Optimización robusta de carteras de inversión usando muestreo estocástico y metaheurísticas

Francisco Luna[1], David Quintana[2], Sandra Rodríguez[3], and Pedro Isasi[2]

[1] Dpto. de Lenguajes y Ciencias de la Computación, Universidad de Málaga
flv@lcc.uma.es
[2] Dpto. de Informática, Universidad Carlos III de Madrid
{david.quintana,pedro.isasi}@uc3m.es
[3] CEA Saclay, DRT/LIST/DM2I/LADIS
sandra.garciarodriguez@cea.fr

Resumen La optimización de carteras de inversión presenta un gran número de dificultades, siendo una de las más importantes la gestión de la incertidumbre que surge en diferentes aspectos del proceso. Este trabajo propone matizar esta última mejorando la robustez de las soluciones mediante el uso de ϵ-vecindarios combinados con un mecanismo de remuestreo basado en marcas temporales. La aproximación se ha incorporado a cuatro metaheurísticas multi-objetivo del estado del arte, que se han evaluado sobre un histórico bastante amplio de activos. Los resultados han mostrado una mejora significativa de la fiabilidad de la frontera eficiente estimada.

Keywords: Gestión de carteras de inversión, robustez, optimización multi-objetivo

1. Introducción

El problema de la optimización de carteras de inversión consiste en identificar las mejores formas de distribuir capital entre un conjunto alternativas de inversión. Entre los diferentes aspectos de este problema clásico en finanzas, hay uno que está recibiendo especial atención en la actualidad, que es el de la robustez. La literatura académica dedicada a la optimización de carteras es abundante, y está basada, principalmente, en el trabajo de Markowitz [14,15], que abordaba este problema como una tarea de optimización multiobjetivo donde los inversores debían ponderar dos objetivos contrapuestos: la minimización del riesgo y la maximización de la rentabilidad. La solución al problema no es única, si no un conjunto de soluciones de compromiso entre ambos objetivos que definen el frente de Pareto (conocido en este contexto como la frontera eficiente). Dado este frente, el inversor deberá escoger aquella solución que más se adecue a su perfil.

La formulación más básica del problema se puede resolver eficientemente utilizando Programación Cuadrática. No obstante, a medida que se incorporan restricciones propias del mundo real [1], la utilización de aproximaciones algorítmicas más flexibles es indispensable. De entre estas, las metaheurísticas en general, y la computación evolutiva en particular, han sido una de las más utilizadas [16]. Dentro de este contexto, el desarrollo reciente experimentado por los algoritmos evolutivos multi-objetivo (MOEAs por sus siglas en inglés) ha captado la atención de multitud de investigadores. Algunos ejemplos de esto serían [11], donde se presentan mejoras en el operador de mutación usado en los MOEAs; en [12], los autores proponen un nuevo MOEA basado en aprendizaje MODEwAwL y se compara con cuatro algoritmos del estado del arte; [2] introduce un nuevo algoritmo metaheurístico basado en enjambres de partículas (MOPSO por sus siglas en inglés) para abordar una formulación basada en programación lineal entera del problema; o, [21] en el que se presenta un MOEA que usa lógica difusa para mantener la diversidad entre las carteras. Un artículo de revisión relevante sobre el uso de MOEAs en la gestión carteras sería [16].

A pesar del número de trabajos existentes sobre optimización de carteras, aún quedan muchas líneas prometedoras. Entre ellas, este trabajo se centra en dotar de robustez a las soluciones proporcionadas por las metaheurísticas, entendiendo por robustez la fiabilidad de las carteras en escenarios con incertidumbre [20,9]. Dentro del marco de trabajo de Markowitz, la mayor fuente de incertidumbre tiene que ver con la calidad de las estimaciones de las ganancias de los activos y de sus matrices de varianzas-covarianzas, que hacen que el comportamiento previsto de las soluciones obtenidas por los algoritmos quede lejos del real. Un mecanismo posible para abordar este problema es el remuestreo [17], que se basa en la idea de exponer soluciones intermedias a distintos escenarios para obtener carteras que no estén tan especializadas en una previsión única de las ganancias y las matrices de varianzas-covarianzas. Esta técnica de remuestreo, junto con un mecanismo de marcas de tiempo, son la base para manejar este tipo de incertidumbre y se han aplicado ya en [6,7].

