

Application of Copula-Based Correlation Coefficients and Dynamic Programming-Based Approaches to Determining Similarity between Time Series for Clustering and Index Tracking¹

Shapour Mohammadi², Farid Tondnevis³,
Elnaz Mahmoudi SaeidAbad⁴

Received: 2022/06/23

Accepted: 2022/11/26

Research Paper

Abstract

Index tracking, known as one of the passive management methods of investing, seeks to form portfolios in such a way that replicate the performance of market index. In this study, the application of a binary optimization model in time series clustering to form an index tracker portfolio is investigated. For the clustering process, various time series similarity measures including Copula-based correlation coefficients and dynamic distance based approaches have been used. Out-of-sample test on market ratios and tracking error of portfolios based on 50 More Active Company index of Tehran Stock Exchange in the period from beginning of 1394 till the end of spring 1397 indicates that portfolios have been successful in tracking index and average tracking error of portfolios did not differ significantly. Also, pairwise comparisons test on portfolio tracking error indicates that portfolios tracking error does not differ significantly.

Key Words: Index Tracking, Tracking Error, Copula, DTW, EDR.

JEL Classification: G11.

1. DOI: 10.22034/JSE.2020.11394.1562

2. Associate Professor, Department of Finance and Insurance, Faculty of management, Tehran University. Tehran, Iran. (shmohmad@ut.ac.ir).

3. Ph.D. Department of Finance and Insurance, Faculty of Management, Tehran University. Tehran, Iran. (Corresponding Author). (farid.tondnevis@ut.ac.ir).

4. M.Sc. Department of Industrial Engineering, Faculty of Engineering, Alzahra University, Tehran, Iran. (mahmoudi.fe88@gmail.com).



سازمان بورس و اوراق بهادار، مرکز پژوهش، توسعه و مطالعات اسلامی

فصلنامه بورس اوراق بهادار، سال پانزدهم، شماره ۶۰، زمستان ۱۴۰۱، صص ۷۲-۴۷

کاربرد ضرایب همبستگی مبتنی بر کاپولا و رویکردهای مبتنی بر برنامه‌ریزی پویا در تعیین شباهت میان سری‌های زمانی به منظور خوشه‌بندی و تشکیل پرتفوی مبتنی بر شاخص^۱

شاپور محمدی^۲، فرید تندنویس^۳، الناز محمودی سعیدآباد^۴

تاریخ دریافت: ۱۴۰۱/۰۴/۰۲

تاریخ پذیرش: ۱۴۰۱/۰۹/۰۵

مقاله پژوهشی

چکیده

ردیابی شاخص که به عنوان یکی از روش‌های مدیریت غیرفعال سرمایه‌گذاری شناخته شده است، به دنبال تشکیل پرتفوی، به گونه‌ای است که در طول زمان با کمترین میزان خطا، بازدهی نزدیک به شاخص داشته باشد. در این پژوهش به بررسی کاربرد یک مدل برنامه‌ریزی صفر و یک در خوشه‌بندی سری‌های زمانی به منظور تشکیل پرتفوی ردیاب شاخص پرداخته شده است. به منظور فرایند خوشه‌بندی، معیارهای متنوع سنجش شباهت سری‌های زمانی از جمله، ضرایب همبستگی استوار بر کاپولا همچنین فاصله استوار بر رویکردهای پویا در سنجش شباهت سری‌های زمانی، مورد استفاده قرار گرفته است. آزمون خارج از نمونه بر روی نسبت بازار و خطای ردیابی پرتفوها استوار بر شاخص ۵۰ شرکت فعال‌تر بورس اوراق بهادار تهران در بازه زمانی ابتدای سال ۱۳۹۴ تا پایان بهار سال ۱۳۹۷ نشان از این موضوع دارد که پرتفوها در ردیابی شاخص موفق عمل کرده‌اند و متوسط خطای ردیابی روزانه پرتفوها دارای تفاوت معنی‌داری با یکدیگر نیست. همچنین آزمون سخت‌گیرانه مقایسات زوجی بر روی خطای ردیابی پرتفوها نیز نشان از این موضوع دارد که خطای ردیابی پرتفوها با یکدیگر تفاوت معنی‌دار ندارد.

واژه‌های کلیدی: ردیابی شاخص، خطای ردیابی، کاپولا، EDR و DTW.

طبقه‌بندی موضوعی: G11.

DOI: 10.22034/JSE.2020.11394.1562

۲. دانشیار، گروه مالی و بیمه، دانشکده مدیریت، دانشگاه تهران، تهران، ایران. (shmohmad@ut.ac.ir)

۳. دکتری، گروه مدیریت مالی و بیمه، دانشکده مدیریت، دانشگاه تهران، تهران، ایران. (نویسنده مسئول). (farid.tondnevis@ut.ac.ir)

۴. کارشناسی ارشد، گروه مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه الزهراء، تهران، ایران. (mahmoudi.fe88@gmail.com)

مقدمه

روش‌های مدیریت سرمایه‌گذاری شامل یک طیف گسترده هستند که رویکردهای غیرفعال در ابتدای آن رویکردهای فعال در سوی دیگر این طیف قرار دارند. ردیابی شاخص، یکی از پرکاربردترین روش‌هایی است، که در ابتدای طیف یادشده قرار گرفته‌است. به طور کلی رویکرد غیرفعال که براساس فرض کارایی اطلاعاتی بازار سرمایه استوار است با ایجاد محدودیت در میزان انعطاف‌پذیری عملیاتی مدیر، ضوابط مشخصی را برای مدیر در نظر می‌گیرد. ردیابی شاخص به عنوان یکی از رویکردهای منفعلانه مدیریت سرمایه‌گذاری، به دنبال تشکیل پرتفوی به گونه‌ای است که با سرمایه‌گذاری در تعدادی از سهم‌های تشکیل‌دهنده یک شاخص، عملکرد شاخص را بازسازی کند.

بدون تردید هرچه تعداد سهم‌های تشکیل‌دهنده پرتفوی، از سهام‌های تشکیل‌دهنده شاخص کمتر باشد، عملکرد پرتفوی از شاخص فاصله می‌گیرد. کمینه کردن این اختلاف با استفاده از یک مدل پیچیده بهینه‌سازی ریاضی می‌تواند به عنوان یکی از روش‌های تشکیل پرتفوی استوار بر شاخص مورد استفاده قرار گیرد. اما این مدل‌ها با توجه به نیازشان به تخمین پارامترهای متعدد و پیش‌بینی آنها، به خصوص در زمانی که در سائز واقعی مورد استفاده قرار گیرند، دارای پیچیدگی زمانی و محاسباتی بالایی خواهند بود. استفاده مستقیم از مدل میانگین واریانس و همچنین استفاده غیر مستقیم از مدل‌های عاملی که در ادبیات موضوع به عنوان روش‌های ردیابی شاخص معرفی شده‌اند و در قسمت بعدی مورد بررسی قرار می‌گیرند نیز، پژوهشگران را با محدودیت‌های بالا روبرو خواهند کرد.

یکی دیگر از رویکردهای مورد استفاده برای تشکیل پرتفوی استوار بر شاخص، رویکرد خوشه‌بندی سری‌های زمانی است. در این رویکرد، سهم‌هایی از شاخص که دارای روند عملکردی مشابه هستند، در یک گروه قرار می‌گیرند و از هر گروه یک نماینده برای تشکیل پرتفوی انتخاب می‌شود. روشن است که این روش، تنها به تخمین میزان شباهت میان سری‌های زمانی به منظور خوشه‌بندی نیاز دارد و در نتیجه پیچیدگی محاسباتی رویکردهای قبلی را نخواهد داشت.

این پژوهش به ارائه یک مدل بهینه‌سازی صفر و یک، برای خوشه‌بندی سهم‌های تشکیل‌دهنده شاخص و در نتیجه تشکیل پرتفوی ردیاب شاخص پرداخته‌است. خوشه‌بندی، بر اساس معیارهای شباهت متنوعی انجام می‌شود که در این میان می‌توان به رویکردهای استوار بر

برنامه‌ریزی پویا (مانند DTW و EDR) و رویکردهای استوار بر کاپولا اشاره کرد. پژوهش حاضر به دنبال پاسخ‌دهی به این پرسش است که کدام یک از رویکردهای مورد استفاده در تعیین شباهت میان سری‌های زمانی، به تشکیل پرتفوی ردیاب شاخص با عملکرد بهتری منتج می‌شود؟ و این پرسش با آزمون مقایسه خطای ردیابی پرتفوهای حاصل، پاسخ داده می‌شود. بدون تردید، خطای ردیابی پایین‌تر، نشان‌دهنده برتری پرتفوی ایجاد شده در ردیابی شاخص است.

در ادامه متن حاضر، در قسمت بحث و بررسی به بیان مساله و اهمیت آن و همچنین، مروری بر پیشینه پژوهش و مبانی نظری پرداخته می‌شود. همچنین پرسش‌ها و فرضیه‌های پژوهش در این قسمت معرفی شده و چگونگی پاسخ‌دهی به آنها در بخش روش پژوهش و فنون تجزیه و تحلیل اطلاعات مورد اشاره قرار خواهند گرفت. در ادامه پس از ارائه قسمت بحث در یافته‌های پژوهش که نتایج آزمون فرضیه‌ها را توسعه می‌دهد، بخش نتیجه‌گیری و ارائه پیشنهادها مورد بررسی قرار می‌گیرد.

