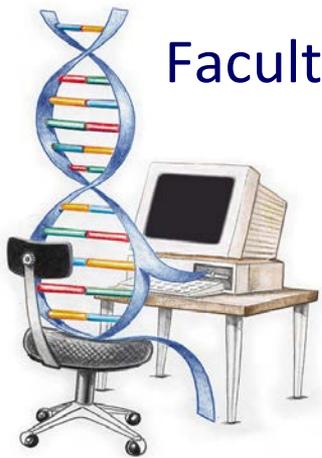


ALGORITMOS EVOLUTIVOS

Curso 2024

Tema 10: AE para optimización multiobjetivo

Centro de Cálculo, Instituto de Computación
Facultad de Ingeniería, Universidad de la República, Uruguay



cecal



Contenido

1. Problemas de optimización multiobjetivo
2. AE para optimización multiobjetivo (MOEA)
 - Non Pareto based MOEAs
 - Pareto based MOEAs
3. Antecedentes: VEGA
4. MOEAs de primera generación:
 - NSGA, NPGA, MOGA
5. MOEAs de segunda generación:
 - SPEA y SPEA-2, PAES, NSGA-II, NPGA-2, MICRO GA
6. Métricas para optimización multiobjetivo

Definiciones

- En general se hace referencia al problema de optimización multiobjetivo por su sigla en inglés MOP (Multiobjective Optimization Problem).
- Las principales características de un MOP son:
 - trabajan sobre un espacio **multidimensional** de funciones.
 - **no existe una única solución** al problema.
 - es necesario un proceso de **toma de decisiones** en cual se decide que tipo de compromisos son más convenientes desde la perspectiva del tomador de decisiones. Este proceso puede ser realizado **a priori** o **a posteriori**.

Formulación

- Problema de optimización multiobjetivo

$$\text{Min / Max} \quad \vec{F}(x) = (f_1(x), f_2(x), \dots, f_n(x))$$

$$\text{sujeto a} \quad \vec{G}(x) = (g_1(x), g_2(x), \dots, g_s(x)) \geq 0$$

$$\vec{H}(x) = (h_1(x), h_2(x), \dots, h_r(x)) = 0$$

siendo $x = (x_1, x_2, \dots, x_m)$ las variables de decisión del problema
y $x \in \Omega$ (el espacio de soluciones factibles)

- Las funciones son **vectoriales**

Conceptos

- La noción de óptimo se modifica, ya que no es posible encontrar una solución única que sea óptima para todas las funciones a optimizar.
- Una solución es un **óptimo de Pareto** si no existe ningún vector factible que decremente algún criterio sin causar un incremento simultáneo en al menos otro criterio.

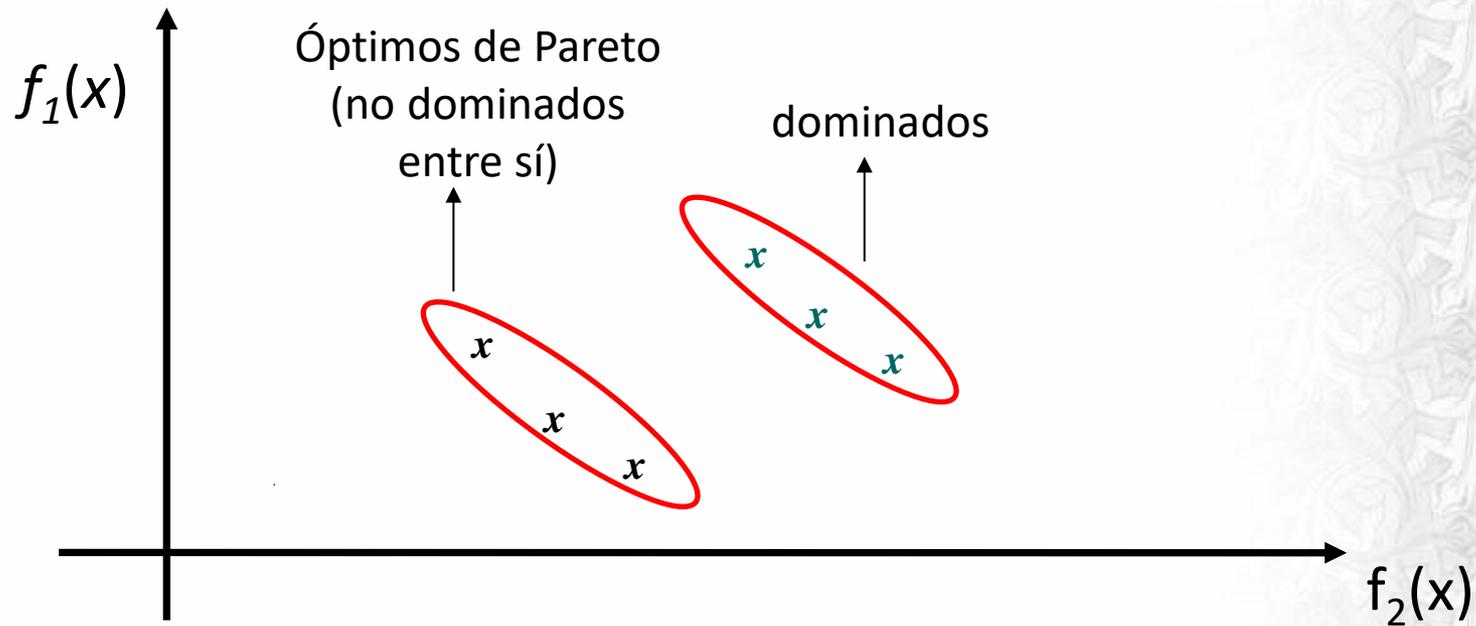
x^* es un óptimo de Pareto si $\forall x \in \Omega$ e $I = 1, \dots, n$:

$$\Rightarrow \begin{cases} \forall i \in I (f_i(x) = f_i(x^*)) & \text{ó} \\ f_i(x) > f_i(x^*) & \text{para al menos un índice } i \end{cases}$$

- Los vectores correspondientes a las soluciones no incluidas en el conjunto de óptimos de Pareto son llamados soluciones **no dominadas**.
- La **dominancia de Pareto** es una relación de orden parcial entre las soluciones factibles ($x \in \Omega$).

$w = (w_1, w_2, \dots, w_n)$ domina a $v = (v_1, v_2, \dots, v_n)$ si:

$$w_i \leq v_i \quad \forall i = 1, \dots, n \wedge \exists j / w_j < v_j, j \in \{1, \dots, n\}$$



Dominancia entre soluciones de un problema multiobjetivo

Óptimos de Pareto y frente de Pareto

- Las soluciones del MOP están dadas por el conjunto de soluciones factibles no dominadas al que se denomina conjunto de **óptimos de Pareto**.
- Al conjunto de valores funcionales de los óptimos de Pareto se le denomina **frente de Pareto (FP)**.

