2.

écart type				A et B	B et C
Année	(RA-RA)	(RB-RB)	RC-RC)	(RA-RA)(RB-RB)	(RB-RB)(RC-RC)
2003	11%	-1%	-12%	0,00	0,0012
2004	20%	11%	-15%	0,02	-0,0165
2005	-3%	-3%	-1%	0,00	0,0003
2006	-15%	-12%	11%	0,02	-0,0132
2007	-12%	-15%	20%	0,02	-0,03
2008	-1%	20%	-3%	0,00	-0,006
somme=				0,0558	-0,0642
covariance=				0,0112	-1,0128
corrélacion=				62%	-71%

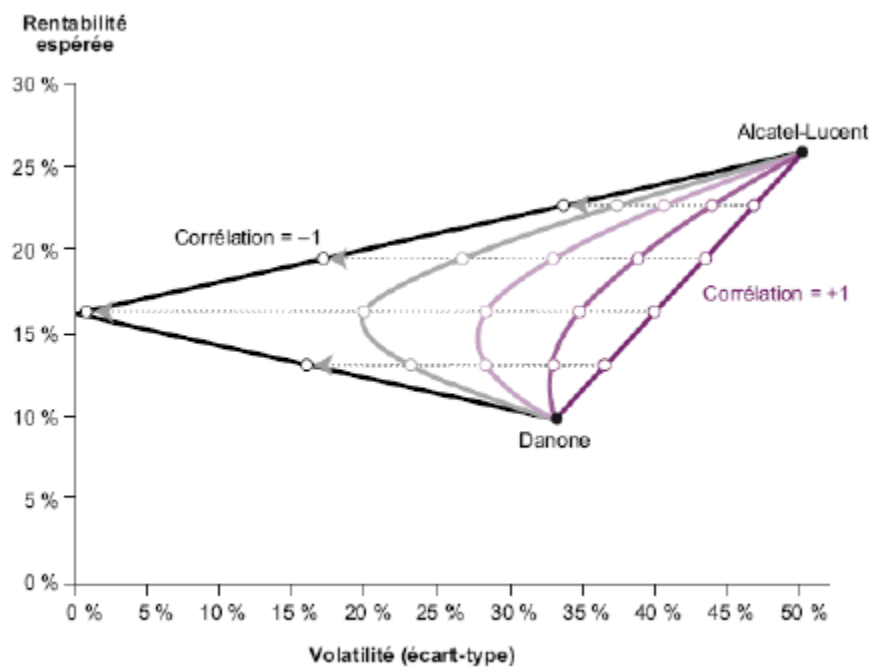


FIGURE 3.1 – les différents niveaux de corrélation entre les rentabilités des deux titres.

4. Détermination d'un portefeuille efficient composé de deux actions :

On prend deux actions a et b des deux entreprises (A et B).

Avec :

a : c'est l'action Alcatel-Lucent.

b : c'est l'action Danone.

Soit le tableau suivant :

action	rendement	écart type
a	26 %	50 %
b	6 %	25 %

On suppose que la corrélation est nulle :

Détermination de portefeuille efficient :

Portefeuille composé de 40% de l'action a et de 60% de l'action b donc :

$$E[R_{40-60}] = x_a E[R_a] + x_b E[R_b].$$

$$var[R_{40-60}] = x_a^2 \sigma_{R_a}^2 + x_b^2 \sigma_{R_b}^2 + 2 * x_a * x_b * \sigma_{R_a} * \sigma_{R_b} * cor(R_a, R_b).$$

$$\sigma_{R_{40-60}} = \sqrt{var[R_{40-60}]}.$$

Application numérique :

$$E[R_{40-60}] = 0.4 * 0.26 + 0.6 * 0.06 = 14\%$$

$$var[R_{40-60}] = 0.4^2 * 0.50^2 + 0.6^2 * 0.25^2 + 2 * 0.4 * 0.6 * 0.50 * 0.25 * 0 = 0.0625.$$

$$\sigma_{R_{40-60}} = \sqrt{0.0625} = 25\%.$$

A partir du tableau 3 on aura le graphe qui montre les portefeuilles efficients composés de deux actions a et b .