بحث و بررسی

بیان مساله و اهمیت آن

ردیابی شاخص، به عنوان یکی از روش‌های غیرفعال در مدیریت پرتفوی، از اهمیت بالایی در ادبیات مالی برخوردار است. این پژوهش به ارائه یک مدل بهینه‌سازی صفر و یک به منظور خوشه‌بندی سری‌های زمانی، و تشکیل پرتفوی استوار بر شاخص می‌پردازد. مهم‌ترین موضوع در خوشه‌بندی سری‌های زمانی تعیین معیار سنجش شباهت در فرایند خوشه‌بندی است. مهم‌ترین هدف این پژوهش بررسی معیارهای سنجش میزان شباهت پویا (مانند DTW، EDR) و ضرایب همبستگی استوار بر کاپولا منظور خوشه‌بندی سری‌های زمانی و تحلیل تاثیر استفاده از آنها در مدل تشکیل پرتفوی استوار بر شاخص است.

مروری بر پیشینه پژوهش

از آنچه که در بررسی ادبیات موضوع به دست می‌آید می‌توان روش‌های ردیابی شاخص و تشکیل پرتفوی ردیاب را می‌توان به ۴ گروه کلی، طبقه‌بندی کرد. طبقه اول، به طور مستقیم به کمینه‌سازی اختلاف میان عملکرد پرتفوی و شاخص - که با عنوان خطای ردیابی معروف است

می‌پردازند (چن و روی^۱، ۲۰۱۲). در مورد خطای ردیابی تعاریف گوناگونی در ادبیات موضوع ارائه شده است. برخی خطای ردیابی را به عنوان قدرمطلق اختلاف بازده مورد انتظار شاخص پرتفوی معرفی کرده‌اند (کراس، استیتمن، کلارک^۲، ۱۹۹۴؛ میهوا^۳ و همکاران، ۲۰۱۱). برخی دیگر از پژوهش‌ها واریانس اختلاف بین بازده شاخص و پرتفوی را به عنوان معیار خطای ردیابی معرفی کرده‌اند (مید و سالکین^۴، ۱۹۸۹؛ مید و سالکین، ۱۹۹۰؛ رول و رونالد^۵، ۲۰۱۲). پژوهشگران دیگری به معرفی مجموع مجذور اختلاف بازده شاخص و پرتفوی به عنوان خطای ردیابی و پرتفوی پرداختند (بیسلی^۶ و همکاران، ۲۰۱۲). همانطور که در تعریف خطای ردیابی به عنوان واریانس اختلاف بین بازده پرتفوی، شاخص و یا مجذور این اختلاف روشن است، مساله کمینه‌سازی خطای ردیابی یک مساله برنامه‌ریزی درجه دوم است از نظر محاسباتی پیچیدگی‌های خاص خود را دارد. برخی پژوهشگران به بررسی رویکردهای ابتکاری برای تحلیل مساله پرداختند و در واقع بر روی پیچیدگی محاسباتی مساله تمرکز کردند (بیسلی و همکاران، ۲۰۱۲). گیلی و کلزی (۲۰۰۲) نیز به بررسی الگوریتم ابتکاری پذیرش آستانه‌ای پرداختند و این الگوریتم را به عنوان رویکردی موثر در ردیابی شاخص معرفی کردند (گیلی و کلزی^۷، ۲۰۰۲).

کنو و ویجایانایاکا (۲۰۰۱) به جای واریانس، به کمینه‌سازی قدرمطلق اختلافات بازده شاخص و بازده پرتفوی با استفاده از الگوریتم شاخه و کران پرداختند (کنو و ویجایانایاک^۸، ۲۰۰۱). سانت آنا^۹ و همکاران (۲۰۱۷) با اشاره به پیچیدگی زمانی الگوریتم‌های استاندارد تحلیل مساله برنامه‌ریزی درجه دوم به ارائه یک الگوریتم ابتکاری مبتنی بر الگوریتم ژنتیک پرداختند و سه شاخص اس اندپی ۵۰۰، فتسی ۱۰۰ و دکس را با موفقیت و در زمانی مناسب‌تر از الگوریتم‌های استاندارد ردیابی کردند (سانت آنا و همکاران، ۲۰۱۷). گارسیا و همکاران (۲۰۱۷) به مقایسه دو الگوریتم فراابتکاری جستجوی ممنوع و الگوریتم ژنتیک در تشکیل پرتفوی ردیاب شاخص استوار بر ۷ شاخص واقعی پرداختند و برتری الگوریتم جستجوی ممنوع را در

1. Chen & Roy
2. Krase, Statman, Clarke
3. Meihua
4. Meade & Salkin
5. roll & Ronald
6. Beasley
7. Gilli & Kellezi
8. Konno & Wijayanayake
9. Sant'Anna

این خصوص اثبات کردند (گارسیا، گوئیجارو و اُلور^۱، ۲۰۱۷). رودلف، مارکوس و پورگن^۲ (۱۹۹۹) نیز به معرفی معیارهایی خطی برای اندازه گیری خطای ردیابی پرداختند و با استفاده از برنامه ریزی خطی مقدار خطای ردیابی را کمینه کردند و جواب مساله را استخراج کردند (رودلف، مارکوس و پورگن، ۱۹۹۹). سانت آنا و همکاران (۲۰۱۷) به تحلیل مدل با استفاده از رویکردهای همبستگی و هم انباشتگی پرداختند و عملکرد مشابه این دو رویکرد را در تشکیل پرتفوی استوار بر شاخص بر اساس اطلاعات بازار سرمایه کشور برزیل نشان دادند (سانتا آنا و همکاران، ۲۰۱۷). عیوض لو، شفیق زاده و قهرمانی (۱۳۹۶) نیز با مقایسه این دو رویکرد بر اساس اطلاعات بازار سرمایه ایران، برتری مدل استوار بر هم انباشتگی در تشکیل پرتفوی ردیاب شاخص را اثبات کردند. سجادی و سبزواری (۱۳۹۴) به تحلیل مساله با استفاده از الگوریتم کرم شبتاب پرداختند و پیچیدگی زمانی چند جمله ای برای تحلیل مساله ایجاد کردند. امیری، کرمی و ناصر پور (۱۳۹۵) نیز با استفاده از الگوریتم بیگ بنگ، بیگ کرانچ و الگوریتم دیفرانسل تکاملی به تحلیل مساله بالا پرداختند.

رویکرد دیگری که در ادبیات به آن اشاره شده است، استفاده از مدل های عاملی به منظور ایجاد همخوانی بین مشخصه های پرتفوی ردیاب و شاخص است (چن و روی، ۲۰۱۲). مدل های عاملی بازده یک سهم را به یک یا چند عامل اقتصادی مربوط می سازند. برای مثال یک مدل تک شاخصی به منظور به دست آوردن بازده سهم i از یک مدل رگرسیون خطی استفاده می کند (بیسلی، مید و چانگ^۳، ۲۰۰۳) به این ترتیب که:

$$r_i = \alpha_i + \beta_i \cdot r_M + \varepsilon \quad \text{رابطه (۱)}$$

بتای سهم i شاخص ارزیابی ریسک سیستماتیک است. در علوم مالی ریسک را به دو دسته سیستماتیک و غیر سیستماتیک تقسیم بندی می کنند. ریسک غیر سیستماتیک به مشخصات شرکت باز می گردد و ریسک سیستماتیک نیز به مشخصات کلی بازار مربوط است. یک مدل کمینه سازی واریانس تحت شرایطی که بتای پرتفوی برابر با یک باشد، می تواند به عنوان یک روش تشکیل پرتفوی ردیاب مورد استفاده قرار گیرد. از آنجا که پرتفوی ردیاب باید یکسان با شاخص حرکت کند و ضریب رگرسیون شاخص نسبت به خودش برابر ۱

1. Garcia, Gujjarro and Oliver

2. Rudolf, Markus & Jurgen

3. Beasley, Meade & Change

خواهد بود (رود^۱، ۱۹۸۰). اردوگان^۲ و همکاران (۲۰۰۴) با در نظر گرفتن چندین عامل و هزینه‌های معاملاتی به پیشینه‌سازی شاخص شارپ پرداختند (اردوگان و همکاران، ۲۰۰۴). این پژوهشگران به منظور برآورد دقیق‌تری از بازده دارایی‌ها از مدل قیمت‌گذاری دارایی‌های سرمایه‌ای (شارپ^۳، ۱۹۶۳) استفاده کردند. کوریلی و مارسلینو (۲۰۰۶) با ارائه یک مدل عاملی، پویایی شاخص که به دلیل تغییر قیمت سهم‌ها و وزن آنها در شاخص اتفاق می‌افتد را در نظر گرفته و با استفاده از این مدل پویا به تحلیل مساله ردیابی شاخص پرداختند (کوریلی و مارسلینو^۴، ۲۰۰۶). جنگ^۵ و همکاران (۲۰۱۳) با استفاده از یک مدل چندعاملی برای انتخاب سهم‌ها و تکنیک ادهاک برای وزن‌دهی، به تشکیل پرتفوی شاخصی ارتقایافته استوار بر شاخص ۵۰ شرکت شانگهای پرداختند و اثبات کردند که این مدل به نسبت اطلاعاتی بالاتر نسبت به رویکردهای استاندارد منتج می‌شود (جنگ و همکاران، ۲۰۱۳). فلاح‌پور، تندنیس و هاشمی (۱۳۹۴) با استفاده کمینه‌سازی واریانس جزء اخلاص به تشکیل پرتفوی با بتای نزدیک به ۱ پرداختند و عدم قطعیت موجود در پارامترهای مدل را با استفاده از رویکرد بهینه‌سازی استوار تحلیل کردند.

