

---

# EXÁMEN DE CANDIDATURA - OPTIMIZACIÓN ESTOCÁSTICA

---

A PREPRINT

**Joel Chacón Castillo**  
Department of Computer Science  
Centro de Investigación en Matemáticas  
joel.chacon@cimat.mx

June 12, 2019

**Keywords** Multi-objective · Directed Search Method · Optimization

## 1 Introduction

En muchas aplicaciones es necesario lidiar con problemas que involúcran varios objetivos, estos objetivos usualmente están en conflicto, esto es conocido como un Problema de Optimización Multi-objetivo (*Multi-objective Optimization Problem*- MOP). Una característica importante de un MOP, es que el conjunto de soluciones, mejor conocido como conjunto de Pareto y su respectiva imagen frente de Pareto, no consisten de un sólo punto.

Este trabajo utiliza como referencia *The Directed Search Method for Unconstrained Multi-Objective Optimization Problems* by Oliver et al. [1], donde se utiliza un método de búsqueda directa, el cual implica la resolución de un Problema de Valor Inicial (*Initial Value Problem* - IVP) y además de propone un procedimiento Predictor-Corrector para tener una convergencia al frente de Pareto y un recorrido a lo largo del frente de Pareto.

## 2 Literature Review

### 2.1 Survey Methods of Multi-objective Optimization

Con el propósito de proveer una solución a un Problema de Optimización Multi-Objetivo (MOP - Multi-Objective Optimization Problem) se han propuesto una gran variedad de métodos. Una clasificación ampliamente utilizada en el ámbito multi-objetivo [1, 2] identifica a los métodos de escalarización, técnicas de subdivisión y métodos de búsqueda estocásticos. A su vez los métodos de escalarización y las técnicas de subdivisión son tratados por medio de procesos determinísticos.

#### 2.1.1 Métodos de escalarización

En su mayoría estos métodos se basan en la escalarización del problema de optimización, es decir utilizan un principio para transformar un MOP en un problema de optimización escalar.

- **Weighting Method:**

- Este método, fue propuesto por Zadeh et. al [3].
- Su principio fundamental se basa en asignar un peso  $\alpha_i \geq 0$  para cada función objetivo  $k$ , donde el conjunto de vectores de pesos son normalizados  $\sum_{i=1}^k \alpha_i = 1$  y resolver el siguiente, problema escalar:

$$\min_{x \in \mathbb{R}} \sum_{i=1}^k \alpha_i f_i(x) \quad (1)$$

- Depende de la geometría del frente de Pareto, es decir no es posible encontrar regiones en geometrías irregulares.

- **Goal Programming:**

- El propósito de este método es encontrar el punto cuya imagen esté más cercano a otro punto ideal  $Z \in \mathbb{R}^k$ .
- La formulación de este problem se define a continuación.

$$\min_{x \in \mathbb{R}^n} d(Z, F(x)) \quad (2)$$

donde  $d(., .)$  es una métrica de distancia escogida.

- **Weigthed  $L_p$ -metric method:**

- Escoge un punto  $\hat{y}$  en el espacio objetivo para la búsqueda de soluciones eficientes.
- La forma escalar que se considera en este problema es:

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^k \omega_i |f_i(x) - \hat{y}_i|^p \quad (3)$$

- **$\epsilon$  - Constraint Method:**

- Propuesto por Marglin et. al. [4].
- Minimiza un objetivo de forma individual  $f_j, j \in \{1, \dots, k\}$ .
- El problema de optimización escalar tiene la siguiente forma:

$$\begin{aligned} \min_{x \in \Omega} f_j(x) \quad & j \in \{1, \dots, k\} \\ \text{s.t.} \quad & \\ \Omega = \{x \in \mathbb{R}^n | f_i(x) \leq \epsilon_i \forall i \in \{1, \dots, k\}, i \neq j\} \end{aligned} \quad (4)$$

- **Method of equality constraints:**
- **Normal-Boundary Intersection:**
- **Homotopy Approach**

### 2.1.2 Técnicas de subdivisión

Este tipo de estrategias se basan en algoritmos de subdivisión para la aproximación del conjunto de Pareto, particularmente se basa en cajas que cubren el espacio de los objetivos ver [5] para mayor información.

### 2.1.3 Métodos estocásticos

- **Stochastic Search According to Timmel.**
- **Evolutionary Algorithms.**

## 3 Related Works

En esta sección se realiza una revisión de las estrategias que se basan en el cálculo de las direcciones de búsqueda para optimización multi-objetivo. El punto de referencia es el trabajo propuesto por Oliver Shutze et al. [1].

