

حل مسئله ردیابی شاخص با استفاده از الگوریتم فراابتکاری ترکیبی کرم شب تاب

سید جعفر سجادی^۱

سید مجتبی سجادی^۲

ملیحه سبزواری^۳

چکیده

سبد ردیاب شاخص، مدلی از سرمایه‌گذاری منفعل در بازار سرمایه است که با محبوبیت زیادی مواجه شده است. هدف از این نوع سرمایه‌گذاری، خرید سبد سهامی است که رفتاری مشابه شاخص دارد بدون آن که مجبور باشیم تمام سهام تشکیل دهنده شاخص را بخریم. محدودیت‌های مسئله شامل محدودیت کف و سقف سرمایه‌گذاری و همچنین محدودیت عدد صحیح تعداد سهام خریداری شده است. در این تحقیق از الگوریتم ترکیبی فراابتکاری کرم شب تاب برای حل مسئله بهره گرفته شده است. شاخص‌های مورد بررسی شاخص بورس اوراق بهادار تهران، S&P500 و Hang Seng هستند. نتایج تحقیق، کارایی بالای این الگوریتم در ردیابی شاخص را نشان می‌دهد.

واژگان کلیدی: الگوریتم کرم شب تاب، ردیابی شاخص، مدیریت منفعل سرمایه‌گذاری

sjsadjadi@iust.ac.ir

msajadi@ut.ac.ir

malihe.sabzevari@yahoo.com

۱. استاد، گروه صنایع دانشکده فنی و مهندسی، دانشگاه علم و صنعت، تهران، ایران

۲. استادیار، دانشکده کارآفرینی دانشگاه تهران، تهران، ایران

۳. دانشجوی دکتری، دانشکده فنی و مهندسی، دانشگاه علوم و تحقیقات، تهران، ایران

تاریخ دریافت مقاله: ۹۴/۸/۱۹؛ تاریخ پذیرش مقاله: ۹۴/۱۱/۱۰

مقدمه

یکی از استراتژی‌های سرمایه‌گذاری، ردیابی شاخص است که متشکل از سبد سهامی است که حرکت شاخص را شبیه‌سازی می‌کند چرا که داده‌های تاریخی و تجربه ثابت کرده است که بازار همواره بازده مثبتی در بلندمدت دارند. ردیابی شاخص بعنوان یک استراتژی محبوب، به دلیل بازده مثبت و ریسک کمتر بطور گسترده توسط مدیران پورتنفوی در سال‌های اخیر بکار گرفته می‌شود. برای یافتن این سبد سهام بهینه، دو رویکرد قابل بررسی است: یکی یافتن این دارایی‌ها که جستجو در فضایی گسسته است و دیگری وزن نسبت داده شده به دارایی‌های این سبد است که جستجو در فضایی پیوسته است. این قبیل مسائل بعنوان مسائل حل نشدنی در زمان چند جمله‌ای (NP-Hard) شناخته می‌شوند یعنی هیچ الگوریتم قطعی شناخته شده‌ای که یک راه حل دقیق را در زمان چند جمله‌ای بیابد وجود ندارد. بمنظور نشان دادن پیچیدگی مسئله بهینه‌سازی پورتنفوی، شرایطی را که در ادامه مطرح می‌شود بررسی می‌کنیم. اگر هدف پیدا کردن پورتنفوی بهینه است که متشکل از ۱۰ دارایی از میان ۱۰۰ دارایی باشد پس تمام ترکیبات ممکنه $\binom{100}{10} = 1.73 \times 10^{13}$ خواهند شد. که برای هر یک از این ترکیبات، وزن‌های بهینه نیز باید پیدا شود (درفضای حل پیوسته). همانطور که مشاهده می‌شود، این ترکیبات در فضای حل پیوسته به بی‌نهایت میل خواهند کرد. در نتیجه الگوریتم‌های جستجوی کامل یا دیگر روش‌های سنتی در تحقیق در عملیات برای یافتن راه حل بهینه ناکارآمد هستند، یا در بهترین حالت در بهینه محلی گیر می‌کنند.

تحقیق حاضر با هدف ایجاد ترکیبی بهینه از دارایی‌های ردیابی کننده شاخص، بمنظور کاهش ریسک و بازده مثبت انجام شده است. بدین منظور، الگوریتم فراابتکاری ترکیبی کرم شب‌تاب برای یافتن این پورتنفوی مورد استفاده قرار گرفته است. در ادامه ابتدا پیشینه تحقیق را بیان کرده و سپس به روش تحقیق و نتایج تجربی خواهیم پرداخت. در پایان نیز نتیجه گیری ارائه خواهد شد.

پیشینه پژوهش

تحقیقات مختلفی برای حل این مسئله انجام شده است. در تحقیقی که فنگمین و ژوسونگ و زونگبن (۲۰۱۵) ارائه کرده‌اند از یک روش گرادیان تصویر شده غیریکنواخت^۱ (NPG) برای حل این مسئله استفاده شد. در این روش در هر تکرار چندین زیرمسئله گرادیان تصویر شده حل می‌شود و نشان داده شده که هر زیرمسئله یک حل فرم بسته دارد که می‌تواند در زمان خطی حل شود. در انتها این مدل با دو مدل دیگر که از الگوریتم‌های تکاملی ترکیبی برای حل مسئله ردیابی شاخص استفاده می‌کنند مقایسه شده است. در این مقاله ثابت شده که روش ارائه شده سریع‌تر از دو روش دیگر به سمت بهینه خواهد رفت.