روش دیگری که برای ردیابی شاخص مورد استفاده قرار می‌گیرد، روش میانگین واریانس مارکوویتز است. رچارد رول (۱۹۹۲) با استفاده از رویکرد بهینه‌سازی مارکوویتز^۶ (۱۹۵۹) به تشکیل پرتفوی ردیاب شاخص پرداخته و راودر (۱۹۹۸) هزینه‌های معاملاتی را نیز به این مدل اضافه کرد (روودر^۷، ۱۹۹۸). روی و دخیانگ (۲۰۱۶) با ارائه مدل پایدار ۳ عاملی فاما و فرنچ به ترکیب مدل‌های عاملی و رویکرد میانگین واریانس برای تشکیل پرتفوی شاخصی افزوده پرداختند و با بازنویسی مدل و تبدیل آن به یک مدل بهینه‌سازی ترکیبی مخروطی درجه دوم نشان دادند که عملکرد این مدل پایدار نسبت به مدل‌های استوار بر شاخص شارپ بهتر خواهد بود (روی و دکسیانگ^۸، ۲۰۱۶). این رویکرد به منظور تشکیل پرتفوی شاخصی بهبود یافته نیز مورد بهره‌برداری قرار می‌گیرد (انصاری، بهزادی و تندنیس؛ ۱۳۹۸). صندوق شاخصی بهبود یافته از یکسو در تلاش برای دستیابی به بازده بیشتر نسبت به شاخص است و از سوی دیگر، برای کاهش انحراف از شاخص تلاش می‌کند (حجازی، جعفری، دلشادی، ۱۳۹۰).

1. Rudd
2. Erdogan
3. Sharpe
4. Corielli & Marcellino
5. jeng
6. Markowitz
7. Rohweder
8. Roy & Dexiang

رویکرد خوشه‌بندی سری‌های زمانی برای تشکیل پرتفوی استوار بر شاخص نیز مورد توجه بسیاری از پژوهشگران قرار گرفته است. فوکاردی و فبوزی^۱ (۲۰۰۴) به ارائه مدل خوشه‌بندی بر پایه هم‌جمعی پرداختند و پرتفوی استوار بر شاخص را تشکیل دادند. داز و کینوتی (۲۰۰۵) با استفاده از رویکرد برنامه‌ریزی تصادفی، به تحلیل مساله ردیابی شاخص با استفاده از خوشه‌بندی سری‌های زمانی پرداخته‌اند (داس و سینکاتی^۲، ۲۰۰۵). کورنجلوس و توتونچی^۳ (۲۰۰۵) مدل بهینه‌سازی صفر و یک را برای خوشه‌بندی سری‌ها زمانی بر مبنای ضریب همبستگی پیرسون را معرفی کردند (کورنجلوس و توتونچی، ۲۰۰۵)، فلاچ‌پور و تندنویس (۱۳۹۴) و امیری و شهسوار و چالسپاری (۱۳۹۹) عدم قطعیت ضریب همبستگی به عنوان پارامتر این مدل را با استفاده از رویکرد بهینه‌سازی استوار مورد بررسی قرار دادند.

ادبیات و چارچوب نظری

این بخش از پژوهش به معرفی شاخص‌های مورد استفاده برای سنجش میزان شباهت در سری‌های زمانی می‌پردازد:

رویکرد DTW

رویکرد پیچش زمانی پویا بهترین نگاشت بین دو سری زمانی را با توجه به تغییرات پویای زمان، مورد محاسبه قرار می‌دهد. دو سری زمانی به صورت $x(i)$ برای $i = \{1, \dots, m\}$ و $y(j)$ برای $j = \{1, \dots, n\}$ را در نظر بگیرید. برای این دو سری، ماتریس $D_{m \times n}$ با اجزای $d(i, j)$ مورد محاسبه قرار می‌گیرد. $d(i, j)$ به عنوان فاصله محلی دو جزء $x(i)$ و $y(j)$ معرفی می‌شود. به طوری که:

$$d(i, j) = (x(i) - y(j))^2 \quad \text{رابطه (۲)}$$

مسیر پیچشی W به عنوان نگاشتی بین x و y به صورت زیر معرفی می‌شود:

$$W = \begin{pmatrix} w_x(k) \\ w_y(k) \end{pmatrix} \quad k = 1, 2, \dots, p \quad \text{رابطه (۳)}$$

در رابطه بالا $w_x(k)$ و $w_y(k)$ به ترتیب مرتبه اجزا در سری زمانی x و y را نشان می‌دهند. به عبارتی $(w_x(k) \text{ و } w_y(k))'$ نشان دهنده این موضوع است که $w_x(k)$ امین جزء سری زمانی x با $w_y(k)$ امین جزء سری زمانی y هم تراز شده است. این مسیر پیچشی W باید در شرایط زیر صدق کند:

1. Focardi & Fabozzi
2. Dos & Cincotti
3. Cornuejols, Tutuncu

- شرایط مرزی: یک مسیر پیچشی باید از نقطه $W(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ آغاز شده و در نقطه $W(p) = \begin{pmatrix} m \\ n \end{pmatrix}$ پایان یابد. به عبارتی نقاط ابتدایی و انتهایی مسیرها متناظر هستند.
- شرایط پیوستگی مسیر: اجزای مسیر W باید به گونه‌ای قرار گیرند که:

$$w_x(k+1) - w_x(k) \leq 1$$

$$w_y(k+1) - w_y(k) \leq 1$$

رابطه (۴)

به عبارتی این شرط بیان می‌کند که در هر پیش‌روی مسیر، هر یک از ابعاد حداکثر یک واحد تغییر خواهند کرد.

- شرط یکنواختی: اجزای مسیر W باید به گونه‌ای قرار گیرند که:

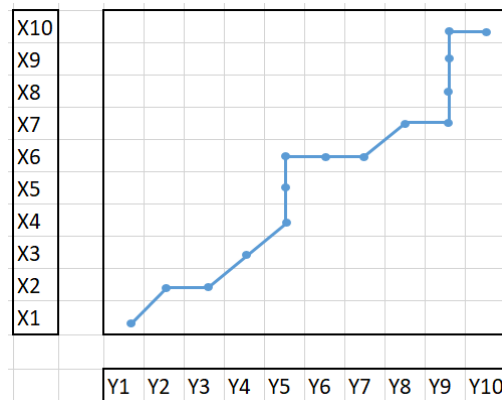
$$w_x(k+1) - w_x(k) \geq 0$$

رابطه (۵)

$$w_y(k+1) - w_y(k) \geq 0.$$

به عبارتی این شرط بیان می‌کند که در هر گام مسیر، هیچ یک از سری‌ها حرکت به سمت عقب نخواهند داشت.

تصویر زیر یک مسیر پیچشی را به صورت نمادین نشان می‌دهد:



تصویر ۱. نمونه‌ای از یک نگاشت بین دو سری x و y

همانطور که در تصویر و توضیحات بالا مشخص است، بی‌شمار مسیر بین دو سری زمانی X و Y قابل معرفی است. در رویکرد پیچش زمانی پویا، مسیری با کمترین مسافت بین X و Y به عنوان مسیر بهینه معرفی می‌شود. این مسیر با استفاده از تکنیک برنامه‌ریزی پویا به شکل زیر قابل یافتن است:

$$\begin{cases} r(i, j) = d(i, j) + \min\{r(i-1, j-1), r(i-1, j), r(i, j-1)\} \\ DTW(x, y) = \min\{r(m, n)\} \end{cases} \quad \text{رابطه (۶)}$$

در رابطه برنامه‌ریزی پویا بالا $r(i, j)$ نشان‌دهنده کمترین فاصله تجمعی میان نقاط (۰ و ۰) و (i و j) است. اگر با تحلیل مساله برنامه‌ریزی پویای بالا بتوان مسیر بهینه W را مشخص کرد، دو سری زمانی اولیه X و Y به دو سری زمانی جدید تبدیل می‌شوند:

$$\begin{cases} \bar{x}_k = x(w_x(k)) \\ \bar{y}_k = y(w_y(k)) \end{cases} \quad k = 1, 2, \dots, p \quad \text{رابطه (۷)}$$

