

l'optimisation multiobjectif

Tiré des cours de:

Alain Berro

El Ghazali Talbi

Plan

- Définitions et problématiques
- Les méthodes agrégées
- Les méthodes non agrégées et non Pareto
- Les méthodes basées sur Pareto
- Synthèse
- Discussion

Optimisation multi-objectifs

- Nombreux secteurs de l'industrie concernés (**Télécommunications, Transport, Environnement, Mécanique, Aéronautique, ...**).
- Racines dans le 19^{ième} siècle dans les travaux en **économie de Edgeworth et Pareto (Sciences de l'ingénieur, Management)**.
- Optimisation multi-critères linéaire ou non-linéaire en variables continues [Steuer86, White 90].
- **Optimisation combinatoire multi-critères.**

Les problèmes d'optimisation multiobjectifs

- [Un producteur de vin

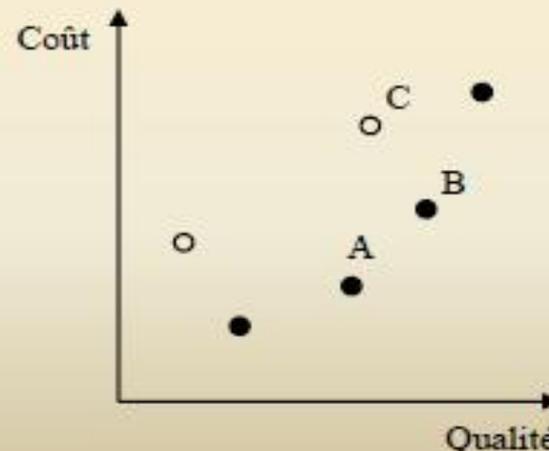
- ✓ Maximiser sa qualité
- ✓ Minimiser son coût de production

↗ la qualité → ↗ des coûts moyens

↘ les coûts moyens → ↘ de la qualité

↘ les coûts marginaux par ↗ du volume de production →

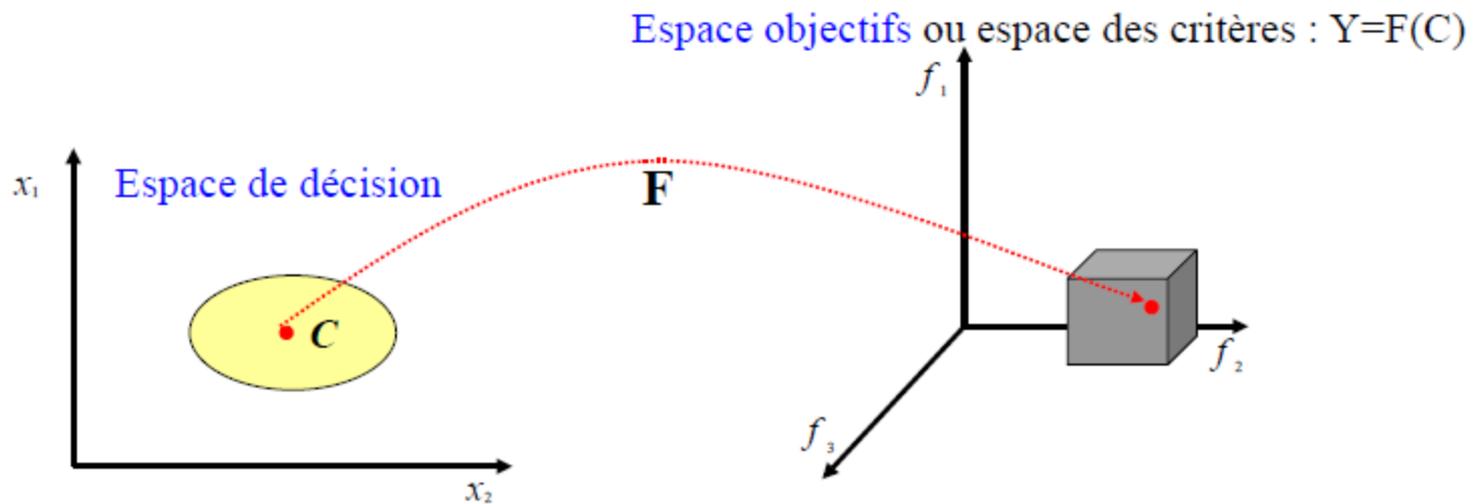
↘ coûts moyens mais une perception en ↘ de la qualité



Optimisation multi-objectifs

$$\text{(PMO)} \left\{ \begin{array}{l} \min F(x) = (f_1(x), f_2(x), \dots, f_n(x)) \quad n \geq 2 \\ \text{s.c. } x \in C \end{array} \right.$$

Variables de décision $x = (x_1, x_2, \dots, x_k)$



Famille de problèmes

- Ordonnancement[Sayinet al. 99]
- Cheminement[Warburton87],
- Arbre recouvrant[Zhouet Gen99]
- Voyageur de commerce[Serafini92],
- Affectation[Teghem95]
- Routage de véhicules[Park et Koelling89], ...

Domaine d'application

- Télécommunications:
 - Design d'antennes [Sandlinet al. 97],
 - affectation de fréquences [Dahl et al. 95],...
- Aéronautique:
 - ailes d'avions [Obayashi98],
 - moteurs [Fujita98],...
- Environnement:
 - gestion de la qualité de l'air [Loughlin98],
 - distribution de l'eau, ...
- Transport:
 - tracé autoroutier,
 - gestion de containers [Tamaki96], ...
- Finances, Productique, Robotique, Mécanique...

Exemple

- **Design de réseaux de radiocommunications mobiles** [Thèse H. Meunier, OPAC/FT R&D]
- **Objectifs et/ou contraintes:**
 - Min (Nombre de sites candidats utilisés)
 - Min (Interférence)
 - Min (Overhead)
 - Max (Trafic)
 - Couverture, Handover, Connectivité, ...

Définitions

- Une **action (ou un vecteur de décisions)** sera notée $x = (x_1, x_2, \dots, x_n)$ avec x_i les variables du problème et n le nombre de variables
- Les contraintes seront notées $g_i(x)$ avec $i = 1, \dots, m$ et m le nombre de contraintes.
- Le **vecteur de fonctions objectifs** sera noté $f : f(x) = (f_1(x), f_2(x), \dots, f_k(x))$ avec f_i les objectifs (ou critères de décision) et k le nombre d'objectifs.
- On recherche l'action x^* telle que les contraintes $g_i(x^*)$ soient satisfaites pour $i = 1, \dots, m$ et qui optimise la fonction $f : f(x^*) = (f_1(x^*), f_2(x^*), \dots, f_k(x^*))$