El objetivo de este trabajo va un paso más allá, y trata de abordar el problema que se presenta cuando los gestores emiten órdenes de compra/venta. En general, es muy habitual que estas no se ejecuten inmediatamente debido a restricciones de liquidez, necesidades de ocultar intenciones, etc. En este caso, el resultado supone tener carteras intermedias que son muy similares a la cartera original, pero con pequeñas modificaciones. Si estas conllevasen divergencias importantes en el perfil riesgo-beneficio, los inversores estarían expuestos a un riesgo imprevisto e indeseable. En este trabajo se muestra cómo añadiendo un objetivo adicional basado en ϵ-vecindarios permite atajar de forma efectiva este problema, permitiendo alcanzar soluciones más robustas a las metaheurísticas que lo incorporan.

El uso de ϵ-vecindarios para optimización robusta con MOEAs no es nuevo [3,4,8], pero, hasta donde conocemos, su aplicación a la limitación del riesgo en optimización de carteras es novedoso. La idea es definir un vecindario alrededor de cada solución tentativa (cartera), que se muestrea aleatoriamente para después analizar la estructura de dicha muestra en el espacio de objetivos. Así, se añade objetivo adicional al problema donde se favorezcan aquellas que están en zonas estables del espacio, i.e., aquellas carteras menos sensibles a las pequeñas variaciones que pueden sufrir cuando se ejecutan las distintas órdenes de mercado. Este mecanismo se ha incorporado a varios MOEAs habituales en el dominio, en particular NSGA-II [5], SPEA2 [22], SMPSO [19] y GDE3 [10].

El resto del artículo se organiza como sigue. En la próxima sección se incluye una descripción formal del problema de optimización de carteras de inversión, y se detalla la aproximación basada en ϵ-vecindarios descrita anteriormente. La Sec 3 está dedicada a la validación experimental de la propuesta. Finalmente, la última sección resume las principales conclusiones del trabajo y propone posibles lineas de trabajo futuro.

2. El problema de la carteras robustas de inversión

Una cartera de inversión se puede definir como una colección de activos. En general, el problema de la optimización de carteras consiste en seleccionar un conjunto óptimo de activos para incluir en la cartera así como la distribución del capital del inversor entre ellas. En los trabajos originales de Markowitz [15] se asumía que resolver el problema requería la satisfacción simultánea de maximizar el rendimiento esperado de la cartera, $E(R_p)$, y minimizar el riesgo de la misma. En este trabajo se usa el modelo de riesgo canónico que viene determinado por la varianza, σ_p^2. Formalmente, el problema multiobjetivo resultante se define como:

$$\text{mín}\, \sigma_p^2 = \Sigma_{i=1}^n \Sigma_{j=1}^n w_i w_j \sigma_{ij} \tag{1}$$

$$\text{máx}\, E(R_p) = \Sigma_{i=1}^n w_i \mu_i \tag{2}$$

$$\text{sujeto a:} \tag{3}$$

$$\Sigma_{i=1}^n w_i = 1 \tag{4}$$

$$0 \leq w_i \leq 1; i = 1...n \tag{5}$$

$$C_{min} \leq \Sigma(w_i \neq 0) \leq C_{max} \tag{6}$$

$$0.0 \leq lim_{inf} \leq w_i \leq lim_{sup} \leq 1.0 \tag{7}$$

donde n es el número de activos disponibles, μ_i es la rentabilidad esperada del activo i, σ_{ij} es la covarianza entre los activos i y j, y w_i son las variables de decisión que determinan la composición de la cartera (activos que están presentes en la misma). Las restricciones básicas de la formulación (Ec. 4 y 5) obligan, respectivamente, a que se inviertan todos los fondos y a prevenir que el inversor tome posiciones cortas. En este trabajo también se han considerados dos restricciones del mundo real: la cardinalidad (Ec. 6), que determina número mínimo y máximo, C_{min} y C_{max} respectivamente, de activos en los que se permite invertir simultáneamente; y los valores límites de cada peso w_i (Ec. 7), que deben estar en el rango $[lim_{inf}, lim_{sup}]$.

