Introducción a la Computación Evolutiva

Carlos A. Coello Coello

carlos.coellocoello@cinvestav.mx
CINVESTAV-IPN
Evolutionary Computation Group (EVOCINV)
Departamento de Computación
Av. IPN No. 2508, Col. San Pedro Zacatenco
México, D.F. 07360, MEXICO

Clase 14

Motivación

La mayor parte de los problemas del mundo real tienen varios objetivos que requerimos satisfacer, varios de los cuales suelen estar en conflicto. Por ejemplo, si queremos optimizar el diseño de un puente, normalmente queremos minimizar su costo y maximizar su seguridad. Estos objetivos se contraponen, pues un mayor peso hace el puente más seguro, pero más caro.

Definición Formal

Encontrar el vector $\vec{x}^* = \left[x_1^*, x_2^*, \ldots, x_n^*\right]^T$ que satisfaga las m restricciones de desigualdad:

$$g_i(\vec{x}) \leq 0 \quad i = 1, 2, \ldots, m \tag{1}$$

las p restricciones de igualdad

$$h_i(\vec{x}) = 0 \quad i = 1, 2, \ldots, p \tag{2}$$

y que optimice la función vectorial

$$\vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \ldots, f_k(\vec{x})]^T \tag{3}$$

Vector Objetivo Ideal

Es un vector que contiene los mínimos de las funciones objetivo, consideradas por separado. Los componentes z_i^* del **vector objetivo ideal** $\mathbf{z}^* \in \mathbb{R}^k$ se obtienen minimizando cada función objetivo individualmente, sujetas a las restricciones del problema.

Es decir, se obtienen resolviendo:

$$\text{minimizar } f_i(\vec{x}) \tag{4}$$

$$\text{sujeta a } \vec{x} \in \mathcal{F}, \text{ for } i = 1, \ldots, k \tag{5}$$

Vector Objetivo Utópico

Algunos autores, como Kaisa Miettinen consideran el concepto de **vector objetivo utópico**.

El vector objetivo utópico se define como: $\mathbf{z}^{**} \in \mathbb{R}^k$ y es un vector objetivo infactible cuyos componentes se forman mediante:

$$z^{**} = z_i^* - \varepsilon_i \tag{6}$$

para cada $i = 1, \ldots, k$, donde z_i^* es un componente del vector objetivo ideal y $\varepsilon_i > 0$ es un escalar que es relativamente pequeño, pero computacionalmente significativo. Claramente, el vector objetivo utópico es estrictamente mejor (es decir, estrictamente domina) a cada solución que sea un óptimo de Pareto.

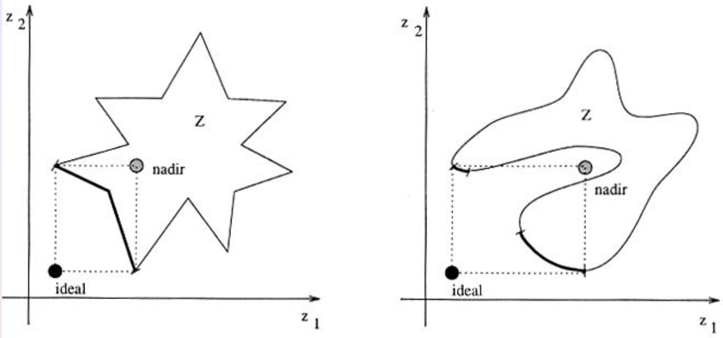
Vector Objetivo de Nadir

Se refiere a los límites superiores del conjunto de óptimos de Pareto. Se le denota normalmente como $\mathbf{z}^{nad}$ y sus componente suelen ser muy difíciles de obtener.

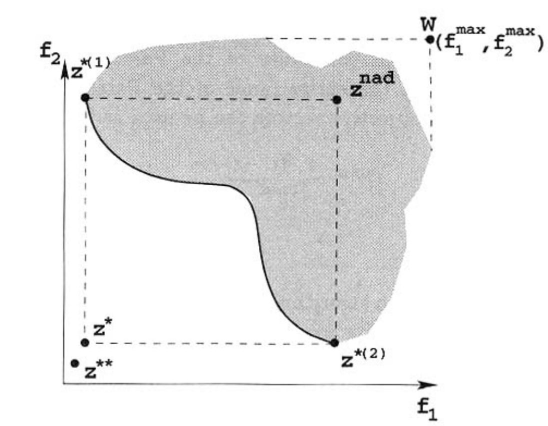
Para problemas con sólo dos objetivos, es posible utilizar un procedimiento determinista para encontrar el vector objetivo de Nadir (utilizando tablas).

Sin embargo, no existe procedimiento que asegure obtener dicho vector para problemas con 3 o más objetivos.

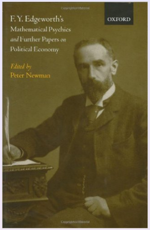
Optimización Evolutiva Multi-Objetivo



En esta figura, se usa un círculo negro para indicar los Vectores Objetivo Ideales y un círculo gris para indicar los vectores objetivo de Nadir. Es importante resaltar el hecho de que el Vector Objetivo de Nadir puede ser infactible.



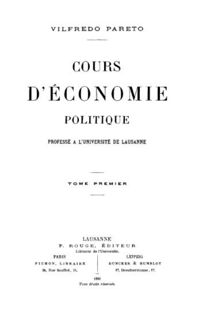
Esta figura muestra el Vector Objetivo Ideal (z^*), el Vector Objetivo Utópico (z^{**}) y el Vector Objetivo de Nadir (z^{nad}).



Noción de Optimalidad

Cuando tenemos varias funciones objetivo, la noción de "óptimo" cambia, porque en este caso lo que intentamos es encontrar buenos compromisos entre los objetivos, en vez de un óptimo global.

La noción de óptimo que suele adoptarse más comúnmente es la que propuso originalmente Francis Ysidro Edgeworth (en 1881) en su libro titulado **Mathematical Psychics**.



VILFREDO PARETO

COURS
D'ÉCONOMIE
POLITIQUE

PROFESSÉ A L'UNIVERSITÉ DE LAUSANNE

TOME PREMIER

LAUSANNE
F. ROUGE, ÉDITEUR
Librairie de l'Université.

PARIS
PICHON, LIBRAIRE
24, Rue Soufflot, 24.

LEIPZIG
DUNCKER & HUMBLOT
17, Dresdnerstrasse, 17.

1896
Tous droits réservés.

Optimalidad de Pareto

Esta noción fue propuesta de una forma más general por el economista italiano Vilfredo Pareto (en 1896) en su libro **Cours d'Economie Politique**.

Aunque algunos autores denominan *optimalidad de Edgeworth-Pareto* a esta noción que Pareto llamara **ophelimity**, el término más común es el de **óptimo de Pareto**.

Optimalidad de Pareto

Se dice que un vector de variables de decisión $\vec{x}^* \in \mathcal{F}$ es **Pareto óptimo** si no existe otro vector $\vec{x} \in \mathcal{F}$ tal que $f_i(\vec{x}) \leq f_i(\vec{x}^*)$ para toda $i = 1, \ldots, k$ y $f_j(\vec{x}) < f_j(\vec{x}^*)$ para al menos una j (suponiendo que estamos miniizando todos los objetivos).

En palabras esta definición dice que $\vec{x}^*$ es **Pareto óptima** si no existe otro vector de variables de decisión $\vec{x} \in \mathcal{F}$ que decremente algún criterio sin hacer que se incremente al mismo tiempo otro. Este concepto normalmente produce un conjunto de soluciones a las que se denomina **conjunto de óptimos de Pareto**. Los vectores $\vec{x}^*$ correspondientes a las soluciones incluidas en el conjunto de óptimos de Pareto se denominan **no dominadas**. La imagen del conjunto de óptimos de Pareto se denomina **frente de Pareto**.

Dominancia de Pareto

Se dice que un vector $\vec{u} = (u_1, \ldots, u_k)$ **domina a** $\vec{v} = (v_1, \ldots, v_k)$ (denotado mediante $\vec{u} \preceq \vec{v}$) si y sólo si u es parcialmente menor que v, es decir, $\forall i \in \{1, \ldots, k\},\ u_i \leq v_i \wedge \exists i \in \{1, \ldots, k\} : u_i < v_i$.

Conjunto de Óptimos de Pareto

Para un problema multi-objetivo $\vec{f}(x)$ dado, el conjunto de óptimos de Pareto ($\mathcal{P}^*$) se define como:

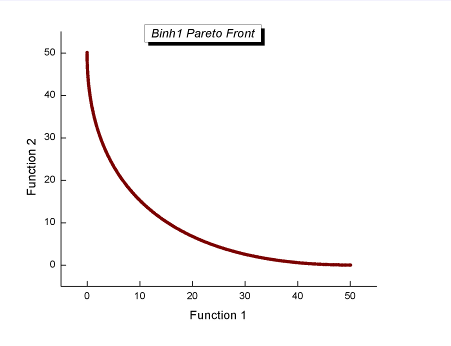
$$\mathcal{P}^* := \{x \in \mathcal{F} \mid \neg\exists\ x' \in \mathcal{F}\ \ \vec{f}(x') \preceq \vec{f}(x)\}. \tag{7}$$

Frente de Pareto

Para un problema multi-objetivo $\vec{f}(x)$ y un conjunto de óptimos de Pareto $\mathcal{P}^*$ dados, el frente de Pareto ($\mathcal{PF}^*$) se define como:

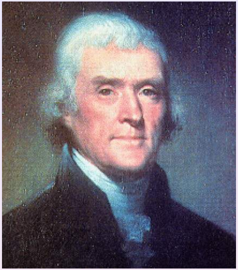
$$\mathcal{PF}^* := \{\vec{u} = \vec{f} = (f_1(x), \ldots, f_k(x)) \mid x \in \mathcal{P}^*\}. \tag{8}$$

Optimización Evolutiva Multi-Objetivo



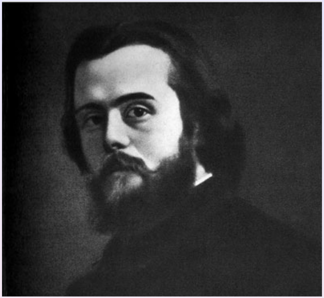
Frente de Pareto

En general, es imposible encontrar una expresión analítica que represente la línea o hiper-superficie correspondiente al frente óptimo de Pareto. Esto sólo es posible para problemas sumamente sencillos.