Optimisation multi-objectif

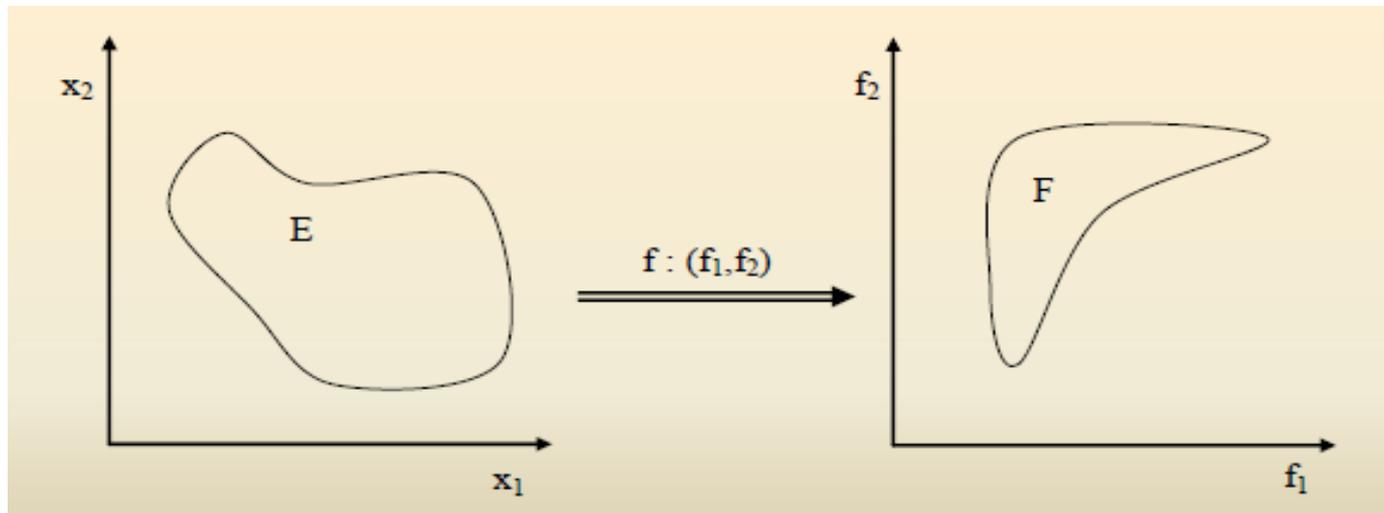
D'une façon générale, un problème d'optimisation multi – objectifs est exprimé par l'équation :

$$\left\{ \begin{array}{ll} \min F(x) = (f_1(x), f_2(x), \dots, f_n(x)), & x \in E = \mathbb{R}^m \\ g_i(x) \leq 0, & i = 1, \dots, p \\ h_j(x) = 0, & j = 1, \dots, q \\ x_{k \min} \leq x_k \leq x_{k \max}, & k = 1, \dots, m \end{array} \right.$$

où $n > 2$ est le nombre de fonctions objectif et $F(x) = (f_1(x), \dots, f_n(x))$ est le vecteur des fonctions à optimiser. Le vecteur $x = (x_1, \dots, x_m)$ est le vecteur des variables de décision. $(g_i(x) \leq 0, h_j(x) = 0, x_{k \min} \leq x_k \leq x_{k \max})$ sont respectivement les contraintes d'inégalités, les contraintes d'égalités et les contraintes de domaine.

Définitions

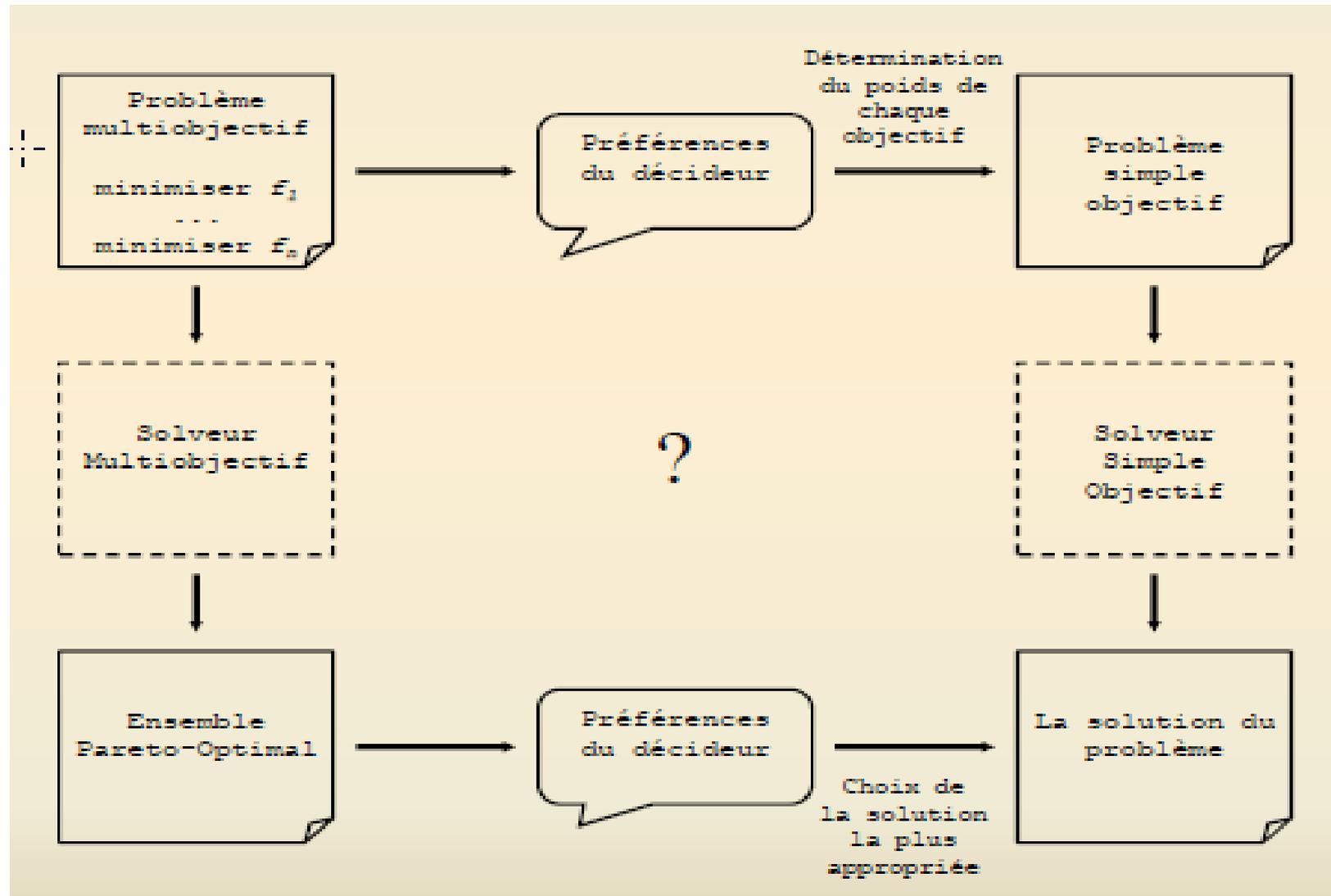
- L'ensemble des domaines de définition de chaque variable et les contraintes forment un ensemble E que nous appellerons **l'ensemble des actions réalisables**.
- **Nous appellerons F l'ensemble des objectifs réalisables.**



Problématiques

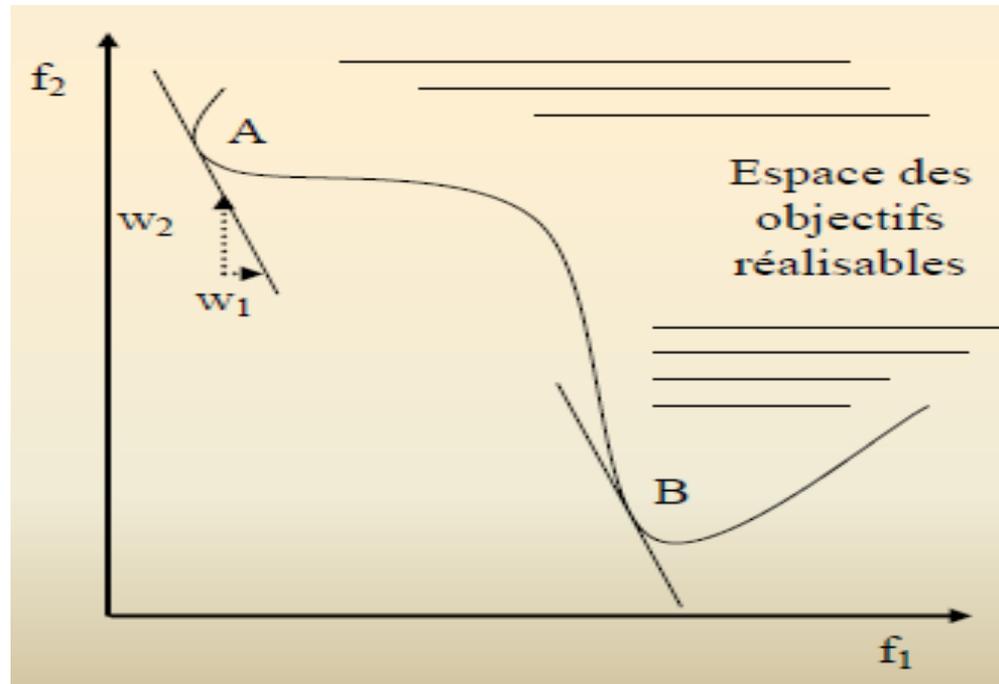
- Difficulté à établir une définition de l'optimum
Le décideur peut simplement exprimer le fait qu'une solution est préférable à une autre mais il n'existe pas une solution meilleure que toutes les autres.
- Recherche d'un ensemble de solutions satisfaisantes
Cet ensemble doit être représentatif.