2.1. Enfoque robusto

En este trabajo se introduce una formulación extendida del modelo de media-varianza para carteras de inversión que include un tercer objetivo de estabilidad basado en ϵ-vecindarios. Este mecanismo se añade al mencionado anteriormente de remuestreo con marcas de tiempo (R+T, del inglés *time-stamped resampling*) [7], que ha mostrado ser efectivo. El R+T genera un conjunto de escenarios similares y filtra aquellas carteras que son muy sensibles a las desviaciones respecto a los rendimientos y la matriz de varianzas-covarianzas esperadas. También se considera la antiguedad de la solución tentativa (marcas de tiempo), y se favorecen las carteras que han tenido un mejor rendimiento durante un periodo de tiempo más largo.

El uso de ϵ-vecindarios para mejorar la robustez de las carteras se ha desarrollado como sigue. Cada cartera p es evaluada con respecto a los objetivos tradicionales, rendimiento, $E(R_p)$, y riesgo, $E(R_p)$, pero ahora se le añade uno nuevo, Rob_p. El cálculo de este último requiere la generación de H carteras vecinas, $p_1, \ldots, p_H$, usando la aproximación del Hipercubo Latino (HL) [18]. HL opera dividiendo el dominio de cada activo alrededor de $[-\epsilon, \epsilon]$ en H cuadrículas, dividiendo el ϵ-vecindario en H^n hipercubos. Así, la intensidad de la perturbación aplicada a las carteras está determinada por ϵ. Una vez que se han elegido los hipercubos (aleatoriamente), se genera un punto dentro de cada uno de ellos y se utiliza para calcular el rendimiento, $E(R_{p_i})$, y el riesgo, $\sigma^2_{p_i}$, de la cartera p_i. La media de la distancia de Mahalanobis [13] entre p y todos los p_i se considera como la función objetivo que se ha de minimizar, ya que esas distancias miden la dispersión de los rendimientos y riesgos del conjunto de carteras, i.e., cuanto mayor es la distancia de Mahalanobis, mayor es la dispersión y, por tanto, más sensible es la solución a pequeñas perturbaciones. La elección de la distancia de Mahalanobis está justificada por algunas propiedades deseables, como la independencia de la escala o el hecho de que tiene en cuenta estructuras de correlación para evaluar la similaridad. Todo esto reduce las diferencias en riesgo y rendimiento a un único número. Más formalmente, la nueva función objetivo es:

$$Rob_p = \frac{1}{H}\sum_{i=1}^{H} md(p, p_i) \tag{8}$$

donde $md(p, p_i)$ es la distancia de Mahalanobis en términos de riesgo y rendimiento entre las carteras p y p_i. Si se define el par $(E(R_{p_i}), \sigma^2_{p_i})$ como $\overline{x}$, $(E(R_p), \sigma^2_p)$ como $\overline{y}$, y Σ como la matriz de varianzas-covarianzas, la distancia se puede calcular usando la expresión:

$$md(\overline{x}, \overline{y}) = \sqrt{(\overline{x} - \overline{y})^T \Sigma^{-1} (\overline{x} - \overline{y})} \tag{9}$$

Dado que estos tres objetivos no son comparables con la formulación clásica biobjetivo, se ha realizado una proyección en el plano rentabilidad-riesgo. Así, los conjuntos de carteras se han dividido en subconjuntos de acuerdo con el nuevo objetivo, al que nos referiremos como ϵ-sensibilidad. Esto nos permitirá analizar la conexión entre los valores del tercer objetivo y la sensibilidad a la incertidumbre respecto a los valores reales de los parámetros de optimización. El proceso funciona como sigue: al final de cada ejecución de cada MOEA, el conjunto final de carteras se ordena según el objetivo de ϵ-sensibilidad, Rob_p, y se divide en tres subconjuntos. El primero de ellos contiene un tercio de las carteras con los valores más bajos, y clasificados como de "Alta estabilidad". El segundo se etiqueta como “Media estabilidad”, para el segundo tercio de carteras y, finalmente, el último tercio que se corresponde con las carteras localizadas en la porción menos estable del espacio de soluciones, considerados como los de “Baja estabilidad”. El proceso termina con la identificación de las carteras no dominadas en términos ya sólo de riesgo y rentabilidad, dentro de cada subconjunto anterior. Estos subfrentes ya son comparables con las soluciones proporcionadas por cualquier otro optimizador.

