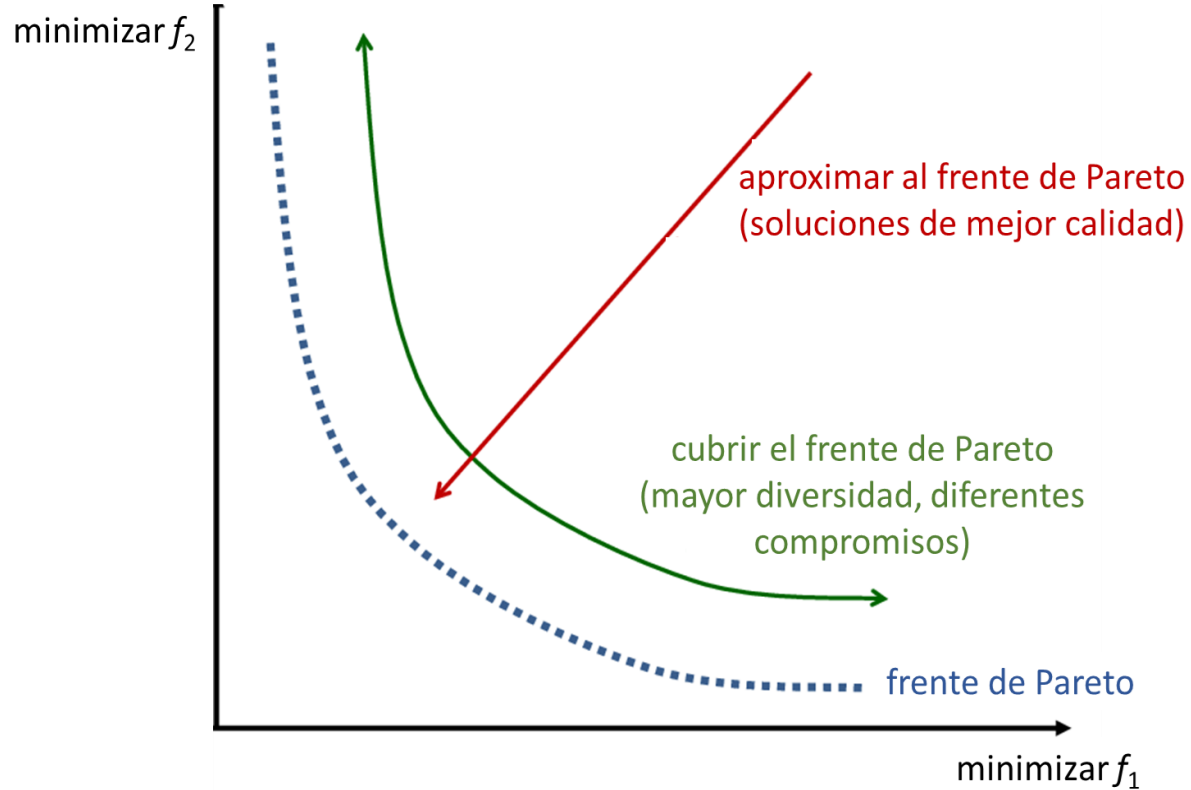


Conceptos y herramientas para la resolución de problemas de optimización multiobjetivo

Métricas para la evaluación de algoritmos para optimización multiobjetivo

SERGIO NESMACHNOW (Universidad de la República, Uruguay)
DIEGO ROSSIT (Universidad Nacional del Sur, Argentina)

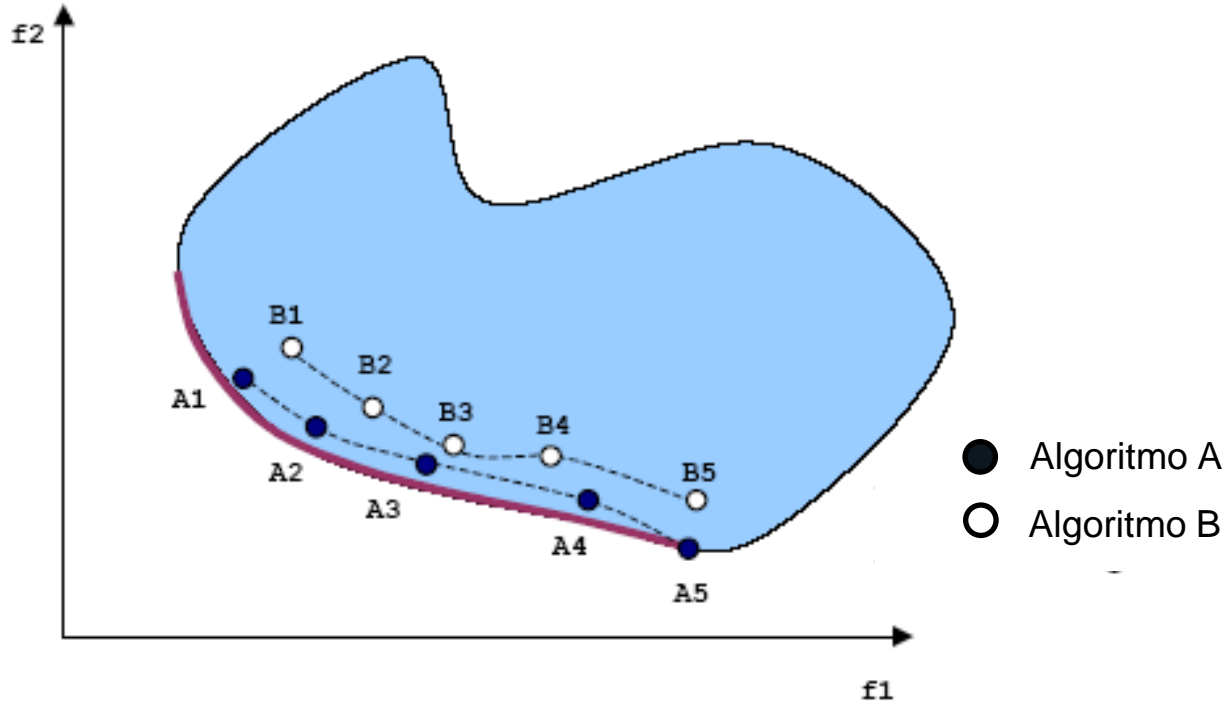
Propósitos de un algoritmo multiobjetivo