Problématiques



Problématiques

- Minimisation de f_1 et f_2
- Si le décideur opte pour une méthode agrégée avec w_1 et w_2 les poids des objectifs
- Toutes les solutions sur la portion entre A et B de courbe peuvent satisfaire le décideur



Les méthodes basées sur Pareto

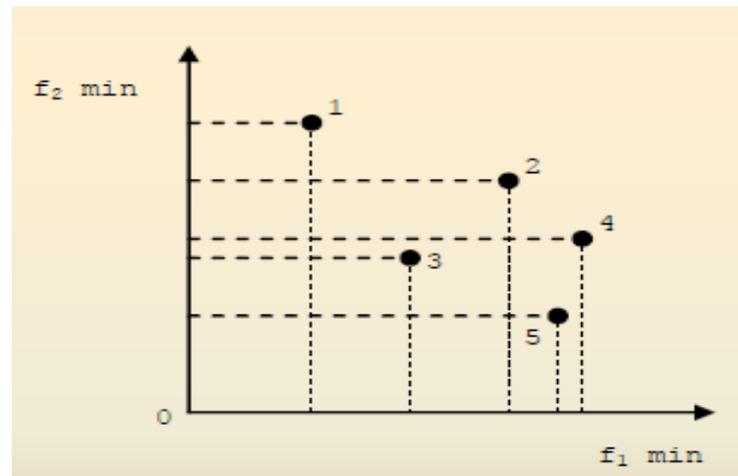
- Vilfredo Pareto : « Il existe un équilibre tel que l'on ne peut pas améliorer un critère sans détériorer au moins un des autres critères »

- La notion de dominance

$$x_1 \succ x_2 \Leftrightarrow f_i(x_1) \geq f_i(x_2) \quad \forall i \in \{1, \dots, k\} \wedge \exists j \in \{1, \dots, k\} \Leftrightarrow f_j(x_1) > f_j(x_2)$$

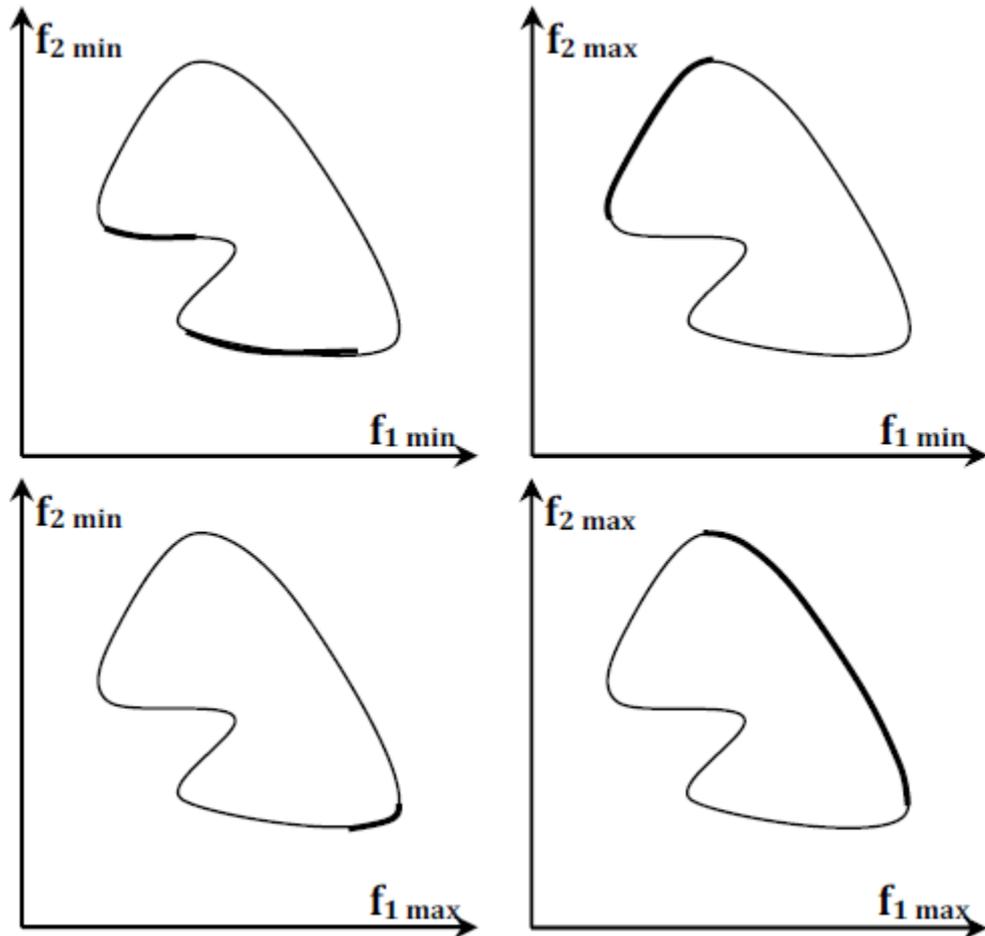
- Optimalité au sens de Pareto

$x^* \in \Omega \subset \mathcal{R}^n$ est optimal au sens de Pareto, ou efficace, ou non-dominé, ou non-inférieur s'il n'existe pas de $x \in \Omega$ tel que $x \succ x^*$.



La frontière de Pareto

- La frontière de Pareto est l'ensemble de tous les points Pareto-Optimum



Classification des méthodes

- Décideur
 - Les méthodes a priori (décideur \rightarrow recherche)
 - Les méthodes a posteriori (recherche \rightarrow décideur)
 - Les méthodes progressives ou interactives (décideur \leftrightarrow recherche)

Classification des méthodes

- Concepteur

- Transformation de PMO en uni-objectif (agrégation)

- Méthode par moyenne pondérée
 - Méthode du but programmé [Charmes 1961]
 - Méthode par but à atteindre [Chen et Liu 1994]
 - Le min-max [Coello 1995]
 - Méthode ε -contrainte [Ritzel 1994]

- Approches Pareto

- Utilisent la notion de dominance

- Approches non Pareto (Traitent séparément les différents objectifs)

- Sélection parallèle (VGA)
 - Utilisation des genres
 - La méthode lexicographique

Les méthodes agrégées

Axiome fondamental

Tout décideur essaye inconsciemment de maximiser une fonction d'utilité.

Les méthodes agrégées

Somme pondérée

Cette méthode consiste à additionner tous les objectifs en affectant un coefficient de poids. Ce coefficient représente l'importance relative que le décideur attribue à l'objectif. Le problème devient ainsi un problème mono-objectif

$$\min \sum_{i=1}^k \lambda_i f_i(x) \quad \text{avec } \lambda_i \geq 0$$

Où λ_i représente le poids affecté à chacun des objectifs i et $\sum_{i=1}^k \lambda_i = 1$

Les méthodes agrégées

Méthode du but programmé, appelée aussi **target vector optimization**

Le décideur fixe **un but Z_j** à atteindre pour chaque objectif f_j .

