

Portfolio Optimization with Clustering Methods

Mohammad Soltani-Nejad¹, Maryam Davallou²

1- MSc Student, Financial Management Dept., Faculty of Management and Accounting Shahid Beheshti University, Tehran, Iran.

m.soltaninejad@ut.ac.ir

2- Assistant Professor, Financial Management Dept., Faculty of Management and Accounting Shahid Beheshti University, Tehran, Iran.

m_davallou@sbu.ac.ir

Abstract

The portfolio selection problem and optimizing it is an important issue in investment field. This process aims to create A portfolio that incur the minimum amount of risk to the investor in different levels of return. The existence of statistical uncertainty and noise in information, which is used in the process of optimization, would hamper the performance of optimized portfolios. The purpose of this article is to eliminate the noise of correlation coefficient matrix by using clustering method. In order to do so, two clustering methods, Single Linkage and Average Linkage, are applied, and the optimized portfolio based on the Tehran Stock Exchange information from March 2007 to March 2013 is obtained. The results show that such portfolio is more reliable and provides less risk to the capital owner. For analyzing the robustness of the results in a wide range of different investment horizons and numbers of assets, a Bootstrap analysis with resampling is performed and in most cases a significant improvement was observed.

Keywords: Clustering methods, Correlation Coefficient Matrix, Portfolio Optimization.

بهینه‌سازی سبد سرمایه‌گذاری با استفاده از روش‌های خوشه‌بندی

محمد سلطانی‌نژاد^{۱*}، مریم دولو^۲

۱- دانشجوی کارشناسی ارشد گروه مدیریت مالی دانشکده مدیریت و حسابداری دانشگاه شهید بهشتی، تهران، ایران.

m.soltaninejad@ut.ac.ir

۲- استادیار گروه مدیریت مالی دانشکده مدیریت و حسابداری دانشگاه شهید بهشتی، تهران، ایران.

m_davallou@sbu.ac.ir

چکیده

مسئله انتخاب سبد سرمایه‌گذاری و بهینه‌سازی، یکی از موضوعات مهم در حوزه سرمایه‌گذاری است. در این فرایند سعی بر آن است که سرمایه‌گذار بازای هر نرخ از بازده، کمترین ریسک ممکن را متحمل شود. اگر اطلاعاتی که در جریان بهینه‌سازی سبد سرمایه‌گذاری استفاده شده‌اند، دچار عدم قطعیت‌های آماری و یا در معرض نویز باشند، به کارایی سبد سرمایه‌گذاری لطمه وارد می‌شود. هدف این مقاله، حذف نویز اطلاعات ماتریس ضرایب همبستگی از طریق روش‌های خوشه‌بندی است. برای این منظور، دو روش خوشه‌بندی اتصال واحد و اتصال میانگین به کار گرفته شد و براساس اطلاعات بازده روزانه ۸۰ شرکت بورس اوراق بهادار تهران در بازه فروردین ۱۳۸۵ تا اسفند ۱۳۹۲، سبد سرمایه‌گذاری بهینه به دست آمد. نتایج نشان می‌دهد میزان اختلاف ریسک پیش‌بینی شده و ریسک مشاهده شده در این سبد سرمایه‌گذاری کمتر است و از طرفی در مجموع، ریسک کمتری را به سرمایه‌گذار تحمیل می‌کند. همچنین با هدف تحلیل حساسیت نتایج دو روش، یک مطالعه بوت استرپینگ در تعداد سهام و افق‌های زمانی مختلف انجام گرفت، در بیشتر حالت‌ها بهبود عملکرد سبد سرمایه‌گذاری مشاهده می‌شود.

واژه‌های کلیدی: بهینه‌سازی سبد سرمایه‌گذاری، روش‌های خوشه‌بندی، ماتریس ضرایب همبستگی.

مقدمه

یکی از مسائل مهم در سرمایه گذاری، انتخاب سبد سرمایه گذاری به صورت بهینه و به حداقل رساندن ریسک در سطوح مختلف بازده است. مارکوویتز (۱۹۵۳) در الگوی کلاسیک خود نشان داد چگونه با توجه به دو عامل ریسک و بازده، یک سبد سرمایه گذاری بهینه می توان تشکیل داد. سرمایه گذاران برای استفاده از روش مارکوویتز در عمل از اطلاعات تاریخی برای تخمین پارامترهای الگو از جمله ماتریس ضرایب همبستگی استفاده می کنند. یکی از مسائل مهم، وجود عدم قطعیت در این تخمین ها است. با توجه به داده های سری زمانی محدود، پیش بینی ماتریس ضرایب همبستگی در افق زمانی سرمایه گذاری، همواره با عدم قطعیت آماری همراه است. یکی از روش های کاهش عدم قطعیت، حذف نویز از اطلاعاتی است که بر اساس آن ها ماتریس ضرایب همبستگی تخمین زده می شود.