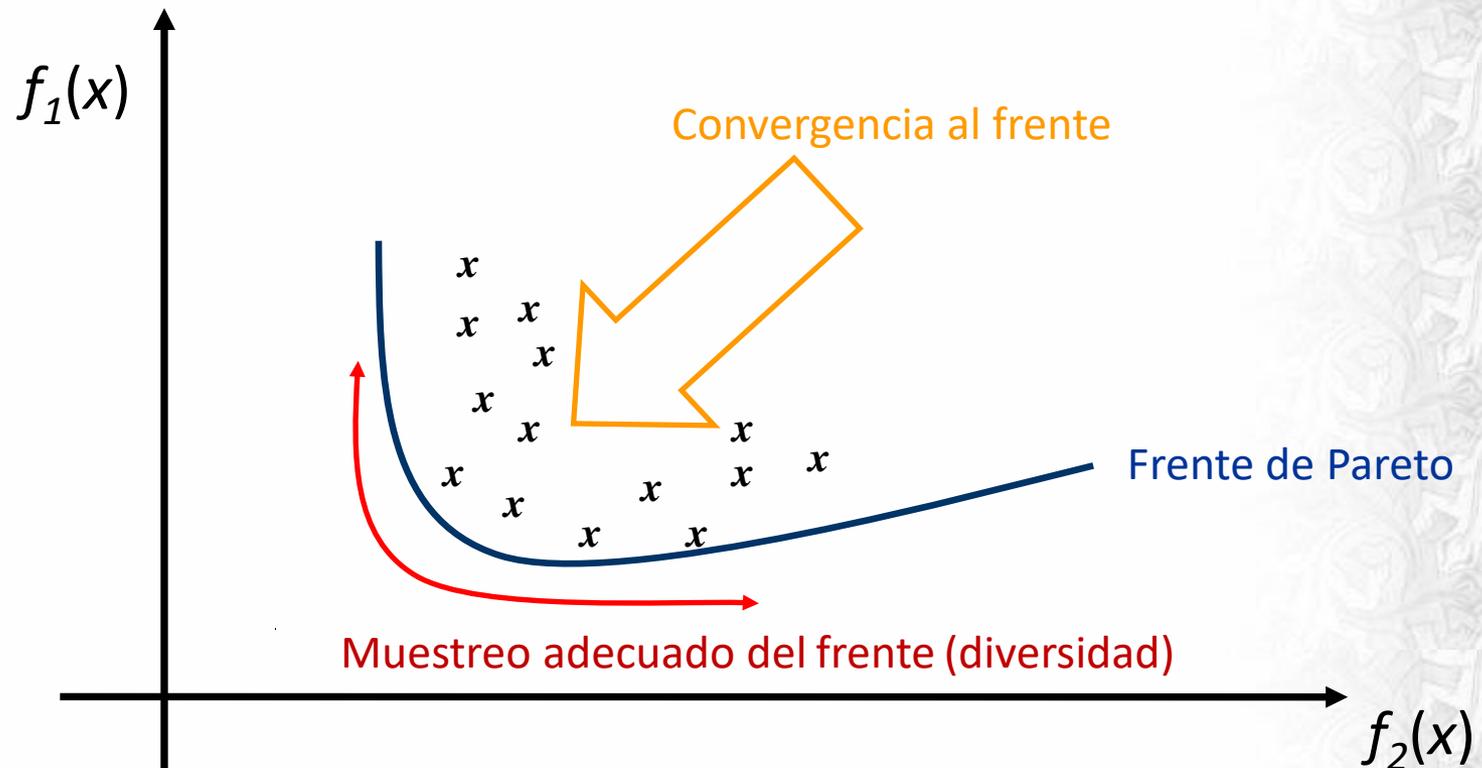
$$P^* = \{x \in \Omega \mid \neg \exists x' \in \Omega : f(x') \text{ domina a } f(x)\}$$
$$FP = \{u = (f_1(x), f_2(x), \dots, f_n(x)) \mid x \in P^*\}$$

- MOEA corresponde a la sigla (en inglés) de MultiObjective Evolutionary Algorithm.
- La primer referencia existente sobre una sugerencia de aplicación de AE a los MOPs corresponde a Rosenberg en 1967.
- Sin embargo, el primer MOEA recién fue propuesto por Schaffer en 1984.
- Ventajas de un EA para resolver problemas multiobjetivo:
 - al trabajar en paralelo sobre un conjunto de soluciones tienen la potencialidad de tratar problemas con múltiples objetivos, hallando en cada ejecución un conjunto de soluciones aproximadas al frente de Pareto.
 - son menos sensibles a la forma o a la continuidad del FP.
 - permiten abordar problemas con espacio de soluciones de gran dimensión.

Propósitos de un MOEA

- Los propósitos de un MOEA son:
 - **aproximarse al frente de Pareto** del MOP.
 - **muestrear adecuadamente el frente de Pareto**, hallando alternativas que expresen diferentes compromisos entre las funciones a optimizar (permitiendo realizar la toma de decisiones a posteriori).
 - **tener buenos valores de eficiencia computacional** (aunque en general se considera, erróneamente, un propósito secundario).

Propósitos de un MOEA



Propósitos en la resolución de un problema multiobjetivo.

Esquema algorítmico

```
Inicializar(P( $\theta$ ))
generacion = 0
mientras (no CriterioParada) {
    Evaluar(P(generacion))
    Operador de diversidad(P(generacion))
    Asignar fitness(P(generacion))
    Padres = Seleccion(P(generacion))
    Hijos = Operadores de Reproduccion(Padres)
    NuevaPop = Reemplazar(Hijos,P(generacion))
    generacion ++
    P(generacion) = NuevaPop
} retornar frente de Pareto
```

Esquema de un AE para optimización multiobjetivo.

Esquema algorítmico

- En el algoritmo previo se observan dos operadores que son característicos de los MOEAs y que no aparecen en la estructura genérica de un EA: el **operador de diversidad** y el **operador de asignación de fitness**.
- El operador de diversidad aplica una técnica para evitar la convergencia a un sector del frente de Pareto.
- La asignación de fitness está orientada a brindar una mayor chance de perpetuarse a aquellos individuos con mejores características, considerando los valores de las funciones objetivo y los resultados de la métrica utilizada para evaluar la diversidad.
- Algunas propuestas de MOEAs no siguen exactamente el esquema algorítmico presentado.

Clasificación

- De acuerdo a su mecanismo de asignación de fitness, los MOEAs se clasifican en **non-Pareto based MOEAs** y **Pareto based MOEAs**.
- *Non-Pareto based MOEAs*
 - Proponen mecanismos de asignación de fitness relativamente sencillos, que no reflejan la independencia entre los objetivos del problema y en general no garantizan una correcta resolución del MOP.
- *Pareto based MOEAs*
 - Utilizan explícitamente la dominancia de Pareto en el procedimiento de asignación de fitness.

Métodos no basados en Pareto

- Proponen mecanismos de asignación de fitness relativamente simples
- En la práctica, no garantizan en general una correcta resolución del MOP: solo son adecuados para problemas con no más de 3 funciones objetivo y espacio de soluciones sencillo.
- Existen varias propuestas de mecanismos de asignación:
- **Agregación o combinación lineal de objetivos**

- El fitness se calcula como una suma ponderada:

$$F(x) = \sum_{i=1}^n w_i f_i(x)$$

- Los pesos se fijan a priori.
- La expresión **no refleja el carácter multiobjetivo** del problema.