Métricas para evaluar algoritmos multiobjetivo

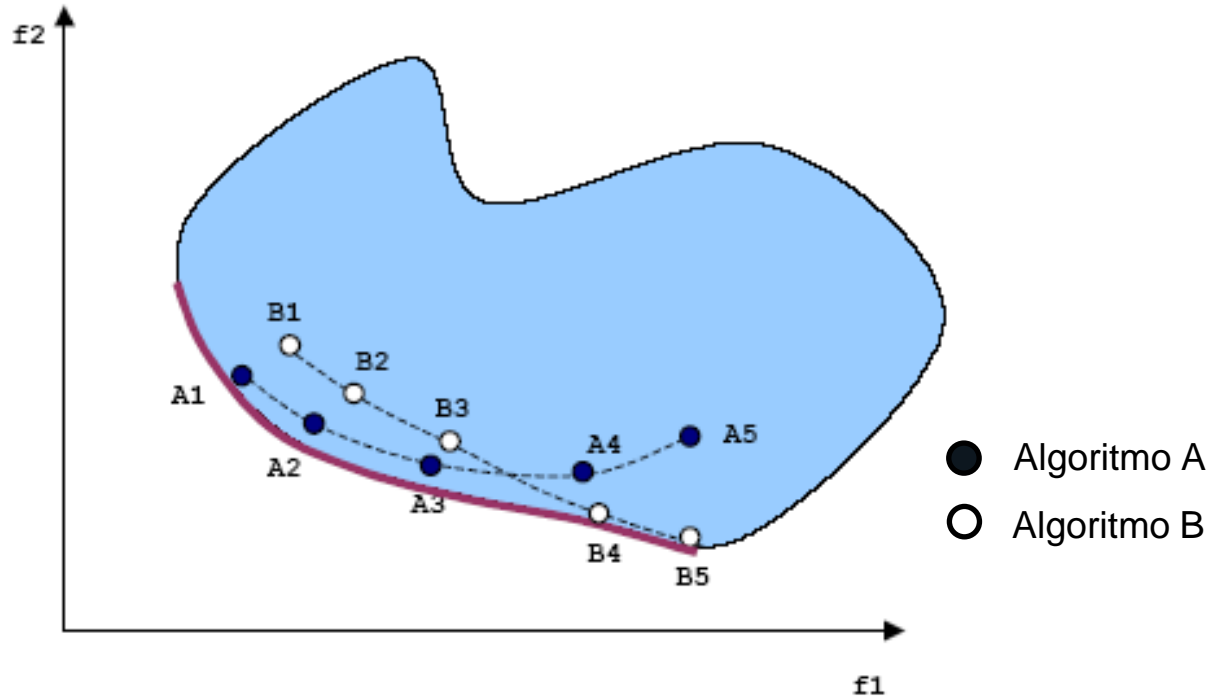
- Es necesario utilizar métricas bien definidas para evaluar el desempeño de los algoritmos.
- Las métricas deben evaluar:
 - la calidad de los resultados (cercanía al frente de Pareto).
 - la diversidad de las soluciones encontradas.
 - la cantidad de elementos del conjunto de óptimos de Pareto.
 -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

Métricas para evaluar algoritmos multiobjetivo



- El algoritmo A obtiene mejores resultados que el algoritmo B.

Métricas para evaluar algoritmos multiobjetivo



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

Métricas para evaluar algoritmos multiobjetivo

- Propiedades deseables de las métricas
- Monotonía (débil): Dado un conjunto de soluciones no dominadas, agregar un punto no dominado mejora (no degrada) su evaluación.
- Relatividad (débil): La evaluación del verdadero frente de Pareto es óptima de manera (no) única, i.e., todos los otros conjuntos de soluciones no dominadas tienen una evaluación estrictamente inferior (no superior)
- Independencia: no dependen de propiedades ni necesitan valores o conjuntos de referencia

Métricas para evaluar algoritmos multiobjetivo

- Número de puntos no dominados
 - Evalúa la cantidad efectiva de puntos no dominados que encuentra el algoritmo.
 - No determina si los puntos pertenecen al frente de Pareto.
 - Mayores valores de la métrica corresponden a mejores resultados.
 - Peor valor: 1.
 - Mejor valor: n , número de soluciones con las que trabaja el algoritmo.

