

Métodos de escalarización en optimización multiobjetivo

Rocío Guadalupe Fonseca

Director: Elvio Angel Pilotta

Trabajo especial de
Licenciatura en matemática



Facultad de Matemática, Astronomía, Física y Computación
Universidad Nacional de Córdoba
Argentina

Esta obra está bajo una licencia Creative Commons
“Reconocimiento-NoCommercial-CompartirIgual
4.0 Internacional”.



Métodos de escalarización en optimización multiobjetivo

Fonseca Rocío Guadalupe

Resumen

La vida inevitablemente involucra la toma de decisiones y elecciones y es natural querer que estas sean las mejores posibles, en otras palabras, sean óptimas. La dificultad aquí radica en el conflicto (al menos parcial) entre nuestros diversos objetivos y metas. Los problemas con múltiples objetivos y criterios son generalmente conocidos como problemas de optimización multiobjetivo. A lo largo de este trabajo, se presentaron los conceptos necesarios y algunos métodos para resolver problemas de optimización multiobjetivo. Resolver un problema de optimización multiobjetivo significa encontrar el conjunto de soluciones Pareto optimal o frente de Pareto. Los métodos fueron divididos en cuatro categorías según la articulación de preferencias de un tomador de decisiones. De cada método se estudiaron las ventajas y desventajas y se seleccionaron tres métodos para estudiar con mayor profundidad. Los métodos seleccionados fueron métodos de escalarización; el método de sumas ponderadas, restricción ε y métricas ponderadas, que además fueron implementados para generar una aproximación del frente de Pareto. Se seleccionaron problemas test para generar aproximaciones de sus frentes de Pareto y analizar los resultados obtenidos. Se encontró que ningún método es superior que otro. La selección de un método específico depende del tipo de información que proporciona el problema, las preferencias del usuario, los requisitos de la solución y la capacidad de cómputo.

Palabras Claves:

- Optimización multiobjetivo
- Frente de Pareto
- Métodos de escalarización
- Método de sumas ponderadas
- Método de restricción ε
- Método de métricas ponderadas

Métodos de escalarización en optimización multiobjetivo

Fonseca Rocío Guadalupe

Abstract

Life inevitably involves decision making and choices and it is natural to want them to be the best possible, in other words, to be optimal. The difficulty here lies in the (at least partial) conflict between our various goals and objectives. Problems with multiple objectives and criteria are generally known as multiobjective optimization problems. Throughout this work, the necessary concepts and some methods to solve multiobjective optimization problems were presented. Solving a multiobjective optimization problem means finding the optimal Pareto solution set or Pareto front. The methods were divided into four categories according to the articulation of preferences of a decision maker. The advantages and disadvantages of each method were studied, and three methods were selected for further study. The selected methods were scalarization methods; the method of weighted sums, ε -constraint, and weighted metrics, which were also implemented to generate an approximation of the Pareto front. Test problems were selected to generate approximations of their Pareto fronts and to analyze the results obtained. It was found that no one method is superior to another. Selecting a specific method depends on the type of information the problem provides, the user preferences, the solution requirements, and the computational capacity.

Key words:

- multiobjective optimization
- Pareto front
- Scalarization methods
- Weighted sum method
- ε -constraint method
- Weighted metrics method

Índice general

1. Introducción	5
2. Resultados preliminares de optimización multiobjetivo	7
2.1. Algunos ejemplos y aplicaciones	7
2.2. Problema de optimización multiobjetivo	9
2.3. Optimización de Pareto	10
2.3.1. Vector objetivo ideal	13
2.3.2. Vector objetivo nadir	14
2.4. Optimalidad débil de Pareto	15
3. Métodos numéricos	22
3.1. Métodos de no preferencia	24
3.1.1. Método del criterio global	24
3.1.2. Método de haces proximal multiobjetivo (MPB)	26
3.2. Métodos a posteriori	27
3.2.1. Suma ponderada	28
3.2.2. Restricción ε	28
3.2.3. Restricción elástica	29
3.2.4. Métricas ponderadas	29
3.2.5. Programación por metas (Goal programming)	30
3.3. Métodos a priori	31
3.3.1. Método de función de valor	31
3.3.2. Ordenamiento lexicográfico	33
3.4. Métodos interactivos	35
3.4.1. Método de compensación interactiva de valor sustituto (ISWT)	37
3.4.2. Método de punto de referencia	39
4. Métodos de escalarización	42
4.1. Sumas ponderadas	43
4.2. Restricción ε	45
4.3. Métricas ponderadas	51
4.4. Escalarización por combinaciones convexas	55

5. Resultados numéricos	59
5.1. Ejemplo 4	59
5.2. Problema de Schaffer	64
5.3. Problema de Hazen	77
5.4. Problema de Fonseca - Fleming	82
5.5. Problema de Poloni	87
6. Conclusiones	92

Capítulo 1

Introducción

En cada momento, quizás sin pensarlo, estamos forzados a tomar decisiones que involucran varios factores, que pueden ser contradictorios entre sí, entonces necesitamos poder decidir basándonos en algún criterio, después de todo siempre buscamos la mejor solución, pero ¿cuál es la mejor solución cuando se tienen tantos factores a considerar? Esta pregunta no sólo aparece en nuestra vida cotidiana, sino que es recurrente en las ciencias por eso para poder conseguir soluciones óptimas bajo cada uno de los criterios considerados se ha desarrollado la *optimización multicriterio* o *multiobjetivo*.

En este trabajo se estudió la teoría de optimización multicriterio, la cual consiste en encontrar soluciones óptimas a problemas en el que la función objetivo es una función vectorial. El objetivo es estudiar métodos de optimización para resolver problemas multicriterio. Además, se implementarán algunos métodos de optimización multicriterio para resolver algunos problemas test.

Este trabajo se estructura de la siguiente manera:

En el Capítulo 2 consideramos algunos ejemplos motivadores y presentamos los conceptos necesarios para desarrollar la teoría de la optimización multicriterio. Además, probamos algunos resultados básicos y definimos la notación que utilizaremos en todo el trabajo.

En el Capítulo 3 presentamos varios métodos para resolver problemas de optimización multicriterio clasificándolos según la intervención de un tomador de decisiones que nos ha presentado el problema que deseamos resolver siendo los analistas, los que efectivamente resolverán el problema y le presentarán las soluciones de una manera conveniente. Además, realizamos comparaciones entre los métodos y elegimos algunos para estudiar en detalle.

En el Capítulo 4 desarrollamos con mayor profundidad tres de los métodos que elegimos en el Capítulo 3 para emplearlos para resolver nuestros proble-

mas test. Además, presentamos un nuevo método de escalarización planteado por Victor Pereyra [11] que también vamos a aplicar en nuestros problemas test.

En el Capítulo 5 mostramos cuatro problemas test, cada uno tiene un atractivo particular y nos permite explorar algún aspecto característico de los métodos estudiados anteriormente. Además, realizamos experimentos numéricos con cada problema.

Finalmente, en el Capítulo 6 presentamos nuestras conclusiones sobre la efectividad de los distintos métodos estudiados, sus ventajas y algunos aspectos en los que se podría seguir trabajando en un futuro.

Capítulo 2

Resultados preliminares de optimización multiobjetivo

2.1. Algunos ejemplos y aplicaciones

Consideremos los siguientes ejemplos:

Ejemplo 1: Si se quiere responder a la pregunta ¿Cuál es el animal que nada, corre y vuela más rápidamente simultáneamente? se puede pensar antes en estas preguntas que ayudarán a responder la pregunta inicial:

- ¿Cuál es el animal más veloz en la tierra?
- ¿Cuál es el animal más veloz volando?
- ¿Cuál es el animal más veloz nadando?

Las respuestas a estas preguntas son el **guepardo** o **chita** es el animal más veloz sobre la tierra, el **halcón peregrino** es el ave más veloz y el **pez espada** es el pez más rápido del mundo. Pero si se busca un animal que cumpla las tres condiciones al mismo tiempo, el animal que puede hacer las tres cosas es **pato** que es capaz de correr, aunque es menos veloz que un guepardo, es capaz de volar, aunque no tan rápido como un halcón peregrino y es capaz de nadar, aunque no con la habilidad de un pez espada.

Ejemplo 2: Queremos comprar un auto nuevo y encontramos cuatro modelos entre los cuales nos gustaría elegir: un Volkswagen Golf, un Opel Astra, un Ford Focus y un Toyota Corolla. La decisión va a depender del precio del auto, el consumo de combustible y la potencia. Buscamos el auto más barato, con más potencia y que consuma menos combustible. En este caso, nos enfrentamos a un problema con tres funciones criterio y cuatro opciones posibles. Las características de los cuatro autos se muestran en la tabla (2.1) a continuación.

		Alternativas			
		VW	Opel	Ford	Toyota
Criterio	Precio (1000 Euros)	16.2	14.2	14.0	15.2
	Consumo ($\frac{l}{100km}$)	7.2	7.0	7.5	8.2
	Potencia (kW)	66.0	62.0	55.0	71.0

Tabla 2.1: Criterios y alternativas del Ejemplo 2

Cómo decidir cuál de los cuatro autos es el “mejor” si el más potente es al mismo tiempo el que más combustible consume, entonces no podemos conseguir un auto que sea el más barato, el más potente y el que menos combustible consume al mismo tiempo. Sin embargo, podemos observar que si sólo miramos uno de los criterios la elección es simple.

Ejemplo 3: Para la construcción de una represa de agua un proveedor de electricidad está interesado en maximizar la capacidad de almacenamiento y al mismo tiempo minimizar la pérdida de agua debido a la evaporación y el costo de construcción. Se debe tomar una decisión sobre los meses de mano de obra humana utilizada para la construcción, así como la medida del radio del lago, y también debe respetar ciertas limitaciones, tales como la fuerza mínima de la represa. Aquí el conjunto de alternativas (posibles diseños de la represa) permite infinitas opciones diferentes. Los criterios son funciones de las variables de decisión a maximizar o minimizar y están claramente en conflicto: una presa con una gran capacidad de almacenamiento ciertamente no implicará un pequeño costo de construcción.

Ejemplo 4: Consideremos el problema matemático con dos criterios y una variable de decisión. Las funciones objetivo o criterios, que queremos minimizar simultáneamente sobre la recta real no negativa son:

$$f_1(x) = \sqrt{x+1} \quad f_2(x) = x^2 - 4x + 5 = (x-2)^2 + 1 \quad (2.1)$$

Queremos resolver el problema de optimización

$$\min_{x \geq 0} (f_1(x), f_2(x)) \quad (2.2)$$

La pregunta es, ¿cuáles son los mínimos y los minimizadores en este problema? Notemos que, de nuevo, para cada función individualmente, el correspondiente problema de optimización es fácil: $x_1 = 0$ y $x_2 = 2$ son los minimizadores (únicos) de f_1 y f_2 sobre $x \in \mathbb{R} : x \geq 0$, respectivamente, aunque $f_2(x_1) = 0^2 - 4 * 0 + 5 = 5 > 1 = 2^2 - 4 * 2 + 5 = f_2(x_2)$ y $f_2(x_2) = \sqrt{2+1} \approx 1,732 > 1 = \sqrt{0+1} = f_1(x_1)$ entonces cuando una de las funciones encuentra su mínimo la otra no puede hacerlo, pero entonces, ¿cómo trabajo con las dos funciones al mismo tiempo?

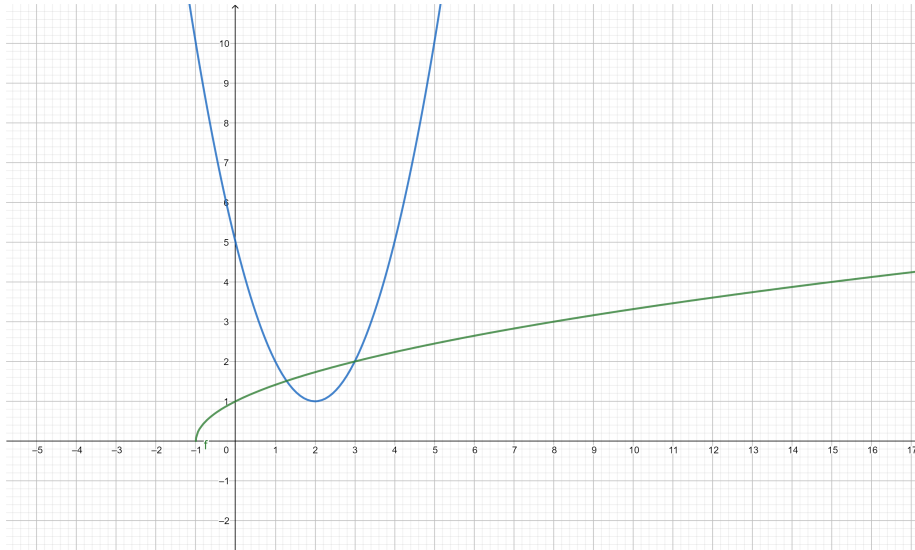


Figura 2.1: Funciones objetivo del Ejemplo 3

2.2. Problema de optimización multiobjetivo

Recordaremos algunas definiciones de optimización a valores escalares para luego comparar con la optimización multiobjetivo.

Definición 2.2.1 *Dado el problema*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & x \in S \subset \mathbb{R}^m \end{aligned}$$

• Decimos que x^* es un **minimizador global** si $f(x^*) \leq f(x)$ para todo $x \in S$, en este caso x^* es llamado el **mínimo** de f en S .

• Decimos que x^* es un **minimizador local** si existe $\varepsilon > 0$ tal que $f(x^*) \leq f(x)$ para todo $x \in S$ tal que $\|x - x^*\| \leq \varepsilon$.

También se acostumbra decir que x^* es **minimizador local estricto** si existe $\varepsilon > 0$ tal que $f(x^*) < f(x)$ para todo $x \in S$ tal que $0 < \|x - x^*\| \leq \varepsilon$

Estudiaremos un *problema de optimización multiobjetivo* de la forma:

$$\begin{aligned} \min \quad & \{f_1(x), f_2(x), \dots, f_k(x)\} \\ \text{sujeto a} \quad & x \in S \end{aligned} \tag{2.3}$$

donde tenemos $k \geq 2$ funciones objetivo $f_i : \mathbb{R}^m \rightarrow \mathbb{R}$.

Denotaremos el vector de funciones objetivo por $f(x) = (f_1(x), f_2(x), \dots, f_k(x))^T$.

El *vector de decisión (variable)* $x = (x_1, x_2, \dots, x_m)^T$ que pertenece a la (no vacía) *región factible* (conjunto) S , que es un subconjunto del *espacio de decisión*

\mathbb{R}^m . Todavía no fijamos la forma de las *restricciones* (*funciones restricción*) que forman S , pero nos referimos a S en general.

La palabra “*min*” significa que queremos minimizar todas las funciones objetivo simultáneamente. Si no hay conflicto entre las funciones objetivo, es decir que las funciones no están acopladas, entonces se puede encontrar una solución donde cada función objetivo alcanza su óptimo. En este caso no son necesarios métodos especiales. Para evitar estos casos triviales, suponemos que no existe una solución única que sea óptima con respecto a cada función objetivo. Esto significa que las funciones objetivo son, al menos, parcialmente conflictivas. También pueden ser inconmensurables (es decir, tener diferentes unidades).

De aquí en adelante, denotaremos a la imagen de la región factible por $Z (= f(S))$ y la llamamos *región objetivo factible*. Es un subconjunto del *espacio objetivo* \mathbb{R}^k . Los elementos de Z se denominan *vectores (función) objetivo* o *vectores criterio* y son denotados por $f(x)$ ó $z = (z_1, z_2, \dots, z_k)^T$, donde $z_i = f_i(x)$ para todo $i = 1, \dots, k$ son *valores (función) objetivo* o *valores criterio*. Las palabras entre paréntesis, son usualmente omitidas para acortar.

Para mayor simplicidad y claridad del tratamiento del problema asumimos que todas las funciones objetivo deben ser minimizadas. Si una función objetivo f_i debe ser maximizada, es equivalente a minimizar la función $-f_i$, así en el ejemplo 3 que se busca maximizar la capacidad de almacenamiento mientras que se minimiza la pérdida de agua por la evaporación a la función que se refiere a la capacidad de almacenamiento $f_{\text{almacenamiento}}$ se la reemplaza y se minimiza $-f_{\text{almacenamiento}}$ y así todas las funciones que están involucradas en ese problema serán minimizadas.

2.3. Optimización de Pareto

Debido al conflicto y a la posible inconmensurabilidad de las funciones objetivo, no es posible encontrar una solución única que sea óptima para todos los objetivos simultáneamente. Los problemas de optimización multiobjetivo, están, en cierto modo mal definidos. No hay un orden natural definido en el espacio objetivo, pues este es sólo parcialmente ordenado (esto significa, por ejemplo, que $(1, 1)^T$ es menor que $(3, 3)^T$, pero ¿cómo comparo $(1, 3)^T$ y $(3, 1)^T$?).

De todas formas, algunos de los vectores objetivos se pueden extraer para ser examinados. Estos son aquellos en los que ninguna de las componentes puede ser mejorada sin deteriorarse otra de las componentes. Edgeworth (1987) presentó esta definición en 1881. Sin embargo, la definición es usualmente llamada Pareto optimal después de que el economista y sociólogo italo-francés Vilfredo Pareto en 1896 lo desarrollara aún más. Una definición más formal de la *Optimalidad de Pareto* es la siguiente:

Definición 2.3.1 *Un vector de decisión $x^* \in S$ es Pareto optimal si no existe otro vector de decisión $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y $f_j(x) < f_j(x^*)$ para al menos un índice j .*

Un vector objetivo $z^* \in Z$ es Pareto optimal si no existe otro vector objetivo $z \in Z$ tal que $z_i \leq z_i^*$ para todo $i = 1, \dots, k$ y $z_j < z_j^*$ para al menos un índice j ; o equivalentemente, z^* es Pareto optimal si el correspondiente vector de decisión es Pareto optimal.

Esta definición (2.3.1) introduce la noción de *globalmente Pareto optimal*.

En optimización de un sólo objetivo, el foco principal está en el espacio de decisión. En el contexto multiobjetivo, a menudo estamos más interesados en el espacio objetivo, que generalmente es de menos dimensión que el espacio de decisión.

En la Figura (2.2) la región factible $S \subset \mathbb{R}^3$ y su imagen, la región objetivo factible $Z \subset \mathbb{R}^2$ están ilustrados. La línea azul contiene a todos los vectores objetivo Pareto optimal. El vector z^* es un ejemplo de ellos.

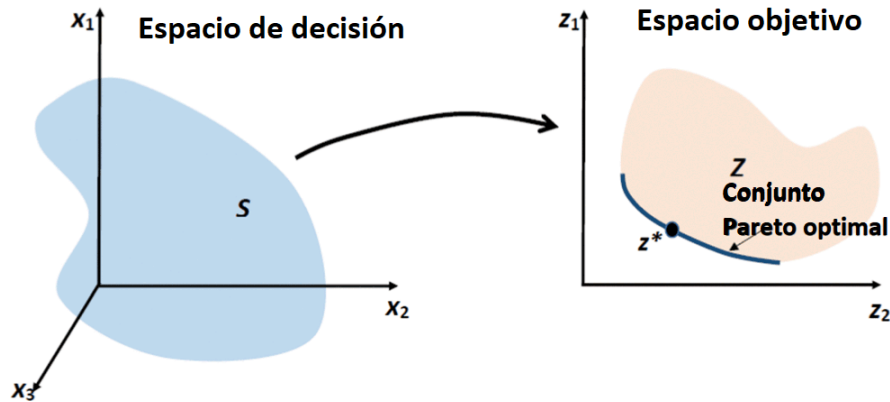


Figura 2.2: Los conjuntos S y Z y los vectores objetivo Pareto optimal

Generalmente hay muchas soluciones Pareto optimal (un número infinito). Podemos hablar de un conjunto de soluciones Pareto optimal o el *conjunto Pareto optimal* o *frente de Pareto*. El conjunto de Pareto o frente de Pareto se define entonces como el conjunto de puntos en el espacio objetivo que contiene a todos los vectores objetivo Pareto optimal.

Recordemos estas definiciones.

Definición 2.3.2

- Un conjunto $S \subset \mathbb{R}^m$ es convexo si para todo $x, y \in S$ el segmento $\alpha x + (1 - \alpha)y \in S$ para todo $0 \leq \alpha \leq 1$.
- Si $S \subset \mathbb{R}^m$ es un conjunto convexo, decimos que $f : S \rightarrow \mathbb{R}$ es una función convexa si para todo $x, y \in S$, $x \neq y$ $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ para todo $0 \leq \lambda \leq 1$.

- Un espacio topológico X es conexo si no existen U y V abiertos en X , no vacíos y disjuntos tales que $X = U \cup V$.
- Un conjunto $S \subset \mathbb{R}^m$ es conexo si como subespacio es un espacio topológico conexo.

En optimización multiobjetivo, a diferencia de la optimización a valores escalares, donde estamos interesados en encontrar una solución óptima, estamos interesados en encontrar todo el conjunto Pareto optimal, ya que, en principio todas las soluciones Pareto optimal son igualmente importantes. En los próximos capítulos veremos cómo generar una aproximación del conjunto de Pareto completo.

La frontera de Pareto obtenida puede ser lineal, cóncava, convexa, continua o discontinua dependiendo de las funciones objetivo integrantes del problema. Todas las soluciones pertenecientes a la frontera son igualmente buenas, y no se puede especificar si alguna de las soluciones es preferible a las otras. En la siguiente figura tenemos algunos frentes de Pareto, el primero es cóncavo o no convexo, el segundo es no conexo y el tercero es convexo.

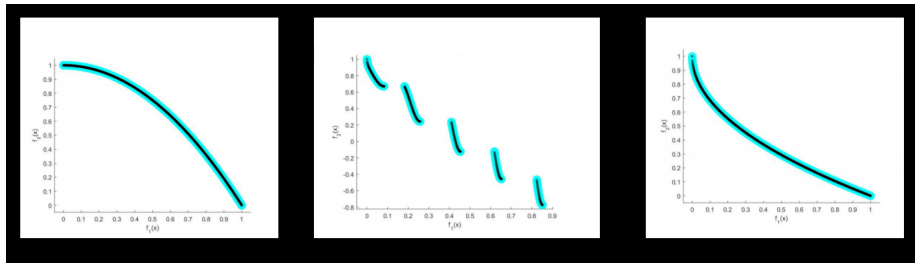


Figura 2.3: Algunos frentes de Pareto.

Otro concepto particularmente importante es el de *localmente Pareto optimal*.

Definición 2.3.3 Un vector de decisión $x^* \in S$ es *localmente Pareto optimal* si existe $\delta > 0$ tal que x^* es Pareto optimal en $S \cap B(x^*, \delta)$.

Un vector objetivo $z^* \in Z$ es localmente Pareto optimal si el correspondiente vector de decisión es localmente Pareto optimal.

Naturalmente, cualquier solución global Pareto optimal es localmente Pareto optimal. La recíproca, es válida para problemas de optimización multiobjetivo convexos. Recordemos que, en optimización a valores escalares un problema de programación convexa se define de la siguiente manera:

Definición 2.3.4

$$\begin{array}{ll} \min & f(x) \\ \text{s.a.} & x \in S \end{array}$$

donde S es un conjunto convexo y f es una función convexa.

Al igual que en el caso de problemas de optimización a valores escalares, un problema de optimización multiobjetivo convexo se define de la siguiente manera:

Definición 2.3.5 *Un problema de optimización multiobjetivo es convexo si todas las funciones objetivo son convexas y el conjunto factible es convexo.*

El siguiente resultado nos da la recíproca a la afirmación anterior.

Teorema 2.3.6 *Si un problema de optimización multiobjetivo es convexo, entonces toda solución localmente Pareto optimal es también una solución globalmente Pareto optimal.*

Demostración:

Sea $x^* \in S$ localmente Pareto optimal. Entonces, existe un $\delta > 0$ y un entorno $B(x^*, \delta)$ de x^* tal que, no existe ningún $x \in B(x^*, \delta)$ de x^* tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y para al menos un índice j $f_j(x) < f_j(x^*)$.

Supongamos que x^* no es globalmente Pareto optimal. En este caso, existe otro punto $x^o \in S$ tal que $f_i(x^o) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y $f_j(x^o) < f_j(x^*)$ para algún j .

Ahora definimos $\hat{x} = \beta x^o + (1 - \beta)x^*$ con $0 < \beta < 1$ elegido de tal forma que $\hat{x} \in B(x^*, \delta)$.

La convexidad de S nos implica que $\hat{x} \in S$. Por la convexidad de las funciones objetivo tenemos que

$$f_i(\hat{x}) \leq \beta f_i(x^o) + (1 - \beta)f_i(x^*) \leq \beta f_i(x^*) + (1 - \beta)f_i(x^*) = f_i(x^*)$$

para todo $i = 1, \dots, k$ pues, x^* es localmente Pareto optimal y $\hat{x} \in B(x^*, \delta)$ entonces $f_i(\hat{x}) = f_i(x^*)$ para todo i . Luego

$$f_i(x^*) \leq \beta f_i(x^o) + (1 - \beta)f_i(x^*) \forall i = 1, \dots, k$$

pues $\beta > 0$, dividiendo por β obtenemos $f_j(x^*) \leq f_j(x^o)$ para todo i . Luego, tenemos que $f_j(x^*) > f_j(x^o)$ para algún j lo que contradice que x^* es localmente Pareto optimal. Entonces x^* es globalmente Pareto optimal. \square

2.3.1. Vector objetivo ideal

Un vector objetivo que minimiza cada una de las funciones objetivo se denomina *vector objetivo ideal (o perfecto)*.

Definición 2.3.7 Las componentes z_i^* de un vector objetivo ideal $z^* \in \mathbb{R}^k$ son obtenidas al minimizar cada una de las funciones objetivo individualmente sujeto a las restricciones, o sea resolviendo para todo $i = 1, \dots, k$

$$\begin{aligned} \min \quad & f_i(x) \\ \text{s.a.} \quad & x \in S \end{aligned} \tag{2.4}$$

Es obvio que, si un vector objetivo ideal fuera factible (o sea si $z^* \in Z$) este sería la solución del problema de optimización multiobjetivo (y el conjunto Pareto optimal se reduciría a este vector). Esto en general no es posible ya que hay un conflicto entre las funciones objetivo. Aunque el vector objetivo ideal no es alcanzable puede considerarse como un punto de referencia, algo que buscar. Del vector objetivo ideal se obtienen cotas inferiores para el conjunto Pareto optimal para cada función objetivo.

En la práctica, se debe tener especial cuidado con los problemas no convexos. La definición de vector objetivo ideal asume que se conoce el mínimo global de cada una de las funciones objetivo. Garantizar la optimización global en los cálculos numéricos no es tan simple. Esto debe ser tenido en cuenta con los problemas prácticos.