Un poco de historia

La optimización multi-objetivo es una parte intrínseca de la teoría de equilibrio económico y, como tal, puede decirse que sus orígenes se remontan al famoso tratado de Adam Smith titulado **An Inquiry into the Nature and Causes of the Wealth of Nations**, que data de 1776.



Un poco de historia

El concepto de equilibrio económico en general se suele atribuir a Léon Walras (1834-1910). Dentro de la teoría de equilibrio económico, los trabajos más relevantes (además de los de Walras) son los de Jevons y Menger, sobre teoría de la utilidad, así como el trabajo de Edgeworth y Pareto en torno al bienestar. Estos trabajos abarcan el período de 1874 a 1906.

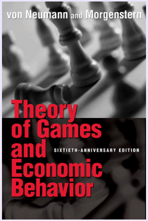
Un poco de historia

Una segunda área que se considera como una de las precursoras de la optimización multi-objetivo es la teoría sicológica de juegos y la noción de estrategias (de juego).

Los juegos de azar tienen una historia muy antigua. Sin embargo, es a Félix Édouard Justin Émile Borel (1871-1956) a quien normalmente se le considera como el iniciador de la teoría sicológica de los juegos. Fue también él el que introdujo la definición formal de estrategias que se basan en el análisis de la sicología del oponente.

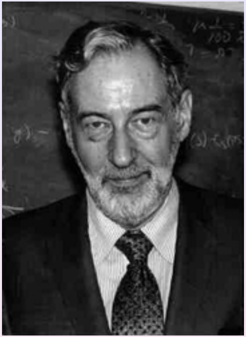
Un poco de historia

Los orígenes de la denominada **teoría de juegos** se remontan al trabajo realizado por Borel en 1921. Sin embargo, muchos historiadores normalmente atribuyen los orígenes de la teoría de juegos a un artículo escrito por el famoso matemático húngaro John von Neumann, el cual fue presentado oralmente en 1926 y se publicó en 1928.



Un poco de historia

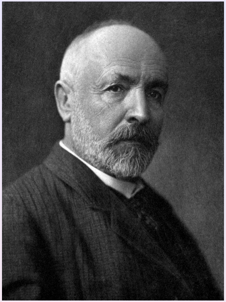
En 1944, John von Neumann y Oskar Morgenstern mencionaron (en su famoso libro sobre **Game Theory**) que habían encontrado un problema en economía que era una "mezcla peculiar y desconcertante de varios problemas en conflicto entre sí", el cual no podía resolverse con los métodos clásicos de optimización conocidos en esa época. Se desconocen las razones por las que Von Neumann no se interesó en resolver este problema tan peculiar.



Un poco de historia

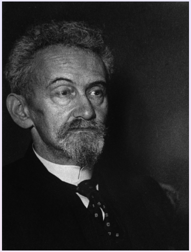
En 1951, Tjalling C. Koopmans editó un libro titulado **Activity Analysis of Production and Allocation**, en el cual se usa por primera vez el concepto de vector **eficiente** (el cual es el mismo que el de **vector no dominado**) de manera significativa.

Optimización Multi-Objetivo



Fundamentos matemáticos

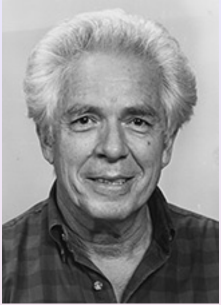
Los orígenes de los fundamentos matemáticos de la optimización multi-ojetivo se remontan al período de 1895 a 1906, en el que Georg Cantor y Felix Hausdorff establecieron los fundamentos de los espacios ordenados de infinitas dimensiones.



Fundamentos matemáticos

Cantor también introdujo las clases de equivalencia y estableció el primer conjunto de condiciones de suficiencia para la existencia de una función de utilidad.

Hausdorff proporcionó el primer ejemplo de un ordenamiento total.



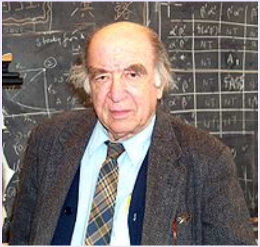
Fundamentos matemáticos

Sin embargo, fue el concepto del **problema del vector máximo**, introducido por Harold W. Kuhn y Albert W. Tucker (1951) el que le permitió a la optimización multi-objetivo convertirse en una disciplina matemática en sí misma.

Fundamentos matemáticos

Es bien sabido que las ahora famosas condiciones de optimalidad que suelen atribuirse a Kuhn y Tucker fueron planteadas y demostradas previamente por W. Karush en su tesis de maestría (cuyo resultado principal nunca se publicó) en 1939.

Kuhn y Tucker le dieron el crédito correspondiente a Karush, por lo cual actualmente se utiliza el término **Karush-Kuhn-Tucker (KKT) Conditions**.



Fundamentos matemáticos

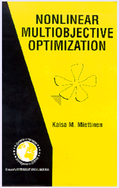
Sin embargo, los fundamentos teóricos de la optimización multi-objetivo permanecieron prácticamente inexplorados durante los 1950s. Fue hasta los 1960s en que los fundamentos matemáticos del área se consolidaron cuando Leonid Hurwicz generalizaron los resultados de Kuhn y Tucker a espacios vectoriales topológicos.

Kenneth J. Arrow realizó trabajo pionero muy importante en los 1950s usando el concepto de puntos admisibles y planteando su famoso *Teorema de la Imposibilidad*, el cual se relaciona con un dilema que se presenta dentro de la toma de decisiones multi-criterio de un grupo.

Quizás el desarrollo más importante de los 1950s fue **Goal Programming**, que fue introducida originalmente por Abraham Charnes y William Wager Cooper en 1957. Sin embargo, esta técnica se volvió popular en los 1960s.

La primera aplicación de optimización multi-objetivo fuera de la economía la realizó Koopmans (1951) en el área de teoría de la producción. Posteriormente, Marglin (1967) desarrolló la primera aplicación de optimización multi-objetivo en el área de recursos hidráulicos.

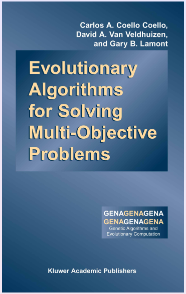
Optimización Multi-Objetivo



Estado Actual

Actualmente, existen unas 30 técnicas de programación matemática diseñadas para resolver problems de optimización multi-objetivo. Sin embargo, dichas técnicas tienen varias limitantes, de entre las cuales destaca su requerimiento de información específica del problema (p.ej., varias requieren que los objetivos y las restricciones sean diferenciable). Otras no pueden lidiar con frentes de Pareto no convexos o desconectados. Adicionalmente, la mayoría de estas técnicas generan una sola solución por cada ejecución algorítmica.

Optimización Evolutiva Multi-Objetivo



Los algoritmos evolutivos presentan varios ventajas al usárseles para resolver problemas multi-objetivo. La principal es que el uso de una población para realizar la búsqueda permite generar varios elementos del conjunto de óptimos de Pareto en una sola ejecución. Adicionalmente, son menos susceptibles a la forma y a la continuidad deol frente de Pareto que las técnicas de programación matemática y requieren de menos información específica del problema para operar.

Un poco de historia

El potencial de los algoritmos evolutivos para resolver problemas de optimización multi-objetivo se remonta a la tesis doctoral de Richard Rosenberg, que data de 1967.

Richard Rosenberg, **Simulation of genetic populations with biochemical properties**, PhD thesis, University of Michigan, Ann Arbor, Michigan, USA, June 1967.

Un poco de historia

Aunque la tesis doctoral de Rosenberg plantea, por primera vez, la posibilidad de usar algoritmos genéticos para resolver un problema multi-objetivo, no proporciona una implementación. La razón es que el problema multi-objetivo que intentaba resolver es transformado en un problema mono-objetivo (el segundo objetivo es transformado en una restricción).

La sugerencia de Rosenberg era usar múltiples *propiedades* (cercanía a una composición química específica) en su simulación de la genética y química de una población de organismos unicelulares. Puesto que su implementación adoptó sólo una propiedad, no se requirió un enfoque multi-objetivo.

VEGA

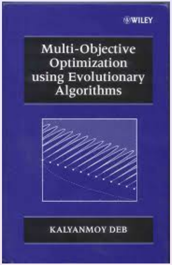
John David Schaffer fue el primero en desarrollar una implementación de un algoritmo evolutivo multi-objetivo (AEMO), la cual fue presentada en su tesis doctoral, que data de 1984.

John David Schaffer, **Multiple Objective Optimization with Vector Evaluated Genetic Algorithms**, PhD thesis, Vanderbilt University, USA, 1984.

