Optimización Evolutiva Multi-objetivo

Dr. Carlos Segura González

Centro de Investigación de Matemáticas

December 16, 2019

Esquema General

Introducción y antecedentes

Conceptos

Paradigmas de algoritmos multi-objetivo

Problemas de prueba

Convergencia en algoritmos evolutivos

Algoritmo de diversidad basado en dominancia

Optimizadores estocásticos poblacionales

- Una meta-heurística es un proceso maestro que de forma iterativa guía y modifica un conjunto de operaciones heurísticas con el fin de producir soluciones de calidad de forma eficiente[1].
- Las meta-heurísticas poblacionales consideran múltiples soluciones en cada iteración.
- Algunas meta-heurísticas poblacionales más populares son las siguientes:
  - Algoritmos genéticos (GA).
  - Programación genética (GP).
  - Programación evolutiva (EP).
  - Evolución diferencial (DE).
  - Estrategia evolutiva (ES).
  - Algoritmos de estimación de distribución (EDA).
  - Optimización por enjambre de particulas (PSO).
  - Algoritmos de optimización por colonias de hormigas (ACO).

[1] Stefan Voß et al. *Meta-heuristics: Advances and trends in local search paradigms for optimization*. Springer Science & Business Media, 2012.

Optimizadores estocásticos poblacionales

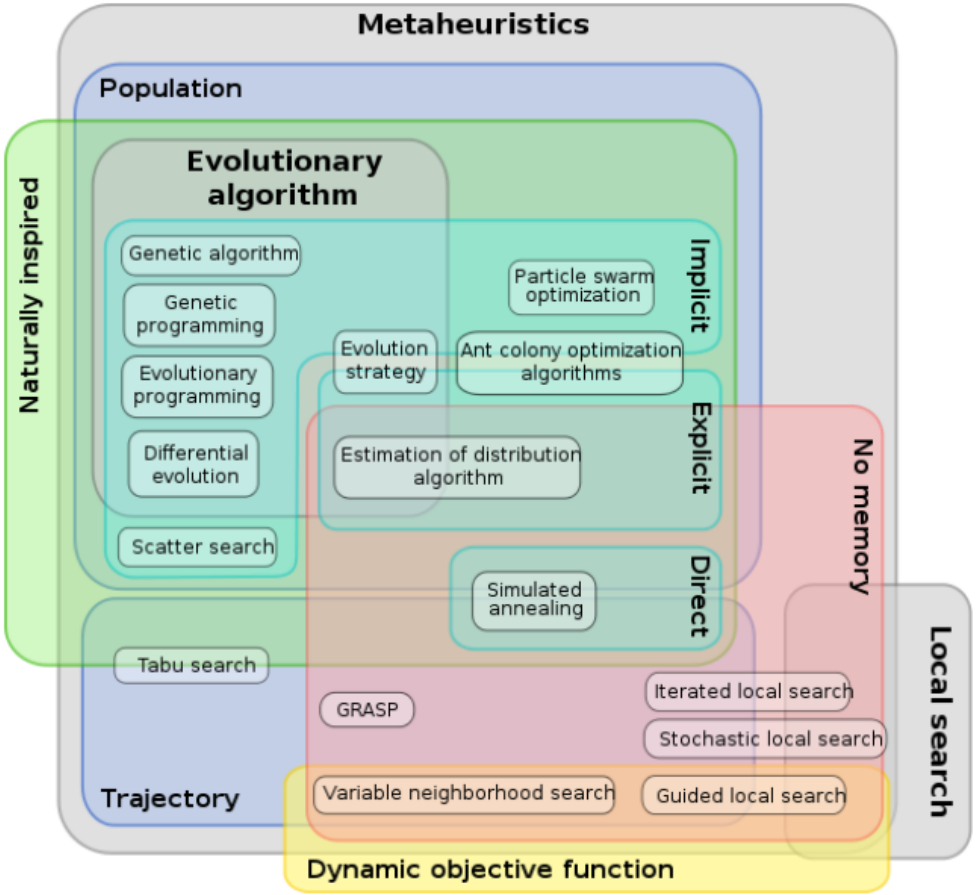


Figure 1: Clasificación de las meta-heurísticas más populares[2].

[2] Zahra Beheshti and Siti Mariyam Hj Shamsuddin. "A review of population-based meta-heuristic algorithms". In: *Int. J. Adv. Soft Comput. Appl* 5.1 (2013), pp. 1–35.

Problemas de las meta-heurísticas poblacionales

- La mayoría de las meta-heurísticas poblacionales tienen problemas de convergencia, y no es fácil controlar la velocidad en que convergen.
- Se dice que un algoritmo converge de forma prematura cuando mucho antes de alcanzar el criterio de paro, todas las soluciones están en una zona muy pequeña del espacio de búsqueda[3].

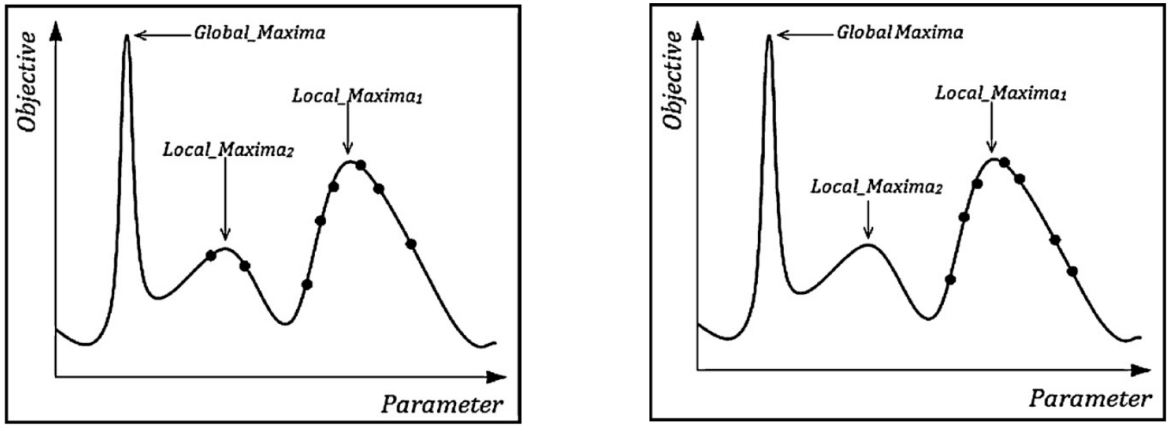


Figure 2: Escenarios de convergencia de una meta-heurística poblacional[4].

[3] Matej Črepinšek, Shih-Hsi Liu, and Marjan Mernik. "Exploration and Exploitation in Evolutionary Algorithms: A Survey". In: *ACM Computing Surveys* 45.3 (July 2013), 35:1–35:33.

[4] Mohammad Shafiul Alam et al. "Diversity Guided Evolutionary Programming: A novel approach for continuous optimization". In: *Applied soft computing* 12.6 (2012), pp. 1693–1707.

Algoritmos Evolutivos (EAs)

- Son un tipo de meta-heurísticas poblacionales.
- Son un enfoque ampliamente utilizado para resolver distintos tipos de problemas de optimización.
- Se han desarrollado diversas variantes que han sido aplicadas en múltiples campos, como en transporte, economía e ingeniería[5].
- Los EAs pueden ser aplicados en problemas de dominio continuo y/o de dominio discreto.
- Este tipo de algoritmos han sido exitosos principalmente en problemas del tipo NP-completo cuya solución exacta no es conocida.

[5] Dipankar Dasgupta and Zbigniew Michalewicz. *Evolutionary algorithms in engineering applications*. Springer Science & Business Media, 2013.

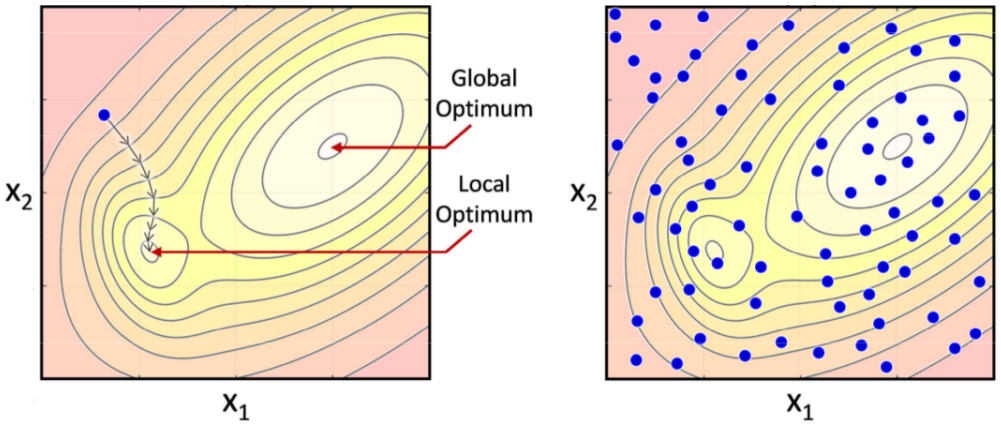
Esquemas de los algoritmos evolutivos

Algorithm 1 Evolutionary Algorithm

```
Initialization: Generate an initial population P_0 with N individuals.
Evaluation: Evaluate all individuals in the population.
Assign t = 0
while (not stopping criterion) do
    Mating selection: Fill the mating pool performing selection on P_t.
    Variation: Apply crossover and mutation operators to the mating pool to create
    an offspring population Q_t with N individuals.
    Evaluation: Evaluate all individuals in Q_t.
    Survivor selection: Generate P_{t+1} by applying the replacement operator using
    P_t and Q_t.
    t = t + 1
end while
```

Principios de diseño de un algoritmo poblacional

- Como parte del diseño de una meta-heurística poblacional, para tratar adecuadamente la convergencia debería existir un balanceo entre exploración e intensificación del espacio de búsqueda[6].
- La exploración del espacio de búsqueda evalúa regiones que no han sido muestreadas con el fin de detectar regiones promisorias.
- La intensificación realiza una búsqueda más profunda con el fin de encontrar soluciones de mayor calidad.



[6] Francisco Herrera and Manuel Lozano. "Adaptation of genetic algorithm parameters based on fuzzy logic controllers". In: *Genetic Algorithms and Soft Computing* 8 (1996), pp. 95–125.

¿Cómo tratar la convergencia?

- Una grado de exploración elevado puede causar una falta de intensificación en las regiones más prometedoras (**alta diversidad**).
- Por otra parte, una intensificación elevada provoca una pérdida de **diversidad** en la población.
- ¿Cómo fomentar este balanceo en una meta-heurística poblacional? (línea de investigación activa).

¿Cómo tratar la convergencia?

- Algunas técnicas controlan de manera directa o indirecta la cantidad de diversidad mantenida por el algoritmo[7].
- Éstas técnicas son identificadas de forma general como enfoques uni-proceso y multi-proceso.
- Una taxonomía ampliamente utilizada de las estrategias para promover la diveridad son:
  - Enfoques basados en la selección[8].
  - Enfoques basados en población[9].
  - Enfoques basados en la cruza y/o la mutación[10].

[7] Črepinšek, Liu, and Mernik, "Exploration and Exploitation in Evolutionary Algorithms: A Survey".

[8] Agoston E Eiben, James E Smith, et al. *Introduction to evolutionary computing*. Vol. 53. Springer, 2003.

[9] Enrique Alba. *Parallel metaheuristics: a new class of algorithms*. Vol. 47. John Wiley & Sons, 2005.

[10] Wei-Jie Yu et al. "Differential evolution with two-level parameter adaptation". In: *IEEE Transactions on Cybernetics* 44.7 (2014), pp. 1080–1099.

Esquemas de reemplazamiento basados en diversidad

- Los esquemas de reemplazamiento se utilizan para diversificar a los individuos sobrevivientes, de forma que los operadores de reproducción puedan generar nuevas soluciones en diferentes regiones[11].
- Existen varias estrategias de reemplazo para forzar la diversidad en la población, algúnas populares son:
  - *Restricted Tournament Selection* - RTS (1995) por Harik.
  - *Contribution of Diversity and Replacement of the Worst* - CD/RW (2008) por Lozano, Herrera, and Cano.
  - *Hybrid Genetic Search with Adaptive Diversity Control* - HGSADC (2013) por Vidal et al.
  - *Multi-dynamic* (2016) por Segura et al.
  - *Best-Non Penalized* (2018) por Romero Ruiz and Segura.

[11] Agoston E Eiben and Cornelis A Schippers. "On evolutionary exploration and exploitation". In: *Fundamenta Informaticae* 35.1-4 (1998), pp. 35–50.

Definición de un problema mono-objetivo

- Un problema de optimización mono-objetivo puede ser definido de la siguiente forma[12]:

$$
\begin{aligned}
Minimizar/Maximizar \quad & f(x) \\
sujeto \quad a \quad & g_j(x) \geq 0 \quad j = 1, 2, ...J \\
& h_k(x) = 0, \quad k = 1, 2, ..., K \\
& x_i^{(L)} \leq x_i \leq x_i^{(U)} \quad i = 1, 2, ..., n
\end{aligned}
\tag{1}
$$

donde $x_i^{(L)}, x_i^{(U)}$ son el límite inferior y superior de la i-ésima variable, $g_j(x)$ y $h_k(x)$ son restricciones de desigualdad e igualdad respectivamente.

[12] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

Definición de un problema multi-objetivo

- Un problema de optimización multi-objetivo tiene un número de funciones objetivo en conflicto las cuales se deben minimizar o maximizar.
- Similarmente a un problema de optimización mono-objetivo, un problema multi-objetivo tiene un número de restricciones que cualquier solución factible debería satisfacer.

La forma general de un problema multi-objetivo se puede definir de la siguiente forma[13]:

$$\begin{aligned} Minimizar/Maximizar \quad & f_m(x), \quad m = 1, 2, ..., M \\ sujeto \quad a \quad & g_j(x) \geq 0 \quad j = 1, 2, ...J \\ & h_k(x) = 0, \quad k = 1, 2, ..., K \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)} \quad i = 1, 2, ..., n \end{aligned} \tag{2}$$

[13] Kalyanmoy Deb and Deb Kalyanmoy. *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, NY, USA: John Wiley & Sons, Inc., 2001. ISBN: 047187339X.

Definición de dominancia

- A diferencia del caso mono-objetivo, en el ámbito multi-objetivo no es posible comparar de forma directa dos soluciones $x, y \in \Omega$.
- Una alternativa para comparar soluciones es considerando la relación de dominancia.

Definition 2.1: Dominancia

Se dice que una solución x domina a otra solución y ($x \prec y$), si se cumplen las siguientes dos condiciones:

- La solución x no es peor que la solución y en ninguno de los objetivos.
  $\forall m \in \{1, 2, ..., M\} : f_m(x) \leq f_m(y)$
- La solución x es estrictamente mejor que la solución y en al menos un objetivo.
  $\exists m \in \{1, 2, ..., M\} : f_m(x) < f_m(y)$

Dominancia

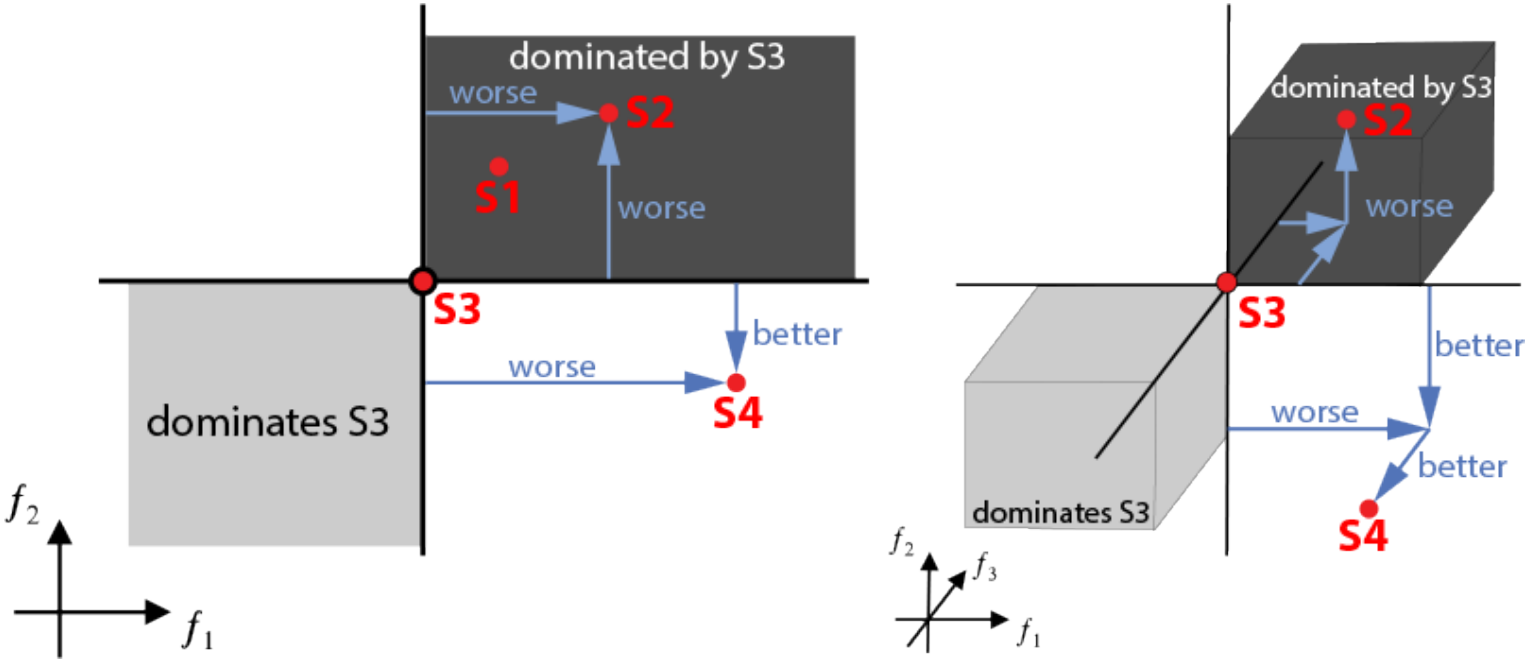


Figure 3: Esquema de dominancia en dos y tres objetivos respectivamente[14].