Algunas veces, para realizar cálculos numéricos, se necesita un vector que sea estrictamente mejor que el vector objetivo ideal y por eso surge la necesidad de definir el *vector objetivo utópico*

Definición 2.3.8 el vector objetivo utópico $z^{**} \in \mathbb{R}^k$ es un vector objetivo no factible cuyas componentes son de la forma

$$z_i^{**} = z_i^* - \varepsilon_i$$

para todo $i = 1, \dots, k$ donde z_i^* es la componente del vector objetivo ideal y $\varepsilon_i > 0$ es un escalar relativamente pequeño, pero computacionalmente significativo.

2.3.2. Vector objetivo nadir

La cota superior del conjunto Pareto optimal, es decir, las componentes de un *vector objetivo nadir* (o vector objetivo imperfecto) z^{nad} , son mucho más difíciles de encontrar. Sin embargo, pueden estimarse. El vector objetivo nadir puede ser un punto factible o no serlo. Para problemas no lineales, no existe un método constructivo para calcular el vector objetivo nadir.

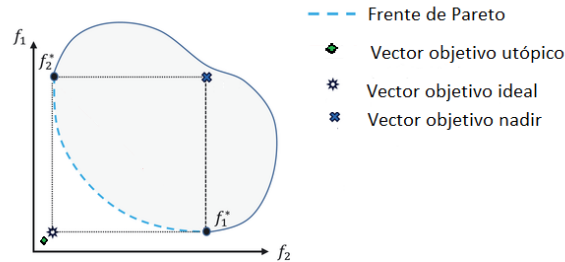


Figura 2.4: Vector objetivo ideal y vector objetivo nadir.

2.4. Optimalidad débil de Pareto

Además de la optimización de Pareto, otros conceptos relacionados son ampliamente utilizados, como el caso de la optimalidad débil de Pareto. La relación entre ellas es que Pareto optimal es un subconjunto de los puntos débilmente Pareto optimal.

Un vector es débilmente Pareto optimal si no existe ningún otro vector tal que todas sus componentes sean mejores.

Definición 2.4.1 *Un vector de decisión $x^* \in S$ es débilmente Pareto optimal si no existe otro vector de decisión $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$.*

Un vector objetivo $z^* \in Z$ es débilmente Pareto optimal si no existe otro vector $z \in Z$ tal que $z_i < z_i^*$ para todo $i = 1, \dots, k$, o equivalentemente, si el correspondiente vector de decisión es débilmente Pareto optimal.

De la misma manera que antes definimos localmente Pareto optimal, se puede definir además la optimalidad débil local de Pareto y la optimalidad débil global de Pareto.

Si bien, las soluciones débilmente Pareto optimal son importantes desde un punto de vista técnico, ya que son fáciles de generar, hay que tener en cuenta que, en algunos casos, pueden no ser tan útiles debido al gran tamaño del conjunto débilmente Pareto optimal.

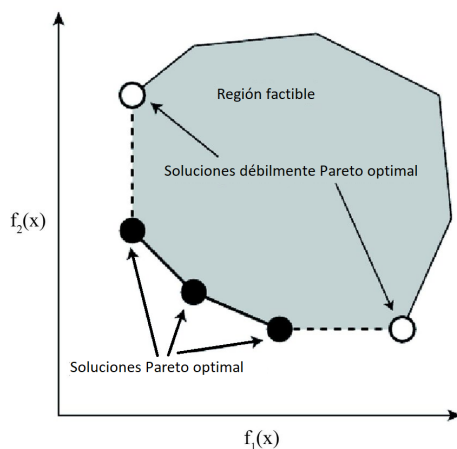


Figura 2.5: Vectores débilmente Pareto optimal y vectores Pareto optimal.

En los problemas de optimización multicriterio, el interés está en encontrar todo el conjunto de Pareto o al menos una gran parte de mismo, ya que son todas las posibles soluciones al problema que queremos resolver y al menos en principio, no hay una solución que sea mejor que otra. Si sólo hay dos funciones objetivo, en ocasiones, el conjunto óptimo de Pareto se puede generar paramétricamente. Cuando hay más objetivos, el problema se vuelve más complicado.

Ahora recordando el ejemplo 4, generaremos el frente de Pareto de manera paramétrica.

En este ejemplo queremos resolver un problema con dos funciones objetivo, que queremos minimizar simultáneamente sobre la línea real no negativa.

$$f_1(x) = \sqrt{x + 1}$$

$$f_2(x) = x^2 - 4x + 5 = (x + 2)^2 - 1$$

Y queremos resolver el problema de optimización

$$\min_{x \geq 0} (f_1(x), f_2(x))$$

Para cada función objetivo individualmente, el correspondiente problema de optimización es simple de resolver siendo $x_1 = 0$ y $x_2 = 2$ son los minimizadores (únicos) de f_1 y f_2 en $x \in \mathbb{R} : x \geq 0$, respectivamente.

El espacio de decisión es \mathbb{R} y el espacio criterio es \mathbb{R}^2 pues $f(X) \subset \mathbb{R}^2$. Para obtener la imagen del conjunto factible en el espacio de criterio, hay que sustituir y_1 por $f_1(x)$ e y_2 por $f_2(x)$ para obtener $x = (y_1)^2 - 1$ (resolviendo $y_1 = \sqrt{1 + x}$ para x).

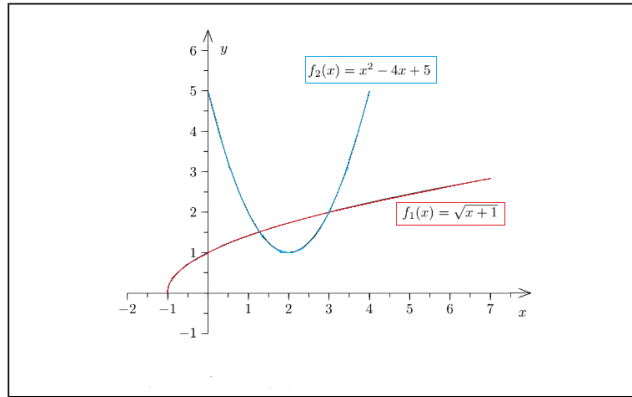


Figura 2.6: Funciones objetivo.

Por lo tanto, obtenemos

$$y_2 = ((y_1)^2 - 1)^2 + 4 - 4(y_1)2 + 5 = (y_1)^4 - 6(y_1)^2 + 10$$

El gráfico de esta función es

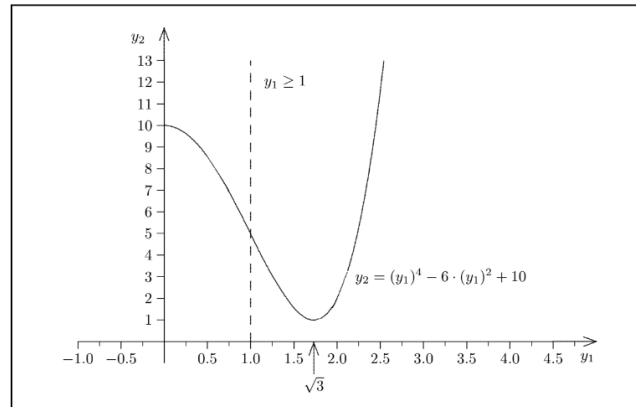


Figura 2.7: Espacio criterio.

Notemos que $x \geq 0$ se traduce como $y_1 \geq 1$, entonces el conjunto de Pareto es la parte del gráfico que se encuentra a la derecha de la línea vertical $y_1 = 1$. Calculando el mínimo de y_2 en función de y_1 , vemos que las soluciones $x \in [0, 2]$ encontradas antes corresponden a valores de $y_1 = f_1(x)$ en $[1, \sqrt{3}]$ y $y_2 = f_2(x) \in [1, 5]$. Estos puntos en la gráfica de $y_2(y_1)$ con $1 \leq y_1 \leq \sqrt{3}$ son los puntos que pertenecen al frente de Pareto. Podemos verlo en color rosa en la siguiente figura:

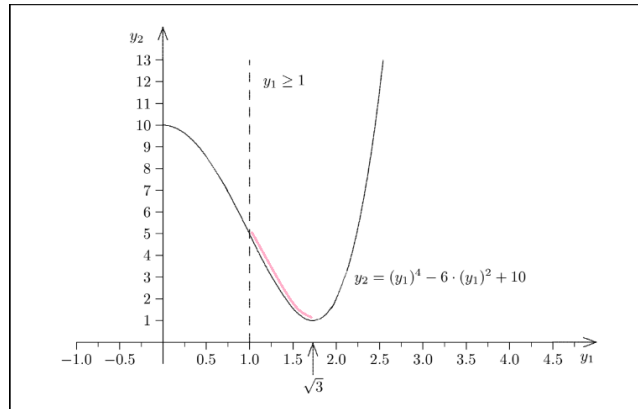


Figura 2.8: Frente de Pareto.

Ahora que ya hemos definido qué es un problema de optimización multicriterio vamos a presentar un último ejemplo.

Ejemplo 5: La radioterapia es el tratamiento de los tejidos cancerosos con radiación ionizante que daña el ácido desoxirribonucleico (ADN) de las células. Mientras que las células no cancerosas son capaces de reproducirse incluso con su ADN ligeramente dañado, el estado de reproducción elevado en el que se encuentran las células cancerosas significa que pequeñas cantidades de ADN dañadas las hace incapaces de reproducirse. El objetivo de la radioterapia es explotar esta ventaja terapéutica y enfocar la radiación de modo que la dosis suficiente se entrega a la región objetivo para extinguir las células cancerosas mientras que no se afecte demasiado a las estructuras circundantes.

El proceso desde el diagnóstico hasta el tratamiento contiene muchos pasos. Comienza con la obtención de imágenes médicas, como tomografías computadas (CT) o resonancias magnéticas (MRI) que se utilizan para diagnosticar y localizar tumores. Estas imágenes permiten la reconstrucción 3D del tumor y de otros órganos con el propósito de planificar el tratamiento.

Diseñar un tratamiento óptimo significa decidir un gran número de parámetros. El proceso de planificación suele dividirse en tres fases:

1. La selección del número de haces y las direcciones desde las cuales enfocar la radiación en el paciente. El **problema geométrico**.
2. El cálculo de un mapa de intensidad (o patrón de intensidad) para cada una de las direcciones seleccionadas en la fase 1. El **problema de intensidad**.

3. Encontrar una secuencia de configuraciones del colimador multilámina (MLC) para administrar el tratamiento. El **problema de realización**.

Consideramos el problema de intensidad, al que modelaremos como un problema de optimización multiobjetivo. El objetivo es encontrar un plan de irradiación para un paciente con cáncer de próstata.

El tumor será irradiado con 5 haces equidistantes que pueden descomponerse en 400 haces controlables distintos. Asumimos que ya resolvimos el problema geométrico, por lo que la geometría de los haces está fija.

Las partes relevantes del cuerpo del paciente son mapeadas con la ayuda de la tomografía computada y de acuerdo al grosor de las rebanadas que dividiremos en cubos llamados voxels.

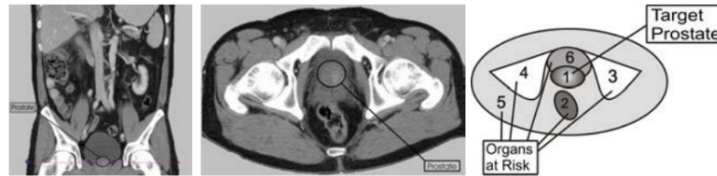


Figura 2.9: Tomografía computada, corte coronal y axial. Esquema del corte axial del cuerpo con estructuras de interés.

Utilizando métodos de clustering, podemos agrupar los voxels que están expuestos a la misma radiación, así reducimos los 435501 voxels originales en 11877 clusters, c_j con $j = \{1, \dots, 11877\}$. Luego a cada cluster se le asigna una de las estructuras de volumen de interés V_0, \dots, V_6 , esta asignación es realizada por un médico. En este ejemplo el tumor es V_0 , V_1 , el recto es V_2 los huesos izquierdo y derecho de la cadera son V_3 y V_4 respectivamente, el tejido circundante es V_5 y la vejiga es V_6 .

Exámenes médicos han demostrado que la vejiga y el recto son órganos sensibles mientras que los otros órganos siguen a estos órganos dominantes en su estrés causado por diferentes planes de irradiación. Para más información consultar ([3], [5]).

La emisión de los haces B_i con $i \in \{1, \dots, 400\}$ en los clusters c_j es descrita por la matriz $P = (P_{ij})$, $j = 1, \dots, 11877$, $i = 1, \dots, 400$. Sea $x \in \mathbb{R}^{400}$ es el perfil de intensidad, entonces $P_j x$ denota la dosis de radiación en el cluster c_j causado por los haces B_i , $i = 1, \dots, 400$ siendo P_j la j -ésima fila de la matriz P .

Para evaluar y comparar el estrés producido por la radiación en los órganos usamos el concepto de *dosis uniforme equivalente* basado en p -normas propuesto por Nimierko.

$$EUD_k(x) = \frac{1}{U_k} \left(\frac{1}{N(V_k)} \sum_{j|c_j \in V_k} N(c_j) \cdot (P_j x)^{p_k} \right)^{\frac{1}{p_k}} - 1$$

Para $k = 2, \dots, 6$ y el escalar $p_k \in [1, \infty]$ es una constante dependiente del órgano que refleja la estructura más paralela o en serie del órgano, $N(V_k)$ es el número de voxels en el órgano V_k y $N(c_j)$ es el número de voxels en el cluster c_j , entonces $\sum_{j|c_j \in V_k} N(c_j) = N(V_k)$. El valor U_k es la dosis límite de cada órgano, que no debe superarse y es un valor evaluado estadísticamente.

Un plan de tratamiento factible ahora tiene que satisfacer varias limitaciones. En primer lugar, debe evitarse una sobredosis peligrosa del tejido crítico y, por tanto, no debe superarse el valor máximo Q_k para todos los órganos de riesgo V_k , $k = 2, \dots, 6$, i. e.

$$U_k(EUD_k(x) + 1) \leq Q_k \quad k = 2, \dots, 6.$$

Estas restricciones pueden ser reescritas como:

$$\sum_{j|c_j \in V_k} N(c_j)(P_j x)^{p_k} \leq Q_k^{p_k} \cdot N(V_k) \quad k = 2, \dots, 6.$$

También es importante que la dosis en el tejido tumoral permanezca por debajo de un valor máximo para evitar lesiones en el cuerpo del paciente y para lograr la homogeneidad de la irradiación. Además, para tener el efecto deseado de destruir todas las células tumorales, se debe alcanzar una dosis curativa. Aquí, diferenciamos entre el llamado V_0 tejido objetivo y el V_1 tejido de refuerzo, que es tejido tumoral que tiene que ser irradiado especialmente alto. Esas condiciones se traducen en las siguientes restricciones para cada cluster del objetivo y del refuerzo:

$$L_0(1 - \varepsilon_0) \leq P_j x \leq L_0(1 + \delta_0) \quad \forall j \text{ con } c_j \in V_0$$

$$L_1(1 - \varepsilon_1) \leq P_j x \leq L_1(1 + \delta_1) \quad \forall j \text{ con } c_j \in V_1$$

Donde $L_0, L_1, \varepsilon_0, \varepsilon_1, \delta_0$ y δ_1 son constantes dadas por un profesional. Por otra parte, hay que asegurarse que la intensidad de los haces es no negativa. Resumiendo esto, tenemos el siguiente conjunto factible:

$$S = \{x \in \mathbb{R}_+^{400} \mid \begin{array}{l} U_k(EUD_k(x) + 1) \leq Q_k, \quad k = 2, \dots, 6, \\ L_0(1 - \varepsilon_0) \leq P_j x \leq L_0(1 + \delta_0), \quad \forall j \text{ con } c_j \in V_0 \\ L_1(1 - \varepsilon_1) \leq P_j x \leq L_1(1 + \delta_1), \quad \forall j \text{ con } c_j \in V_1 \end{array} \}$$

con 17795 restricciones y 400 variables. El objetivo ahora es mantener la peligrosa sobredosis de los órganos en riesgo, el recto (V_2) y la vejiga (V_6), tan bajo como sea posible, es decir que nuestras dos funciones objetivo a minimizar

son EUD_2 y EUD_6 . Las investigaciones han demostrado que estos dos órganos son los órganos dominantes y que los otros órganos en riesgo siguen en el nivel de sus valores EUD de estos órganos con un valor inferior.

Entonces tenemos el siguiente problema de optimización bi-objetivo:

$$\min_{x \in S} \begin{pmatrix} f_2(x) \\ f_1(x) \end{pmatrix} = \min_{x \in S} \begin{pmatrix} EUD_2(x) \\ EUD_6(x) \end{pmatrix}$$

Es muy importante tener en cuenta que para resolver este tipo de problemas, se necesitan datos numéricos proporcionados por profesionales (médicos, bioingenieros, físicos, etc).

Capítulo 3

Métodos numéricos

En este capítulo presentaremos varios métodos para resolver problemas de optimización multiobjetivo. Por lo general, esto significa encontrar la solución Pareto optimal que mejor satisfaga las necesidades del tomador de decisiones.

En la mayoría de los métodos, estamos interesados en el espacio objetivo en lugar del espacio de decisión. Una de las razones por las que nos interesa el espacio objetivo es que la dimensión de este suele ser considerablemente más pequeña que la dimensión del espacio de decisión. Otra razón es que los tomadores de decisiones a menudo están más interesados en los valores objetivo. Sin embargo, el cálculo tendrá lugar en el espacio de decisión ya que generalmente no conocemos la forma explícita de la región objetivo factible. En resumen, los tomadores de decisiones generalmente manejan valores objetivos, mientras que la programación matemática se lleva a cabo en el espacio de decisión.

En general, los problemas de optimización multiobjetivo se resuelven mediante la escalarización, aunque hay excepciones como el **método de haces proximal multiobjetivo** para problemas no diferenciables que no se basa en la escalarización en el sentido tradicional.

Escalarizar significa convertir el problema multiobjetivo en un solo problema o en una familia de problemas de optimización de un solo objetivo con una función objetivo a valores reales, denominada *función de escalarización*, que posiblemente dependa de algunos parámetros. Esto permite el uso de la teoría y los métodos de optimización escalar, es decir, programación no lineal. Debido a que las funciones de escalarización generalmente dependen de ciertos parámetros auxiliares, pueden aparecer algunas dificultades numéricas, por lo tanto, la idea aparentemente prometedora de simplificar el problema en optimizaciones de un sólo objetivo también tiene sus debilidades.

Un hecho importante a tener en cuenta es que las rutinas estándar para problemas de optimización de un único objetivo solamente pueden encontrar

óptimos locales. Esta es la razón por la que sólo localmente se obtienen y manejan soluciones óptimas de Pareto cuando se trata de funciones de escalarización. Si las funciones objetivo y la región factible son convexas entonces se pueden garantizar soluciones Pareto optimal globales.

Una cuestión a considerar es la posibilidad de que la función de escalarización tenga varias soluciones óptimas alternativas. En este caso, el vector objetivo producido depende de la solución elegida.

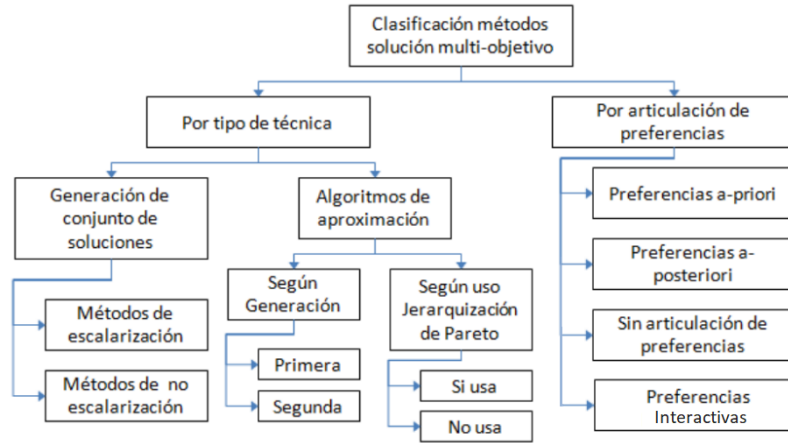
Al seleccionar un método de solución, se deben tener en cuenta las características específicas del problema a resolver. Además, las opiniones del tomador de decisiones son importantes, el tomador de decisiones, es la persona que presenta el problema al analista, la persona que efectivamente lo resolverá, en algunos casos puede que ambos actores sean la misma persona, pero en este trabajo pensaremos al tomador de decisiones como una persona ajena y nosotros asumiremos el rol de analista. No es suficiente que el analista simplemente prefiera algún método. Puede suceder que el tomador de decisiones no pueda o no quiera usarlo. Se puede decir que seleccionar un método de optimización multiobjetivo apropiado en sí mismo es un problema con múltiples objetivos.

Los métodos de optimización multiobjetivo pueden clasificarse de muchas maneras según diferentes criterios. La clasificación de Hwang y Masud (1979) [8] donde los métodos se clasifican de acuerdo con la manera en que se expresen las preferencias del tomador de decisiones. Según esta clasificación las clases son:

1. métodos de no preferencia: métodos donde el responsable de tomar las decisiones no brinda información de preferencia, sólo se necesita una solución.
2. métodos a posteriori: métodos donde el tomador de decisiones aporta información de preferencia después de ser encontradas las posibles soluciones.
3. métodos a priori: métodos donde el encargado de tomar las decisiones comunica información de preferencia que se deberá tener en cuenta antes de resolver el problema.
4. métodos interactivos: métodos en los que se utiliza preferiblemente la información de preferencia de manera progresiva.

Hay que tener en cuenta que las clasificaciones no son absolutas. Otra clasificación puede ser teniendo en cuenta el tipo de técnica utilizado para resolver el problema. La superposición y las combinaciones de clases son posibles y se puede considerar que algunos métodos pertenecen a más de una clase. La agrupación presentada es sólo para orientación.

En el siguiente cuadro se muestran dos clasificaciones, por tipo de técnica y teniendo en cuenta las preferencias del tomador de decisiones.



A continuación, veremos varios métodos para resolver problemas de optimización multiobjetivo. No se pretende abarcar todos los métodos existentes, sino formas de abordar la resolución de estos problemas.

3.1. Métodos de no preferencia

En los métodos de no preferencia, donde las opiniones de los tomadores de decisiones no son tenidas en consideración, el problema de optimización multiobjetivo es resuelto utilizando algún método relativamente simple y la solución obtenida es presentada al encargado de tomar las decisiones. Luego, este puede o bien aceptar la solución o rechazarla. Parece bastante improbable que la solución que mejor satisfaga al tomador de decisiones pueda encontrarse con estos métodos. Es por eso que los métodos sin preferencia son adecuados para situaciones en las que, quien toma las decisiones no responde a ninguna expectativa especial de la solución y está satisfecho simplemente con alguna solución óptima.

3.1.1. Método del criterio global

El método del criterio global también se llama a veces *programación de compromiso*. En este método, la distancia entre algún punto de referencia y la región objetivo factible se minimiza. El analista debe seleccionar el punto de referencia y la métrica para medir las distancias. Se cree que todas las funciones objetivo son igualmente importantes.

Diferentes métricas

Examinaremos el método donde el vector objetivo ideal es usado como punto de referencia y las métricas L_p se usan para medir. En este caso, el problema L_p a resolver es:

$$\begin{aligned} \min & \left(\sum_{i=1}^k |f_i(x) - z_i^*|^p \right)^{1/p} \\ \text{s.a.} & x \in S \end{aligned} \quad (3.1)$$

De la definición de vector objetivo ideal z^* sabemos que $f_i(x) \geq z_i^*$ para todo $i = 1, \dots, k$ y para todo $x \in S$. Es por eso que no se necesitan valores absolutos si conocemos el vector objetivo ideal global. Si el vector objetivo ideal global no se conoce, el método no necesariamente funciona como debería. Si el vector objetivo ideal se reemplaza por otro vector, debe seleccionarse cuidadosamente. Se deben evitar los puntos de referencia pesimistas ya que el método no puede encontrar soluciones mejores que el punto de referencia.

El exponente $1/p$ puede descartarse. Los problemas con o sin el exponente $1/p$ son equivalentes para $1 \leq p < \infty$, ya que el problema L_p es una función creciente del problema correspondiente sin el exponente.

Si $p = \infty$ la métrica L_∞ también es llamada *métrica de Tchebycheff* y el problema es de la forma:

$$\begin{aligned} \min & \max_{i=1, \dots, k} [|f_i(x) - z_i^*|] \\ \text{s.a.} & x \in S \end{aligned} \quad (3.2)$$

Observe que el problema (3.2) no es diferenciable incluso en ausencia de valores absolutos. En este caso, sin embargo, puede transformarse en un formato diferenciable si el objetivo y las funciones de restricción son diferenciables. Así el problema será de la forma:

$$\begin{aligned} \min & \alpha \\ \text{s.a.} & \alpha \geq f_i(x) - z_i^* \quad \forall i = 1, \dots, k \\ & x \in S \end{aligned} \quad (3.3)$$

se resuelve, siendo $x \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ variables.

La solución obtenida depende en gran medida del valor elegido para p . Las opciones ampliamente utilizadas son $p = 1, 2$ o ∞ . En la Figura (3.1), se muestran los contornos de estas tres métricas diferentes. El punto negro es el vector objetivo ideal y la línea negrita representa el conjunto Pareto optimal. Vale la pena notar que, si el problema original es lineal, la opción $p = 1$ mantiene la linealidad. A medida que aumenta el valor p , el problema de minimización no lineal se vuelve más difícil y mal condicionado para resolverlo.

El método del criterio global es un método simple de usar si el objetivo es simplemente obtener una solución sin ninguna expectativa. Las propiedades de las métricas implican que, si las funciones objetivo no se normalizan de ninguna

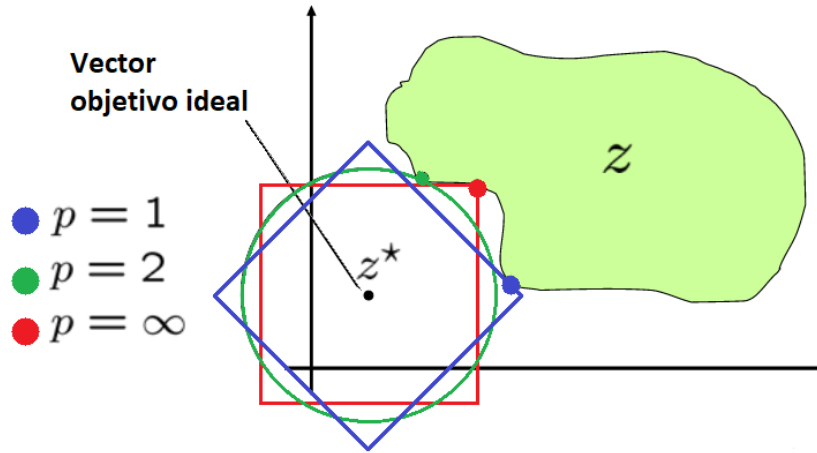


Figura 3.1: Diferentes métricas.

manera, entonces una función objetivo cuyo valor objetivo ideal se sitúa más cerca de la región objetivo factible recibe más importancia. Se garantiza que la solución obtenida con la métrica L_p ($1 \leq p < \infty$) es la Pareto optimal. Si se utiliza la métrica de Tchebycheff, la solución puede ser débilmente Pareto optimal. Depende del analista seleccionar una métrica apropiada.