گیو و لی (۲۰۱۳) انتخاب سبد سهامی که محدودیت عدد صحیح آن متفاوت است را مورد مطالعه قرار دادند که در آن واریانس هر سهم در سبد، با بازده مورد انتظار و محدودیت عدد صحیح روی تعداد سهام را مینیمم می‌کند. آن‌ها حد پایین کارایی را بکار بردند و از یک الگوریتم شاخه و کران برای حل آن استفاده کردند.

تحقیقی تحت عنوان "مطالعات تجربی در مورد الگوریتم‌های ترکیبی الهام گرفته از طبیعت برای بهینه‌سازی پورتفولیوی مالی" توسط گیاناکوریس، واسیلیادیس و دانیاس (۲۰۱۰) انجام شد که در آن از الگوریتم کرم شب تاب و کلونی مورچگان بهره جسته است و از این دو الگوریتم برای حل مسئله بهینه‌سازی پورتفوی که هدف آن یافتن ترکیب بهینه از دارایی‌ها و وزن مربوط به آن‌ها استفاده کرده‌اند که محدودیت‌ها ارضا و تابع هدف برآورده شود. در تحقیق آن‌ها، تابع هدف حداکثر نسبت‌های مالی است که ترکیبی از فاکتورهای ریسک و بازده است. در عین حال، یک محدودیت مشکل نیز مدنظر قرار گرفته است که اشاره به توانایی ردیابی پورتفوی ساخته شده بسوی یک شاخص سهام معیار دارد.

وو (۲۰۱۱) از یک مدل صفر و یک برای ردیابی شاخص S&P100 استفاده کرده است و نتایج به دست آمده از شاخص هدف و پورتفوی را با یکدیگر مقایسه کرده است.

1. no monotone projected gradient

توریانو و سوارز (۲۰۰۸) یک راهبرد ترکیبی متشکل از یک الگوریتم تکاملی و برنامه‌ریزی کوادراتیک را بعنوان استراتژی هیبریدی بمنظور حل مسئله ردیابی شاخص طراحی کردند. در رویکرد آن‌ها، تابع شایستگی^۱ برای حل مدل توسط الگوریتم ژنتیک، بهینه‌سازی درجه دو تعریف شده است. تابع هدف پیشنهادی آن‌ها میانگین مربع خطا بین بازده‌های شاخص و پورتنفوی ردیابی کننده بود. آن‌ها همچنین محدودیت‌های میزان سرمایه‌گذاری در مجموعه دارایی‌های پورتنفوی، تعداد سهام تشکیل دهنده و میزان سرمایه قابل تخصیص به هر دارایی را نیز مد نظر قرار دادند. ایشان در تحلیل روش خود از داده‌های بیزلی استفاده کرده و نتایج یافته‌ها و زمان حل روش خود را ارائه داده‌اند.

بیزلی، مید و چانگ (۲۰۰۳) از رویکرد تکاملی الگوریتم ژنتیک استفاده کرده و علاوه بر مسئله ایجاد پورتنفوی ردیابی کننده، تعدیل ترکیب آن و هزینه‌های معاملاتی را نیز مدنظر قرار دادند و علاوه بر این، مسئله را بدون در نظر گرفتن هزینه‌های معاملاتی نیز حل کردند.

کرینک، میتیک و پاترلینی (۲۰۰۹) از الگوریتم فراابتکاری تکامل دیفرانسیلی (DE^۲) استفاده کردند. ایشان ابتدا بیان می‌دارند که از روش‌های معمول نمی‌توان مسئله را حل کرد و باید قطعاً برای حل مسئله از روش‌های فراابتکاری استفاده کرد. سپس به مقایسه این روش‌ها می‌پردازند و بیان می‌دارند که در مسائل پیچیده، DE مناسب‌تر از الگوریتم ژنتیک، شبیه‌سازی تبرید (SA) و بهینه‌سازی بر اساس الگوریتم PSO^۳، با توجه به تحقیقات اخیر عمل کرده است.

مارینگر (۲۰۰۷) نیز روش حل فراابتکاری DE را برای حل سبد ردیاب شاخص متوسط صنعت سهام داو جونز^۴ آمریکا بکار برده است. وی به بیان مزایای روش حل فراابتکاری DE می‌پردازد و بیان می‌کند که اولین مزیت DE، داشتن پارامترهای کم و دومین مزیت راحتی استفاده است. بعلاوه چگونگی تقویت الگوریتم برای جلوگیری از

1. Fitness Function
2. Differential Evolution
3. Particle swarm algorithm
4. Down Jones

همگرایی زودرس را تبیین می‌کند. وی همچنین از داده‌های ۱۶۴۸ روز سهام بخش داوجونز آمریکا از مارس ۲۰۰۰ تا نوامبر ۲۰۰۶ استفاده کرده است در حالی که داده‌های مفقوده با متوسط نزدیکترین قیمت‌ها جایگزین شده‌اند.

فرهاد حنیفی و محمد مهدی بحرالعلوم (۱۳۸۸) برای اولین بار مسئله پورتفوی ردیابی کننده شاخص را در ایران مطرح ساختند. در مسئله مورد نظر با محدودیت عدد صحیح انتخاب سهام ۵، ۱۰، ۱۵ و ۲۰ از بازار سهام و با استفاده از الگوریتم ژنتیک به حل آن پرداختند. آن‌ها سه رویکرد الگوریتم ژنتیک کلاسیک، الگوریتم ژنتیک چند مرحله‌ای و الگوریتم ژنتیک بهبود یافته را برای حل مسئله شاخص محور بکار گرفتند. نتایج حاصل، برتری الگوریتم ژنتیک چند مرحله‌ای را به دلیل دستیابی به پورتفوهایی با خطای ردیابی کمتر به اثبات رساند.