Ces valeurs sont ajoutées au problème comme des contraintes supplémentaires. La nouvelle fonction objectif est modifiée de façon à **minimiser la somme des écarts (norme)** entre les résultats et les buts à atteindre. On peut pondérer cette norme par des poids λ_j

$$\min \sum_{j=1}^k \lambda_j |f_j(x) - Z_j| \quad \text{avec } x \in F$$

Les méthodes agrégées

Méthode par but à atteindre

Dans cette approche le décideur spécifie l'ensemble des buts Z_j qu'il souhaite atteindre et les poids associés λ_j . La solution optimale est trouvée en résolvant le problème suivant :

minimiser α tel que $Z_j + \alpha \cdot \lambda_j \geq f_j(x)$ avec $\sum_{j=1}^k \lambda_j = 1$

Les objectifs Z_j représentent le point de départ de la recherche dans l'espace de définition, et les poids w_i indiquent la direction de la recherche dans l'espace

Les méthodes agrégées

Le min-max

Elle minimise le maximum de l'écart relatif entre un objectif et son but associé par le décideur.

$$\min \max \left(\frac{f_i(x) - Z_i}{Z_i} \right) \quad \text{avec } i = 1, \dots, k$$

Z_i représente la valeur à atteindre pour le $i^{\text{ème}}$ objectif

Les méthodes agrégées

La méthode e-contrainte

Cette méthode est basée sur la minimisation d'un objectif f_i en considérant les autres objectifs contraints par une valeur ε_j .

Résolution par les métaheuristiques

Méthode multi-objectif → Méthode mono-objectif

Technique de résolution

- Recuit simulé SA
- Recherche tabou TS
- Algorithme génétique GA
- Algorithme colonie de fourmis ACO
- Algorithme par essaim particulaire PSO

La somme pondérée

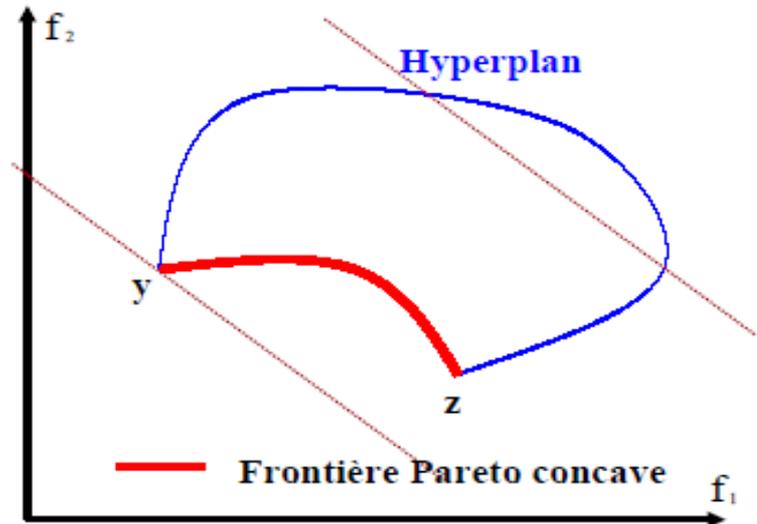
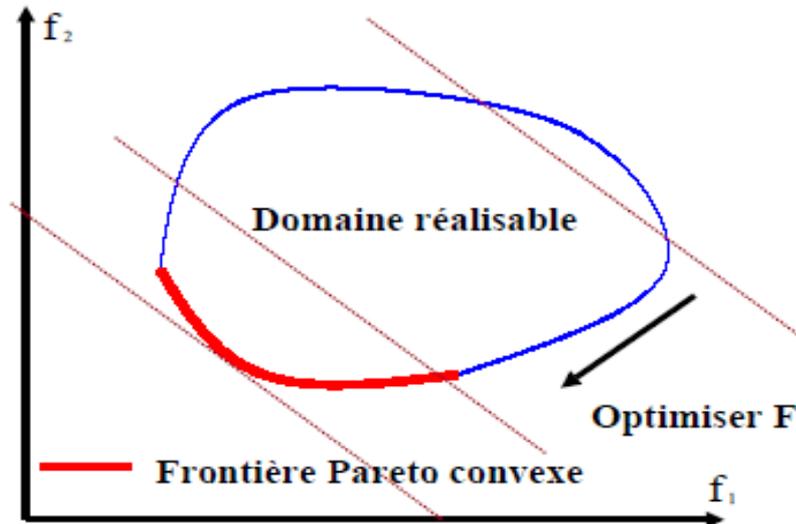
Cette méthode consiste à additionner tous les objectifs en affectant un coefficient de poids. Ce coefficient représente l'importance relative que le décideur attribue à l'objectif.

$$(PMO_{\lambda}) \begin{cases} \min F(x) = \sum_{i=1}^n \lambda_i f_i(x) & \text{Agrégation linéaire} \\ \text{s.c. } x \in C \quad \lambda_i \geq 0, i = 1, \dots, n \quad \sum_{i=1}^n \lambda_i = 1 \end{cases}$$

Complexité = problème combinatoire sous-jacent.

Ex : Polynomial = flot, cheminement, ...

NP-complet = routage, affectation, ...



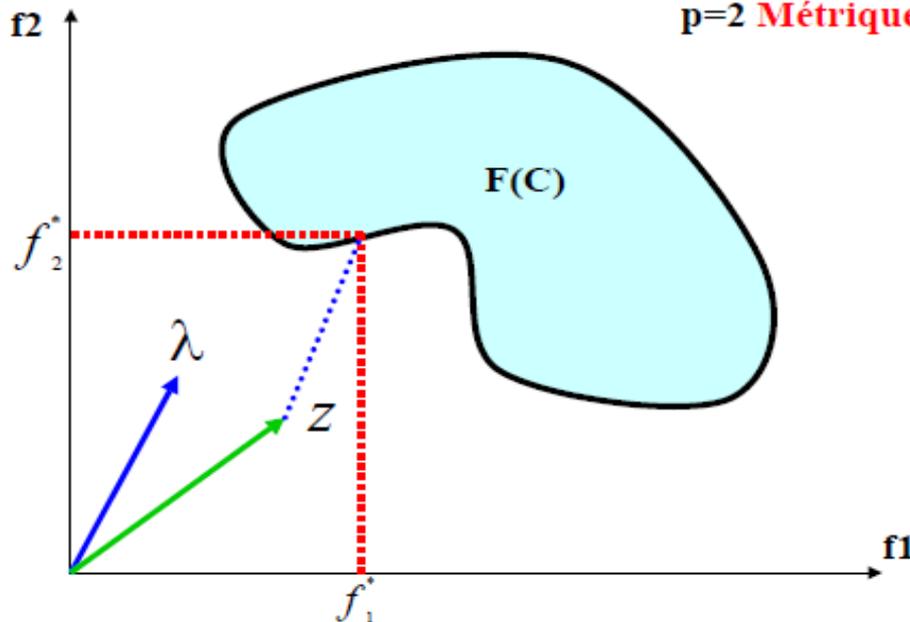
Métaheuristiques

- **Algorithmes génétiques** [Hajela et Lin 92]
 - Codage d'un individu = Solution + λ
 - But = Générer une variété de solutions Pareto
- **Recuit simulé** [Serafini 92]
 - Fonction d'acceptation
$$p_{xy}(T) = \min\left(1, e^{\frac{\sum_{i=1}^n \lambda_i (f_i(x) - f_i(y))}{T}}\right)$$
- **Recherche tabou** [Dahl et al. 95]
 - Poids = priorité de l'objectif
- **Métaheuristiques hybrides** [Talbi 98]
 - ◆ **Algorithme glouton + Recuit simulé** [Tuyttens 98]
 - ◆ **Algorithme génétique(Recherche locale)** [Ishibuchi et Murata 98]
 - Sélection avec des poids différents
 - Recherche locale sur l'individu produit (même poids)