3.1.2. Método de haces proximal multiobjetivo (MPB)

La idea del método de haces proximal multiobjetivo (MPB), es moverse en una dirección donde los valores de todas las funciones objetivo mejoren simultáneamente.

El método MPB es capaz de resolver problemas con funciones no lineales (posiblemente no diferenciables). Se supone que todos los objetivos y las funciones de restricción son localmente Lipschitz. El método MPB no es como los otros métodos de escalarización. Los métodos de escalarización ordinarios transforman el problema en uno con una sola función objetivo. Este nuevo problema se puede resolver con cualquier método apropiado para la programación no lineal. En el método MPB, la función de escalarización está formulada dentro de un optimizador especial (no diferenciable). Supongamos ahora que la región factible es de la forma

$$S = \{x \in \mathbb{R}^n \mid g(x) = (g_1(x), g_2(x), \dots, g_m(x))^T \leq 0\}.$$

Como se mencionó anteriormente, el método MPB no se basa directamente en el empleo de ninguna función de escalarización. Sin embargo, se necesita

algún tipo de escalarización para derivar el método de minimización para todas las funciones objetivas. Teóricamente, utilizamos una *función de mérito* sin restricciones $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. definida por:

$$H(x^1, x^2) = \max [f_i(x^1) - f_i(x^2), g_l(x^1)] | i = 1, \dots, k, l = 1, \dots, m]$$

El algoritmo MPB

En el método MPB, la solución se busca de forma iterativa, hasta que se cumpla algún criterio de parada.

Presentamos los pasos básicos del algoritmo MPB:

1. El contador de iteraciones h se refiere a las iteraciones internas del método MPB. Sea x^h la aproximación actual a la solución del problema de optimización multiobjetivo en la iteración h .
2. Buscar la dirección de búsqueda d^h como una solución del problema de optimización sin restricciones

$$\begin{array}{ll} \min & H(x^h + d, x^h) \\ \text{s.a.} & d \in \mathbb{R}^n \end{array} \quad (3.4)$$

3. Aproximar el problema teniendo en cuenta que es no diferenciable, la forma de aproximarlo dependerá de si el problema es convexo o no convexo y definirá la estrategia a seguir.

El método MPB se puede utilizar como un método en el que no se busca opinión del tomador de decisiones. En este caso, debemos seleccionar el punto de partida para que no sea (débilmente) Pareto optimal, sino que cada componente del vector objetivo pueda ser mejorada. La debilidad de la rutina MPB es que no se puede garantizar que las soluciones obtenidas sean Pareto optimal. En la práctica, sin embargo, es muy probable que las soluciones sean al menos débilmente Pareto optimal.

3.2. Métodos a posteriori

Los métodos a posteriori también podrían llamarse *métodos para generar soluciones Pareto optimales*. Una vez que se ha generado el conjunto Pareto optimal (o una parte de él), se presenta al tomador de decisiones, quien seleccionará entre las alternativas su preferida, la mejor. Los inconvenientes aquí son que el proceso de generación suele ser computacionalmente costoso y, a veces, difícil. Por otro lado, es difícil para el tomador de decisiones seleccionar entre un amplio conjunto de alternativas. Una pregunta importante es cómo presentar o mostrar las alternativas al encargado de tomar las decisiones de manera efectiva.

Si sólo hay dos funciones objetivo, el conjunto óptimo de Pareto se puede generar paramétricamente. Cuando hay más objetivos, el problema se vuelve

más complicado. En problemas no lineales, la distinción se encuentra entre problemas convexos y no convexos. En otras palabras, algunos métodos sólo pueden generar soluciones Pareto optimales para problemas convexos.

3.2.1. Suma ponderada

En el método de las sumas ponderadas la idea es asociar a cada función objetivo con un coeficiente de peso y minimizar la suma ponderada de las funciones objetivo. De esta manera, las múltiples funciones objetivo se transforman en una sola función objetivo. Los coeficientes de peso w_i son números reales tales que $w_i \geq 0$ para todo $i = 1, \dots, k$ y se supone que los pesos están normalizados, es decir $\sum_{i=1}^k w_i = 1$. Para ser más exactos el problema de optimización multiobjetivo se modifica en el siguiente problema:

$$\begin{aligned} \min \quad & \sum_{i=1}^k w_i f_i(x) \\ \text{s.a.} \quad & x \in S \end{aligned} \tag{3.5}$$

donde $0 \leq w_i \leq 1$ para $i = 1, \dots, k$ $\sum_{i=1}^k w_i = 1$. En el capítulo 4 se explicará cómo elegir los coeficientes de ponderación w_i para generar el frente de Pareto.

Ventajas y desventajas: Este método es computacionalmente eficiente para generar soluciones débilmente Pareto optimal sin cambiar la estructura del problema manteniendo el mismo conjunto de restricciones. La debilidad del método es que no se pueden encontrar todas las soluciones Pareto optimal a menos que el problema sea convexo.

3.2.2. Restricción ε

En el método de restricción ε , una de las funciones objetivo es seleccionada para ser optimizada y las otras funciones objetivo se convierten en restricciones eligiendo una cota superior para cada una de ellas. El problema a resolver ahora es de la forma

$$\begin{aligned} \min \quad & f_l(x) \\ \text{s.a.} \quad & f_j(x) \leq \varepsilon_j \quad \forall j = 1, \dots, k; j \neq l \\ & x \in S \end{aligned} \tag{3.6}$$

donde $l \in \{1, \dots, k\}$.

Ventajas y desventajas: Este método nos conduce a una solución débilmente Pareto optimal. Sin embargo, si sólo existe una solución óptima esta resulta ser Pareto optimal. Su desventaja es que es un proceso que consume mucho tiempo, y en el caso de demasiados objetivos, es difícil formular el problema en sí. La elección de ε afecta significativamente a las soluciones de problemas cuando se trata de valores objetivos continuos y también modificar la estructura del problema mediante la incorporación de restricciones adicionales.

3.2.3. Restricción elástica

Este método se desarrolló con el fin de incorporar los beneficios del método de la suma ponderada y de la restricción ε , evitando al mismo tiempo sus debilidades. Este método está concebido para resolver un único objetivo como beneficio de una suma ponderada, teniendo en cuenta todas las soluciones eficientes, como en el método de restricción ε . La idea subyacente es que las restricciones elásticas permiten resolver fácilmente un problema estableciendo límites superiores en los valores objetivos violados que se utilizan para penalizar la violación de las restricciones. El problema quedará de la siguiente forma

$$\begin{aligned} \min \quad & f_j(x) + \sum_{k \neq j} \mu_k s_k \\ \text{s.a.} \quad & f_k(x) - s_k \leq \varepsilon_k \quad k \neq j \\ & s_k \geq 0 \quad k \neq j \\ & x \in S \end{aligned} \tag{3.7}$$

donde ε_k es un parámetro y μ_k es el coeficiente de penalización para una dada función objetivo k . Este método utiliza variables de holgura s_k con el fin de transformar los límites superiores de los valores objetivos en restricciones de igualdad para cualquier $x \in S$ basado en una selección adecuada de s .

Ventajas y desventajas: Como solución a la dificultad de elegir adecuadamente los ε_j en el problema de restricción ε , el método de restricción elástica relaja estas restricciones permitiendo que sean violadas y penalizando cualquier violación en la función objetivo, lo que aumenta la eficiencia en la búsqueda de soluciones. La desventaja es que la complejidad de la búsqueda de soluciones del método de restricción elástica se ve significativamente afectada por el valor de μ_k en el que un pequeño μ_k incurre en mayor complejidad computacional. En consecuencia, la existencia de la optimalidad de Pareto se ve notablemente afectada por el valor de μ_k .

3.2.4. Métricas ponderadas

Este método utiliza un criterio global al minimizar la distancia entre un determinado punto de referencia, z_i^* , y una región objetivo factible donde cada métrica puede ser ponderada con un grado diferente. El problema para minimizar la distancia es ahora de la forma:

$$\begin{aligned} \min \quad & (\sum_{i=1}^k w_i |f_i(x) - z_i^*|^p)^{1/p} \\ \text{s.a.} \quad & x \in S \end{aligned} \tag{3.8}$$

para $1 \leq p < \infty$, asumiendo que el vector objetivo ideal global es conocido. Una reformulación basada en el problema ponderado de Tchebyshev es dada por

$$\begin{aligned} \min \quad & \max_{i=1, \dots, k} (\sum_{i=1}^k w_i |f_i(x) - z_i^*|) \\ \text{s.a.} \quad & x \in S \end{aligned} \tag{3.9}$$

El primer problema (3.8) es Pareto optimal si la solución es única o si $w_i > 0$ mientras que el segundo problema (3.9) da una solución débilmente Pareto optimal para $w_i > 0$.

Ventajas y desventajas: Si se conocen los puntos de referencia z^* , un vector de valores objetivos utópicos, el segundo problema (3.9) puede identificar al menos una solución Pareto optimal para cada función objetivo. Sin embargo, el primer problema (3.8) no es diferenciable, lo que excluye el uso de optimizadores objetivos individuales que utilicen información de gradiente.

3.2.5. Programación por metas (Goal programming)

Este método es utilizado por los tomadores de decisiones para asignar metas específicas que quieren alcanzar para cada objetivo. La función objetivo tiene por objeto minimizar la diferencia absoluta entre los objetivos previstos y el rendimiento alcanzado, y se formula como:

$$\min \sum_{i=1}^k |f_i(x) - T_i| \quad (3.10)$$

donde T_i denota el conjunto de metas establecidas para la i -ésima función objetivo. El método de programación por metas puede ser usado como método a posteriori o como método a priori dependiendo del conjunto de metas T_i . Si las metas fueron establecidas por el tomador de decisiones antes de resolver el problema entonces será un método a priori. Si por el contrario estas metas fueron establecidas por el analista usando, por ejemplo, el vector objetivo ideal o el vector objetivo utópico como meta, entonces será tomado como método a posteriori.

Ventajas y desventajas: Este método es computacionalmente eficiente si el objetivo se selecciona en el espacio de solución factible. Sin embargo, el encargado de tomar las decisiones debe establecer metas específicas adecuadamente con el fin de eliminar las características del problema que no son proporcionales. Este no es un proceso trivial a menos que la información sobre la forma del espacio de búsqueda se conozca a priori. Si la región factible no es identificada, este método es ineficiente. Además, una debilidad bien conocida de este método de programación por metas es generar soluciones que no sean Pareto optimales.

A continuación, presentamos una tabla que resume los métodos a posteriori, haciendo énfasis en las ventajas y desventajas de cada método. Además, agregamos a la tabla que tipo de soluciones obtenemos con cada método, si es que tenemos alguna garantía de obtener soluciones Pareto optimal o al menos débilmente Pareto optimal.

Método	Ventajas	Desventajas	Optimalidad
SUMA PONDERADA	Computacionalmente eficiente generando soluciones débilmente Pareto optimal.	Dependiente de los coeficientes de ponderación, frentes de Pareto concavos no pueden ser generados.	Pareto optimal o débilmente Pareto optimal para funciones objetivo convexas.
RESTRICCIÓN ε	Puede identificar frentes de Pareto concavos.	Proceso que lleva mucho tiempo; difícil formular un problema con demasiados objetivos.	Débilmente Pareto optimal.
RESTRICCIÓN ELÁSTICA	Transforma algunas restricciones para una mejor eficiencia de búsqueda de soluciones	Impacto significativo por un factor de penalización	No hay garantía.
MÉTRICAS PONDERADAS	Permite encontrar un conjunto de soluciones Pareto optimal para cada objetivo si se conocen puntos de referencia como el vector objetivo utópico	No puede resolver problemas de optimización multiobjetivo para funciones no diferenciables.	Débilmente Pareto optimal.
PROGRAMACIÓN POR METAS	Computacionalmente eficiente si el espacio factible es encontrado.	Computacionalmente ineficiente si el espacio factible no es encontrado.	No hay garantía.

Tabla 3.1: Ventajas, desventajas y optimalidad de los métodos a posteriori

3.3. Métodos a priori

En el caso de los métodos a priori, el tomador de decisiones debe especificar sus preferencias, esperanzas y opiniones antes del proceso de solución. La dificultad es que la persona que toma las decisiones no necesariamente sabe de antemano que es posible alcanzar en el problema y qué tan realistas son sus expectativas.

3.3.1. Método de función de valor

En el método de la función de valor, el tomador de decisiones puede dar una forma matemática precisa y explícita de la función de valor $U : \mathbb{R}^k \rightarrow \mathbb{R}$ que representa sus preferencias a nivel global. Esta función proporciona un orden

completo del espacio objetivo. Entonces, el *problema de la función de valor*

$$\begin{aligned} \max \quad & U(f(x)) \\ \text{s.a.} \quad & x \in S \end{aligned} \tag{3.11}$$

está listo para ser resuelto por algún método para la optimización de un solo objetivo como se ve en la figura :

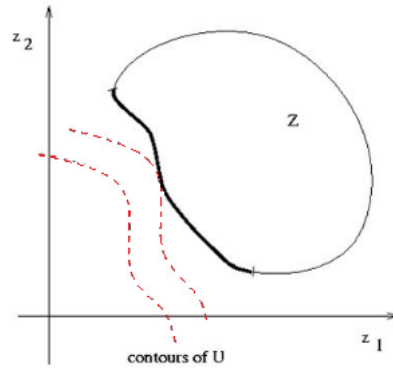


Figura 3.2: Curvas de nivel de la función de valor.

La línea en negrita representa el conjunto Pareto optimal y las líneas rojas son las curvas de nivel de la función de valor, la solución del problema es Pareto optimal si la función de valor está disminuyendo fuertemente. El método de la función de valor parece ser un método muy simple, pero la dificultad radica en especificar la expresión matemática de la función de valor. Se muestra que los métodos de codificación que teóricamente deberían producir funciones de valor idéntico fallan: las funciones pueden diferir entre sí en más del 50 %. También se señala que no existe un análisis real de la precisión de la evaluación de la función de valor. Las verificaciones de consistencia, es decir, si los tomadores de decisiones brindan respuestas consistentes a preguntas similares, no son adecuadas: un instrumento sesgado puede proporcionar datos consistentes. Por otro lado, incluso si fuera posible que la persona que toma las decisiones exprese sus preferencias globalmente, la estructura de preferencias resultante podría ser demasiado simple, ya que las funciones de valor no pueden representar intransitividad o incompatibilidad.

El método de la función de valor podría llamarse una forma “óptima” de resolver problemas de optimización multiobjetivo si el responsable de la toma de decisiones pudiera expresar de manera confiable la función de valor. El uso del método de la función de valor está restringido en la práctica a los problemas

de análisis de decisiones de múltiples atributos con un conjunto discreto de alternativas factibles.

El método de la función de valor es un método excelente si el tomador de decisiones conoce una formulación matemática explícita para la función de valor y si esa función representa totalmente las preferencias del tomador de decisiones. Estas dos condiciones previas cruciales son las dificultades del enfoque. Hay ciertas condiciones que las preferencias del tomador de decisiones pueden satisfacer para que se pueda definir una función de valor en ellas. Por lo tanto, puede que no exista necesariamente una función de valor que imponga un orden total en el conjunto de posibles vectores objetivos. La suposición de un orden total a menudo es contraria a nuestros objetivos intuitivos y, por lo tanto, es muy probable que conduzca a selecciones menos que ideales. Una cosa importante a tener en cuenta en la práctica es que las aspiraciones del tomador de decisiones pueden cambiar durante el proceso de solución.

3.3.2. Ordenamiento lexicográfico

En el ordenamiento lexicográfico, el tomador de decisiones debe organizar las funciones objetivo de acuerdo con su importancia absoluta. Este orden significa que un objetivo es infinitamente más importante que un objetivo menos importante. Después de ordenar, la función objetivo más importante se minimiza sujeta a las restricciones originales. Si este problema tiene una solución única, es la solución de todo el problema de optimización multiobjetivo. De lo contrario, la segunda función objetivo más importante se minimiza. Ahora, además de las restricciones originales, se agrega una nueva restricción. Esta nueva restricción está ahí para garantizar que la función objetivo más importante conserve su valor óptimo. Si este problema tiene una solución única, es la solución del problema original. De lo contrario, el proceso continúa como se indicó anteriormente.

Dadas las funciones objetivo organizadas de acuerdo con el orden lexicográfico desde f_1 la función objetivo más importante hasta f_k la menos importante. Escribimos el *problema lexicográfico* como:

$$\begin{array}{ll} \text{lex min} & f_1(x), f_2(x), \dots, f_k(x) \\ \text{s.a.} & x \in S \end{array} \quad (3.12)$$

Teorema 3.3.1 *La solución del problema lexicográfico es Pareto optimal*

La demostración del teorema (3.3.1) puede consultarse en [9] página 119.

El ordenamiento lexicográfico corresponde al método de ponderación cuando los coeficientes de ponderación son de magnitud muy diferente. La pregunta de si existen vectores de ponderación tales que la solución óptima del método de ponderación sea idéntica a la solución obtenida por ordenamiento lexicográfico, la respuesta es positiva para problemas lineales y varios problemas discretos. En la práctica, esto significa que el problema del ordenamiento lexicográfico puede

resolverse como un problema de sumas ponderadas con optimizadores estándar.

La justificación para utilizar el orden lexicográfico es su simplicidad y el hecho de que las personas suelen tomar decisiones sucesivamente. Sin embargo, este método tiene varios inconvenientes. El tomador de decisiones puede tener dificultades para poner las funciones objetivas en un orden absoluto de importancia.

Por ejemplo, si pensamos nuevamente en el ejemplo 2, que presentamos al inicio del trabajo, donde comparábamos autos, si definimos el problema lexicográfico como:

$$\begin{array}{ll} \text{lex min} & (-\text{potencia}(x), \text{precio}(x), \text{consumo de combustible}(x)) \\ \text{s.a.} & x \in \{VW, Opel, Ford, Toyota\} \end{array} \quad (3.13)$$

Teniendo en cuenta que tomamos $-\text{potencia}$ porque lo que queremos es maximizar la potencia del auto, entonces para poder minimizar trabajamos con $-\text{mayor potencia}$. En este caso, el auto que elegiríamos es el Toyota, ya que tiene potencia 71.0 mientras que VW, Opel y Ford tienen 66.0, 62.0, 55.0 respectivamente. entonces con analizar sólo la primera función fue suficiente y las demás funciones objetivo no tuvieron ningún peso al momento de tomar la decisión. Lo mismo ocurriría si el problema lexicográfico fuera definido de la siguiente manera:

$$\begin{array}{ll} \text{lex min} & (\text{precio}(x), -\text{potencia}(x), \text{consumo de combustible}(x)) \\ \text{s.a.} & x \in \{VW, Opel, Ford, Toyota\} \end{array} \quad (3.14)$$

En este caso elegiríamos el Ford Focus ya que su precio es de 14000 euros, mientras que VW, Opel y Toyota cuestan 16200, 14200 y 15200 euros respectivamente y nuevamente analizando una sola función objetivo ya obtengo el mínimo y no influyen las otras funciones objetivo.

Ahora si además de comparar VW, Opel, Ford y Toyota comparamos el Fiat Argo que cuesta 14000 euros, 54.0 de potencia y un consumo de 7.3 litros cada 100 kilómetros, tomando nuevamente el mismo orden lexicográfico:

$$\begin{array}{ll} \text{lex min} & (\text{precio}(x), -\text{potencia}(x), \text{consumo de combustible}(x)) \\ \text{s.a.} & x \in \{VW, Opel, Ford, Toyota, Fiat\} \end{array} \quad (3.15)$$

Tanto el Fiat Argo como el Ford Focus cuestan lo mismo, entonces es necesario minimizar la segunda función objetivo, o sea, $-\text{potencia}(x)$ y así nuevamente elegiríamos el Ford Focus por tener mayor potencia que el Fiat Argo, pero si el problema fuera:

$$\begin{array}{ll} \text{lex min} & (\text{precio}(x), \text{consumo de combustible}(x), -\text{potencia}(x)) \\ \text{s.a.} & x \in \{VW, Opel, Ford, Toyota, Fiat\} \end{array} \quad (3.16)$$

Elegiríamos el Fiat Argo que consume menos combustible que el Ford Focus, pero no se tiene en cuenta la tercera función objetivo en la toma de decisión.

Teniendo en cuenta el ejemplo anterior vemos que el método suele ser robusto. Los objetivos menos importantes tienen muy pocas posibilidades de afectar la solución final. Si la función objetivo más importante tiene una solución única, los otros objetivos no tienen ninguna influencia en la solución. Incluso si el objetivo más importante tuviera óptimos alternativos y fuera posible usar el segundo objetivo más importante, es muy poco probable que este problema tenga óptimos alternativos, y se puedan utilizar el tercer u otros objetivos menos importantes.

El orden lexicográfico no permite intercambiar un pequeño incremento de una función objetivo importante con una gran disminución de una función objetivo menos importante. Sin embargo, este tipo de comercio a menudo puede ser atractivo para el tomador de decisiones.

Vamos a presentar una tabla que resume los métodos a priori, al igual que como hicimos con los métodos a posteriori, haciendo énfasis en las ventajas y desventajas de cada método.

Método	Ventajas	Desventajas
FUNCIÓN DE VALOR	Si U presenta la estructura de preferencia global del tomador de decisiones la solución obtenida es la mejor.	Es muy difícil para el tomador de decisiones especificar la formulación matemática de U . Incluso si se conociera explícitamente U , el tomador de decisiones puede tener dudas o cambiar sus preferencias.
ORDENAMIENTO LEXICOGRÁFICO	La solución es Pareto optimal.	Dificultad para especificar un orden absoluto de importancia. El método es robusto, los objetivos menos importantes tienen muy pocas posibilidades de afectar a la solución final.

Tabla 3.2: Ventajas y desventajas de los métodos a priori

3.4. Métodos interactivos

El interés dedicado a esta clase puede explicarse por el hecho de que, suponiendo que el tomador de decisiones tenga suficiente tiempo y capacidades para la cooperación, se puede suponer que los métodos interactivos producen

los resultados más satisfactorios. Es decir, sólo se debe generar y evaluar una parte de los puntos Pareto optimal, y el tomador de decisiones puede especificar y corregir sus preferencias y selecciones a medida que continúa el proceso de solución y así se conoce mejor el problema y sus potencialidades. Además, se puede suponer que el tomador de decisiones tiene más confianza en la solución final ya que está involucrado en todo el proceso de solución.

En los métodos interactivos, el tomador de decisiones trabaja junto con un analista o un programa informático interactivo. Se puede decir que el analista intenta determinar la estructura de preferencias del tomador de decisiones de manera interactiva. Se forma un patrón de solución y se repite varias veces. Después de cada iteración se proporciona cierta información al responsable de la toma de decisiones y se le pide que responda algunas preguntas o proporcione algún otro tipo de información. Después de un número razonable (finito) de iteraciones, cada método interactivo debe proporcionar una solución que el tomador de decisiones pueda estar satisfecho y convencido de que no existe una solución considerablemente mejor. Los pasos básicos en algoritmos interactivos se pueden expresar como:

1. Encontrar una solución inicial factible
2. Interactuar con el tomador de decisiones
3. obtener una nueva solución (o un conjunto de nuevas soluciones). Si la nueva solución (o una de ellas) o una de las soluciones anteriores es aceptable para el tomador de decisiones, deténgase. De lo contrario, vaya al paso 2.

Un problema a resolver al diseñar un método interactivo es qué tipo de datos se deben usar para interactuar con el tomador de decisiones. Los tomadores de decisiones deben comprender el significado de los parámetros para los cuales se le pide que proporcione valores. Para garantizar que se pueda obtener el mayor beneficio posible del método interactivo, el responsable de la toma de decisiones debe considerar que el método vale la pena y es aceptable y debe poder usarlo correctamente. Esto generalmente significa que el método debe ser comprensible y suficientemente fácil de usar. El analista puede apoyar al tomador de decisiones de muchas maneras y, en el mejor de los casos, explicarle al tomador de decisiones el comportamiento del problema.

Los métodos interactivos se han clasificado de muchas maneras, principalmente de acuerdo con sus enfoques de solución.

Repetidamente, se ha asumido y se asumirá que el tomador de decisiones toma decisiones consistentes o que tiene una función de valor subyacente (conocida implícitamente) sobre la cual se toman sus decisiones. El propósito no es profundizar en las teorías de la toma de decisiones. Sin embargo, vale la pena

mencionar que esas suposiciones pueden cuestionarse porque son difíciles de verificar.

Para resolver problemas prácticos, se necesita conocimiento sobre los procesos de decisión y el análisis de decisiones para garantizar una cooperación fructífera entre el tomador de decisiones y el analista. Para provocar una toma de decisiones intuitiva, los analistas deben proporcionar información en términos de gráficos ricos y multidimensionales y evitar requerir consistencia.

Un concepto más interesante es la convergencia de un método interactivo. Uno puede entender varias características diferentes como convergencia. Por un lado, se puede decir que el método converge en puntos Pareto optimales si se puede demostrar que la solución final es Pareto optimal. También se puede decir que el método converge en una solución satisfactoria, si la solución final es satisfactoria. Encontrar la mejor solución de compromiso Pareto optimal puede entenderse como convergencia.

El único criterio práctico de detención es la satisfacción del tomador de decisiones con la solución obtenida. Esto generalmente significa que el tomador de decisiones debe sentir que se ha recibido suficiente información sobre el problema para ser resuelto.

A pesar de que los métodos interactivos pueden considerarse los métodos de solución más prometedores para la optimización de objetivos múltiples, todavía hay casos en los que estos métodos no son practicables independientemente de la disponibilidad del tomador de decisiones.

Veremos dos métodos interactivos, el método de compensación interactiva de valor sustituto (ISWT) y el método de punto de referencia.

3.4.1. Método de compensación interactiva de valor sustituto (ISWT)

El método de restricción ε , es un elemento fundamental del método de compensación interactiva de valor sustituto ISWT. La idea es maximizar una función de valor subyacente (implícitamente conocida). Las opiniones del tomador de decisiones sobre las tasas de compensación en el punto de solución actual detectan una dirección de búsqueda. El tamaño del paso a tomar en la dirección de búsqueda se determina resolviendo varios problemas de restricción ε y pidiéndole al tomador de decisiones que seleccione la solución más satisfactoria para la continuación del proceso.