حال می‌توان با محاسبه فاصله اقلیدسی دو سری زمانی جدید بالا، شاخص DTW را به عنوان شاخص شباهت دو سری زمانی مورد محاسبه قرار داد (برنت و کلیوفورد^۱، ۱۹۹۴؛ شن^۲ و همکاران، ۲۰۱۶):

$$DTW(x, y) = \sum_{k=1}^p (\bar{x}(k) - \bar{y}(k))^2 \quad \text{رابطه (۸)}$$

رویکرد EDR

رویکرد EDR یکی از رویکردهای نوینی است که می‌تواند رویکرد DTW را با چالش جدی روبرو سازد. محاسبه EDR می‌تواند با استفاده از رویکرد برنامه‌ریزی پویا به شکل زیر انجام شود:

رابطه (۹)

$$D(i, j) = \begin{cases} D(i-1, j-1) & \text{if } (m(x(i), y(i)) = 1) \\ 1 + \min\{D(i-1, j-1), D(i-1, j), D(i, j-1)\} & \text{if } (m(x(i), y(i)) = 0) \end{cases}$$

$$EDR(x, y) = \sum_{k=1}^p (\bar{x}(k) - \bar{y}(k))^2$$

که در رابطه بالا:

1. Berndt & Clifford
2. Shen

$$m(x(i).y(i)) = \theta(\varepsilon - f(x(i).y(i))) \quad \text{رابطه (۱۰)}$$

تابع θ به صورتی بیان می‌شود که اگر z مقداری مثبت و یا صفر را به خود تخصیص دهد، $\theta(z) = 1$ خواهد بود. در غیر این صورت $\theta(z) = 0$. (وانگ^۱ و همکاران، ۲۰۱۳؛ چن و همکاران، ۲۰۰۵؛ سِرا و اِکروس^۲، ۲۰۱۴).

مقدار ε هم آستانه‌ای است که اختلاف فاز سری‌های زمانی را به عنوان ورودی مورد کنترل قرار می‌دهد.

رویکرد کاپولا

این رویکرد می‌تواند توزیع حاشیه‌ای سری‌های زمانی مختلف را به توزیع مشترک آنها مرتبط سازد تا همبستگی میان دو متغیر تصادفی را تشریح کند. بر اساس تئوری اسکالر برای توزیع مشترک دو بعدی $F(x_1, x_2)$ برای دو متغیر تصادفی X_1 و X_2 با توزیع‌های حاشیه‌ای $F(x_1)$ و $F(x_2)$ یک تابع کاپولا C وجود دارد که:

$$F(x_1, x_2) = C[F_1(x_1), F_2(x_2)] \quad \text{رابطه (۱۱)}$$

با مشتق گرفتن از دو طرف معادله بالا داریم:

$$\begin{aligned} \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} &= \frac{\partial^2 C[F_1(x_1), F_2(x_2)]}{\partial F_1 \partial F_2} f_1(x_1) \cdot f_2(x_2) \\ &= \frac{\partial^2 C[u_1, u_2]}{\partial u_1 \partial u_2} \times \prod_i \frac{\partial F_i(x_i)}{\partial x_i} \\ &= c(\tilde{u}) \times \prod_i f_i(x_i) \end{aligned} \quad \text{رابطه (۱۲)}$$

در رابطه بالا، f_i نشان‌دهنده تابع چگالی $F_i(x_i)$ و u_i برای هر $i = 1, 2$ و $\tilde{u} = (u_1, u_2)$ و $c(\tilde{u})$ تابع چگالی کاپولا است. هنگامی که متغیرهای پیوسته باشند، برپایه قضیه اسکالر هر تابع توزیع احتمال چند متغیره می‌تواند با یک توزیع حاشیه‌ای و یک ساختار وابستگی به صورت معادله بالا نشان داده شود.

اگر همه حاشیه‌ها پیوسته باشند، کاپولا یکتاست و به صورت یکتا می‌توان آن را با استفاده از مقادیر مختلف تابع حاشیه‌ای بیان کرد. مهم‌ترین ویژگی نتیجه بیان شده این است که هیچ الزامی مبنی بر تشابه توزیع‌های حاشیه‌ای وجود ندارد.

1. Wang

2. Serra & Acros

این تئوری، پژوهشگران را قادر می‌سازد که بتوانند با توابع کاپولای متنوع توزیع حاشیه‌ای متغیرهای تصادفی را با توزیع مشترک آنها مرتبط سازند و از طریق این ارتباط، الگوهای همبستگی میان این متغیرها را تشخیص دهند (پاتون^۱، ۲۰۰۷).

کاپولای نرمال

سونگ^۲ (۲۰۰۰) تابع توزیع خانواده کاپولای نرمال را به صورت زیر معرفی کرد (سانگ^۳، ۲۰۰۰):

$$C^{Ga}(u_1, u_2; \rho) = \Psi_p(\Psi^{-1}(u_1), \Psi^{-1}(u_2)) \quad \text{رابطه (۱۳)}$$

که در رابطه بالا Ψ_p تابع توزیع نرمال استاندارد دو متغیره، با ضریب همبستگی $\rho \in (0, 1)$ است. تابع چگالی کاپولای گوسین به صورت زیر معرفی می‌شود:

$$C^{Ga}(u_1, u_2, \dots, u_n; \Sigma) = |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \xi' (\Sigma^{-1} - I) \xi\right) \quad \text{رابطه (۱۵)}$$

کاپولای تی استیودنت

امبرجت و همکاران (۲۰۰۲) تابع توزیع تابع کاپولای تی استیودنت را به صورت زیر معرفی کردند (امبرچتس^۳، ۲۰۰۲):

$$C^T(u_1, u_2; \rho) = T_{v, \rho}(t_v^{-1}(u_1), t_v^{-1}(u_2)) \quad \text{رابطه (۱۶)}$$

در اینجا $T_{v, \rho}$ توزیع دو متغیره تی استیودنت، ρ ضریب همبستگی، v درجه آزادی و t_v^{-1} معکوس توزیع تک متغیره تی استیودنت است. امبرجت و همکاران نشان دادند که تابع کاپولای تی استیودنت بیانگر وابستگی دنباله بالایی و پایینی به طور همزمان است. تابع چگالی کاپولای تی استیودنت به شرح زیر قابل ارائه است (استولتر^۳، ۲۰۰۸):

$$C_v(u_1, u_2, \dots, u_n; \Sigma) \quad \text{رابطه (۱۷)}$$

$$= K |\Sigma|^{\frac{1}{2}} \left(1 + v^{-1} \xi \Sigma \xi \right)^{-(v+2)/2} \prod_{i=1}^n \left(1 + v^{-1} \xi_i^2 \right)^{(v+1)/2}$$

1. Patton
2. Song
3. Embrechts

در تابع بالا، K نشان‌دهنده تابع گاما و به صورت زیر است:

$$K = \Gamma\left(\frac{v}{2}\right)^{n-1} \Gamma\left(\frac{v+1}{2}\right)^{-n} \Gamma\left(\frac{v+n}{2}\right) \quad \text{رابطه (۱۸)}$$

کاپولاهای ارشمیدسی

کاپولای ارشمیدسی یک دسته مهم از توابع کاپولا با ساختار ساده و خصوصیات تحلیلی بسیاری است. تابع کاپولای ارشمیدسی دو متغیره به صورت $C(u_1, u_2) = \varphi^{[-1]} \{ \varphi(u_1) + \varphi(u_2) \}$ است که پیوسته، بی چون و چرا کاهشی و دارای تابع مولد $\varphi: [0, 1] \rightarrow [0, \infty]$ به طوری که $\varphi(1) = 0$ و تابع شبه معکوس $\varphi^{[-1]}$ به صورت معادله زیر است (کشاورز حداد، حیرانی، ۱۳۹۳):

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & 0 \leq t \leq \varphi(0) \\ 0 & \varphi(0) \leq t \leq \infty \end{cases} \quad \text{رابطه (۱۹)}$$

سه نوع کاپولای ارشمیدسی به صورت رایج مورد استفاده قرار می‌گیرند که عبارتند از، کاپولای کلایتون، کاپولای فرانک و کاپولای گامبل که در ادامه مورد اشاره قرار خواهند گرفت:

کاپولای فرانک به صورت زیر ارائه می‌شود (خورابی و مایور، ۲۰۱۳):

$$C^{Frank}(u_1, u_2; \alpha) = -\frac{1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}{e^{-\alpha} - 1}\right) \quad \text{رابطه (۲۰)}$$

تابع چگالی کاپولای فرانک به صورت زیر معرفی می‌شود:

$$C^{Frank} = \frac{\alpha(1 - e^{-\alpha})e^{-\alpha(u_1 + u_2)}}{((1 - e^{-\alpha}) - (1 - e^{-\alpha u_1})(1 - e^{-\alpha u_2}))^2} \quad \text{رابطه (۲۲)}$$

کاپولای گامبل به صورت زیر ارائه می‌شود: (خورابی و مایور، ۲۰۱۳)

$$C^{Gumble}(u_1, u_2; \alpha) = -\frac{1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}{e^{-\alpha} - 1}\right) \quad \text{رابطه (۲۳)}$$