در مبانی نظری اکونوفیزیک^۱، روش های متعددی برای کاهش نویز مذکور معرفی شده است. یکی از این روش ها، استفاده از نظریه ماتریس تصادفی^۲ است. ماتریس همبستگی فیلتر شده حاصل از این روش، در مطالعات متعددی برای بهینه سازی سبد سرمایه گذاری به کار رفته است و شواهد به دست آمده نشان می دهد در صورت دسترسی به پیش بینی کامل بازده و نوسان های آینده، سبد سرمایه گذاری بهینه حاصل از این ماتریس به سبد سرمایه گذاری بهینه واقعی نزدیک تر است [۷]. همچنین در سال های اخیر، روش های جدیدی برای فیلتر کردن ماتریس ضرایب همبستگی مبتنی بر روش های خوشه بندی^۳ در ادبیات

اکونوفیزیک مطرح شده است. در این روش ها برای بخشی از اطلاعات که کمتر در معرض نویز قرار دارند، ماتریس ضرایب همبستگی استخراج می شود. این اطلاعات به گونه ای انتخاب می شوند که بتوانند کل ماتریس را نمایندگی کنند و بقیه اطلاعات حذف شوند. روش های خوشه بندی متعددی در دسترس است؛ اما بسیاری از آن ها با قیود لازم در بهینه سازی سبد سرمایه گذاری بر اساس ماتریس همبستگی، سازگاری ندارند. در این مقاله از دو روش متداول - که در صورت استفاده از آن ها، در ماتریس همبستگی فیلتر شده همچنان قیود لازم برقرار باشد - بهره گرفته شده است؛ سپس سبد سرمایه گذاری بهینه تعیین می شود.

مبانی نظری پژوهش

در مدل مارکوویتز اگر N دارایی ریسک پذیر در اختیار باشد و ماتریس کوواریانس بازده این دارایی ها C و بردار $R = [r_1, r_2, \dots, r_N]$ ، بازده دارایی ها باشد، با ترکیب این دارایی ها یک سبد سرمایه گذاری با اوزان p_i می توان به دست آورد که در بردار $W = [p_1, p_2, \dots, p_N]$ قرار دارد. بر این اساس، مسأله بهینه سازی مارکوویتز، یافتن بردار اوزان W به گونه ای است که بازای بازده r_p ، کمترین مقدار σ_p برای سبد سرمایه گذاری ایجاد شود. در صورتی که r_p و σ_p به ترتیب، به صورت روابط (۱) و (۲) تعریف شوند:

$$r_p = \sum_{i=1}^N p_i r_i \quad (1)$$

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N p_i p_j \sigma_{ij}, \quad (2)$$

تابع هدف و قید بهینه سازی به شرح رابطه (۳)

خواهد بود:

1. Econophysics
2. Random Matrix Theory
3. Clustering Methods

پیشینه تجربی پژوهش

در ادبیات اکونومیک، استفاده از نظریه ماتریس تصادفی برای کاربردهای متنوع از جمله فیلتر کردن ماتریس همبستگی، سابقه زیادی دارد و تاکنون از این روش در پژوهش‌های بسیاری برای حذف نویز اطلاعات مالی استفاده شده است. از جمله این مطالعات به گالوچیو و همکاران (۱۹۹۸)، لالوکس و همکاران (۲۰۰۰)، دورزد و همکاران (۲۰۰۱)، پلرو و همکاران (۲۰۰۲)، پافکا و همکاران (۲۰۰۳) و شریفی و همکاران (۲۰۰۴) می‌توان اشاره کرد. پلرو و همکاران (۲۰۰۱) برای سه دوره زمانی متفاوت و سه مجموعه متفاوت از سهام شرکت‌های امریکایی، ماتریس همبستگی را محاسبه کردند و ویژگی‌های مقادیر ویژه این ماتریس‌ها را بررسی کردند. آن‌ها با تعیین ماتریس تصادفی متناظر با ماتریس‌های همبستگی به این نتیجه رسیدند که برخی بردارهای ویژه، داخل و برخی خارج از محدوده ماتریس تصادفی قرار می‌گیرند. در ادامه با بررسی بیشتر بردارهای ویژه خارج از محدوده نشان دادند این انحرافات یا مربوط به یک اثر مشترک در سهام همه شرکت‌ها و یا مربوط به یک بخش اقتصادی خاص است. این پژوهشگران در نهایت نتیجه گرفتند این روش برای تشکیل یک سبد سرمایه‌گذاری می‌تواند استفاده شود که از نسبت ریسک به بازده پایداری برخوردار باشد [۱۳]. شریفی و همکاران (۲۰۰۴) با استفاده از نظریه ماتریس تصادفی نشان دادند در ماتریس همبستگی سهام شرکت‌های اس‌اند پی ۵۰۰، حجم زیادی از نویز اطلاعاتی را می‌توان مشاهده کرد. آن‌ها با شبیه‌سازی یک مجموعه داده مجزا و اضافه کردن مقداری نویز تصادفی به آن، میزان حساسیت به نویز مشخصات طیفی ماتریس تصادفی را نیز بررسی کردند؛

$$\begin{aligned} \min \sigma_p^2 \\ = W^T C W \\ S.T: \sum_{i=1}^N p_i = 1 \end{aligned} \quad (3)$$