Se supone que:

1. La función de valor subyacente $U : \mathbb{R}^k \rightarrow \mathbb{R}$ existe y es implícitamente conocida por el tomador de decisiones. Además, U es continuamente

diferenciable y decreciente.

2. Las funciones son dos veces continuamente diferenciables.
3. La región factible S es compacta.

Algoritmo ISWT

Las principales características del método ISWT se pueden presentar de manera gradual con cuatro pasos:

1. Seleccionar una función de referencia f_l para ser minimizada y establecer cotas superiores para las restantes funciones objetivo. Definir $h = 1$ que será el número de la iteración realizada.
2. Resolver el problema de restricción ε para obtener una solución Pareto optimal x^h . La tasa de información se obtiene de los multiplicadores KKT correspondientes.
3. Solicitar las opiniones del tomador de decisiones con respecto a las tasas de compensación en z^h correspondientes a x^h .
4. Si se cumple algún criterio de detención, deténgase con x^h como la solución final. De lo contrario, actualice los límites superiores de la función objetivo con la ayuda de las respuestas obtenidas en el paso (3) y resuelve varios problemas de restricción ε (para determinar un tamaño de paso apropiado). Deje que el tomador de decisiones elija la alternativa preferida. Denote el vector de decisión correspondiente por x^{h+1} y defina $h = h + 1$. Vaya al paso (3).

En la práctica, cuando se le pide al tomador de decisiones que exprese sus preferencias con respecto a las tasas de compensación, se le pide implícitamente que compare las tasas de compensación con sus tasas marginales de sustitución. Por esto es sumamente importante que el responsable de tomar las decisiones comprenda bien el significado de las compensaciones.

Es necesario tener en cuenta que el rol de f_l es importante y esta debe ser elegida con sumo cuidado.

La facilidad de comparación depende de la cantidad de objetivos y del tomador de decisiones.

Una desventaja de este método es que puede ser difícil para el responsable de tomar las decisiones especificar valores de valor sustituto.

Una ventaja de este método es que todas las soluciones son Pareto optimal.

3.4.2. Método de punto de referencia

Como su nombre indica, el método del punto de referencia, se basa en un punto de referencia de los niveles de aspiración. Un nivel de aspiración, es un valor en una variable de objetivo (por ejemplo, ganancias o participación de mercado) o, en el caso de objetivos múltiples, un vector de valores de objetivos que es satisfactorio para el tomador de decisiones. El punto de referencia es un punto factible o inviable en el espacio objetivo que es razonable o deseable para el tomador de decisiones. El punto de referencia se utiliza para derivar funciones de escalarización de méritos que tienen soluciones mínimas en puntos débilmente o Pareto optimal. En este método, la generación de soluciones Pareto optimal se basa en puntos de referencia, no en funciones de valor o vectores de ponderación. No se establecen suposiciones específicas sobre el problema a resolver.

Los puntos de referencia son intuitivos y fáciles de especificar para el tomador de decisiones y su consistencia no es un requisito esencial.

Se clasifican las funciones objetivo en aceptables o inaceptables. Si el nivel de aspiración es inferior al valor objetivo actual, esa función objetivo es actualmente inaceptable, y si el nivel de aspiración es igual o superior al valor objetivo actual, esa función es aceptable. La diferencia aquí es que el punto de referencia puede no ser factible en todas las componentes. En otras palabras, donde el conjunto de funciones objetivas aceptables está vacío, el enfoque basado en puntos de referencia todavía se puede utilizar. Naturalmente, esto no significa que se puedan disminuir todos los valores objetivos, sino que se puede generar una solución diferente.

La técnica de optimización multiobjetivo interactiva del método de punto de referencia es muy simple y práctica. Antes de que comience el proceso de solución, se proporciona cierta información al tomador de decisiones sobre el problema. Si es posible, el vector objetivo ideal y el vector objetivo nadir aproximado se presentan para ilustrar los rangos del conjunto óptimo de Pareto. Otra posibilidad es minimizar y maximizar las funciones objetivas individualmente en la región factible (si está limitada). Se presentan tanto la variable de decisión como los valores objetivos. También se debe seleccionar una forma apropiada para la función de mérito. Un ejemplo de función de mérito que utiliza tanto niveles de aspiración como niveles de reserva \check{z}_i , que son valores de funciones objetivo que se desean evitar, donde z_i^* son las componentes del vector objetivo ideal, $\rho > 0$ es un argumento y

$$\nu = \min_{i=1, \dots, k} \frac{\bar{z}_i - z_i^*}{\check{z}_i - \bar{z}_i}$$

Es la siguiente:

$$\min_{i=1, \dots, k} \left[\min \left[\frac{\check{z}_i - f_i(x)}{\check{z}_i - \bar{z}_i}, 1 + \frac{\nu(\bar{z}_i - f_i(x))}{\bar{z}_i - z_i^*} \right] \right] +$$

$$\rho \left(\sum_{i=1}^k \min \left[\frac{\check{z}_i - f_i(x)}{\check{z}_i - \bar{z}_i}, 1 + \frac{\nu(\bar{z}_i - f_i(x))}{\bar{z}_i - z_i^*} \right] \right)$$

Las funciones de mérito pueden ser computacionalmente muy costosas.

Los pasos básicos del método del punto de referencia son los siguientes:

1. Presentar información sobre el problema al responsable de la toma de decisiones. Definir $h = 1$ que será el número de iteraciones realizadas.
2. Solicitar al responsable de la toma de decisiones que especifique un punto de referencia $\bar{z}^h \in \mathbb{R}^k$ (un nivel de aspiración para cada función objetivo).
3. Minimizar la función de logro y obtener una solución Pareto optimal x^h y la z^h correspondiente. Presentar z^h al responsable de la toma de decisiones.
4. Calcular otras k soluciones Pareto optimal minimizando la función de logro con puntos de referencia perturbados

$$\bar{z}(i) = z^h + d^h e^i$$

donde $d^h = \|\bar{z}^h - z^h\|$ y e^i es el i -ésimo vector unitario para $i = 1, \dots, k$.

5. Presentar las alternativas al tomador de decisiones. Si encuentra una de las $k+1$ soluciones satisfactoria, la correspondiente x^h es la solución final. De lo contrario, solicite al tomador de decisiones que especifique un nuevo punto de referencia \bar{z}^{h+1} , establecer $h = h + 1$ y volver al paso (3).

La ventaja de perturbar el punto de referencia en el paso (4) es que el fabricante de decisiones obtiene una concepción más clara de las posibles soluciones. Si el punto de referencia está lejos del conjunto Pareto optimal, el tomador de decisiones obtiene una descripción más amplia del conjunto Pareto optimal y si el punto de referencia está cerca del conjunto Pareto optimal, se proporciona una descripción más precisa del conjunto Pareto optimal. Los efectos de la perturbación y los puntos de referencia cercanos y distantes se ilustran en la Figura (3.3)

Un punto de referencia no se puede definir directamente con la ayuda de la función de valor. Sin embargo, es posible probar si un nuevo punto de referencia tiene un valor de función de valor más alto que las soluciones anteriores.

El método de puntos de referencia es bastante fácil de entender para el tomador de decisiones. El tomador de decisiones sólo tiene que especificar los niveles de aspiración apropiados y comparar los vectores objetivos.

Nuevamente presentamos una tabla que resume los métodos interactivos que consideramos en esta sección, haciendo énfasis en las ventajas y desventajas de cada método.

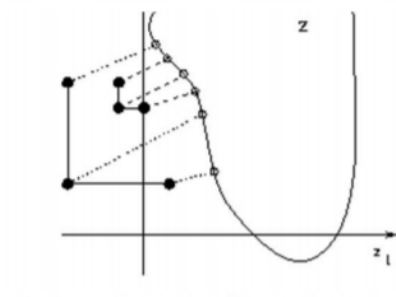


Figura 3.3: Alterando puntos de referencia.

Método	Ventajas	Desventajas
COMPENSACIÓN INTERACTIVA DE VALOR SUSTITUTO (ISWT)	Todas las soluciones son Pareto optimal.	Puede ser difícil para el tomador de decisiones especificar coherentemente valores de valor sustituto.
PUNTO DE REFERENCIA	Para el tomador de decisiones es simple de entender, tiene que especificar los niveles de aspiración y comparar vectores objetivo. Funciona para problemas no diferenciables.	La facilidad de comparación depende del problema. No hay una estrategia clara para producir la solución final.

Tabla 3.3: Ventajas y desventajas de los métodos interactivos

Capítulo 4

Métodos de escalarización

El enfoque tradicional para resolver problemas de optimización multicriterio es mediante algún **método de escalarización**, que implica formular un único problema de optimización a valores reales relacionado con el problema multiobjetivo original. La función de escalarización suele ser una función de las funciones objetivo y de las variables del problema de optimización multiobjetivo original utilizando variables vectoriales o escalares auxiliares y/o parámetros. A veces el conjunto factible del problema de optimización multiobjetivo original está restringido adicionalmente por nuevas restricciones relacionadas con las funciones objetivo del problema original o con nuevas variables.

Cómo vimos en el capítulo 3 los métodos de **sumas ponderadas**, **restricción ε** y **métricas ponderadas** son métodos a posteriori, por lo que, en principio, no necesitamos que el responsable de tomar las decisiones nos facilite ninguna información antes de enfrentar el problema, y estos tres métodos nos garantizan encontrar puntos en el conjunto Pareto optimal o en el conjunto débilmente Pareto optimal. En este capítulo vamos a desarrollar más en detalle estos métodos.

Además, como lo que buscamos es todo el frente de Pareto o al menos una gran parte de él, lo que vamos a presentar es como utilizar los métodos de sumas ponderadas, restricción ε y métricas ponderadas para generar aproximaciones del frente de Pareto, dicho en otras palabras, cómo elegir adecuadamente los parámetros y cómo modificarlos para generar una buena aproximación del conjunto Pareto optimal o del conjunto débilmente Pareto optimal. Además, siguiendo con este análisis, presentaremos un método basado en la escalarización por combinaciones convexas que también nos permite generar una aproximación del conjunto Pareto optimal.

4.1. Sumas ponderadas

En el método de sumas ponderadas las múltiples funciones objetivo se transforman en una sola al ser multiplicadas cada una por un coeficiente de ponderación no negativo y luego sumadas entre sí. Supondremos que los coeficientes de ponderación están normalizados, de esta forma el problema de optimización multiobjetivo se transforma en el siguiente *problema de ponderación*

$$\begin{aligned} \min \quad & \sum_{i=1}^k w_i f_i(x) \\ \text{s.a.} \quad & x \in S \end{aligned} \tag{4.1}$$

donde los coeficientes de ponderación w_i son números reales tales que $0 \leq w_i \leq 1$ para $i = 1, \dots, k$ y $\sum_{i=1}^k w_i = 1$.

Para encontrar una aproximación del frente de Pareto, lo que hacemos es variar los coeficientes de ponderación y resolver sucesivos problemas.

Algorithm 1 Aproximación del conjunto eficiente para $k = 2$ usando el método de sumas ponderadas

Entrada: $r =$ cantidad de coeficientes de ponderación w .

1. Generar los coeficientes de ponderación w_1 y w_2 tales que $w_1 + w_2 = 1$ y $w_i \geq 0$

2. Resolver

$$\min_{x \in S} w_1 f_1(x) + w_2 f_2(x)$$

con solución x^i con $i = 1, \dots, r$.

Salida: El conjunto $A := \{f(x^1), f(x^2), \dots, f(x^r)\}$ que es la aproximación del frente de Pareto.

A continuación, veremos varios resultados teóricos relacionados con el método de las sumas ponderadas.

Teorema 4.1.1 *La solución del problema de ponderación (4.1) es débilmente Pareto optimal.*

Demostración:

Sea $x^* \in S$ solución del problema de ponderación. Supongamos que no es débilmente Pareto optimal. En este caso, existe al menos una solución $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$. De acuerdo con los supuestos establecidos para los coeficientes de ponderación, $w_j > 0$ para al menos un j . Así tenemos

$$\sum_{i=1}^k w_i f_i(x) < \sum_{i=1}^k w_i f_i(x^*)$$

Lo que nos lleva a una contradicción ya que x^* era solución del problema de ponderación. De esta manera x^* es débilmente Pareto optimal. \square

Teorema 4.1.2 *La solución del problema de ponderación (4.1) es Pareto optimal si todos los coeficientes de ponderación son positivos, es decir $w_i > 0$ para todo $i = 1, \dots, k$.*

Demostración:

Sea $x^* \in S$ solución del problema de ponderación con coeficientes de ponderación positivos. Supongamos que no es Pareto optimal. En este caso, existe al menos una solución $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y $f_j(x) < f_j(x^*)$ para al menos un j . Como $w_i > 0$ para todo $i = 1, \dots, k$ se tiene que

$$\sum_{i=1}^k w_i f_i(x) < \sum_{i=1}^k w_i f_i(x^*)$$

Lo que contradice que x^* era solución del problema de ponderación y, entonces, x^* debe ser Pareto optimal. \square

Teorema 4.1.3 *Si la solución del problema de ponderación (4.1) es única entonces es Pareto optimal.*

Demostración:

Sea $x^* \in S$ la única solución del problema de ponderación. Supongamos que no es Pareto optimal. En este caso, existe al menos una solución $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y $f_j(x) < f_j(x^*)$ para al menos un índice j . Debido a que todos los coeficientes de ponderación w_i son no negativos tenemos que

$$\sum_{i=1}^k w_i f_i(x) \leq \sum_{i=1}^k w_i f_i(x^*)$$

Por otro lado, por la unicidad de x^*

$$\sum_{i=1}^k w_i f_i(x^*) < \sum_{i=1}^k w_i f_i(\hat{x})$$

para todo $\hat{x} \in S$. Las dos desigualdades anteriores son contradictorias y, por lo tanto, x^* resulta ser Pareto optimal. \square

Los dos teoremas anteriores afirman que la solución del problema de ponderación es siempre Pareto optimal si los coeficientes de ponderación son todos positivos o si la solución es única, sin ningún otro supuesto. La debilidad del método de sumas ponderadas es que no se pueden encontrar todas las soluciones Pareto optimal a menos que el problema sea convexo.

El siguiente teorema nos mostrará que cualquier solución Pareto optimal de un problema de optimización multiobjetivo convexo puede ser encontrada utilizando el método de sumas ponderadas.

Teorema 4.1.4 *Dado un problema de optimización multiobjetivo convexo, si $x^* \in S$ es Pareto optimal, entonces existen vectores de ponderación w ($w_i \leq 0$, $i = 1, \dots, k$, $\sum_{i=1}^k w_i = 1$) tal que x^* es solución del problema de ponderación.*

La demostración del teorema (4.1.4) puede consultarse en [9] página 89.

Hay que tener en cuenta que el vector de ponderación puede no ser único.

En conclusión, el método de sumas ponderadas es una forma sencilla de generar diferentes soluciones Pareto optimales y estas se garantizan si los coeficientes de ponderación son positivos o si la solución es única. Se pueden obtener diferentes soluciones Pareto optimales mediante el método de sumas ponderadas alterando los coeficientes de ponderación positivos. Para generar una aproximación de todo el frente de Pareto sólo se necesita variar los coeficientes de ponderación para así generar diferentes soluciones. Una debilidad de este método es que no se pueden encontrar todos los puntos Pareto optimales si el problema no es convexo. Esto lo veremos mejor en el capítulo 5 con algunos problemas test.

4.2. Restricción ε

Otra escalarización, especialmente utilizada para calcular puntos Pareto optimales, se basa en la minimización de una sola de las k funciones objetivo, mientras que todas las demás se transforman en restricciones mediante la introducción de cotas superiores, este es el llamado método de *restricción ε* . El problema a resolver con este método es de la forma

$$\begin{aligned} \min \quad & f_l(x) \\ \text{s.a.} \quad & f_j(x) \leq \varepsilon_j \quad \forall j = 1, \dots, k; j \neq l \\ & x \in S \end{aligned} \tag{4.2}$$

Aquí los parámetros son las cotas superiores ε_i con $i \in \{1, \dots, k\} \setminus \{l\}$ para $l \in \{1, \dots, k\}$. Estos parámetros ε_i son difíciles de determinar y para poder generar una aproximación del frente de Pareto es necesario variar convenientemente el vector ε .

A continuación, probaremos un resultado relacionado con los puntos débilmente Pareto optimal.

Teorema 4.2.1 *Una solución del problema de restricción ε (4.2) es débilmente Pareto optimal.*

Demostración:

Sea $x^* \in S$ solución del problema de restricción ε . Supongamos que x^* no es débilmente Pareto optimal. En este caso, existe al menos una solución $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$.

Esto significa que $f_j(x) < f_j(x^*) \leq \varepsilon_j$ para todo $j = 1, \dots, k$, $j \neq l$. Entonces x es un punto factible con respecto al problema de restricción ε , mientras que $f_l(x) < f_l(x^*)$, nos conduce a una contradicción al supuesto que x^* es una solución al problema de restricción ε . Entonces, x^* deberá ser débilmente Pareto optimal. \square

Ahora vamos a relacionar los puntos Pareto optimal con el método de restricción ε . Para esto veremos algunos teoremas.

Teorema 4.2.2 *Un vector de decisión $x^* \in S$ es Pareto optimal si y sólo si es solución del problema de restricción ε para todo $l = 1, \dots, k$, donde $\varepsilon_j = f_j(x^*)$ para $j = 1, \dots, k$, $j \neq l$.*

Demostración:

\Rightarrow) Sea $x^* \in S$ Pareto optimal. Supongamos que x^* no resuelve el problema de restricción ε para algún l donde $\varepsilon_j = f_j(x^*)$ para $j = 1, \dots, k$, $j \neq l$. Entonces existe una solución $x \in S$ tal que $f_l(x) < f_l(x^*)$ y $f_j(x) \leq f_j(x^*)$ con $j \neq l$. Esto contradice que x^* es Pareto optimal. En otras palabras, x^* tiene que resolver el problema para alguna función objetivo.

\Leftarrow) Como $x^* \in S$ es, por hipótesis, una solución del problema de restricción ε para cada $l = 1, \dots, k$, no existe otro $x \in S$ tal que $f_l(x) < f_l(x^*)$ y $f_j(x) \leq f_j(x^*)$ donde $j \neq l$. Esta es la definición de que x^* es Pareto optimal. \square

Notemos que por la condición necesaria del teorema anterior es posible encontrar cada solución Pareto optimal de cualquier problema de optimización multiobjetivo utilizando el método de restricción ε .

Teorema 4.2.3 *Un punto $x^* \in S$ es Pareto optimal si es la única solución del problema de restricción ε (4.2) para algún l con $\varepsilon_l = f_l(x^*)$ para $j = 1, \dots, k$, $j \neq l$.*

Demostración:

Sea $x^* \in S$ la única solución del problema de restricción ε para algún l . Supongamos que no es Pareto optimal. En otras palabras, existe algún punto $x^o \in S$ tal que $f_i(x^o) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y para al menos un índice j es válido que $f_j(x^o) < f_j(x^*)$. La unicidad de x^* implica que para todo $x \in S$ tal que $f_i(x) \leq f_i(x^*)$, $i \neq l$, es $f_l(x^*) < f_l(x)$. Lo que nos lleva a una contradicción que provino de las desigualdades, y así x^* debe ser Pareto optimal. \square

Teorema 4.2.4 *La única solución del problema de restricción ε (4.2) es Pareto optimal para cualquier cota superior $\varepsilon = (\varepsilon_1, \dots, \varepsilon_{l-1}, \varepsilon_{l+1}, \dots, \varepsilon_k)^T$.*

Demostración:

Sea $x^* \in S$ la única solución del problema de restricción ε . Esto significa que $f_l(x^*) < f_l(x)$ para todo $x \in S$ cuando $f_j(x^*) \leq \varepsilon_j$ para cada $j = 1, \dots, k, j \neq l$. Supongamos que x^* no es Pareto optimal. En este caso, existe un vector $x^o \in S$ tal que $f_i(x^o) \leq f_i(x^*)$ para todo $i = 1, \dots, k$, y la desigualdad es estricta para al menos un índice j .

Si $j = l$, entonces $f_l(x^o) < f_l(x^*)$ y $f_i(x^o) \leq f_i(x^*) \leq \varepsilon_j$ para todo $i \neq l$. Lo que contradice que x^* es una solución del problema de restricción ε .

Por otro lado, si $j \neq l$, entonces $f_l(x^o) < f_l(x^*)$ y $f_i(x^o) \leq f_i(x^*) \leq \varepsilon_i$ para todo $i \neq j$ y l , y $f_l(x^o) \leq f_l(x^*)$. Esto contradice que x^* es la única solución del problema de restricción ε , y así x^* debe ser Pareto optimal. \square

Como antes dijimos, es difícil determinar el vector ε de manera conveniente, pero el siguiente lema nos presenta cotas superiores e inferiores para ε para problemas bi-criterio.

Lema 4.2.5 *Dado $k = 2$ y sean \bar{x}^1 solución minimal de $\min_{x \in S} f_1(x)$ y \bar{x}^2 solución minimal de $\min_{x \in S} f_2(x)$. Si \bar{x} es solución Pareto optimal del problema de optimización multiobjetivo, entonces existe un parámetro $\varepsilon \in \mathbb{R}$ tal que \bar{x} es solución del problema de restricción ε*

$$\begin{aligned} \min \quad & f_2(x) \\ \text{s.a.} \quad & f_1(x) \leq \varepsilon \\ & x \in S \end{aligned} \tag{4.3}$$

$$y f_1(\bar{x}^1) \leq \varepsilon \leq f_1(\bar{x}^2).$$

La demostración del lema (4.2.5) puede consultarse en [4].

Si nos concentramos en el caso bi-criterio podemos utilizar el lema anterior para elegir el parámetro ε de forma conveniente pero todavía queda por aclarar cómo elegir el parámetro ε en el intervalo $[f_1(\bar{x}^1), f_1(\bar{x}^2)]$ para poder generar una aproximación del frente de Pareto. Asumimos que las funciones objetivo f_1 y f_2 son dos veces continuamente diferenciables. Además de eso, suponemos que ya hemos resuelto

$$\begin{aligned} \min \quad & f_2(x) \\ \text{s.a.} \quad & f_1(x) \leq \varepsilon^0 \\ & x \in S \end{aligned} \tag{4.4}$$

para algún parámetro $\varepsilon^0 \in \mathbb{R}$ con $x^0 := x(\varepsilon^0)$ una solución con multiplicador de Lagrange $\mu^0 \geq 0$ de las restricción $f_1(x) - \varepsilon^0 \leq 0$. El punto x^0 es débilmente Pareto optimal y $f(x^0)$ sirve como un punto de aproximación del conjunto Pareto optimal ε .

Ahora queremos encontrar el parámetro ε^1 con

$$\|f(x(\varepsilon^1)) - f(x^0)\| = \alpha \tag{4.5}$$

para algún valor $\alpha > 0$ dado para generar otra solución al menos débilmente Pareto optimal, a distancia α de la solución anterior. En todo momento, suponemos que $x(\varepsilon)$ denota una solución óptima del problema de restricción ε (4.4) para algún $\varepsilon \in \mathbb{R}^{k-1}$. Suponemos que la restricción $f_1(x) \leq \varepsilon^0$ está activa en x^0 , o sea si, se cumple con $f_1(x^0) = \varepsilon^0$. De lo contrario, podemos encontrar fácilmente un parámetro $\tilde{\varepsilon}^0$ con $f_1(x^0) = \tilde{\varepsilon}^0$. Bajo los supuestos anteriores, la función de valor mínimo local $\tau^\delta : \mathbb{R} \rightarrow \mathbb{R}$,

$$\tau^\delta(\varepsilon) := \inf\{f_2(x) \mid f_1(x) \leq \varepsilon, x \in S, x \in B_\delta(x^0)\},$$

con $B_\delta(x^0)$ la bola cerrada centrada en x^0 con radio δ para algún $\delta > 0$ pequeño, es diferenciable en un entorno de ε^0 con

$$(\tau^\delta)'(\varepsilon^0) = -\mu^0$$

Usando la derivada de la función local de valor mínimo para una aproximación de Taylor (suponiendo que esto sea posible) y suponiendo que la restricción $f_1(x) \leq \varepsilon$, sigue estando activa, obtenemos

$$f_2(x(\varepsilon^1)) \approx f_2(x^0) - \mu^0(\varepsilon^1 - \varepsilon^0)$$

Como consecuencia la ecuación (4.5) es aproximadamente satisfecha por

$$\varepsilon^1 = \varepsilon^0 \pm \alpha \left(\sqrt{1 + (\mu^0)^2} \right)^{-1}$$

.

Para el caso de un problema bi-criterio ($k = 2$) tenemos el algoritmo (2) que, nos sirve para escoger de manera conveniente el parámetro ε para resolver el problema de restricción ε y así aproximar el conjunto eficiente.

En el caso de tener 3 o más funciones objetivo no podemos dar cotas tan fuertes para el parámetro ε , pero podemos reducir el espacio del parámetro a un conjunto compacto (para $f(S)$ compacto), por ejemplo en el caso $m = 3$ de la siguiente manera: Resuelva $\min_{x \in S} f_i(x)$ para $i = 1, 2$ con solución minimal $x^{min,i}$ y valores mínimos $f_i(x^{min,i}) =: \varepsilon_i^{min}$ tanto como el $\max_{x \in S} f_i(x)$ para $i = 1, 2$ con solución maximal $x^{max,i}$ y valores máximos $f_i(x^{max,i}) =: \varepsilon_i^{max}$. Entonces, para cada solución \bar{x} Pareto optimal de un problema de optimización multicriterio existe un $\varepsilon \in \mathbb{R}^2$ que es solución minimal del problema

$$\begin{aligned} \min \quad & f_3(x) \\ \text{s.a} \quad & f_1(x) \leq \varepsilon_1 \\ & f_2(x) \leq \varepsilon_2 \\ & x \in S \end{aligned}$$

con $\varepsilon_i^{min} \leq \varepsilon_i \leq \varepsilon_i^{max}$, $i = 1, 2$.

Algorithm 2 Aproximación del conjunto eficiente para $k = 2$ usando el método de restricción ε

Entrada: Distancia $\alpha > 0$, distancia inicial $\beta \in (0, \alpha)$.

1. Resolver $\min_{x \in S} f_2(x)$ con solución minimal x^1 .

2. Definir $\varepsilon^2 := f_1(x^1) - \beta$ y $l := 2$.

3. Resolver $\min_{x \in S} f_1(x)$ con solución minimal x^E .

4. **while** $\varepsilon^l \geq f_1(x^E)$ **do**

5. Resolver

$$\begin{array}{ll} \min & f_2(x) \\ \text{s.a.} & f_1(x) \leq \varepsilon^l \\ & x \in S \end{array}$$

con solución x^l y multiplicador de Lagrange μ^l .