ناصر شمس و محسن ورسه‌ای (۱۳۸۹) یک روش حل ابتکاری بمنظور بهینه‌سازی مسئله سبد ردیاب شاخص پیشنهاد کردند. محدودیت‌های مسئله شامل محدودیت عدد صحیح انتخاب تعداد سهام و همچنین محدودیت کف و سقف سرمایه‌گذاری است. در روش مورد استفاده آن‌ها مسئله به دو زیرمسئله انتخاب سهام و تخصیص اوزان بهینه تقسیم شده است و با محدود کردن فضای مسئله به ۳۰ سهم از کل سهام بازار مسئله بصورت قطعی حل شده و در نهایت دو زیرمسئله با هم ترکیب شده‌اند.

روش شناسی پژوهش

در سبد ردیاب شاخص بطور مشخص به دنبال پاسخ به این دو سوال هستیم که در درجه اول کدام k سهم از مجموعه n سهم مناسب‌تر هستند و ثانیاً سرمایه‌گذاری در هر کدام از k سهم در سبد ردیاب شاخص با چه وزنی صورت بگیرد. به بیان دیگر، کدام k سهم از n سهم و با چه وزنی در سبد ردیاب شاخص سبب می‌شوند انحراف بازده سبد از بازده شاخص زمانی صفر تا T کمترین باشد؟ ابتدا برای بازه زمانی $(0, T)$ مجموعه K سهم را یافته و سپس میزان بازدهی آن را در بازه زمان $(T, T+\delta)$ بررسی می‌کنیم. مسئله بصورت زیر فرموله می‌شود:

$$\text{minimize } \frac{[\sum_{t=1}^T ((\sum_{i=1}^k z_i r_{it} w_i) - R_t)^2]^{1/2}}{T} \quad \text{رابطه (۱)}$$

subject to:

$$\sum_{i=1}^k z_i = k \quad \text{رابطه (۲)}$$

$$\sum_{i=1}^k z_i w_i = 1 \quad \text{رابطه (۳)}$$

$$0 \leq w_i \leq 1 \quad i = 1, 2, \dots, n \quad \text{رابطه (۴)}$$

$$x^l z_i \leq w_i \leq x^u z_i \quad \text{رابطه (۵)}$$

$$z_i \in \{0, 1\}$$

پارامترها:

T : بیانگر دوره مشاهده ارزش گذشته سهام و شاخص بر اساس واحد زمان انتخابی است. که می تواند روزانه یا هفتگی باشد.

R_t : بیانگر عدد شاخص در زمان t است.

r_{it} : بازده هر سهم در زمان t است.

k : تعداد سهام های انتخابی در سبد ردیاب است.

متغیرهای تصمیم:

z_i : نشان دهنده انتخاب سهم است که اگر مقدار آن از صفر به یک تغییر کند بیانگر انتخاب آن در پورتفوی است.

w_i : نشان دهنده وزن هر سهم در زمان t است.

تابع هدف در این تحقیق، حداقل ساختن خطای ردیابی این زیرمجموعه از سهام است که بصورت تابعی از اختلاف میان بازده پورتفوی زمانی ردیابی کننده و شاخص در نظر گرفته شده است.

$$E = \left[\left[\sum_{t \in S} \Delta_t |r_t - R_t|^\alpha \right]^{1/\alpha} \right] / T \quad \text{(رابطه ۶)}$$

Δ_t : پارامتری که برای بازده‌ها با دوره زمانی نزدیک تر و دورتر وزن متفاوتی قائل می‌شود.
E: خطای ردیابی

در این تحقیق آلفا برابر دو و Δ_t برابر یک لحاظ شده است.

محدودیت عدد صحیح اصلی ترین محدودیت مسئله و نشان دهنده تعداد سهام پورتفوی است (رابطه ۲). (رابطه ۳ و ۴) محدودیت وزن اقلام تشکیل دهنده پورتفوی، اشاره به کسری از بودجه است که به سرمایه گذاری در سهام تشکیل دهنده پورتفوی اختصاص می‌یابد. متغیر پیوسته w_i را به گونه‌ای تعریف می‌کنیم که نمایانگر وزن سرمایه گذاری در سهم i ام از میان کل سرمایه گذاری صورت گرفته در π سهم باشد. محدودیت کف و سقف سرمایه گذاری اشاره به حداقل و حداکثر سرمایه‌ای است که به سهام i اختصاص می‌دهیم (رابطه ۵). x^l حداقل درصد مقدار سرمایه گذاری در سهم شرکت i است بطوری که اگر سهم شرکت i برای قرار گرفتن در سبد انتخاب شد، مدیریت یا شرایط بازار یا قوانین، ما را موظف به سرمایه گذاری در حداقل x^l درصد از کل حجم سرمایه گذاری سبد در سهم i کند. x^u : بیانگر حداکثر درصد سرمایه گذاری سهم شرکت i است که مدیریت یا شرایط بازار یا قوانین، ما را موظف به سرمایه گذاری در حداکثر x^u درصد از کل حجم سرمایه گذاری سبد در سهم i کند. در این تحقیق، حداقل سرمایه گذاری ۰/۰۱ و برای حداکثر آن، مقدار یک در نظر گرفته شده است.

الگوریتم کرم شب تاب

این الگوریتم بر پایه رفتار اجتماعی کرم‌های شب‌تاب و حل مسائل بهینه‌یابی استوار است. بطور خاص برای مسائل NP-Hard پیوسته کاربرد دارد و رفتار کرم‌ها را شبیه‌سازی

کرده است. کرم‌های شب‌تاب از خود نور ساطع می‌کنند و به سمت کرمی که نور بیشتری تولید می‌کند جذب می‌شوند. درخشندگی کرم‌های شب‌تاب با استفاده از میزان آن در تابع هدف مشخص می‌شود. در استفاده از این الگوریتم، این کرم‌ها با نرخ مشخصی حرکت تصادفی دارند تا در جواب بهینه محلی گیر نکنند.