روش مارکوویتز برای یافتن سبد سرمایه‌گذاری بهینه، فرض‌های متعددی دارد. نخست اینکه بازده دارایی‌ها، توزیع نرمال دارند. دوم، فرض می‌شود بردار بازده R و ماتریس کوواریانس C در افق سرمایه‌گذاری، ثابت باقی می‌ماند. هر دو فرض فوق در عمل چندان دقیق نیستند. حتی با قبول این فرض‌ها و با فرض در اختیار داشتن اطلاعات دوره زمانی T ، برای محاسبه ماتریس کوواریانس به $\frac{N(N-1)}{2} \sim \frac{N^2}{2}$ داده نیاز است و این مقادیر فقط بر اساس NT داده در دسترس، تخمین زده می‌شوند. برای جلوگیری از پدیده «نفرین چندبعدی بودن» به سری زمانی با طول $N \gg T$ نیاز است تا خطای موجود در ماتریس کوواریانس کوچک باشد. هرچه ابعاد ماتریسی که قرار است با استفاده از داده‌های تاریخی تخمین زده شود بیشتر باشد، به طور تصادفی تعداد مشاهدات بیشتری لازم خواهد بود. این پدیده را اولین بار بلمن (۱۹۶۱) شناسایی و تحت عنوان نفرین چندبعدی بودن معرفی کرد [۱]. با این حال برای T ‌های بزرگ، اطلاعات در معرض نویز و عدم قطعیت‌های آماری قرار می‌گیرند. به این دلیل بخشی از ماتریس کوواریانس که نویز کمتری دارد، با روش‌های خوشه‌بندی اتصال واحد و اتصال میانگین جدا می‌شود و از آن برای تشکیل سبد سرمایه‌گذاری استفاده می‌شود.

درخت تجربی، بیشتر با یک شبکه پیچیده منطبق است و الگوهای تک عاملی برای توصیف تمامی مشخصات بازار، کارایی مناسبی ندارند [۲]. تولا و همکاران (۲۰۰۸) از سه روش نظریه ماتریس تصادفی، خوشه بندی اتصال میانگین و خوشه بندی اتصال واحد برای فیلتر کردن ماتریس همبستگی استفاده کردند و نتایج بهینه سازی سبد سرمایه گذاری براساس هر کدام از روش ها را با هم مقایسه کردند. آن ها این گونه نتیجه گرفتند که روش های خوشه بندی، عملکرد بهتری نسبت به روش نظریه ماتریس تصادفی داشته است. همچنین به دلیل تحلیل حساسیت نتایج برای مجموعه های مختلفی از سهام و افق های زمانی متفاوت، بهینه سازی سبد سرمایه گذاری تکرار شده است و نتایج مشابهی به دست آمده است [۱۷]. سانداوال و همکاران (۲۰۱۴) در بازار کشور برزیل، با استفاده از نظریه ماتریس تصادفی، به پالایش و حذف نویز از ماتریس همبستگی اقدام کردند و براساس آن سبد سرمایه گذاری بهینه سهام را تشکیل دادند. همچنین برای محاسبه ماتریس همبستگی نیز از رگرسیون استفاده کردند و این کار را برای دو بازه زمانی با نوسان های کم و زیاد تکرار کردند. نتایج آن ها نشان می دهد استفاده از رگرسیون و روش ماتریس تصادفی، قدرت پیش بینی ریسک را افزایش می دهد؛ اما در زمان هایی که بازار، نوسان های شدید دارد، این روش کارایی چندانی از خود نشان نمی دهد [۱۵].

پژوهش های اکونومیک در ایران، همچنان نوپا و محدود است. راعی و همکاران (۱۳۸۹) در پژوهشی با عنوان «تحلیل بازار بورس اوراق بهادار تهران با استفاده از شبکه های پیچیده مبتنی بر روش حد آستانه»، شبکه سهام موجود در بورس اوراق بهادار تهران و کمیت های مرتبط با آن را بررسی کرده اند. آن ها با

سپس با روش های معرفی شده مبتنی بر ماتریس تصادفی در پیشینه پژوهش، بخش نویزی اطلاعات را جدا کردند و نشان دادند این روش پیشنهادی برای فیلتر کردن ماتریس همبستگی مزیت بیشتری دارد [۱۶]. اما در کنار روش های مبتنی بر ماتریس تصادفی، در سال های اخیر مطالعاتی در زمینه فیلتر کردن ماتریس ضرایب همبستگی براساس روش های خوشه بندی نیز انجام شده است که از آن جمله به اونلا و همکاران (۲۰۰۲)، بونانو و همکاران (۲۰۰۳)، تولا و همکاران (۲۰۰۸) و سانداوال و همکاران (۲۰۱۴) می توان اشاره کرد. اونلا و همکاران (۲۰۰۲) برای ۱۱۶ شرکت از مجموعه اس اند پی ۵۰۰، با روش های خوشه بندی، درخت پوششی کمینه را بر ماتریس همبستگی تشکیل دادند و یک بخش بندی اقتصادی معنادار از آن استخراج کردند. یک درخت پوششی کمینه، ارتباط میان شرکت هایی را که با یکدیگر همبستگی بیشتری دارند با ترسیم یک یال میان آن ها مشخص می کند. هر چه شرکت ها بر درخت به یکدیگر نزدیک تر باشند، نشان دهنده تعلق آن ها به یک بخش اقتصادی مشترک است. با تعیین مرکز درخت پوششی کمینه و بررسی حرکت و ساختار آن در طول زمان، رفتار همبستگی میان شرکت ها را می توان بررسی کرد. اونلا و همکاران نشان دادند سبد سرمایه گذاری بهینه مارکوویتز در عمل همواره در حاشیه درخت قرار دارد. این یافته آنان نشان می دهد یافتن سبد سرمایه گذاری بهینه در صورت استفاده از یک روش خوشه بندی مناسب امکان پذیر است [۱۱]. بونانو و همکاران (۲۰۰۳) نیز با تشکیل درخت پوششی کمینه برای سهام ۱۰۷۱ شرکت در بورس اوراق بهادار نیویورک در بازه ای ۱۲ ساله و مقایسه آن با درخت حاصل از یک بازار شبیه سازی شده براساس الگوهای تک عاملی، نشان دادند مشخصات

نسبت به ماتریس همبستگی فیلترنشده، کمتر است. آن‌ها در پژوهش خود، حساسیت نتایج این روش را در افق‌های زمانی و تعداد سهام متفاوت بررسی نکردند و فقط برای یک پنجره زمانی خاص ادعا کردند که نتایج آزمون در سطح ۹۹ درصد تأیید می‌شود [۱۰].