Métricas para evaluar algoritmos multiobjetivo

- Tasa de error (error ratio)
 - Indica el porcentaje de soluciones halladas que no son miembros del frente de Pareto.

- Tasa de error $ER = \frac{\sum_{i=1}^n e_i}{n}$

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

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

- 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 algoritmos multiobjetivo

- La métrica no es monótona y puede dar resultados contradictorios con la evaluación de calidad.
- Si un algoritmo A encuentra dos soluciones no dominadas, de las cuales una está en el verdadero frente de Pareto y la otra está lejos del frente de Pareto, su error ratio es 0.5.
- Si el algoritmo B obtiene 100 soluciones, de las cuales 99 están muy cerca del verdadero frente de Pareto (y posiblemente distribuidas uniformemente), pero solo una de ellas está en el verdadero frente de Pareto, el error ratio es 0.99.
- Claramente, el segundo conjunto de soluciones es mejor, pero el primero tiene mucho mejor valor de tasa de error.

Métricas para evaluar algoritmos multiobjetivo

- Tasa de error: no es una métrica monótona.
- Dado un conjunto no dominado A conteniendo uno o más puntos Pareto óptimos, agregar más puntos no dominados que no sean Pareto óptimos causa que el valor de la tasa de error sea peor.
- Viola también la relatividad, puesto que cualquier subconjunto no vacío del conjunto de óptimos de Pareto tiene un valor óptimo de tasa de error (a pesar de no ser el frente de Pareto real, ni siquiera muestrearlo apropiadamente).
- El frente de Pareto es evaluado como no peor que cualquier otro conjunto.

Métricas para evaluar algoritmos multiobjetivo

- Tasa de error
 - Menores valores de la métrica corresponden a mejores resultados.
 - Peor valor: 1, mejor valor: 0
- Ventajas:
 - Es fácil de entender y fácil de calcular.
 - Es independiente de la escala de las funciones objetivo.
 - Puede usarse como un mecanismo rápido (pero crudo) de evaluar el progreso hacia el verdadero frente de Pareto.
- Desventajas:
 - Requiere conocer el verdadero frente de Pareto.
 - No es monótona ni relativa

Métricas para evaluar algoritmos multiobjetivo

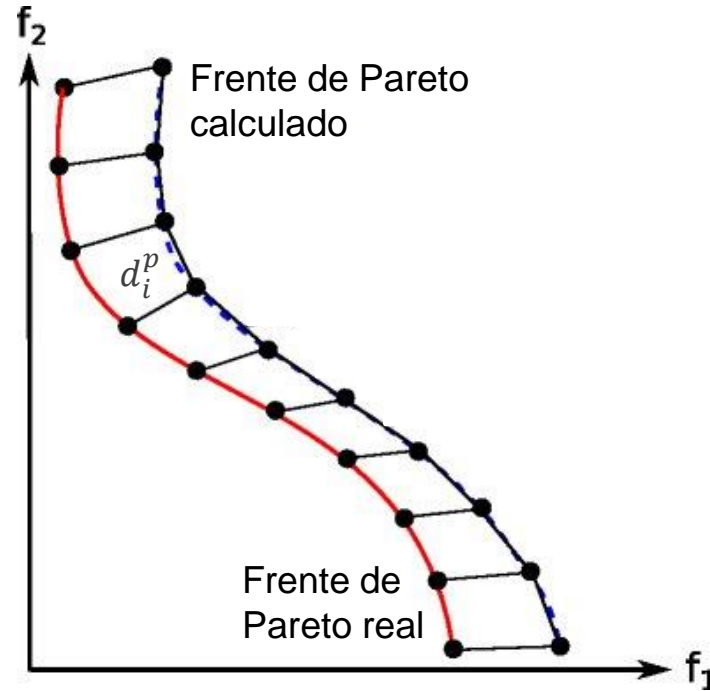
- Distancia generacional
 - Estima qué tan lejos están las soluciones no dominadas obtenidas como resultado, de los elementos del verdadero frente de Pareto.

$$GD = \frac{\sqrt[p]{\sum_i^n d_i^p}}{n}$$

- n es 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.

Métricas para evaluar algoritmos multiobjetivo

- Distancia generacional



Métricas para evaluar algoritmos multiobjetivo

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

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

- f_k^i es 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).

Métricas para evaluar algoritmos multiobjetivo

- Distancia generacional
- Es una medida que usa el verdadero frente de Pareto como referencia. Induce un ordenamiento total y es no cardinal.
- Puede ser engañosa, dado que son comunes los ejemplos en los que el indicador produce un resultado inapropiado para evaluar resultados.

Métricas para evaluar algoritmos multiobjetivo

- Distancia generacional
 - Menores valores de la métrica corresponden a mejores resultados.
 - Mejor valor: 0, peor valor: no determinado.
- Ventajas:
 - Es sencilla de calcular (computacionalmente)
- Desventajas:
 - Requiere conocer el verdadero frente de Pareto.
 - No es directamente aplicable para evaluar conjuntos no dominados que cambian de cardinalidad.
 - Depende de las escalas de los objetivos, requiere información adicional sobre las preferencias del tomador de decisiones.