El concepto de aplicar el cálculo de las direcciones de búsqueda y considerar múltiples objetivos no es nuevo. Por su parte Brown y Smith (2005) [6] aplicaron el gradiente en el ámbito multi-objetivo, por otra parte Fliege y Fux (2000) [7, 8] utilizaron el gradiente negativo dentro un simplejo con el propósito de obtener soluciones a lo largo del frente de Pareto. Además, Bosman (2012) [9] utiliza una procedimiento de búsqueda en línea como búsqueda local. Finalmente, una variante del método de búsqueda directa es propuesto por Shutze et. al (2016) [10] el cual utiliza una idea similar que la de Bosman (2012), sin embargo el primero utiliza este último para dirigir la búsqueda tanto hacia el frente de Pareto (convergencia) como a lo largo del frente.

Particularmente, Shutze et. al ha incorporado el trabajo de referencia [1] en dos categorías de algoritmos evolutivos multi-objetivo (Multi-objective Evolutionary Algorithms - MOEA), estos son basados en indicadores [11] y basados en descomposición [10].

### 3.1 Gradient Descent Methods for MOPs

En los métodos de optimización basados en gradiente más usuales, los algoritmos primero evalúan el gradiente perteneciente a una función. Este gradiente es utilizado para identificar una dirección de búsqueda. La búsqueda lineal pasa por el espacio de decisión, con el propósito de identificar el punto con mínimo valor objetivo. Particularmente, en el campo multi-objetivo se han propuesto varios métodos basados en el descenso de gradiente [8, 12, 13].

Los métodos de descenso por gradiente son principalmente utilizados en los problemas de escalarización.

#### 3.1.1 Dirección de descenso en *Weighted Sum Approach*

Dado un conjunto de pesos  $\alpha$  convexos, este método define la escalarización del MOP como se indica a continuación.

$$\min_{x \in \mathbb{R}} \sum_{i=1}^k \alpha_i f_i(x) \quad (5)$$

Este problema se puede resolver por medio del método de mayor descenso (*steepest descent method*) escogiendo como dirección  $v(x) = -\sum_{i=1}^k \alpha_i \nabla f_i(x)$ .

#### 3.1.2 Exploiting Gradient Information in Numerical Multi-Objective Evolutionary Optimization

Por su parte Bosman et. al. [13] consideran esquemas de hibridización empleados en forma generacional, es decir, algoritmos meméticos, estos procedimientos consisten que al final de cada generación se aplica la búsqueda local con cada individuo como punto inicial. Particularmente se definen las siguientes categorías para el cálculo del gradiente. Aparentemente, estas categorías están desarrolladas para tratar la correlación que existe en el espacio de los objetivos.

- *Random-Objective Conjugate Gradient - ROCG*: En este enfoque se aplica el algoritmo de gradiente conjugado a un objetivo escogido de forma aleatoria. Este enfoque depende en la correlación que existe entre los objetivos, es decir si mejorar un objetivo implica una mejora en los otros objetivos.
- *Alternating-Objective Repeated Line Search - AORL*: Para reducir el efecto de empeorar un subconjunto de objetivos al mejorar otro, se puede alternar el objetivo en el proceso de optimización. En AORL, la dirección del gradiente negativo es utilizado en un objetivo, esta dirección es cambiada al alternar de objetivo.
- *Combined-Objective Repeated Line Search - CORL*: Esta estrategia utiliza una dirección de interés del cual es calculado el gradiente en el caso multi-objetivo.

#### 3.1.3 Directed Search Methods - DSM

Los algoritmos de derivación libre *Derivative-free* para el caso de optimización mono-objetivo son típicamente divididos en tres grupos [14]: Métodos de búsqueda directa (*Search Methods - DSM*), algoritmos de búsqueda en línea (*line-search algorithms*) y métodos de interpolación basados en regiones de confianza (*trust-region interpolation based methods*).

Los métodos de búsqueda directa son caracterizados por no considerar algún modelo explícito o implícito de la función objetivo. La minimización de una función es logrado por medio de un proceso iterativo de la evaluación de la función en un conjunto de puntos finitos, de esta forma se utilizan los resultados para determinar cuales puntos deben ser evaluados en las siguientes iteraciones.

Los métodos de búsqueda dirigida en el ámbito multi-objetivo son relativamente recientes. Zhong et al. [15] propuso un algoritmo empírico basado en la búsqueda de compás o brújula, el cual carece de un análisis de convergencia. Únicamente dos DSM han sido propuestos para optimización multi-objetivo, estos son: Búsqueda directa multi-objetivo de malla adaptativa (*Multiobjective Mesh Adaptive Direct Search - MULTIMADS*) [16] y Búsqueda múltiple directa (*Direct Multisearch - DMS*) [17].