6. Definir

$$\varepsilon^{l+1} := \varepsilon^l - \frac{\alpha}{\sqrt{1 + (\mu^l)^2}}$$

y $l := l + 1$

7. **end while**

Salida: El conjunto $A := \{f(x^1), f(x^2), \dots, f(x^E)\}$ que es la aproximación de ε , el conjunto eficiente.

Ahora veremos la relación entre el método de sumas ponderadas y el método de restricción ε .

Teorema 4.2.6 *Sea $x^* \in S$ una solución del problema de sumas ponderadas (4.1) y $0 \leq w \in \mathbb{R}^k$ el correspondiente vector de ponderación. Entonces*

1. *Si $w_l > 0$, x^* es una solución del problema de restricción ε para f_l como función objetivo y $\varepsilon_j = f_j(x^*)$ para $j = 1, \dots, k$, $h \neq l$.*
2. *Si x^* es la única solución del problema de ponderación, entonces x^* es una solución del problema de restricción ε cuando $\varepsilon_j = f_j(x^*)$ para $j = 1, \dots, k$, $j \neq l$ y para cada f_l , $l = 1, \dots, k$, como la función objetivo.*

Demostración:

Sea $x^* \in S$ una solución del problema de ponderación para $0 \leq w \in \mathbb{R}^k$ un vector de ponderación.

(1) Suponemos que $w_l > 0$. Tenemos que

$$\sum_{i=1}^k w_i f_i(x) \geq \sum_{i=1}^k w_i f_i(x^*)$$

para todo $x \in S$.

Supongamos que x^* no es una solución del problema de restricción ε . Entonces existe un punto $\hat{x} \in S$ tal que $f_l(\hat{x}) < f_l(x^*)$ y $f_j(\hat{x}) \leq f_j(x^*)$ cuando $j = 1, \dots, k$, $j \neq l$. Suponemos que $w_l > 0$ y $w_o \geq 0$ si $i \neq l$. Ahora se tiene que

$$0 > w_l(f_l(\hat{x}) - f_l(x^*)) + \sum_{i=1, i \neq l}^k w_i(f_i(\hat{x}) - f_i(x^*))$$

lo que contradice la desigualdad anterior. Entonces x^* es una solución del problema de restricción ε .

(2) Si x^* es la única solución del problema de sumas ponderadas, entonces para todo $x \in S$

$$\sum_{i=1}^k w_i f_i(x^*) < \sum_{i=1}^k w_i f_i(x)$$

si existiera una función objetivo f_l tal que x^* no resuelve el problema de restricción ε cuando f_l es minimizada, entonces podemos encontrar una solución $\bar{x} \in S$ tal que $f_l(\bar{x}) < f_l(x^*)$ y $f_i(\bar{x}) \leq f_i(x^*)$ si $j \neq l$. Esto significa que para algún $w \geq 0$

$$\sum_{i=1}^k w_i f_i(\bar{x}) \leq \sum_{i=1}^k w_i f_i(x^*)$$

Lo que contradice la desigualdad anterior. Entonces x^* es una solución del problema de restricción ε para toda f_l para ser minimizada. \square

Por último, enunciaremos un teorema válido para problemas convexos.

Teorema 4.2.7 *Dado un problema de optimización multiobjetivo convexo, si $x^* \in S$ es solución del problema de restricción ε (4.2) para alguna función dada f_l que será minimizada y $\varepsilon_j = f_l(x^*)$ para $j = 1, \dots, k, j \neq l$, entonces existe un vector de ponderación $0 \leq w \in \mathbb{R}^k, \sum_{i=1}^k w_i = 1$, tal que x^* también es solución del problema de ponderación (4.1).*

La demostración del teorema (4.2.7) puede consultarse en [13] página 88.

Teóricamente, cada solución Pareto optimal de cualquier problema de optimización multiobjetivo se puede encontrar mediante el método de restricción ε al alterar los límites superiores y la función que se minimizará, sin importar que el problema sea o no convexo. Computacionalmente, el método de restricción ε es más laborioso que el método de sumas ponderadas ya que se aumenta el número de restricciones, pero puede ser difícil especificar límites superiores apropiados para las funciones objetivo.

4.3. Métricas ponderadas

Obtendremos diferentes soluciones al alterar los coeficientes de ponderación w_i , suponiendo $w_i \geq 0$ para todo $i = 1, \dots, k$ y $\sum_{i=1}^k w_i = 1$ en la métrica L_p ponderada y en la métrica L_∞ ponderada ó métrica de Tchebycheff ponderada. El problema L_p ponderado para minimizar distancias es de la forma:

$$\begin{aligned} \min & \left(\sum_{i=1}^k w_i |f_i(x) - z_i^*|^p \right)^{1/p} \\ \text{s.a.} & x \in S \end{aligned} \quad (4.6)$$

para $1 \leq p < \infty$ y siendo z_i^* la componente i -ésima del vector objetivo ideal. El problema ponderado de Tchebycheff es de la forma:

$$\begin{aligned} \min & \max_{i=1, \dots, k} [w_i |f_i(x) - z_i^*|] \\ \text{s.a.} & x \in S \end{aligned} \quad (4.7)$$

Este problema fue introducido originalmente por Bowman [1] en 1976.

Los signos de valor absoluto pueden eliminarse debido a la definición de vector objetivo ideal, si este se conoce globalmente.

Si $p = 1$, la suma ponderada de las desviaciones es minimizada y el problema es igual al problema de sumas ponderadas (4.1) salvo por una constante (si z^* es conocido globalmente). Si $p = 2$, tenemos un método de cuadrados mínimos. Cuando p se va incrementando, la minimización de la desviación mayor se vuelve

más y más importante. Finalmente, cuando $p = \infty$, lo único que importa es la desviación relativa ponderada de una función objetivo.

El problema ponderado de Tchebycheff (4.7) es no diferenciable. Puede resolverse de manera diferenciable siempre que las funciones objetivo y las restricciones sean diferenciables y que z^* sea conocido globalmente. En este caso, se tiene el siguiente problema:

$$\begin{aligned} \min & \quad \alpha \\ \text{s.a.} & \quad \alpha \geq w_i(f_i(x) - z_i^*) \quad \forall i = 1, \dots, k \\ & \quad x \in S \end{aligned} \quad (4.8)$$

Se resuelve, donde tanto $x \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ son variables.

Veremos a continuación algunos resultados relacionados al método de métricas ponderadas. Supondremos que conocemos el vector objetivo ideal z^* globalmente.

Teorema 4.3.1 *La solución del problema L_p ponderado (5.27) es Pareto optimal si sucede alguna de estas afirmaciones:*

1. *La solución es única.*
2. *Todos los coeficientes de ponderación son positivos.*

Demostración:

Esta prueba es análoga a las realizadas en el método de sumas ponderadas, teniendo en cuenta que, si x^* , solución del problema L_p ponderado no fuese Pareto optimal, entonces existiría otro punto $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y $f_j(x) < f_j(x^*)$ para al menos un índice j . Ahora $(f_i(x) - z_i^*)^p \leq (f_i(x^*) - z_i^*)^p$ para todo i y $(f_j(x) - z_j^*)^p < (f_j(x^*) - z_j^*)^p$. De esto tenemos

$$\sum_{i=1}^k (f_i(x) - z_i^*)^p < \sum_{i=1}^k (f_i(x^*) - z_i^*)^p$$

y al elevar a ambos miembros a $1/p$ la desigualdad se mantiene, pero esto sería una contradicción, entonces resulta que x^* debe ser Pareto optimal, y con esto se completa la prueba. \square

Teorema 4.3.2 *La solución del problema ponderado de Tchebycheff (4.7) es débilmente Pareto optimal si todos los coeficientes de ponderación son positivos.*

Demostración:

Sea $x^* \in S$ solución del problema ponderado de Tchebycheff. Supongamos que x^* no es débilmente Pareto optimal. En este caso, existe un punto $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$. Esto implica que $f_i(x) - z_i^* <$

$f_i(x^*) - z_i^*$ para todo i y al mismo tiempo que $w_i(f_i(x) - z_i^*) < w_i(f_i(x^*) - z_i^*)$ siendo $w_i > 0$ para todo i . Entonces x^* no puede ser solución del problema ponderado de Tchebycheff, lo que nos lleva a una contradicción, entonces x^* debe ser débilmente Pareto optimal. \square

Teorema 4.3.3 *EL problema ponderado de Tchebycheff (4.7) tiene al menos una solución Pareto optimal.*

Demostración:

Supongamos que no hay ninguna solución Pareto optimal del problema ponderado de Tchebycheff. Dada $x^* \in S$ una solución óptima del problema de Tchebycheff, como asumimos que no hay ninguna solución Pareto optimal, entonces existe una solución $x \in S$ que no es óptima para el problema de Tchebycheff pero para que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ y $f_j(x) < f_j(x^*)$ para al menos un j .

Tenemos ahora que $f_i(x)z_i^* \leq f_i(x^*) - z_i^*$ para todo i con desigualdad estricta para al menos un índice j , también $w_i(f_i(x)z_i^*) \leq w_i(f_i(x^*) - z_i^*)$ para todo i y además $\max[f_i(x)z_i^*] \leq \max[f_i(x^*) - z_i^*]$. Como x^* es solución óptima del problema de ponderación de Tchebycheff, x tiene que ser una solución optimal también, esta contradicción completa la prueba. \square

Como consecuencia directa de este teorema tenemos el siguiente corolario.

Corolario 4.3.4 *Si el problema ponderado de Tchebycheff (4.7) tiene solución única entonces esta es Pareto optimal.*

La convexidad del problema de optimización multiobjetivo es necesaria para garantizar que cada solución Pareto optimal pueda ser encontrada por el problema L_p ponderado (5.27). El siguiente teorema nos mostrará que toda solución Pareto optimal puede ser encontrada utilizando el problema ponderado de Tchebycheff (4.7).

Teorema 4.3.5 *Sea $x^* \in S$ Pareto optimal. Entonces existe un vector de ponderación $0 < w \in \mathbb{R}^k$ tal que x^* es solución del problema ponderado de Tchebycheff (4.7), cuando el punto de referencia es el vector objetivo utópico z^{**} .*

Demostración:

Sea $x^* \in S$ Pareto optimal. Supongamos que no existe un vector de ponderación $w > 0$ tal que x^* sea solución del problema ponderado de Tchebycheff. Sabemos que $f_i(x) > z_i^{**}$ para todo $i = 1, \dots, k$, para todo $x \in S$. Ahora, elegimos $w_i = \beta / (f_i(x^*) - z_i^{**})$ para todo $i = 1, \dots, k$ siendo $\beta > 0$ algún factor de normalización.

Si x^* no es solución del problema ponderado de Tchebycheff, entonces existe

otro punto $x^o \in S$ tal que es solución del problema ponderado de Tchebycheff, esto significa que

$$\begin{aligned} \max_i [w_i(f_i(x^o) - z_i^{**})] &< \max_i [w_i(f_i(x^*) - z_i^{**})] \\ &= \max_i \left[\frac{\beta}{f_i(x^*) - z_i^{**}} (f_i(x^*) - z_i^{**}) \right] = \beta \end{aligned}$$

Entonces $w_i(f_i(x^o) - z_i^{**}) < \beta$, y después de simplificar la expresión tenemos

$$f_i(x^o) \leq f_i(x^*)$$

para todo $i = 1, \dots, k$, lo que nos lleva a una contradicción del hecho que x^* es Pareto optimal. \square

El teorema anterior suena muy prometedor para el problema ponderado de Tchebycheff. Desafortunadamente, esta no es toda la verdad. Como vimos antes en el teorema (4.3.2) el problema ponderado de Tchebycheff encuentra soluciones débilmente Pareto optimal, por lo tanto, encuentra las soluciones Pareto optimal, pero se necesita un cálculo auxiliar para poder identificar los puntos débiles. Por otro lado, el problema L_p ponderado con $1 \leq p < \infty$, produce soluciones Pareto optimal, pero no necesariamente las encuentra a todas.

Otro factor importante es la selección del valor para el exponente p , una pauta a tener en cuenta es que para un número menor de funciones objetivo se recomienda seleccionar mayores valores de p .

Particularmente, el método de las métricas ponderadas de Tchebycheff son métodos populares para obtener soluciones débilmente Pareto optimal. Funcionan tanto para problemas convexos como no convexos (a diferencia del método de las sumas ponderadas) y la alteración de los parámetros es más simple que en el método de restricción ε . La desventaja de este método es que, si bien encuentra todas las soluciones Pareto optimal, junto con ellas encuentra soluciones débilmente Pareto optimal.

En lo que respecta a las métricas L_p ponderadas, las soluciones generadas son Pareto optimal, pero no necesariamente se encuentran todas.

Para encontrar la aproximación del frente de Pareto que estamos buscando, al igual que en el método de sumas ponderadas lo que hacemos es variar los coeficientes de ponderación para obtener diferentes soluciones, como se ve en el siguiente algoritmo (3)

Algorithm 3 Aproximación del conjunto eficiente para $k = 2$ usando el método de métricas ponderadas

Entrada: Peso $1 \leq p < \infty$, r = cantidad de coeficientes de ponderación w .

1. Generar los r coeficientes de ponderación w_1 y w_2 tales que $w_1 + w_2 = 1$ y $w_i \geq 0$.

2. Resolver

$$\min_{x \in S} (w_1 |f_1(x) - z_1^*|^{1/p} + w_2 |f_2(x) - z_2^*|^{1/p})^p$$

con solución x^i con $i = 1, \dots, r$ siendo r la cantidad de coeficientes generados en 1.

Salida: El conjunto $A := \{f(x^1), f(x^2), \dots, f(x^r)\}$ que es la aproximación del frente de Pareto.

La desventaja del método de métricas L_p ponderadas con $1 \leq p < \infty$ de no encontrar todas las soluciones del problema lo veremos más claramente en el capítulo 5 con algunos problemas test que hemos implementado.

4.4. Escalarización por combinaciones convexas

Víctor Pereyra en 2009 y 2013 ([10], [11]) introdujo una nueva idea para problemas multiobjetivo sin restricciones que produce una representación discreta equiespaciada del frente de Pareto usando pocas evaluaciones funcionales. El enfoque utilizado allí consistió la implementación del método de Newton para resolver las condiciones de optimización de primer orden. Comenzando desde un punto óptimo de los objetivos individuales (ellos mismos puntos Pareto optimales) y utilizando la continuación homotópica de longitud de arco, el método resuelve una sucesión de problemas de optimización con restricciones que implican las variables del problema, la variable de homotopía λ y una restricción no lineal añadida que requiere que el vector de valores objetivo esté a una distancia dada del valor anterior. Esta es la clave para producir un frente Pareto equiespaciado.

Aunque la escalarización por combinaciones convexas tiene algunas limitaciones, es uno de los enfoques más comunes y naturales utilizado para generar puntos Pareto.

La dificultad hasta ahora ha sido que no es obvio cómo elegir los valores de λ con el fin de generar un frente equiespaciado. Esto al igual que en el caso del método de restricción ε , nuevamente la dificultad no está en encontrar una solución sino en encontrar todas las soluciones necesarias para generar una aproximación aceptable del frente de Pareto completo, o al menos, una parte considerablemente útil del mismo para entregar al tomador de decisiones.

Algoritmo de optimización multiobjetivo

Por simplicidad consideraremos el problema bi-objetivo

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & g(x) \leq 0 \\ & x \in D \end{aligned} \tag{4.9}$$

donde $x \in \mathbb{R}^n$, $f \in \mathbb{R}^2$, $g \in \mathbb{R}^m$, D es un conjunto convexo y las desigualdades son aplicadas componente a componente. Asumimos que las funciones no lineales f y g son suaves. Introduciremos la siguiente función objetivo a valores escalares

$$f(x, \lambda) = (1 - \lambda)f_1(x) + \lambda f_2(x) \tag{4.10}$$

donde $0 \leq \lambda \leq 1$ y las mismas constantes del problema original. La homotopía estándar comienza con $\lambda = 0$ y luego los pasos λ de alguna manera, al resolver los subproblemas sucesivos para obtener un muestreo discreto del conjunto de Pareto. Una crítica frecuente y válida de este método es que no hay manera segura de obtener un muestreo uniforme con él porque la parametrización del frente de Pareto por λ es generalmente un mapa desconocido muy no lineal.

Sean x_1^* y x_2^* los minimizadores de $f_1(x)$ y $f_2(x)$ respectivamente. Definimos $x_0 = x_1^*$ y $x_{l+1} = x_2^*$ los puntos extremos del conjunto de Pareto que vamos a construir, donde l es el número deseado de puntos en el frente de Pareto que une los puntos extremos.

Dada una colección de puntos $f_k = f(x_k)$, $k = 0, 1, \dots, l+1$. Si los puntos son muestras de una curva suave, entonces cuando el espacio entre los puntos tiende a cero.

Ahora definimos

$$\gamma = \alpha \|f(x_0) - f(x_{l+1})\| / l \tag{4.11}$$

donde la distancia entre las imágenes de los minimizadores está multiplicada por un factor $\alpha > 1$ es una estimación de la longitud total del acorde del frente de Pareto (que representa la curvatura).

En el improbable caso en el que $f(x_0) = f(x_{l+1})$ ya encontramos todo nuestro frente de Pareto que está formado por un sólo punto. Si esto no ocurre, $\gamma \neq 0$ y le imponemos al problema la siguiente restricción de equiespaciado:

$$\|f(x) - f_{prev}\|^2 = \gamma^2 \tag{4.12}$$

donde f_{prev} es el punto anterior que conseguimos en el proceso de homotopía. Entonces minimizaremos la función escalarizada sujeta a todas las constantes con λ como una variable adicional:

$$\begin{aligned} \min \quad & (1 - \lambda)f_1(x) + \lambda f_2(x) \\ \text{s.a.} \quad & g(x) \leq 0 \\ & x \in D \\ & \|f(x) - f_{prev}\|^2 = \gamma^2 \end{aligned} \tag{4.13}$$

Siendo $f_{prev} = f(x_0)$, la solución que obtendremos será (x_1, λ_1) . Repetimos el proceso con $f_{prev} = f(x_1)$ obteniendo (x_2, λ_2) y se sigue el proceso. El correspondiente conjunto de Pareto discreto obtenido es definido por el conjunto $\{f(x_k), k = 0, \dots, l + 1\}$, donde los $\{x_k, k = 0, \dots, l + 1\}$ son Pareto. Si el proceso es exitoso, obtenemos una representación discreta del conjunto de Pareto con valores equiespaciados $f(x_k)$. Observe que con el uso intrínseco de esta parametrización, la parametrización λ del frente pierde importancia.

Si l es grande suficiente para hacer que γ sea relativamente pequeño, podemos argumentar heurísticamente que para cada problema

$$\begin{aligned}
 \min \quad & (1 - \lambda)f_1(x) + \lambda f_2(x) \\
 \text{s.a.} \quad & g(x) \leq 0 \\
 & x \in D \\
 & \|f(x) - f_{prev}\|^2 = \gamma^2
 \end{aligned} \tag{4.14}$$

en la homotopía es probable que iniciar el optimizador en el minimizador anterior es probable que dé un minimizador global para el problema actual.

La extensión a $p > 2$ objetivos es sencilla, aunque la complejidad aumenta considerablemente.

Algorithm 4 Aproximación del conjunto eficiente para $k = 2$ usando una combinación convexa de los objetivos

Entrada: Cantidad de puntos l , factor $\alpha > 1$.

1. Resolver $\min_{x \in S} f_1(x)$ con solución minimal x_1^* y $\min_{x \in S} f_2(x)$ con solución minimal x_2^* .

2. Definir $x_0 = x_1^*$, $x_{l+1} = x_2^*$ y $\lambda = 0$.

3. Calculo $\gamma = \alpha \|f(x_0) - f(x_{l+1})\|/l$.

4. **if** $\gamma \neq 0$ **do**

5. **for** $i = 1, \dots, l$ **do**

6. Resolver

$$\begin{array}{ll} \min_{x, \lambda} & (1 - \lambda)f_1(x) + \lambda f_2(x) \\ \text{s.a.} & g(x) \leq 0 \\ & \|f(x) - f_{prev}\|^2 = \gamma^2 \\ & x \in S \end{array}$$

 con solución (x_i, λ_i)

7. **end for**

Salida: El conjunto $A := \{f(x_0), f(x_1), \dots, f(x_{l+1})\}$ que es la aproximación del frente de Pareto equiespaciado.

Capítulo 5

Resultados numéricos

En este capítulo realizamos experimentos numéricos con un conjunto de problemas test para estudiar los métodos de **sumas ponderadas**, **restricción ε** , **escalarización por combinaciones convexas** y **métricas ponderadas**, presentados en el capítulo 4.

Todos los algoritmos fueron implementados en lenguaje Python versión 3.8. Fueron ejecutados en una notebook acer aspire 5, con procesador 17-8550U de 1.80 GHz, 12 Gb de memoria RAM, y sistema operativo Windows 10 de 64 bits, procesador x64.

Primero retomaremos un ejemplo convexo en el que calculamos de manera explícita el frente de Pareto en el capítulo 2. Luego en segundo lugar el ejemplo de Schaffer que en el que aproximamos un frente de Pareto convexo. En tercer lugar, analizamos el ejemplo de Hazen, que también tiene un frente de Pareto convexo, pero este problema a diferencia del problema de Schaffer, es un problema irrestricto. Luego estudiamos el problema de Fonseca - Fleming que presenta un conjunto de Pareto no convexo y podemos observar que el método de sumas ponderadas no es el más conveniente en este tipo de problemas y para terminar con los problemas test implementamos el problema de Poloni, donde el conjunto de Pareto es no convexo y no conexo.

5.1. Ejemplo 4

Consideramos nuevamente el ejemplo 4 del capítulo 2, página 8, que era un problema con dos funciones objetivo, que queremos minimizar simultáneamente sobre la recta real no negativa.

$$f_1(x) = \sqrt{x+1}$$
$$f_2(x) = x^2 - 4x + 5 = (x+2)^2 - 1$$

Y queremos resolver el siguiente problema de optimización:

$$\min_{x \geq 0} (f_1(x), f_2(x))$$

En el capítulo 2, generamos paramétricamente el frente de Pareto que es el siguiente, podemos verlo en color rosa en la siguiente figura:

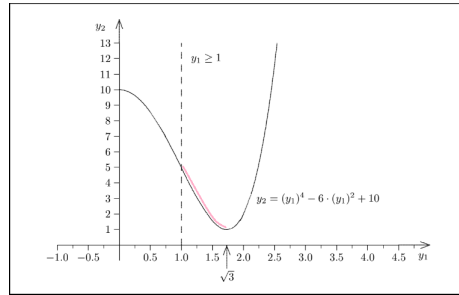


Figura 5.1: Frente de Pareto.

Ahora, resolveremos este problema utilizando los distintos métodos que estudiamos antes utilizando la parametrización del frente de Pareto que conseguimos explícitamente.

Sumas ponderadas

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\min_{x \geq 0} w_1 \sqrt{x+1} + w_2 (x^2 - 4x + 5) \quad (5.1)$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$. Variando los parámetros w_1 y w_2 y resolviendo el problema de la suma ponderada se pueden generar aproximaciones del conjunto de Pareto.

Para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.2)$$

Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.2), siendo la línea celeste la expresión obtenida antes y cada una de las estrellas la aproximación generada por el método de sumas ponderadas.

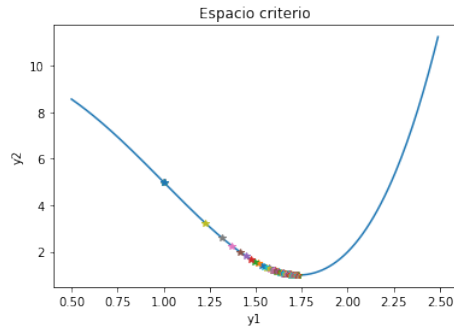


Figura 5.2: Aproximación generada por Sumas Ponderadas.

Restricción ε

Aplicamos el algoritmo (2) con diferentes α y β y conseguimos la siguiente aproximación del frente de Pareto, ver Figura (5.3), nuevamente la línea naranja es la expresión del conjunto de Pareto calculada explícitamente, las estrellas las aproximaciones generadas por el método de restricción ε y esta vez agregamos un + que es el vector objetivo ideal que, como dijimos antes, es un punto de referencia como cotas inferiores del frente de Pareto.

En la figura de arriba a la izquierda tenemos 40 puntos, en la de arriba a la derecha 11 puntos, en la de abajo a la izquierda 8 puntos y en la de abajo a la derecha 102 puntos. Para generar estos puntos utilizamos los siguientes valores.

	Gráfico 1	Gráfico 2	Gráfico 3	Gráfico 4
# Puntos	40	11	8	102
α	0.1	0.4	0.4	0.04
β	0.1	0.1	0.4	0.04

Tabla 5.1: Valores utilizados para generar la imagen (5.3)

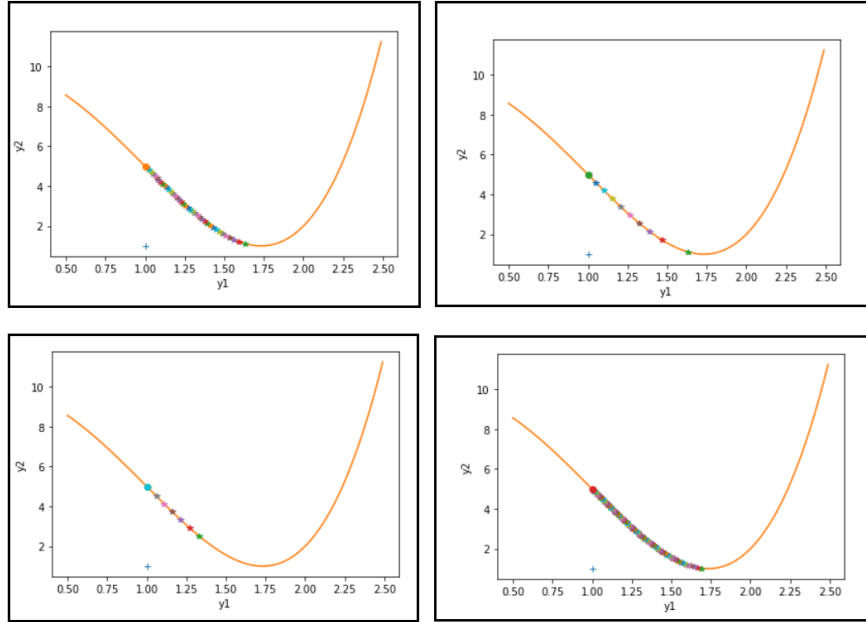


Figura 5.3: Aproximaciones generadas por Restricción ϵ

Métricas ponderadas

Para aplicar el método de las métricas ponderadas resolveremos el problema L_p ponderado:

$$\begin{aligned} \min \quad & (w_1|f_1(x) - z_1^*|^p + w_2|f_2(x) - z_2^*|^p)^{1/p} \\ \text{s.a.} \quad & x \in S \end{aligned} \quad (5.3)$$

para $1 \leq p < \infty$ y siendo z_i^* la componente i -ésima del vector objetivo ideal.