Programmation par but programmé

$$(PMO_k(\varepsilon)) \left\{ \begin{array}{l} \min \left(\sum_{j=1}^n \lambda_j |f_j(x) - z_j|^p \right)^{\frac{1}{p}} \\ \text{s.c. } x \in C \end{array} \right. \quad \mathbf{Z} : \text{Vecteur idéal ou de référence}$$

Norme utilisée : **Métrique de Tchebycheff** L_p métrique
 $p=2$ **Métrique euclidienne** $p=\infty$ **Min-Max**



Métaheuristique

- Algorithmes génétiques

- Fonction min-max, AG parallèle avec différents poids [Coello 98]
- Règles floues dans l'évaluation de F [Reardon 98]

- Recuit simulé [Serafini 92]

- Fonction d'acceptation $p_{xy}(T) = \min\left(1, e^{-\frac{\max_i(\lambda_i(f_i(x)-z_i)) - \max_i(\lambda_i(f_i(y)-z_i))}{T}}\right)$
- Itération $\Rightarrow \lambda_i = \lambda_i \pm [-0.05, +0.05]$

- Recherche tabou [Gandibleux 96]

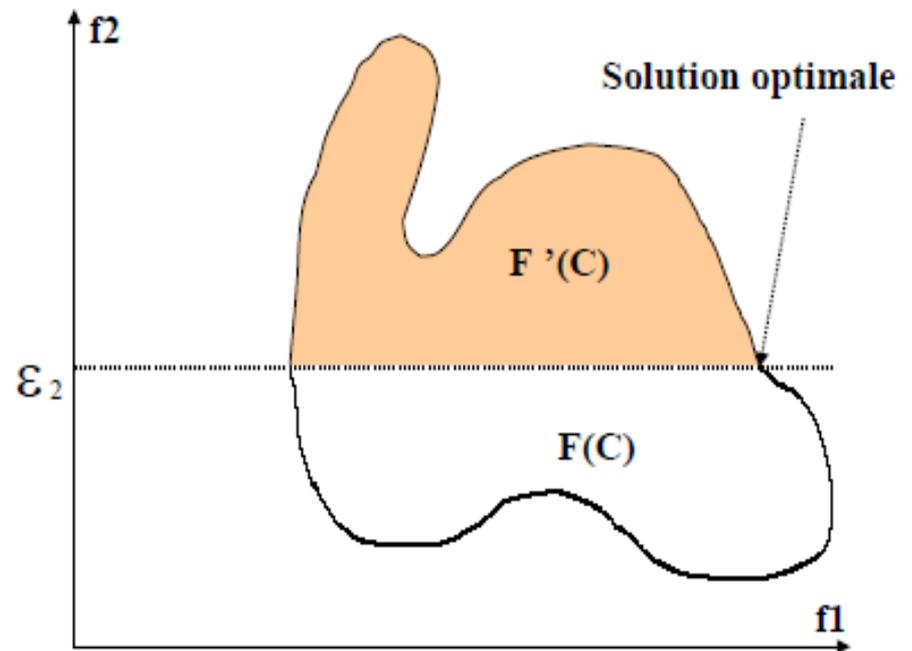
- Meilleur voisin = Meilleur compromis non tabou
- Compromis = Norme L_p de Tchebycheff par rapport au vecteur idéal
- Mémorisation des M meilleures solutions

La méthode e-contrainte

$$(PMO_k(\varepsilon)) \begin{cases} \min f_k(x) \\ \text{s.c. } x \in C \\ f_j(x) \leq \varepsilon_j, j = 1, \dots, n, j \neq k \end{cases} \quad \varepsilon = (\varepsilon_1, \dots, \varepsilon_{k+1}, \dots, \varepsilon_n)$$

- Variation de ε
- $F(C)$ = espace objectifs PMO
- $F'(C)$ = espace objectifs du problème transformé

$$\begin{cases} \max f_1 \\ f_2 \geq \varepsilon_2 \end{cases}$$



Métaheuristiques

- Algorithmes génétiques

- Individu = ε [Loughlin 98]

$$\varepsilon_i = \varepsilon_{\min} + \frac{(i-1)(\varepsilon_{\max} - \varepsilon_{\min})}{(k-1)}$$

k : taille de la population

- Recherche tabou [Hertz et al. 94]

- Suite de sous-problèmes
- Ordre de priorité
- Seuil (f') = Optimum (f^*)
- + Déterioration acceptée

$$(PMO_q) \left\{ \begin{array}{l} f^*_q = \min f_q(x) \\ s.c. f_r(x) \leq f'_r, r = 1, \dots, q-1 \\ x \in C \end{array} \right.$$

- Métaheuristiques hybrides [Quagliarella et al. 97]

- Algorithme génétique + Recherche locale

Les difficultés

- Déterminer les différents paramètres
 - Connaissance à priori
 - Tests pour déterminer l'influence de chaque objectif
- Exprimer l'interaction entre critères
 - Corrélation, interchangeabilité / complémentarité, dépendance préférentielle
 - Intégrale de Choquet ou de Sugeno [Marichal 1999]

Analyse critique

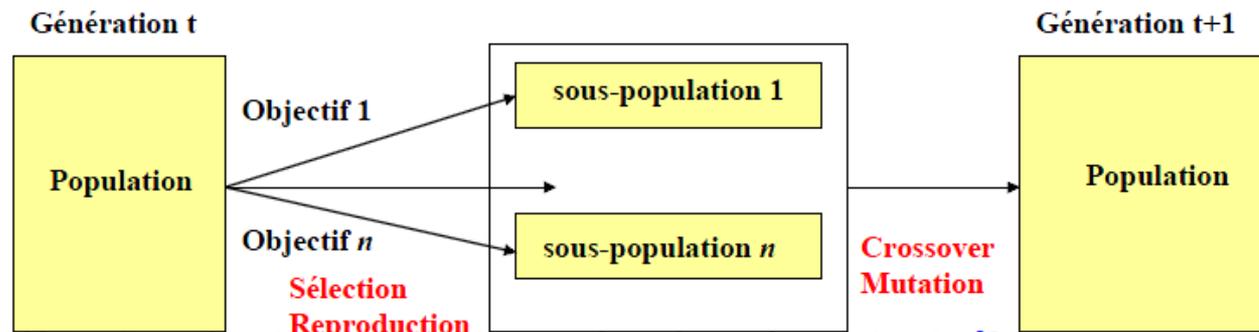
- Espace de recherche \leftrightarrow Problème initial. Le problème perd ses éventuelles propriétés.
- Une exécution produit **une seule solution**.
- **Connaissances a priori** (détermination des paramètres).
- Sensibles au **paysage** de la frontière Pareto (convexité, discontinuité, ...), et à différents **paramètres** (poids, contraintes, buts, ...).
- Objectifs **bruités** ou données **incertaines**.
- **Coût** associé aux algorithmes (exécution multiple).