Métricas para evaluar algoritmos multiobjetivo

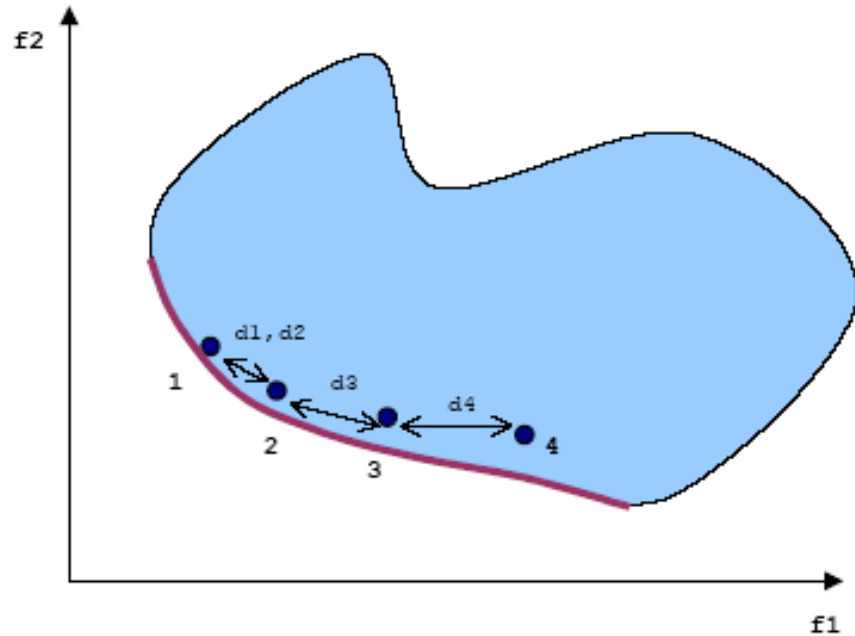
- Spacing
 - Evalúa la distribución de los puntos no dominados hallados como solución del problema mediante el algoritmo aplicado.

$$\Delta = \sum_1^{|FP|} |d_i - \bar{d}|$$

- 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} es promedio de los d_i .
- Un valor ideal de cero para spacing indica que todas las soluciones no dominadas obtenidas por el algoritmo están equiespaciados.

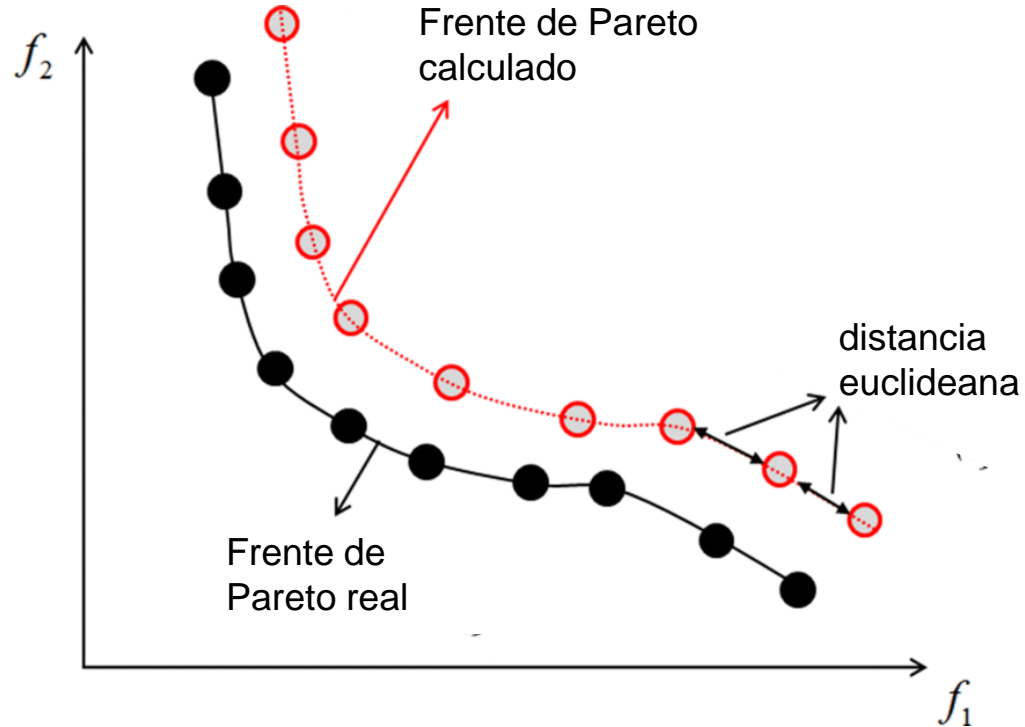
Métricas para evaluar algoritmos multiobjetivo

- Distancias involucradas en el cálculo de la métrica spacing.