#### Multi-objective Mesh Adaptive Direct Search - MULTIMADS

Este DSM convierte el problema multi-objetivo en uno mono-objetivo, esto en base a la combinación de los componentes de las funciones objetivo. Un enfoque ampliamente utilizado es la suma ponderada de los objetivos. Específicamente Audet et al. [18] desarrolló una variante del DSM para optimización de dos objetivos. En este último trabajo, se propusieron dos funciones de agregación las cuales definen diferentes formulaciones mono-objetivo.

## Direct Multisearch for Multi-objective Optimization

Este tipo de métodos no consideran la agregación de los componentes, tampoco definen prioridades de los objetivos involucrados. En la propuesta de Cutódio et al. [19] la meta era generalizar todos los tipos de DSM en el ambito multi-objetivo. Similarmente a otros trabajos de Audet et al. [18, 16], este algoritmo mantiene una lista de puntos no dominados que representan una aproximación del frente de Pareto, estos puntos son utilizados para generar otros puntos nuevos.

## 4 Analyses The Directed Search Method for Unconstrained Multi-Objective Optimization Problems

En este trabajo se propone un procedimiento iterativo para el tratamiento numérico de problemas de optimización multi-objetivo. Particularmente, se realiza una dirección de búsqueda en el espacio de decisión dado una dirección en el espacio de los objetivos. La principal aportación de este trabajo se basan en el ensamblado de dos métodos. En el primero se propone un método de descenso de búsqueda directo el cual se puede interpretar como una categoría de los métodos de descenso, este método se enfatiza la convergencia hacia el frente de Pareto dado un vector de direcciones. El segundo método se basa en un procedimiento de continuación. Es decir, dado un punto ubicado en el frente de Pareto, este procedimiento de continuación realiza un recorrido a lo largo del frente de Pareto. A diferencia de otros métodos [20], el método propuesto no requiere información del 2nd gradiente (segundas derivadas).

Por lo tanto la idea es ajustar la dirección de búsqueda en el espacio objetivo de forma a priori para obtener un esquema numérico .

### 4.1 Notations and background

La notación que se utiliza está basada en el trabajo de Schutze et al.[10] y no en el trabajo de referencia [1] ya que el primero es más reciente y se puede considerar como una continuación. En la ecuación 6 se define un Problema de Optimización Multi-objetivo (*Multi-objective Optimización Problem - MOP*).

$$\min_{x \in \mathbb{R}^n} \{F(x)\} \quad (6)$$

donde  $F$  es definido como un vector de funciones objetivo  $F : Q \rightarrow \mathbb{R}^k$ ,  $F(x) = (f_1(x), \dots, f_k(x))$  . Además, cada objetivo  $f_i : Q \rightarrow \mathbb{R}$  es suficientemente suave. Así, la optimalidad de un MOP es definido por el concepto de dominancia.

**Definition 1** Un vector  $y \in \mathbb{R}^n$  es dominado por un vector  $x \in \mathbb{R}^n$  ( $x \preceq y$ ) con respecto a un MOP si  $f_i(x) \leq f_i(y)$ ,  $i = 1, \dots, k$  y si existe un índice  $j$  tal que  $f_j(x) < f_j(y)$ , en caso contrario  $y$  es no dominado por  $x$ .

Un punto  $x \in \mathbb{R}^n$  es definido como punto de Pareto si no hay un vector  $y \in \mathbb{R}^n$  el cual domine a  $x$ . El conjunto de todas las soluciones de Pareto óptimas es definido como *Conjunto de Pareto*, y es definido por  $P$ . La imagen  $F(P)$  de un conjunto de Pareto es conocido como *frente de Pareto*

El Jacobiano de  $F$  en un punto  $x$  es dado por:

$$J(x) = \begin{pmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_k(x)^T \end{pmatrix} \in \mathbb{R}^{k \times n}$$

donde  $\nabla f_i(x)$  denota el gradiente del objetivo  $f_i$ . Si todos los objetivos del MOP son diferenciables, entonces el teorema de Kuhn Ticks [21] establece una condición necesaria para la optimalidad de Pareto en problemas sin restricciones

**Definition 2** Si  $x^*$  es un punto de Pareto de un MOP, entonces existe un vector  $\alpha \in \mathbb{R}^k$  con  $\alpha_i \geq 0, i = 1, \dots, k$ , y  $\sum_{i=1}^k \alpha_i = 1$  tal que

$$\sum_{i=1}^k \alpha_i \nabla f_i(x^*) = J(x)^T \alpha = 0 \quad (7)$$

Así, los puntos que satisfacen a la ecuación 7 son llamados como los puntos de Karush-Kuhn-Tucker (KKT).