Al igual que en el método de sumas ponderadas, para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.4)$$

Usamos el algoritmo (3) para aproximar los frentes de Pareto. Vamos a variar el p tomando $p = 1$, $p = 2$, $p = 4$ y $p = 10$ sólo utilizamos el problema L_p ponderado ya que genera puntos Pareto optimal a diferencia del problema ponderado de Tchebycheff genera puntos débilmente Pareto optimal.

Notemos que cómo vimos en el capítulo 4, el método de métricas ponderadas nos asegura conseguir soluciones Pareto optimal, pero puede que no encuentre

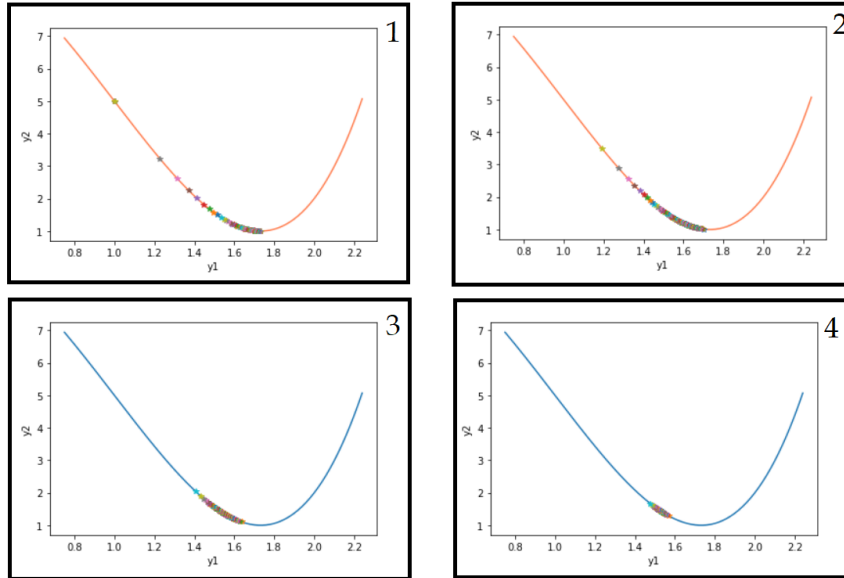


Figura 5.4: Aproximaciones generadas por métricas ponderadas

todas las soluciones como vemos en los casos $p = 4$ y $p = 10$ que corresponde a los frentes 3 y 4 de la Figura (5.4).

Escalarización por combinaciones convexas

Vamos a usar el algoritmo (4) variando la cantidad de puntos l que vamos a generar y el factor α que estima la curvatura total del conjunto de Pareto.

En la figura de arriba a la izquierda tenemos 30 puntos, al igual que en la de arriba a la derecha y en las figuras de abajo hay 60 puntos. Para generar estos puntos utilizamos los siguientes valores.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4
# Puntos	30	30	60	60
α	1.5	3	1.5	3

Tabla 5.2: Valores utilizados para generas la imagen (5.3)

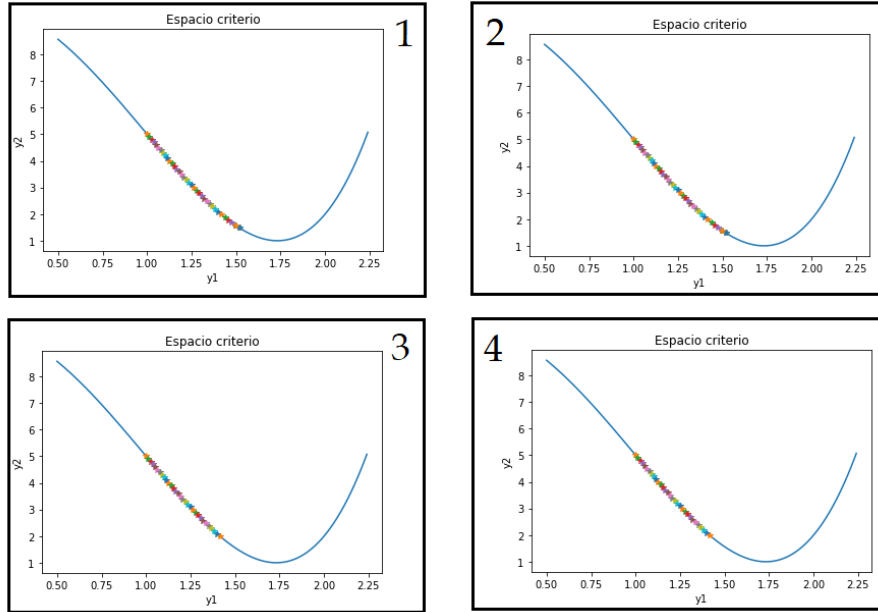


Figura 5.5: Aproximaciones generadas por combinaciones convexas

Este problema test ([2]) nos ayuda a ver que efectivamente nuestros cuatro algoritmos aproximan al conjunto de Pareto, aprovechando que teníamos parametrizado a nuestro frente de Pareto.

Ahora vamos a ver otros problemas test y aplicar a cada problema los cuatro algoritmos. Analizaremos como resuelve cada algoritmo a cada ejemplo.

5.2. Problema de Schaffer

En este problema, se garantiza que el conjunto de Pareto es el gráfico de una función convexa, si las funciones objetivo y el conjunto de restricciones son convexos.

El problema de Schaffer [14] es el siguiente:

$$\begin{aligned}
 \min \quad & f(x) = \begin{cases} f_1(x) = x^2 \\ f_2(x) = (x - 2)^2 \end{cases} \\
 \text{s.a.} \quad & -\mathcal{A} \leq x \leq \mathcal{A}
 \end{aligned} \tag{5.5}$$

para valores de \mathcal{A} variando entre 10 y 10^5 . A medida que \mathcal{A} aumenta su valor la dificultad del problema aumenta.

Al ser un problema convexo, los cuatro métodos que utilizamos no deberían tener ninguna complicación para generar el frente de Pareto, salvo quizás el método de métricas ponderadas que puede que no encuentre todas las soluciones.

Como en este problema se puede variar el valor de \mathcal{A} entre 10 y 10^5 vamos a tomar $\mathcal{A} \in \{10, 10^2, 10^3, 10^4, 10^5\}$, por lo que repetiremos el análisis 5 veces.

Caso $\mathcal{A} = 10$

Sumas ponderadas

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1 x^2 + w_2 (x - 2)^2 \\ \text{s.a.} \quad & -10 \leq x \leq 10 \end{aligned} \tag{5.6}$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$. Variando los parámetros w_1 y w_2 y resolviendo el problema de la suma ponderada se pueden generar aproximaciones del conjunto de Pareto.

Para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \tag{5.7}$$

Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.6), siendo cada una de las estrellas la aproximación generada por el método de sumas ponderadas.

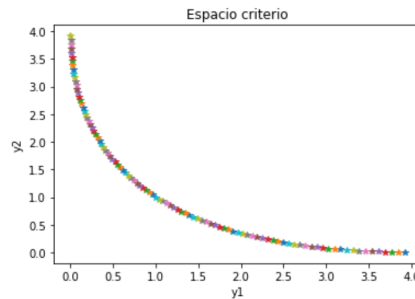


Figura 5.6: Aproximación generada por Sumas Ponderadas.

Caso $\mathcal{A} = 10^2$

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1 x^2 + w_2 (x - 2)^2 \\ \text{s.a.} \quad & -10^2 \leq x \leq 10^2 \end{aligned} \tag{5.8}$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$. Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.7).

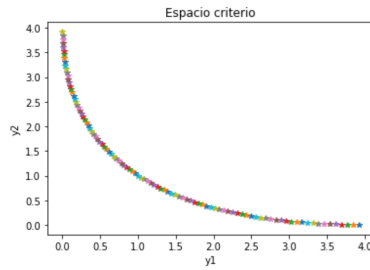


Figura 5.7: Aproximación generada por Sumas Ponderadas.

Caso $\mathcal{A} = 10^3$

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1 x^2 + w_2 (x - 2)^2 \\ \text{s.a.} \quad & -10^3 \leq x \leq 10^3 \end{aligned} \tag{5.9}$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$. Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.8).

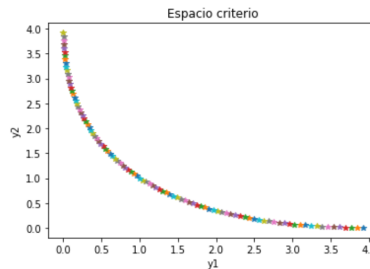


Figura 5.8: Aproximación generada por Sumas Ponderadas.

Caso $\mathcal{A} = 10^4$

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1 x^2 + w_2 (x - 2)^2 \\ \text{s.a.} \quad & -10^4 \leq x \leq 10^4 \end{aligned} \quad (5.10)$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$.

Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.9).

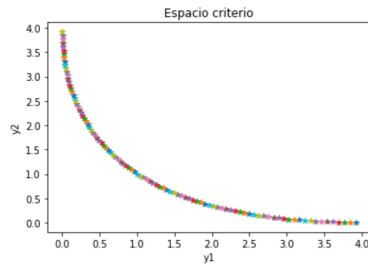


Figura 5.9: Aproximación generada por Sumas Ponderadas.

Caso $\mathcal{A} = 10^5$

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1 x^2 + w_2 (x - 2)^2 \\ \text{s.a.} \quad & -10^5 \leq x \leq 10^5 \end{aligned} \quad (5.11)$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$.

Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.10).

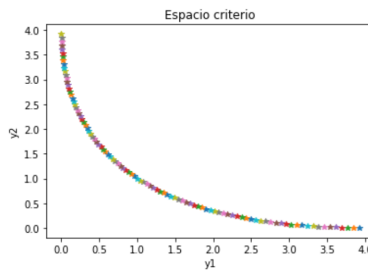


Figura 5.10: Aproximación generada por Sumas Ponderadas.

Restricción ε

Aplicamos el algoritmo (2) con diferentes α y β y conseguimos las siguientes aproximaciones del frente de Pareto, nuevamente, las estrellas son las aproximaciones generadas por el método de restricción ε del frente de Pareto.

Para todos los casos de \mathcal{A} podemos observar que en la figura de arriba a la izquierda tenemos 63 puntos, en la de arriba a la derecha 15 puntos, en la de abajo a la izquierda 15 puntos y en la de abajo a la derecha 160 puntos. Para generar estos puntos utilizamos los siguientes valores.

	Gráfico 1	Gráfico 2	Gráfico 3	Gráfico 4
# Puntos	63	15	15	160
α	0.1	0.4	0.4	0.04
β	0.1	0.1	0.4	0.04

Tabla 5.3: Valores utilizados para generar las imágenes

Podemos ver que a pesar de modificar los valores de \mathcal{A} los frentes de Pareto generados por este método son los siguientes:

Caso $\mathcal{A} = 10$

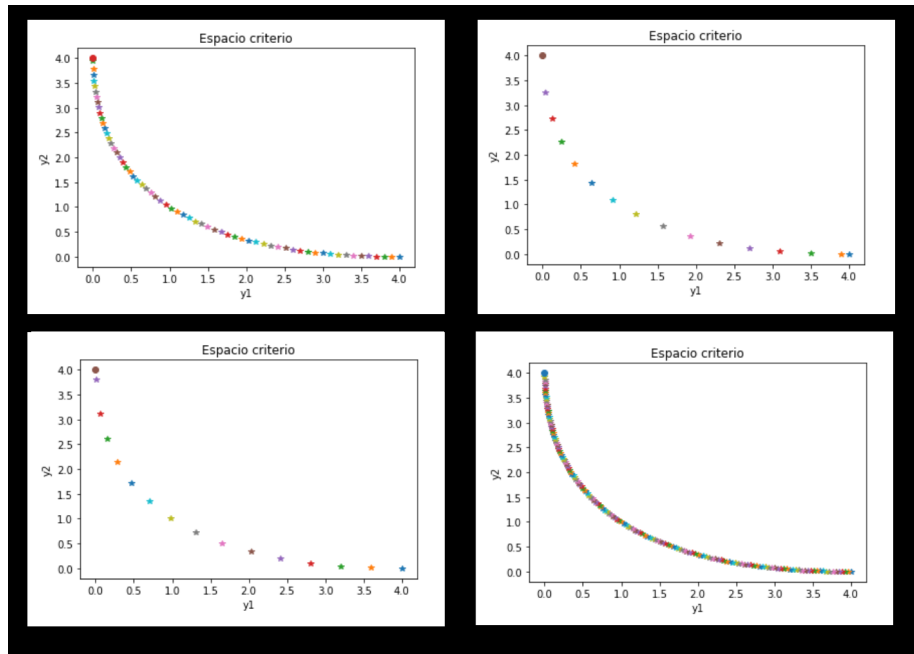


Figura 5.11: Aproximaciones generadas por Restricción ε

Caso $\mathcal{A} = 10^2$

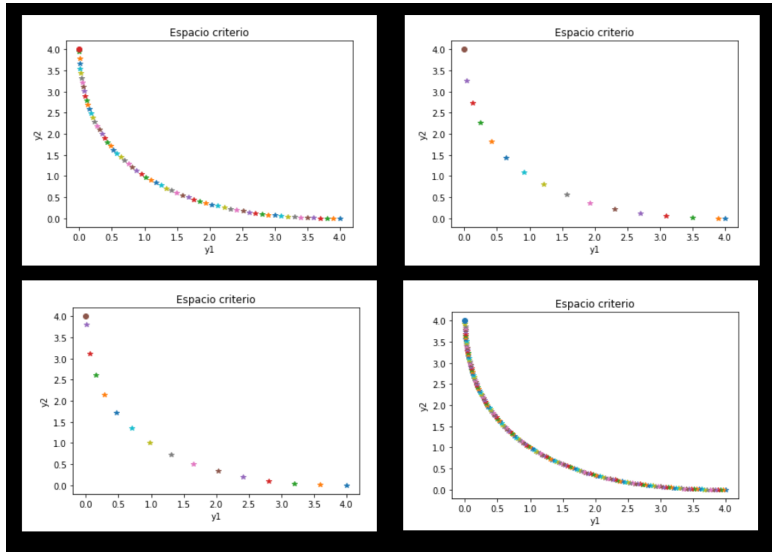


Figura 5.12: Aproximaciones generadas por Restricción ε

Caso $\mathcal{A} = 10^3$

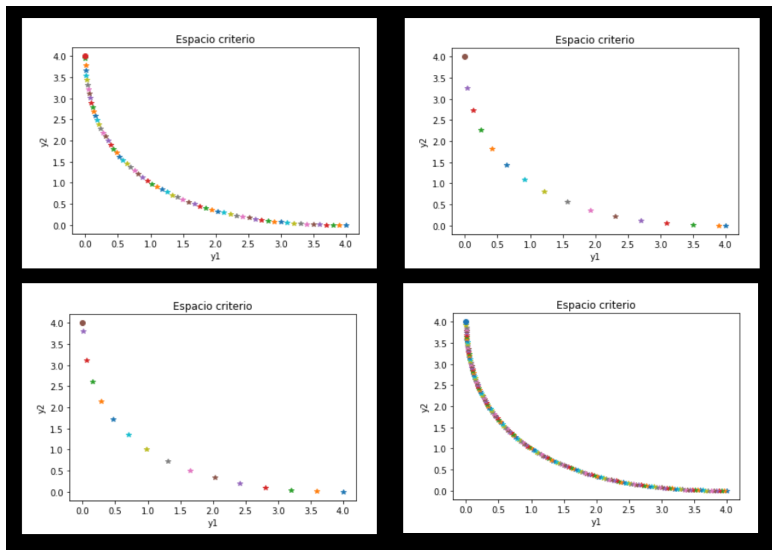


Figura 5.13: Aproximaciones generadas por Restricción ε

Caso $\mathcal{A} = 10^4$

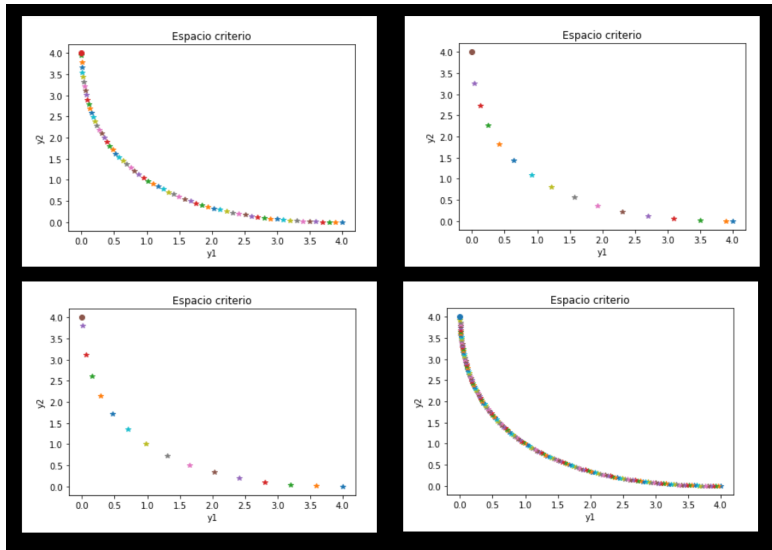


Figura 5.14: Aproximaciones generadas por Restricción ε

Caso $\mathcal{A} = 10^5$

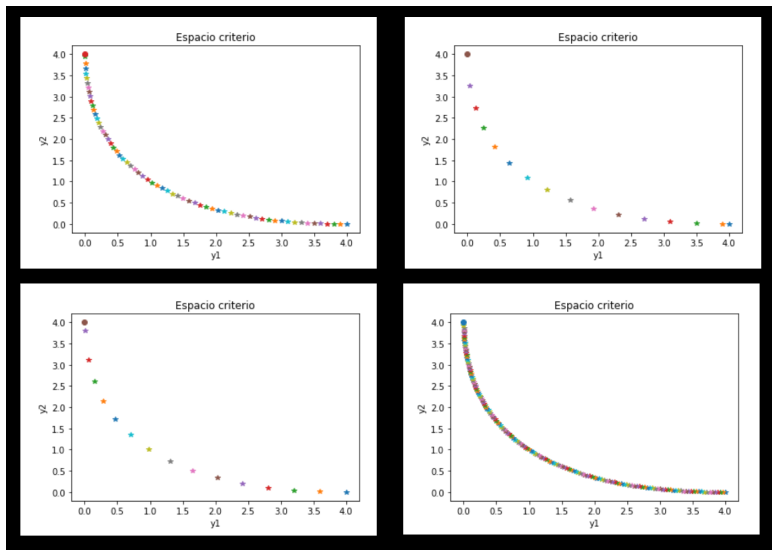


Figura 5.15: Aproximaciones generadas por Restricción ε

Escalarización por combinaciones convexas

Vamos a usar el algoritmo (4) variando la cantidad de puntos l que vamos a generar y el factor α que estima la curvatura total del conjunto de Pareto. En la figura de arriba a la izquierda tenemos 30 puntos, al igual que en la de arriba a la derecha y en las figuras de abajo hay 60 puntos. Para generar estos puntos utilizamos los siguientes valores.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4
# Puntos	30	30	60	60
α	1.5	3	1.5	3

Tabla 5.4: Valores utilizados para generar las imagenes

En todos los casos utilizaremos los mismos parámetros, sólo vamos a variar el valor de \mathcal{A} .

Caso $\mathcal{A} = 10$

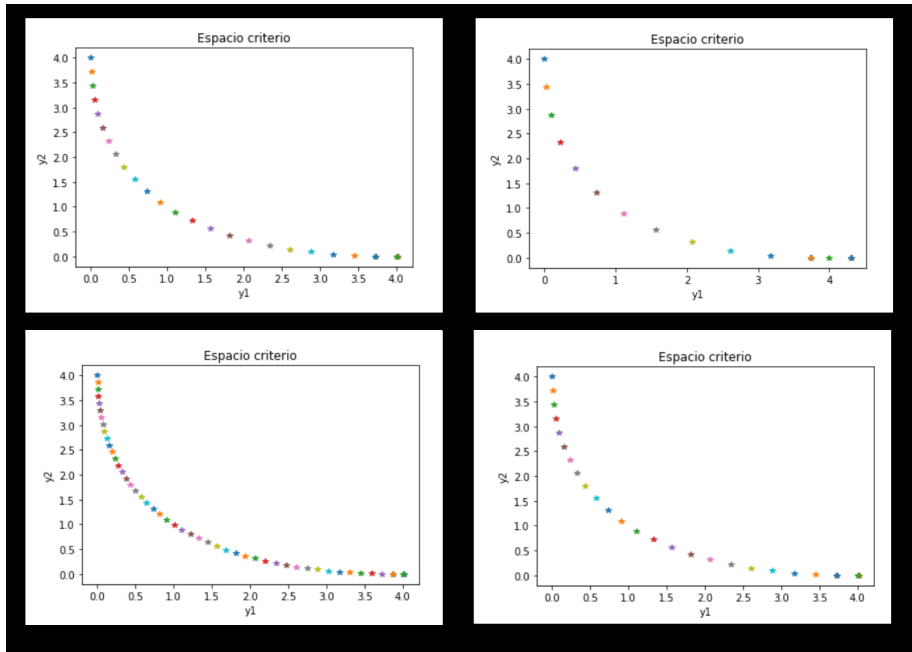


Figura 5.16: Aproximaciones generadas por combinaciones convexas

Caso $\mathcal{A} = 10^2$

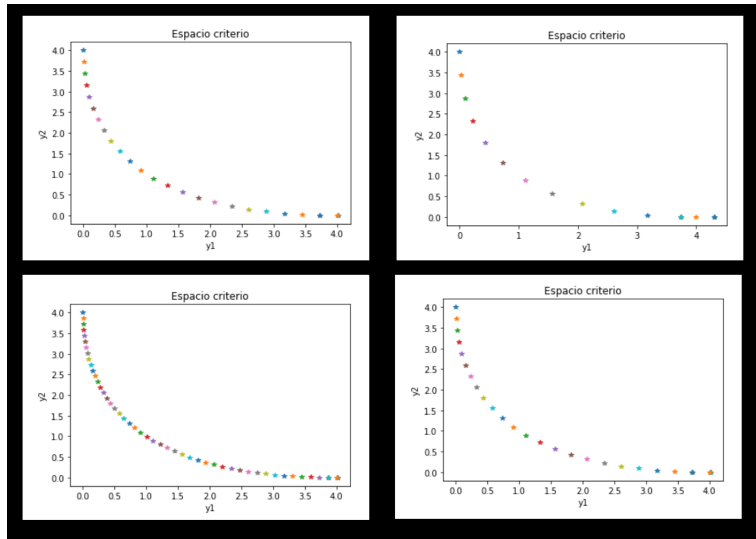


Figura 5.17: Aproximaciones generadas por combinaciones convexas

Caso $\mathcal{A} = 10^3$

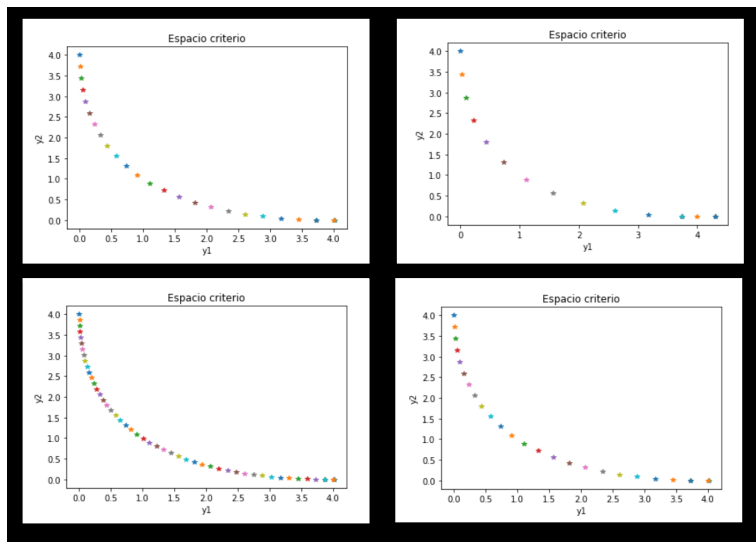


Figura 5.18: Aproximaciones generadas por combinaciones convexas

Caso $\mathcal{A} = 10^4$

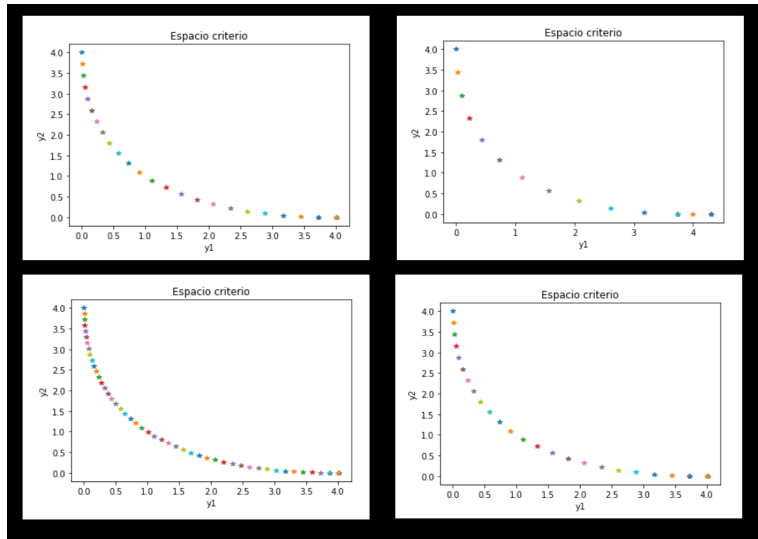


Figura 5.19: Aproximaciones generadas por combinaciones convexas

Caso $\mathcal{A} = 10^5$

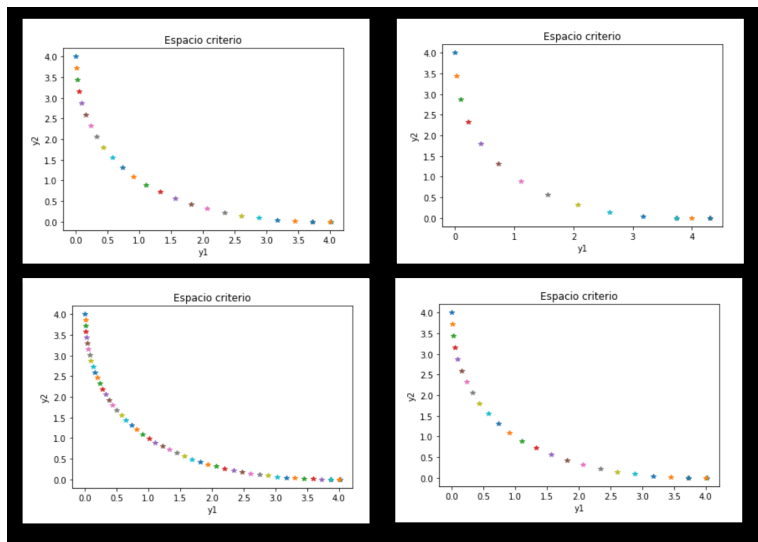


Figura 5.20: Aproximaciones generadas por combinaciones convexas

Métricas ponderadas

Para aplicar el método de las métricas ponderadas resolveremos el problema L_p ponderado:

$$\begin{aligned} \min \quad & (w_1|f_1(x) - z_1^*|^p + w_2|f_2(x) - z_2^*|^p)^{1/p} \\ \text{s.a.} \quad & -\mathcal{A} \leq x \leq \mathcal{A} \end{aligned} \quad (5.12)$$

para $1 \leq p < \infty$ y siendo z_i^* la componente i -ésima del vector objetivo ideal.