VEGA

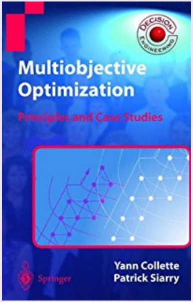
El algoritmo propuesto por Schaffer se llama **Vector Evaluated Genetic Algorithm** (VEGA) y fue publicado en la *First International Conference on Genetic Algorithms* en 1985. VEGA usa un método sub-poblacional de selección y no adopta la optimalidad de Pareto.

J. David Schaffer, "**Multiple Objective Optimization with Vector Evaluated Genetic Algorithms**" , in *Genetic Algorithms and their Applications: Proceedings of the First International Conference on Genetic Algorithms*, pp. 93–100, Lawrence Erlbaum, New Jersey, USA, 1985.



Los Viejos Días

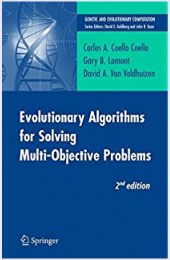
- Métodos No Elitistas y No Basados en Optimalidad de Pareto
  - Ordenamiento Lexicográfico
  - Funciones Agregativas Lineales
  - VEGA
  - Método de las restricciones-ε
  - Métodos Basados en Vectores Meta



Los Viejos Días

- Métodos No Elitistas Basados en Optimalidad de Pareto
  - Jerarquización de Pareto Pura
  - MOGA
  - NSGA
  - NPGA y NPGA 2

Optimización Evolutiva Multi-Objetivo



Algoritmos Contemporáneos

- Métodos Elitistas y Basados en la Optimalidad de Pareto
  - SPEA y SPEA2
  - NSGA-II
  - PAES, PESA y PESA II
  - Micro-genetic Algorithm for Multi-Objective Optimization and μGA^2
  - Los algoritmos que el mundo olvidó

Algoritmos Recientes

- MOEA/D
- Algoritmos Basados en Indicadores
  - SMS-EMOA
  - HyPE
  - Otros Algoritmos
- NSGA-III

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Dentro de este grupo, se incluyen los AEMOs más antiguos de la literatura especializada, los cuales no incorporan la optimalidad de Pareto en su mecanismo de selección y no retienen las soluciones no dominadas generadas durante el proceso evolutivo (es decir, no son elitistas).

Estos algoritmos son simples y eficientes, pero son también poco efectivos (sobre todo si se intentan aplicar a problemas con más de 3 funciones objetivo).

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Ordenamiento Lexicográfico

En este caso, las funciones objetivo deben ordenarse con base a su importancia.

El problema multi-objetivo original se re-plantea como:

$$\text{Minimizar} \quad f_1(\vec{x}) \tag{9}$$

sujeto a:

$$g_j(\vec{x}) \leq 0; \quad j = 1, 2, \ldots, m \tag{10}$$

y obtenemos $\vec{x}_1^*$ y $f_1^* = f(\vec{x}_1^*)$.

Ordenamiento Lexicográfico

Posteriormente, el segundo problema se plantea como:

$$\text{Minimizar} \quad f_2(\vec{x}) \tag{11}$$

sujeto a:

$$g_j(\vec{x}) \leq 0; \quad j = 1, 2, \ldots, m \tag{12}$$

$$f_1(\vec{x}) = f_1^* \tag{13}$$

y obtenemos $\vec{x}_2^*$ y $f_2^* = f_2(\vec{x}_2^*)$.

Ordenamiento Lexicográfico

Este procedimiento se repite hasta que se hayan considerado todos los objetivos.

Una variante popular de este algoritmo consiste en seleccionar aleatoriamente un objetivo a cada generación.

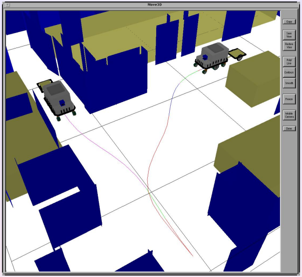
Otros autores han adoptado un esquema en el cual la selección se realiza comparando sólo con respecto al objetivo más importante. En este caso, si hay un empate, se adopta el segundo objetivo en importancia, y así sucesivamente.

Ordenamiento Lexicográfico

La **principal ventaja** de este algoritmo es su simplicidad, la cual conlleva una muy buena eficiencia computacional.

Su **principal desventaja** es que el desempeño de este algoritmo depende en el orden impuesto en los objetivos. Así mismo, este tipo de enfoques no es adecuado para problemas con más de 2 objetivos.

Métodos No Elitistas y No Basados en Optimalidad de Pareto



Ordenamiento Lexicográfico

Algunas aplicaciones de este algoritmo son las siguientes:

- Compactación de circuitos usando codificación simbólica [Fourman, 1985]
- Planeación de movimientos de robots [Gacôgne, 1999]
- Planeación de tripulaciones de una línea aérea [El Moudani et al., 2001]

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Funciones Agregativas Lineales

Este tipo de enfoques fue muy popular en los orígenes de la optimización evolutiva multi-objetivo y algunos investigadores todavía los utilizan (p.ej., en ingeniería y en Investigación de Operaciones).

La idea principal de este tipo de enfoque es muy simple: transformar un problema multi-objetivo en un problema escalar, realizando una suma ponderada de los objetivos:

$$\min \sum_{i=1}^{k} w_i f_i(\vec{x}) \tag{14}$$

donde $w_i \geq 0$ son los pesos que representan la importancia relativa de los k objetivos del problema (los objetivos deben escalarse). Normalmente se pre-supone que: $\sum_{i=1}^{k} w_i = 1$.

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Funciones Agregativas Lineales

Sus **ventajas principales** es que son un enfoque simple y eficiente.

Sus **principales desventajas** son: (1) es difícil definir un conjunto de pesos que permita la generación de la mayor parte del frente de Pareto, y (2) el hecho de que, sin importar qué pesos se adopten, este enfoque no puede generar porciones no convexas del frente de Pareto.

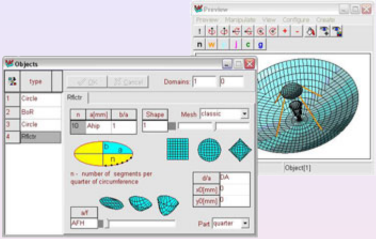
I. Das and J. Dennis, J. (1997), “**A Closer Look at Drawbacks of Minimizing Weighted Sums of Objectives for Pareto Set Generation in Multicriteria Optimization Problems**”, *Structural Optimization*, **14**(1):63–69.

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Funciones Agregativas Lineales

Durante varios años, el uso de funciones agregativas lineales dentro de un AEMO se ha visto como una mala idea. Sin embargo, existe evidencia sólida respecto a su utilidad en algunas clases de problemas (p.ej., en optimización combinatoria multi-objetivo).

Métodos No Elitistas y No Basados en Optimalidad de Pareto



Funciones Agregativas Lineales

Algunas aplicaciones de este tipo de técnica son las siguientes:

- Planeación de tareas [Jakob, 1992].
- Diseño de controladores [Donha, 1997].
- Diseño de filtros ópticos para lámparas [Eklund, 2001].
- Diseño de geometrías de antenas [Van Veldhuizen et al., 1998].

Métodos No Elitistas y No Basados en Optimalidad de Pareto

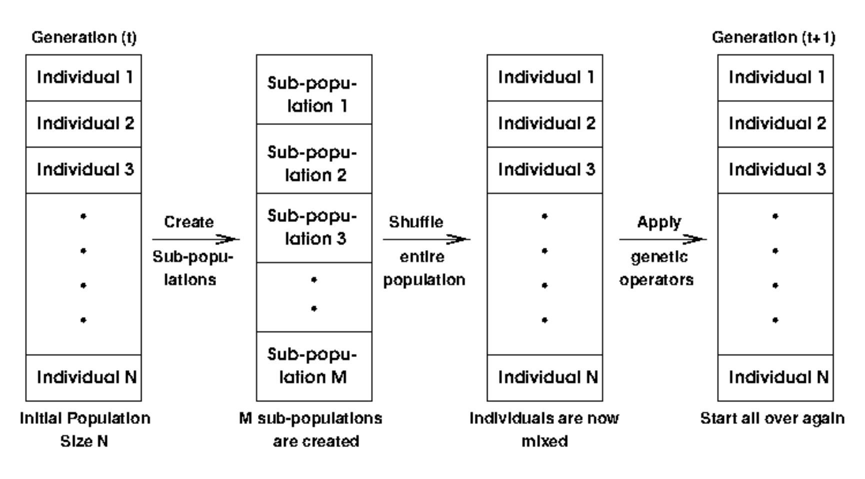
VEGA

Como se indicó anteriormente, el **Vector Evaluated Genetic Algorithm** (VEGA) fue propuesto por David Schaffer en 1984, como parte de su tesis doctoral titulada “Multiple Objective Optimization with Vector Evaluated Genetic Algorithms”.

La descripción de VEGA se publicón en las memorias de la *First International Conference in Genetic Algorithms* (ICGA’1985).

VEGA se aplicó originalmente a un problema de aprendizaje de máquina.

Métodos No Elitistas y No Basados en Optimalidad de Pareto



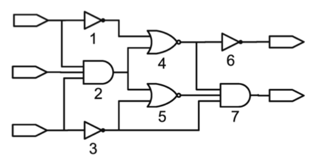
Métodos No Elitistas y No Basados en Optimalidad de Pareto

VEGA

Sus **ventajas principales** son: (a) es fácil de implementar (sólo debe modificarse el mecanismo de selección de un algoritmo genético simple y (2) es eficiente.