3. Experimentación

Esta sección está dedicada a presentar la experimentación realizada para evaluar el rendimiento del nuevo objetivo basado en ϵ-vecindarios. En primer lugar, se detalla el conjunto de datos y la metodología utilizada. Seguidamente, se presentan los algoritmos y su parametrización, así como las métricas que se usan para su evaluación. Finalmente se valoran los resultados obtenidos.

3.1. Muestra

Los experimentos se han realizado usando una muestra de las ganancias en 240 meses según ocho índices financieros que representan ocho clases de activos. Las series de ganancias mensuales, proporcionados por la firma *Datastream*, cubren el periodo que va desde enero de 1990 a diciembre de 2009. La lista de los índices se incluye en la Tabla 1.

Para asegurar que los algoritmos se enfrentan a un abanico amplio de situaciones, se han generado 120 problemas distintos optimización de un solo periodo. Para esto, una ventana de tamaño $n = 120$, que se corresponde con 10 años de datos, se ha desplazado 120 veces. Esto significa que los algoritmos dependen de la información relativa a los periodos t_1 a t_n para identificar la mejor asignación posible en el periodo t_{n+1} (t_1 a t_n se usan para hacer las estimaciones de los parámetros usados en la optimización, que pretende predecir la frontera eficiente en t_{n+1}. La ventana de 10 años se mueve de mes en mes hasta un total de 120 veces en total, para cubrir el intervalo de 01/01/1990 a 31/12/2009. Dada la naturaleza estocástica de los algoritmos, para cada ventana se han realizado 100 ejecuciones independiente de los algoritmos.

Tabla 1. Conjunto de datos

Nombre	Código
Frank Russell 1000 Growth	FRUS1GR
Frank Russell 1000 Value	FRUS1VA
Frank Russell 2000 Growth	FRUS2GR
Frank Russell 2000 Value	FRUS2VA
S&P GSCI Commodity Total Return	GSCITOT
MSCI EAFE	MSEAFEL
BOFA ML CORP MSTR ($)	MLCORPM
BOFA ML US TRSY /AGCY MSTRAAA($)	MLUSALM

Tabla 2. Parametrización de los algoritmos. L es la longitud del individuo.

Parametrización de NSGA-II		Parametrización de SPEA2	
Población	500 individuos	*Población*	500 individuos
Selección	torneo binario	*Selección*	torneo binario
Recombinación	SBX, $p_c = 0.9$	*Recombinación*	SBX, $p_c = 0.9$
Mutación	Polinomial, $p_m = 1.0/L$	*Mutación*	Polinomial, $p_m = 1.0/L$
Parametrización de SMPSO		**Parametrización de GDE3**	
Swarm	500 partículas	*Población*	500 individuos
Líderes	500 partículas	*Recombinación*	$CR = 0.1,\ F = 0.5$
Mutación	Polinomial, $p_m = 1.0/L$		

3.2. Algoritmos, representación y parametrización

Los MOEAs utilizados están entre los más conocidos: NSGA-II [5], SPEA2 [22], SMPSO [19] y GDE3 [10]. Todos ellos han utilizado una codificación real donde $w_i \in [0,1]$ representa la proporción invertida en el activo i. La gestión de las restricciones se ha hecho en todos los casos mediante un operador de reparación [7].

La comparación entre los algoritmos es justa ya que todos realizan 150000 evaluaciones, y todos las aproximaciones a los frentes contienen, como máximo, 500 soluciones no dominadas. Los parámetros empleados para calcular el tercer objetivo son: $H = 1000$ y $\epsilon = 0.01$. La configuración detallada de los algoritmos se incluye en la Tabla 2. La parametrización elegida es resultado de un estudio preliminar previo no detallado por limitaciones de espacio.

3.3. Indicadores de calidad

Los indicadores de calidad más usuales en la comunidad multi-objetivo (Hipervolumen, cubrimiento de conjuntos, etc.) no son adecuados para medir la calidad de las soluciones desde el punto de vista de la robustez. Estos indicadores son válidos para identificar dominancia o diversidad de las soluciones a lo largo del frente, pero no son capaces de capturar la divergencia entre resultados esperados y obtenidos. Por esta razón se han utilizado otro tipo de indicadores especialmente diseñados para esta tarea. Se han considerado cuatro: Estimación del error, Estabilidad, Rentabilidad perdida y Riesgo extremo [7].