ساختار کرم‌های شب‌تاب

هر جواب بالقوه که در ساختار الگوریتم به آن‌ها کرم‌های شب‌تاب می‌گوییم، متشکل از یک بردار n بعدی است که k سهم از کل n سهام توسط تابع ایجاد کننده جمعیت اولیه به گونه‌ای تصادفی از آن انتخاب خواهد شد که این k سهم مقادیر مثبت گرفته و مابقی سهام مقادیری صفر خواهند داشت (محدودیت $\sum z_i = k$)، هر یک از سهام انتخاب شده مقادیری بین 0.01 و 1 اختیار کرده تا بدین ترتیب محدودیت کف و سقف سرمایه‌گذاری برآورده شود ($0.01 \leq w_i \leq 1$) و این سهام‌ها به گونه‌ای وزن‌دهی می‌شوند که مجموع آن‌ها برابر یک شود ($\sum w_i = 1$). همانطور که مشاهده می‌شود تمام محدودیت‌ها در تابع ایجاد کننده جمعیت اولیه ارضا خواهد شد.

جذابیت کرم‌های شب‌تاب

بعد از ایجاد جمعیت اولیه از کرم‌های شب‌تاب (جواب‌های بالقوه)، حال باید در پی یافتن درخشندگی هر یک از کرم‌ها بمنظور یافتن درخشان‌ترین آن‌ها باشیم. همانطور که پیش‌تر اشاره شد، در ساده‌ترین حالت می‌توان درخشش یک کرم شب‌تاب را با مقدار آن در تابع هدف برآورد مرد. با توجه به تابع هدف مسئله که یک مسئله مینیمم‌سازی است بهترین جواب، جوابی است که کمترین مقدار را در تابع هدف داشته باشد. بعبارت دیگر، خطای ردیابی آن کمترین باشد. در این مرحله به هر یک از کرم‌ها درخشندگی (مقدار در تابع هدف) نسبت می‌دهیم.

فاصله کرم‌های شب تاب

با توجه به این که هر یک از جواب‌ها بصورت آرایه‌ای از مقادیر صفر و یک هستند برای محاسبه فاصله دو کرم شب تاب از یکدیگر می‌توان از رابطه جاکارد استفاده کرد. برای بدست آوردن اندازه فاصله بین دو کرم شب تاب x_i و x_j ، با ویژگی‌های دودویی جدول ۱ را محاسبه می‌کنیم:

جدول ۱. جدول احتمالی 2×2

		x_j	
		1	0
x_i	1	a	b
	0	c	d

a: تعداد ویژگی‌های دودویی از نمونه x_i و x_j ، بطوری که $x_{i,m} = 1, x_{j,m} = 1$. (شباهت)

b: تعداد ویژگی‌های دودویی از نمونه x_i و x_j ، بطوری که $x_{i,m} = 1, x_{j,m} = 0$. (عدم شباهت)

c: تعداد ویژگی‌های دودویی از نمونه x_i و x_j ، بطوری که $x_{i,m} = 0, x_{j,m} = 1$. (عدم شباهت)

d: تعداد ویژگی‌های دودویی از نمونه x_i و x_j ، بطوری که $x_{i,m} = 0, x_{j,m} = 0$. (شباهت)

رابطه جاکارد:

$$r_{ij} = a / (a + b + c)$$

رابطه (۷)

که در آن x_i و x_j هر یک از جواب‌های بالقوه و $x_{i,m}$ ، $x_{j,m}$ هر یک از دارایی‌ها هستند. r_{ij} هم فاصله هر دو کرم از یکدیگر است.

حرکت کرم‌ها

حرکت کرم‌ها از دو بخش تشکیل می‌شود. یکی نزدیک شدن به بهینه محلی و دیگری مرحله تصادفی. در این قسمت نزدیک شدن به سمت بهینه محلی صورت می‌گیرد اما در بخش مرحله تصادفی از روشی مشابه به جهش در الگوریتم ژنتیک بهره گرفته شده است.

مرحله حرکت به سوی بهینه:

در صورتی که کرم با درخشندگی کمتر را x_i بنامیم که به سمت کرمی با درخشندگی بیشتر به نام x_j حرکت می‌کند، برای هر دارایی $x_{i,m}$ از کرم x_i داریم:

$$x_{i,m} = x_{i,m} + \beta \cdot e^{-\gamma r_{ij}} (x_{j,m} - x_{i,m}) \quad \text{رابطه ۸}$$

عبارت $\beta \cdot e^{-\gamma r_{ij}} (x_{j,m} - x_{i,m})$ نشان دهنده میزان جذب شدن کرم x_i به سمت x_j است. که این میزان جذب شدن به فاصله دو کرم و درجه جذب نور توسط محیط، γ ، بستگی دارد. β نشان دهنده درخشندگی در $t=0$ است.

مرحله حرکت تصادفی:

در این مرحله بعد از حرکت هر یک از کرم‌ها عددی بصورت تصادفی بین $(1, n)$ تولید می‌شود. سپس دارایی مربوط به آن عدد در x_i مقداری تصادفی در بازه $(0.01, 1)$ می‌گیرد. (۰/۰۱) حداقل دارایی با توجه به کف سرمایه‌گذاری).

مجموعه پارامترها:

در بیشتر موارد می‌توان β را برابر یک گرفت (وراسی و شمس، ۲۰۱۱). که در این تحقیق نیز این مقدار یک در نظر گرفته شده است.