## 4.2 The Directed Search Method

Principalmente se considera un punto  $x_0 \in \mathbb{R}^n$  con rango  $J(x_0) = k$  y un vector  $d \in \mathbb{R}^k$  que representa la dirección de búsqueda deseada en el espacio de los objetivos. Entonces, una dirección de búsqueda  $v \in \mathbb{R}^n$  en el espacio de decisión es buscado tal que  $y_0 = x_0 + hv$ , donde  $h \in \mathbb{R}_+$  es el tamaño de paso. Se mantiene lo siguiente si  $\|v\| = 1$ :

$$\lim_{h \rightarrow 0} \frac{f_i(y_0) - f_i(x_0)}{h} = \langle \nabla f_i(x_0), v \rangle = d_i, \quad i = 1, \dots, k, \quad (8)$$

En base a la ecuación 8 se puede establecer lo siguiente en notación matricial

$$J(x_0)v = d \quad (9)$$

Así la dirección de búsqueda  $v$  puede ser calculada resolviendo el sistema lineal. Sin embargo es muy probable que el sistema de la ecuación 9 pueda ser no determinado, esto implica que la solución no es única. Una selección posible puede ser como se indica a continuación.

$$v_+ = J(x_0)^+ d \quad (10)$$

donde  $J(x_0)^+ \in \mathbb{R}^{n \times k}$  denota la pseudo inversa del Jacobiano  $J(x_0)$ . Así, dada una solución candidata  $x_0$  se puede generar una nueva solución de la forma  $x_1 = x_0 + hv$ , donde  $v \in \mathbb{R}^n$  es un vector que satisface a la ecuación 9. Entre las soluciones de la ecuación 9,  $v_+$  es aquella con menor norma euclídeana. Por lo tanto, dado  $h$  se espera que para un paso en la dirección  $v_+$  en el espacio de decisión se tenga un mayor progreso en la dirección  $d$  del espacio objetivo.

### 4.2.1 The Set of Descent Direction

El conjunto de direcciones de descenso están definidas por la relación de dominancia, así el cono de descenso en el punto  $x_0$  está dado por

$$D(x_0) = \{v \in \mathbb{R}^n : \langle \nabla f_i(x_0), v \rangle \leq 0, \forall i = 1, \dots, k, \langle \nabla f_j(x_0), v \rangle < 0, \text{ for } j \in \{1, \dots, k\}\} \quad (11)$$

Utilizando el Jacobiano la ecuación 11 se puede escribir de la siguiente forma:

$$J(x_0)v \leq 0^J(x_0)v \neq 0 \quad (12)$$

Después de una posible normalización la ecuación 12 se puede escribir de la siguiente forma

$$J(x)v = -d, \quad i = 1, \dots, k \quad (13)$$

### 4.2.2 A Curve of Dominating Points

La curva de puntos dominados es aquella en la que cada punto que la conforma es dominada por sus sucesores [2]. Particularmente, esta curva es modelada como una ecuación diferencial mejor conocido como un el problema de valor inicial (*Initial Value Problem - IVP*). Dada una dirección  $d \in \mathbb{R}^k$  con  $d_i \leq 0, i = 1, \dots, k$  y un punto  $x_0 \in \mathbb{R}^n$  con rango  $J(x_0) = k$ . Una búsqueda greedy en la dirección  $d$  por medio de la ecuación 10 conduce a una solución numérica del siguiente problema de valor inicial.

$$Q(t, x_0, d) = \begin{cases} x(0) = x_0 \in \mathbb{R}^n \\ \frac{d(x(t))}{dt} = J(x(t))^+ d, t > 0 \end{cases} \quad (14)$$

donde  $t$  denota el tiempo. Así, la ecuación 14 define una familia de curvas conformadas por puntos dominantes. Dado que la ecuación 14 define una curva de puntos dominantes a partir de un valor inicial, se convierte en el punto principal de estudio. Una propiedad muy importante de la ecuación 14, es lo siguiente:

$$f(x(t)) < f(x(s)) \quad \forall 0 \leq s < t < \infty \quad (15)$$

así  $\lim_{t \rightarrow T} Q(t, x_0, d) = 0$ , donde  $T$  es el número real más largo que conforma una solución de  $Q(\cdot)$ . Ahora, dado  $\gamma : [0, t_f] \rightarrow \mathbb{R}^n$  sea una solución, donde  $t_f$  es el valor final, y  $t_c$  (ver figura 2) es el valor más pequeño  $t \geq 0$  tal que  $v \in \mathbb{R}^n : J(x(t))v = d$ .