Métricas para evaluar algoritmos multiobjetivo

- Spacing

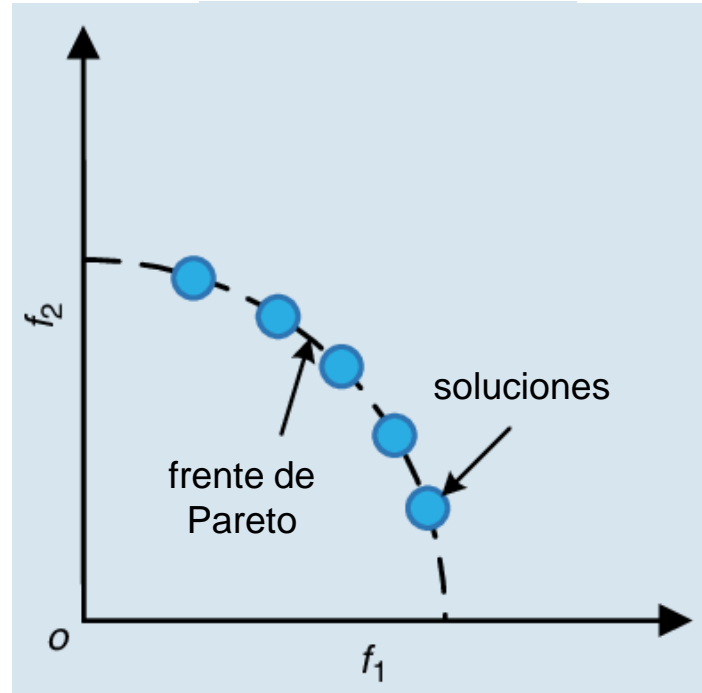


Métricas para evaluar algoritmos multiobjetivo

- Spacing
 - Menores valores de la métrica corresponden a mejores resultados.
 - Mejor valor: 0, peor valor: no determinado.
- Ventajas:
 - Es independiente, no requiere conocer el verdadero frente de Pareto.
 - Es sencilla de calcular (computacionalmente)
- Desventajas:
 - No es monótona ni relativa (el verdadero frente de Pareto puede ser no uniforme).
 - No considera los extremos del frente de Pareto.
 - Requiere normalización

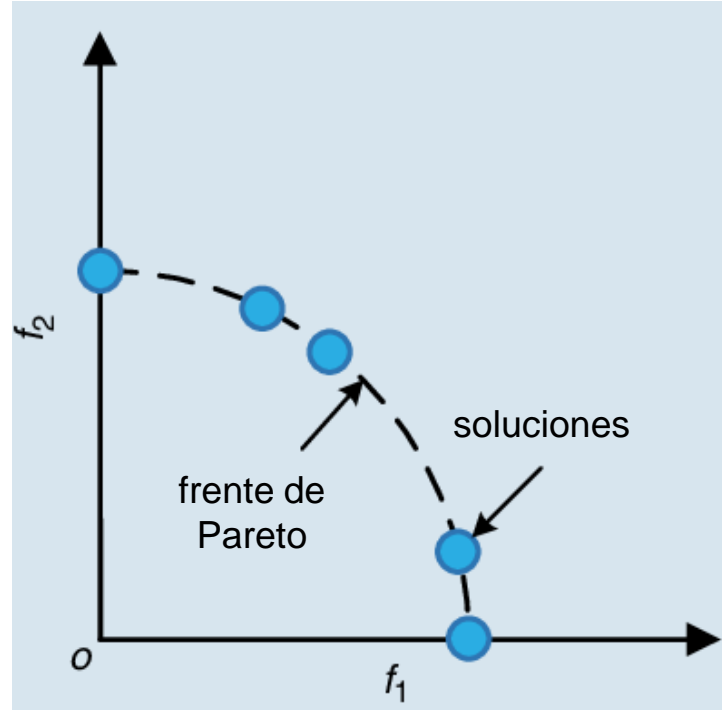
Métricas para evaluar algoritmos multiobjetivo

- Buena equidistribución pero bajo cubrimiento



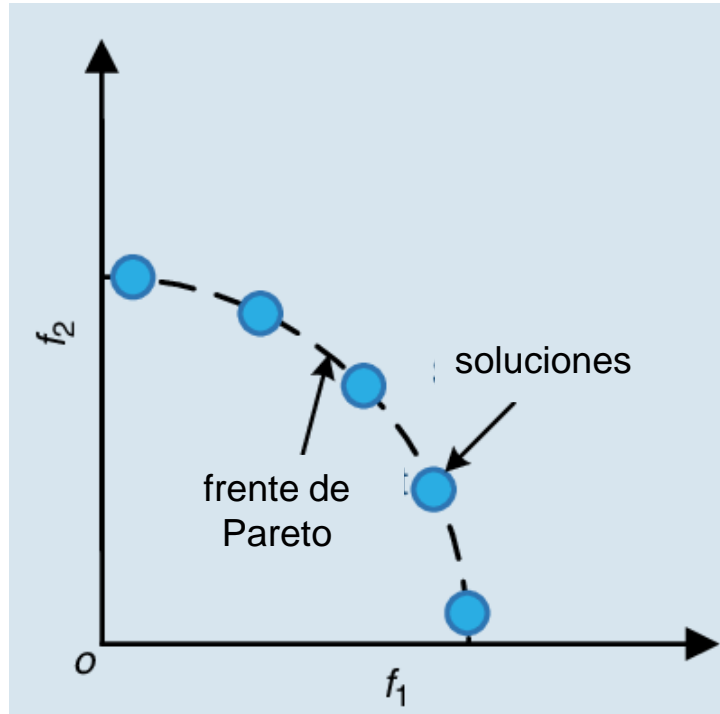
Métricas para evaluar algoritmos multiobjetivo

- Buen cubrimiento, pero baja equidistribución



Métricas para evaluar algoritmos multiobjetivo

- Buen cubrimiento y buena equidistribución



Métricas para evaluar algoritmos multiobjetivo

- Spread
 - Utiliza como información adicional a spacing la distancia a los extremos del frente de Pareto [los puntos con menores valores en cada función objetivo], para tener una medida más precisa de la cobertura del frente.

$$spread = \frac{\sum_{k=1}^M d_k^e + \sum_{k=1}^{|FP|} |\bar{d} - d_i|}{\sum_{k=1}^M d_k^e + n\bar{d}}$$

- Tomando en cuenta los extremos del frente de Pareto se evita obtener valores aceptables para la medida de distribución en casos en los que no se está muestreando adecuadamente el frente de Pareto (buena equidistribución pero bajo cubrimiento).