[14] *Generative Components Community Wiki*. https://communities.bentley.com/products/products_generativecomponents/w/generative_components_community_wiki/37421/b---non-dominated-solutions-and-the-pareto-frontier. Accessed: 2019-12-13.

Conceptos multi-objetivo

Definition 2.2: Conjunto no dominado

Es la colección de todas la soluciones incomparables y no dominadas relacionadas al conjunto $S \subseteq \Omega$.

$$NDS(S) = \{s \in S : \nexists s\prime \in S, s\prime \prec s\}$$

Definition 2.3: Conjunto Óptimo de Pareto

Es un conjunto de vectores de decisión óptimos no dominados en Ω.

$$POS = \{x^* \in \Omega : \nexists x \in \Omega \subset \Re^n, x \prec x^*\}$$

Definition 2.4: Frente de Pareto

La imágen del conjunto óptimo de Pareto es el frente de Pareto óptimo.

$$POF = \{f(x^*) \in Z \subset \Re^m : x^* \in POS\}$$

Conceptos multi-objetivo

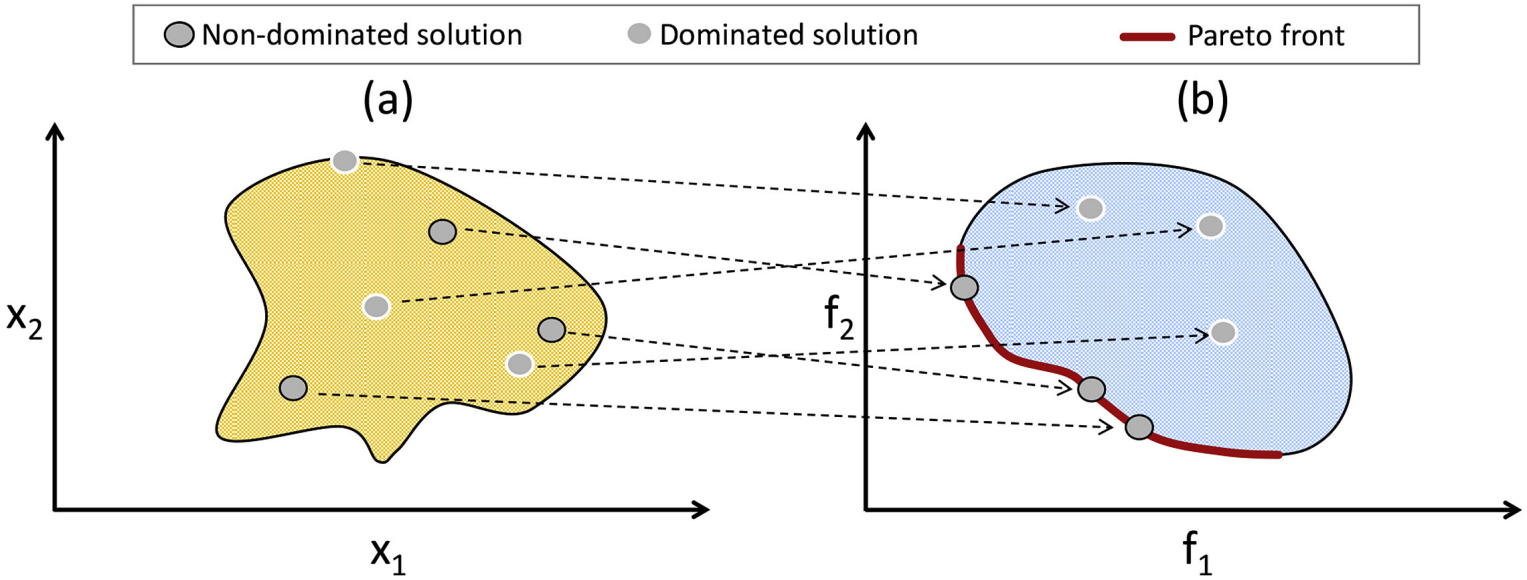


Figure 4: Mapeo del espacio de las variables (a) en el espacio objetivo (b)[15].

[15] Alam et al., "Diversity Guided Evolutionary Programming: A novel approach for continuous optimization".

Paradigmas de los algoritmos multi-objetivo

Una clasificación muy aceptada de los MOEAs es[16]:

- **Basados en dominancia**.
  - Utilizan la relación de dominancia de Pareto.
- **Basados en descomposición**.
  - Transforman un problema multi-objetivo en un conjunto de problemas de optimización mono-objetivo que son resueltos simultáneamente.
- **Basados en indicadores**.
  - Combinan el grado de convergencia y/o la diversidad en el espacio objetivo con una métrica.

Actualmente no existe un algoritmo el cual sea considerado superior, por lo tanto usualmente se selecciona al menos un algoritmo de cada tipo como parte del estado del arte.

[16] Anupam Trivedi et al. "A survey of multiobjective evolutionary algorithms based on decomposition". In: *IEEE Transactions on Evolutionary Computation* 21.3 (2016), pp. 440–462.

Algoritmos basados en dominancia

Algunos de los clásicos algoritmos multi-objetivo basados en dominancia son:

- 1993, *Multi-objective Genetic Algorithm* (MOGA), Fonseca and Fleming.
- 1994, *Niche Pareto Genetic Algorithm* (NPGA), Horn, Nafpliotis, and Goldberg.
- 2001, *Strength Pareto Evolutionary Algorithm* (SPEA2), Zitzler, Laumanns, and Thiele.
- 2002, ***Non-dominated Sorting Genetic Algorithm II*** **(NSGA-II)**, Deb et al.
- 2005, *Generalized Differential Evolution* (GDE3), Kukkonen and Lampinen.

Non-dominated Sorting Genetic Algorithm II (NSGA-II)

- Es uno de los primeros algoritmos multi-objetivo que introducen elitismo.
- Realiza una clasificación por frentes de baja complejidad $O(MN^2)$, estrategia ampliamente utilizada en multi-objetivo conocida como *fast-non-dominated-sortig* (convergencia al frente de Pareto).
- Considera un operador de amontonamiento o *crowding* (diversidad en el frente de Pareto).

$$d_i = \sum_{j=1}^{m} \frac{f_j^{i+1} - f_j^{i-1}}{f_j^{max} - f_j^{min}}$$

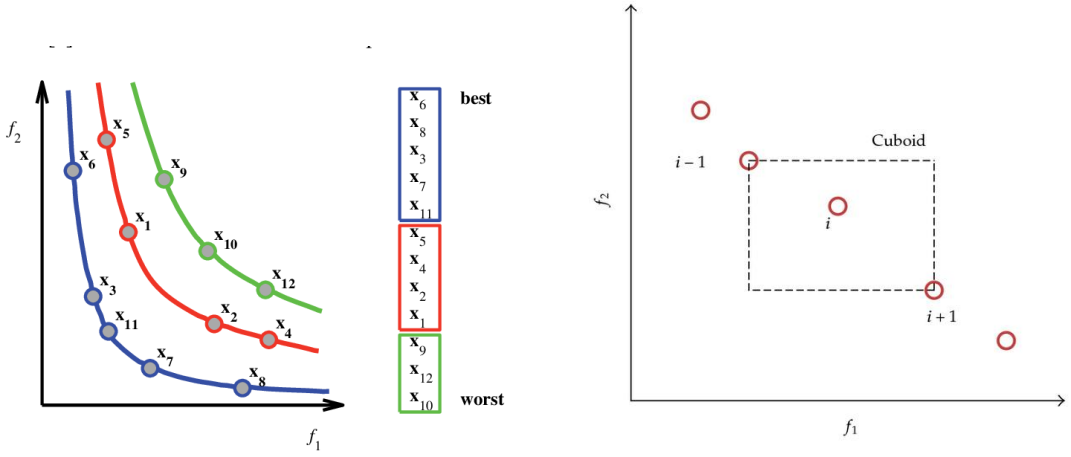


Figure 5: Clasificación por frentes (izquierda) y distancia de crowding (derecha)[17].

[17] Petr Kadlec and Zbyněk Raida. "Multi-objective self-organizing migrating algorithm". In: *Self-Organizing Migrating Algorithm*. Springer, 2016, pp. 83–103.

Non-dominated Sorting Genetic Algorithm II (NSGA-II)

- Es uno de los primeros algoritmos multi-objetivo que introducen elitismo.
- Realiza una clasificación por frentes de baja complejidad $O(MN^2)$, estrategia ampliamente utilizada en multi-objetivo conocida como *fast-non-dominated-sortig*.
- Considera un operador de amontonamiento o *crowding*.

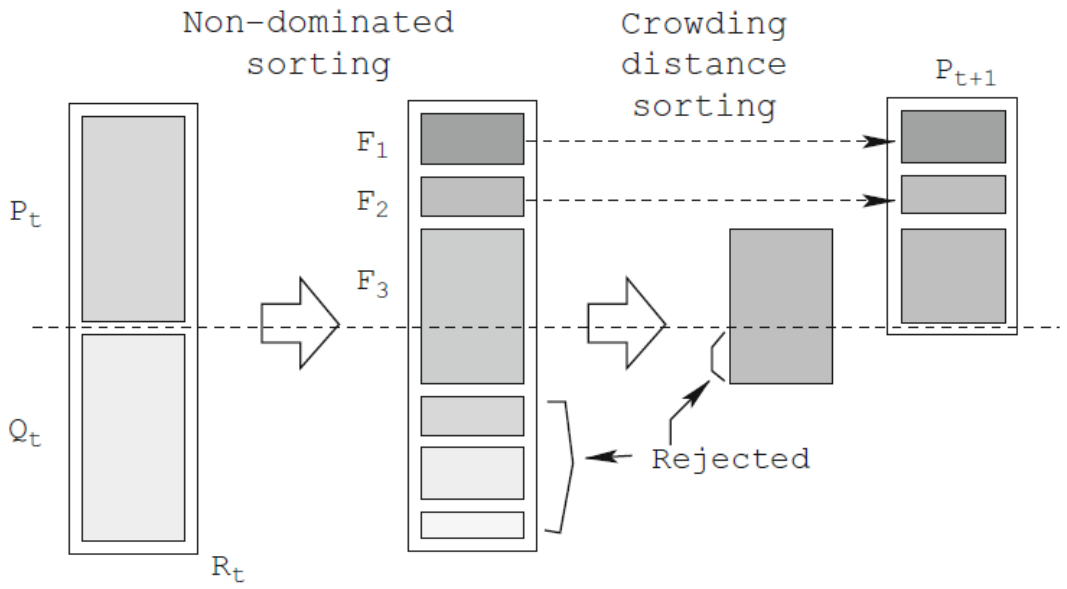


Figure 6: Procedimiento principal del NSGA-II[18].

[18] K. Deb et al. "A fast and elitist multiobjective genetic algorithm: NSGA-II". In: *IEEE Transactions on Evolutionary Computation* 6.2 (Apr. 2002), pp. 182–197. ISSN: 1089-778X.

Algoritmos basado en descomposición

Algunos de los clásicos algoritmos multi-objetivo basados en descomposición son:

- 1998, *Multi-objective Genetic Local Search Algorithm* (MOGLS), Ishibuchi and Murata,
- 2002, *Cellular Multi-objective Genetic Algorithm* (C-MOGA), Murata and Gen,
- 2007, ***Multi-objective Evolutionary Algorithm based in Decomposition* (MOEA/D)**, Zhang and Li.
- 2009, *Multi-objective Evolutionary Algorithm based in Decomposition with DE*, (MOEA/D -CEC 2009), Li and Zhang.
- 2011, *MOEA/D - Adaptive Mating Selection Mechanism* (MOEA/D - AMS) Chiang and Lai.
- 2015, *MOEA/D - Generalized Resource Allocation* (MOEA/D-GRA) Zhou and Zhang.

Multi-objective Evolutionary Algorithm based in Decomposition (MOEA/D)

La versión original del MOEA/D[19] tiene las siguientes particularidades:

- Cada subproblema está definido por un vector de pesos.
- Define vecindades de cada subproblema en relación a su ubicación en el espacio objetivo.
- Sólo se hacen emparejamientos y reemplazamientos en los vecindarios.
- Considera el operador de cruza SBX y el operador de mutación polinomial.

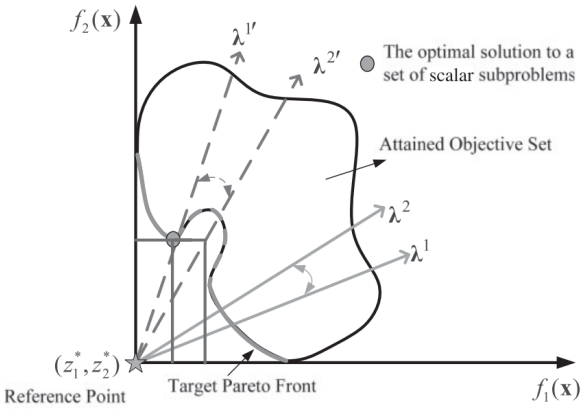


Figure 7: Vectores de pesos y su punto óptimo con el enfoque de Tchebysheff[20]

[19] Hisao Ishibuchi and Tadahiko Murata. "A multi-objective genetic local search algorithm and its application to flowshop scheduling". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 28.3 (1998), pp. 392–403.

[20] Z. Wang et al. "Adaptive Replacement Strategies for MOEA/D". In: *IEEE Transactions on Cybernetics* 46.2 (Feb. 2016), pp. 474–486. ISSN: 2168-2267.

Multi-objective Evolutionary Algorithm based in Decomposition (MOEA/D)

Una versión bastante popular y que tomó el primer lugar en la competancia del CEC 2009 (*Congress on Evolutionary Computation*) es el MOEA/D-DE que además considera las siguientes características[21]:

- Reestricciones de emparejamiento y de reemplazamiento.
- Para mantener diversidad en el espacio de las variables se considera emparejamiento en toda la población con probabilidad $1 - \delta$ y en el vecindario con probabilidad δ.
- Implementa un mecanismo que administra las evaluaciones de cada subproblema *Dynamical Resource Allocation*.
- Utiliza operadores de evolución diferencial y el operador de mutación polinomial.

[21] Hui Li and Qingfu Zhang. "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II". In: *IEEE transactions on evolutionary computation* 13.2 (2009), pp. 284–302.

Funciones de utilidad

Existen una gran variedad de funciones de utilidad para converir un problema multi-objetivo en un conjunto de problemas de un sólo objetivo, algunos de los más utilizados son[22]:

- Suma ponderada (*Weighted Sum*)

$$\min g^{ws}(x|\lambda, z^*) = \sum_{i=1}^{m} \lambda_i f_i(x)$$

- Tchebysheff

$$\min g^{te}(x|\lambda, z^*) = \max_{1 \leq i \leq m} \{\lambda_i |f_i(x) - z_i^*|\}$$

- Función de escalarización alcanzado (*Achieved Scalarized Funcition* - ASF)

$$\min g^{te}(x|\lambda, z^*) = \max_{1 \leq i \leq m} \{\frac{|f_i(x) - z_i^*|}{\lambda_i}\}$$

- Intersección del límite de penalización (*Penalty Boundary Intersection* - PBI)

$$\min g^{bi}(x|\lambda, z^*) = d_1 + \theta d_2$$

$$sujeto \quad a \quad d_1 = \frac{||(F(x) - z^*)^T \lambda||}{||\lambda||}, \quad d_2 = ||F(x) - (z^* + d_1 \lambda)||$$

[22] Raquel Hernández Gómez. "Parallel Hyper-Heuristics for Multi-Objective Optimization (PhD thesis)". In: (2018).