- Ordenamiento lexicográfico

- Las funciones objetivo se jerarquizan en orden de importancia.
- Se resuelven en el orden de la jerarquía, agregando restricciones ficticias para las funciones objetivo resueltas previamente.

- ε -constraint

- Se optimiza una de las funciones objetivo y se considera el resto como restricciones acotadas por ciertos niveles permisibles a los que se denomina ε_i .
- A continuación se modifican los valores de ε_i con el fin de generar el conjunto de Pareto.

- Target vector optimization

- Intenta alcanzar un “vector ideal” mediante la combinación de estrategias de optimización.
- Es el mecanismo utilizado por el algoritmo VEGA (Schaffer, 1984), la propuesta pionera de AE para resolver MOPs.

- Goal attainment (satisfacción de metas)

- Se basa en “alcanzar metas” para cada una de las funciones objetivo.
- Las metas son definidas a priori por el usuario.

Métodos basados en Pareto

- Utilizan explícitamente la dominancia de Pareto para asignar el fitness.
- Suelen utilizar mecanismos para la preservación de la diversidad. Estas técnicas usualmente se utilizan en AE mono-objetivo para problemas multimodales.
- Mecanismos para la preservación de la diversidad usados frecuentemente:
 - *Fitness sharing*
 - Aplicado en el espacio de soluciones o en el espacio de valores funcionales.
 - Depende fuertemente de la función de distancia utilizada.

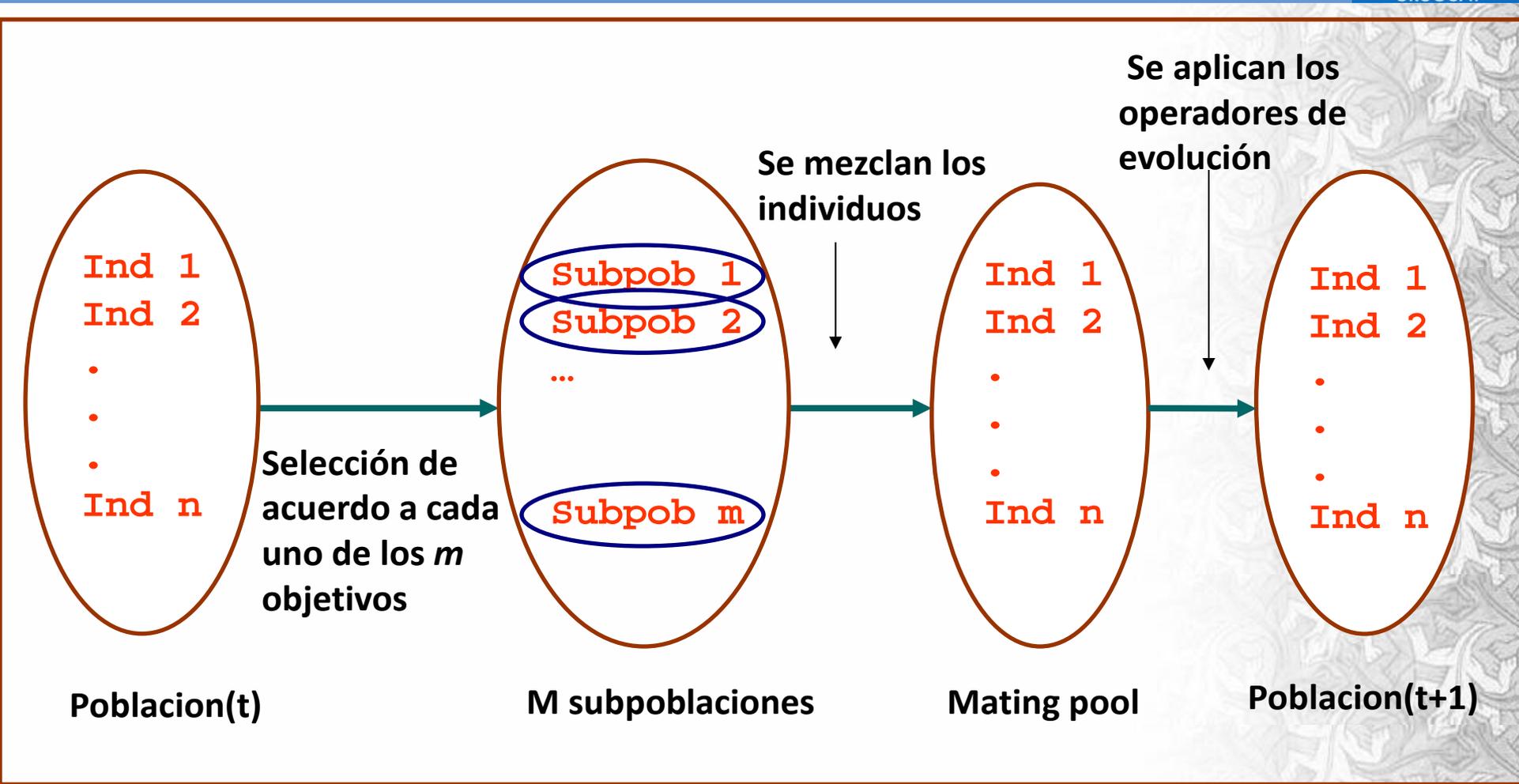
Métodos basados en Pareto

- Mecanismos para la preservación de la diversidad usados frecuentemente:
 - *Restricción al cruzamiento*
 - Impide el cruzamiento de individuos muy similares (cuya distancia sea menor que un valor σ_{MATING})
 - *Crowding*
 - No tiene dependencia paramétrica
 - *Otros mecanismos*
 - *Isolation by distance*, mantiene individuos “alejados”

Antecedentes

- David Schaffer realizó la primer propuesta de MOEA en 1984: el Vector Evaluated Genetic Algorithm (VEGA).
- La idea consiste en incorporar una estrategia de selección especial, que considera los objetivos múltiples en un AE simple.
- Para un problema con m objetivos se generan m subpoblaciones de tamaño $\#poblacion/m$.
- Los individuos son seleccionados de acuerdo a cada uno de los objetivos.
- El mecanismo de selección utilizado se basa en usar soluciones localmente dominadas en cada generación.
- No utiliza dominancia de Pareto.

El algoritmo VEGA



Mecanismo de evolución de VEGA.

El algoritmo VEGA

- Si bien se tienen en cuenta las distintas funciones a optimizar, el mecanismo de búsqueda es unidimensional.
- Sin embargo surge un problema relacionado con la especialización:
 - Los individuos se especializan en optimizar f_1, \dots, f_m **por separado**.
 - Un individuo que tenga buenos valores de compromiso pero que no optimiza en particular ninguna de las funciones objetivo f_1, \dots, f_m **no sobrevive** a la selección.
 - Es contraproducente para hallar los óptimos de Pareto y para muestrear el frente de Pareto.