La **Estimación del error** (EE, por sus siglas en inglés) considera la distancia de Mahalanobis media entre el riesgo y rendimiento esperado de cada cartera en la aproximación del frente y el riesgo y rendimiento de la misma a posteriori, una vez que los valores reales se han observado. La **Estabilidad** (ST, por sus siglas en inglés) mide la diferencia media entre el escenario esperado y 500 escenarios generados usando un bootstrap no paramétrico. El indicador **Rentabilidad perdida** (UR, por sus siglas en inglés) evalúa, para cada cartera, la diferencia entre el rendimiento alcanzado y el rendimiento potencial máximo dado un nivel de riesgo. Finalmente, el **Riesgo extremo** (ER, por sus siglas en inglés) mide la sensibilidad promedio a los peores escenarios posibles. Estos se definen como el 1 % de entre los 500 remuestreados con mayor distancia de Mahalanobis media entre los pares riesgo-rentabilidad de las carteras evaluadas con esos escenarios y los correspondientes en el escenario previsto. En todos los casos, los valores más bajos de los indicadores se asocian a una mayor robustez.

3.4. Resultados

En esta sección se presentan y analizan los resultados de incluir los ϵ-vecindarios como objetivo adicional para aumentar las robustez de las carteras finales alcanzadas por NSGA-II, SPEA2, SMPSO y GDE3. Dicha robustez se ha evaluado usando los cuatro indicadores anteriores, EE, ST, UR y ER. Por cada algoritmo, se han comparado los resultados de la versión canónica con respecto a cuatro alternativas: la versión con remuestreo y marcas de tiempo (R+T), y los tres niveles de fiabilidad procedentes de la extensión de R+T con el nuevo objetivo basado en ϵ-vecindarios (Sect. 2.1). Las tablas de datos incluyen la media, mediana y varianza sobre 100 ejecuciones independientes. Las diferencias en los indicadores con respecto a las versiones básicas se han analizado usando un test de Wilcoxon dado que no son normales.

Los resultados de EE (Tabla 3) son muy variables dependiendo del MOEA considerado. De los cuatro algoritmos básicos, GDE3 es el que ha proporcionado el mejor punto de partida. La inclusión del remuestreo con marcas de tiempo (R+T) permite una mejora considerable en todos los algoritmos, alcanzando el mínimo valor del indicador para SMPSO. La mayor ganancia se obtiene, sin embargo, con SPEA2, pero se explica ya que su versión canónica ha sido la que ha obtenido los peores (mayores) valores de EE. Este algoritmo ha sido también el menos consistente, ya que presenta valores de varianza mucho más altos que el resto. El mejor valor medio para este indicador se ha alcanzado en el subconjunto de alta estabilidad en SMPSO. Aquí, la inclusión del objetivo con la ϵ-sensibilidad ha resultado en una mejora sobre R+T de 12.7 % en términos globales, y 28.7 % en términos relativos. Todas las diferencias son entre el valor del algoritmo básico y las versiones robustas son significativas al 1 %, excepto la comparación entre las soluciones de baja estabilidad en SMPSO y la configuración estándar, que no se puede rechazar la igualdad a un 5 %.