روش پژوهش

محور اصلی پژوهش حاضر، بر تخمین ماتریس ضرایب همبستگی مبتنی است. ماتریس ضرایب همبستگی، ماتریسی است که درایه d_{ij} در آن، ضریب همبستگی بین بازده دارایی i ام و بازده دارایی j ام است. از این ماتریس در بهینه‌سازی سبد سرمایه‌گذاری استفاده می‌شود. اگر با روش‌های فیلتر کردن، از نويز اطلاعات این ماتریس‌ها کاسته شود، بهینه‌سازی نتایج بهتری را به بار می‌آورد. در این پژوهش برای فیلتر کردن ماتریس ضرایب همبستگی از دو روش خوشه‌بندی اتصال میانگین و اتصال واحد استفاده می‌شود. ماتریس فیلترشده در بهینه‌سازی سبد سرمایه‌گذاری به کار برده می‌شود و عملکرد سبد سرمایه‌گذاری با دو معیار ارزیابی می‌شود. برای پاسخ به این پرسش که آیا روش‌های فیلتر کردن از نظر معیارهای معرفی شده، سبدهای سرمایه‌گذاری موفق‌تری نسبت به روش بدون فیلتر کردن تشکیل می‌دهند، یک‌بار در فضای داده‌ها و نمونه، هر سه روش اتصال واحد، اتصال میانگین و بدون فیلتر کردن اعمال می‌شود و مقادیر به‌دست آمده هر یک از دو معیار در این روش‌ها با همدیگر مقایسه می‌شوند؛ سپس برای تحلیل حساسیت این نتایج نسبت به شرایط مختلف (از نظر تعداد سهام و افق‌های سرمایه‌گذاری) یک مطالعه بوت استرپینگ^۱ با نمونه‌برداری داخلی^۲ انجام خواهد شد.

محاسبه بازده لگاریتمی و همبستگی ۳۲۵ سهم بازار، ماتریس همبستگی را تعیین کردند و سپس با اعمال حد آستانه برای قبول یا رد همبستگی میان دو سهم، شبکه‌ای را تشکیل دادند که هر گره آن، یک سهم و هر یال، نشان‌دهنده وجود همبستگی میان آن دو سهم است. آن‌ها در ادامه با تغییر حد آستانه، سه موضوع «توزیع درجات شبکه سهام»، «عناصر مستقل» و «ضریب خوشگی» را در این شبکه بررسی کردند. آن‌ها بیان کردند ریسک نظام‌مند بازار که دو معیار تعداد عناصر مستقل و ضریب خوشگی، سنجه‌ای از آن است، در حد آستانه‌ای در حدود میانگین ضرایب همبستگی، رفتار کاهنده شدیدی دارد [۱۴]. محمدی و خجسته (۱۳۸۸) در پژوهشی با عنوان «بهنه‌سازی سبد سرمایه‌گذاری سهام براساس ماتریس کوواریانس نوفه‌زدایی شده»، با توجه به مسأله وجود نويز در اطلاعات، برای فیلتر کردن ماتریس کوواریانس بازده قیمت سهام، از روشی مبتنی بر نظریه ماتریس تصادفی استفاده کردند و آن را در بهینه‌سازی سبد سرمایه‌گذاری به کار بردند. آن‌ها برای ۷۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران، ماتریس همبستگی را محاسبه کردند؛ سپس با استفاده از یک ماتریس تصادفی در بازه زمانی مشابه، ماتریس همبستگی را فیلتر کرده و از آن در بهینه‌سازی سبد سرمایه‌گذاری بازار در بازه فروردین ۱۳۸۳ تا خرداد ۱۳۸۷ استفاده کردند. آن‌ها برای بررسی عملکرد این روش، بهینه‌سازی را هم با ماتریس همبستگی فیلترنشده و هم با ماتریس همبستگی فیلترشده، فقط در یک افق زمانی انجام دادند و با آزمون فشر برای مقایسه واریانس‌ها به این نتیجه رسیدند که در سطح اطمینان ۹۵ درصد، ریسک مجموعه سبدهای سرمایه‌گذاری به‌دست آمده براساس ماتریس همبستگی فیلترشده