Funciones de utilidad

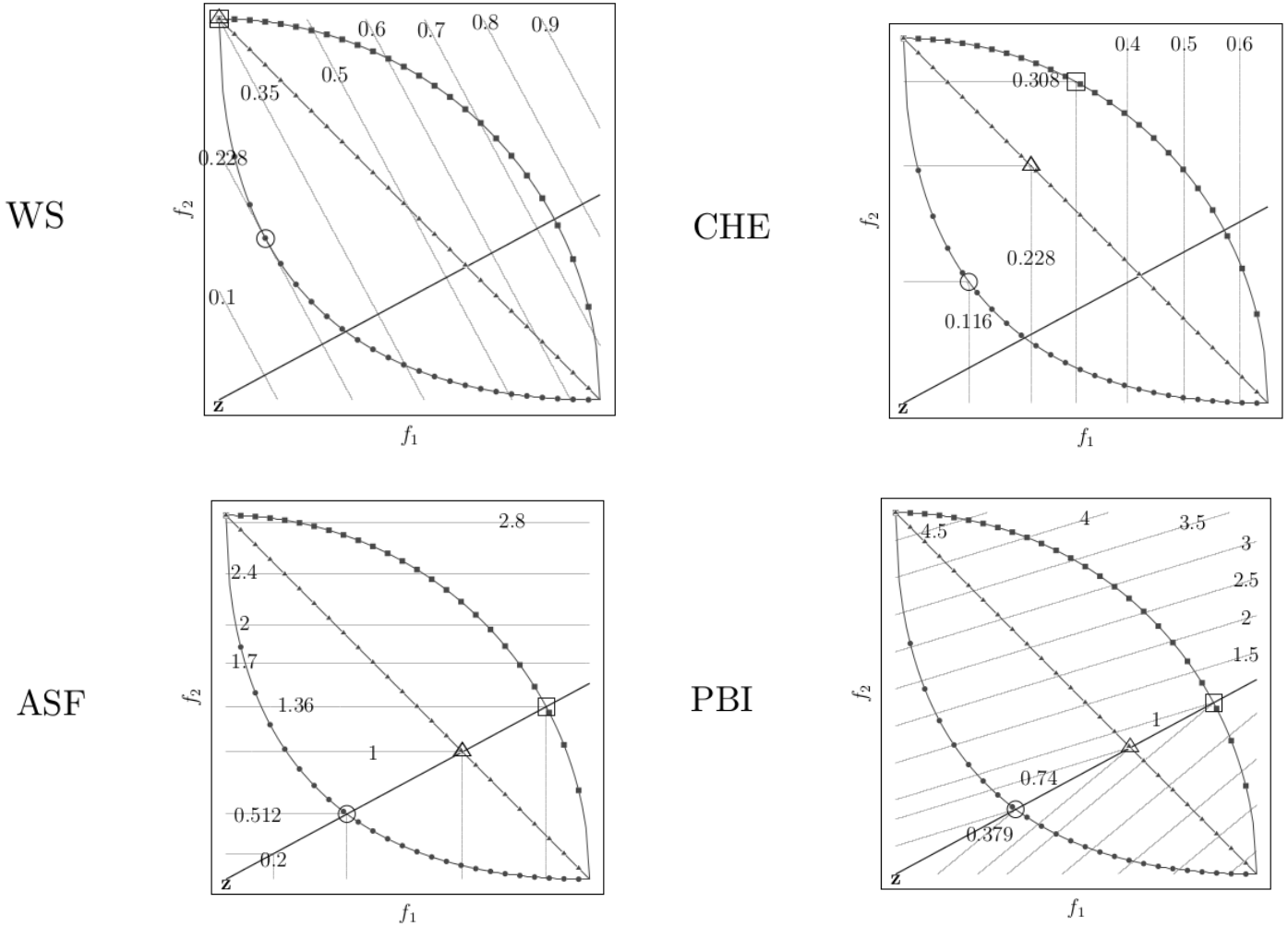


Figure 8: Mapa de contorno de las funciones de utilidad[23].

[23] Gómez, "Parallel Hyper-Heuristics for Multi-Objective Optimization (PhD thesis)".

Algoritmos basados en indicadores

Algunos de los algoritmos clásicos multi-objetivo basados en indicadores son:

- 2004, *Indicator Based-Selection Evolutionary Algorithm* (IBEA), Zitzler and Künzli.
- 2007, ***S-Metric Selection Evolutionary Multi-objective Optimization Algorithm* (SMS-EMOA)**, Beume, Naujoks, and Emmerich.
- 2015, *Fast Hypervolume Multi-objective Evolutionary Algorithm* FV-MOEA, Jiang et al.
- 2015, *Many-Objective Metaheuristic Based on* $R2$ *Indicador* MOMBI-II, Hernández Gómez and Coello Coello.

S-Metric Selection Evolutionary Multi-objective Optimization Algorithm (SMS-EMOA)

- Implementa un operador de selección especial que combina la métrica de hipervolumen con el concepto de dominancia de Pareto.
- Considera un esquema de selección de estado estable (*Steady State*).
- Una propuesta incorpora la clasificación por frentes y sólo considera el HV cuando la población está compuesta por individuos no dominados.

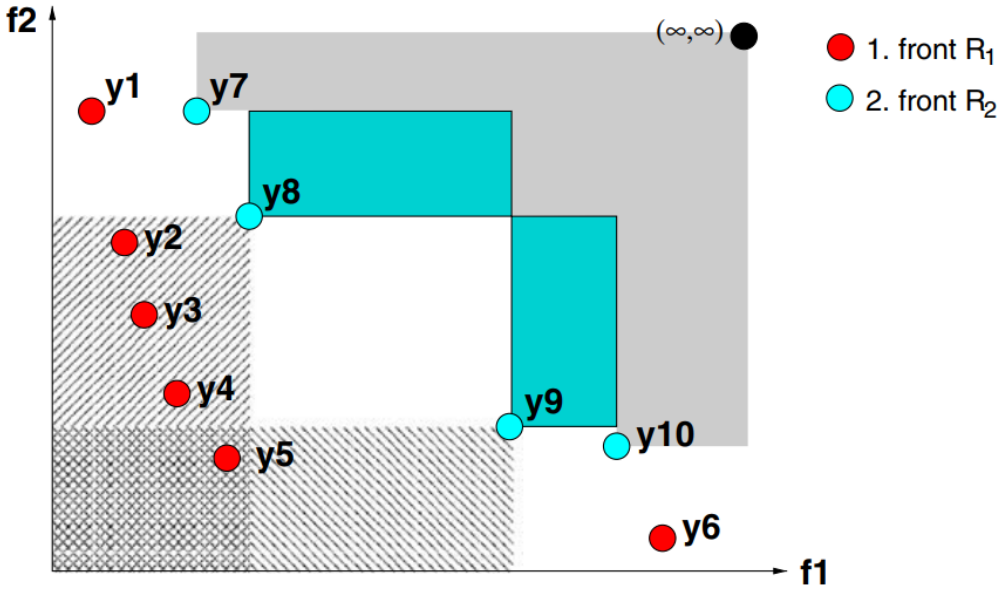


Figure 9: Contribución del HV de cada solución[24].

[24] Nicola Beume, Boris Naujoks, and Michael Emmerich. "SMS-EMOA: Multiobjective selection based on dominated hypervolume ". In: *European Journal of Operational Research* 181.3 (2007), pp. 1653–1669. ISSN: 0377-2217.

Análisis de los resultados

Existen distintas herramientas para verificar la calidad de las soluciones de un optimizador, se enlistan las siguientes:

- Métricas de desempeño: son métricas para evaluar la calidad de un conjunto de soluciones no dominadas.
  - La Distancia Generacional Invertida (IGD).
  - La Distancia Generacional Invertida Modificada (IGD+).
  - El hipervolumen (HV).
- Análisis estadísticos:
  - Pruebas estadísticas.
  - Tabla general estadística.
- Superficies de cubrimiento

Poner diagrama de las pruebas estadísticas.

Métricas de desempeño

- En optimización multi-objetivo se han propuesto varios indicadores o métricas para evaluar la calidad de un conjunto de soluciones.
- En general los indicadores son métricas que determinan el grado de convergencia al frente de Pareto y/o diversidad de las soluciones.
- El indicador más frecuente es el hipervolumen debido a que es el único del tipo *Pareto compliant*[25].
- Ser *Pareto compliant*[26] quiere decir que el indicador no contradice el orden inducido por la relación de dominancia.

$$A \succeq B \wedge B \not\succeq A \rightarrow I(A) > I(B)$$

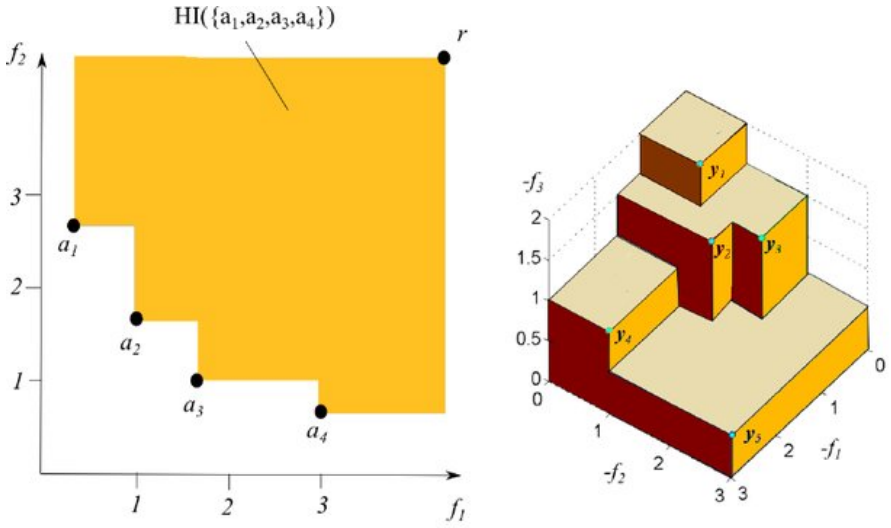
[25] Eckart Zitzler and Lothar Thiele. "Multiobjective optimization using evolutionary algorithms—a comparative case study". In: *International Conference on Parallel Problem Solving from Nature*. Springer. 1998, pp. 292–301.

[26] J. Knowles and D. Corne. "On metrics for comparing nondominated sets". In: *Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on*. Vol. 1. May 2002, pp. 711–716. DOI: 10.1109/CEC.2002.1007013.

Hipervolumen

- Es un indicador unario conocido como “Pareto Compliant”[27].
- Este indicador determina la porción del espacio objetivo que es dominado por las soluciones de un conjunto A, el cual es limitado por un punto de referencia $r \in \Re^n$ indicado en la ecuación (3), donde Λ denota la medición de Lebesgue y r debería ser dominado por todos los elementos en A.

$$HV(A; r) = \Lambda \left(\bigcup_{a \in A} \{x | a \prec x \prec r\} \right) \tag{3}$$



[27] Eckart Zitzler. *Evolutionary algorithms for multiobjective optimization: Methods and applications*. Vol. 63. Citeseer, 1999.

Distancia generacional

- En el 1998 Veldhuizen and Lamont propuso la distancia generacional (GD), el cual es un indicador unario del tipo Pareto *non-compliant*.
- Requiere de un conjunto de referencia, el cual es una discretización del frente de Pareto.
- El GD calcula la distancia promedio de cada solución obtenido al punto de referencia más cercano.

$$GD(A;R) = \frac{1}{|A|}\left(\sum_{a\in A} d(a,R)^p\right)^{1/p}, \quad d(a,R) = \min_{r\in R}\sqrt{\sum_{i=1}^{m}(a_i - r_i)^2}$$

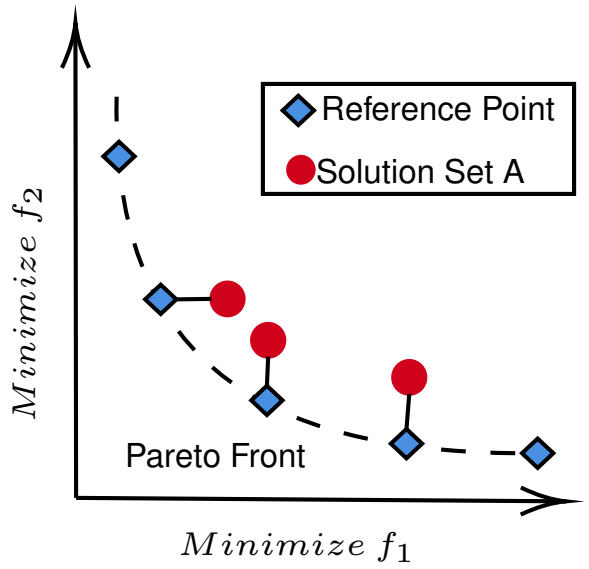


Figure 10: Ejemplo de la distancia generacional y un conjunto de referencia.

Desventajas de la distancia generacional

- El conjunto aproximado puede no converger a todo el frente de Pareto y aún así se tendría una medición “buena” (ver figura 11).
- Existe un sesgo al incrementar el número de puntos aproximados, es decir genera valores menores de GD al aumentar la cantidad de soluciones pero sin mejorar la calidad de las aproximaciones[28].
- ¿Cómo se puede mejorar?

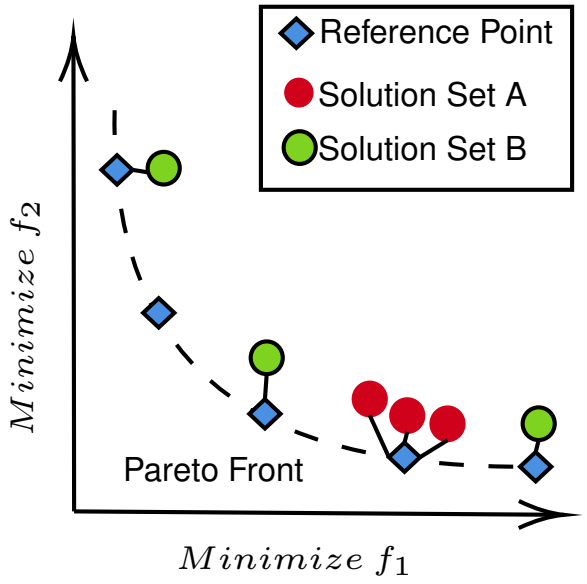


Figure 11: Ejemplo de una desventaja de la distancia generacional.

[28] Oliver Schutze et al. “Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization”. In: *IEEE Transactions on Evolutionary Computation* 16.4 (2012), pp. 504–522.

Distancia generacional invertida

- A diferencia de GD, la Distancia Generacional Invertida (IGD) es la distancia promedio de cada punto de referencia a su solución más cercana del conjunto aproximado.

$$IGD(A;R) = \frac{1}{|R|}\left(\sum_{r\in R} d(r,A)^p\right)^{1/p}, \quad d(r,A) = \min_{a\in A}\sqrt{\sum_{i=1}^{m}(r_i - a_i)^2}$$

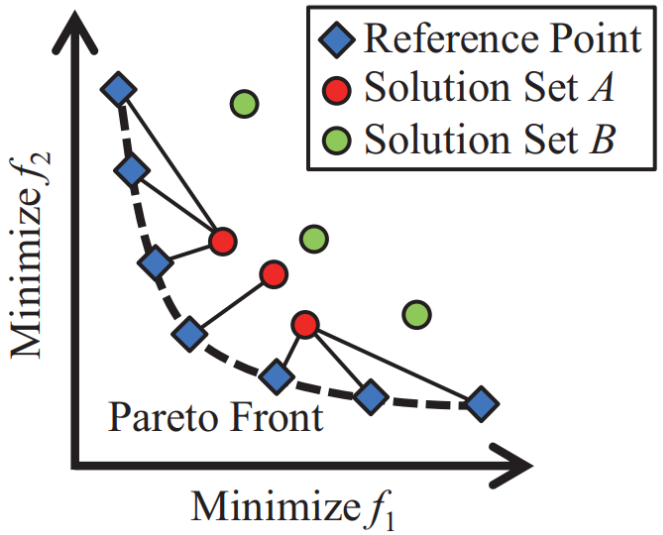


Figure 12: Ejemplo de la distancia generacional invertida.[29].

[29] Hisao Ishibuchi, Hiroyuki Masuda, and Yusuke Nojima. "Sensitivity of performance evaluation results by inverted generational distance to reference points". In: *2016 IEEE Congress on Evolutionary Computation (CEC)*. IEEE. 2016, pp. 1107–1114.

Desventajas de la distancia generacional invertida

- La métrica sigue siendo Pareto *non-compliant*, por lo tanto existen situaciones en que un conjunto de soluciones alejado del frente de mejores valores que un conjunto más cercano del frente.
- Se puede observar en la figura 13 que a pesar de que la solución $\boldsymbol{b}$ es dominada por la solución $\boldsymbol{a}$, la distancia de un punto de referencia $\boldsymbol{z}$ a $\boldsymbol{b}$ puede ser menor que la distancia de $\boldsymbol{z}$ a $\boldsymbol{a}$.
- ¿Cómo se puede mejorar?

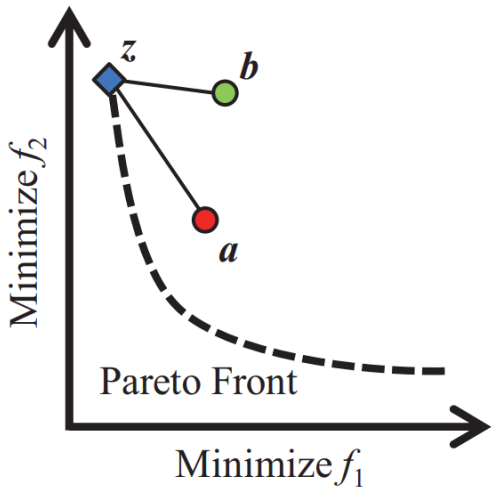


Figure 13: Desventaja de la métrica IGD.[30].

[30] Ishibuchi, Masuda, and Nojima, "Sensitivity of performance evaluation results by inverted generational distance to reference points".

Distancias

- Propuesta por Ishibuchi et al. en el 2015, es considerada como débilmente *Pareto Compliant*.
- Si una solución es dominada por un punto de referencia se utiliza la distancia euclideana.
- Si la solución y el punto de referencia son no dominados entre sí, entonce se calcula la distancia mínima desde el punto de referencia a la región no dominada.

$$IGD(A;R) = \frac{1}{|R|}\left(\sum_{r\in R} d^{+}(r,A)^{p}\right)^{1/p}, \quad d^{+} = \min_{a\in A}\sqrt{\sum_{i=1}^{m}(max\{a_i - z_i, 0\})^2}$$

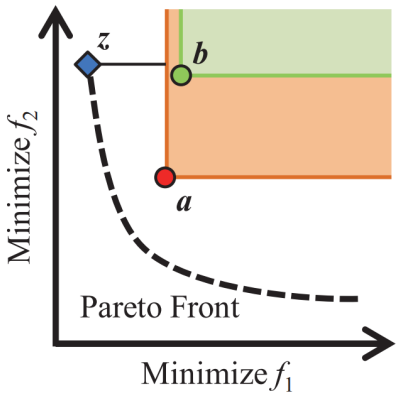


Figure 14: Distancia generacional invertida modificada (IGD+).[31].