تابع چگالی کاپولای گامبل به شرح زیر است:

$$C^{Gumble} = \frac{e^{-((-\ln u_1)^\alpha + (-\ln u_2)^\alpha)^{\frac{1}{\alpha}}} ((-\ln u_1)^\alpha + (-\ln u_2)^\alpha)^{\alpha-1}}{u_1 u_2 ((-\ln u_1) + (-\ln u_2))^{\frac{2}{\alpha}}} \quad \text{رابطه (۲۵)}$$

$$\times (((-\ln u_1)^\alpha + (-\ln u_2)^\alpha)^{\frac{1}{\alpha}} + \alpha - 1)$$

کاپولای کلایتون نیز به شکل زیر تعریف می شود (خاروبی و مورر، ۲۰۱۳):

$$C^{Clayton}(u_1, u_2; \alpha) = (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{\frac{1}{\alpha}} \quad \text{رابطه (۲۶)}$$

در نتیجه تابع چگالی این کاپولا به صورت زیر تعریف می شود:

$$C^{Clayton} = (1 + \alpha)(u_1 u_2)^{-\alpha-1} (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{\frac{2}{\alpha}} \quad \text{رابطه (۲۸)}$$

فرضیه ها و پرسش های پژوهش

با توجه به موضوعات ارائه شده در بالا، این پژوهش به دنبال پاسخ دادن به این پرسش است که «استفاده از کدام یک شاخص های ضرایب همبستگی مبتنی بر کاپولا، رویکرد DTW و رویکرد EDR در خوشه بندی سری های زمانی به پرتفوی ردیاب شاخص با خطای ردیابی پایین تر منتج خواهد شد؟». همچنین این پژوهش به دنبال آزمون کردن این فرضیه است که «پرتفوی ردیاب شاخص مبتنی بر رویکرد EDR به عنوان یک رویکرد بهبود دهنده نسبت به رویکرد DTW خطای ردیابی پایین تری نسبت به پرتفوی ردیاب مبتنی بر شاخص DTW دارد».

روش پژوهش

مدل استفاده شده در این پژوهش برای خوشه بندی سری های زمانی، این امکان را به تصمیم گیرنده می دهد که تعداد سهم های مورد نظر خود برای تشکیل پرتفوی استوار بر شاخص را با پارامتر q مشخص کند. در واقع، این مدل سهم های تشکیل دهنده شاخص را به q گروه تقسیم کرده و از هر گروه ۱ نماینده انتخاب می کند. در مدل بهینه سازی صفر و یک زیر، اگر مقدار z_{ij} برابر با ۱ باشد، سهم j به عنوان نماینده انتخاب شده است. همچنین اگر مقدار متغیر x_{ij} برابر با ۱ باشد، سهم j به عنوان نماینده سهم i انتخاب شده است.

$$\max \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \cdot x_{ij}$$

Subject to

$$(1) \sum_{j=1}^n y_j = q$$

رابطه (۲۹)

$$(2) \sum_{j=1}^n x_{ij} = 1$$

$$(3) x_{ij} \leq y_j$$

$$x_{ij}, y_j \in (0,1)$$

محدودیت‌های مدل بالا به صورت زیر قابل بررسی هستند:

۱. محدودیت اول تعداد سهم‌های انتخابی برای تشکیل پرتفوی را مشخص می‌کند.
 ۲. مجموعه محدودیت‌های دوم نشان می‌دهند که هر سهم فقط یک نماینده خواهد داشت.
 ۳. مجموعه محدودیت‌های سوم نشان می‌دهند که اگر سهم j برای تشکیل پرتفوی انتخاب شده باشد، می‌تواند نماینده سهم‌های دیگر نیز باشد.
- با حل کردن مدل بالا، فرایند انتخاب سهم‌های مورد نظر برای تشکیل پرتفوی تکمیل می‌شود. گام بعدی تعیین وزن این سهم‌ها در پرتفوی و یا نسبت سرمایه‌گذاری در هر کدام از این سهم‌های انتخابی است. مقدار w_j یا ارزش بازار برای هر سهم انتخاب شده به عنوان نماینده، به صورت زیر محاسبه می‌شود:

$$w_j = \sum_{i=1}^n v_i \cdot x_{ij} \quad \text{رابطه (۳۰)}$$

در رابطه فوق مقدار v_i نشان‌دهنده ارزش بازار سهم i ام است. با این توضیحات در واقع w_j مقدار بیانگر مجموع ارزش بازار سهم‌هایی است که سهم j به عنوان نماینده آن‌ها انتخاب شده است. حال که ارزش بازار هر کدام از سهم‌های انتخابی برای پرتفوی محاسبه شد، نسبت سرمایه‌گذاری در هر کدام از سهم‌ها به صورت زیر تعریف می‌شود:

$$\Phi_j = \frac{w_j}{\sum_{f=1}^q w_f} \quad \text{رابطه (۳۱)}$$

فنون تجزیه و تحلیل اطلاعات

در این پژوهش، مدل ارائه شده در رابطه ۲۹ با استفاده از شاخص‌های اندازه‌گیری شباهت ارائه شده قسمت پیشینه نظری پژوهش اجرا شده و پس از خوشه‌بندی سهم‌های تشکیل دهنده شاخص، با استفاده از رابطه ۳۱، وزن هم سهم در پرتفوی مبتنی بر شاخص محاسبه می‌شود.

داده‌های مورد استفاده

شاخص پنجاه شرکت فعال‌تر، یکی از شاخص‌های پر کاربرد در بورس اوراق بهادار تهران است. ترکیب این شاخص هر ۳ ماه به روز می‌شود و سهم‌های تشکیل‌دهنده شاخص با توجه به معیارهای زیر انتخاب می‌شوند (راعی و پویان‌فر، ۱۳۸۴).

- نقدشوندگی سهام
- تأثیر سهام بر ارزش جاری بازار
- نسبت‌های مالی، به خصوص نسبت‌های سودآوری

با توجه به اینکه برای سنجش میزان شباهت میان دارایی‌ها، سری زمانی بازده روزانه مورد نیاز است و با توجه به معیارهای بالا و پرمعامله‌بودن سهم‌های تشکیل‌دهنده شاخص پنجاه شرکت فعال‌تر، این شاخص به عنوان شاخص هدف به منظور تشکیل پرتفوی ردیاب شاخص انتخاب شده است. به منظور تحلیل مدل از اطلاعات تاریخی در بازه زمانی فروردین ماه تا اسفندماه ۱۳۹۶ مربوط به شرکت‌هایی که در سه ماه نخست سال ۱۳۹۷ به عنوان ۵۰ شرکت فعال انتخاب شده‌اند، استفاده شده است. بازده روزانه این سهم‌ها در بازه زمانی فروردین لغایت اسفند ۹۶ برای محاسبه معیارهای شباهت مورد استفاده قرار گرفته است.

معیارهای ارزیابی عملکرد

در این قسمت معیارهای استاندارد که برای ارزیابی عملکرد پرتفوی استخراج شده از مدل مورد استفاده قرار می‌گیرند، معرفی می‌شود.

به منظور سنجش موفقیت یک پرتفوی در ردیابی شاخص از معیار نسبت بازار استفاده می‌شود. این معیار که با استفاده از رابطه ۱۸ محاسبه می‌شود، نشان می‌دهد که آیا پرتفوی در ردیابی شاخص موفق بوده است یا خیر. مقدار مطلوب این معیار برای پرتفوی ردیاب شاخص معادل عدد ۱ است.

$$MR^i = \frac{1 + r_p^i}{1 + r_I^i} \quad \text{رابطه (۳۲)}$$

در رابطه بالا، r_I^i نشان‌دهنده مقدار بازده شاخص در زمان i و r_p^i نشان‌دهنده بازده پرتفوی در زمان i است. به منظور مقایسه عملکرد پرتفوهای ردیاب از معیار خطای ردیابی استفاده می‌شود. این معیار که از طریق رابطه زیر محاسبه می‌شود، به بررسی اختلاف عملکرد بازده شاخص و بازده پرتفوی، در دوره زمانی مورد آزمون خواهد پرداخت (چن و روی، ۲۰۱۲).

$$TE^i = \frac{\sum_{i=1}^n |r_p^i - r_l^i|}{n} \quad \text{رابطه (۳۳)}$$

پارامترهای موجود در رابطه بالا، معادل پارامترهای معرفی شده در خصوص رابطه ۳۲ هستند.

بحث در یافته‌های پژوهش

در این پژوهش از داده‌های ۵۰ شرکت فعالتر بورس اوراق بهادار تهران، در ۱۳ دوره سه ماهه از ابتدای سال ۱۳۹۴ تا پایان ۳ ماه نخست سال ۱۳۹۷ به منظور سنجش عملکرد مدل در ردیابی شاخص استفاده شده است. در ابتدا با استفاده از قیمت‌های روزانه یک سال قبل از ۳ ماهه هدف، بازده سهم‌های تشکیل دهنده شاخص ۵۰ شرکت فعالتر بورس اوراق بهادار تهران محاسبه و با استفاده از شاخص‌های همبستگی استوار بر الگوریتم‌های DTW و EDR و ضرایب همبستگی استوار بر کاپولا مقدار شاخص‌های سنجش شباهت میان سری‌های بازده، مورد محاسبه قرار گرفت. با به کار بردن شاخص‌های سنجش شباهت، در مدل بهینه‌سازی ارائه شده در رابطه ۲۹ سهم‌ها و وزن هر یک از آن‌ها در پرتفوی بهینه محاسبه شد. تعداد سهم‌های تشکیل دهنده پرتفوی ردیاب، ۵ سهم در نظر گرفته شد.