1. Bootstrapping
2. In-sample

در روش اتصال واحد، ابتدا لازم است معیار شباهت میان اعضا، S تعیین شود. این معیار در اینجا ضریب همبستگی میان هر زوج از اعضای بازده سهام است. پس از محاسبه شباهت میان هر دو عضو، S_{ij} ها به صورت نزولی مرتب می شوند. در فرایند خوشه بندی با شروع از دو عضو با بیشترین شباهت و تکرار مراحل که توضیح داده خواهند شد، اعضا در خوشه های مختلف قرار می گیرند. در هر مرحله وقتی دو عضو، یا یک عضو و یک خوشه یا دو خوشه p و q در یک خوشه جدید قرار می گیرند، شباهت میان سایر اعضا با خوشه جدید باید دوباره محاسبه شود. شباهت میان خوشه های t و r با استفاده از معیار ضریب همبستگی و با رابطه (۴) محاسبه می شود:

$$S_{tr} = \max(S_{pr}, S_{qr}) \quad (4)$$

با تکرار این روند $N - 1$ عضو منحصربه فرد از $\frac{N(N-1)}{2}$ عضو اولیه انتخاب می شوند. روش اتصال میانگین نیز مشابه روش اتصال واحد است؛ با این تفاوت که شباهت بین یک عضو و نزدیک ترین خوشه با محاسبه میانگین شباهت های میان آن عضو و هر یک از اعضای آن خوشه به دست می آید؛ به بیان دیگر، S_{tr} بین دو خوشه t و r از میانگین شباهت هریک از اعضای این دو خوشه به دست می آید. اگر ماتریس ورودی C و ماتریس فیلتر شده، روش اتصال واحد C_{SL} و روش اتصال میانگین C_{AL} باشد، مراحل اجرای روش اتصال میانگین و اتصال واحد با جزئیات بیشتر به صورت زیر خواهد بود [۹]:

۱- ماتریس B را برابر با ماتریس C در نظر بگیرید.

۲- بیشترین همبستگی b_{hk} را در ماتریس B پیدا کنید. یادآوری می شود که h و k می توانند عضوی ساده (یعنی خوشه ای با یک عضو) و یا خوشه

هدف، خوشه بندی تفکیک N متغیر و قراردادن آن ها در بخش های جداگانه ای است که براساس مشخصات هر متغیر و شباهت های آن ها، معنادار باشند؛ برای مثال در خوشه بندی مبتنی بر ضرایب همبستگی که در پژوهش حاضر استفاده می شود، ضریب همبستگی دو سری زمانی، نشان دهنده میزان شباهت آن ها به یکدیگر است. خوشه بندی مبتنی بر همبستگی در مطالعات متعددی برای یافتن ساختار سلسله مراتبی سهام موجود در سبد سرمایه گذاری استفاده شده است [۳ و ۲]. این خوشه بندی به عنوان روشی برای فیلتر کردن ماتریس همبستگی نیز می تواند استفاده شود؛ به این معنا که پس از اجرای این روش، درایه های با احتمال نویزی بودن، حذف می شوند و تعداد درایه های منحصربه فرد ماتریس مذکور، کاهش می یابد؛ برای مثال در روش خوشه بندی، اتصال واحد تعداد درایه های منحصربه فرد ماتریس فیلتر شده $N - 1$ است، در حالی که تعداد درایه های منحصربه فرد ماتریس اصلی $\frac{N(N-1)}{2}$ است [۶]. تاکنون روش های خوشه بندی متنوعی معرفی شده اند؛ اما از آنجا که لازم است در بهینه سازی سبد سرمایه گذاری، ماتریس همبستگی فیلتر شده مثبت معین باشد، بسیاری از این روش ها قابل استفاده نخواهند بود. تامینلو و همکاران (۲۰۱۰) نشان داده اند در هر دو روش خوشه بندی اتصال واحد و اتصال میانگین، ماتریس فیلتر شده، مثبت معین باقی می ماند و به همین علت از این دو روش برای فیلتر کردن ماتریس ضرایب همبستگی استفاده می شود.

۱. ماتریس ضرایب همبستگی و ماتریس کوواریانس طبق تعریف مثبت معین هستند و ماتریس های تجربی آن ها نیز باید مثبت معین باشند. این شرط در صورتی صادق است که تمامی بردارهای ویژه این ماتریس بزرگتر از صفر باشند.

$$\begin{cases} b_{qj} = \frac{n_h b_{hj} + n_k b_{kj}}{n_h + n_k}, & \text{if } j \notin h \text{ and } (6) \\ b_{ij} = b_{ij}, & \text{otherwise} \end{cases}$$

در روابط (۵) و (۶) n_h و n_k به ترتیب، تعداد اعضای خوشه‌های h و k پیش از ادغام است. باید دقت داشت اگر اندازه ماتریس B ، برابر با $m \times m$ باشد، پس از بازتعریف، این اندازه معادل $(m - 1) \times (m - 1)$ خواهد بود؛ زیرا خوشه‌های h و k در هم ادغام می‌شوند و خوشه q را می‌سازند.