Sus **desventajas principales** son: (1) si se usa selección proporcional, VEGA tiene un desempeño similar al de una función agregativa lineal y (2) presenta un problema que Schaffer denominó “middling”. Básicamente, el mecanismo de selección de VEGA no sólo omite el uso de la optimalidad de Pareto, sino que además se opone a ésta. Adicionalmente, no cuenta con un mecanismo explícito para mantener diversidad.

Métodos No Elitistas y No Basados en Optimalidad de Pareto



VEGA

Algunas aplicaciones de este algoritmo son las siguientes:

- Contención de la contaminación de mantos acuíferos [Ritzel, 1994].
- Selección óptima de actuadores [Rogers, 2000].
- Diseño de circuitos lógicos combinatorios a nivel de compuertas [Coello et al., 2000].
- Manejo de restricciones [Surry, 1997; Coello, 2000].

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Método de las Restricciones-ε

En este caso, optimizamos primero la función objetivo de nuestra elección y los demás objetivos son considerados como restricciones acotadas por ciertos niveles permisibles, a los que denominamos ε_i.

Posteriormente, efectuamos una optimización mono-objetivo sujeta a dichas restricciones. Después de eso, modificamos los niveles ε_i y efectuamos otra optimización mono-objetivo. Este proceso se repite varias veces. Esto nos permitirá generar todo el conjunto de óptimos de Pareto aún en casos en que el frente de Pareto no sea convexo o esté desconectado. Evidentemente, en este caso, se usa un algoritmo evolutivo para realizar las optimizaciones mono-objetivo.

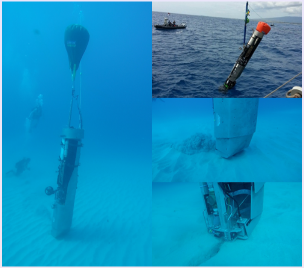
Métodos No Elitistas y No Basados en Optimalidad de Pareto

Método de las Restricciones-ε

La **principal ventaja** de esta técnica es su relativa simplicidad y su amplia aplicabilidad, pues puede resolver cualquier tipo de problema multi-objetivo.

Sus **principales desventajas** son: (1) requiere obtener el punto de Nadir, lo cual es difícil para problemas con más de 2 objetivos, y (2) es costoso computacionalmente, puesto que requiere muchas optimizaciones mono-objetivo.

Métodos No Elitistas y No Basados en Optimalidad de Pareto



Método de las Restricciones-ε

Algunas aplicaciones de este método son las siguientes:

- Diseño preliminar de vehículos marinos [Lee, 1997].
- Contaminación de mantos acuíferos [Chetan, 2000].
- Diseño tolerante a fallas [Schott, 1995].
- Ingeniería ambiental [Kumar, 2002].

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Métodos Basados en Metas

En este caso, se presupone que el usario (o tomador de decisiones) define un conjunto de metas que pretende cumplir para cada función objetivo. Posteriormente, el algoritmo evolutivo intenta minimizar las diferencias entre las soluciones actuales y este conjunto de metas (hay diferentes métricas que pueden usarse para determinar la cercanía con las metas).

Estrictamente hablando, estas técnicas también son funciones agregativas. Sin embargo, se consideran como un grupo aparte porque se basan en funciones agregativas no lineales y, por tanto, bajo ciertas condiciones, pueden generar porciones no convexas del frente de Pareto.

Métodos No Elitistas y No Basados en Optimalidad de Pareto

Métodos Basados en Metas

Los híbridos más populares en esta categoría son los siguientes:

- EA + Goal Programming [Deb, 1999; Wienke, 1992; Sandgren, 1994]
- EA + Goal Attainment [Wilson, 1993; Zebulum, 1998]
- EA + Min-Max Optimum [Hajela, 1992; Coello, 1998]

Métodos No Elitistas y No Basados en Optimalidad de Pareto

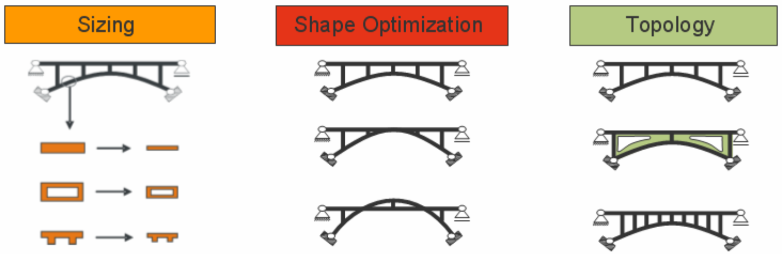
Métodos Basados en Metas

Sus **principales ventajas** son su eficiencia computacional y su simplicidad.

Sus **principales desventajas** tienen que ver con la dificultad para generar las metas a las que se aspira. Así mismo, bajo ciertas circunstancias, este tipo de técnica puede presentar un comportamiento ambiguo.

Adicionalmente, algunas de estas técnicas requieren que las metas se definan en la región factible a fin de garantizar que las soluciones generadas sean óptimos de Pareto.

Métodos No Elitistas y No Basados en Optimalidad de Pareto

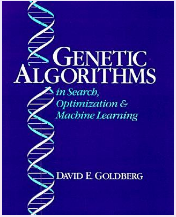


Métodos Basados en Metas

Algunas aplicaciones de este tipo de técnicas son las siguientes:

- Diseño de filtros digitales IIR [Wilson, 1993].
- Optimización estructural [Sandgren, 1994; Hajela, 1992]
- Balance de contrapesos de brazos robóticos [Coello, 1998].

Métodos No Elitistas Basados en la Optimalidad de Pareto



Orígenes

David Goldberg critica VEGA en su famoso libro de algoritmos genéticos [Goldberg, 1989] y propone la idea de usar un esquema de asignación de aptitud basado en la optimalidad de Pareto para un algoritmo evolutivo multi-objetivo.

Métodos No Elitistas Basados en la Optimalidad de Pareto

Orígenes

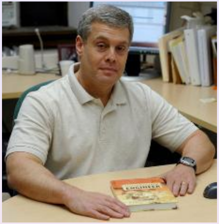
Goldberg propuso identificar las soluciones no dominadas en la población. A estos individuos se les asignaría una mejor aptitud que a los individuos dominados. Posteriormente, los individuos no dominados serían removidos de la población temporalmente de manera que pudiera efectuarse una nueva jerarquización. Este proceso se repetiría hasta que toda la población hubiese recibido un valor de aptitud. A este algoritmo se le conoce actualmente como **ordenamiento no dominado** y al proceso que realiza se le conoce como **jerarquización de Pareto**.

Métodos No Elitistas Basados en la Optimalidad de Pareto

Orígenes

Goldberg también indicó que, debido al ruido estocástico, un algoritmo genético multi-objetivo convergería eventualmente a una solución única. De tal forma, sugirió bloquear el proceso de selección a fin de mantener diversidad. Esto permitiría la generación de varios elementos del conjunto de óptimos de Pareto en una sola ejecución. A este mecanismo se le conoce hoy en día como **estimador de densidad**.

Métodos No Elitistas Basados en la Optimalidad de Pareto



Orígenes

Goldberg propuso utilizar **fitness sharing**, que es una técnica que él mismo introdujo en 1987. La idea es subdividir la población en varias subpoblaciones con base en la similitud de sus individuos. La similitud se puede medir en el espacio fenotípico (o sea, los valores decodificados) o en el espacio genotípico.

Métodos No Elitistas Basados en la Optimalidad de Pareto

Orígenes

Se utiliza la siguiente ecuación:

$$\phi(d_{ij}) = \begin{cases} 1 - \left(\frac{d_{ij}}{\sigma_{sh}}\right)^{\alpha}, & d_{ij} < \sigma_{share} \\ 0, & \text{de lo contrario} \end{cases} \tag{15}$$

Donde: $\alpha = 1$, d_{ij} es una métrica que indica la distancia entre las soluciones i y j, y σ_{share} es un parámetro definido por el usuario (también denominado **radio de nicho**).

La aptitud de un invidiuo i se modifica usando:

$$f_{s_i} = \frac{f_i}{\sum_{j=1}^{M} \phi(d_{ij})} \tag{16}$$

dobde M es el número de individuos que se localizan en el vecindario del individuo i. Idealmente, un individuo debería estar solo en su nicho.

Métodos No Elitistas Basados en la Optimalidad de Pareto

Jerarquización de Pareto Pura

Se refiere al uso de jerarquización de Pareto sin adoptar un estimador de densidad. Este tipo de enfoque ya no se utiliza en la actualidad, dado que se considera al estimador de densidad como un componente indispensable de un algoritmo evolutivo multi-objetivo.

El uso principal de este tipo de técnica es para aplicaciones en las que el usuario se interesa en obtener una sola solución (o unas pocas) no dominada(s) y no el frente de Pareto completo. Se adoptón sobre todo en aplicaciones de ingeniería

Métodos No Elitistas Basados en la Optimalidad de Pareto

Jerarquización de Pareto Pura

Ejemplos de aplicaciones:

- Monitoreo de mantos acuíferos [Cieniawski, 1995].
- Horarios de bombeo para suministro de agua [Schwab, 1996;Savic, 1997].
- Diseño preliminar de submarinos [Thomas, 1998].
- Planeación de sistemas de distribución de potencia [Ramírez, 1999].

MOGA

La propuesta algorítmica de Goldberg puede mejorarse y eso fue precisamente lo que hizo Carlos M. Fonseca en el denominado **Multi-Objective Genetic Algorithm** (MOGA), que propuso (junto con Peter Fleming) en 1993.