Al valor  $t_c$  se le identifica como valor crítico, y a  $\gamma(t_c)$  el punto crítico de  $Q(\cdot)$ . Así  $\gamma$  puede ser dividido en dos partes:  $\gamma([0, t_c])$  y  $\gamma([t_c, t_f])$ . En la primera parte  $F(\gamma(t))$  alcanza el decremento deseado en la dirección  $d$ . Así, el cálculo del punto crítico  $\gamma(t_c)$  se puede calcular en base a la dirección definida de la siguiente forma:

$$v_+(x(t)) = J(x(t))^+ d = \arg \min_{v \in \mathbb{R}^n} \|J(x(t))v - d\| \quad (16)$$

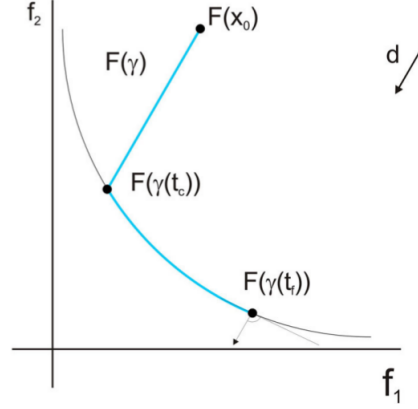


Figure 1: Una posible solución que define una curva  $F(\gamma)$  de  $Q(\cdot)$ .

La segunda parte implica el cálculo de  $\gamma$  en el segmento  $[t_c, t_f]$  (en la curva del frente de Pareto), esto podría requerir un costo computacional significativo desde que  $Q(\cdot)$  es numéricamente inestable *stiff*.

Este problema se puede resolver alternativamente con un método de discretización directo. En el algoritmo 1 se describe un posible algoritmo para trazar la solución de la curva del problema IVP ( $Q(\cdot)$  - ecuación 14). Es interesante notar que el algoritmo 1 que es equivalente al método de descenso de gradiente, en este se calcula la integral de la curva en base al método de Euler donde en la línea 5 se considera una discretización del dominio, este dominio es evaluado en la línea 3. El criterio de paro que se aplica en este algoritmo se basa en el número de condición del Jacobiano  $k_2(J(x_i))$ , esto ya que conforme el algoritmo alcanza un punto crítico (no necesariamente puede ser una solución del conjunto óptimo de Pareto) existe una inestabilidad numérica ya que la norma del gradiente de varios objetivos tiende a cero (condición de optimalidad de primer orden).

Una solución al problema de valor inicial IVP no garantiza tener una convergencia a un punto de Pareto óptimo ya que depende principalmente de la dirección  $\alpha$  en el espacio de los objetivos. Sin embargo dado el teorema establecido por *Kuhn and Tucker* se puede seleccionar una dirección en el espacio objetivo que sí garantice ser un punto óptimo de Pareto resolviendo la siguiente ecuación.

$$\begin{aligned} \alpha^* = \arg \min_{\alpha} & \left\| \sum_{i=1}^k \alpha_i \nabla f_i(x) \right\|_2^2 \\ \text{s.t. } & \alpha_i \geq 0, \quad i = 1, \dots, k \\ & \sum_{i=1}^k \alpha_i = 1 \end{aligned} \quad (17)$$

Debido a que la elección de la dirección de búsqueda en el espacio de los objetivos (cambio de notación  $\alpha = d$ ),

### 4.3 A Continuation Algorithm Based on the Directed Search Approach

El procedimiento de continuación asume lo siguiente como resultado de aplicar el procedimiento predictor que se menciona con anterioridad. Dado un punto de Pareto (local)  $x$  y un vector de pesos convexo tal que

$$\sum_{i=1}^k \alpha_i \nabla f_i(x) = 0 \quad (18)$$

y además se asume que  $\text{rank}(J(x)) = k - 1$  ( $k$  la dimensión del espacio objetivo). Entonces una dirección ortogonal a la dirección de búsqueda  $\alpha$  podría ser prometedora para obtener nuevos puntos predictores. De esta forma si se considera que la dirección de búsqueda es equivalente la descomposición QR

$$\alpha = QR \quad (19)$$

donde  $Q = (q_1, \dots, q_k) \in \mathbb{R}^{k \times k}$  es una matriz ortogonal y  $q_i, i = 1, \dots, k$ , son los vectores columna, y  $R = (r_{11}, 0, \dots, 0)^T \in \mathbb{R}^{k \times 1}$  con  $r_{11} \in \mathbb{R}$ . Entonces a dirección de búsqueda promisorio  $v_i$  podría ser aquella

que satisfice

$$J(x)v_i = q_i, \quad i = 2, \dots, k \quad (20)$$

El efecto de esto es encontrar la dirección de descenso en el espacio de las variables que sea ortogonal en el espacio de los objetivos. Es importante notar que el nuevo punto  $x$  podría no pertenecer al frente de Pareto.