به منظور محاسبه معیارهای ارزیابی عملکرد، بازده روزانه پرتفوها با بازده شاخص ۵۰ شرکت فعالتر مورد مقایسه قرار گرفت. نمادهای مورد استفاده در جداول نتایج به شرح زیر هستند:

جدول ۱. نام‌گذاری پرتفوی‌های ردیاب شاخص

نماد پرتفوی	ویژگی
پرتفوی A	پرتفوی بهینه بر اساس شاخص "همبستگی مبتنی بر کاپولای گاوسی" به عنوان معیار سنجش شباهت سری‌های زمانی
پرتفوی B	پرتفوی بهینه بر اساس شاخص "همبستگی مبتنی بر کاپولای تی" به عنوان معیار سنجش شباهت سری‌های زمانی
پرتفوی C	پرتفوی بهینه بر اساس شاخص "همبستگی مبتنی بر کاپولای گامیل" به عنوان معیار سنجش شباهت سری‌های زمانی
پرتفوی D	پرتفوی بهینه بر اساس شاخص "همبستگی مبتنی بر کاپولای فرانک" به عنوان معیار سنجش شباهت سری‌های زمانی
پرتفوی E	پرتفوی بهینه بر اساس شاخص "همبستگی مبتنی بر کاپولای کلایون" به عنوان معیار سنجش شباهت سری‌های زمانی
پرتفوی F	پرتفوی بهینه بر اساس شاخص "DTW" به عنوان معیار سنجش شباهت سری‌های زمانی
پرتفوی G	پرتفوی بهینه بر اساس شاخص "EDR" به عنوان معیار سنجش شباهت سری‌های زمانی

در این مرحله به محاسبه مقدار نسبت بازار برای پرتفوها پرداخته می‌شود. همانطور که پیش از این اشاره شد، مقدار مطلوب برای این معیار در خصوص پرتفوی ردیاب شاخص معادل ۱ است. به منظور بررسی مقادیر، از آزمون آماری سنجش مقدار میانگین جامعه به صورت زیر استفاده می‌شود:

نتیجه آزمون نشان می‌دهد که هر ۷ پرتفوی در ردیابی شاخص موفق عمل کرده‌اند. در ادامه، به مقایسه عملکرد پرتفوها پرداخته می‌شود. پیش از این اشاره شد که به منظور مقایسه عملکرد پرتفوها از معیار خطای ردیابی استفاده می‌شود. متوسط خطای ردیابی ۷ پرتفوی به شرح جدول زیر است:

جدول ۴. متوسط خطای ردیابی روزانه پرتفوها

پرتفوی A	پرتفوی B	پرتفوی C	پرتفوی D	پرتفوی E	پرتفوی F	پرتفوی G	
متوسط خطای ردیابی بهار ۹۴	-۰.۰۰۱٪	۰.۱۳۵٪	-۰.۰۶۷٪	-۰.۰۲۸٪	-۰.۱۰۱٪	-۰.۰۷۳٪	-۰.۱۵۹٪
متوسط خطای ردیابی تابستان ۹۴	۰.۶۱۷٪	۰.۳۳۳٪	۰.۹۶۸٪	۰.۸۸۸٪	۱.۰۶۵٪	۰.۹۳۰٪	۲.۶۴۸٪
متوسط خطای ردیابی پاییز ۹۴	۰.۰۴۴٪	-۰.۰۷۰٪	-۰.۰۲۵٪	۰.۰۱۹٪	-۰.۰۱۸٪	-۰.۰۲۶٪	۰.۰۵۱٪
متوسط خطای ردیابی زمستان ۹۴	۰.۱۵۷٪	۰.۹۴۱٪	۰.۴۷۵٪	۰.۴۸۳٪	۰.۶۱۱٪	۰.۸۳۳٪	۱.۲۲۴٪
متوسط خطای ردیابی بهار ۹۵	۰.۰۱۰٪	-۰.۴۱۳٪	-۰.۳۷۷٪	-۰.۳۱۴٪	-۰.۳۰۳٪	-۰.۲۷۳٪	-۰.۳۸۳٪
متوسط خطای ردیابی پاییز ۹۵	۰.۲۰۱٪	۰.۱۵۶٪	۰.۱۱۰٪	۰.۱۵۷٪	۰.۲۶۸٪	۰.۱۳۳٪	۰.۱۵۴٪
متوسط خطای ردیابی زمستان ۹۵	-۰.۰۰۳٪	-۰.۰۵۸٪	۰.۱۷۰٪	۰.۲۰۰٪	۰.۱۸۷٪	۰.۱۷۳٪	-۰.۰۴۸٪
متوسط خطای ردیابی بهار ۹۶	۰.۰۷۴٪	-۰.۲۵۵٪	-۰.۰۳۵٪	-۰.۰۲۷٪	-۰.۰۳۹٪	-۰.۰۰۴٪	-۰.۲۴۳٪
متوسط خطای ردیابی تابستان ۹۶	۰.۰۲۳٪	-۰.۲۱۸٪	-۰.۰۲۳٪	۰.۰۹۵٪	۰.۱۲۸٪	۰.۱۳۶٪	۰.۱۳۶٪
متوسط خطای ردیابی پاییز ۹۶	-۰.۰۳۵٪	-۰.۰۶۹٪	-۰.۰۹۸٪	۰.۱۶۹٪	۰.۱۰۵٪	۰.۰۱۷٪	-۰.۱۷۸٪
متوسط خطای ردیابی زمستان ۹۶	-۰.۰۴۴٪	۰.۰۵۴٪	۰.۴۲۷٪	۰.۱۵۱٪	۰.۱۸۱٪	۰.۱۵۸٪	-۰.۰۴۹٪
متوسط خطای ردیابی بهار ۹۷	-۰.۰۷۱٪	۰.۲۲۸٪	۰.۰۱۱٪	-۰.۰۰۳٪	۰.۱۰۵٪	۰.۰۸۴٪	۰.۲۳۴٪
متوسط کل خطای ردیابی	۰.۰۷۷٪	۰.۰۸۱٪	۰.۱۴۷٪	۰.۱۴۸٪	۰.۱۷۵٪	۰.۱۸۰٪	۰.۳۷۰٪

قابل مشاهده است که متوسط خطای ردیابی روزانه، در پرتفوی A مقداری پایین تر از سایر پرتفوها است. به منظور سنجش دقیق و مقایسه عملکرد این پرتفوها، از آزمون مقایسات زوجی استفاده می شود. در این آزمون مقدار خطای ردیابی دو پرتفوی مورد مقایسه قرار می گیرد. در مورد دو پرتفوی A و B، به منظور مقایسه مقدار خطای ردیابی، این آزمون به صورت زیر انجام می شود:

$$\left\{ \begin{array}{l} TE_A > TE_B : \text{فرض صفر} \\ TE_A \leq TE_B : \text{فرض یک} \end{array} \right. \quad \text{رابطه (۳۶)}$$

این آزمون نسبت به این فرضیه که خطای ردیابی پرتفوی A از پرتفوی B کمتر است، آزمونی سخت گیرانه است چرا که فرضیه بالا، در فرض یک (در مقابل فرض صفر) قرار گرفته است. برای انجام این آزمون، ابتدا اختلاف خطای ردیابی دو پرتفوی، در بازه زمانی i به صورت زیر محاسبه می شود:

$$D_i^{A,B} = TE_i^A - TE_i^B \quad \text{رابطه (۳۷)}$$

پس از محاسبه این اختلاف به ازای تمامی i ها، مقدار متوسط و انحراف استاندارد این مقادیر محاسبه شده و آماره مورد بررسی قرار خواهد گرفت:

$$t = \frac{\bar{D}^{A,B}}{\left(\frac{S^{A,B}}{\sqrt{n}} \right)} \quad \text{رابطه (۳۸)}$$

در عبارت بالا $\bar{D}^{A,B}$ نشان دهنده متوسط و $S^{A,B}$ نشان دهنده انحراف استاندارد اختلاف های میان خطای ردیابی دو پرتفوی و n نشان دهنده تعداد نمونه ها (در اینجا، تعداد بازه های زمانی) است. در صورتی که توزیع اختلاف ها ($D_i^{A,B}$ ها) نرمال باشد، این آماره دارای توزیع t با n-1 درجه آزادی است و با مقدار بحرانی این توزیع مقایسه می شود. رد شدن فرض صفر، در این آزمون نشان می دهد که مقدار خطای ردیابی پرتفوی A از پرتفوی B به صورتی معناداری پایین تر است.