Al igual que en el método de sumas ponderadas, para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.13)$$

Vamos a variar el p tomando $p = 1$, $p = 2$, $p = 4$ y $p = 10$ sólo utilizamos el problema L_p ponderado, siguiendo el algoritmo (3), ya que genera puntos Pareto optimal a diferencia del problema ponderado de Tchebycheff genera puntos débilmente Pareto optimal.

Notemos que cómo vimos en el capítulo 4, el método de métricas ponderadas nos asegura conseguir soluciones Pareto optimal, pero puede que no encuentre todas las soluciones como vemos en los casos $p = 2$, $p = 4$ y $p = 10$ que corresponde a los frentes 2, 3 y 4 de las siguientes Figuras.

Caso $\mathcal{A} = 10$

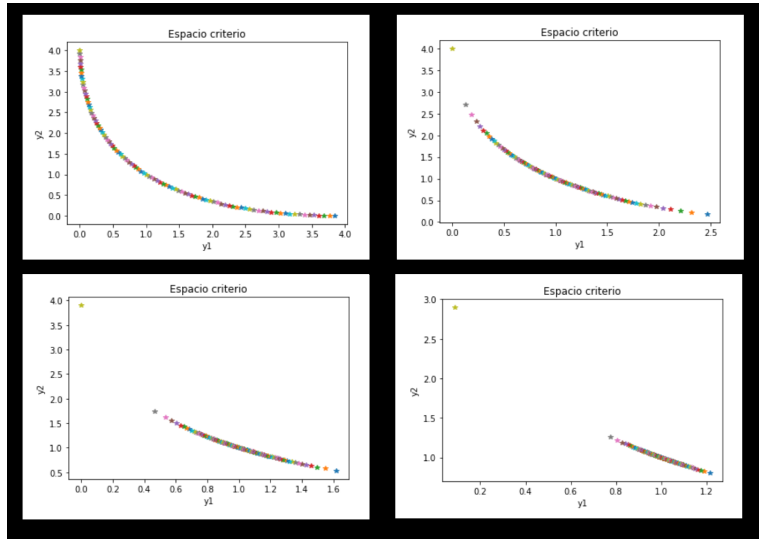


Figura 5.21: Aproximaciones generadas por métricas ponderadas

Caso $\mathcal{A} = 10^2$

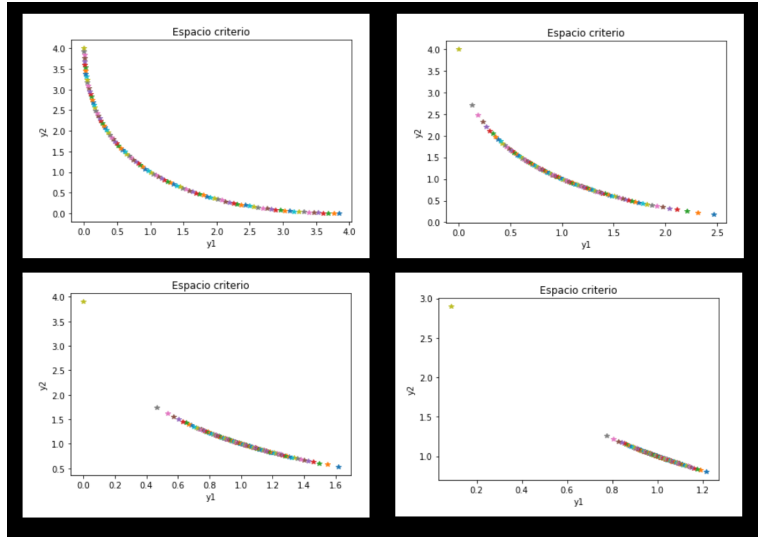


Figura 5.22: Aproximaciones generadas por métricas ponderadas

Caso $\mathcal{A} = 10^3$

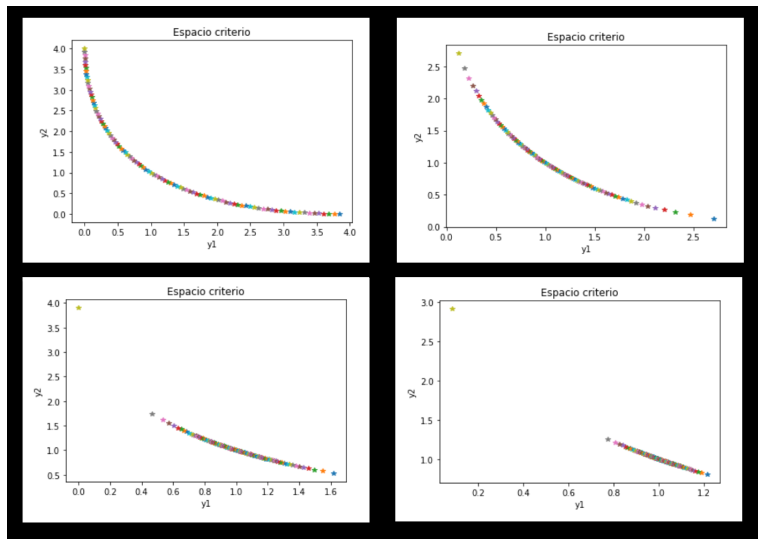


Figura 5.23: Aproximaciones generadas por métricas ponderadas

Caso $\mathcal{A} = 10^4$

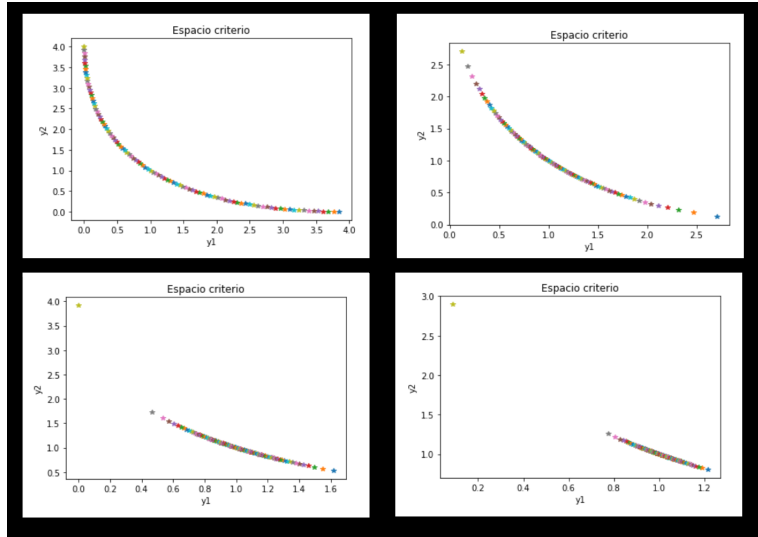


Figura 5.24: Aproximaciones generadas por métricas ponderadas

Caso $\mathcal{A} = 10^5$

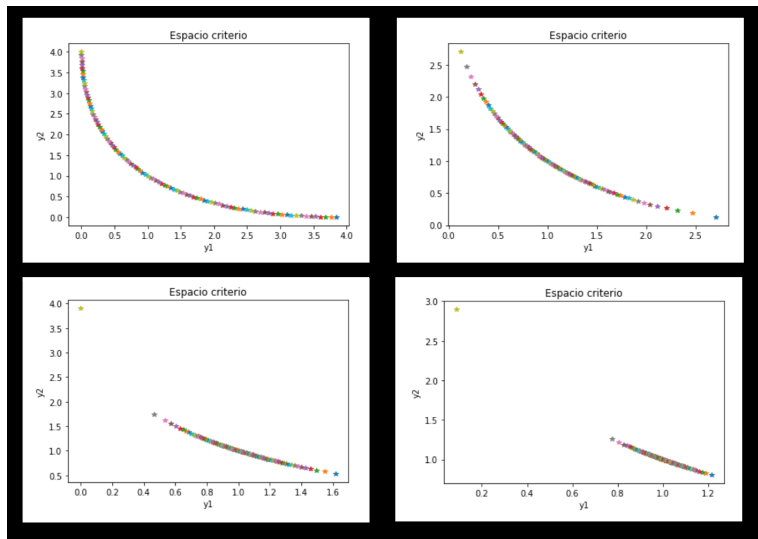


Figura 5.25: Aproximaciones generadas por métricas ponderadas

Este ejemplo es el que uno esperaría que al variar el valor de \mathcal{A} obtendría diferentes frentes de Pareto, podemos ver que siempre obtengo el mismo al ir variando los diferentes algoritmos que implementamos. En este caso vemos que los que mejor aproximan el frente de Pareto son la **suma ponderada** con 100 valores, el método de **restricción** ε con $\alpha = 0,1$ y $\beta = 0,1$, el algoritmo de **escalarización por combinaciones convexas** para 30 puntos con $\gamma = 1,5$ y el método de **métricas ponderadas** para $p = 1$.

5.3. Problema de Hazen

El problema de Hazen [7], al igual que el problema de Schaffer, es un problema con frente de Pareto convexo y conexo. Además, el problema de Hazen es un problema sin restricciones con funciones objetivo convexas.

El problema de Hazen es el siguiente:

$$\begin{aligned} \min \quad & f(x) = \begin{cases} f_1(x) = x_1^2 + \frac{x_2^2}{2} - 10x_2 - 100 \\ f_2(x) = \frac{x_1^2}{2} + x_2^2 - 10x_1 - 100 \end{cases} \\ \text{s.a.} \quad & x \in \mathbb{R}^2 \end{aligned} \quad (5.14)$$

Sumas ponderadas

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1(x_1^2 + \frac{x_2^2}{2} - 10x_2 - 100) + w_2(\frac{x_1^2}{2} + x_2^2 - 10x_1 - 100) \\ \text{s.a.} \quad & x \in \mathbb{R}^2 \end{aligned} \quad (5.15)$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$. Variando los parámetros w_1 y w_2 y resolviendo el problema de la suma ponderada se pueden generar aproximaciones del conjunto de Pareto.

Para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.16)$$

Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.26), siendo cada una de las estrellas la aproximación generada por el método de sumas ponderadas.

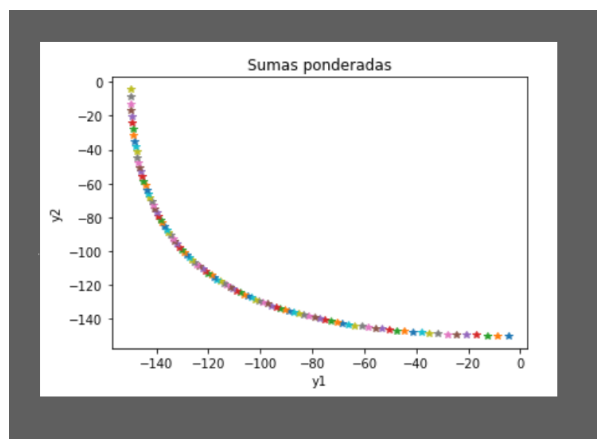


Figura 5.26: Aproximación generada por Sumas Ponderadas.

Restricción ε

Aplicamos el algoritmo (2) con diferentes α y β y conseguimos la siguiente aproximación del frente de Pareto, ver Figura (5.27), nuevamente, las estrellas son las aproximaciones generadas por el método de restricción ε . Esta vez, a diferencia de los valores de α y β que usamos en el problema de Schaffer, que eran $\alpha, \beta \in [0, 1]$, usamos valores mayores, ya que, con valores pequeños generábamos demasiados puntos y si bien estos ejemplos son simples, tenemos que tener en cuenta que este método a diferencia de los otros es más costoso computacionalmente, es decir, se necesitan realizar más operaciones, de hecho, este método necesita del orden de nm^{n-1} operaciones siendo n la cantidad de funciones objetivo y m la cantidad de soluciones factibles, mientras que, el método de sumas ponderadas necesita del orden de nm operaciones, siendo n y m los mismos que antes.

En la figura de arriba a la izquierda tenemos 2484 puntos, en la de arriba a la derecha 246 puntos, en la de abajo a la izquierda 23 puntos y en la de abajo a la derecha 23 puntos. Para generar estos puntos utilizamos los siguientes valores.

	Gráfico 1	Gráfico 2	Gráfico 3	Gráfico 4
# Puntos	2484	246	23	23
α	0.1	1.0	10.0	10.0
β	0.1	1.0	10.0	5.0

Tabla 5.5: Valores utilizados para generar la imagen (5.27)

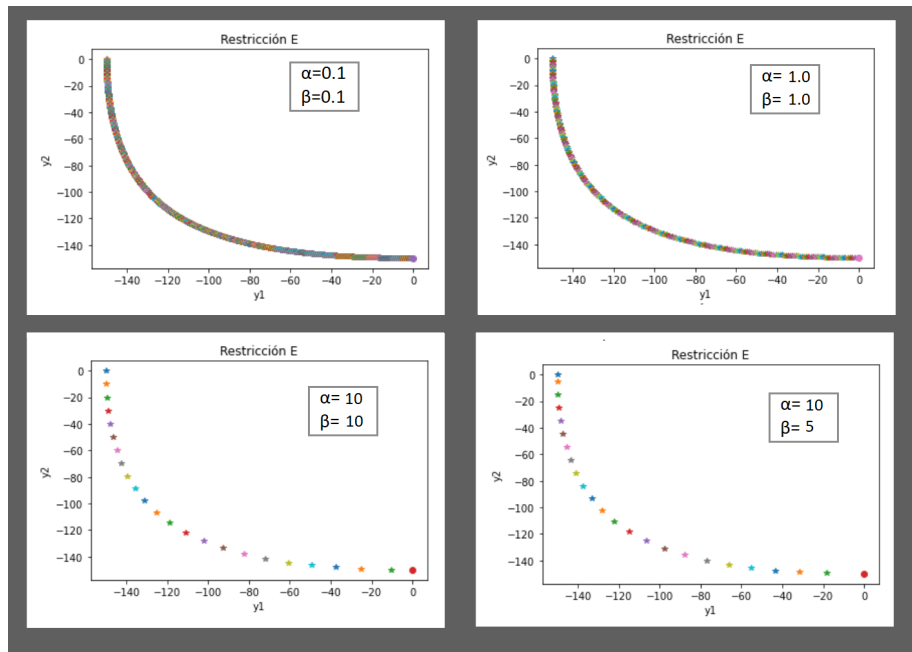


Figura 5.27: Aproximaciones generadas por Restricción ε

Escalarización por combinaciones convexas

Vamos a usar el algoritmo (4) variando la cantidad de puntos l que vamos a generar y el factor α que estima la curvatura total del conjunto de Pareto

En la figura de arriba a la izquierda tenemos 30 puntos, al igual que en la de arriba a la derecha y en las figuras de abajo hay 60 puntos. Para generar estos puntos utilizamos los siguientes valores. Notemos que este algoritmo no pudo encontrar todos los valores del frente de Pareto, a diferencia del método de sumas ponderadas y restricción ε que encontraron también el vector de decisión Pareto óptimo $(8.348, 0.471)$ que corresponde al vector objetivo Pareto óptimo $(-34.91, -148.41)$ que no está en ninguno de los frentes generados por el método de escalarización por combinaciones convexas.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4
# Puntos	30	30	60	60
α	1.5	3	1.5	3

Tabla 5.6: Valores utilizados para generas la imagen (5.28)

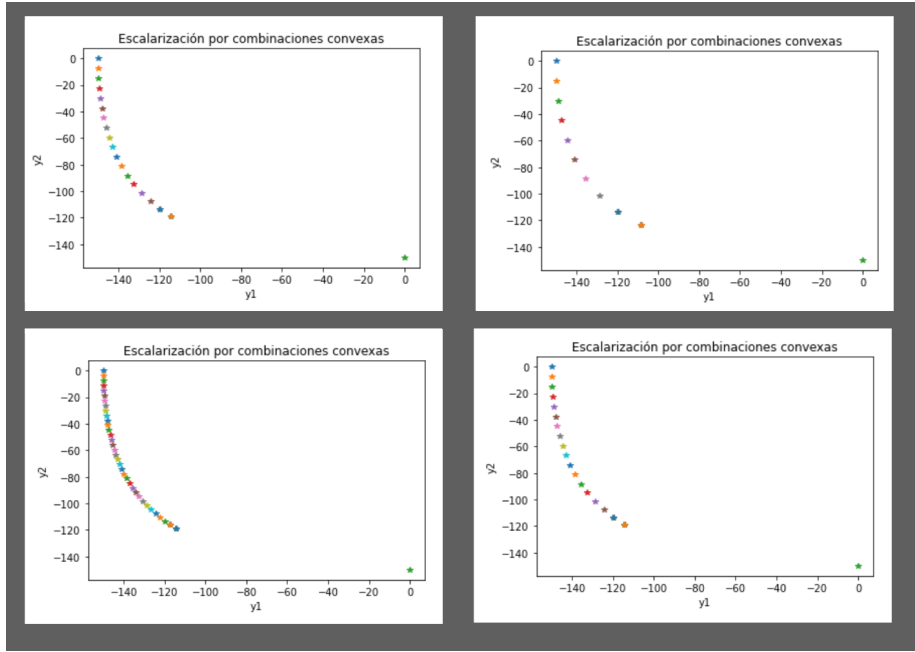


Figura 5.28: Aproximaciones generadas por combinaciones convexas

Métricas ponderadas

Para aplicar el método de las métricas ponderadas resolveremos el problema L_p ponderado:

$$\begin{aligned} \min \quad & (w_1|f_1(x) - z_1^*|^p + w_2|f_2(x) - z_2^*|^p)^{1/p} \\ \text{s.a.} \quad & x \in S \end{aligned} \quad (5.17)$$

para $1 \leq p < \infty$ y siendo z_i^* la componente i -ésima del vector objetivo ideal.

Al igual que en el método de sumas ponderadas, para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.18)$$

Vamos a variar el p tomando $p = 1$, $p = 2$, $p = 4$ y $p = 10$ sólo utilizamos el problema L_p ponderado, siguiendo el algoritmo (3), ya que genera puntos Pareto optimal a diferencia del problema ponderado de Tchebycheff genera puntos débilmente Pareto optimal.

Notemos que cómo vimos en el capítulo 4, el método de métricas ponderadas nos asegura conseguir soluciones Pareto optimal, pero puede que no encuentre

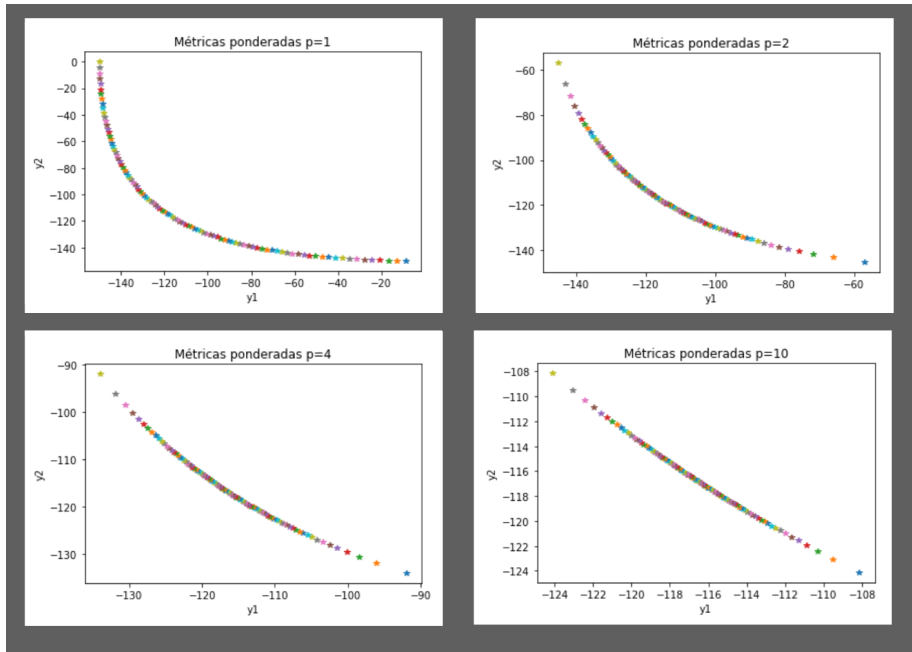


Figura 5.29: Aproximaciones generadas por métricas ponderadas

todas las soluciones como vemos en los casos $p = 2$, $p = 4$ y $p = 10$ que corresponde a los frentes 2, 3 y 4 de la Figura (5.29).

En este ejemplo vemos que tanto el método de sumas ponderadas como el de restricción ε encuentran una buena aproximación del frente de Pareto, siempre que se elijan adecuadamente los parámetros de entrada de los algoritmos para no realizar demasiadas operaciones, como ocurrió en el método de restricción ε al usar como parámetros $\alpha = 0,1$ y $\beta = 0,1$ que generó 2484 soluciones, lo que requiere del orden de $2 * 2484^{2-1}$ operaciones mientras que el método de sumas ponderadas utilizando del orden de $2 * 100$ operaciones generó un frente de Pareto similar. Si bien el método de métricas ponderadas con $p = 1$ generó una solución casi idéntica a la generada por el método de sumas ponderadas, ya que sus formulaciones son muy similares, en el caso $p = 2$, $p = 4$ y $p = 10$ sólo generó una parte del frente de Pareto. Por último, el método de escalarización por combinaciones convexas vimos que sólo generó una parte del frente de Pareto.

5.4. Problema de Fonseca - Fleming

El problema de Fonseca - Fleming [6], a diferencia del problema de Scheffer, nos generará un frente de Pareto no convexo. Es un problema de optimización bi-criterio que puede ser definido para $x \in \mathbb{R}^n$ para cualquier $n \in \mathbb{N}$ y es de la siguiente forma:

$$\begin{aligned} \min \quad & f(x) = \begin{cases} f_1(x) = 1 - \exp \left[-\sum_{i=1}^n (x_i - \frac{1}{\sqrt{n}})^2 \right] \\ f_2(x) = 1 - \exp \left[-\sum_{i=1}^n (x_i + \frac{1}{\sqrt{n}})^2 \right] \end{cases} \\ \text{s.a.} \quad & -4 \leq x_i \leq 4 \quad 1 \leq i \leq n \end{aligned} \quad (5.19)$$

Al implementar los métodos decidimos tomar $n = 2$. Con este ejemplo veremos que el método de sumas ponderadas no encontrará un frente de Pareto como en el caso anterior, sino encontrará sólo algunos puntos que también fueron encontrados por otros métodos, a esto lo veremos en la figura (5.32) que combina los puntos encontrados por el método de restricción ε y sumas ponderadas.

Sumas ponderadas

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1 \left(1 - \exp \left[-(x_1 - \frac{1}{\sqrt{2}})^2 - (x_2 - \frac{1}{\sqrt{2}})^2 \right] \right) + \\ & w_2 \left(1 - \exp \left[-(x_1 + \frac{1}{\sqrt{2}})^2 - (x_2 + \frac{1}{\sqrt{2}})^2 \right] \right) \\ \text{s.a.} \quad & x \in \mathbb{R}^2 \end{aligned} \quad (5.20)$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$. Variando los parámetros w_1 y w_2 y resolviendo el problema de la suma ponderada se pueden generar aproximaciones del conjunto de Pareto.

Para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.21)$$

Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.30), siendo cada una de las estrellas la aproximación generada por el método de sumas ponderadas.

Notemos que, como dijimos antes, al ser este un frente no convexo, el método de sumas ponderadas no genera todas las soluciones Pareto optimal, sino que sólo genera 3 soluciones.

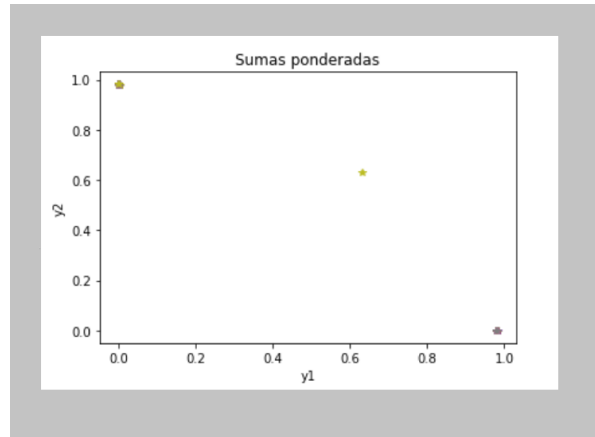


Figura 5.30: Aproximación generada por Sumas Ponderadas.

Restricción ε

Aplicamos el algoritmo (2) con diferentes α y β y conseguimos la siguiente aproximación del frente de Pareto, ver Figura (5.31), nuevamente las estrellas son las aproximaciones generadas por el método de restricción ε . Esta vez volvimos a utilizar valores de $\alpha, \beta \in [0, 1]$, ya que con valores mayores generamos muy pocos puntos y lo que queremos es tener una aproximación que nos permita tener una idea completa de como es el frente de Pareto.

En la figura de arriba a la izquierda tenemos 30 puntos, en la de arriba a la derecha 3 puntos, en la de abajo a la izquierda 12 puntos y en la de abajo a la derecha 59 puntos. Para generar estos puntos utilizamos los siguientes valores.

	Gráfico 1	Gráfico 2	Gráfico 3	Gráfico 4
# Puntos	30	3	12	59
α	0.04	0.4	0.1	0.02
β	0.04	0.4	0.1	0.02

Tabla 5.7: Valores utilizados para generar la imagen (5.31)

Como antes anticipamos, generamos un gráfico donde están las soluciones obtenidas por el método de sumas ponderadas junto con el método de restricción ε con $\alpha = 0,04$ y $\beta = 0,04$, siendo las estrellas azules las soluciones del método de sumas ponderadas y las estrellas rojas las generadas por el método de restricción ε en la Figura (5.32).