En el caso de dos objetivos la orientación en el espacio objetivo es denotada por el signo del segundo componente en relación al segundo componente del segundo vector propio ( $q_{2,2}$ ), en resultado se tiene define el paso predictor de la siguiente forma

$$p = x_0 - \text{sgn}(q_{2,2})tv \quad (21)$$

En base a varias asunciones de la naturaleza del problema y en debido a que cada objetivo es Lipshitz continuo (ver [22] para mayor información sobre la continuidad Lipshitz) y que el tamaño de paso es suficientemente pequeño se tiene que  $|f_i(p) - f_i(x)| \approx L_{i,x}||p - x||_2$ , donde  $L_{i,x}$  puede ser aproximado por la norma de la derivada direccional, en resultado se tiene para cada objetivo que  $t_i = \frac{\epsilon}{|\langle \nabla f_i(x), v \rangle|}$  en resultado se tiene que un tamaño adecuado como sigue

$$t = \min_{i=1, \dots, k} t_i \quad (22)$$

---

#### Algorithm 1 Directed Search Method

---

**Require:** starting point  $x_0 \in \mathbb{R}^n$  with  $\text{rank}(J(x_0)) = k$ ,  $\text{tol} \in \mathbb{R}_+$  convex weight  $\alpha_0 \in \mathbb{R}^k$

- 1:  $i=0$
  - 2: **while**  $k_2(J(x_i)) < \text{tol}$  **do**
  - 3:   compute  $v_i = -J(x_i)^+ \alpha_i$
  - 4:   compute step-size  $t_i \in \mathbb{R}_+$ , Armijo conditions can be used.
  - 5:   set  $x_{i+1} = x_i + t_i v_i$
  - 6:   set  $i = i + 1$
- 

---

#### Algorithm 2 Bi-Objective Continuation

---

**Require:** Initial Solution  $(x_0, \alpha_0)$  with  $\alpha_0 \approx 1$ ,  $\text{threshold} \epsilon \in \mathbb{R}_+$ , tolerance  $\delta \in \mathbb{R}_+$

- 1:  $i=0$
  - 2: **while**  $1 - \alpha_1 > \delta$  **do**
  - 3:   compute  $q_2$  as in eqn. 19
  - 4:   compute  $v$  as in eqn. 20
  - 5:    $p_i = x_i - \text{sgn}(q_{2,2})tv$
  - 6:   compute  $x_{i+1}$  by solving IVP  $(Q(p_i, \alpha_i))$  with the initial value  $p_i$  and using  $\alpha_i$
  - 7:   compute  $\alpha_{i+1}$  as in eqn. 17
  - 8:   set  $i = i + 1$
- 

## 5 Detalles de implementación

Se implementaron los algoritmos 1 y 2 y probaron con distintos puntos iniciales y vectores de pesos. Principalmente, se observó que el tamaño de paso influye mucho en estos algoritmos de descenso, particularmente en el algoritmo 2, donde se observó que en ocasiones el condicionamiento de la matriz Jacobiana no era muy elevado, sin embargo ya se tenía una convergencia al frente de Pareto. Para acelerar el procedimiento en el algoritmo 1 se agregó un criterio de paro en base a las iteraciones, diez en este caso.

Como se espera el algoritmo de búsqueda directo realiza un recorrido en el espacio de los objetivos de izquierda a derecha y de arriba hacia abajo. El problema multi-objetivo que se utiliza se define a continuación<sup>1</sup>

$$\begin{aligned} f_1, f_2 : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ f_1(x, y) &= \frac{1}{2}(\sqrt{1 + (x + y)^2} + \sqrt{1 + (x - y)^2} + x - y) + \lambda e^{-(x-y)^2} \\ f_2(x, y) &= \frac{1}{2}(\sqrt{1 + (x + y)^2} + \sqrt{1 + (x - y)^2} - x + y) + \lambda e^{-(x-y)^2} \\ \lambda &= 0.85 \end{aligned} \quad (23)$$

---

<sup>1</sup>Esta función está mal definida en el artículo de referencia en un signo, en su lugar se utiliza la que se indica en [23]