actions		rentabilité	Ecart- type
x_a	x_b	E[R]	σ_R
100%	0%	26%	50%
80%	20%	22%	40%
60%	40%	18%	32%
40%	60%	14%	25%
20%	80%	10%	22%
0%	100%	6%	25%

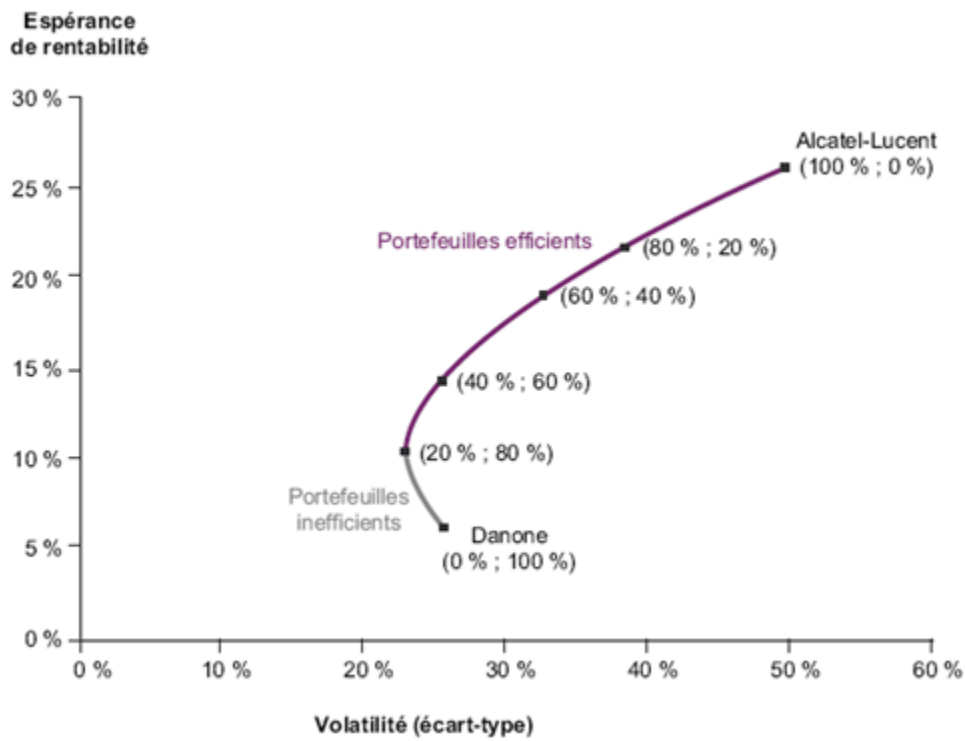


FIGURE 3.2 – Les portefeuilles efficaces composés de deux actions

Conclusion générale

La conclusion du modèle de Markowitz, c'est de conceptualiser et quantifier le couple risque-rentabilité, rendant ainsi possible la détermination des proportions optimales à investir dans les différents actifs financiers pris en considération par l'investisseur ou le gestionnaire de fortune. C'est toutefois depuis le milieu des années 1960, avec les travaux de Sharpe, Lintner et Mossin (sur les conditions d'équilibre des marchés financiers) et de Fama (sur l'efficience de ces mêmes marchés), que la littérature relative à la gestion de portefeuille connaît un extraordinaire développement qui semble encore loin de son terme. La communauté des investisseurs avait commencé à parler de risque sans pour autant pouvoir exprimer la notion de risque en des termes spécifiques. Mais, pour construire leur portefeuille, les investisseurs ont besoin de quantifier le risque qui est défini comme l'incertitude quant aux résultats futurs. Harry Markowitz, a construit le modèle de base d'un portefeuille avec la mesure du taux de rendement espéré et du risque espéré. Il montre que la variance (ou l'écart type) du taux de rendement constitue une bonne mesure du risque d'un portefeuille sous certaines hypothèses et donne une formulation de la variance d'un portefeuille. Ainsi, il ouvre la voie à l'optimisation de portefeuilles. Un portefeuille est dit efficient s'il admet un taux de rendement maximal pour un risque donné ou un risque minimal pour un niveau de rendement donné. L'ensemble des portefeuilles efficients définit la frontière efficiente.