Antecedentes

- Durante la década de 1980 y principios de la década del 1990 hubo muy poco avance en el área de los MOEAs.
- En general los algoritmos propuestos se caracterizaban por utilizar AE monoobjetivos incorporando alguna técnica de agregación para considerar las múltiples funciones objetivo, como por ejemplo: suma ponderada, target vector optimization u ordenamiento lexicográfico.
- En 1989 Goldberg sugirió incorporar dominancia de Pareto en el mecanismo de selección de VEGA.

MOEAs de primera generación

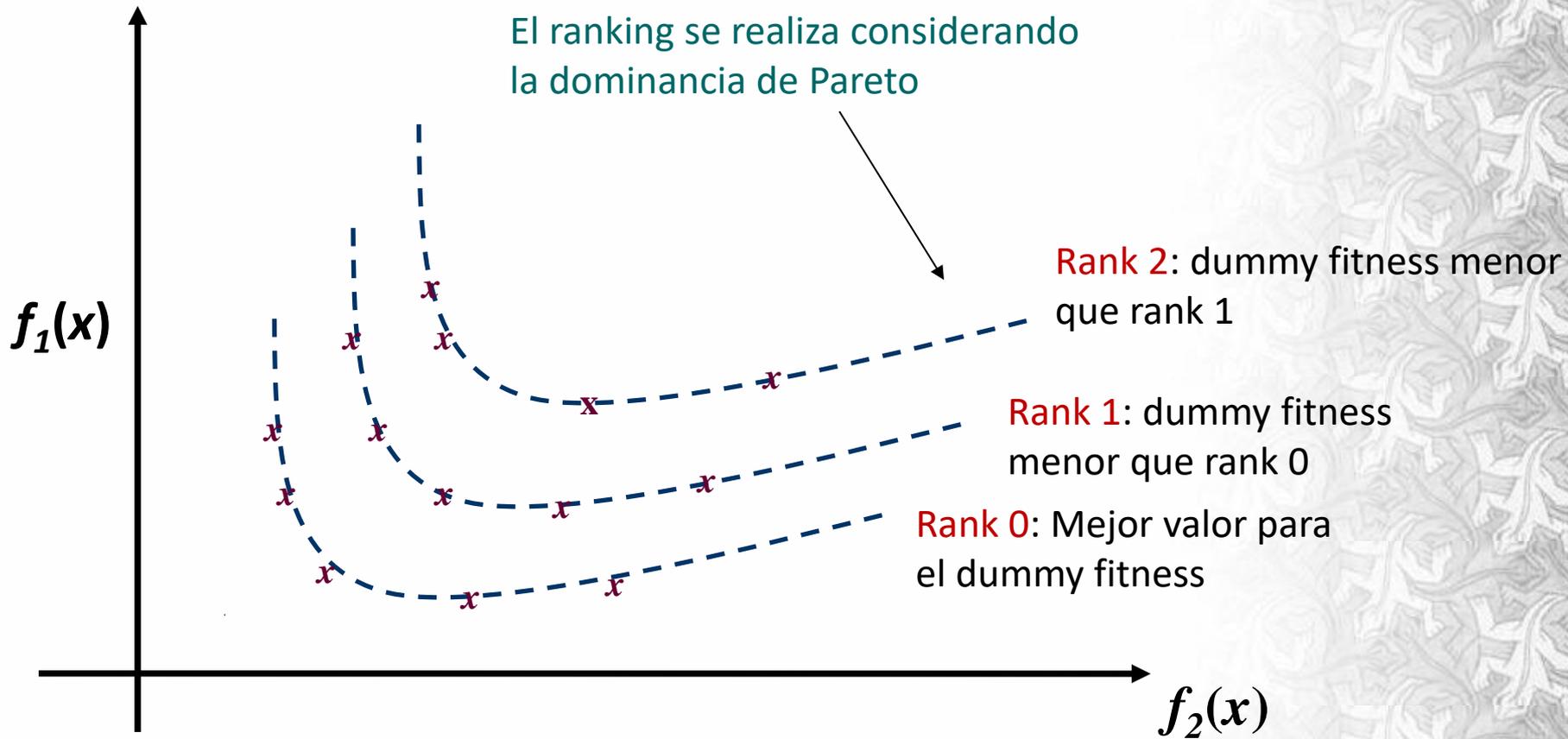
- La primera generación de propuestas de MOEAs transcurrió en el período 1994-1999.
- Se caracterizaban por:
 - la utilización de dominancia de Pareto.
 - el empleo de nichos y sharing para proveer diversidad.
 - en general no utilizaban elitismo.
- Se presentan a continuación tres algoritmos representativos de esta generación:
 - NSGA
 - NPGA
 - MOGA

NSGA

- Shrinivas y Deb propusieron el Nondominated Sorting Genetic Algorithm (NSGA) en 1994.
- Se caracteriza porque la asignación de fitness se realiza por rango de dominancia.
- No trabaja con el valor funcional, sino con un fitness ficticio constante (*dummy fitness*) que se establece a partir de la posición en el ranking de dominancia.
- Se sigue un esquema de asignación de fitness por frentes:
 - Se asigna el fitness para los individuos no dominados de la población.
 - Se remueven los individuos no dominados de la población.
 - Se repite el proceso hasta que no queden más individuos.
- Utiliza la técnica de *sharing* sobre el *dummy fitness*.

MOEAs de primera generación

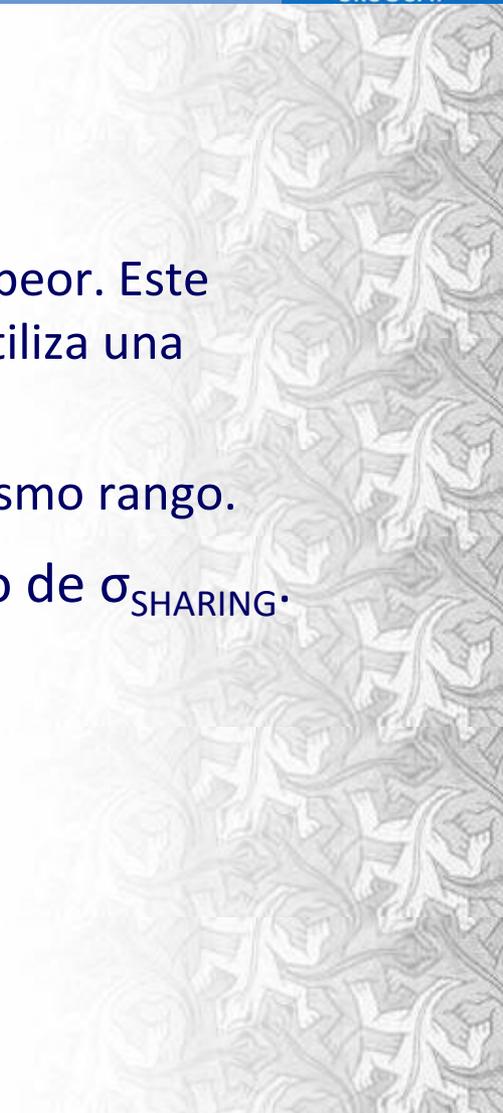
NSGA: asignación de fitness



Mecanismo asignación de fitness de NSGA

- El Niche-Pareto Genetic Algorithm (NPGA) fue propuesto por Horn et al. en 1993.
- La idea del método consiste en utilizar selección mediante torneo basada en la dominancia de Pareto con respecto a un subconjunto de la población.
- En caso de empates se decide por fitness sharing, tanto en el espacio de las funciones objetivo como en el de las variables de decisión.