Les méthodes non agrégées et non Pareto

- Principe
Elles possèdent un processus de recherche qui traite séparément les objectifs.
- Techniques
 - Vector Evaluated Genetic Algorithm (VEGA) [Schaffer 1985]
 - Utilisation de genres [Allenson 1992, Lis et Eiben 1996]
 - La méthode lexicographique [Fourman 1985]
- Les objectifs sont classés par ordre de préférence
 - A Non Generational Genetic Algorithm [Valenzuela et Uresti 1997]
- La fitness est calculée de façon incrémentale (domination et voisinage) : réduisent le problème à deux objectifs

Approche évolutionnaire non Pareto

- Sélection parallèle (VEGA) [Schaffer 85]

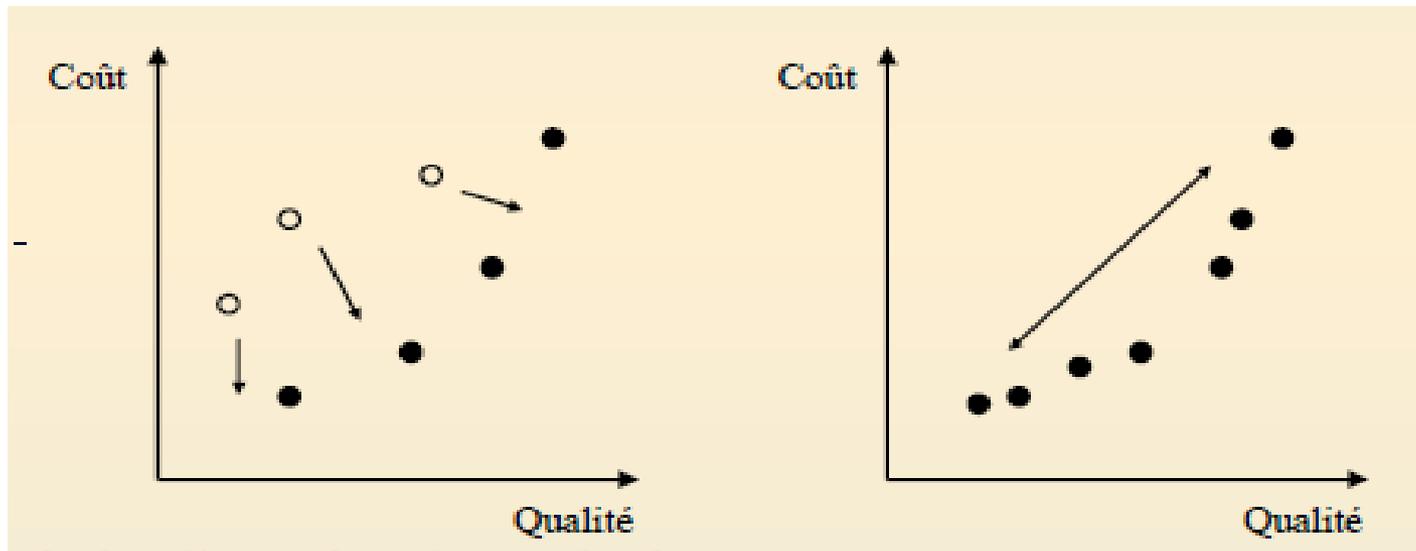


- Sélection lexicographique (ordre de priorité)
 - Algorithmes génétiques [Fourman 85]
 - Stratégies évolutionnistes [Kursawe 91]
- Reproduction multi-sexuelle [Lis et Eiben 96]
 - ◆ Plusieurs classes. Une classe = Un objectif
 - ◆ Reproduction (crossover) sur plusieurs individus

Tendance à ignorer les solutions compromis

Les méthodes basées sur Pareto

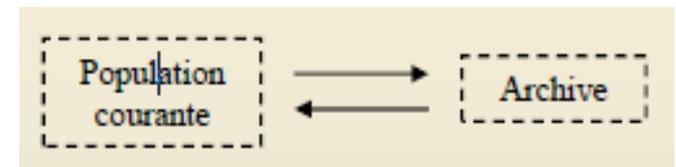
- Objectifs
- Guider le processus de recherche vers la frontière de Pareto



- Maintenir la diversité des solutions
- Assurer une bonne répartition des solutions sur la frontière

Les méthodes basées sur Pareto

- Les méthodes non élitistes
 - Ne conservent pas les individus Pareto-optimaux
 - Maintiennent difficilement la diversité sur la frontière de Pareto
 - La convergence vers la frontière de Pareto est lente
- Les méthodes élitistes
 - Utilisent une population externe (archive)
 - Procédures de mise à jour et de réinjection
 - Utilisent des techniques de
 - Formation de niches
 - Partitionnement en cluster
 - Maillage
 - Préfèrent les solutions non dominées

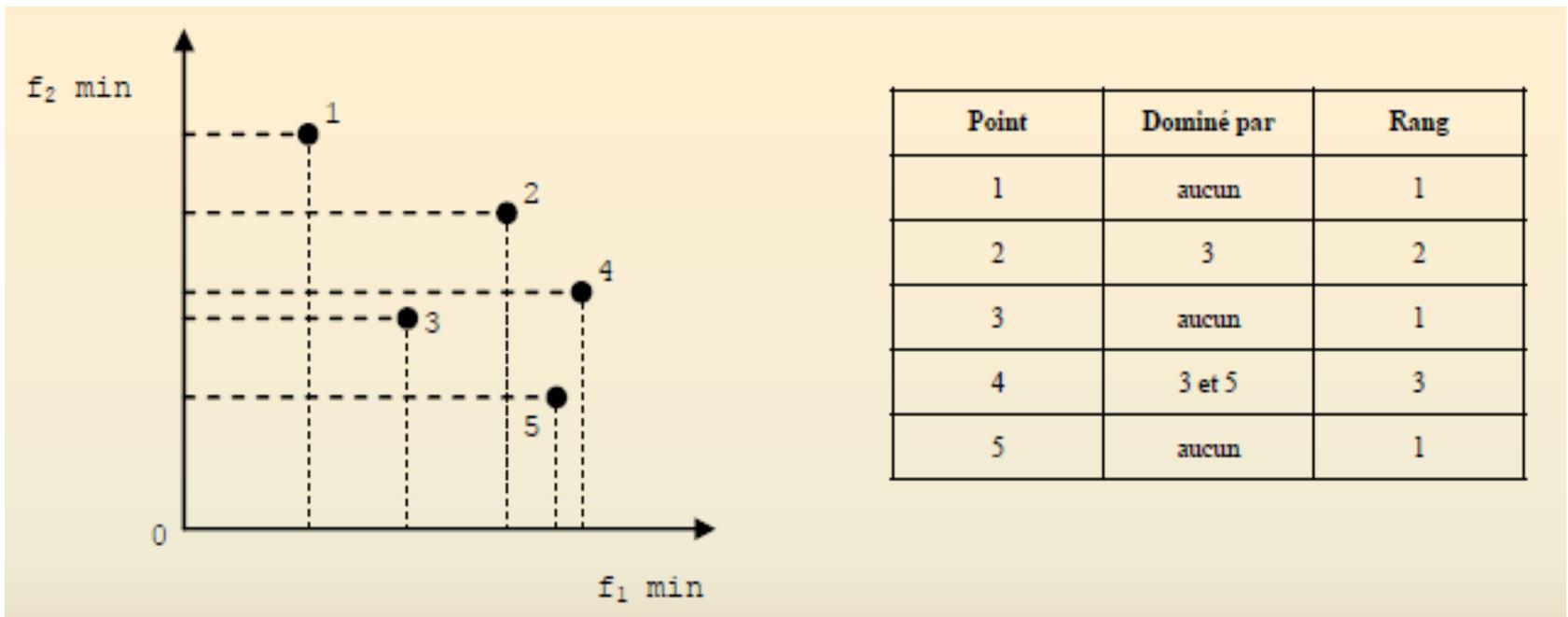


Les méthodes basées sur Pareto

- Les méthodes non élitistes
 - Multiple Objective Genetic Algorithm (MOGA) [Fonseca et Fleming 1993]
 - **Notation, calcul du rang de dominance + interpolation**
 - Non dominated Sorting Genetic Algorithm (NSGA) [Deb 1993]
 - **Notation, tri par rapport aux différents degrés de dominance**
 - Niche Pareto Genetic Algorithm (NPGA) [Horn et Nafpliotis 1993]
 - **Sélection, tournoi de dominance**