Carlos M. Fonseca and Peter J. Fleming, "**Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization**, in Stephanie Forrest (Ed), *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp. 416–423, Morgan Kauffman Publishers, San Mateo, California, 1993.

MOGA

En MOGA, la jerarquía de un individuo está dada por:

$$\mathrm{rank}(x_i, t) = 1 + p_i^{(t)} \tag{17}$$

A cada individuo no dominado se le asigna una jerarquía de 1 y los individuos dominados son penalizados con base en el número de soluciones que lo dominan.
La aptitud se asigna usando el procedimiento siguiente:

- Ordenar la población con base en su jerarquía.
- Asignar aptitud a cada individuo interpolando de la mejor jerarquía (1) a la peor ($n \leq M$), donde M es el tamaño de población. La interpolación es usualmente, pero no necesariamente, lineal.
- Promediar los valores de aptitud de los individuos con la misma jerarquía, de forma que todos sean muestreados de la misma forma.
- MOGA usa fitness sharing pero Fonseca proporciona un procedimiento para calcular σ_{share}).

Métodos No Elitistas Basados en la Optimalidad de Pareto

MOGA

MOGA se popularizón a mediados de los 1990s, porque era no sólo relativamente eficiente ($O(N^2)$), sino también muy efectivo. Adicionalmente, fue implementado en Matlab, lo que motivó a varios investigadores de control automático a utilizarlo.

Algunos estudios comparativos de mediados de los 1990s mostraron que MOGA era el mejor AEMO de la primera generación.

La implementación de MOGA no es tan simple, puesto que contiene diversos detalles complicados. De hecho, Fonseca afirma que su implementación incluía elitismo, pero ese mecanismo no se describe en su artículo de 1993.

Métodos No Elitistas Basados en la Optimalidad de Pareto

MOGA

Algunas aplicaciones de MOGA son las siguientes:

- Co-síntesis de sistemas embebidos distribuidos heterogéneos jerárquicos [Dick, 1998].
- Optimización de controladores magnéticos activos distribuidos [Schroder, 1997].
- Diagnóstico de fallas de procesos [Marcu, 1999].
- Optimización de armaduras planas [Narayanan, 1999].
- Manejo forestal [Ducheyne, 2001].
- Diseño de turbinas de gas [Fonseca, 1995].

Métodos No Elitistas Basados en la Optimalidad de Pareto

NSGA

El **Nondominated Sorting Genetic Algorithm** (NSGA) fue propuesto por Srinivas y Deb en 1994. Fue el primer AEMO publicado en una revista especializada (*Evolutionary Computation*).

NSGA sigue fielmente la descripción informal del proceso de jerarquización de Pareto que un AEMO debe adoptar.

Métodos No Elitistas Basados en la Optimalidad de Pareto

NSGA

El NSGA se publicó en:
N. Srinivas and Kalyanmoy Deb, **Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms**, *Evolutionary Computation*, **2**(3):221-248, Fall 1994.

Métodos No Elitistas Basados en la Optimalidad de Pareto

NSGA

El NSGA identifica primero a los individuos que son no dominados con respecto a toda la población. A tales individuos se les asigna un valor de aptitud "falso" y luego se les remueve de la población. Este proceso se repite con el resto de la población, asignando valores falsos de aptitud que se decrementan con cada capa (de tal forma que la primera capa tenga los valores más elevados). Este proceso termina cuando todos los individuos cuentan con un valor de aptitud.

El estimador de densidad también es *fitness sharing* (usando los valores falsos de aptitud asignados por el procedimiento de ordenamiento no dominado). Sin embargo, en este caso, la similitud se mide en el espacio de las variables de decisión (a diferencia de MOGA, que lo hace en el espacio de las funciones objetivo).

Métodos No Elitistas Basados en la Optimalidad de Pareto

NSGA

El algoritmo de ordenamiento no dominado adoptado por el NSGA tiene $O(N^3)$, el cual contrasta con el procedimiento de jerarquización de Pareto (de una pasada) de MOGA, el cual es $O(N^2)$.

Los pocos estudios comparativos de los 1990s indican que el NSGA era lento y no producía buenos resultados. Así mismo, era muy sensible al valor de σ_{share}.

El MOGA es discutido el artículo del NSGA. Sin embargo, los resultados del NSGA sólo se compararon con respecto a los de VEGA. La mayoría de los estudios comparativos de la época indicaron que MOGA le ganaba consistentemente al NSGA.

Métodos No Elitistas Basados en la Optimalidad de Pareto

NSGA

Algunas aplicaciones del NSGA son las siguientes:

- Monitoreo de mantos acuíferos [Reed et al., 2001].
- Optimización de controladores difusos para diseño de sistemas para guiar misiles [Blumel, 2001]
- Diseño de constelaciones de satélites [Mason, 1999].
- Optimización multi-objetivo en dinámica de fluidos [Marco, 1999].

Métodos No Elitistas Basados en la Optimalidad de Pareto

NPGA

El *Niched-Pareto Genetic Algorithm* (NPGA) fue propuesto por Jeffrey Horn en un reporte técnico de 1993 y fue publicado en un congreso internacional en 1994.

Jeffrey Horn, Nicholas Nafpliotis and David E. Goldberg, "**A Niched Pareto Genetic Algorithm for Multiobjective Optimization**", in *Proceedings of the First IEEE Conference on Evolutionary Computation*, Vol. 1, pp. 82–87, IEEE Press, Piscataway, New Jersey, USA, June 1994.

Métodos No Elitistas Basados en la Optimalidad de Pareto

NPGA

El NPGA adopta una variante del torneo binario en la que dos individuos seleccionados aleatoriamente compiten en términos de la dominancia de Pareto. Cada uno de estos dos individuos se compara con respecto a una muestra de la población cuyo tamaño es un parámetro definido por el usuario (normalmente, se adopta el 10% del tamaño total de la población). El torneo sólo tiene dos resultados posibles:

1. Uno de los dos individuos es no dominado y el otro es dominado. En este caso, el individuo no dominado gana el torneo y, por tanto, es seleccionado.
2. Hay un empate (los dos individuos son no dominados o ambos son dominados). En este caso, se aplica *fitness sharing* a los dos individuos. La técnica adoptada se denomina *equivalence class sharing* y se aplica tanto en el espacio de las variables de decisión como en el de las funciones objetivo.

Métodos No Elitistas Basados en la Optimalidad de Pareto

NPGA

Horn mostró que una pequeña muestra de la población era suficiente para estimar la optimalidad de Pareto de un individuo y produjo el AEMO más rápido de su época.

Éste es el único AEMO en el que David Goldberg aparece como co-autor (Goldberg fue el asesor de tesis doctoral de Horn en la Universidad de Illinois en Urbana-Champaign).

Los pocos estudios comparativos de finales de los 1990s indican que el NPGA era mejor que el NSGA, pero no mejor que el MOGA. De hecho, si el tamaño de torneo se hace igual que el tamaño de la población, el NPGA se volvería el MOGA.

Métodos No Elitistas Basados en la Optimalidad de Pareto

NPGA

Algunas aplicaciones del NPGA son las siguientes:

- Espectroscopía de plasma [Golovkin, 2000].
- Selección de características [Emmanouilidis, 2000].
- Diseño tolerante a fallas [Schott, 1995].
- Reducción de contaminación generada por el tráfico [Haastrup & Pereira, 1997].

Métodos No Elitistas Basados en la Optimalidad de Pareto

NPGA 2

Erickson et al. [2001] propusieron una variante del NPGA llamada NPGA 2.

Mark Erickson, Alex Mayer and Jeffrey Horn, "**The Niched Pareto Genetic Algorithm 2 Applied to the Design of Groundwater Remediation Systems**", in Eckart Zitzler, Kalyanmoy Deb, Lothar Thiele, Carlos A. Coello Coello and David Corne (Eds), *First International Conference on Evolutionary Multi-Criterion Optimization*, pp. 681–695, Springer-Verlag, Lecture Notes in Computer Science No. 1993, 2001.

NPGA 2

NPGA 2 adopta la jerarquización de Pareto usando la selección mediante torneo binario del NPGA. Sin embargo, se usa un esquema diferente de *fitness sharing* en el cual el conteo de nichos se calcula usando individuos de la siguiente generación (parcialmente llena), en vez de usar individuos de la generación actual. A este esquema se le denomina **continuously updated fitness sharing** [Oei, 1991].

NPGA 2 se ha utilizado sólo para el diseño de sistemas de remediación de acuíferos [Erickson, 2001].

Métodos Elitistas Basados en la Optimalidad de Pareto

¿Qué es el Elitismo?

En optimización mono-objetivo, el **elitismo** es un operador que hace que la mejor solución de la población pase intacta a la siguiente generación (es decir, no es perturbado ni por la cruza ni por la mutación).

En optimización multi-objetivo, el elitismo opera de forma similar, pero en este caso, necesitamos retener las soluciones no dominadas generadas por el AEMO. Puesto que es impráctico retener todas estas soluciones, normalmente se establece un límite al número máximo de soluciones que se retienen. Esto es particularmente importante en los AEMOs en los que el elitismo juega un papel primordial en el mecanismo de selección (p.ej., SPEA).