- El Multi-Objective Genetic Algorithm (MOGA) fue propuesto por Fonseca y Fleming en 1993.
- Implementa una variante de la jerarquización de Pareto propuesta por Goldberg.
- Utiliza rangos de dominancia para asignar el fitness:
 - p_i^t = cantidad de individuos que dominan a al individuo i en la generación t .
 - $\text{rango}(i, t) = p_i^t + 1$
- Los individuos no dominados tendrán un rango de 1.
- Los individuos dominados son penalizados de acuerdo a la densidad de las regiones que lo dominan.



- El procedimiento de asignación de fitness es:
 - ordenar por rango(i,t).
 - asignar fitness mediante interpolación desde el mejor al peor. Este esquema fue propuesto por Goldberg y típicamente se utiliza una interpolación lineal.
 - se promedia el fitness para individuos que están en el mismo rango.
- Utiliza nichos y propone una metodología para el cálculo de σ_{SHARING} .
- Utiliza restricciones al cruzamiento.

- La aplicación de los MOEAs de primera generación abrió las siguientes preguntas en la comunidad científica:
 - ¿es la agregación efectivamente una mala idea?
 - ¿es posible mejorar la eficiencia computacional y los resultados?
 - ¿es posible mantener la diversidad sin utilizar nichos?
- Algunos aspectos importantes que no fueron abordados en la etapa conocida como “primera generación”:
 - la fundamentación teórica de los MOEAs.
 - métricas para la evaluación de resultados.
 - funciones estándar de prueba.

MOEAs de segunda generación

- La segunda generación de propuestas de MOEAs transcurre desde 1999 a la actualidad.
- La característica fundamental consiste en la incorporación de elitismo a los algoritmos.
- Se propusieron básicamente dos mecanismos para la incorporación de elitismo:
 - mediante una selección $(\mu+\lambda)$.
 - mediante la utilización de una población externa (que almacena soluciones no dominadas).

Conceptos

- El esquema que utiliza población externa ha sido exitosamente implementado por varios MOEAs.
- Algunos aspectos a considerar en este enfoque son:
 - La interacción entre la población y la población externa.
 - Los criterios para realizar las inserciones de elementos en la población externa.
 - Las acciones a tomar cuando se completa la población externa.
- Se presentan a continuación seis algoritmos representativos de esta generación:
 - SPEA, SPEA-2, PAES, NSGA-2, NPGA-2 y MICRO-GA

- El Strength Pareto Evolutionary Algorithm (SPEA) fue propuesto por Zitzler y Thiele en 1998.
- Utiliza una población externa de soluciones no dominadas obtenidas previamente.
- Se basa en el concepto de fuerza (*strength*) que tiene un rol similar al rango de dominancia en otros MOEAs. La fuerza de un individuo es proporcional al número de individuos que son dominados por él.
- La asignación de fitness se basa en la medida de “fuerza”.
- Se utiliza clustering para mantener la diversidad.

- El SPEA-2 es una mejora al algoritmo original propuesta por Zitzler y Thiele en 2001.
- Presenta tres diferencias sustanciales con SPEA:
 - la asignación de fitness considera la cantidad de individuos dominados por cada solución (como SPEA), pero además toma en cuenta la cantidad de individuos que la dominan.
 - para mantener diversidad utiliza una técnica de estimación de densidad de individuos vecinos.
 - utiliza un esquema de truncamiento de la población externa que le garantiza la preservación de soluciones de frontera (bordes).

- El algoritmo Pareto Archived Evolution Strategy (PAES) fue propuesto por Knowles en 1999.
- Es uno de los métodos más simples, ya que incorpora el elitismo mediante una selección (1+1).
- Utiliza un archivo histórico de individuos no dominados.
- Realiza una división del espacio de funciones objetivo aplicada en forma recursiva, generando una **grilla** adaptativa.
- Para preservar la diversidad utiliza crowding, evaluando las soluciones en cada uno de los cuadrantes de la grilla.

NSGA-II

- Es una versión mejorada del NSGA propuesta por Deb et al. en el 2000.
- No utiliza población secundaria.
- Incorpora elitismo mediante un esquema de selección ($\mu+\lambda$).
- El chequeo de dominancia fue mejorado para aumentar la performance computacional.
- En lugar de la técnica de sharing del NSGA, utiliza un mecanismo de crowding que no requiere parámetros.
- En general funciona mejor para representación real que para representación binaria.
- Es uno de los algoritmos que en la práctica ha mostrado mejores resultados.

- Es una versión mejorada del NPGA propuesta por Erickson et al. en 2001.
- Utiliza jerarquización de Pareto pero mantiene la selección mediante torneo.
- Incorpora elitismo mediante un esquema de selección ($\mu+\lambda$).
- Utiliza sharing mediante el conteo en los nichos de individuos de la siguiente generación (*partial niching*). Esta propuesta se basa en la idea de *continuously updated fitness sharing* propuesta por Oei en 1991.