در مورد پارامتر γ که تعیین کننده میزان جذب کنندگی نور بواسطه محیط است، مقدار آن در سرعت همگرایی و رفتار الگوریتم تأثیر بسزایی دارد. هر چه γ بزرگتر باشد حرکت کرم‌ها بیشتر تصادفی خواهد بود و هر چه این مقدار به سمت صفر برود شدت نور با فاصله کاهش نمی‌یابد و الگوریتم بسرعت به یک جواب بهینه همگرا خواهد شد. مقدار γ در بازه $[0.1, 0.01]$ است.

برای تعیین مقدار γ از روش تیونینگ استفاده شد. بدین صورت که چند عدد بصورت تصادفی در بازه $0/01$ و 100 انتخاب شده و با توجه به خروجی برنامه و مشاهده بهترین جواب‌ها سعی شد که مقدار تصادفی بعدی بیشتر به γ نزدیک باشد. جدول ۲ برخی از این مقادیر را نشان می‌دهد.

۷۴۹۲,۵

جدول ۲. تعیین مقدار پارامترها

مقادیر γ	خطای ردیابی دوره آموزش
$\gamma=0/2$	$7492/5 \times 10^{-4}$
$\gamma=0/8$	$5/5330 \times 10^{-4}$
$\gamma=1$	$5/6752 \times 10^{-4}$
$\gamma=2$	$5/2326 \times 10^{-4}$
$\gamma=3$	$5/6486 \times 10^{-4}$

تعیین مقدار γ

در جدول ۲ برای شاخص S&P500 مقادیر مورد آزمایش قرار گرفته‌اند و بعد از ۵ بار اجرای برنامه، میانگین این مقادیر در این جدول ثبت شده است. مقدار k نیز برابر ۱۰ در نظر گرفته شده است. چنان که مشاهده می‌کنید مقدار $\gamma = 2$ بهترین جواب را دارد لذا در این تحقیق این مقدار برابر ۲ در نظر گرفته شده است.

ارضای محدودیت‌ها

همانطور که توضیح داده شد، تابع ایجاد کننده جمعیت اولیه تمام محدودیت‌ها را ارضا کرد. در مرحله حرکت نیز برنامه به گونه‌ای نوشته شده است که محدودیت تعداد سهام پورتنفوی یعنی k تأمین خواهد شد. بمنظور برآورده ساختن محدودیت وزن سهام تشکیل دهنده پورتنفوی بعد از هر مرحله از حرکت، وزن هر یک از سهام را بر مجموع کل وزن‌های سهام تقسیم می‌کنیم تا بدین ترتیب این محدودیت نیز تأمین شود یعنی:

$$w_i = \frac{v_i}{\sum v_i} \quad \text{رابطه ۹}$$

در مورد محدودیت کف و سقف سرمایه‌گذاری می‌توان بعد از اجرای الگوریتم با توجه به دستور lsqin در نرم‌افزار متلب این محدودیت را برآورده کرد، تا هم این محدودیت ارضا شود و هم بهترین وزن‌ها به پورتنفوی اختصاص یابد.

شرط توقف الگوریتم

در این تحقیق مانند تحقیقات صورت گرفته پیشین این شرط برابر ۱۰۰ در نظر گرفته شده است.

تعداد کرم‌های شب‌تاب

کل جمعیت اولیه برای کرم‌های شب‌تاب برابر ۵۰ است.

خروجی الگوریتم

بعد از اجرای الگوریتم به جواب بهینه یا نزدیک به بهینه خواهیم رسید، میزان خطای ردیابی شاخص در آن کوچک خواهد بود.

شبه کد الگوریتم کرم شب‌تاب

B begin

$$\text{Objective function } f(x) = \frac{[\sum_{i=1}^n (\sum_{j=1}^m |x_{ij} - R_{ij}|^2)^2]}{T}$$

x $x_i = (x_{i1}, \dots, x_{in})^T$

Generate initial population of 50 fireflies x_i ($i = 1, 2, \dots, 50$)

Formulate light intensity I so that it is associated with $f(x)$

While (MaxGeneration=100)

Define absorption coefficient $\gamma = 2$

for $i=1:100$ (100 fireflies)

for $j=1:100$ (100 fireflies)

if ($I_j > I_i$),

move firefly i towards j

end if

Vary attractiveness with distance r via $\exp(-\gamma r^2)$

Evaluate new solutions and update light intensity

end for j

end for i

Rank the fireflies and find the current best

end while

Post-processing the results and visualization

E end

یافته‌های پژوهش

در این تحقیق از شاخص بورس آمریکا S&P500 و شاخص هنگ کنگ Hang Seng و نیز بازار سهام ایران TEPIX استفاده شده است. در مورد بورس‌های خارجی، داده‌های مربوط به ۲۹۱ هفته، از مارس ۱۹۹۲ تا دسامبر ۱۹۹۷ مورد استفاده قرار گرفته است. این داده‌ها از تنوع خوب و اعتبار بالایی برخوردارند که توسط دو تن از اساتید دانشگاه College London Imperial جمع‌آوری شده‌اند و در OR-Library در دسترس محققان هستند. داده‌های مربوط به بورس تهران از نرم‌افزار سازمان بورس، بصورت روزانه اخذ شد. بازه زمانی مورد نظر از فرودین سال ۱۳۸۷ تا اسفند سال ۱۳۸۸ و ۴۸۳ روز است. به دلیل عدم ثبات و نوسانات زیاد شاخص در این بازه زمانی، بازه زمانی هفته آخر مرداد ۱۳۸۱ تا هفته آخر خرداد ۱۳۸۲ طی ۲۰۵ روز نیز مورد مطالعه قرار گرفت. دلیل این انتخاب شرایط مناسب اقتصادی و سیاسی در آن دوره است.