Tabla 3. Resultados de EE y ST

	EE				ST			
	Media	**Mediana**	**Varianza**	**Mejora**	**Media**	**Mediana**	**Varianza**	**Mejora**
NSGAII	2.2613	1.8071	3.9781		6.5078	6.1306	13.0883	
NSGAII R+T	1.2216	0.6357	2.5894	45.98 %	2.6528	2.4189	2.3038	59.24 %
NSGAII Alta	1.4160	0.6662	4.4399	37.38 %	2.5566	2.0398	3.8033	60.71 %
NSGAII Media	1.3058	0.5213	3.9894	42.26 %	2.4027	1.9133	3.1669	63.08 %
NSGAII Baja	1.5023	0.7374	3.6167	33.57 %	3.2265	2.8121	4.2622	50.42 %
SPEA2	2.5196	1.8162	5.1751		7.4009	7.0380	20.8347	
SPEA2 R+T	1.1016	0.5270	2.4888	56.28 %	2.2201	1.8669	2.2534	70.00 %
SPEA2 Alta	1.1699	0.4355	4.0348	53.57 %	1.7877	1.7189	0.9625	75.84 %
SPEA2 Media	1.1485	0.4245	3.9249	54.42 %	1.7290	1.5949	0.9446	76.64 %
SPEA2 Baja	1.2720	0.4907	4.1532	49.52 %	2.0624	1.8056	1.5423	72.13 %
SMPSO	1.4939	1.2324	1.7256		5.2076	4.8454	7.0554	
SMPSO R+T	1.0234	0.7044	0.9181	31.49 %	3.2402	2.8408	3.7394	37.78 %
SMPSO Alta	0.8334	0.5243	1.2212	44.21 %	2.2559	2.0676	1.6274	56.68 %
SMPSO Media	1.0864	0.7259	1.1941	27.28 %	3.7002	3.1791	5.6781	28.95 %
SMPSO Baja	1.5237	1.1091	1.9836	-2.00 %	5.1610	4.4768	9.7711	0.89 %
GDE3	1.4807	1.1839	1.6799		5.1506	4.7763	6.9177	
GDE3 R+T	1.1001	0.6798	1.3447	25.71 %	3.3691	3.0426	3.9137	34.59 %
GDE3 Alta	1.0567	0.6538	1.6219	28.63 %	3.1418	2.7035	5.7365	39.00 %
GDE3 Media	1.2213	0.7590	1.7732	17.52 %	3.9129	3.2678	6.4121	24.03 %
GDE3 Baja	1.4567	0.9770	2.1729	1.62 %	4.9855	4.1357	10.3618	3.21 %

Tabla 4. Resultados de ER y UR.

	ER				UR			
	Media	**Mediana**	**Varianza**	**Mejora**	**Media**	**Mediana**	**Varianza**	**Mejora**
NSGAII	3.2148	2.8235	3.8975		3.4788	2.7252	7.9201	
NSGAII R+T	1.7640	1.1865	2.6789	45.13 %	2.5265	1.9556	4.2799	27.37 %
NSGAII Alta	2.0821	1.2042	5.2588	35.23 %	2.5961	2.0124	6.5110	25.37 %
NSGAII Media	1.8916	1.1513	3.8595	41.16 %	2.4815	1.8862	5.5478	28.67 %
NSGAII Baja	2.2715	1.5507	4.0121	29.34 %	2.7626	2.1029	5.9501	20.59 %
SPEA2	3.5315	3.0472	4.9711		3.5859	2.7953	8.6577	
SPEA2 R+T	1.5991	1.0266	2.5787	54.72 %	2.3506	1.7922	4.1225	34.45 %
SPEA2 Alta	1.6560	0.8300	4.2195	53.11 %	2.2857	1.7652	5.0552	36.26 %
SPEA2 Media	1.6658	0.8852	3.9287	52.83 %	2.2563	1.6774	4.9483	37.08 %
SPEA2 Baja	1.9061	1.0162	4.5500	46.03 %	2.4650	1.8161	5.6578	31.26 %
SMPSO	2.2303	1.9779	1.8379		2.9953	2.4891	5.3718	
SMPSO R+T	1.4872	1.2325	1.0194	33.32 %	2.2702	1.7245	3.0980	24.21 %
SMPSO Alta	1.3263	0.9891	1.2227	40.53 %	1.8383	1.4231	2.2844	38.63 %
SMPSO Media	1.7204	1.3656	1.5185	22.86 %	2.2866	1.7815	3.5252	23.66 %
SMPSO Baja	2.2431	1.9261	2.4140	-0.57 %	2.8306	2.2418	5.1108	5.50 %
GDE3	2.2279	1.9722	1.8898		3.3427	2.8714	6.0005	
GDE3 R+T	1.6970	1.2665	1.6804	23.83 %	2.6886	2.1756	4.6265	19.57 %
GDE3 Alta	1.6523	1.2325	1.8790	25.84 %	2.4671	1.9619	3.7148	26.20 %
GDE3 Media	1.9049	1.4598	2.2477	14.50 %	2.7911	2.2465	4.7696	16.50 %
GDE3 Baja	2.2192	1.8683	2.7076	0.39 %	3.1406	2.5447	5.9849	6.05 %