Ensuite, pour palier le problème de dimension, principal inconvénient du modèle de Markowitz, Sharpe a développé un modèle de marché à indice unique pour expliquer le rendement des titres.

L'évolution des opportunités d'investissement peut avoir des effets importants sur les portefeuilles optimaux d'investisseurs avec différents horizons de placements (Campbell & Viceira). Sur ce dernier point, plusieurs recherches empiriques (Fama & French (1971), Campbell & al), dans (le modèle de fixation de prix d'un actif Financier) de William Sharpe ont démontré que l'espérance de rendements des titres semble varier avec le temps, ce qui implique donc que les opportunités d'investissement ne sont pas constantes.

Ce sont ces même professionnels qui par leurs travaux génèrent et diffusent l'information permettant aux marchés modernes d'approcher l'état d'efficience.

Bibliographie

- [1] :A.Farber, M.P.Laurent, K.Oosterlinck, Finance du marché 2^e edition, Pearson, 2010.
- [2] :A.Cabrol, Gestion de portefeuille, ENSAE, 2006.
- [3] :B.Jacquillat, B.Solnik, C.Perignon, Gestion des portefeuilles et des risques,DUNOD,2014.
- [4] :D.C.Dang, Composition d'un portefeuille optimal,l'IAE de Paris,1998.
- [5] :D.Verwaerde, P.Laurent-Gengoux, Optimisation,Central Paris, 2006-2007.
- [6] :F.Rossi, Optimisation,Paris Tech,2009-2010.
- [7] :F.Moraux, Finance du marché,Person,2010.
- [8] :G.Gasso, Introduction à l'optimisation avec contraintes,INSA Rouen,2016.
- [9] :H.Hocquard, Optimisation des fonctions de plusieurs variables,Bordeaux.France,2013.
- [10] :J.Lehec, Analyse convexe approfondie,Paris-Dauphine,2013-2014.
- [11] :J.Grenet, Lagrangien et conditions de Kuhn et Tucker,TECK,2007-2008.
- [12] :L.Guillop, Optimisation sous contraintes, Nantes,2015-2016.
- [13] :M.Bierlaire, Méthode de Newton pour l'optimisation quadratique,EPFL, 2010.
- [14] :M.Minoux, Programmation mathématique : théorie et algorithmes,Tec et Doc-Lavoisier,2007.
- [15] :M.Ouriemchi, Résolution des problèmes non linéaires,Havre,2005.
- [16] :M.Posson, Économie financière et gestion de portefeuille,HEC,2005-2006.
- [17] :O.Bourachnikova, la théorie comportementale du portefeuille :(le modèle moyenne-variance),IFS,2009.

- [18] :P.Poncet, R.Portait, La théorie moderne du portefeuille, STDI,2009.
- [19] :P.Ciarlet, Hasnaa.Zidani, Optimisation quadratique,AO101,2010-2011.
- [20] :P.Bernard, Le modèle d'équilibre des actifs financiers,Paris-Dauphine,2007.
- [21] :R.Omheni, Méthodes primales-duales pour l'optimisation, archives-ouvertes,2015.
- [22] :S.Blais, Théories avancés du portefeuille,UQO,2014.
- [23] :S.Mottelet, Optimisation non linéaire,Tec,2003.