با توجه به جدول ۳، با افزایش k میزان خطای دوره ردیابی و سنجش کاهش می‌یابد. این رویکرد منطقی بنظر می‌رسد چرا که با افزایش تعداد سهام پورتنوی، به دلیل تنوع در آن این سبد به شاخصی که متشکل از کل سهام بازار است نزدیک‌تر می‌شود. زمان هم با افزایش تعداد سهام بیشتر می‌شود که این نیز به دلیل افزایش تعداد محاسبات دور از انتظار نیست.

جدول ۳. ردیابی شاخص برای بازار سهام S&P500 و Hang Seng

شاخص	تعداد سهم در سبد (k)	خطای ردیابی (TE) دوره آموزش		خطای ردیابی (TE) دوره سنجش	زمان اجرا (ثانیه)
		ریشه میانگین مربع خطا	بهترین ریشه میانگین مربع خطا		

S&P500 N=۹۸	۵	$\times 10^{-4}$ ۷/۲۲۱	$۶/۶۰۷ \times 10^{-4}$	$۷/۳۰۵ \times 10^{-4}$	۲/۳۸۲
	۱۰	$۵/۱۸۷ \times 10^{-4}$	$۴/۷۰۶ \times 10^{-4}$	$۷/۱۷۸ \times 10^{-4}$	۲/۶۰۸
	۱۵	$۴/۴۹۰ \times 10^{-4}$	$۳/۸۱۱۹ \times 10^{-4}$	$۴/۶۲۰ \times 10^{-4}$	۲/۶۸۶
Hang seng N=۳۱	۵	$۶/۳۷۰ \times 10^{-4}$	$۶/۱۸۷ \times 10^{-4}$	$۵/۳۳۸ \times 10^{-4}$	۱/۳۷
	۱۰	$۴/۶۰۳ \times 10^{-4}$	$۴/۵۴۵ \times 10^{-4}$	$۳/۶۶۰ \times 10^{-4}$	۱/۳۴۱
	۱۵	$۳/۵۹۰ \times 10^{-4}$	$۳/۳۳۹ \times 10^{-4}$	$۳/۴۱۲ \times 10^{-4}$	۱/۸۲۶

نتایج پیاده‌سازی مدل در دو شاخص اول

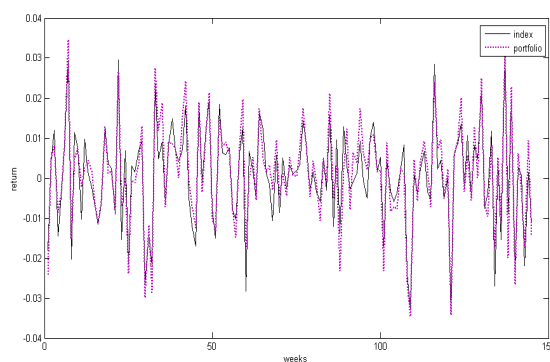
برای تست مدل در بازارهای داخلی بازه زمانی فروردین ۸۷ تا اسفند ۸۸ به مدت ۴۸۲ روز در نظر گرفته شده است. تعداد شرکت‌های انتخاب شده ۹۰ شرکت است. نتایج عددی این پیاده‌سازی در جدول ۴ آمده است.

جدول ۴. ردیابی شاخص برای بازار سهام ایران سال ۸۸-۸۷

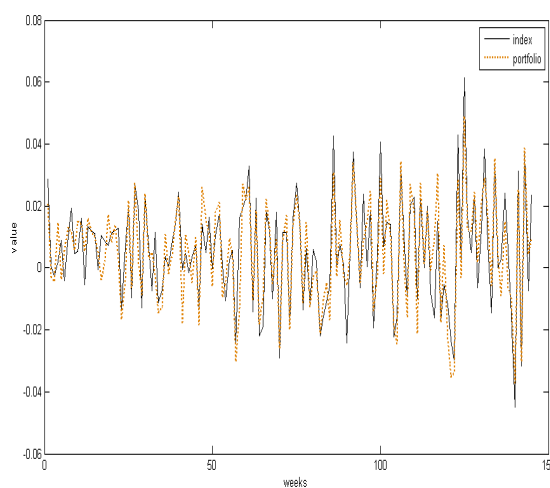
شاخص	(k) تعداد سهام در سبد	خطای ردیابی (TE) دوره آموزش		خطای ردیابی (TE) دوره سنجش	زمان اجرا (ثانیه)
		ریشه میانگین مربع خطا	بهترین ریشه میانگین مربع خطا		
بورس تهران ۸۷	۵	$۸/۵۰۴ \times 10^{-4}$	$۷/۵۶۶ \times 10^{-4}$	۰/۰۰۱۳	۴/۰۶۱۷
	۱۰	$۸/۳۴۶ \times 10^{-4}$	$۷/۰۴۴۸ \times 10^{-4}$	۰/۰۰۱۲	۴/۲۱۳۱
	۱۵	$۷/۹۰۳ \times 10^{-4}$	$۶/۵۱۱۶ \times 10^{-4}$	۰/۰۰۱۰	۵/۰۰۱۳
	۲۰	$۶/۸۲۵ \times 10^{-4}$	$۵/۷۵۰ \times 10^{-4}$	۰/۰۰۰۹	۹/۲۵
بورس تهران ۸۸	۵	$۹/۵۵ \times 10^{-4}$	$۸/۹۵۹ \times 10^{-4}$	$۸/۹۸ \times 10^{-4}$	۴/۲۶۲۶
	۱۰	$۷/۹۵۱ \times 10^{-4}$	$۷/۷۶۰ \times 10^{-4}$	$۷/۹۵ \times 10^{-4}$	۴/۵۲۲۶

	۱۵	$6/552 \times 10^{-4}$	$6/322 \times 10^{-4}$	$6/335 \times 10^{-4}$	۵/۱۲۲
	۲۰	$6/908 \times 10^{-4}$	$6/635 \times 10^{-4}$	$6/225 \times 10^{-4}$	۵/۸۸۶

در ادامه نمودارهای مربوط به ردیابی شاخص در دوره‌های آموزش و سنجش در شاخص S&P500 نمایش داده می‌شود.