En términos de estabilidad, SPEA2 es el algoritmo base que obtiene tanto el mejor valor de ST como la mayor mejora. Sorprendentemente, la mejora de 76.64 % no se consigue con las carteras de más estabilidad, si no con las de estabilidad media. Asi, la Tabla 3 muestra que la inclusión del objetivo basado en ϵ-vecindarios a la versión R+T de SPEA2 mejora el rendimiento en los tipos de estabilidad. Esta observación se mantiene para los resultados de la mediana. Si bien los conjuntos de carteras identificados por GDE3 muestran una gran estabilidad en su configuración básica, el algoritmo no se beneficia tanto como el resto de la inclusión del remuestreo y el tercer objetivo. Esto también ocurre con SMPSO. De nuevo, la única excepción de significación estadística de las diferencias al 1 % entre las versiones robustas y el algoritmo base es SMPSO Baja. Aunque en esta ocasión la hipótesis nula si se rechaza al nivel convencional del 5 %.

La Tabla 4 incluye los resultados de ER. En este caso, la introducción de la nueva aproximación resulta en una clara mejora sobre los algoritmos básicos. La contribución sobre la versión R+T es, no obstante, mínima. La excepción es SMPSO, para el que la porción del frente de alta estabilidad proporcionada por el nuevo tercer objetivo muestra una mejora del 40.53 % respecto al 33.32 %. Esta combinación es también la que obtiene los mejores resultados globales. Los algoritmos genéticos han funcionado bastante mal aquí. Así, aunque los resultados mejoran al algoritmo base, la nueva aproximación empeora la versión R+T independientemente del nivel de estabilidad elegido. Los test estadísticos han sacado exactamente los mismos resultados que para la métrica anterior.

La aproximación propuesta también ha tenido impacto en el indicador UR. En todos los casos, como se observa en la Tabla 4, la inclusión del tercer objetivo ha reducido la cantidad de dinero perdido. La cantidad concreta varía con el algoritmo de base, pero los mejores resultados se han obtenido con SMPSO.

Este algoritmo también ha resultado ser el más consistente, como indican las varianzas. Una vez más, el frente de mayor estabilidad (el tercio inferior en términos de la ϵ-sensibilidad) es el que ha alcanzado las mayores reducciones en UR. Comparado con la versión básica de PSO, que es el algoritmo con el mejor punto de partida, llega a una caída del 38.63 %. Dado que los valores de este indicador hay que minimizarlos, los datos muestran una contribución importante del nuevo objetivo en la mejora de la robustez de las carteras. El estudio estadístico pertinente ha resultado en que las diferencias de las versiones básicas respecto de las versiones robustas son todas significativas al 1 %.

Como conclusión, los resultados de la experimentación sugieren que la inclusión del tercer objetivo basado en ϵ-vecindarios permite mejorar la robustez de las soluciones. Es, por tanto, un mecanismo fiable para conseguir esta meta que se puede incorporar a cualquier MOEA. Si bien la magnitud de la ganancia varía dependiendo del indicador en cuestión y del algoritmo básico elegido, hay una tendencia clara que indica que cuanto más bajo es el valor del tercer objetivo (subfrentes de alta estabilidad), más robusta es la cartera.

4. Conclusiones y trabajo futuro

En este trabajo se ha presentado un método para mejorar la robustez de carteras de inversión. La aproximación se basa en la extensión de la formulación clásica en media/varianza con un tercer objetivo que modela la estructura del landscape alrededor de cada cartera candidata, y favorece aquellas que están en áreas fiables (secciones del espacio de soluciones donde carteras similares están también cerca en el espacio de objetivos, i.e., tiene perfiles esperados similares en término de riesgo y ganancia).

La implementación de esta estrategia se basa en el concepto de ϵ-vecindarios que, junto con el uso de la distancia de Mahalanobis, definen una métrica de dispersión que se debe minimizar. Este objetivo ha mostrado ser un complemento perfecto para una estrategia de remuestreo basado en marcas de tiempo (R+T) y se ha incorporado a cuatro MOEAs muy diferentes (NSGA-II, SPEA2, SMPSO y GDE3). Estos, junto sus versiones con R+T, han sido probados sobre una muestra de rendimientos históricos y evaluados a a través de cuatro métricas de robustez. La contribución del nuevo objetivo ha sido sustancial, contribuyendo a un incremento importante de la fiabilidad de las soluciones.