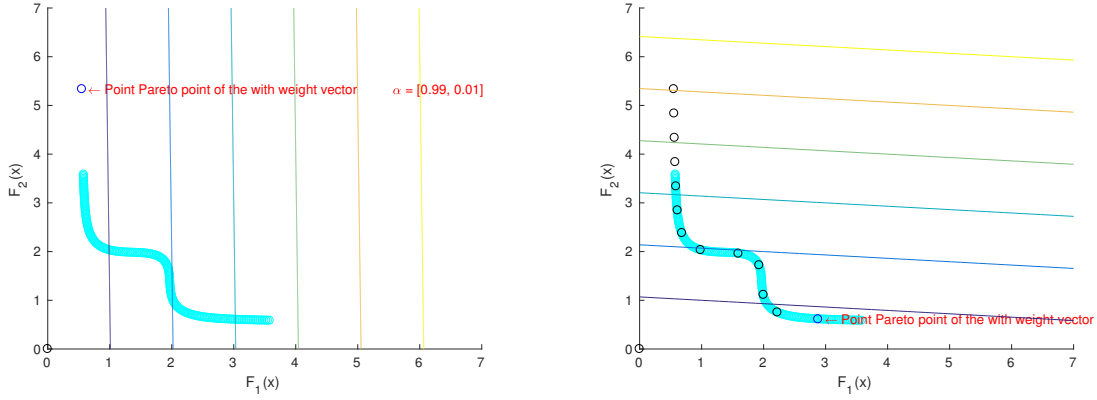


Figure 2: Soluciones encontradas por el algoritmo en una ejecución, en la figura de la izquierda se presenta el primer punto encontrado y en la figura derecha el último punto encontrado. Además se ilustra la gráfica de contorno dado el vector de pesos indicado, el frente de Pareto está indicado en azul.

En la figura 2 se presenta una simulación del algoritmo al inicio y al fin de la ejecución. Además se ilustran las curvas de contorno del problema escalarizado como resultado de aplicar la ponderación a los objetivos<sup>2</sup>. El lenguaje utilizado es MATLAB.

## 6 Conclusión y trabajos futuros

En este trabajo se hizo una revisión de las clasificaciones de optimización multi-objetivo, específicamente los métodos cuyo procedimiento de optimización se basan en estrategias determinísticas. Quizás una de las estrategias más utilizadas son la de descenso por gradiente, particularmente los métodos de búsqueda directa. Estos últimos se ajustan adecuadamente al problema multi-objetivo, ya que calculan una dirección en el espacio de las variables de decisión en base a una dirección de descenso a uno o varios objetivos en el espacio de los objetivos. La principal aportación del trabajo revisado que pertenece a la familia de algoritmos de optimización predictor-corrector, donde la sección predictora es aproximada mediante la resolución de una familia de curvas modeladas como un Problema del Valor Inicial, en especial esta fase conduce a la convergencia del frente de Pareto dada una dirección. Por otra parte, la fase correctora realiza un recorrido del frente de Pareto por medio en base al principio de realiza trayectorias ortogonales a direcciones previas. Esta última es resuelta sin la necesidad de calcular segundas derivadas que en resultado es otra aportación significativa. Algunas ideas de trabajo futuro son listadas a continuación

- El principio de seguir una trayectoria ortogonal para la fase correctora no podría ser lo mejor en geometrías irregulares, en este caso se sugiere replantear el problema para seguir el cono tangente al frente.
- El recorrido que se sigue a lo largo del frente de Pareto consiste en la resolución de un problema para hacer el cálculo de los vectores de pesos, el problema de esto es que no se asegura una diversidad en el espacio objetivo, es decir se pueden realizar recorridos agresivos a lo largo del frente de Pareto. Para resolver esto se sugiere replantear el modelo y agregar un estimador de densidad en el espacio objetivo (e.g. el estimador min-max ).
- En otros trabajos del mismo autor se incorpora este mecanismo como búsqueda local, en específico se estima el Jacobiano en base a los vecinos de cada solución, parece que sería útil utilizar este concepto de aproximación para ajustar la parte de la recombinación.
- Sería útil incorporar el criterio de paro en parte del cálculo del gradiente, ya que esta estrategia podría tener dificultades en problemas multi-frontales, multi-modales y/o deceptivos, esto se podría incorporar a partir del mecanismo utilizado en el NES de mono-objetivo.
- Particularmente, se observa que si el tamaño de paso es muy elevado el punto a actualizarse puede ciclarse a lo largo de las iteraciones, se podría buscar otra alternativa para calcular el tamaño de paso ya que es crítico en las estrategias de gradiente.
- Se pueden explorar otras estrategias de gradientes con principios similares, es decir, sólo actualizar ciertas variables en distintas iteraciones de forma aleatoria.