Métodos Elitistas Basados en la Optimalidad de Pareto

Formas de Elitismo

Las dos formas principales en las que se suele implementar el elitismo son las siguientes:

1. A través del uso de un **archivo externo** (llamado también **población secundaria**), el cual es una estructura de datos que reside en memoria principal y que almacena las soluciones no dominadas generadas durante el proceso evolutivo.
2. Usando un **mecanismo de selección más** en el que la población de padres se combina con la población de hijos y se retiene sólo a la mejor mitad.

Métodos Elitistas Basados en la Optimalidad de Pareto

¿Por qué es importante el elitismo?

Se requiere el elitismo para garantizar convergencia de un AEMO al verdadero conjunto de óptimos de Pareto de un problema de optimización multi-objetivo. Esto fue demostrado por Rudolph y Agapie [2001].

Günter Rudolph and Alexandru Agapie, "**Convergence Properties of Some Multi-Objective Evolutionary Algorithms**", in *Proceedings of the 2000 IEEE Conference on Evolutionary Computation*, Vol. 2, pp. 1010–1016, IEEE Press, Piscataway, New Jersey, USA July 2000.

Métodos Elitistas Basados en la Optimalidad de Pareto

SPEA

Eckart Zitzler [1998,1999] propuso en su tesis doctoral el **Strength Pareto Evolutionary Algorithm** (SPEA) como un AEMO que integra diferentes mecanismos de algoritmos propuestos previamente.

Métodos Elitistas Basados en la Optimalidad de Pareto

SPEA

Eckart Zitzler and Lothar Thiele, "**Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach**", *IEEE Transactions on Evolutionary Computation*, **3**(4):257-271, November 1999.

SPEA utiliza un archivo externo y popularizó esta forma de elitismo en los AEMOs. A cada generación, las soluciones no dominadas de la población se copian a este archivo, y el archivo participa en el proceso de selección. Para cada individuo en el archivo externo, se calcula un valor de "fortaleza". Este valor es similar a la jerarquía del MOGA, puesto que es proporcional al número de soluciones a las que domina un cierto individuo.

Métodos Elitistas Basados en la Optimalidad de Pareto

SPEA

La aptitud de cada individuo en la población se calcula con base en las fortalezas de todos los individuos del archivo externo a las cuales domina un cierto individuo.

Zitzler se percató de que si el tamaño del archivo no se acotaba, la presión de selección se diluiría dado que el número de soluciones no dominadas crecía muy rápidamente. Por tanto, decidió podar el archivo usando un algoritmo de clustering llamado **average linking method** [Morse, 1989], una vez que se alcanzaba un cierto límite pre-establecido.

Métodos Elitistas Basados en la Optimalidad de Pareto

SPEA

Algunas aplicaciones de SPEA son las siguientes:

- Exploración de horarios para software usado en procesadores digitales de señales [Zitzler, 1999].
- Planeación de tratamientos médicos [Petrovski, 2001].
- Optimización de dosis de medicamentos [Lahanas, 2001].
- Diagnóstico de enfermedades [de Toro, 2003].
- Rehabilitación de un sistema de distribución de agua [Cheung, 2003].

Métodos Elitistas Basados en la Optimalidad de Pareto

SPEA2

Eckart Zitzler y sus colegas propusieron una revisión mejorada del SPEA (llamada SPEA2) en 2001.

Eckart Zitzler, Marco Laumanns and Lothar Thiele, "**SPEA2: Improving the Strength Pareto Evolutionary Algorithm**, in K. Giannakoglou, D. Tsahalis, J. Periaux, P. Papailou and T. Fogarty (eds.), *EUROGEN 2001, Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems*, pp. 95–100, Athens, Greece, 2002.

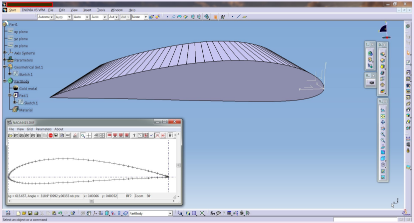
Métodos Elitistas Basados en la Optimalidad de Pareto

SPEA2

SPEA2 tiene 3 diferencias principales con respecto al SPEA original:

- Incorpora una estrategia de asignación de aptitud de grano fino que considera tanto al número de individuos a los que una solución domina como el número de soluciones por las que el individuo es dominado.
- Un estimador de densidad más eficiente (un mejor algoritmo de clustering).
- Un mecanismo para truncar el archivo externo, el cual garantiza que la retención de las soluciones en los extremos del frente de Pareto.

Métodos Elitistas Basados en la Optimalidad de Pareto



SPEA2

Algunas aplicaciones del SPEA2 son las siguientes:

- Reducción del “bloat” en programación genética [Bleuler, 2001].
- Diseño de perfiles aerodinámicos [Willmes, 2003].
- Optimización de portafolios de inversión [Garcia, 2011].
- Optimización de las emisiones de los motores diesel [Hiroyasu, 2005].

Métodos Elitistas Basados en la Optimalidad de Pareto

NSGA-II

El *Nondominated Sorting Genetic Algorithm II* (NSGA-II) fue propuesto por Kalyanmoy Deb y sus estudiantes en el 2000. Sin embargo, la versión más conocida es la publicada en una revista en 2002.

Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal and T. Meyarivan, "**A Fast and Elitist Multiobjective Genetic Algorithm: NSGA–II**, *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, pp. 182–197, April 2002.

NSGA-II

NSGA-II es un algoritmo muy distinto del NSGA original. Aunque sigue utilizando el ordenamiento no dominado, clasifica en una sola pasada (como MOGA). Así mismo, usa una selección más (la población de padres se une a la de hijos y se retiene a la mejor mitad), la cual es implícitamente elitista.

Un elemento fundamental del NSGA-II es su estimador de densidad, el cual es denominado **crowded comparison operator**. Esta técnica requiere ordenar las soluciones con respecto a un objetivo. Posteriormente, cada individuo usa a su vecino anterior y posterior para construir un rectángulo. Al comparar dos soluciones, si hay un empate (ya sea porque las dos son dominadas o ambas son no dominadas), se prefiere a la que tenga un rectángulo con un perímetro más grande (es decir, se prefieren soluciones que están más aisladas en el espacio de las funciones objetivo). Este estimador de densidad no requiere parámetros y es muy eficiente, aunque no es escalable.

Métodos Elitistas Basados en la Optimalidad de Pareto

NSGA-II

La elegancia, efectividad y eficiencia del NSGA-II lo convirtió en un estándar en el área durante más de 10 años.

El hecho de que su código fuente está disponible en internet contribuyó también a su popularidad (tal vez sea el AEMO más popular de todos los tiempos).

Métodos Elitistas Basados en la Optimalidad de Pareto

NSGA-II

Sin embargo, el NSGA-II no escala adecuadamente y su desempeño suele degradarse a partir de 3 objetivos, debido a su estimador de densidad que fue diseñado sólo para 2 objetivos.

Adicionalmente, existe evidencia experimental que indica que el NSGA-II trabaja mejor con números reales que usando codificación binaria.

Métodos Elitistas Basados en la Optimalidad de Pareto

NSGA-II

Algunas aplicaciones del NSGA-II son las siguientes:

- Optimización de formas [Deb, 2001].
- Diseño óptimo de sistemas de seguridad [Greiner, 2003].
- Optimización de las condiciones de procesamiento para la extrusión de polímeros (Gaspar-Cunha, 2002).
- Manejo de la calidad del agua [Dorn, 2003].
- Optimización de dosis para terapias contra el cáncer usando radiación mediante rayos de intensidad modulada [Lahanas, 2003].

Métodos Elitistas Basados en la Optimalidad de Pareto

PAES

La *Pareto Archived Evolution Strategy* (PAES) fue propuesta en 1999, aunque su versión de revista apareció en el 2000.

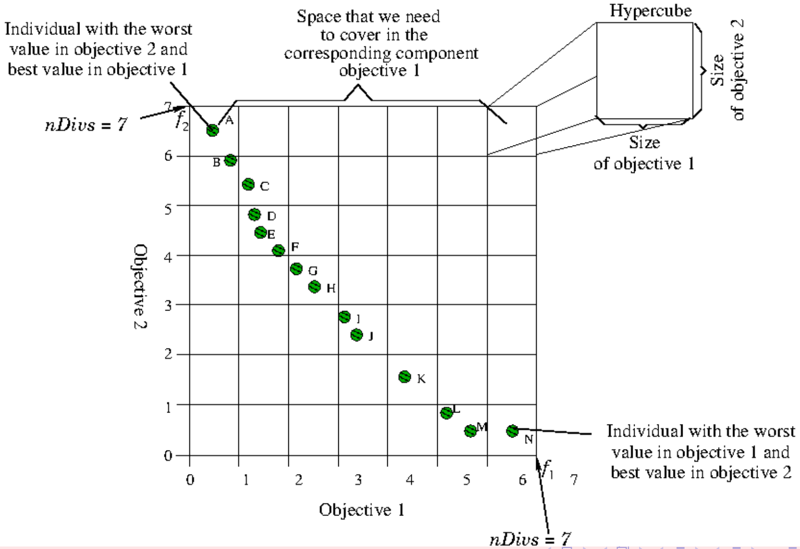
Métodos Elitistas Basados en la Optimalidad de Pareto

PAES

Conceptualmente, PAES es quizás el AEMO más simple que puede diseñarse. Consiste de una (1+1)-ES (o sea, un solo padre que se muta para producir un hijo). Si el hijo domina a su padre, se almacena en un archivo externo y se vuelve padre en la siguiente iteración.