Hay varias extensiones potenciales de este trabajo. Entre ellas, se pueden mencionar estudios de escalabilidad con el número de activos; del rendimiento del nuevo objetivo con otros operadores evolutivos; sobre la exploración de diferentes estrategias de perturbación o ampliar el estudio a más MOEAs.

Agradecimientos. Francisco Luna agradece la ayuda de la Universidad de Málaga, Campus de Excelencia Internacional Andalucía Tech. El trabajo de David Quintana y Pedro Isasi ha sido financiado parcialmente por el Ministerio de Economía y Competitividad bajo el contrato ENE2014-56126-C2-2-R.

Referencias

1. Anagnostopoulos, K., Mamanis, G.: A portfolio optimization model with three objectives and discrete variables. Comp. & Oper. Res. 37(7), 1285–1297 (2010)
2. Babaei, S., Sepheri, M., Babaei, E.: Multi-objective portfolio optimization considering the dependence structure of asset returns. EJOR 244(2), 525–539 (2015)
3. Beyer, H.G., Sendhoff, B.: Robust optimization – a comprehensive survey. Comput. Methods Appl. Mech. Engrg. 196(33 - 34), 3190 – 3218 (2007)
4. Deb, K., Gupta, H.: Introducing robustness in multi-objective optimization. Evolutionary Computation 14(4), 463 – 494 (2006)
5. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. on Ev. Comp. 6(2), 182–197 (2002)
6. García, S., Quintana, D., Galván, I., Isasi, P.: Portfolio optimization using spea2 with resampling. In: IDEAL'11. pp. 127–134 (2011)
7. García, S., Quintana, D., Galván, I., Isasi, P.: Time-stamped resampling for robust evolutionary portfolio optimization. Expert Systems with Applications 39(12), 11722–11730 (2012)
8. Gaspar-Cunha, A., Covas, J.: Robustness in multi-objective optimization using evolutionary algorithms. Comp. Opt. and Appl. 39(1), 75–96 (2008)
9. Huang, Y., Kou, G.: A kernel entropy manifold learning approach for financial data analysis. Decision Support Systems 64, 31–42 (2008)
10. Kukkonen, S., Lampinen, J.: GDE3: The third evolution step of generalized differential evolution. In: IEEE CEC). pp. 443 – 450 (2005)
11. Liagkouras, K., Metaxiotis, K.: A new probe guided mutation operator and its application for solving the cardinality constrained portfolio optimization problem. Expert Systems with Applications 41(14), 6274–6290 (2014)
12. Lwin, K., Qu, R., Kendall, G.: A learning-guided multi-objective evolutionary algorithm for constrained portfolio optimization. Applied Soft Computing Journal 24, 757–772 (2014)
13. Mahalanobis, P.: On the generalized distance in statistics. Proceedings of the National Institute of Sciences 2, 49–55 (1936)
14. Markowitz, H.: Portfolio selection. The Journal of Finance 7(1), 77–91 (1952)
15. Markowitz, H.: Portfolio Selection: efficient diversification of investments. John Wiley & Sons (1959)
16. Metaxiotis, K., Liagkouras, K.: Multiobjective evolutionary algorithms for portfolio management: A comprehensive literature review. Expert Systems with Applications 39(14), 11685–11698 (2012)
17. Michaud, R.: Efficient Asset Management. Harvard Business School Press, Boston, MA (1998)
18. Montgomery, D.: Design and analysis of experiments. Wiley, New York (2001)
19. Nebro, A., Durillo, J., García-Nieto, J., Coello Coello, C., Luna, F., Alba, E.: SMPSO: A new pso-based metaheuristic for multi-objective optimization. In: IEEE MCDM. pp. 66–73 (2009)
20. Peng, Y., Kou, G., Shi, Y., Chen, Z.: A descriptive framework for the field of data mining and knowledge discovery. International Journal of Information Technology & Decision Making 7(4), 639–684 (2008)
21. Yue, W., Wang, Y., Dai, C.: An evolutionary algorithm for multiobjective fuzzy portfolio selection models with transaction cost and liquidity. Mathematical Problems in Engineering p. 569415 (2015)
22. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength Pareto evolutionary algorithms. In: EUROGEN 2001. pp. 95–100 (2002)