[31] Ishibuchi, Masuda, and Nojima, "Sensitivity of performance evaluation results by inverted generational distance to reference points".

Distancia generacional invertida modificada

- Propuesta por Ishibuchi et al. en el 2015, es considerada como débilmente *Pareto Compliant*.
- Si una solución es dominada por un punto de referencia se utiliza la distancia euclideana, pero si la solución y el punto de referencia son no dominados entre sí, entonce se calcula la distancia mínima desde el punto de referencia a la región no dominada.

$$IGD(A;R) = \frac{1}{|R|}\left(\sum_{r\in R} d(r,A)^p\right)^{1/p} \qquad IGD^+(A;R) = \frac{1}{|R|}\left(\sum_{r\in R} d^+(r,A)^p\right)^{1/p}$$

$$d = \min_{a\in A}\sqrt{\sum_{i=1}^{m}(a_i - r_i)^2} \qquad d^+ = \min_{a\in A}\sqrt{\sum_{i=1}^{m}(max\{a_i - z_i, 0\})^2}$$

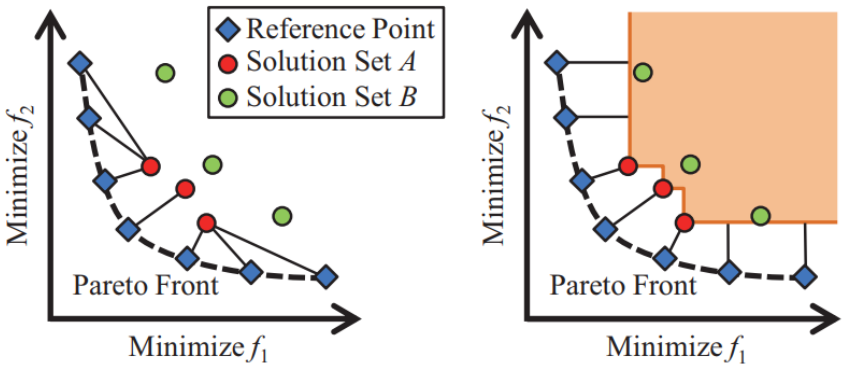


Figure 15: Ejemplo de las métricas IGD (izquierda) e IGD+ (derecha).[32].

[32] Ishibuchi, Masuda, and Nojima, "Sensitivity of performance evaluation results by inverted generational distance to reference points".

Problemas de prueba en el ámbito multi-objetivo

- ZDT, Zitzler, Deb, and Thiele, 2000.
- DTLZ, Deb et al., (versión 1 2002, versión 2 2005).
- WFG, Huband et al., 2005.
- UF, Zhang et al., 2008.

Problemas de prueba DTLZ

- Estos problemas son escalables en los objetivos y las variables.
- Existen dos versiones diseñadas por los mismos autores, en el 2002 y en el 2005.
- Los óptimos globales están ubicados en el centro del espacio de búsqueda o en los límites.
- Para todos los problemas, el óptimo global tiene los mismos valores de parámetros para distintas variables.

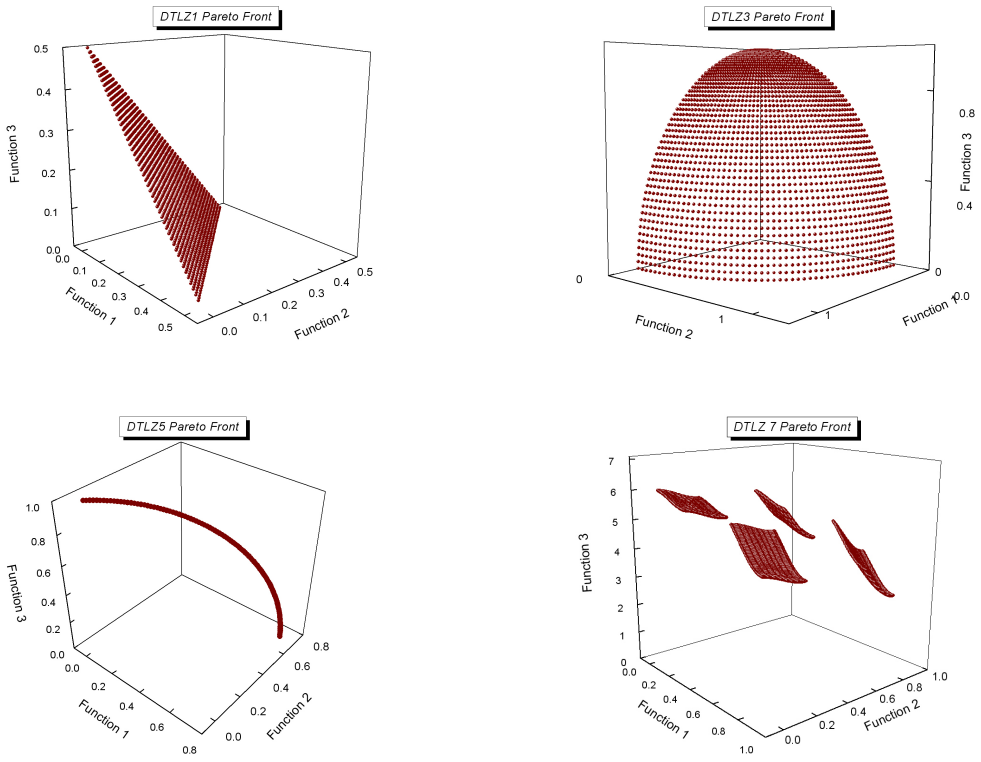


Figure 16: Frente de Pareto discretizado.

Problemas de prueba *Walking Fish Group* WFG

- Estos problemas son escalables en los objetivos y las variables.
- El generador de problemas de prueba WFG fue propuesto por Huband et al. en el 2005.
- A pesar de que es un generador de problemas, popularmente se utilizan nueve problemas de prueba.
- Los problemas WFG dividen el espaco de las variables de decisión en dos subconjuntos: los parámetros de distancia y los parámetros de posición.

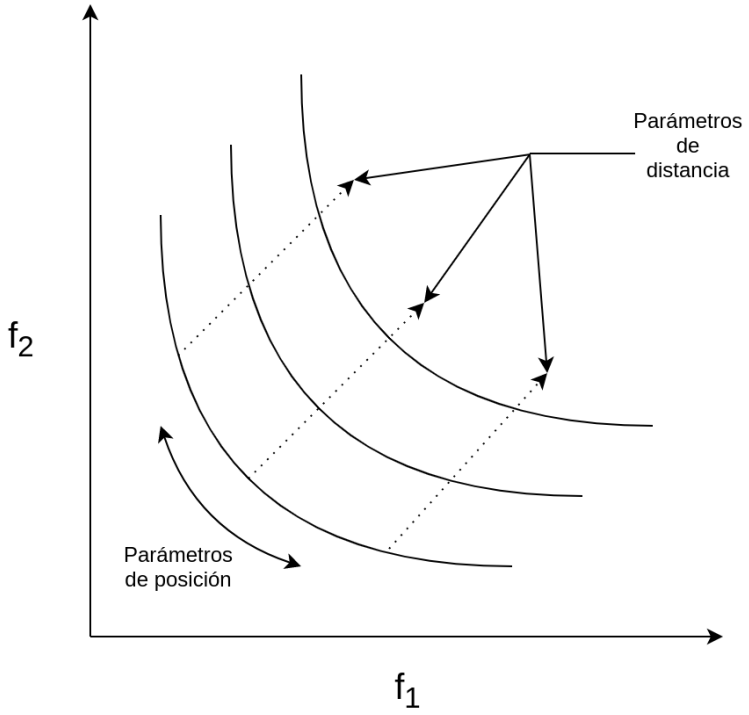


Figure 17: Parámetros de posición y distancia WFG.

Unconstrained Functions UF

- Estos son únicamente de dos objetivos (UF1-UF7) y tres objetivos (UF8-UF10), además son escalables en las variables.
- Este conjunto de prueba fue propuesto en el Congreso de Cómputo Evolutivo (CEC 2009) en el 2008 por Zhang et al.
- Este tipo de problemas aún se consideran un reto para los optimizadores multi-objetivo.

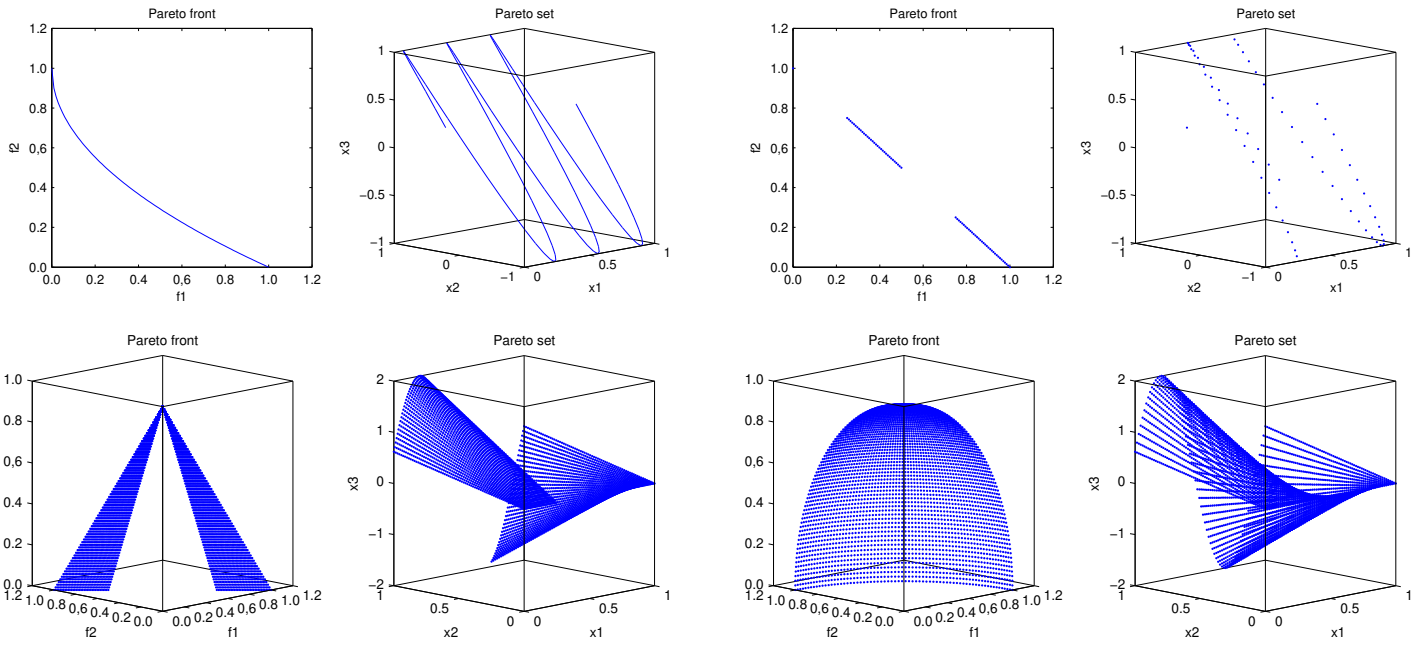


Figure 18: Frente de Pareto y conjunto de Pareto.

Superficie de alcance

- Una superficie de alcance o superficie de cubrimiento es una descripción de la distribución alcanzada en base a un conjunto de soluciones.
- Es un gráfico que muestra un resumen de una cantidad de solucione en varias medidas descriptivas considerando la relación de dominancia.

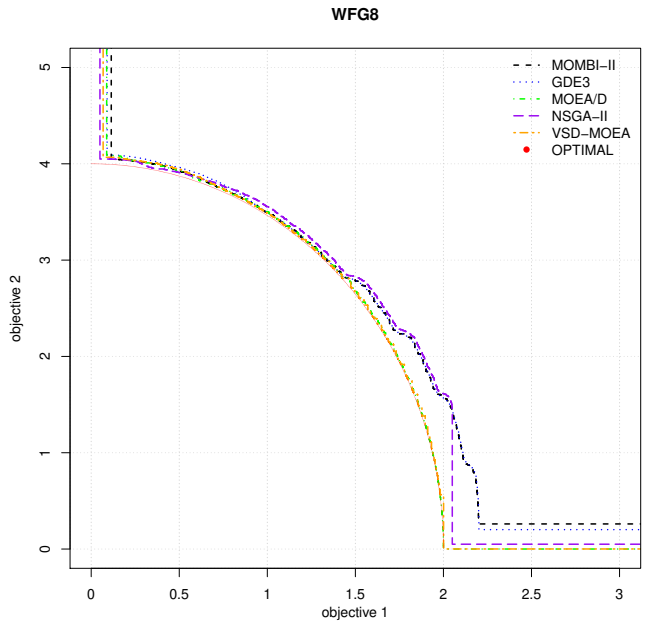


Figure 19: Superficie de alcance de cinco algoritmos considerando 35 ejecuciones y el problema WFG8.

Diferencia de la superficie de alcance

- Proporciona una diferencia estadística entre los resultados generados por dos algoritmos.

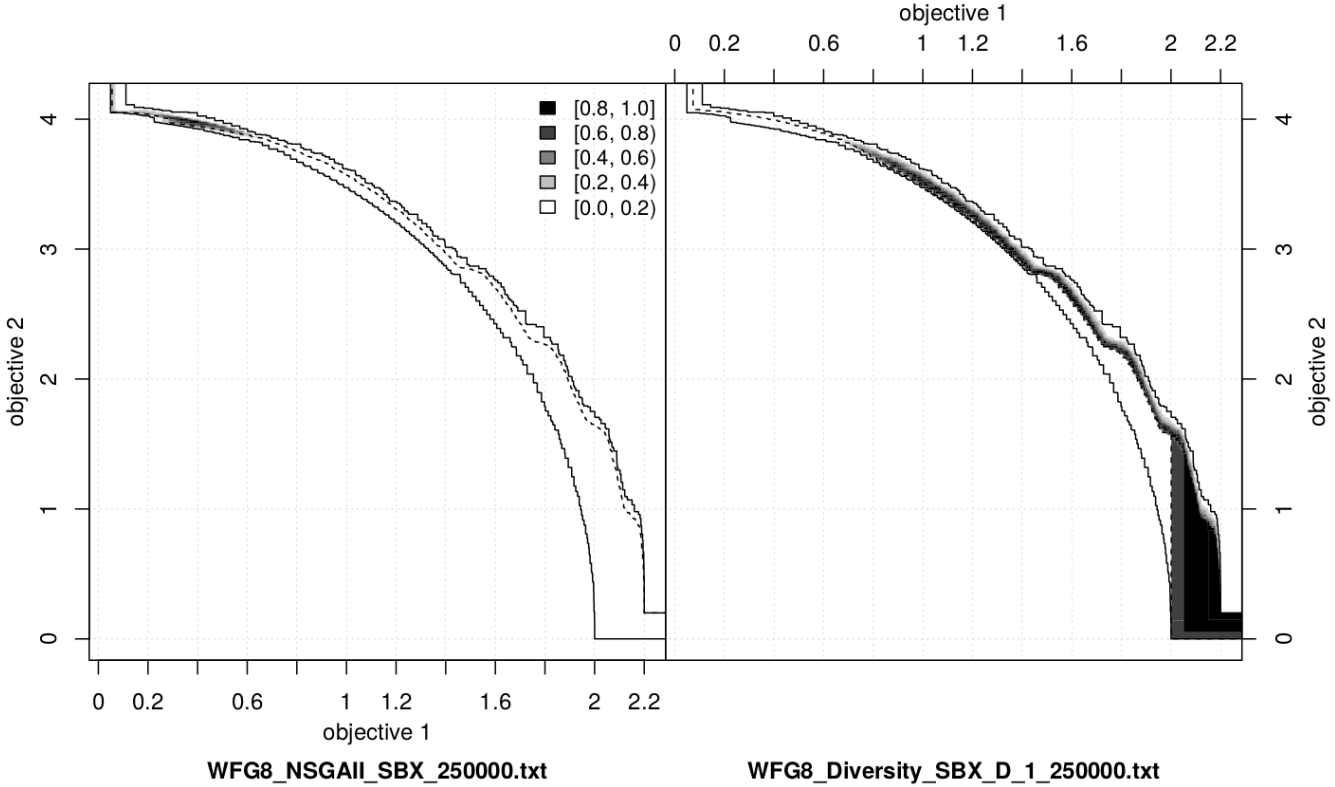


Figure 20: Diferencia estadística de dos algoritmos considerando 35 ejecuciones y el problema WFG8.

Pruebas estadísticas

- Las pruebas estadísticas se realizan comparando resultados de los algoritmos por pares.