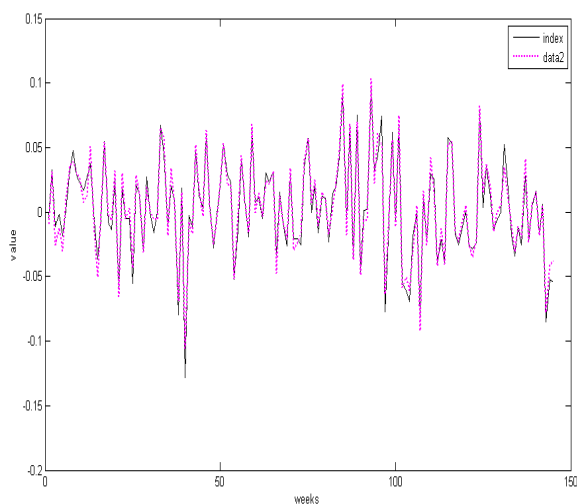
شکل (۱) ردیابی شاخص S&P500 با $k=15$ ، دوره آموزش



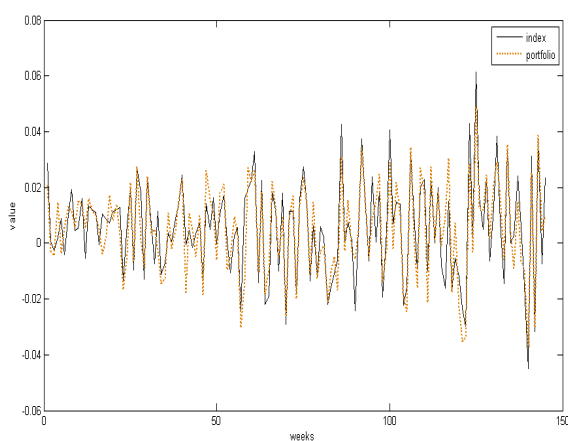
شکل (۲) ردیابی شاخص S&P500 با $k=15$ ، دوره سنجش

شکل‌های (۱) و (۲) ردیابی شاخص در دوره‌های آموزش و سنجش شاخص S&P500 در بازه زمانی ۱۴۵ هفته و تعداد سهام ۱۵ در سید ردیاب را نشان می‌دهد. خطای

ردیابی در دوره آموزش و سنجش به ترتیب برابر 6.187×10^{-4} و 4.620×10^{-4} است. ضریب همبستگی پیرسون، میان سری زمانی پورتفوی و شاخص در دوره آموزش ۹۵/۴٪ در سطح معناداری دو دامنه ۰/۰۱ محاسبه شده و این رقم در دوره سنجش ۹۳/۲٪ در سطح معناداری دو دامنه آن ۰/۰۱ تعیین شد که نشان دهنده همبستگی بالای پورتفوی انتخابی با شاخص است.

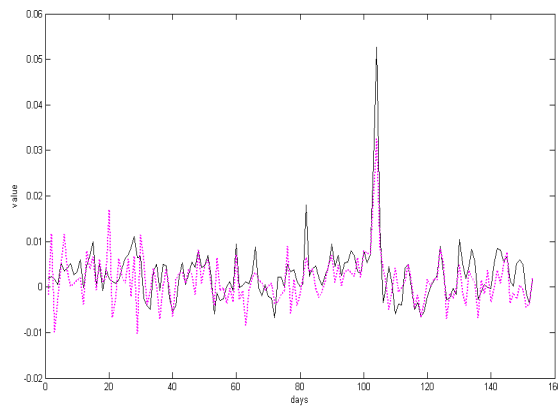


شکل ۳) ردیابی شاخص Hang Seng با $k=5$ آموزش

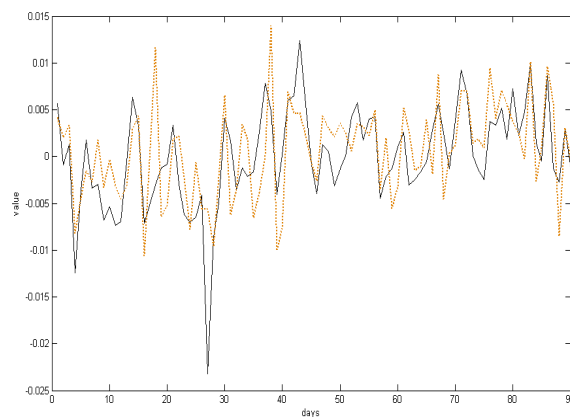


شکل ۴) ردیابی شاخص Hang Seng با $k=5$ دوره سنجش

در شکل‌های (۳) و (۴) شاخص هنگ‌سنگ برای تعداد ۵ سهم در پورتفوی بررسی شد. خطای ردیابی در دوره آموزش برابر 6.187×10^{-4} و خطا در دوره سنجش معادل 5.338×10^{-4} است. با توجه به شکل ۵، خروجی برنامه شاخص را بنحو مؤثری ردیابی می‌کند. همبستگی پیرسون میان سری زمانی پورتفوی و شاخص در دوره ردیابی برابر ۹۷/۷٪ و سطح معنی‌داری دو دامنه آن ۰/۰۱ است و این مقدار برای دوره سنجش برابر ۹۶/۳٪ در سطح معنی‌داری دو دامنه ۰/۰۱ محاسبه شد.