Micro GA

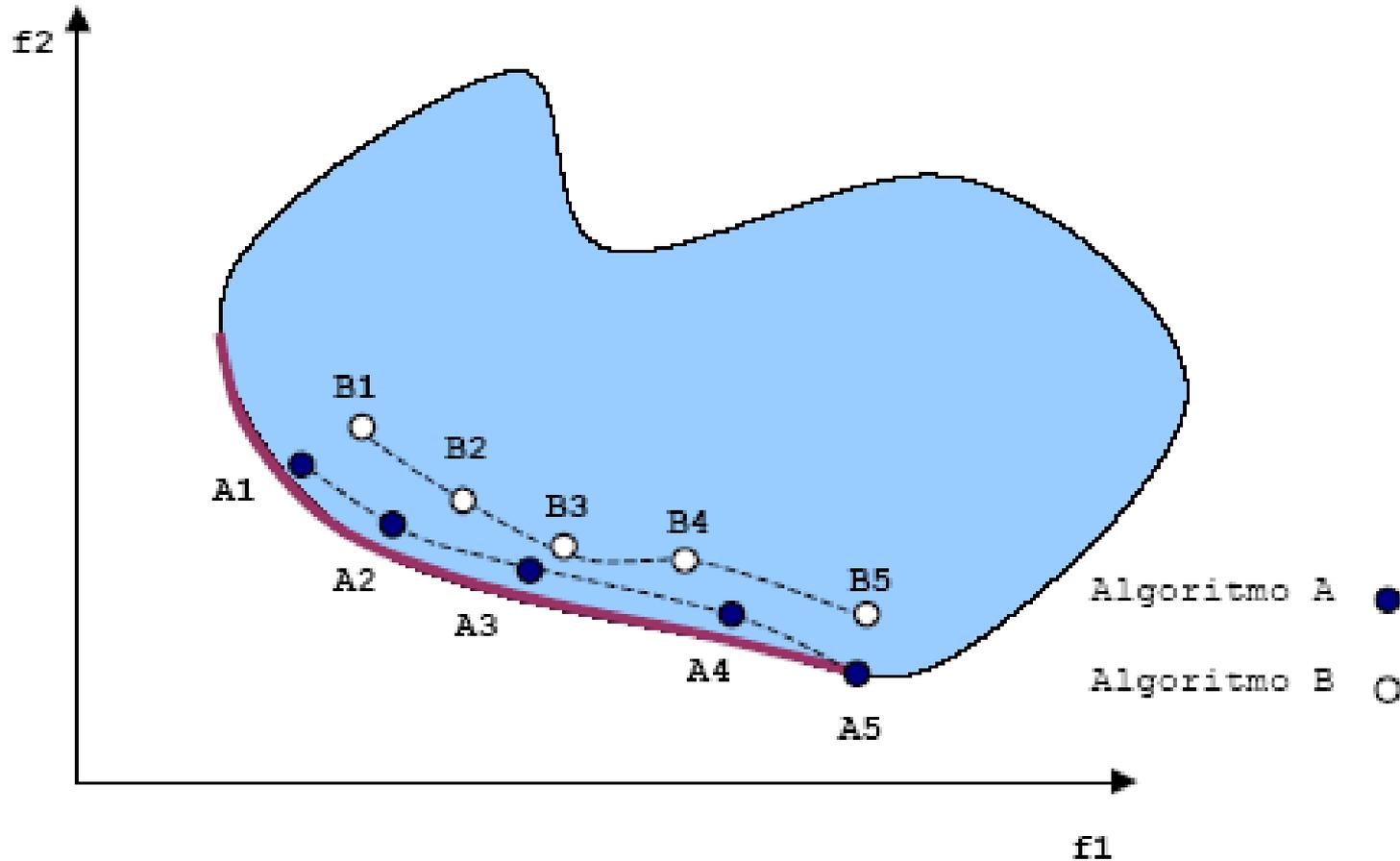
- El Micro Genetic Algorithm for Multiobjective Optimization (Micro GA) fue propuesto por Toscano y Coello en 2001.
- Se caracteriza por utilizar una población muy pequeña e incorporar un mecanismo de reinicialización cuando se alcanza la convergencia.
- Utiliza dos poblaciones, una reemplazable y una no reemplazable con la cual se maneja la dominancia y la diversidad.
- La incorporación del elitismo se basa en la idea de mantener soluciones no dominadas pero con una memoria que se actualiza a plazos.
- Utiliza la grilla adaptativa de PAES.

- En la segunda generación se realizaron avances en la fundamentación teórica de estas técnicas.
- Se realizaron múltiples propuestas de problemas de prueba estándar, dentro de los que destacan:
 - ZDT (Zitzler, Deb, Thiele): proponen un mecanismo para generar problemas de prueba. Diseñan 6 problemas que fueron siguiendo su metodología (ZDT1 a ZDT6).
 - Otros problemas: Ozyczka, Golinski, Viennet.
- Se avanzó en propuestas para mejorar el desempeño computacional (MOEAs paralelos).
- Se propusieron métricas para la evaluación objetiva de los resultados obtenidos (se presentarán a continuación).

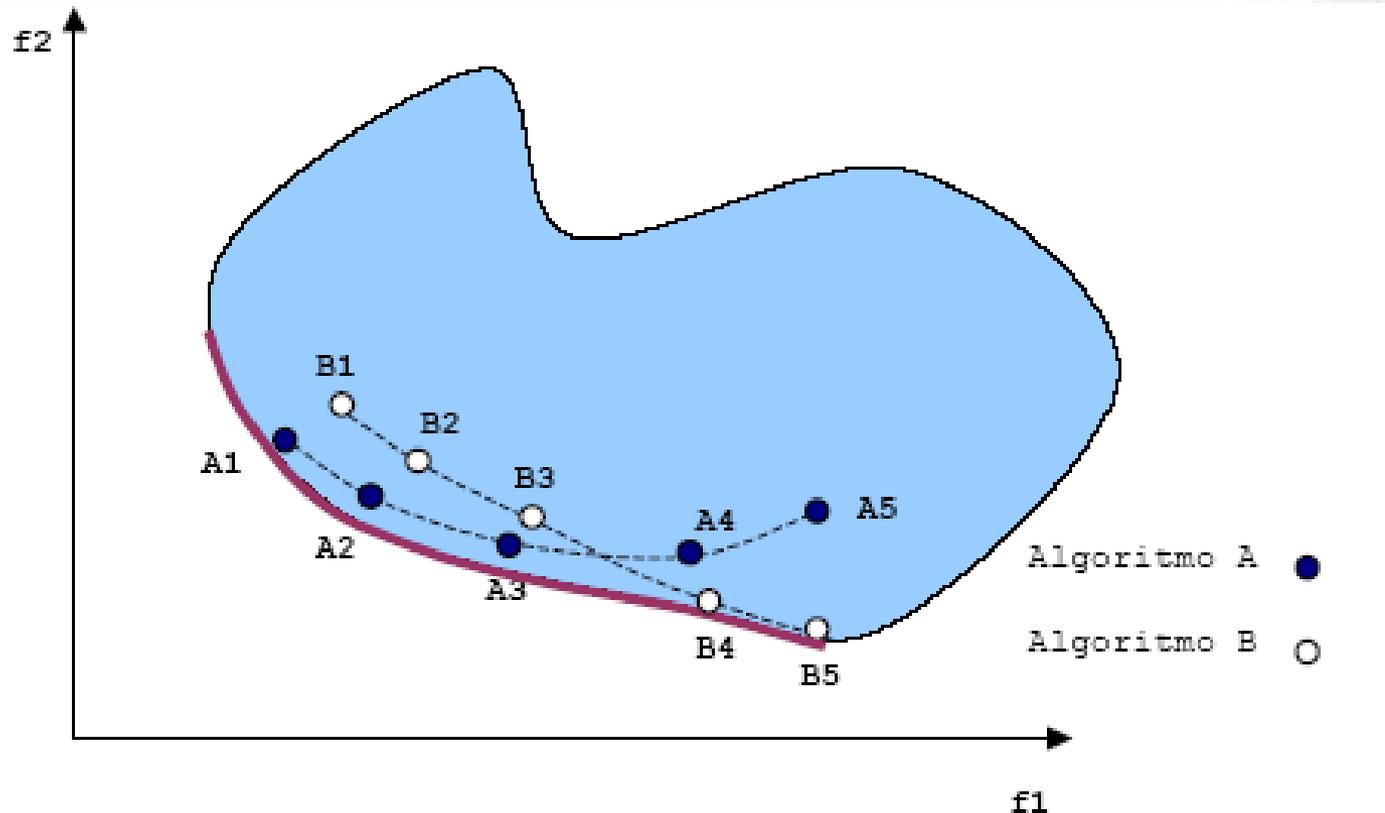
Métricas para evaluar MOEAs

- Es necesario utilizar métricas bien definidas para evaluar el desempeño de los algoritmos.
- Las métricas deben evaluar:
 - la calidad de los resultados (cercanía al frente de Pareto).
 - la diversidad de las soluciones encontradas.
 - la cantidad de elementos del conjunto de óptimos de Pareto.
 - la eficiencia computacional del algoritmo.
- La comparación en forma directa los resultados obtenidos no siempre es posible.