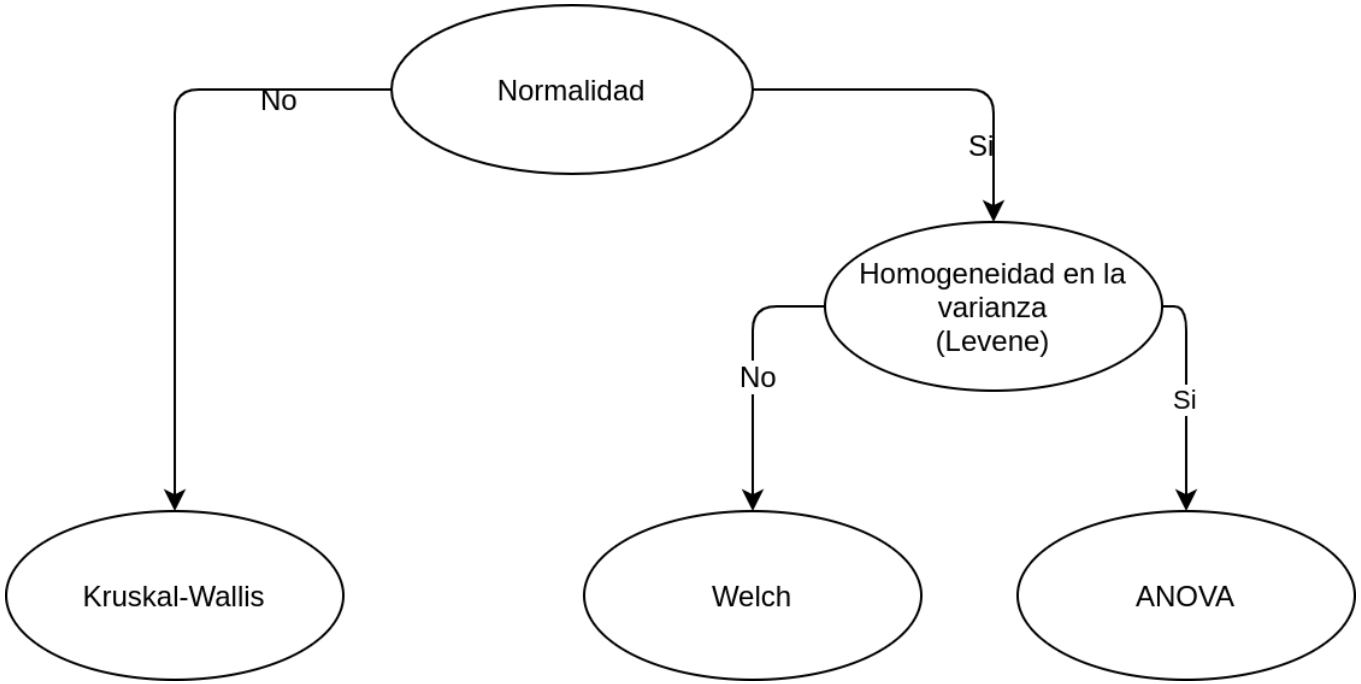


Figure 21: Esquema de las pruebas estadísticas[33].

[33] J. J. Durillo et al. "A Study of Multiobjective Metaheuristics When Solving Parameter Scalable Problems". In: *IEEE Transactions on Evolutionary Computation* 14.4 (Aug. 2010), pp. 618–635. ISSN: 1089-778X.

Operador de reemplazo *Best-Non-Penalized* BNP

Algorithm 2 BNP Survivor Selection Technique

Input: P_t (Population of current generation), Q_t (Offspring of current Generation)
Output: P_{t+1}
$R_t = P_t \cup Q_t$
$P_{t+1} = \emptyset$
$Penalized = \emptyset$
$D_t = D_I - D_I * \frac{G_{Elapsed}}{0.5 * G_{End}}$
while $|P_{t+1}| \leq$ N **do**
  Compute DCS of individuals in R_t, using P_{t+1} as a reference set
  Move the individuals in R_t with $DCS < D_t$ to $Penalized$
  if R_t is empty **then**
    Compute DCS of individuals in $Penalized$, using P_{t+1} as a reference set
    Move the individual in $Penalized$ with the largest DCS to R_t
  end if
  Select the new survivor from R_t as the individual with the best fitness and move it to P_{t+1}
end while
return P_{t+1}

Ejemplo

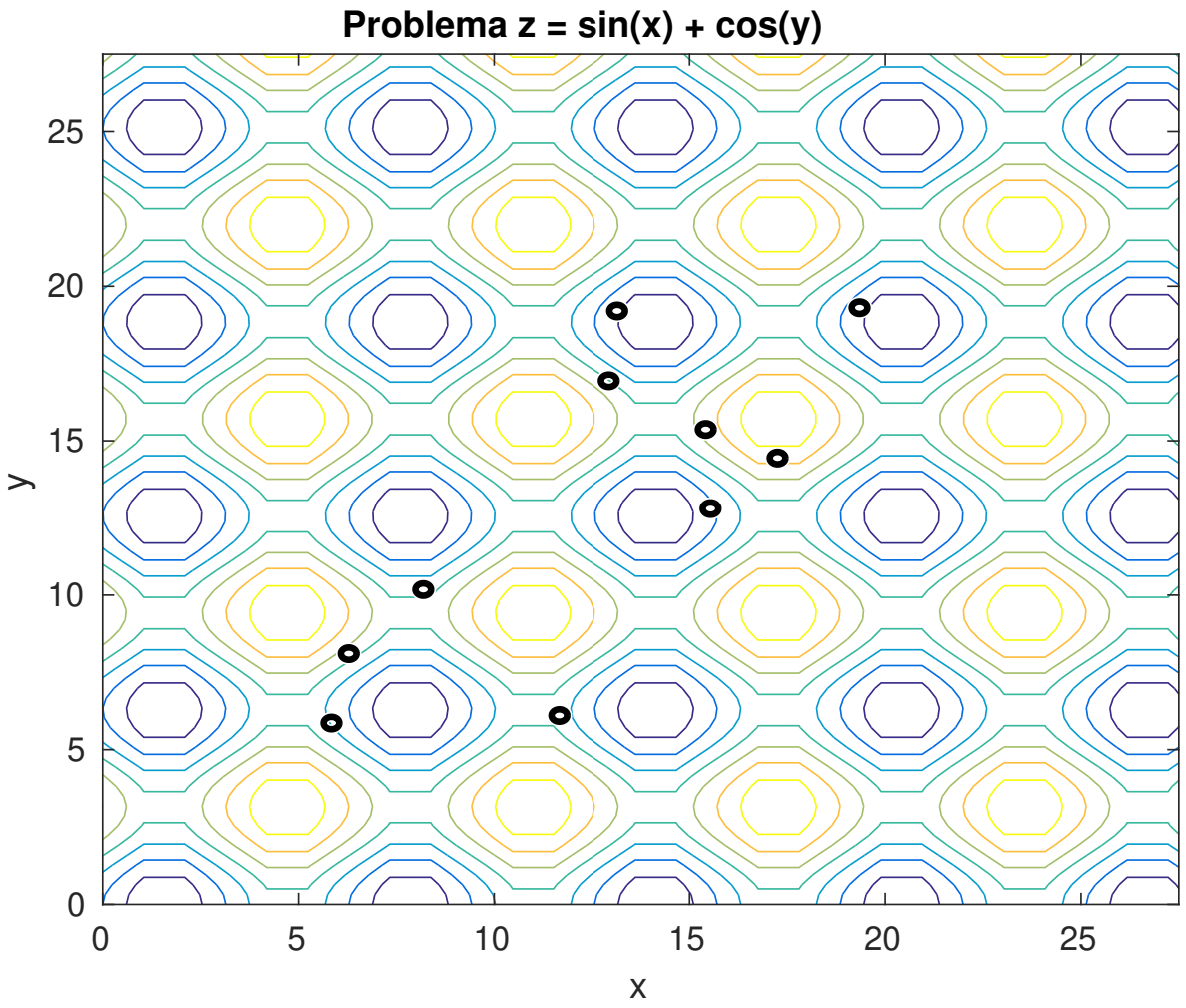


Figure 22: En color negro los candidatos, en rojo los seleccionados y en azul los penalizados.

Ejemplo

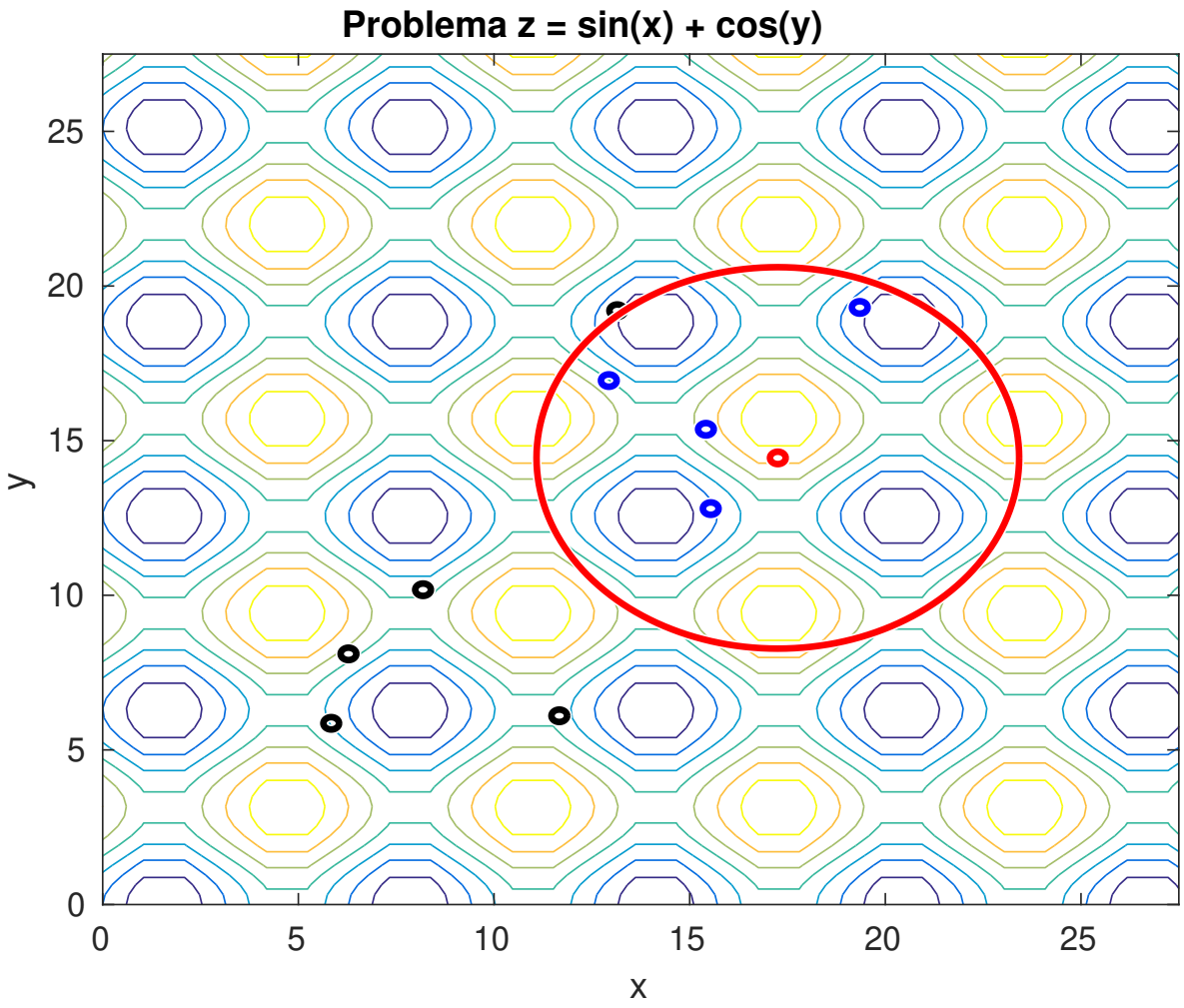


Figure 23: En color negro los candidatos, en rojo los seleccionados y en azul los penalizados.

Ejemplo

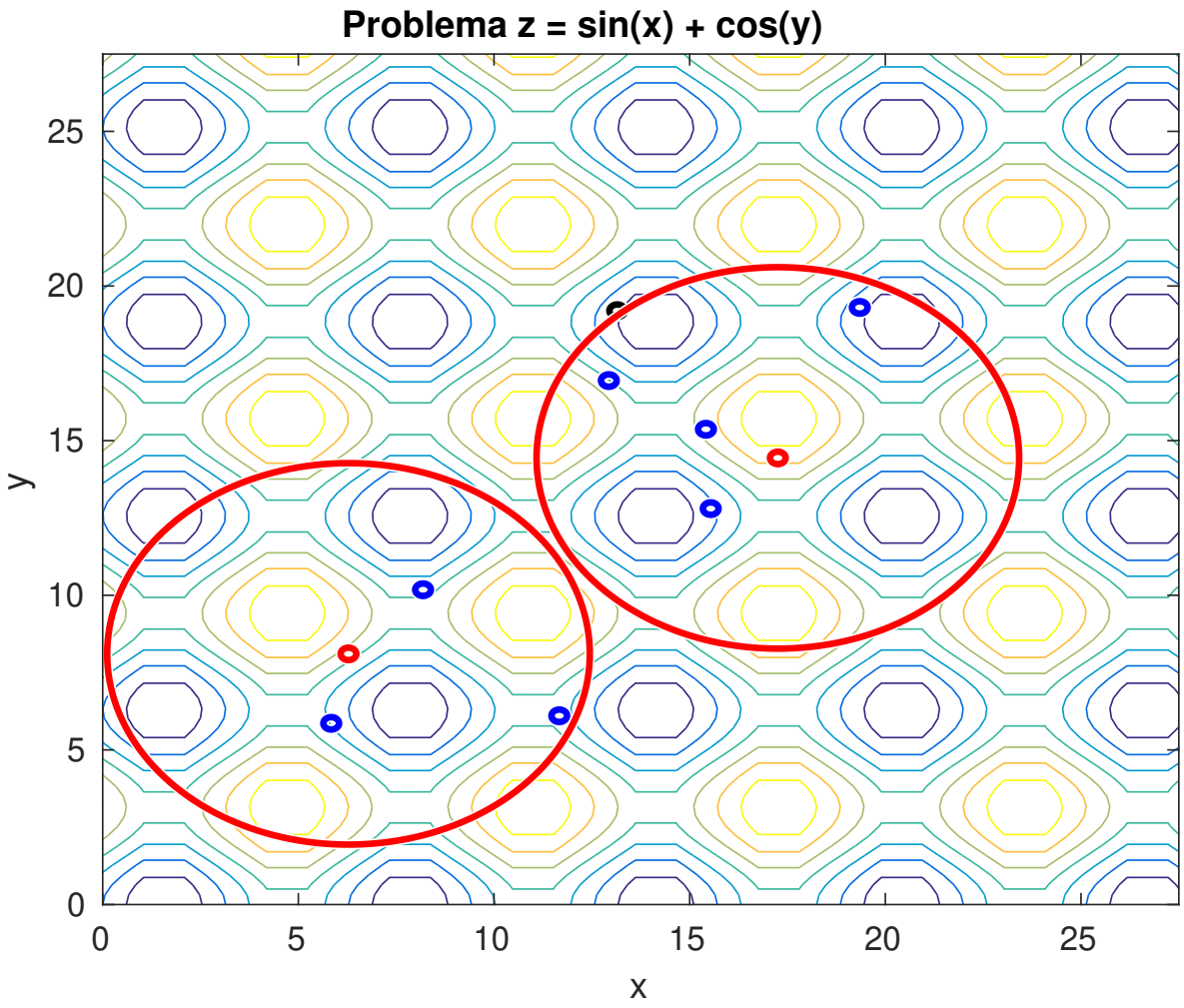


Figure 24: En color negro los candidatos, en rojo los seleccionados y en azul los penalizados.

Ejemplo

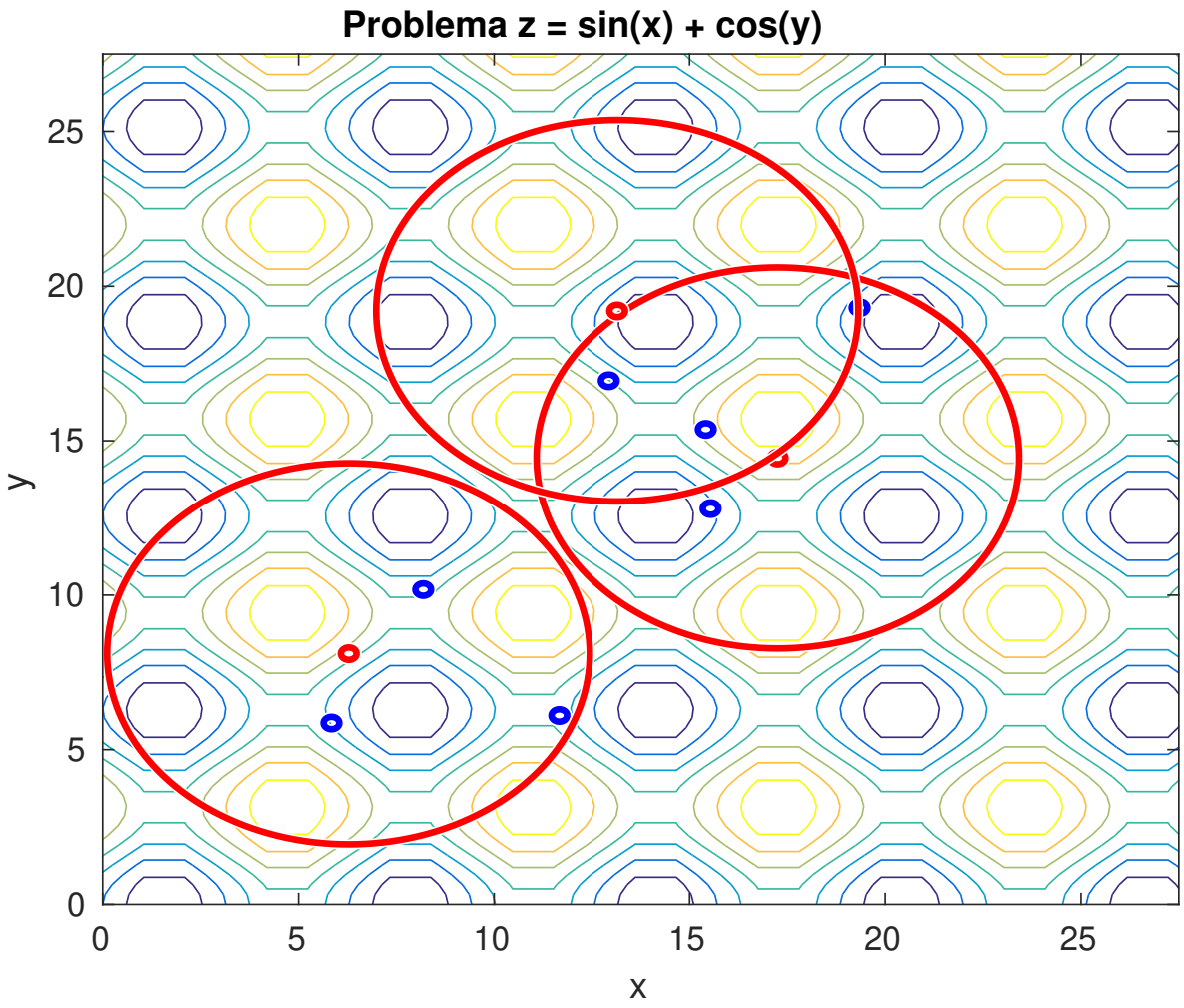


Figure 25: En color negro los candidatos, en rojo los seleccionados y en azul los penalizados.