شکل ۵) ردیابی شاخص تهران در دوره آموزش، $k=15$



شکل ۶) ردیابی شاخص تهران سال ۸۸، دوره سنجش و $k=15$

نتیجه گیری

بطور کلی می توان اظهار داشت که سرمایه گذاری یکی از کلیدهای دستیابی به رشد بلندمدت اقتصاد است. سرمایه گذاری در صورتی منجر به رشد اقتصادی می شود که بتواند تخصیص بهینه منابع را فراهم ساخته و سبب افزایش کارایی سرمایه شود. امروزه در کشورهای توسعه یافته، روند تشکیل سبدهایی برای ردیابی شاخص های معروفی همچون S&P500، Dow Jones، FTSE، DAX، Nikkei و نظیر این ها سرعت گرفته و خرید و سرمایه گذاری منفعل در این سبدها با استقبال بسیار مردم روبرو شده است. شاهد این مدعا این است که حجم سرمایه گذاری در سبد ردیاب شاخص به بیش از ۱۵۰۰ میلیارد دلار تنها در ایالات متحده می رسد. علاوه بر این، ۵۰ درصد حجم سرمایه گذاری در بازار توکیو روی سبد ردیاب شاخص و نیز نزدیک به این مقدار در کشورهای نظیر انگلستان، کانادا و استرالیا نشان دهنده اقبال بسوی سرمایه گذاری در سبد ردیاب شاخص است. از این رو، سرمایه گذاری شاخص محور بعنوان یک روش سرمایه گذاری با سودی نزدیک به شاخص در این تحقیق مورد مطالعه قرار گرفت. این روش خصوصاً برای افرادی که به دلیل کمبود دانش لازم در سرمایه گذاری و عدم دسترسی به افراد خبره، انگیزه ای برای ورود به بازار بورس ندارند، روش کارایی محسوب می شود.

در این مقاله با استفاده از الگوریتم جدیدی که در سال ۲۰۰۸ معرفی شده است سعی در حل مسئله بهینه سازی مدیریت منفعل پورترفوی شد. دقت ردیابی الگوریتم و همبستگی بالای سبد سهام و شاخص، خصوصاً در بازارهای مالی با ثبات بالا، نشان از توانایی حل این مسئله توسط این الگوریتم دارد. الگوریتم کرم شب تاب برای مسائل پیوسته طراحی شده است.

منابع

- Alexander, C., & Dimitriu, A. (2005). Indexing, cointegration and equity market regimes. *International Journal of Finance & Economics*, 10(3), 213-231.
- Beasley, J. E., Meade, N., & Chang, T. J. (2003). An evolutionary heuristic for the index tracking problem. *European Journal of Operational Research*, 148(3), 621-643.
- Besley, S., & Brigham, E. (2011). *Principles of finance*. Cengage Learning.
- Fadaeinejad, E. (2002). Empirical Test of ETH in the Tehran Stock Exchange. *Journal of financial research*. 2 (5): 6-26. (in persian)
- Fengmin X. Zhaosong L. & Zongben X. (2015) An efficient optimization approach for a cardinality-constrained index tracking problem, *Optimization Methods and Software*. (010), 2205-2224.
- Fischer, D. E., & Jordan, R. J. (1975). *Security analysis and portfolio management*. Prentice Hall.
- Frankfurter, G. M., & McGoun, E. G. (1996). *Toward Finance with Meaning: The Methodology of Finance, what it is and what it Can be*. Jai Press.
- Gao, J., & Li, D. (2013). Optimal cardinality constrained portfolio selection. *Operations research*, 61(3), 745-761
- Giannakouris, G., Vassiliadis, V., & Dounias, G. (2010). Experimental study on a hybrid nature-inspired algorithm for financial portfolio optimization. In *Artificial Intelligence: Theories, Models and Applications* (pp. 101-111). Springer Berlin Heidelberg.
- Hanifi, F. Bahrololoum, M. & Javadi B. (2010). Design and comparative analysis of metaheuristic algorithms to implement index-tracking investment in Tehran Stock Exchange. *Journal of business management*. 9 (32): 89-108. (in persian)
- Jahankhani, A. Abdoh Tabrizi, H. (1994). Efficient Capital Market Hypothesis. *Journal of financial research*. 1 (1): 7-23. (in persian)
- Jahankhani, A. Parsanian, A. (1997). *Stock Exchange*. Tehran, University of Tehran press (UTP). (in persian)
- Krink, T., Mittnik, S., & Paterlini, S. (2009). Differential evolution and combinatorial search for constrained index-tracking. *Annals of Operations Research*, 172(1), 153-176.
- Maringer, D. G. (2006). *Portfolio management with heuristic optimization* (Vol. 8). Springer Science & Business Media.
- Ruiz-Torrubiano, R., & Suárez, A. (2009). A hybrid optimization approach to index tracking. *Annals of Operations Research*, 166(1), 57-71.
- Sharpe, W. F. (1991). The arithmetic of active management. *Financial Analysts Journal*, 47(1), 7-9.
- Sharpe, W. F., Alexander, G. J., & Bailey, J. V. (1999). *Investments* (Vol. 6). Upper Saddle River, NJ: Prentice-Hall.
- Sinaei, H. (1994). Performance measurement in Tehran Stock Exchange. *Journal of financial research*. 1 (2): 46-70. (in persian)
- Strong, R. (2008). *Portfolio construction, management, and protection*. Cengage Learning.
- Varasei, M. & Shams, N. (2011). Providing a heuristic solution to optimize index tracking problem and implementing it in Tehran Stock Exchange (TEPIX). *8th International Management conference*. Tehran.
- Wu, D. (2011). An Index Tracking Model: One Application of Integer Programming. In *Modeling Risk Management for Resources and Environment in China* (pp. 77-82). Springer Berlin Heidelberg.