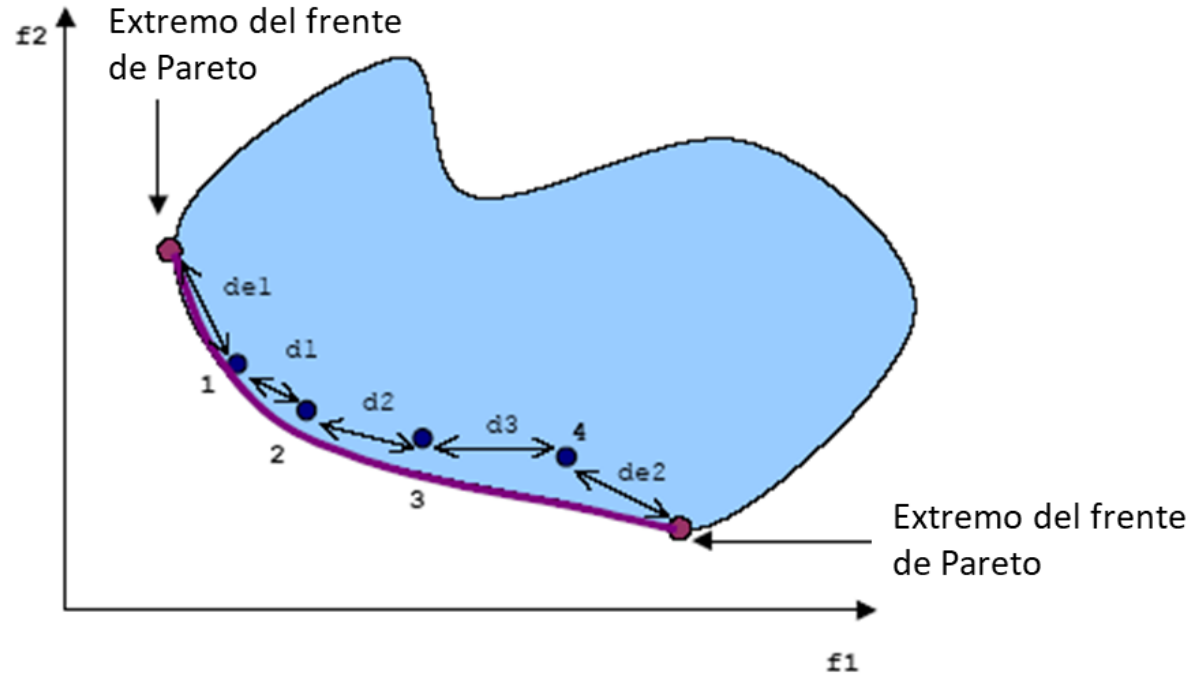
Métricas para evaluar algoritmos multiobjetivo

$$spread = \frac{\sum_{k=1}^M d_k^e + \sum_{k=1}^{|FP|} |\bar{d} - d_i|}{\sum_{k=1}^M d_k^e + n\bar{d}}$$

- 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 los d_i .

Métricas para evaluar algoritmos multiobjetivo

- Distancias involucradas en el cálculo de la métrica spread.