۵- اگر اندازه ماتریس B بزرگ‌تر از ۱ باشد، مراحل بالا باید از گام ۲ دوباره تکرار شوند و در غیر این صورت مراحل در همین گام به پایان می‌رسد. برای مثال، ماتریس ضرایب همبستگی را که در شکل (۱) نشان داده شده است، در نظر بگیرید.

$$C = \begin{pmatrix} 1.000 & 0.413 & 0.518 & 0.543 & 0.529 & 0.341 & 0.271 & 0.231 & 0.412 & 0.294 \\ & 1.000 & 0.471 & 0.537 & 0.617 & 0.552 & 0.298 & 0.475 & 0.373 & 0.270 \\ & & 1.000 & 0.547 & 0.592 & 0.400 & 0.258 & 0.349 & 0.370 & 0.276 \\ & & & 1.000 & 0.664 & 0.422 & 0.347 & 0.351 & 0.414 & 0.269 \\ & & & & 1.000 & 0.533 & 0.344 & 0.462 & 0.440 & 0.318 \\ & & & & & 1.000 & 0.305 & 0.582 & 0.355 & 0.245 \\ & & & & & & 1.000 & 0.193 & 0.533 & 0.591 \\ & & & & & & & 1.000 & 0.258 & 0.166 \\ & & & & & & & & 1.000 & 0.590 \\ & & & & & & & & & 1.000 \end{pmatrix}$$

شکل (۱) ماتریس ضرایب همبستگی

(مجموعه‌ای از اعضا) باشند. حال برای $\forall i \in h$ و $\forall j \in k$ درایه‌های ماتریس C_{SL} یا C_{AL} ، یعنی ρ_{ij}^k برابر خواهند بود با: $\rho_{ij}^k = \rho_{ji}^k = b_{hk}$

۳- خوشه h و خوشه k را در یک خوشه q ادغام کنید. در نتیجه این ادغام، یک گره در درخت سلسله‌مراتبی تعریف می‌شود که دو خوشه h و k را با همبستگی b_{hk} به یکدیگر متصل می‌کند.

۴- ماتریس B را برای روش اتصال واحد با استفاده از رابطه (۵) و برای روش اتصال میانگین با استفاده از رابطه (۶) بازتعریف کنید.

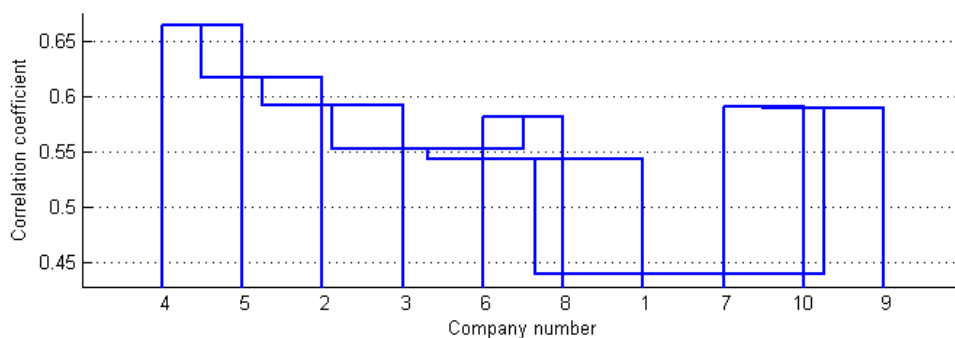
$$\begin{cases} b_{qj} = \max(b_{hj}, b_{kj}), & \text{if } j \notin h \text{ and } j \notin k \\ b_{ij} = b_{ij}, & \text{otherwise} \end{cases} \quad (5)$$

پس از انجام مراحل فوق، درخت سلسله‌مراتبی را براساس خوشه‌بندی انجام‌شده می‌توان ترسیم کرد. در شکل (۲) درخت مربوط به روش اتصال واحد، با نمودار دندروگرام^{۲۱} نشان داده شده است و خوشه‌های تشکیل‌شده را می‌توان مشاهده کرد.

1. Dendrogram

۲. نمودار دندروگرام به صورت بصری درخت سلسله‌مراتبی را نشان می‌دهد و می‌توان خوشه‌های تشکیل‌شده را مشاهده کرد. هر خوشه نشان‌دهنده تمرکز ارتباطات در یک ناحیه است. در محور افقی، شماره هر گره (در اینجا هر سهم) و در محور عمودی، طول یال (در اینجا

ضریب همبستگی) متصل‌کننده دو گره (و یا یک خوشه و یک گره، و یا دو خوشه) به یکدیگر را نشان می‌دهد.



شکل (۲) نمودار دندروگرام درخت خوشه‌ای سلسله‌مراتبی

و اتصال میانگین به ترتیب در شکل (۳) و شکل (۴) آمده است.

برای مثال، سهام شماره ۴، ۵، ۲ و ۳ یک خوشه و سهام شماره ۷، ۱۰، ۹ و نیز یک خوشه دیگر را تشکیل داده‌اند. ماتریس فیلترشده حاصل از روش اتصال واحد

$$C_{SL}^< = \begin{pmatrix} 1.000 & 0.543 & 0.543 & 0.543 & 0.543 & 0.543 & 0.440 & 0.543 & 0.440 & 0.440 \\ & 1.000 & 0.592 & 0.617 & 0.617 & 0.552 & 0.440 & 0.552 & 0.440 & 0.440 \\ & & 1.000 & 0.592 & 0.592 & 0.552 & 0.440 & 0.552 & 0.440 & 0.440 \\ & & & 1.000 & 0.664 & 0.552 & 0.440 & 0.552 & 0.440 & 0.440 \\ & & & & 1.000 & 0.552 & 0.440 & 0.552 & 0.440 & 0.440 \\ & & & & & 1.000 & 0.440 & 0.582 & 0.440 & 0.440 \\ & & & & & & 1.000 & 0.440 & 0.590 & 0.591 \\ & & & & & & & 1.000 & 0.440 & 0.440 \\ & & & & & & & & 1.000 & 0.590 \\ & & & & & & & & & 1.000 \end{pmatrix}$$