El aspecto más interesante de PAES es precisamente su archivo externo, el cual adopta un estimador de densidad llamado **rejilla adaptativa.** Este estimador de densidad sólo requiere un parámetro: el número de sub-divisiones que se aplicarán en el espacio de las funciones objetivo. Su problema principal es que este mecanismo se concibió sólo para dos objetivos y su generalización a un número mayor de objetivos no parece posible.

Métodos Elitistas Basados en la Optimalidad de Pareto



Métodos Elitistas Basados en la Optimalidad de Pareto

PAES

Algunas aplicaciones de PAES son los siguientes:

- Problemas de telecomunicaciones [Knowles, 1999].
- El problema de la base de datos distribuida adaptativa [Knowles, 2000].
- Flexible job shop scheduling [Rabiee, 2012].

PESA

El *Pareto Envelope-based Selection Algorithm* (PESA) fue propuesto por David Corne en 2001.

David W. Corne, Joshua D. Knowles and Martin J. Oates, "**The Pareto Envelope-based Selection Algorithm for Multiobjective Optimization**", in Marc Schoenauer et al. (editors), *Proceedings of the Parallel Problem Solving from Nature VI Conference*, pp. 839–848. Springer, 2000.

Métodos Elitistas Basados en la Optimalidad de Pareto

PESA

PESA usa una pequeña población interna y una población externa más grande (similar a la que adopta PAES).

Al igual que PAES, PESA mantiene implícitamente una división de hiper-rejillas del espacio de las funciones objetivo, lo que le permite llevar un registro del grado de agrupamiento en regiones diferentes del archivo. Sin embargo, en este caso, y a diferencia de PAES, el mecanismo de selección se basa en esta medida de agrupamiento. Esta misma medida de agrupamiento se usa también para decidir qué soluciones ingresarán al archivo externo.

Métodos Elitistas Basados en la Optimalidad de Pareto

PESA-II

El *Pareto Envelope-based Selection Algorithm-II* (PESA II) fue propuesto por David Corne en 2001.

David W. Corne, Nick R. Jerram, Joshua D. Knowles and Martin J. Oates, "**PESA-II: Region-based Selection in Evolutionary Multiobjective Optimization**", in Lee Spector et al. (editors), *Proceedings of the 2001 Genetic and Evolutionary Computation Conference (GECCO'2001)*, pp. 283–290, Morgan Kaufmann Publishers, San Francisco, California, July 2001.

Métodos Elitistas Basados en la Optimalidad de Pareto

PESA-II

La única diferencia de PESA-II con respecto a PESA es que, en este caso, se adopta un esquema de selección basado en regiones.

En PESA-II, en vez de asignar una aptitud selectiva a un individuo, ésta se asigna a las hiper-cajas (en el espacio de las funciones objetivo) que son ocupadas actualmente por al menos un individuo en la aproximación actual al frente de Pareto. De tal forma, se selecciona una hiper-caja y después se selecciona un individuo aleatoriamente de dicha hiper-caja. Los autores de PESA-II argumentan que este esquema produce una mejor distribución de soluciones que el esquema de selección original de PESA.

Puesto que la rejilla adaptativa no es escalable, los esquemas de selección tanto de PESA como de PESA-II no pueden utilizarse con más de 2 funciones objetivo.

PESA y PESA-II se usaron para resolver algunos problemas de telecomunicaciones [Corne, 2000; Corne, 2001].

Métodos Elitistas Basados en la Optimalidad de Pareto

El Micro-Genetic Algorithm for Multi-Objective Optimization

Fue propuesto por Coello and Toscano [2001]. Un micro-algoritmo genético tiene un tamaño de población muy pequeño (no más de 5 individuos) y adopta un proceso de re-inicialización para mantener diversidad.

Carlos A. Coello Coello and Gregorio Toscano Pulido, "**Multiobjective Optimization using a Micro-Genetic Algorithm**, in Lee Spector et al. (editors), *Proceedings of the 2001 Genetic and Evolutionary Computation Conference (GECCO'2001)*, pp. 274–282, Morgan Kaufmann Publishers, San Francisco, California, July 2001.

Métodos Elitistas Basados en la Optimalidad de Pareto

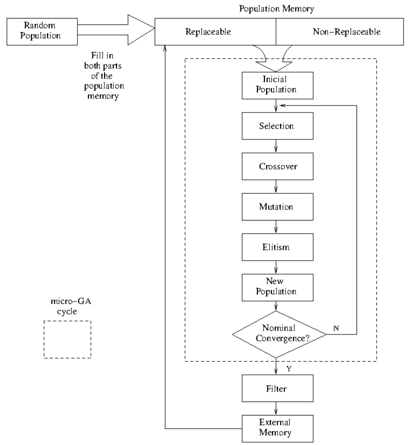
El Micro-Genetic Algorithm for Multi-Objective Optimization

Este algoritmo incorpora 3 formas de elitismo y la rejilla adaptativa de PAES (con una pequeña modificación que la hace más eficiente).

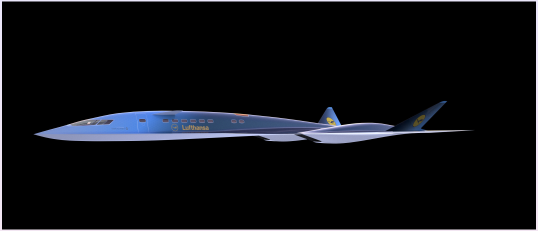
Este AEMO es muy rápido. En los estudios comparativos de la época, se pudo determinar que era hasta un order de magnitud más rápido que el NSGA-II y producía soluciones de calidad similar.

La principal desventaja de esta técnica era que requería un elevado número de parámetros (ocho, de los cuales al menos 3 juegan un papel preponderante en su desempeño).

Métodos Elitistas Basados en la Optimalidad de Pareto



Métodos Elitistas Basados en la Optimalidad de Pareto



El Micro-Genetic Algorithm for Multi-Objective Optimization

Algunas de sus aplicaciones son las siguientes:

- Diseño de jets supersónicos de negocios [Chung et al., 2003].
- Optimización estructural [Coello, 2002].
- Particiones de sistemas de hardware/software [Fornaciari, 2003].
- Localización de reguladores de voltaje automáticos en una red de distribución radial [Mendoza, 2007].

El Micro-Genetic Algorithm for Multi-Objective Optimization 2

Este algoritmo (denominado μGA^2, de forma abreviada) fue propuesto por Toscano y Coello en 2003. A la fecha, es el único AEMO totalmente auto-adaptativo que se ha propuesto en la literatura especializada.

Gregorio Toscano Pulido and Carlos A. Coello Coello, "**The Micro Genetic Algorithm 2: Towards Online Adaptation in Evolutionary Multiobjective Optimization**", in Carlos M. Fonseca et al. (editors), *Evolutionary Multi-Criterion Optimization. Second International Conference, EMO 2003*, pp. 252–266, Springer. Lecture Notes in Computer Science. Volume 2632, Faro, Portugal, April 2003.

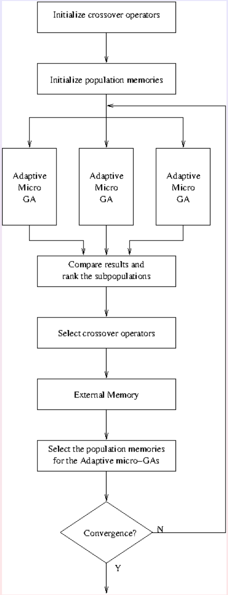
El Micro-Genetic Algorithm for Multi-Objective Optimization 2

La motivación principal del μGA^2 era eliminar los 8 parámetros del algoritmo original. El μGA^2 utiliza mecanismos de adaptación en línea que hacen innecesario el ajuste de cualquiera de sus parámetros.

El μGA^2 puede decidir incluso cuándo detenerse (no se requiere darle el número máximo de generaciones). El único parámetro que requiere es el tamaño del archivo externo (aunque, evidentemente, hay un valor por omisión para este parámetro).

El μGA^2 ha sido utilizado para resolver problemas de reconfiguración considerando pérdidas de potencial e índices de confiabilidad para una red de distribución de voltaje medio [Mendoza, 2009].

El μGA^2



Initialize crossover operators
Initialize population memories
Adaptive Micro GA
Adaptive Micro GA
Adaptive Micro GA
Compare results and rank the subpopulations
Select crossover operators
External Memory
Select the population memories for the Adaptive micro–GAs
Convergence?
N
Y

Los algoritmos que el mundo olvidó

Se han propuesto diversos AEMOs que, por diversas razones, nunca alcanzaron popularidad. Algunos ejemplos son los siguientes:

- El Nash Genetic Algorithm [Sefrioui, 1996].
- La Maximin Fitness Function [Balling, 2000].
- El Incrementing Multi-Objective Evolutionary Algorithm (IMOEA) [Tan et al., 2001].
- El Constraint Method-Based Evolutionary Algorithm (CMEA) for Multiobjective Optimization [Ranjithan et al., 2001].
- El Orthogonal Multi-Objective Evolutionary Algorithm (OMOEA) [Zeng et al., 2004].
- y muchos más ...

MOEA/D

El *Multi-Objective Evolutionary Algorithm based on Decomposition* (MOEA/D) propuesto por Zhang y Li [2007] es uno de los AEMOs más competitivos de la actualidad. MOEA/D transforma un problema multi-objetivo en varios problemas mono-objetivo los cuales son resueltos simultáneamente. Cada subproblema se optimiza utilizando información de sus subproblemas vecinos, en contraste con otros métodos similares (p.ej., MOGLS [Ishibuchi & Murata, 1996]). Este AEMO está basado en una técnica de programación matemática llamada *Normal Boundary Intersection* (NBI) [Das, 1998].