این آزمون برای هر زوج مرتب از ۷ پرتفوی بررسی شده در پژوهش، انجام شده که نتیجه آن به شرح زیر است:

جدول ۵. مقایسات زوجی خطای ردیابی پرتفویهای شاخص

پرتفوی A	پرتفوی G	پرتفوی E	پرتفوی B	پرتفوی C	پرتفوی D	پرتفوی F	
$\bar{D}$ ۰.۰۰٪	$\bar{D}$ ۰.۰۷٪	$\bar{D}$ ۰.۰۷٪	$\bar{D}$ ۰.۱۰٪	$\bar{D}$ ۰.۱۰٪	$\bar{D}$ ۰.۱۰٪	$\bar{D}$ ۰.۱۹٪	پرتفوی D
S ۰.۳۳٪	S ۰.۲۵٪	S ۰.۱۸٪	S ۰.۲۲٪	S ۰.۲۴٪	S ۰.۲۴٪	S ۰.۶۶٪	
t ۰.۰۵	t ۱.۰۳	t ۱.۴۶	t ۱.۶۴	t ۱.۵۶	t ۱.۰۶	t ۱.۰۶	
		$\bar{D}$ ۰.۰۷٪	$\bar{D}$ ۰.۰۹٪	$\bar{D}$ ۰.۱۰٪	$\bar{D}$ ۰.۱۰٪	$\bar{D}$ ۰.۱۹٪	پرتفوی E
		S ۰.۲۸٪	S ۰.۲۷٪	S ۰.۲۸٪	S ۰.۲۳٪	S ۰.۶۶٪	
		t ۰.۸۵	t ۰.۹۱	t ۱.۲۱	t ۱.۶۲	t ۱.۰۳	
			$\bar{D}$ ۰.۰۰٪	$\bar{D}$ ۰.۰۳٪	$\bar{D}$ ۰.۰۳٪	$\bar{D}$ ۰.۱۲٪	پرتفوی G
			S ۰.۱۴٪	S ۰.۱۵٪	S ۰.۱۵٪	S ۰.۵۵٪	
			t ۰.۰۳	t ۰.۶۸	t ۰.۸۲	t ۰.۸۰	
				$\bar{D}$ ۰.۰۳٪	$\bar{D}$ ۰.۰۳٪	$\bar{D}$ ۰.۱۲٪	پرتفوی A
				S ۰.۰۸٪	S ۰.۱۳٪	S ۰.۵۶٪	
				t ۱.۲۰	t ۰.۹۸	t ۰.۷۸	
					$\bar{D}$ ۰.۰۱٪	$\bar{D}$ ۰.۰۹٪	پرتفوی C
					S ۰.۱۰٪	S ۰.۵۰٪	
					t ۰.۱۸	t ۰.۶۸	
						$\bar{D}$ ۰.۰۹٪	پرتفوی F
						S ۰.۵۳٪	
						t ۰.۶۲	

همانطور که در جدول قابل بررسی است، به ازای هر زوج از پرتفوها، مقدار متوسط و انحراف استاندارد اختلاف‌های میان خطای ردیابی و همچنین مقدار آماره ارائه شده در رابطه ۳۸ به منظور سنجش آزمون فرضیه ارائه شده در رابطه ۳۶ قابل مشاهده است. با توجه به مقدار بحرانی توزیع t (در سطح اطمینان ۹۵) که معادل ۱/۶۷ است، برتری هیچ یک از پرتفوها بر دیگری به صورت معنادار اثبات نمی‌شود. به عبارتی همه پرتفوها موفق به انجام فرایند ردیابی شاخص شده‌اند و میزان خطای ردیابی پرتفوها در سطح معنی داری ۹۵ درصد با یکدیگر متفاوت نیست.

نتیجه‌گیری و ارائه پیشنهادها

در این پژوهش تاثیر معیارهای مختلف سنجش شباهت، در خوشه‌بندی سری‌های زمانی و در نهایت تشکیل پرتفوی استوار بر شاخص انجام شد. به این منظور شاخص‌های ضرایب

همبستگی مبتنی بر کاپولای گاوسی، تی، کلایتون، فرانک و گامبل و همچنین رویکردهای سنجش شباهت مبتنی بر برنامه ریزی پویا مانند DTW و EDR به منظور تعیین شباهت میان سری‌های زمانی، مورد استفاده قرار گرفتند.

در ادامه از مقادیر محاسبه شده توسط هر یک از این رویکردها در یک مدل بهینه سازی ریاضی صفر و یک، برای خوشه‌بندی سری‌های زمانی، استفاده شد. این مدل، پس از خوشه‌بندی سری‌های زمانی بازده سهم‌های شاخص، از هر خوشه یک نماینده را انتخاب کرده و وزن هر نماینده را معادل مجموع وزن سهم‌های خوشه در نظر می‌گیرد. با توجه به توضیحات بالا، ۷ پرتفوی ردیاب شاخص به صورت زیر قابل معرفی خواهد بود:

- پرتفوی ردیاب شاخص مبتنی بر رویکرد DTW
 - پرتفوی ردیاب شاخص مبتنی بر رویکرد EDR
 - پرتفوی ردیاب شاخص مبتنی بر ضریب همبستگی بر اساس کاپولای گاوسی
 - پرتفوی ردیاب شاخص مبتنی بر ضریب همبستگی بر اساس کاپولای تی
 - پرتفوی ردیاب شاخص مبتنی بر ضریب همبستگی بر اساس کاپولای گامبل
 - پرتفوی ردیاب شاخص مبتنی بر ضریب همبستگی بر اساس کاپولای کلایتون
 - پرتفوی ردیاب شاخص مبتنی بر ضریب همبستگی بر اساس کاپولای فرانک
- به منظور بررسی فرایند بالا، از اطلاعات شرکت‌های تشکیل دهنده شاخص ۵۰ شرکت فعال‌تر بورس اوراق بهادار تهران در در ۱۳ بازه ۳ ماهه از سال ۱۳۹۴ تا ۱۳۹۷ استفاده شد. پس از محاسبه معیارهای شباهت و تشکیل پرتفوها و انجام آزمون خارج از نمونه، نتایج زیر به دست آمد:
- ✓ آزمون خارج از نمونه روی معیار نسبت بازار پرتفوها نشان از موفقیت همه آنها در تشکیل پرتفوی مبتنی بر شاخص دارد.
- ✓ پرتفوی مبتنی بر رویکرد کاپولای گاوسی بر اساس شاخص نسبت بازار، بیشترین دقت را در تشکیل پرتفوی ردیاب شاخص داشته است.
- ✓ آزمون خارج از نمونه بر روی خطای ردیابی پرتفوها نشان از این موضوع دارد که خطای ردیابی ۷ پرتفوی در سطح اطمینان ۹۵ درصد دارای تفاوت معنی دار نیست.
- در نظر گرفتن هزینه معاملاتی برای تنظیم مجدد پرتفوی در بازه‌های زمانی و همچنین افزایش بازدهی پرتفوی همزمان با حفظ سطح ریسک به عنوان رویکردی برای تشکیل پرتفوی شاخصی ارتقایافته (فعال) به عنوان موضوع‌هایی برای پژوهش‌های آتی پیشنهاد می‌شود.

منابع

- امیری، م؛ شهسواری، ش؛ چالسیاری، س (۱۳۹۹) "تشکیل صندوق شاخصی بهبود یافته با استفاده از رویکرد بهینه سازی امکانی استوار"، *بورس اوراق بهادار*، دوره ۱۳، شماره ۴۹، صفحات ۴۷ تا ۷۲
- امیری، م؛ کرمی، ش؛ ناصرپور، ع (۱۳۹۵). "ردیابی شاخص بورس اوراق بهادار با در نظر گرفتن محدودیت زیان گریزی با استفاده از رویکرد جدید بیگ بنگ بیگ کرانچ"، *دانش سرمایه گذاری*. سال پنجم، شماره ۱۵ صفحات ۸۳ تا ۱۰۵.
- انصاری، ح؛ بهزادی، ع؛ تندنویس، ف. (۱۳۹۸). مقاله پژوهشی: ردیابی شاخص بهبود یافته دو مرحله ای با استفاده از الگوریتم های فراابتکاری. *راهبرد مدیریت مالی*، ۲۲-۱، (۴) ۷.
- حجازی، ر؛ جعفری سرشت، د؛ دلشادی، م، (۱۳۹۰) "تشکیل صندوق شاخصی بهبود یافته با استفاده از الگوریتم ژنتیک"، *بورس اوراق بهادار*، دوره ۴، شماره ۱۴، صفحات ۱۳۵ تا ۱۵۷.
- راعی، ر؛ پویانفر، ا (۱۳۸۴)؛ *مدیریت سرمایه گذاری پیشرفته*، ایران، انتشارات سمت.
- سبزواری، ج؛ سجادی، م؛ سجادی، م (۱۳۹۴) "حل مساله ردیابی شاخص با استفاده از الگوریتم فراابتکاری ترکیبی کرم شب تاب"، *مدیریت مهندسی و رایانش نرم*، دوره ششم، شماره ۱۰، صفحات ۱۲۵ تا ۱۴۳.
- عیوض لو، ر؛ شفیع زاده، م؛ قهرمانی، ع (۱۳۹۶) "ردیابی شاخص و شاخص بهبود یافته با استفاده از رویکردهای هم انباشتگی و همبستگی"، *تحقیقات مالی*، دوره نوزدهم، شماره ۳، صفحات ۴۵۷-۴۷۴.
- فلاح پور، س؛ تندنویس، ف؛ هاشمی، ا، (۱۳۹۴) "بهینه سازی پرتفوی ردیاب شاخص با استفاده از مدل تک شاخصی پایدار بر مبنای شاخص ۵۰ شرکت فعالتر بورس اوراق بهادار تهران"، *مهندسی مالی و مدیریت اوراق بهادار*، دوره ششم، شماره ۲۴، صفحات ۱۱۵ تا ۱۳۴.
- فلاح پور، س؛ تندنویس، ف، (۱۳۹۴) "کاربرد رویکرد بهینه سازی استوار در تشکیل پرتفوی سهام مبتنی بر شاخص با در نظر گرفتن ردیاب پارامترها"، *تحقیقات مالی*، دوره هفدهم، شماره ۲، صفحه ۳۲۵-۳۴۰.
- کشاوری حداد، غ؛ حیرانی، م (۱۳۹۳) "برآورد ارزش در معرض ریسک با وجود ساختار وابستگی بین بازدهی های مالی: رهیافت مبتنی بر توابع کاپولا"، *تحقیقات اقتصادی*، دوره چهل و نهم، شماره ۴، صفحات ۸۶۹-۹۰۲.