Ejemplo

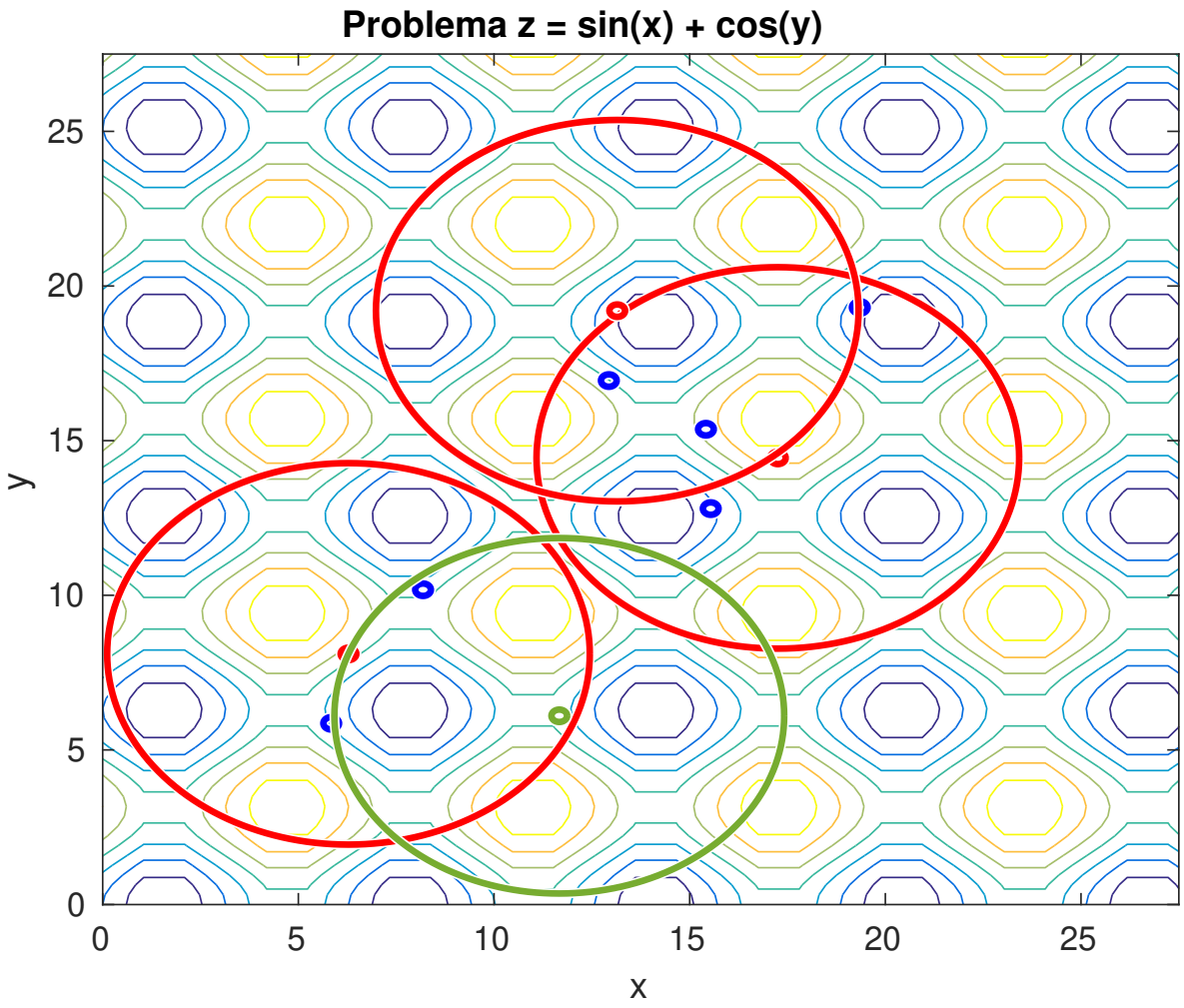


Figure 26: En color negro los candidatos, en rojo los seleccionados y en azul los penalizados.

Ejemplo

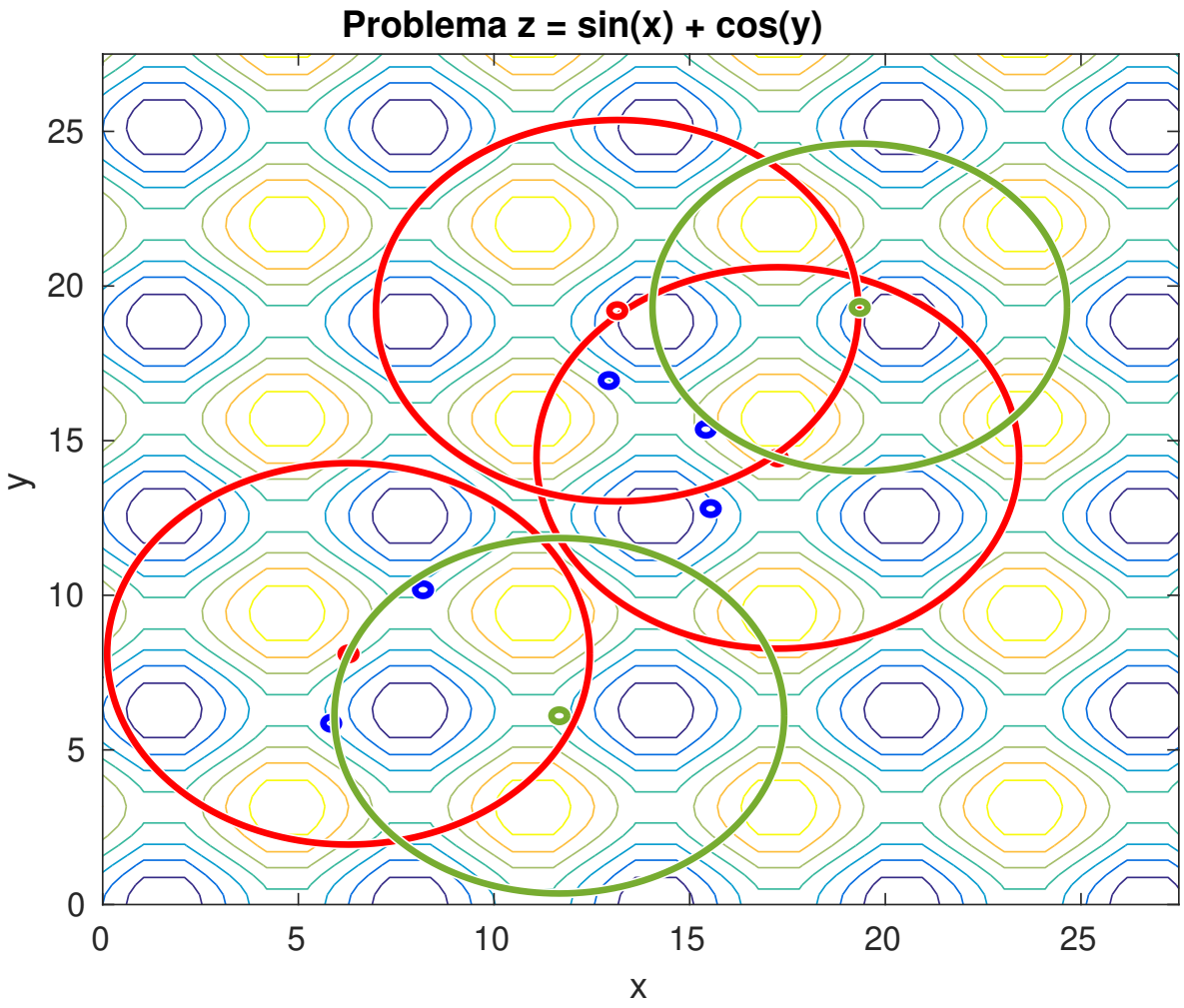
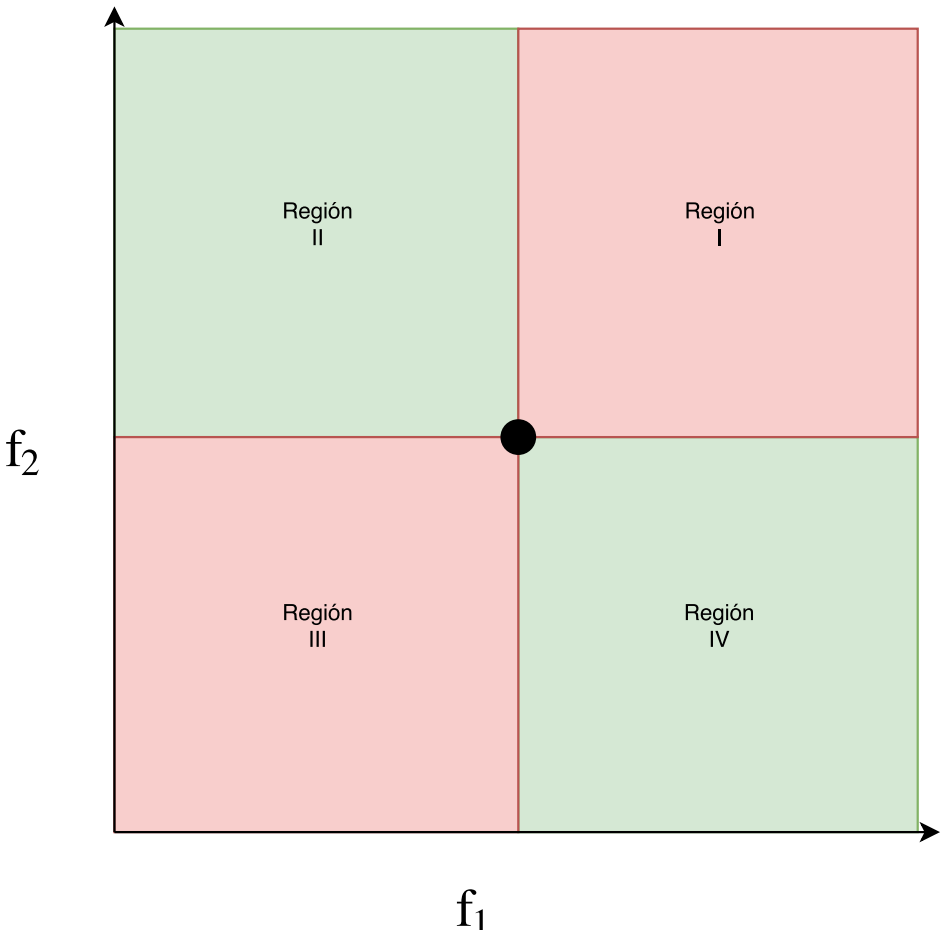


Figure 27: En color negro los candidatos, en rojo los seleccionados y en azul los penalizados.

Análisis preliminar

- Se utilizan los problemas WFG *The Walking Fish Group*.
- Estos problemas dividen a las variables de desición en dos tipos: las variables de distancia y las variables de posición.



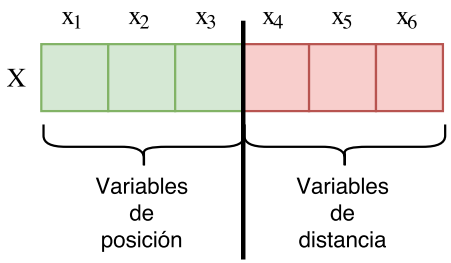


Figure 28: Variables de posición y distancia de los problemas WFG.

Análisis preliminar

- Se eligen los algoritmos NSGA-II, GDE3, MOEA/D y MOMBI-II, siendo basados en dominancia los dos primeros, y basados en descomposición e indicadores los dos últimos.
- Se selecciona el problema de prueba WFG1 que aunque posee una definición simple, muchos MOEAs tienen dificultades de resolverlo.
- Específicamente, la instancia WFG1 es un problema unimodal y separable.
- La región óptima para el WFG1 esta indicada en la ecuación (4), donde k es el número de variables de posición y $k+1$ a n son las variables de distancia.

$$x_{i=k+1:n} = 2i \times 0.35 \tag{4}$$

Análisis preliminar

- Se realizaron 35 ejecuciones con un criterio de paro de $50'000$ generaciones y el tamaño de población de 250.
- Particularmente, el análisis de diversidad es considerando en las variables de distancia.
- Se calcula la distancia promedio entre los individuos (ADI *Average Distance to all Individuals*).

Análisis preliminar

- En relativamente pocas generaciones (aproximadamente $5'000$), las variables de distancia convergen en todos los métodos.
- En algunas generaciones se recupera la diversidad por el efecto del operador de mutación polinomial.

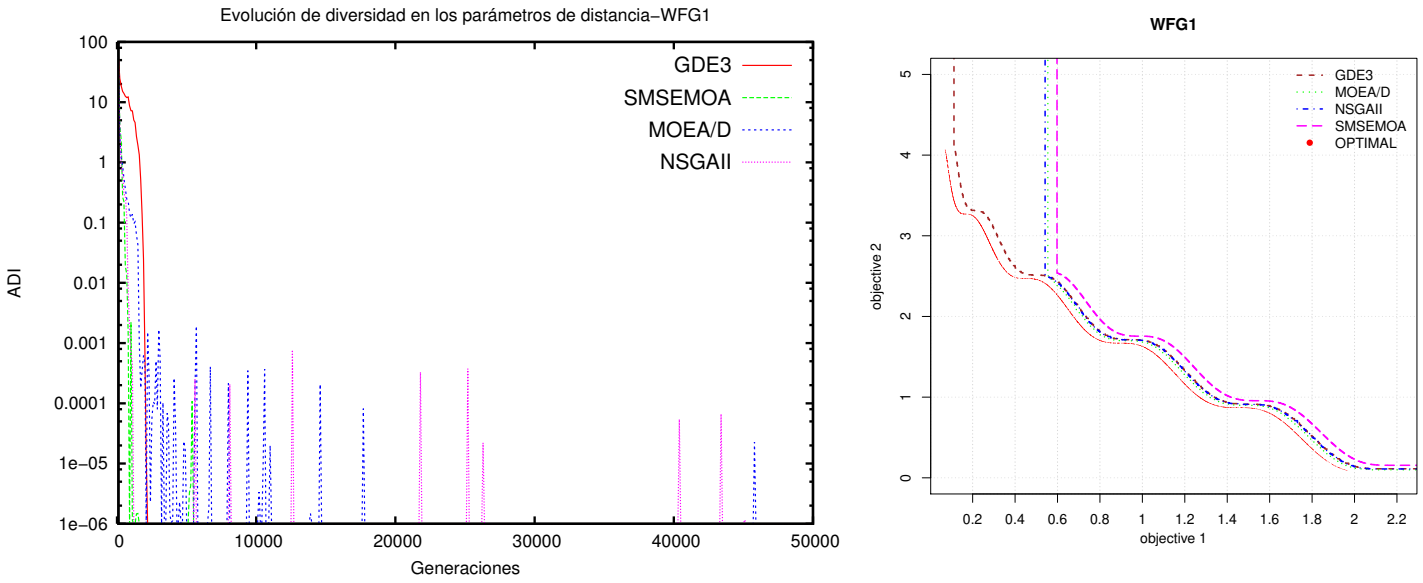


Figure 29: Evolución de la diversidad y superficie de alcance al 50%.

Hipótesis

- Los algoritmos evolutivos tienen un problema de convergencia acelerada en ejecuciones a largo plazo afectando la calidad de las soluciones.
- Este problema podría ser tratado implementando mecanismos para administrar la diversidad considerando el tiempo de ejecución y el criterio de parada.

Procedimiento principal del VSD-MOEA

- Se propone al VSD-MOEA (*Variable Space Diversity - Multi-objective Evolutionary Algorithm*).
- Se utiliza el esquema usual de un Algoritmo Evolutivo (EA - *Evolutionary Algorithm*).
- Se establece una fase especial de reemplazo, donde se define un procedimiento para administrar de forma explícita la diversidad.