Qingfu Zhang and Hui Li, "**MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition**", *IEEE Transactions on Evolutionary Computation*, Vol. 11, No. 6, pp. 712–731, December 2007.

Algoritmos Recientes

Selección Basada en Indicadores

Quizás la tendencia más importante en torno al diseño moderno de AEMOs sea el uso de una medida de desempeño en su mecanismo de selección.

ESP

La **Evolution Strategy with Probability Mutation** adopta una medida basada en el hipervolumen, la cual es independiente de la escala y no requiere de ningún parámetro, para truncar el contenido de un archivo externo [Huband et al., 2003].

Simon Huband, Phil Hingston, Lyndon White and Luigi Barone, "**An Evolution Strategy with Probabilistic Mutation for Multi-Objective Optimisation**", in *Proceedings of the 2003 Congress on Evolutionary Computation (CEC'2003)*, Vol. 3, pp. 2284–2291, IEEE Press, Canberra, Australia, December 2003.

IBEA

El **Indicator-Based Evolutionary Algorithm** en un entorno algorítmico que permite la incorporación de cualquier indicador de desempeño en el mecanismo de selección de un AEMO [Zitzler et al., 2004]. Dicho entorno se validó originalmente con el hipervolumen y el indicador ϵ binario.

Eckart Zitzler and Simon Künzli, "**Indicator-based Selection in Multiobjective Search**, in Xin Yao et al. (editors), *Parallel Problem Solving from Nature - PPSN VIII*, Springer-Verlag, Lecture Notes in Computer Science, Vol. 3242, pp. 832–842, Birmingham, UK, September 2004.

SMS-EMOA

Emmerich et al. [2005] propusieron un algoritmo basado en el NSGA-II el cual utiliza la técnica de archivado propuesta por Knowles, Corne y Fleischer al cual denominaron *S Metric Selection Evolutionary Multiobjective Algorithm*.

SMS-EMOA crea una población inicial y genera una sola solución por iteración (es decir, utiliza una selección de estado uniforme) utilizando los operadores de cruza y mutación del NSGA-II. Posteriormente, aplica ordenamiento no dominado. Cuando el último frente no dominado tiene más de una solución, SMS-EMOA utiliza el hipervolumen para decidir qué solución debe removerse.

Michael Emmerich, Nicola Beume and Boris Naujoks, "**An EMO Algorithm Using the Hypervolume Measure as Selection Criterion**, in Carlos A. Coello Coello et al. (editors), *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005*, pp. 62–76, Springer. Lecture Notes in Computer Science Vol. 3410, Guanajuato, México, March 2005.

SMS-EMOA

Beume et al. [2007] propusieron una nueva versión del SMS-EMOA en la que no se usa la contribución al hipervolumen cuando, en el ordenamiento no dominado, se obtiene más de un frente. En este caso, se utiliza el número de soluciones que dominan a cierto individuo (es decir, la solución que es dominada por la mayor cantidad de individuos es removida).

Los autores de este algoritmo indican que su motivación para utilizar el hipervolumen fue mejorar la distribución de soluciones a lo largo del frente de Pareto (en otras palabras, el hipervolumen se utiliza como un estimador de densidad).

Nicola Beume, Boris Naujoks and Michael Emmerich, "**SMS-EMOA: Multiobjective selection based on dominated hypervolume**", *European Journal of Operational Research*, Vol. 181, No. 3, pp. 1653–1669, 16 September, 2007.

Algoritmos Recientes

MO-CMA-ES

Ésta es una versión multi-objetivo de la **covariance matrix adaptation evolution strategy** (CMA-ES) propuesta por Igel et al. [2007].

Su mecanismo de selección se basa en un ordenamiento no dominado que adopta como criterio secundario de selección ya sea la distancia de agrupamiento o la contribución al hipervolumen (en el artículo se evalúan las dos versiones, pero la basada en el hipervolumen es la que presenta el mejor desempeño).

Este AEMO es invariante a la rotación, igual que el optimizador mono-objetivo en el cual se basa.

Christian Igel, Nikolaus Hansen and Stefan Roth, "**Covariance Matrix Adaptation for Multi-objective Optimization**", *Evolutionary Computation*, Vol. 15, No. 1, pp. 1–28, Spring 2007.

SPAM

El *Set Preference Algorithm for Multiobjective optimization* es una generalización de IBEA que permite adoptar cualquier relación de preferencia de conjuntos en su mecanismo de selección [Zitzler et al., 2008].

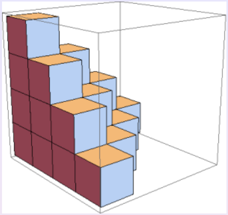
Eckart Zitzler, Lothar Thiele and Johannes Bader, "**SPAM: Set Preference Algorithm for Multiobjective Optimization**", in Günter Rudolph et al. (editors), *Parallel Problem Solving from Nature–PPSN X*, pp. 847–858, Springer, Lecture Notes in Computer Science Vol. 5199, Dortmund, Germany, September 2008.

HyPE

El *hypervolume estimation algorithm for multi-objective optimization*, fue propuesto por Bader y Zitzler [2011]. En este caso, los autores proponen un algoritmo de búsqueda rápida que usa simulaciones de Monte Carlo para aproximar las contribuciones al hipervolumen.

La idea fundamental es que conocer el valor exacto de la contribución al hipervolumen no es realmente tan importante, pues realmente lo que nos interesa es poder producir una jerarquización a partir de ella. Aunque esta propuesta es muy interesante, en la práctica su desempeño es bastante pobre con respecto al de los AEMOs que utilizan las contribuciones exactas al hipervolumen.

Johannes Bader and Eckart Zitzler, "**HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization**", *Evolutionary Computation*, Vol. 19, No. 1, pp. 45–76, Spring, 2011.



El Hipervolumen

El **hipervolumen** (conocido también como la métrica $\mathcal{S}$ o la medida de Lebesgue) de un conjunto de soluciones, mide el tamaño de la porción del espacio de las funciones objetivo que es dominado, colectivamente, por dichas soluciones.

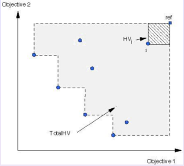
El hipervolumen es el único indicador de desempeño conocido a la fecha que es monotónico con respecto a la dominancia de Pareto. Esto garantiza que el verdadero frente de Pareto tenga el máximo valor posible del hipervolumen y cualquier otro conjunto que lo aproxime producirá un valor inferior para este indicador.

El Hipervolumen

Fleischer [2003] demostró que, dado un espacio de búsqueda finito y un punto de referencia, maximizar el hipervolumen es equivalente a obtener el conjunto de óptimos de Pareto. Por lo tanto, un conjunto acotado que contenga el valor máximo posible del hipervolumen para un cierto tamaño de población, consistirá únicamente de soluciones óptimas en el sentido de Pareto.

Este resultado ha sido validado experimentalmente [Knowles, 2003; Emmerich, 2005], y se ha podido observar que tales soluciones tienen también una buena distribución a lo largo del frente de Pareto.

M. Fleischer, "**The Measure of Pareto Optima. Applications to Multi-objective Metaheuristics**", in Carlos M. Fonseca et al. (editors), *Evolutionary Multi-Criterion Optimization. Second International Conference, EMO 2003*, pp. 519–533, Springer. Lecture Notes in Computer Science. Volume 2632, Faro, Portugal, April 2003.

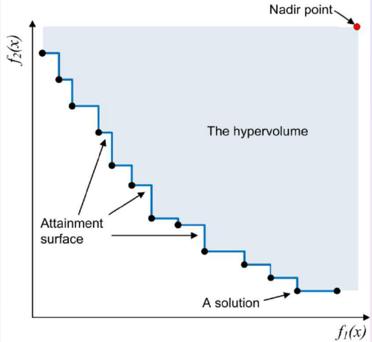


El Hipervolumen

El cálculo del hipervolumen depende del punto de referencia que se adopte y dicho punto puede ejercer una influencia importante en los resultados. Algunos investigadores han propuesto el uso de los peores valores de las funciones objetivo disponibles en la población actual, pero esto requiere que se escalen los objetivos.

Sin embargo, la principal desventaja de usar el hipervolumen es su elevado costo computacional. Los mejores algoritmos que se conocen actualmente para calcular el hipervolumen tienen una complejidad que es polinominal en el número de puntos, pero dicha complejidad crece exponencialmente con el número de objetivos.

Algoritmos Recientes



El Hipervolumen

El hecho de que no exista ningún algoritmo de complejidad polinomial para calcular el hipervolumen de forma exacta, originó la hipótesis de que dicho algoritmo pudiese no ser posible. Esto es algo interesante si consideramos que la cota inferior justa para la complejidad del cálculo del hipervolumen es $O(N \log N)$ [Beume, 2007].

Algoritmos Recientes

El Hipervolumen

Algunos resultados teóricos refuerzan esta hipótesis: Bringmann y Friedrich [2009] demostraron que el cálculo del hipervolumen es #P-Complete. Esto significa que no existe algoritmo de complejidad polinomial para calcularlo, pues de no ser así, esto implicaría que $NP = P$.

Sin embargo, el cálculo de este indicador ha dado pie a una cantidad significativa de trabajos de investigación. Ver por ejemplo:

- `http://ls11-www.cs.uni-dortmund.de/rudolph/hypervolume/start`
- `http://people.mpi-inf.mpg.de/~tfried/HYP/`
- `http://iridia.ulb.ac.be/~manuel/hypervolume`