Métricas para evaluar MOEAs



Métricas para evaluar MOEAs



Los resultados de los algoritmos A y B son difíciles de comparar.

Métricas para evaluar MOEAs

- A continuación se presentan algunas métricas comúnmente utilizadas para la evaluación de los resultados obtenidos por los MOEAs.
- Número de puntos no dominados
 - Evalúa la cantidad efectiva de puntos no dominados que encuentra el algoritmo.
- Tasa de error
 - Indica el porcentaje de soluciones que no son miembros del frente de Pareto.
 - Requiere conocer el frente de Pareto, o por lo menos tener un procedimiento para determinar si un punto pertenece a él o no.

Métricas para evaluar MOEAs

- Tasa de error

$$ER = \frac{\sum_{i=1}^n e_i}{n}$$

- siendo n el número de puntos no dominados de los resultados y

$$e_i = \begin{cases} 1 & \text{si la solución} \notin FP \\ 0 & \text{en otro caso} \end{cases}$$

- Distancia generacional

- Busca estimar qué tan lejos están los elementos no dominados obtenidos como resultado de los elementos del verdadero frente de Pareto.

- Distancia generacional

$$GD = \frac{\left(\sum_{i=1}^n d_i^p \right)^{1/p}}{n}$$

siendo n el número de puntos no dominados obtenidos como resultado y d_i^p la distancia entre i -ésima la solución no dominada obtenida y la solución más próxima del frente de Pareto real.

p es la dimensión del espacio y está dada por la cantidad de funciones objetivo.

- Distancia generacional

- Por ejemplo, para $p = 2$ d_i^p corresponde a la distancia Euclidea:

$$d_i^p = \min_j \sqrt{\sum_{k=1}^n (f_k^i - f_k^{*j})^2}$$

siendo f_k^i el valor de la k -ésima función objetivo evaluada en la solución no dominada i -ésima, y f_k^{*j} el valor de la k -ésima función objetivo evaluada en el punto más cercano del frente de Pareto al punto i (el punto j -ésimo).

- Spacing

- Evalúa la distribución de los puntos no dominados hallados como solución del problema mediante el MOEA.

$$SP = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\bar{d} - d_i)^2}$$

- El término d_i mide la distancia en el espacio de las funciones objetivo entre la solución i -ésima y su vecino más próximo (la solución j -ésima) en el conjunto de soluciones no dominadas obtenidas por el algoritmo

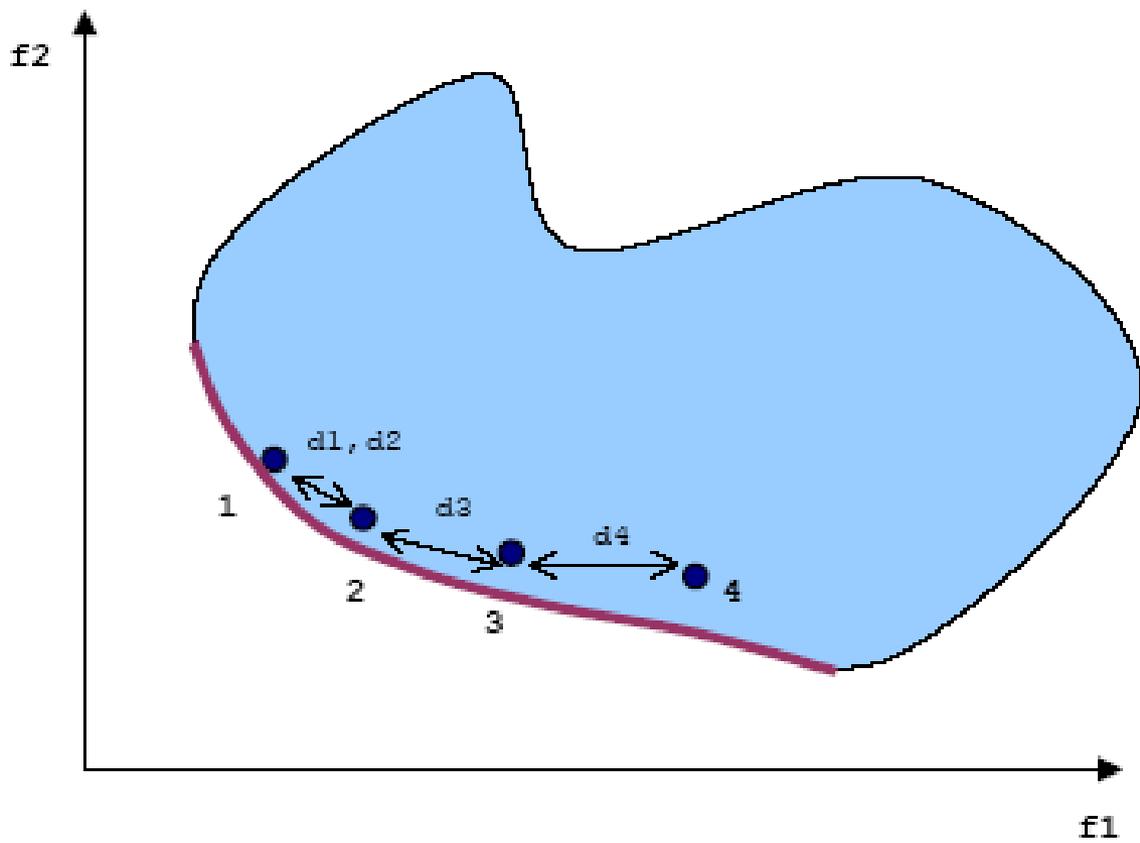
$$d_i = \min \left(\left| f_1^i(x) - f_1^j(x) \right| + \left| f_2^i(x) - f_2^j(x) \right| + \dots + \left| f_M^i(x) - f_M^j(x) \right| \right)$$

$$i = 1, \dots, n$$

- Spacing

- El valor \bar{d} corresponde al promedio de los d_i .
- Un valor ideal de cero en esta métrica indica que todos los puntos no dominados obtenidos por el MOEA como solución al problema, están equiespaciados.

- Spacing



Distancias involucradas en el cálculo de la métrica spacing.

- Spread

- Utiliza como información adicional a spacing la distancia a los “extremos” del frente de Pareto, aquellos puntos con menores valores en cada una de las funciones objetivo, para tener una medida más precisa de la cobertura del frente.
- Tomando en cuenta los “extremos” del FP se evita obtener valores aceptables para la medida de distribución en casos en los cuales no se está muestreando adecuadamente el frente de Pareto.