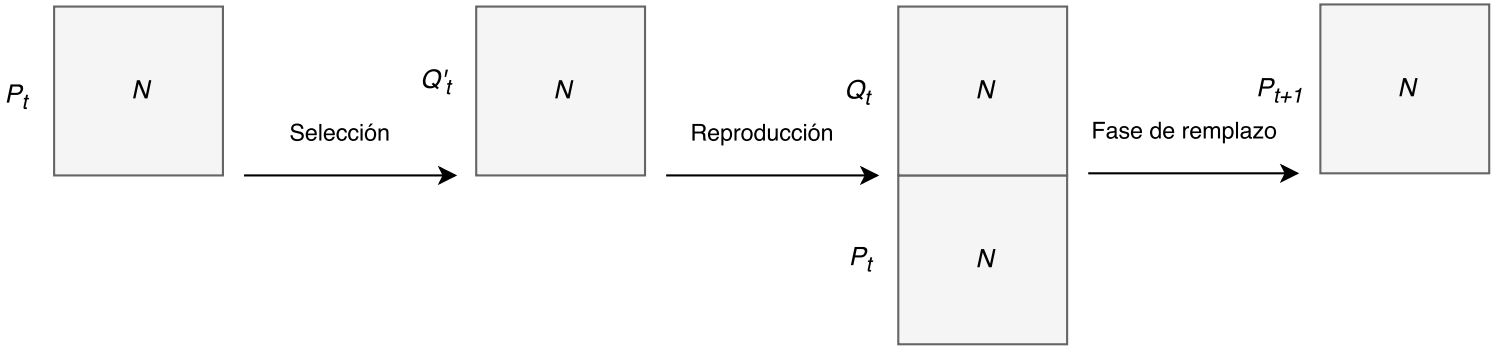
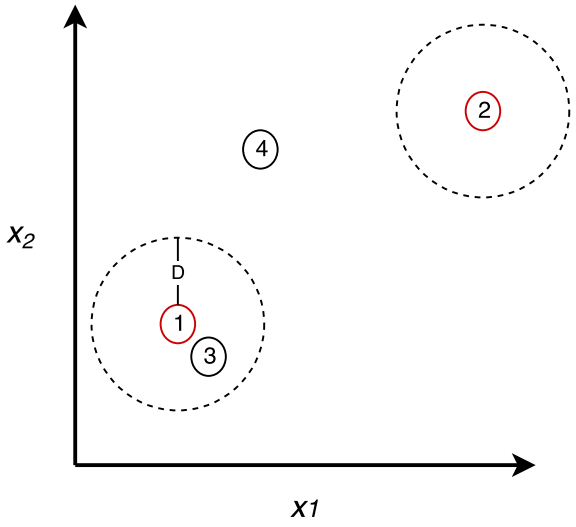
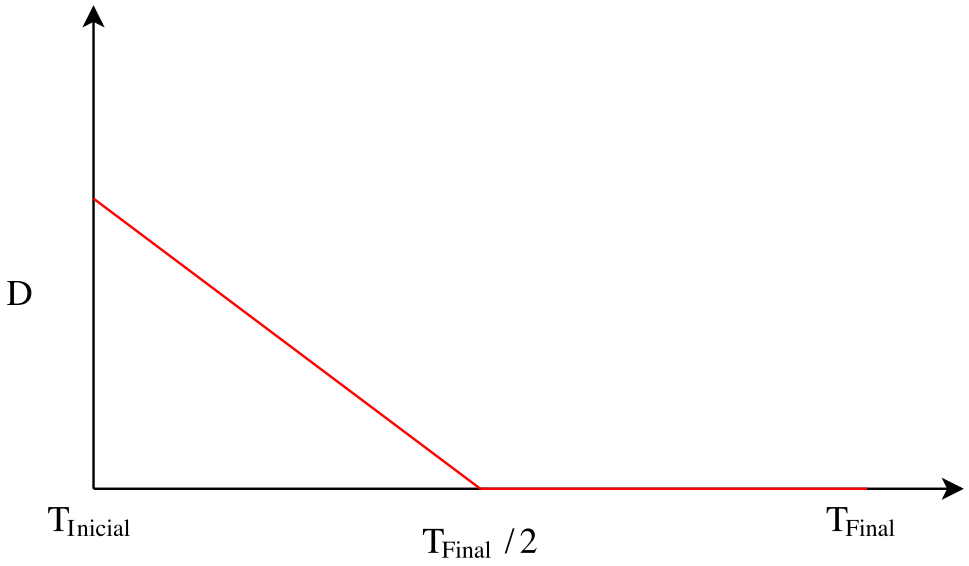


Figure 30: Proceso realizado en cada generación.

Fase de reemplazo

- La fase de reemplazo considera un método de penalización.
- El radio de la hiperesfera es decrementado conforme transcurre la ejecución.





Esquema general de la propuesta basada en dominancia

Algorithm 3 Main procedure of VSD-MOEA

1: **Initialization**: Generate an initial population P_0 with N individuals.
2: **Evaluation**: Evaluate all individuals in the population.
3: Assign $t = 0$
4: **while** (not stopping criterion) **do**
5: **Mating selection**: Fill the mating pool by performing binary tournament selection on P_t, based on the non-dominated ranks (ties are broken randomly).
6: **Variation**: Apply SBX crossover and Polynomial mutation to the mating pool to create a child population Q_t.
7: **Evaluation**: Evaluate all individuals in Q_t.
8: **Survivor selecction**: Generate P_{t+1} by applying the replacement scheme described in Algorithm , using P_t and Q_t as input.
9: $t = t + 1$
10: **end while**

Fase de reemplazo

Principalmente, la fase de reemplazo se basa en un método de penalización que implementa la métrica de mejoría. ($D^b(p_i, r_i) = \sum_{i \in M} max\{0, p_i - r_i\}^2$).

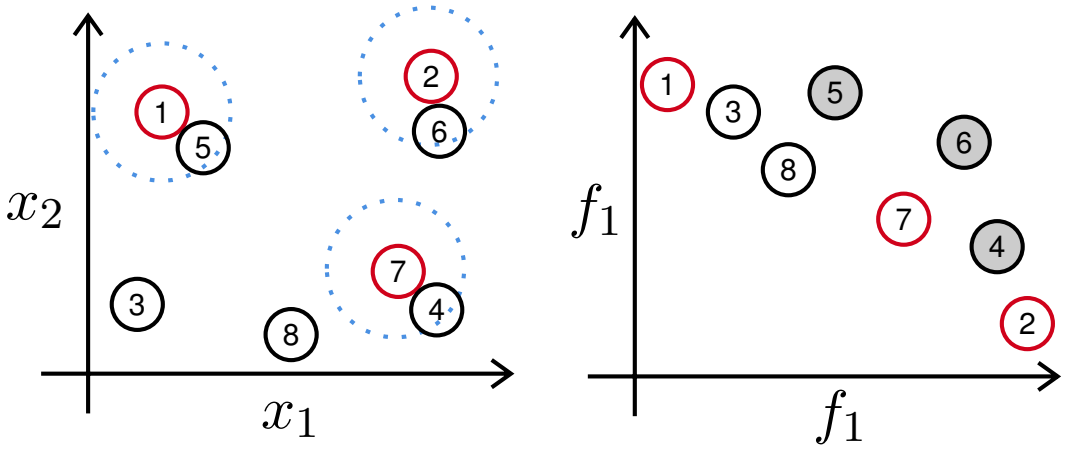


Figure 31: Método de penalización de la fase de reemplazo.

Fase de reemplazo

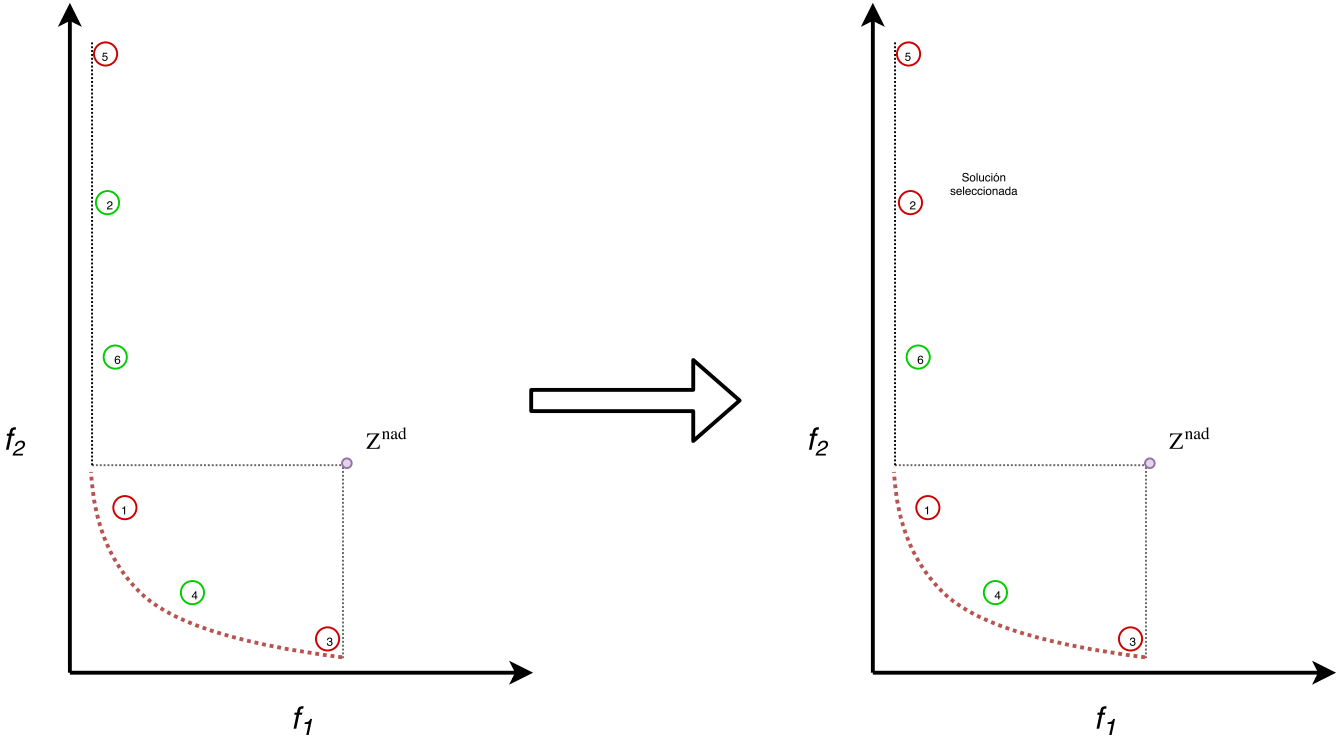
Algorithm 4 Replacement Phase of VSD-MOEA

Input: P_t (Population of current generation), Q_t (Offspring of current Generation)
Output: P_{t+1}
$R_t = P_t \cup Q_t$
$P_{t+1} = \emptyset$
$Penalized = \emptyset$
$D_t = D_I - D_I * \frac{G_{Elapsed}}{0.9 * G_{End}}$
while $|P_{t+1}| \leq$ N **do**
    Compute DCS of individuals in R_t with P_{t+1} used as reference set
    Move to $Penalized$ the individuals in R_t with $DCS < D_t$
    if R_t is empty **then**
        Compute DCS of individuals in $Penalized$ with P_{t+1} used as reference set
        Move to R_t the individual in $Penalized$ with largest DCS
    end if
    Identify the first front (F) in $R_t \cup P_{t+1}$ with an individual $I \in R_t$
    Use the novel density estimator to select a new survivor from F and move it to P_{t+1}
end while
return P_{t+1}

Procedimiento de selección

El estimador de densidad en el espacio objetivo considera el indicador ASF y la distancia de mejoría.

$$AWF_k(\vec{x}) = f_k(\vec{x}) + 10^{-4} \times \sum_{j=1}^{M} f_j(\vec{x}) \tag{5}$$



Procedimiento de selección

Algorithm 5 Density estimator

Input: P_{t+1} (Survivors), R_t (Candidates), F (Current front)
Output: $I \in R_t$
$FP = P_{t+1} \cap F$
$FR = R_t \cap F$
for $k \in$ number of objectives **do**
  Select the best individual $I \in F$ of k according to Eq. 5.
  if $I \in FR$ **then**
    return I
  end if
end for
$MaxID = 0$
for $Ic \in FR$ **do**
  $Improvement = \min_{s \in FP} ID(Ic, s)$
  if $Improvement > MaxID$ **then**
    $MaxID = Improvement$
    $I = Ic$
  end if
end for
return I

Comparación del estado-del-arte con ejecuciones a largo plazo

Pruebas estadísticas considerando al estado-del-arte

Table 1: Statistical Tests and Deterioration Level of the HV ratio for problems with two objectives

	$\uparrow$	$\downarrow$	$\leftrightarrow$	**Deterioration**
MOEA/D	24	36	9	1.615
NSGA-II	13	49	7	1.496
R2-EMOA	34	21	14	1.597
VSD-MOEA	50	15	4	0.059

Table 2: Statistical Tests and Deterioration Level of the HV ratio for problems with three objectives

	$\uparrow$	$\downarrow$	$\leftrightarrow$	**Deterioration**
MOEA/D	16	37	4	1.601
NSGA-II	9	45	3	2.557
R2-EMOA	31	22	4	1.223
VSD-MOEA	52	4	1	0.037

Comparación del estado-del-arte con ejecuciones a largo plazo

Table 3: Summary of the hypervolume ratio results attained for problems with two objectives

	MOEA/D				NSGA-II				R2-EMOA				VSD-MOEA			
	Min	**Max**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**
WFG1	**0.984**	**0.993**	**0.992**	**0.002**	0.987	0.993	0.992	0.002	0.946	0.994	0.988	0.012	0.980	0.994	0.992	0.003
WFG2	0.965	0.996	0.967	0.007	0.966	0.998	0.974	0.014	0.965	0.966	0.966	0.000	**0.998**	**0.998**	**0.998**	**0.000**
WFG3	0.992	0.992	0.992	0.000	0.987	0.988	0.987	0.000	0.991	0.992	0.991	0.000	**0.992**	**0.992**	**0.992**	**0.000**
WFG4	0.988	0.988	0.988	0.000	0.983	0.987	0.985	0.001	**0.991**	**0.991**	**0.991**	**0.000**	0.990	0.990	0.990	0.000
WFG5	0.876	0.893	0.882	0.005	0.884	0.899	0.890	0.002	0.886	0.895	0.891	0.003	**0.894**	**0.928**	**0.914**	**0.010**
WFG6	**0.879**	**0.940**	**0.914**	**0.016**	**0.894**	**0.942**	**0.913**	**0.012**	**0.875**	**0.942**	**0.912**	**0.015**	0.855	0.888	0.868	0.007
WFG7	0.988	0.988	0.988	0.000	0.983	0.987	0.984	0.001	**0.991**	**0.991**	**0.991**	**0.000**	0.990	0.990	0.990	0.000
WFG8	0.800	0.822	0.811	0.006	0.771	0.801	0.789	0.006	0.803	0.824	0.815	0.005	**0.828**	**0.958**	**0.928**	**0.046**
WFG9	0.795	0.972	0.883	0.082	0.793	0.966	0.832	0.070	0.797	0.976	0.884	0.079	**0.963**	**0.975**	**0.970**	**0.004**
DTLZ1	**0.993**	**0.993**	**0.993**	**0.000**	0.990	0.992	0.991	0.000	0.992	0.992	0.992	0.000	0.992	0.992	0.992	0.000
DTLZ2	0.989	0.989	0.989	0.000	0.986	0.988	0.987	0.000	**0.991**	**0.992**	**0.992**	**0.000**	0.990	0.990	0.990	0.000
DTLZ3	0.989	0.989	0.989	0.000	0.987	0.989	0.989	0.001	**0.991**	**0.992**	**0.992**	**0.000**	0.990	0.990	0.990	0.000
DTLZ4	0.259	0.989	0.781	0.330	0.259	0.988	0.863	0.274	0.259	0.992	0.657	0.365	**0.990**	**0.990**	**0.990**	**0.000**
DTLZ5	0.989	0.989	0.989	0.000	0.986	0.988	0.987	0.000	**0.991**	**0.992**	**0.992**	**0.000**	0.990	0.990	0.990	0.000
DTLZ6	0.448	0.910	0.700	0.105	0.138	0.511	0.322	0.075	0.510	0.922	0.691	0.107	**0.990**	**0.990**	**0.990**	**0.000**
DTLZ7	0.996	0.996	0.996	0.000	0.996	0.997	0.996	0.000	**0.997**	**0.997**	**0.997**	**0.000**	0.996	0.996	0.996	0.000
UF1	0.991	0.993	0.992	0.000	0.986	0.989	0.988	0.000	0.978	0.994	0.990	0.005	**0.994**	**0.995**	**0.994**	**0.000**
UF2	**0.987**	**0.993**	**0.991**	**0.002**	0.980	0.983	0.981	0.001	0.984	0.991	0.988	0.002	0.987	0.993	0.990	0.001
UF3	0.481	0.674	0.597	0.043	0.678	0.871	0.784	0.048	0.531	0.704	0.589	0.041	**0.799**	**0.916**	**0.881**	**0.025**
UF4	0.881	0.917	0.908	0.006	0.875	0.910	0.889	0.008	**0.923**	**0.935**	**0.929**	**0.003**	0.923	0.931	0.927	0.002
UF5	0.035	0.792	0.484	0.165	0.256	0.766	0.641	0.104	0.123	0.792	0.566	0.192	**0.582**	**0.763**	**0.647**	**0.040**
UF6	0.255	0.711	0.447	0.114	0.235	0.801	0.635	0.120	0.349	0.767	0.568	0.113	**0.668**	**0.900**	**0.810**	**0.061**
UF7	**0.987**	**0.991**	**0.990**	**0.001**	0.980	0.983	0.981	0.001	0.557	0.991	0.910	0.150	**0.975**	**0.992**	**0.988**	**0.004**
Mean	0.806	0.935	0.881	0.038	0.808	0.927	0.886	0.032	0.801	0.940	0.882	0.048	0.929	0.963	0.949	0.009

Superficies de cubrimiento

Poner superficie de diferencias.