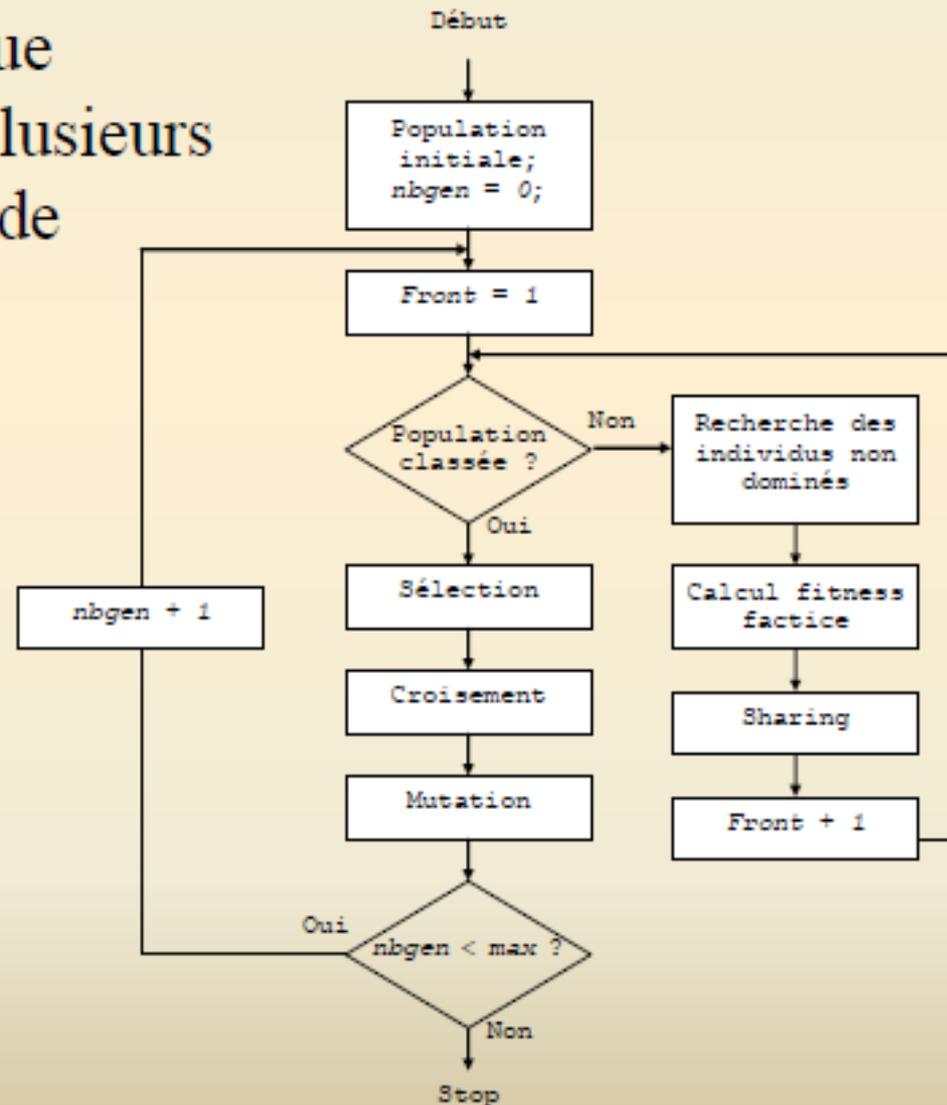
MOGA

- Chaque individu de la population est rangé en fonction du nombre d'individus qui le dominant



NSGA

- Le calcul de la fitness s'effectue en séparant la population en plusieurs groupes en fonction du degré de domination au sens de Pareto de chaque individu



- Critiques

- ✓ Complexe
- ✓ Utilisation du sharing
- ✓ Non élitiste

Les méthodes basées sur Pareto

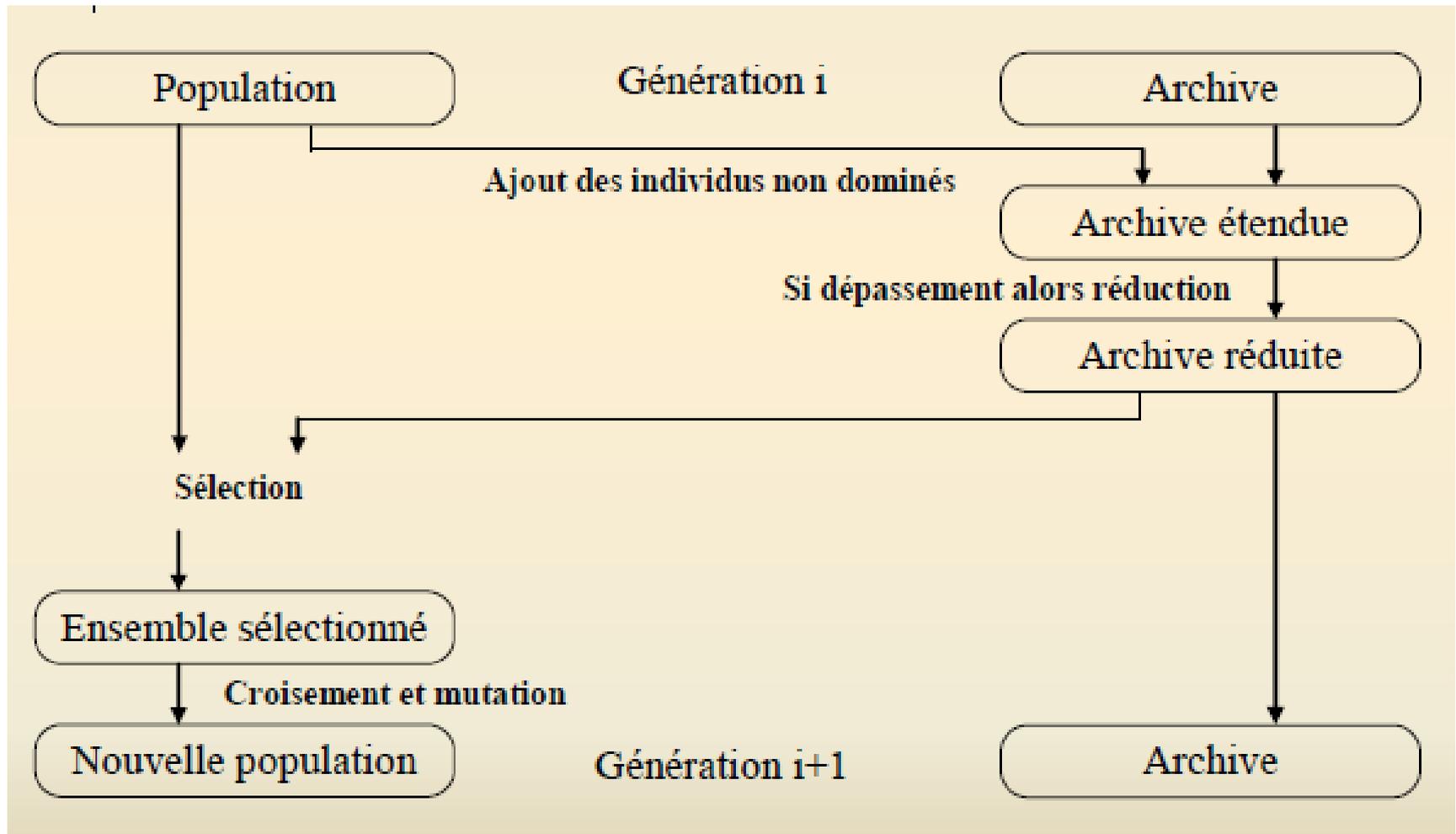
- Les méthodes non élitistes
 - Ne conservent pas les individus Pareto-optimaux
 - Maintiennent difficilement la diversité sur la frontière de Pareto
 - La convergence vers la frontière de Pareto est lente
- Les méthodes élitistes
 - Utilisent une population externe (archive)
 - Procédures de mise à jour et de réinjection
 - Utilisent des techniques de
 - Formation de niches
 - Partitionnement en cluster
 - Maillage
 - Préfèrent les solutions non dominées