<sup>2</sup>ver video de simulación anexo



## References

- [1] Oliver Schütze, Adriana Lara, and CA Coello Coello. The directed search method for unconstrained multi-objective optimization problems. *Proceedings of the EVOLVE—A Bridge Between Probability, Set Oriented Numerics, and Evolutionary Computation*, pages 1–4, 2011.
- [2] Claus Hillermeier. *Nonlinear multiobjective optimization: a generalized homotopy approach*, volume 135. Springer Science & Business Media, 2001.
- [3] LA Zadeh. *DESOER, c. A.: 'Linear system theory'*. McGraw-Hill, 1963.
- [4] Alan S Manne. Multi-sector models for development planning: A survey. *Journal of Development Economics*, 1(1):43–69, 1974.
- [5] Michael Dellnitz, Oliver Schütze, and Stefan Sertl. Finding zeros by multilevel subdivision techniques. *IMA Journal of Numerical Analysis*, 22(2):167–185, 2002.
- [6] Martin Brown and Robert E Smith. Directed multi-objective optimization. *International Journal of Computers, Systems, and Signals*, 6(1):3–17, 2005.
- [7] Joerg Fliege, LM Grana Drummond, and Benar Fux Svaiter. Newton’s method for multiobjective optimization. *SIAM Journal on Optimization*, 20(2):602–626, 2009.
- [8] Jörg Fliege and Benar Fux Svaiter. Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, 51(3):479–494, 2000.
- [9] Peter AN Bosman. On gradients and hybrid evolutionary algorithms for real-valued multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 16(1):51–69, 2011.
- [10] Oliver Schütze, Adanay Martín, Adriana Lara, Sergio Alvarado, Eduardo Salinas, and Carlos A Coello Coello. The directed search method for multi-objective memetic algorithms. *Computational Optimization and Applications*, 63(2):305–332, 2016.
- [11] Oliver Schütze, Víctor Adrián Sosa Hernández, Heike Trautmann, and Günter Rudolph. The hypervolume based directed search method for multi-objective optimization problems. *Journal of Heuristics*, 22(3):273–300, 2016.
- [12] Evan James Hughes. Many-objective directed evolutionary line search. In *Proceedings of the 13th annual conference on Genetic and evolutionary computation*, pages 761–768. ACM, 2011.
- [13] Peter AN Bosman and Edwin D De Jong. Exploiting gradient information in numerical multi-objective evolutionary optimization. In *Proceedings of the 7th annual conference on Genetic and evolutionary computation*, pages 755–762. ACM, 2005.
- [14] AL Custódio, M Emmerich, and JFA Madeira. Recent developments in derivative-free multiobjective optimization. *Computational Technology Reviews*, 5(1):1–31, 2012.
- [15] Xiang Zhong, Wenhui Fan, Jinbiao Lin, and Zuozhi Zhao. A novel multi-objective compass search. In *2010 IEEE International Conference on Progress in Informatics and Computing*, volume 1, pages 24–29. IEEE, 2010.
- [16] Charles Audet, Gilles Savard, and Walid Zghal. A mesh adaptive direct search algorithm for multiobjective optimization. *European Journal of Operational Research*, 204(3):545–556, 2010.
- [17] Ana Luisa Custódio, JF Aguilar Madeira, A Ismael F Vaz, and Luís N Vicente. Direct multisearch for multiobjective optimization. *SIAM Journal on Optimization*, 21(3):1109–1140, 2011.
- [18] Charles Audet, Gilles Savard, and Walid Zghal. Multiobjective optimization through a series of single-objective formulations. *SIAM Journal on Optimization*, 19(1):188–210, 2008.
- [19] Nicola Beume, Marco Laumanns, and Günter Rudolph. Convergence rates of sms-emoa on continuous bi-objective problem classes. In *Proceedings of the 11th workshop proceedings on Foundations of genetic algorithms*, pages 243–252. ACM, 2011.
- [20] Michael E Henderson. Multiple parameter continuation: Computing implicitly defined k-manifolds. *International Journal of Bifurcation and Chaos*, 12(03):451–476, 2002.
- [21] Olvi L Mangasarian. *Nonlinear programming*. SIAM, 1994.
- [22] Larry Armijo. Minimization of functions having lipschitz continuous first partial derivatives. *Pacific Journal of mathematics*, 16(1):1–3, 1966.
- [23] Oliver Schütze, Marco Laumanns, Carlos A Coello Coello, Michael Dellnitz, and El-Ghazali Talbi. Convergence of stochastic search algorithms to finite size pareto set approximations. *Journal of Global Optimization*, 41(4):559–577, 2008.