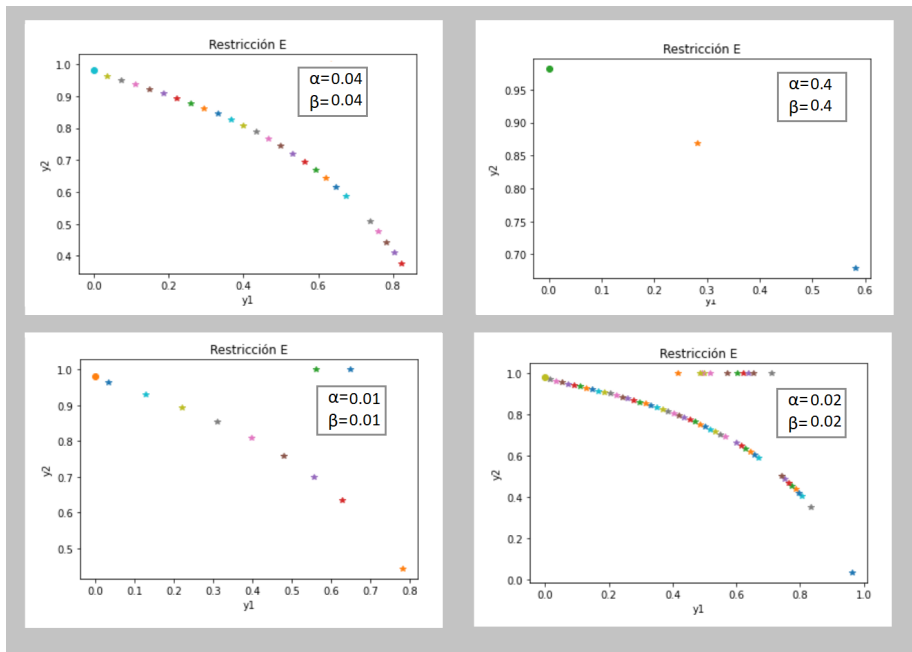


Figura 5.31: Aproximaciones generadas por Restricción ε

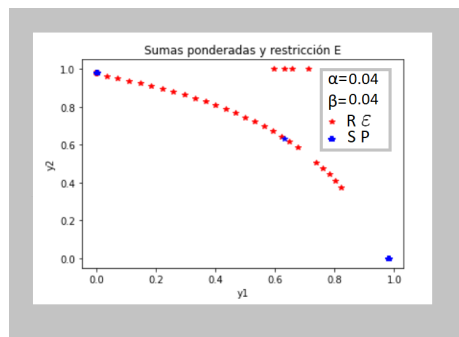


Figura 5.32: Aproximaciones generadas por Sumas ponderadas y Restricción ε

Escalarización por combinaciones convexas

Vamos a usar el algoritmo (4) variando la cantidad de puntos l que vamos a generar y el factor α que estima la curvatura total del conjunto de Pareto

En la figura de arriba a la izquierda tenemos 30 puntos, al igual que en la de arriba a la derecha y en las figuras de abajo hay 60 puntos. Para generar estos

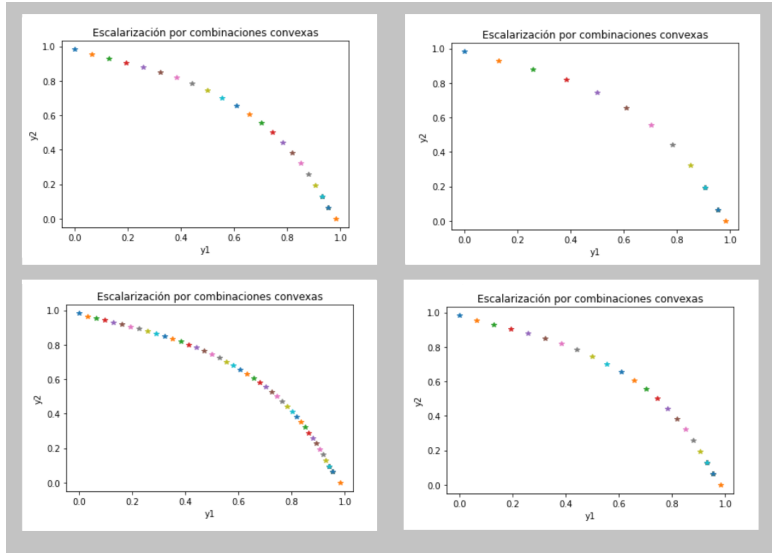


Figura 5.33: Aproximaciones generadas por combinaciones convexas

puntos utilizamos los siguientes valores. Esta vez el método de escalarización por combinaciones convexas genera una buena aproximación del frente de Pareto.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4
# Puntos	30	30	60	60
α	1.5	3	1.5	3

Tabla 5.8: Valores utilizados para generas la imagen (5.33)

Métricas ponderadas

Para aplicar el método de las métricas ponderadas resolveremos el problema L_p ponderado:

$$\begin{aligned} \min \quad & (w_1|f_1(x) - z_1^*|^p + w_2|f_2(x) - z_2^*|^p)^{1/p} \\ \text{s.a.} \quad & x \in S \end{aligned} \quad (5.22)$$

para $1 \leq p < \infty$ y siendo z_i^* la componente i -ésima del vector objetivo ideal.

Al igual que en el método de sumas ponderadas, para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.23)$$

Vamos a variar el p tomando $p = 1$, $p = 2$, $p = 4$ y $p = 10$ sólo utilizamos el problema L_p ponderado, siguiendo el algoritmo (3), ya que genera puntos Pareto optimal a diferencia del problema ponderado de Tchebycheff genera puntos débilmente Pareto optimal.

Notemos que cómo vimos en el capítulo 4, el método de métricas ponderadas

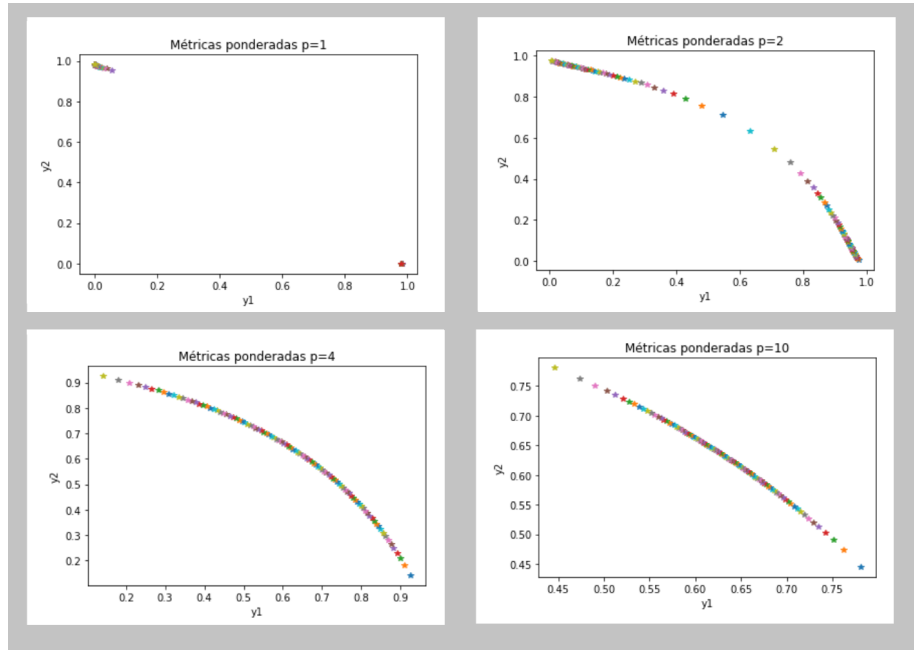


Figura 5.34: Aproximaciones generadas por métricas ponderadas

nos asegura conseguir soluciones Pareto optimal, pero puede que no encuentre todas las soluciones como vemos en los casos $p = 1$, $p = 4$ y $p = 10$ que corresponde a los frentes 1, 3 y 4 de la Figura (5.34). Si bien el frente de Pareto generado por $p = 4$ es muy similar al generado por el método de escalarización por combinaciones convexas, no encuentra las soluciones de los extremos de la figura. Y algo similar ocurre con $p = 10$.

En este ejemplo los métodos de restricción ε , escalarización por combinaciones convexas y métricas ponderadas con $p = 2$ o $p = 4$ resuelven el problema de Fonseca - Fleming generando una mejor aproximación que el método de sumas ponderadas, que tenía como debilidad el no encontrar todas las soluciones Pareto optimal si el frente no era convexo, como ocurre en este ejemplo.

5.5. Problema de Poloni

El problema de Poloni [12] es más desafiante que los anteriores, ya que, es no conexo y no convexo. El problema de Poloni está definido de la siguiente manera:

$$\begin{aligned} \min \quad & f(x) = \begin{cases} f_1(x) = 1 + (A_1 - B_1)^2 + (A_2 - B_2)^2 \\ f_2(x) = (x_1 + 3)^2 + (x_2 + 1)^2 \end{cases} \\ \text{s.a.} \quad & -\pi \leq x_i \leq \pi \quad i = 1, 2 \end{aligned} \quad (5.24)$$

$$\text{donde} \quad \begin{cases} A_1 = 0,5\text{sen}(1) - 2\text{cos}(1) + \text{sen}(2) - 1,5\text{cos}(2) \\ A_2 = 1,5\text{sen}(1) - \text{cos}(1) + 2\text{sen}(2) - 0,5\text{cos}(2) \\ B_1(x) = 0,5\text{sen}(x_1) - 2\text{cos}(x_1) + \text{sen}(x_2) - 1,5\text{cos}(x_2) \\ B_2(x) = 1,5\text{sen}(x_1) - \text{cos}(x_2) + 2\text{sen}(x_2) - 0,5\text{cos}(x_2) \end{cases}$$

Sumas ponderadas

Para este problema de optimización, la escalarización de sumas ponderadas viene dada por:

$$\begin{aligned} \min \quad & w_1(f_1(x) + (A_1 - B_1)^2 + (A_2 - B_2)^2) + w_2((x_1 + 3)^2 + (x_2 + 1)^2) \\ \text{s.a.} \quad & -\pi \leq x_1 \leq \pi \\ & -\pi \leq x_2 \leq \pi \end{aligned} \quad (5.25)$$

con coeficientes de ponderación $w_1, w_2 \in [0, 1]$ y $w_1 + w_2 = 1$. Variando los parámetros w_1 y w_2 y resolviendo el problema de la suma ponderada se pueden generar aproximaciones del conjunto de Pareto.

Para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.26)$$

Aplicamos el algoritmo (1) que aproximará al frente de Pareto de la siguiente forma, ver Figura (5.35), siendo cada una de las estrellas la aproximación generada por el método de sumas ponderadas.

En la Figura (5.35) podemos ver que aproxima una parte del frente de Pareto, pero además genera esos puntos que están en el gráfico, arriba al centro que son puntos que no pertenecen al frente de Pareto ya que, al ser un problema test sabemos que el frente es no conexo y no convexo en $y_1 \in (0, 2)$. Veremos a continuación que el método de restricción sí puede encontrar esta componente conexas del frente de Pareto.

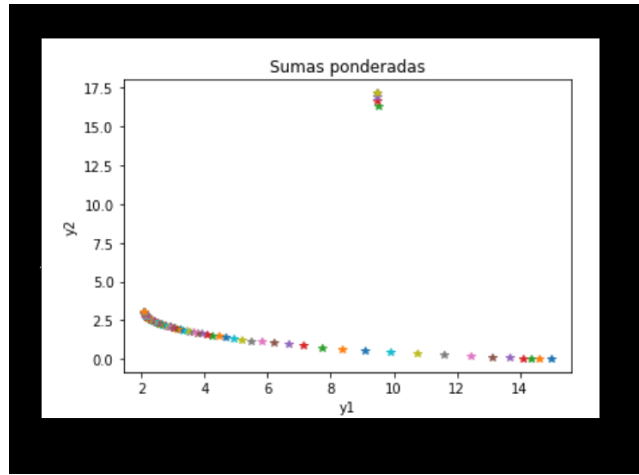


Figura 5.35: Aproximación generada por Sumas Ponderadas.

Restricción ε

Aplicamos el algoritmo (2) con diferentes α y β y conseguimos la siguiente aproximación del frente de Pareto, ver Figura (5.27), nuevamente las estrellas son las aproximaciones generadas por el método de restricción ε .

En la figura de arriba a la izquierda tenemos 132 puntos, en la de arriba a la derecha 337 puntos, en la de abajo a la izquierda 31 puntos y en la de abajo a la derecha 66 puntos. Para generar estos puntos utilizamos los siguientes valores.

	Gráfico 1	Gráfico 2	Gráfico 3	Gráfico 4
# Puntos	132	337	31	66
α	1.0	0.4	4.0	2.0
β	1.0	0.4	4.0	2.0

Tabla 5.9: Valores utilizados para generas la imagen (5.36)

Como anticipamos antes, el método de restricción ε encuentra las dos componente conexas que forman este frente de Pareto, nuevamente hay que ser cuidadosos al elegir las entradas α y β para no generar demasiadas aproximaciones del frente de Pareto para no realizar demasiadas operaciones.

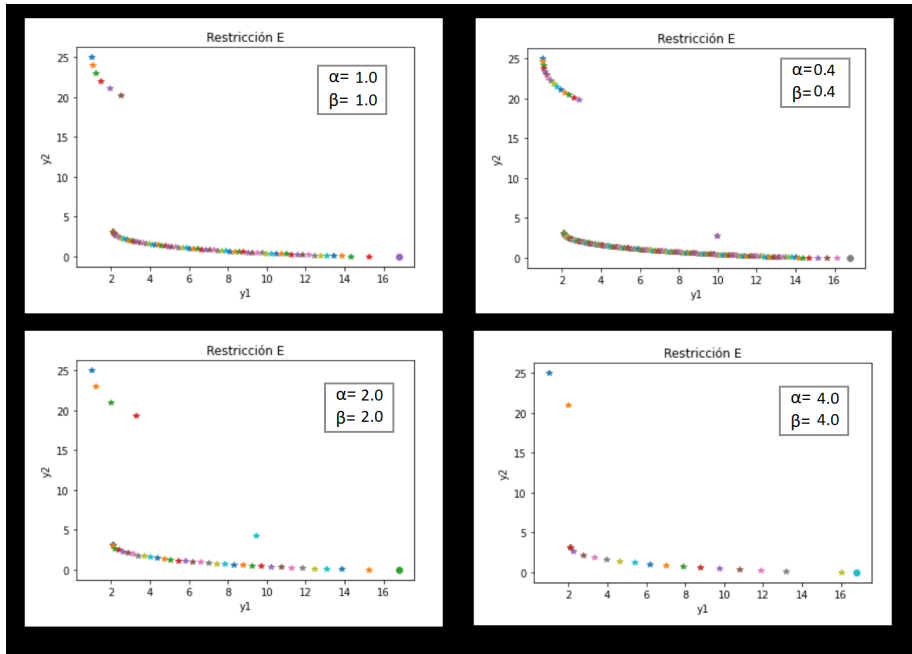


Figura 5.36: Aproximaciones generadas por Restricción ε

Escalarización por combinaciones convexas

Vamos a usar el algoritmo (4) variando la cantidad de puntos l que vamos a generar y el factor α que estima la curvatura total del conjunto de Pareto

En la figura de arriba a la izquierda tenemos 30 puntos, al igual que en la de arriba a la derecha y en las figuras de abajo hay 60 puntos. Para generar estos puntos utilizamos los siguientes valores. Notemos que este algoritmo no pudo encontrar todos los valores del frente de Pareto, pero a diferencia del método de sumas ponderadas encontro la componente conexa de la izquierda del frente de Pareto.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4
# Puntos	30	30	60	60
α	1.5	3	1.5	3

Tabla 5.10: Valores utilizados para generas la imagen (5.37)

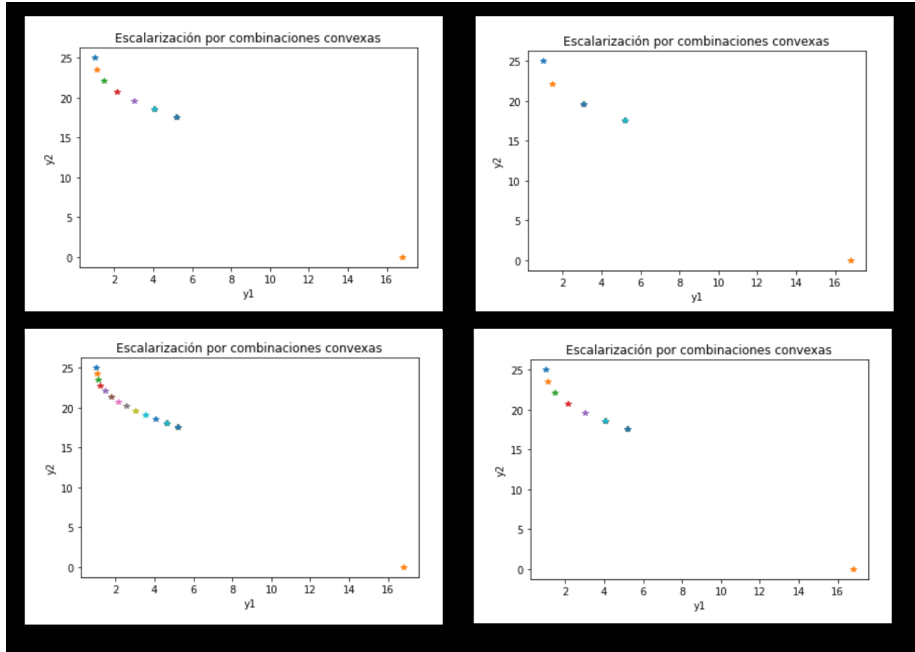


Figura 5.37: Aproximaciones generadas por combinaciones convexas

Métricas ponderadas

Para aplicar el método de las métricas ponderadas resolveremos el problema L_p ponderado:

$$\begin{aligned} \min \quad & (w_1|f_1(x) - z_1^*|^p + w_2|f_2(x) - z_2^*|^p)^{1/p} \\ \text{s.a.} \quad & x \in S \end{aligned} \quad (5.27)$$

para $1 \leq p < \infty$ y siendo z_i^* la componente i -ésima del vector objetivo ideal.

Al igual que en el método de sumas ponderadas, para aproximar el frente de Pareto utilizamos 100 puntos con 100 coeficientes de ponderación equidistantes

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0,01 \\ 0,99 \end{pmatrix}, \begin{pmatrix} 0,02 \\ 0,98 \end{pmatrix}, \dots, \begin{pmatrix} 0,99 \\ 0,01 \end{pmatrix} \right\} \quad (5.28)$$

Vamos a variar el p tomando $p = 1$, $p = 2$, $p = 4$ y $p = 10$ sólo utilizamos el problema L_p ponderado, siguiendo el algoritmo (3), ya que genera puntos Pareto optimal a diferencia del problema ponderado de Tchebycheff genera puntos débilmente Pareto optimal.

Podemos ver que $p = 1$ y $p = 2$ nos dan una aproximación del frente de Pareto que tiene las dos componentes convexas y aunque devuelven algunos va-

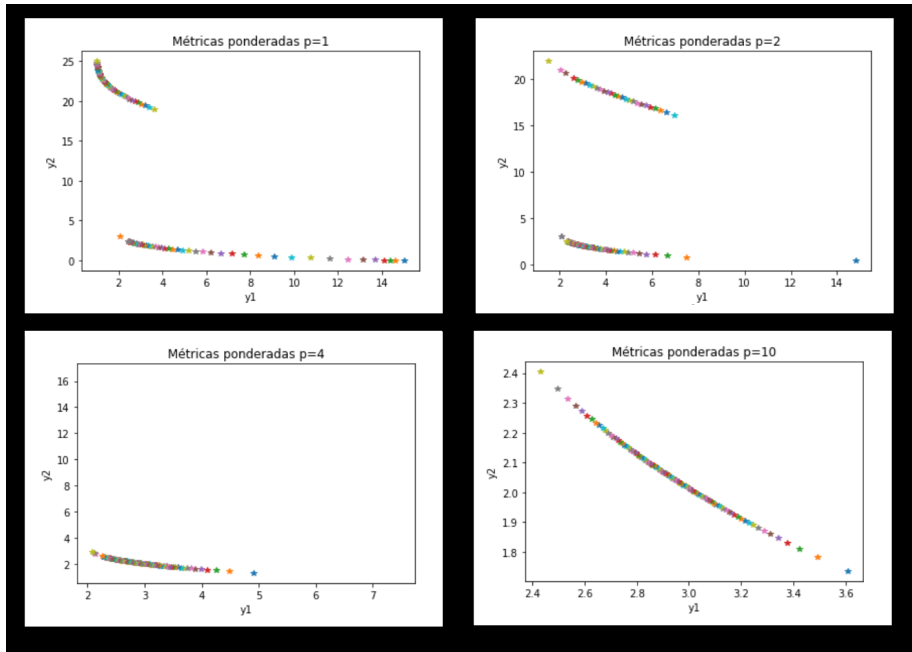


Figura 5.38: Aproximaciones generadas por métricas ponderadas

lores que no pertenecen al frente de Pareto, el analista puede comprender que los valores a la derecha de la recta vertical $y_1 = 2$ no pertenecen al frente ya que la otra componente conexa tiene valores menores en la coordenada y_2 . Por otro lado $p = 4$ y $p = 10$ nos devuelven sólo una parte del frente de Pareto.

Para resolver el problema de Poloni el método de restricción ε o el de métricas ponderadas para $p = 1$ o $p = 2$ son los únicos que encuentran las dos componente conexas del frente de Pareto, pero todos los métodos encuentran al menos una porción del frente.

Con este último ejemplo terminamos de convencernos que ningún método es *el mejor* para resolver todos los problemas.

Capítulo 6

Conclusiones

En este trabajo se estudiaron varios métodos para resolver problemas de optimización multicriterio. Nos concentramos principalmente en los métodos de escalarización que aseguraban generar, al menos, soluciones débilmente Pareto optimal. Además, se eligieron problemas test para realizar experimentos numéricos y así comparar los métodos estudiados anteriormente.

El método de sumas ponderadas es un método sencillo de programar que nos asegura encontrar soluciones Pareto optimal si los coeficientes de ponderación son todos positivos o si la solución es única, lo que no es difícil de obtener, pero como pudimos observar en el ejemplo de Fonseca y Fleming si el conjunto de Pareto no es convexo no encuentra todas las soluciones al problema, entonces este método será muy útil si el problema es convexo.

Por otro lado el método de métricas ponderadas, al igual que el de sumas ponderadas nos garantiza que las soluciones del problema L_p ponderado serán Pareto optimal si los coeficientes de ponderación son todos positivos o si la solución es única y es, aunque un poco más complejo que el método de sumas ponderadas, bastante simple de programar y puede encontrar soluciones incluso si el problema no es convexo, no tuvo grandes inconvenientes para resolver el ejemplo de Fonseca - Fleming que es no convexo, ni el de Poloni que es no convexo y no convexo, pero tiene como flaqueza que no necesariamente encuentra todas las soluciones, como pudimos observar en el ejemplo de Schaffer que al compararlo con los resultados obtenidos con los otros métodos vemos que no encuentra todas las soluciones.

El método de restricción ε puede identificar todas las soluciones Pareto optimal independientemente de la convexidad del problema e incluso identifica las soluciones usando un conjunto de restricciones diferentes. Por su parte, como desventaja, es un método más costoso computacionalmente si lo comparamos con los métodos anteriores, consume un mayor tiempo de cómputo y su formulación es complicado a medida que se aumenta la cantidad de funciones objetivo.

También presentamos ejemplos de la vida real, en particular, una aplicación en terapia de radiación de intensidad modulada, que Gabriele Eichfelder resolvió usando el método de restricción ε , que fue útil para entender la dificultad al encontrar buenos datos reales que se amolden a la teoría que hemos estudiado.

Si bien no resolvimos un problema de mundo real, dejamos las bases para un posible trabajo futuro, una vez que se tengan datos reales, poder usar las comparaciones establecidas en este trabajo para resolverlo de la manera que los analistas junto con el tomador de decisiones consideren adecuadas, ya que después de todo lo estudiado, sabemos que elegir un método para resolver un problema de optimización multicriterio puede ser, en sí mismo, un problema de optimización multiobjetivo.

Bibliografía

- [1] BOWMAN, V. J. On the relationship of the tchebycheff norm and the efficient frontier of multiple-criteria objectives. In *Multiple criteria decision making*. Springer, 1976, pp. 76–86.
- [2] EHRGOTT, M. *Multicriteria optimization*, vol. 491. Springer Science & Business Media, 2005.
- [3] EHRGOTT, M., GÜLER, Ç., HAMACHER, H. W., AND SHAO, L. Mathematical optimization in intensity modulated radiation therapy. *Annals of Operations Research* 175, 1 (2010), 309–365.
- [4] EICHFELDER, G. Adaptive scalarization methods in multiobjective optimization springer. *Berlin* (2008).
- [5] EICHFELDER, G. Vector optimization in medical engineering. In *Mathematics Without Boundaries*. Springer, 2014, pp. 181–215.
- [6] FONSECA, C. M., AND FLEMING, P. J. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary computation* 3, 1 (1995), 1–16.
- [7] HAZEN, G. B. Differential characterizations of nonconical dominance in multiple objective decision making. *Mathematics of operations research* 13, 1 (1988), 174–189.
- [8] HWANG, C.-L., AND MASUD, A. S. M. Methods for multiple objective decision making. In *Multiple objective decision making—methods and applications*. Springer, 1979, pp. 21–283.
- [9] MIETTINEN, K. *Nonlinear multiobjective optimization*, vol. 12. Springer Science & Business Media, 2012.
- [10] PEREYRA, V. Fast computation of equispaced pareto manifolds and pareto fronts for multiobjective optimization problems. *Mathematics and Computers in Simulation* 79, 6 (2009), 1935–1947.
- [11] PEREYRA, V., SAUNDERS, M., AND CASTILLO, J. Equispaced pareto front construction for constrained bi-objective optimization. *Mathematical and Computer Modelling* 57, 9-10 (2013), 2122–2131.

- [12] POLONI, C., GIURGEVICH, A., ONESTI, L., AND PEDIRODA, V. Hybridization of a multi-objective genetic algorithm, a neural network and a classical optimizer for a complex design problem in fluid dynamics. *Computer Methods in Applied Mechanics and Engineering* 186, 2-4 (2000), 403–420.
- [13] SAWARAGI, Y., NAKAYAMA, H., AND TANINO, T. *Theory of multi-objective optimization*. Elsevier, 1985.
- [14] SCHAFFER, J. D. Multiple objective optimization with vector evaluated genetic algorithms. In *Proceedings of the first international conference on genetic algorithms and their applications, 1985* (1985), Lawrence Erlbaum Associates. Inc., Publishers.

Los abajo firmantes, miembros del Tribunal de Evaluación de tesis, damos Fe que el presente ejemplar impreso, se corresponde con el aprobado por éste Tribunal.

Dr. Elvio A. Pilotta Dr. Andrés A. Barrea Dra. A. Georgina Flesia