شکل (۳) ماتریس ضرایب همبستگی فیلترشده با روش اتصال واحد

$$C_{AL}^< = \begin{pmatrix} 1.000 & 0.501 & 0.501 & 0.501 & 0.501 & 0.412 & 0.308 & 0.412 & 0.308 & 0.308 \\ & 1.000 & 0.536 & 0.577 & 0.577 & 0.412 & 0.308 & 0.412 & 0.308 & 0.308 \\ & & 1.000 & 0.536 & 0.536 & 0.412 & 0.308 & 0.412 & 0.308 & 0.308 \\ & & & 1.000 & 0.664 & 0.412 & 0.308 & 0.412 & 0.308 & 0.308 \\ & & & & 1.000 & 0.412 & 0.308 & 0.412 & 0.308 & 0.308 \\ & & & & & 1.000 & 0.308 & 0.582 & 0.308 & 0.308 \\ & & & & & & 1.000 & 0.308 & 0.562 & 0.591 \\ & & & & & & & 1.000 & 0.308 & 0.308 \\ & & & & & & & & 1.000 & 0.562 \\ & & & & & & & & & 1.000 \end{pmatrix}$$

شکل (۴) ماتریس ضرایب همبستگی فیلترشده با روش اتصال میانگین

حاصل شده، قابل اعتمادتر است. با حذف اثر اندازه σ_p سبد سرمایه‌گذاری، زمانی قابل اعتمادتر است که از α کوچک‌تری برخوردار باشد:

$$\alpha = \frac{|\sigma_p - \hat{\sigma}_p|}{\sigma_p} \quad (۸)$$

معیار دوم ارزیابی عملکرد روش‌های فیلتر کردن، ریسک تحقق‌یافته سبد سرمایه‌گذاری، $\hat{\sigma}_p$ است که میزان ریسک پذیر بودن سبد سرمایه‌گذاری را نشان می‌دهد. هرچه ریسک تحقق‌یافته سبد سرمایه‌گذاری کمتر باشد، نشان‌دهنده این است که سبد سرمایه‌گذاری معرفی شده در ابتدای افق سرمایه‌گذاری، در پایان، ریسک کمتری را متحمل شده و عملکرد بهتری داشته است.

یافته‌ها

در این مطالعه ۸۰ شرکت پر معامله بورس اوراق بهادار تهران در بازه زمانی فروردین ۱۳۸۵ تا اسفند ۱۳۹۲ انتخاب شده‌اند. انتخاب شرکت‌های معامله به این دلیل است که بسیاری از شرکت‌ها در طول سال، تعداد روزهای معاملاتی کمی دارند و یا به‌علت قرار گرفتن در صف‌های خرید و فروش و به‌وجود آمدن پدیده گره معاملاتی، قیمت پایانی آن‌ها به‌طور مصنوعی تغییری نمی‌کند و بازه برای چندین روز صفر است. استفاده از این داده‌ها خود به نوبت اطلاعات موجود می‌افزاید؛ بنابراین، در این پژوهش از داده‌های شرکت‌هایی استفاده شده است که از بیشترین تعداد روزهای معاملاتی برخوردار باشند و در این میان، کمتر دچار گره معاملاتی شده باشند. برای جمع‌آوری اطلاعات از نرم‌افزار TSE Client شرکت مدیریت فناوری بورس تهران و برای پردازش و انجام محاسبات از نرم‌افزار MATLAB استفاده شده است.

همان‌طور که ملاحظه می‌شود در هر دو ماتریس فیلتر شده، تعداد درایه‌های منحصربه‌فرد $N - 1$ است که در مقایسه با ماتریس اصلی کمتر است.