Análisis de escalabilidad en el espacio de las variables

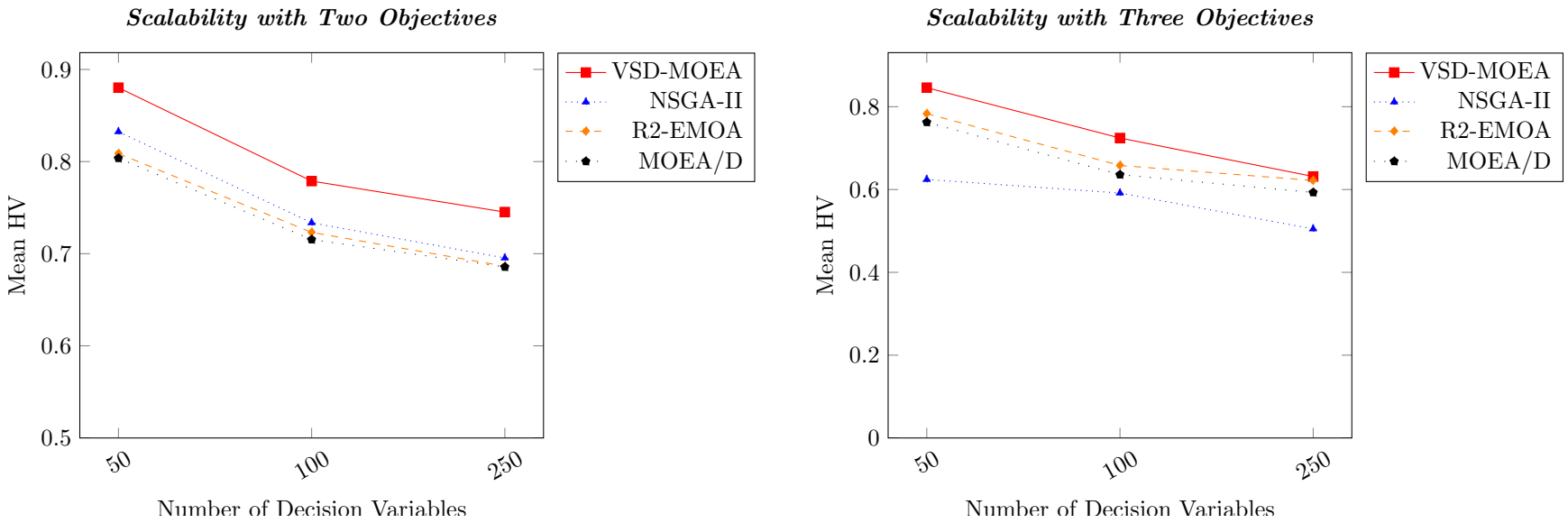


Figure 32: Mean of the HV ratio for 35 runs for the two-objective and three-objective problems considering different numbers of variables

Análisis de diversidad en el espacio de la variables

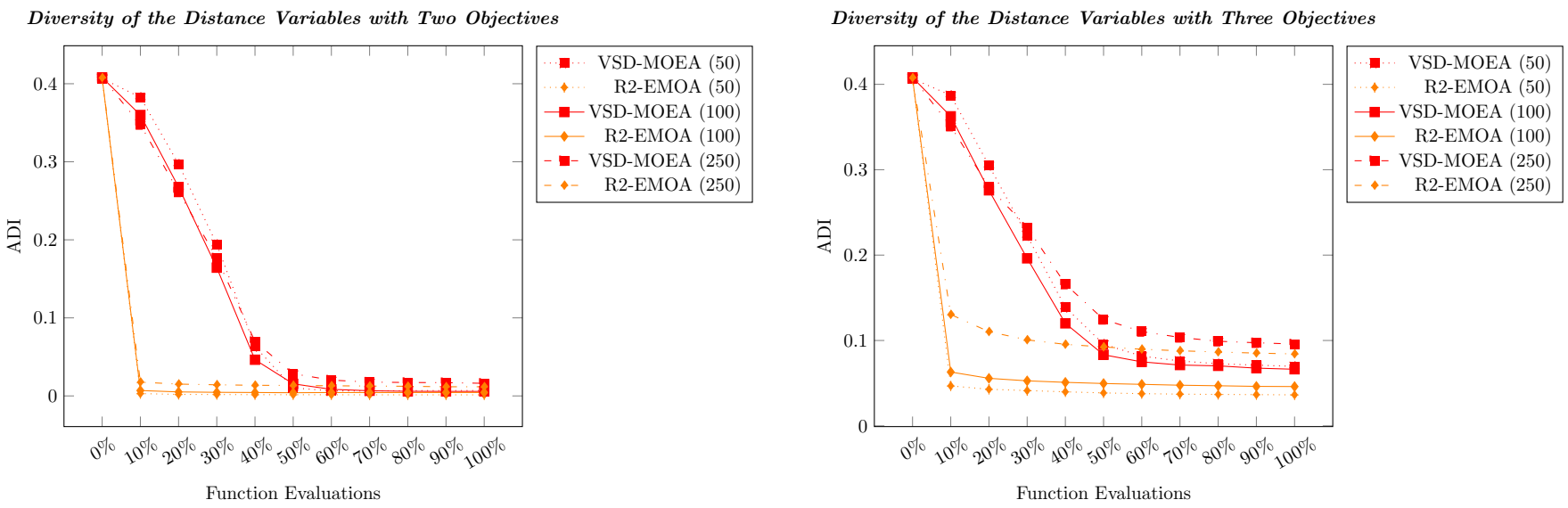


Figure 33: Evolution of ADI for problems WFG1-WFG7 with two objectives considering only the distance variables

Análisis de rendimiento en función al criterio de paro

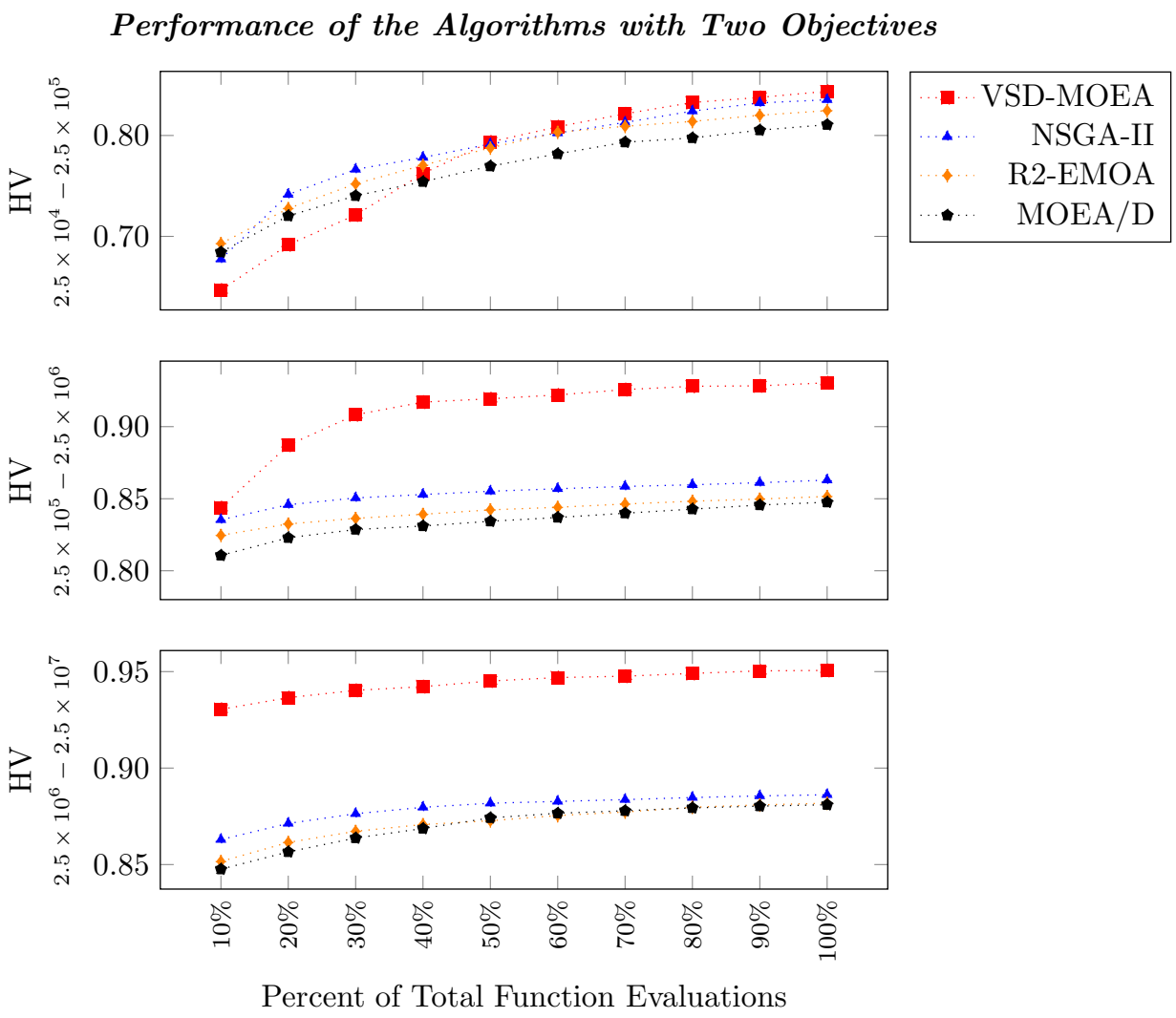


Figure 34: Performance of MOEAs for the problems with two objectives considering three ranges for the stopping criterion: short-term (first row), middle-term (second row) and long-term (third row).

Análisis de rendimiento en función al criterio de paro

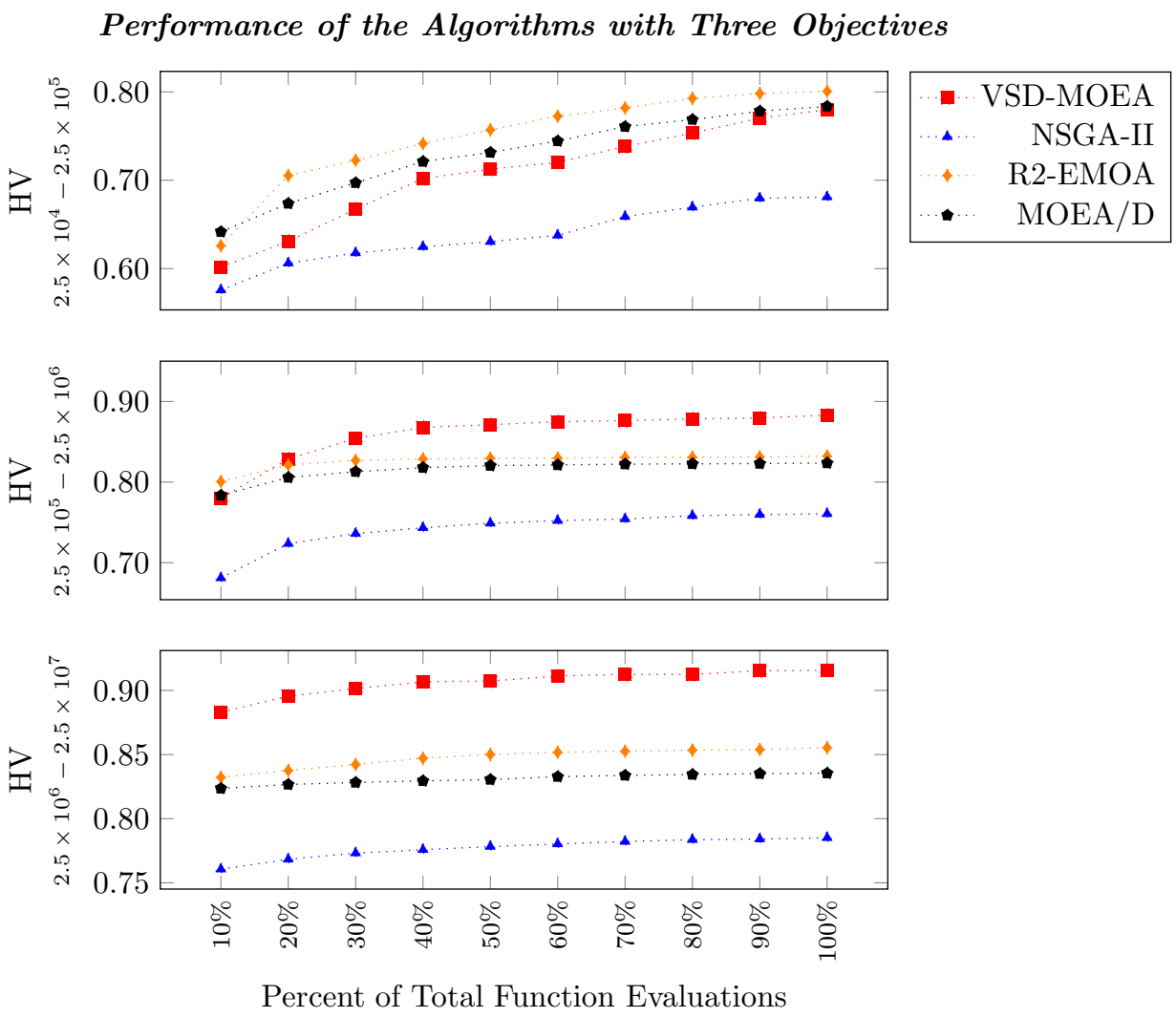


Figure 35: Performance of MOEAs for the problems with three objectives considering three ranges for the stopping criterion: short-term (first row), middle-term (second row) and long-term (third row).

Análisis del umbral de diversidad inicial

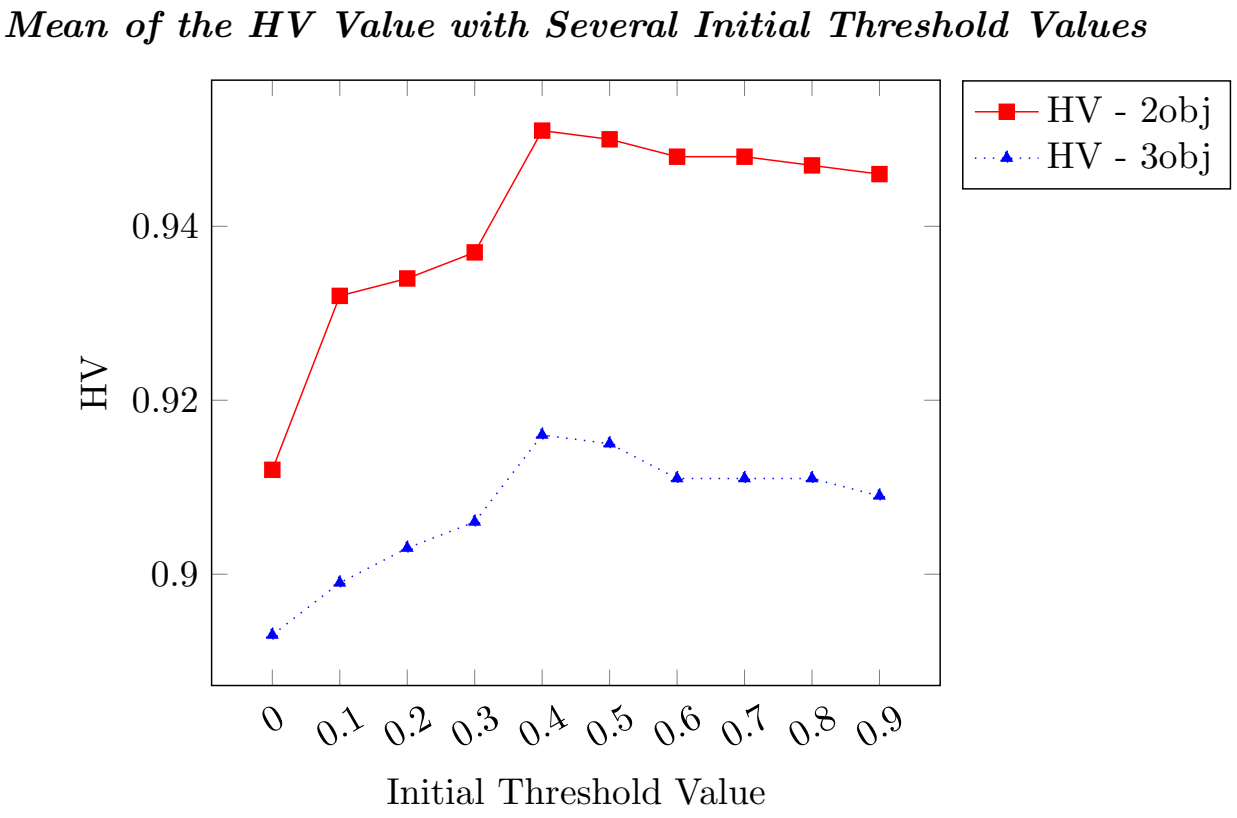


Figure 36: Mean of HV values taking into account all the problems with several initial threshold values

Simulación

https://youtu.be/dbk5DaFJ8y0

Trabajo en progreso

- Algoritmo multi-objectivo basado en decomposición (VSD-MOEA).

Conclusiones y trabajo futuro

- Los algoritmos evolutivos que consideran la diversidad en el espacio de las variables de forma explícita y en función al criterio de parada ofrecen soluciones de calidad en esquemas de largo plazo.
- En el caso multi-objetivo, la cantidad de evaluaciones a función necesarias para mejorar al estado del arte es a partir de 2'500 generaciones.
- Con el fin de obtener un rendimiento superior en los algoritmos de optimización es necesario considerar el criterio de parada para tomar decisiones.

Acknowledgements

This work has been supported by the Center for Research in Mathematics.