References

- Amiri, M., shahsavari, S., lak chalsepary, S. (2020). Enhanced Index Fund Construction using Robust possibilistic programming. *Journal of Securities Exchange*, 13(49), 47-72 (in Persian)
- Amiri, M., Karami, Sh., Naserpour, A. (2107). Tracking Stock Exchange Index with considering the limitation of loss aversion with using the new approach of Big Bang –Big Crunch. *Investment Knowledg*, 5(19), 83-105. (In Persian).
- Ansari, H., Behzadi, A., & Tondnevis, F. (2019). Enhanced Index Tracking with a Two-Stage Mixed Integer Programing Model and Pattern Search Algorithm. *Financial Management Strategy*, 7(4), 1-22 (In Persian).
- Beasley J. Meade N. and Chang T. (2003) "An evolutionary heuristic for the index tracking problem," *European Journal of Operational Research*, p. 148(3): 621–643.
- Berndt D and Clifford J, (1994) "Using dynamic time warping to find patterns in time series," in the AAAI Workshop on Knowledge Discovery in Databases,.
- Chen C. and Roy H. (2012) "Robust portfolio selection for index tracking," *Computers & Operations Research*, p. 39 829–837.
- Chen L. Oszu M. and Oria V. (2005) "Robust and fast similarity search for moving object trajectories," in Int. Conf. on Management of Data.
- Clarke R., Krase C. and Statman M. (1994) "Tracking error, regret and tactical asset allocation," *The Journal of Portfolio Management*, pp. 20, 16–24.
- Corielli F. and Marcellino M. (2006) "Factor based index tracking." *Journal of Banking & Finance* 30.8: 2215-2233.
- Cornuejols and Tutuncu (2005), "Optimization Methods in Finance", Carnegie Mellon University.
- Dos C. and Cincotti S. (2005) "Clustering of Financial time series with application to index and enhanced-index tracking portfolio " Università di Genova, Via Opera Pia 11a, 16145 Genova, Italia
- Embrechts P. McNeil A. and Straumann D. (2002) "Correlation and dependence in risk management: properties and pitfalls," in Risk Management: Value at Risk and Beyond,, Cambridge.
- Erdogan E. Goldfarb V and Iyengar G (2004) "Robust portfolio management" Columbia University, CORC Technical Report TR.
- Eyvazloo, R., Shafizade, M. & Ghahremani, A. (2017). Index Tracking and Enhanced Indexing Using Co-integration and Correlation Approaches. *Financial Researches*, 19(3), 457-474. (In Persian).

- Fallahpor, S., Tondnevis, F., Hashemi, M.A. (2016). Index tracking portfolio optimization by robust sigle factor model based on 50 more active company index of TSE. *Financial Engineering and security Management*, 6(24), 115-134. (In Persian).
- Fallahpor, S. and Tondnevis, F. (2015). Application of an optimization model for constructing an index tracker portfolio and considering the uncertainty of model parameters by using of robust optimization approach. *Financial Researchers*, 17(2), 325-340. (In Persian).
- Focardi S. and Fabozzi F. (2004) "A methodology for index tracking based on timeseries clustering," *Quantitative Finance*, p. 4(4):417-25.
- García, F, Guijarro F, and Oliver J. (2017) "Index tracking optimization with cardinality constraint: a performance comparison of genetic algorithms and tabu search heuristics." *Neural Computing and Applications*, 30 2625-2641.
- Gilli M. and Kellezi E. (2002) "The threshold accepting heuristic for index tracking" in Financial engineering, e-commerce, and supply chain. Kluwer In: Pardalos P, Tsitsiringos VK, editors.
- Hejazi, R., Jafari Seresht, D., Delshadi, M. (2012). To Form Enhanced Index Fund with Genetic Algorithm. *Journal of Securities Exchange*, 4(14), 135-157. (In Persian).
- Jeng, Yih, Chen-Ju Lee, and Shyh-Weir Tzang. (2013) "Application of a multifactor model in enhanced index fund: Performance analysis in China." *Emerging Markets Finance and Trade* 49.sup4 163-183.
- Keshavarz Haddad, Gh., Heyrani, M. (2015). Estimation of Value at Risk in the Presence of Dependence Structure in Financial Returns: A Copula Based Approach. *Economic Researches*, 49(4), 869-902. (In Persian).
- Kharoubi C. and Maurer F. (2013) "Copulas In Finance Ten Years Later,," *Journal of Applied Business Research*, pp. 29(5):1555-1566
- Konno H. and Wijayanayake A. (2001) "Minimal cost index tracking under nonlinear transaction costs and minimal transaction unit constraints." *International Journal of Theoretical and Applied Finance*, p. 4(6):939-958.
- Markowitz H. (1952) "Portfolio selection," *The journal of finance*, pp. 77-91.
- Markowitz H. (1959) "Portfolio selection: efficient diversification of investments," Yale university press,.
- Meade N. and Salkin G. (1990) "Developing and maintaining an equity index fund," *Journal of the Operational Research Society*, p. 41(7):599-607.
- Meade N. and Salkin G. (1989) "Index funds-construction and performance measurement." *J. Opl Res. Soc.*, pp. 40, 871-879.

- Meihua W. Chengxian X. Fengmin X. and Hong X. (2011) "A mixed 0–1 LP for index tracking problem with CVaR risk constraints," *Annals of Operations Research*, 196 591-609
- Patton A. (2007), "Copula-Based Models for Financial Time Series," The Oxford-Man Institute, University of Oxford.
- Raei, R., Pouyanfar, A. (2006). Advanced Investment. Samt publication. (In Persian).
- Rohweder H. (1998) "Implementing stock selection ideas: does tracking error optimization do any good?" *Journal of Portfolio Management*, p. 24(3):49–59.
- Roll J. and Ronald D. (2002) "Optimal benchmark tracking with small portfolios," *Journal of Portfolio Management*, pp. 28(2):9-33.
- Roll R. (1992) "A mean/variance analysis of tracking error," *Journal of Portfolio Management*, p. 18:13–22.
- Roy, K; Dexiang, W; (2016) "Factor-based robust index tracking" *Optimization and Engineering*, pp 443-466.
- Rudd A. (1980) "Optimal selection of passive portfolios," *Financial Management*, p. 9(1):57–66.
- Rudolf, Markus and Jurgen, "A linear model for tracking error minimization," *Journal of Banking & Finance*, p. 23(1): 85–103., 1999.
- Saant'Anna L. Filomena T. and Caldeira J. (2017) "Index tracking and enhanced indexing using cointegration and correlation with endogenous portfolio selection." *The Quarterly Review of Economics and Finance* 65 146-157.
- Saant'Anna L. Filomena T. Guedes P. Borenstein D. (2017) "Index tracking with controlled number of assets using a hybrid heuristic combining genetic algorithm and non-linear programming." *Annals of Operations Research* 258.2 : 849-867.
- Sajjadi, S., Sajjadi, S. and Sabzevari, M. (2016). Index tracking problem solving using metaheuristic firefly algorithm". *Engineering management and soft computing*, 1(1), 125- 143. (In Persian).
- Serra J. and Arcos L.(2014), "An empirical evaluation of similarity measures for time series classification," *Knowledge-Based System*.
- Sharpe W. (1963) "A simplified model for portfolio analysis," *Management science*, pp. vol. 9, pp. 277-293.
- Shen J. Huang W. Zhu D. Liang J. (2016) "A Novel Similarity Measure Model for Multivariate Time Series Based on LMNN and DTW," *Neural Processing Letters*. 45 (3) pp 925–937

Song P. (2000)", Multivariate dispersion models generated from gaussian copula," *Scandinavian Journal of Statistics*, 27(2), p. 305–320.

Wang X. Mueen A., Ding H., Trajcevski G., Scheuermann P. and Keogh J. (2013) "Experimental comparison of representation methods and distance measures for time series data," *Data Min. Knowl. Discov*, p. 26 275–309.

COPYRIGHTS



This is an open access article under the CC BY-NC 4.0 license.