Les méthodes basées sur Pareto

- Les méthodes élitistes
 - Strength Pareto Evolutionary Algorithm (SPEA) [Zitzler et Thiele 1998]
 - **Notation relative au nombre d'individus dominés**
 - Pareto Archived Evolution Strategy (PAES) [Knowles et Corne 1999]
 - Stratégie d'évolution (1+1)
 - Pareto Envelope based Selection Algorithm (PESA) [Knowles et Corne 2000]
 - **Sélection, tournoi en fonction de la mesure d'encombrement**
 - NSGA II [Deb 2000]
 - PESA II

Schéma de fonctionnement



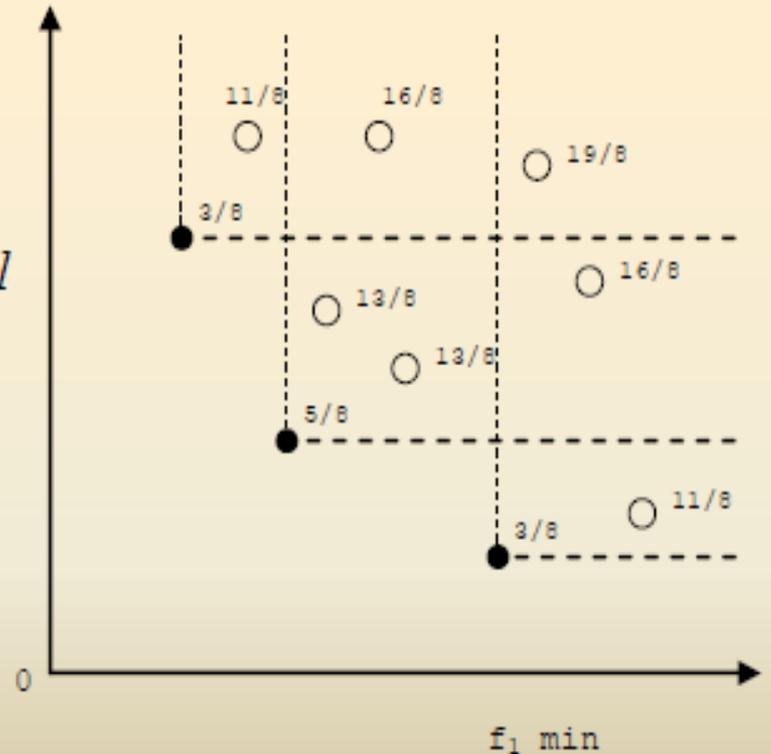
SPEA

- Notation

- ✓ Pour chaque individu de l'ensemble Pareto-optimal. La fitness est proportionnelle au nombre d'individus de la population qu'il domine

- ✓ La fitness f_i d'un individu de la population est égale à la somme des fitness des individus Pareto-optimaux qui le dominent

- ✓ Dépendance par rapport à la taille de l'archive

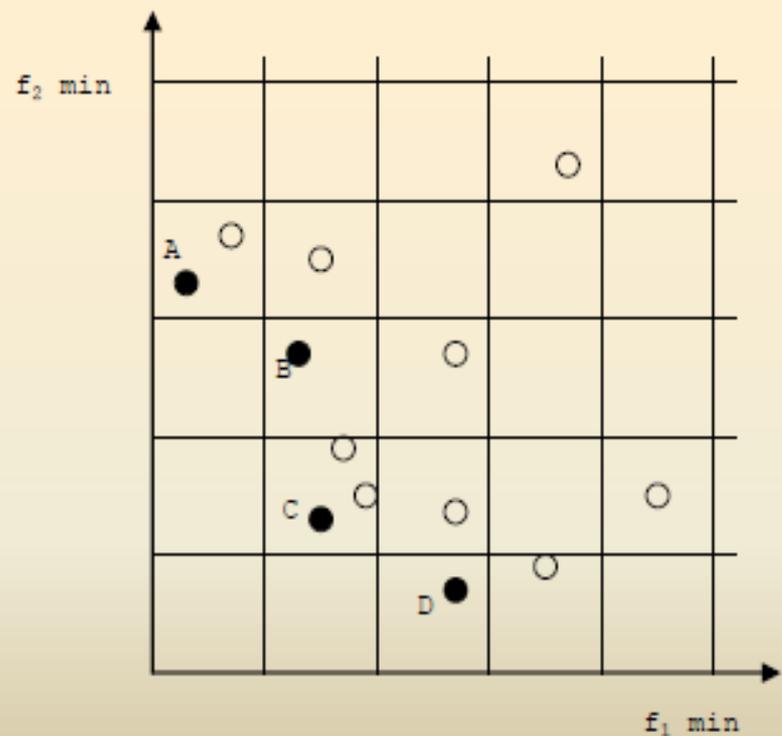


PAES

- Non basée sur une population, elle n'utilise qu'un seul individu à la fois pour la recherche des solutions.
- L'algorithme utilisé est très simple et inspiré d'une stratégie d'évolution ($1+1$) [*Rechenberg 1973*].
- Elle utilise une technique de *crowding* basée sur un découpage en **hypercubes de l'espace des objectifs**.

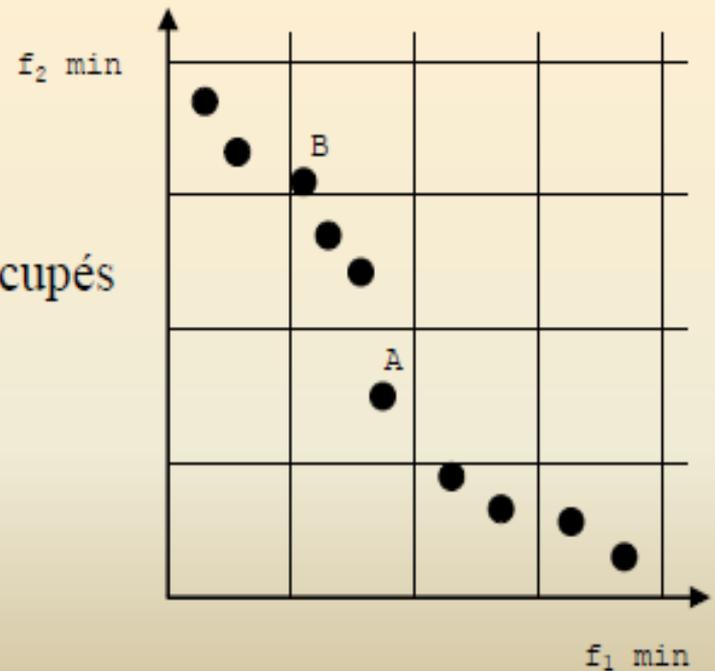
PESA

- Basée sur les AGs
- Reprend le principe de crowding développé dans PAES
 - ✓ Mesure d'encombrement d'une zone de l'espace
 - ✓ Utilisée lors de la sélection et de la mise à jour de l'archive



PESA II

- Sélection par rapport aux hypercubes occupés par au moins un individu.
 - ✓ Après sélection de l'hypercube, on choisit aléatoirement l'individu dans l'hypercube.
 - ✓ Cette méthode se montre plus efficace à répartir les solutions sur la frontière de Pareto
- ✓ Tournoi binaire, 10 points, 6 cubes occupés
 - PESA : $1 - (9/10)^2 = 0,19$
 - PESA II : $1 - (5/6)^2 = 0,31$



Synthèse

- Les méthodes agrégées
 - Nécessitent une connaissance à priori du problème
- Les méthodes non agrégées
 - Paramétrage difficile
 - Algorithmes génétiques
 - Archive
 - Heuristique
 - Interprétation des paramètres
 - Définition d'un critère d'arrêt

Discussion

- Nécessité de maîtriser la technique utilisée
 - Quel génotype, notation, fct de sélection ... ?
- Grand nombre de paramètres
 - Relation entre la valeur d'un paramètre et son action sur la résolution du problème ?
 - Comment arriver à expliquer au décideur (utilisateur) le sens de ces paramètres ?
 - Serait-il possible de mesurer la prise de risque pour le décideur ?
 - B est plus sûr que A !