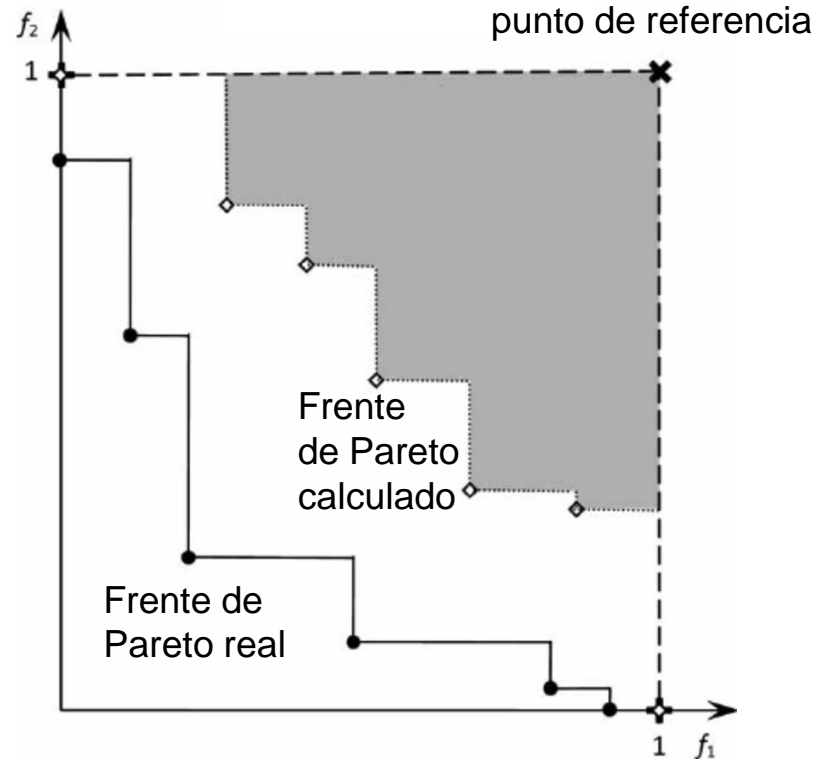


Métricas para evaluar algoritmos multiobjetivo

- Spread
 - Menores valores de la métrica corresponden a mejores resultados.
 - Mejor valor: 0, peor valor: no determinado.
- Ventajas:
 - Es independiente, no requiere conocer el verdadero frente de Pareto (solo los extremos).
 - Es sencilla de calcular (computacionalmente).
- Desventajas:
 - No es monótona ni relativa (el verdadero frente de Pareto puede ser no uniforme).
 - Requiere normalización.

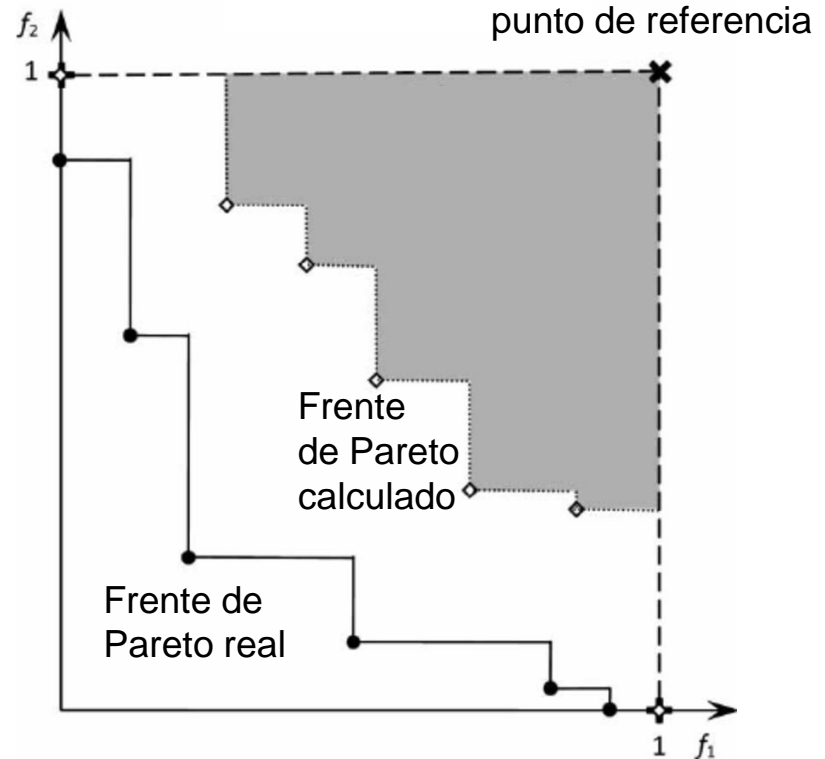
Métricas para evaluar algoritmos multiobjetivo

- Hipervolumen: volumen (en el espacio de funciones objetivo) cubierto por el frente de Pareto calculado por el algoritmo. Puede tomarse como referencia el volumen cubierto por el frente de Pareto real.
- Se considera un punto de referencia para calcular el volumen



Métricas para evaluar algoritmos multiobjetivo

- El punto de referencia se establece de tal forma que todos los conjuntos no dominados sean evaluados como positivos.
- Normalmente, se usa el vector nadir o el vector con los peores valores de las funciones objetivo para definir el punto de referencia