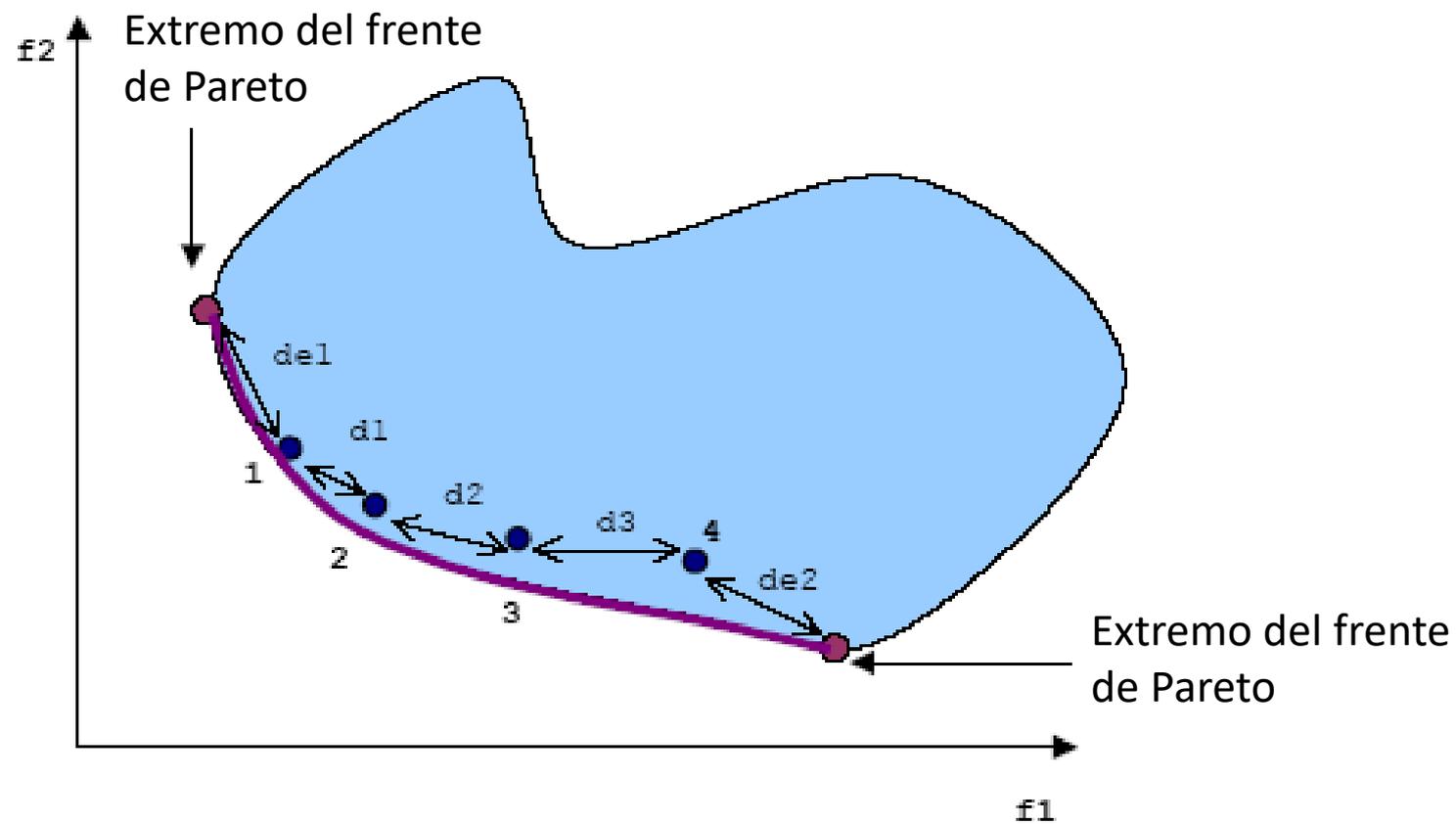
$$Spread = \frac{\sum_{k=1}^M d_k^e + \sum_{k=1}^n \left| \bar{d} - d_i \right|}{\sum_{k=1}^M d_k^e + n \cdot \bar{d}}$$

- Spread

- El término d_k^e mide la distancia entre el punto “extremo” del frente de Pareto, tomando en cuenta la función k -ésima y la solución no dominada más cercana.
- El término d_i mide la distancia en el espacio de las funciones objetivo entre la solución i -ésima y su vecino más próximo (la solución j -ésima) en el conjunto de soluciones no dominadas obtenidas por el algoritmo.
- El valor \bar{d} corresponde al promedio de d_i .

Métricas para evaluar MOEAs

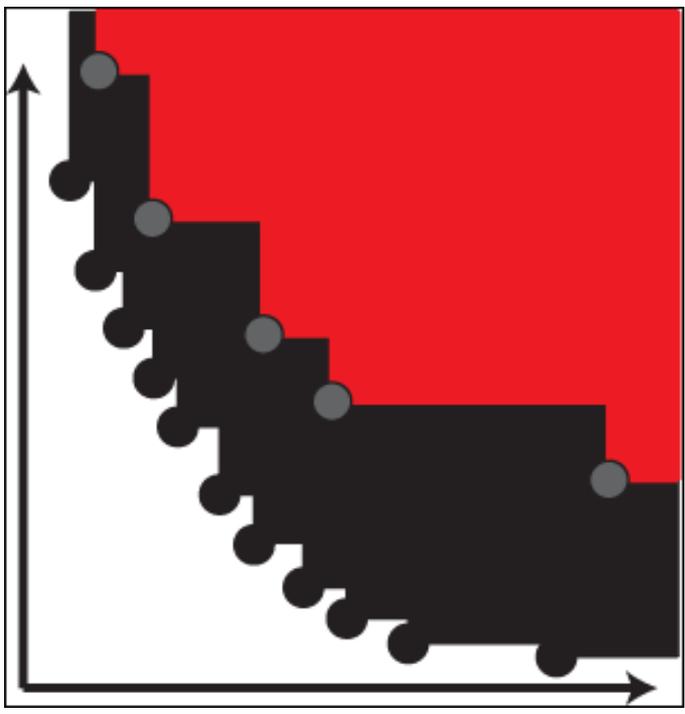
- Spread



Distancias involucradas en el cálculo de la métrica spread.

Métricas para evaluar MOEAs

- Hipervolumen relativo (RHV)
 - Cociente entre el volumen (en el espacio de funciones objetivo) cubierto por el frente de Pareto calculado por el algoritmo y el volumen cubierto por el frente de Pareto real.



- algoritmo A
- algoritmo B
- RHV algoritmo A
- RHV algoritmo B

Métricas para evaluar MOEAs

- Empirical attainment surface (EAS)

- Estimador estadístico de la métrica *attainment surface*, que estima para cada vector en el espacio de funciones objetivo la probabilidad de ser dominado por el frente de Pareto calculado en una ejecución individual del MOEA evaluado.
- EAS no depende del frente de Pareto real, que puede ser desconocido para el problema que se resuelve
- Ejemplo: 50%-EAS (Knowles, 2005), estima la EAS obtenida en al menos 50% de las ejecuciones realizadas. Es el análogo a la media (mediana) en optimización de objetivo único.