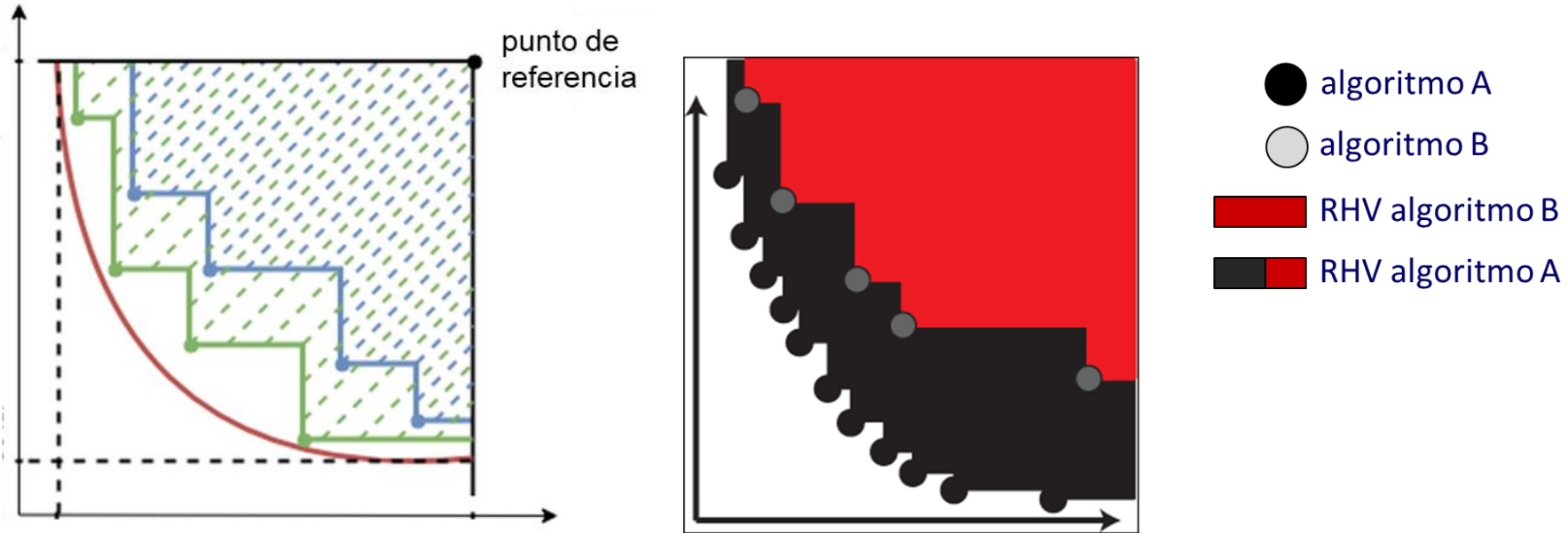


Métricas para evaluar algoritmos multiobjetivo

- Hipervolumen
 - Mayores valores de la métrica corresponden a mejores resultados.
 - Mejor valor: no determinado, peor valor: no determinado.
- Ventajas:
 - Es sencilla de calcular (computacionalmente).
 - Intuitiva e independiente (solo depende del punto de referencia).
- Desventajas:
 - Tiene un costo computacional elevado $O(n^{k+1})$, demanda mucho esfuerzo de cálculo para problemas con muchas funciones objetivo.
 - Multiplica objetivos entre sí, puede sesgar valores si las escalas son muy diferentes.

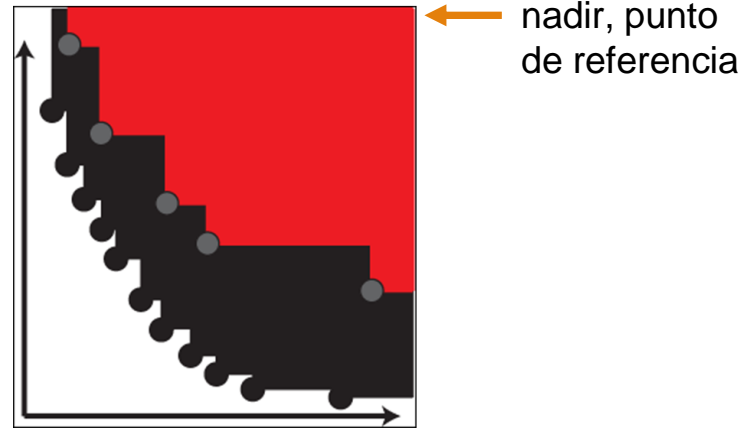
Métricas para evaluar algoritmos multiobjetivo

- Hipervolumen relativo (RHV): cociente entre el hipervolumen de dos o más algoritmos.



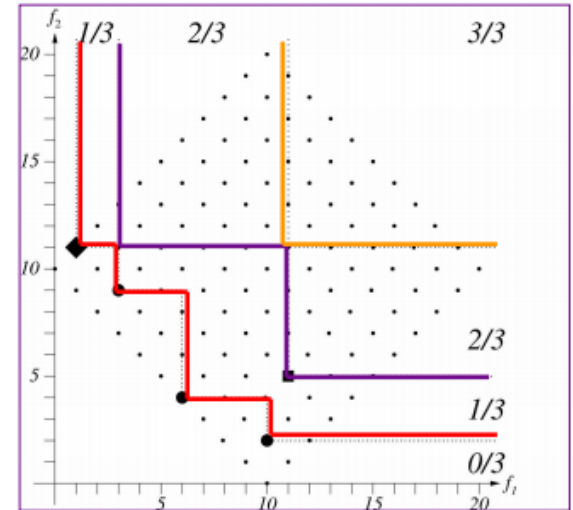
Métricas para evaluar algoritmos multiobjetivo

- Hipervolumen relativo (RHV)
- Es una métrica mixta que evalúa calidad y dispersión de soluciones.
- Depende de un punto de referencia que se utiliza para calcular el volumen
- Punto nadir, punto con los peores valores de las funciones objetivo del frente de Pareto



Métricas para evaluar algoritmos multiobjetivo

- 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 algoritmo evaluado.
- EAS no depende del frente de Pareto real, que puede ser desconocido para el problema.
- Ejemplo: 50%-EAS, estima la EAS obtenida en al menos 50% de las ejecuciones realizadas. Es el análogo a la media (mediana) en optimización de objetivo único.



Métricas para evaluar algoritmos multiobjetivo

- En la práctica:
- Las métricas simples son útiles para analizar la evolución de la calidad de resultados en etapas intermedias del algoritmo.
- Las métricas más comprehensivas son necesarias para una correcta evaluación de los resultados finales.
- Deben considerarse ambos propósitos al resolver un problema multiobjetivo (y también tener en cuenta la eficiencia computacional)
 - Proximidad al frente de Pareto.
 - Diversidad de las soluciones y cobertura del frente de Pareto.
- Puede utilizarse una métrica de proximidad (por ejemplo, distancia generacional) y una métrica de cobertura (spacing/spread) o utilizar una métrica combinada (hipervolumen)