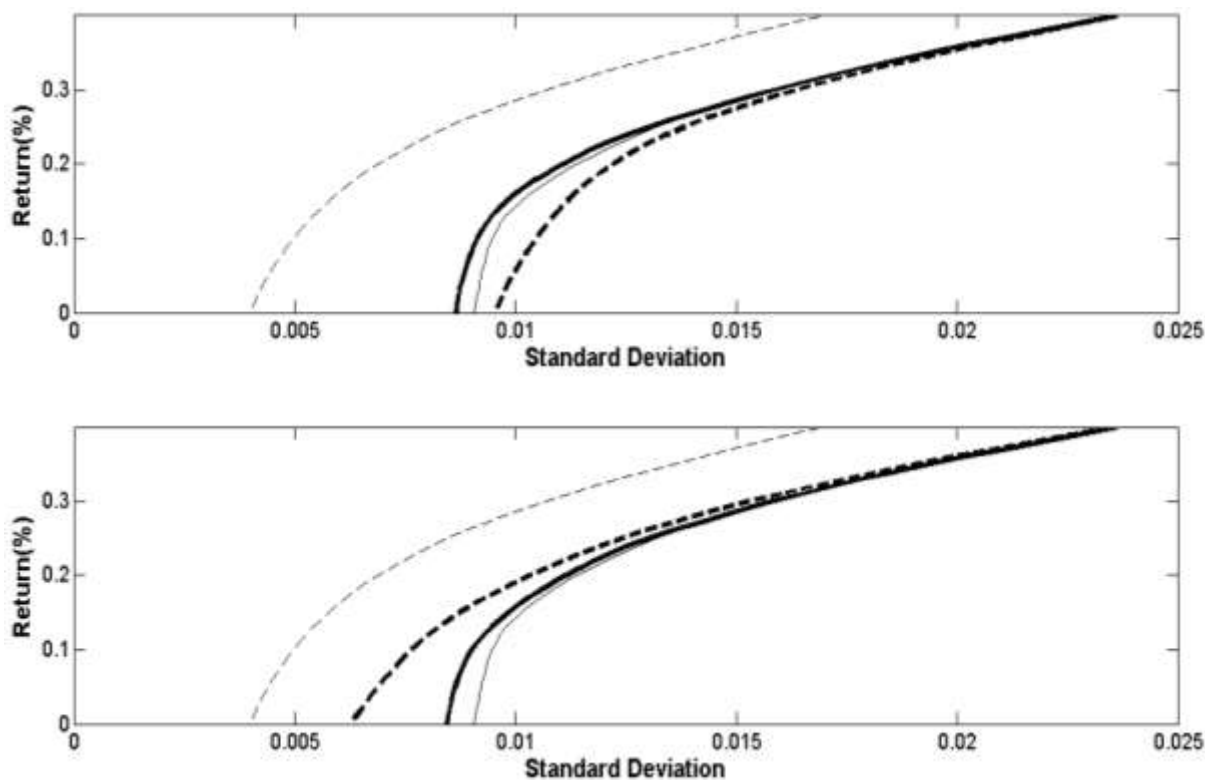
برای اندازه‌گیری کیفیت عملکرد روش‌های خوشه‌بندی، بازه زمانی مورد بررسی به دو قسمت تقسیم می‌شود. به این منظور ماتریس همبستگی براساس اطلاعات T روز پیش از t_0 تخمین زده می‌شود؛ اما میانگین بازده‌ها، r_i و انحراف معیار آن‌ها، σ_i در F روز پس از t_0 محاسبه خواهد شد. با این داده‌ها، ماتریس کوواریانس و سبد سرمایه‌گذاری بهینه متشکل از N دارایی در زمان t_0 تعیین شود. با این کار درحقیقت مسأله تخمین ماتریس همبستگی، از مسأله تخمین میانگین بازده و انحراف معیار بازده جدا می‌شود؛ به بیان دیگر، فرض می‌شود سرمایه‌گذار از میانگین بازده و انحراف معیار بازده، پیش‌بینی کاملی دارد و عدم قطعیت فقط به ماتریس همبستگی مربوط است. با داشتن وزن دارایی‌ها در سبد سرمایه‌گذاری بهینه p_i^* و ماتریس کوواریانس تحقق‌یافته $\hat{\sigma}_{ij}$ که براساس مشاهدات زمان t_0 تا $t_0 + F$ به دست آمده است، ریسک تحقق‌یافته سبد سرمایه‌گذاری را بازای بازده‌های متفاوت با استفاده از رابطه (۷) می‌توان محاسبه کرد.

$$\hat{\sigma}_p^2 = \sum_{i=1}^N \sum_{j=1}^N p_i^* p_j^* \hat{\sigma}_{ij} \quad (۷)$$

برای مقایسه عملکرد روش‌های فیلتر کردن، از دو معیار ارزیابی استفاده می‌شود. معیار نخست، قابل اعتماد بودن سبد سرمایه‌گذاری تشکیل شده را می‌سنجد. هرچه ریسک پیش‌بینی شده سبد سرمایه‌گذاری براساس ماتریس همبستگی σ_p ، به ریسک تحقق‌یافته سبد سرمایه‌گذاری $\hat{\sigma}_p$ نزدیک‌تر باشد، سبد سرمایه‌گذاری

سهام برخی شرکت‌ها در بعضی از روزهای بازه زمانی مدنظر معامله نشده‌اند، برای هر زوج سهام، کل مشاهدات آن بازه زمانی را نمی‌توان به کار برد؛ در نتیجه کوواریانس بازده سهام دوبه‌دو و با حذف روزهای بدون معامله محاسبه شده‌اند. منحنی ریسک-بازده بهینه پیش‌بینی شده با استفاده از ماتریس ضرایب همبستگی عادی (بدون فیلتر کردن) و ماتریس ضرایب همبستگی فیلتر شده و نیز منحنی ریسک-بازده محقق شده هر کدام محاسبه شده است. نتایج دو روش اتصال واحد و اتصال میانگین در شکل (۵) ملاحظه می‌شود.

برای استفاده از ماتریس فیلتر شده با روش‌های خوشه‌بندی باید قید مهم مثبت معین (یا مثبت نیمه‌معین) بودن این ماتریس برقرار باشد. با حل مسأله بهینه‌سازی سبد سرمایه‌گذاری با روش مارکوویتز بازای هر بازه سبد سرمایه‌گذاری، r_p یک مقدار برای ریسک سبد سرمایه‌گذاری محاسبه خواهد شد که ریسک پیش‌بینی شده، σ_p شناخته می‌شود و با داشتن اوزان دارایی‌ها در هر سبد سرمایه‌گذاری سهام، با استفاده از رابطه (۷)، ریسک تحقق یافته، $\hat{\sigma}_p$ به دست می‌آید. پس از محاسبه بازده روزانه لگاریتمی، انحراف معیار بازده‌ها و کوواریانس میان آن‌ها محاسبه می‌شود. از آنجا که



شکل (۵) منحنی‌های ریسک-بازده بدون فیلترینگ (خطوط نازک)، روش اتصال واحد (خطوط ضخیم نمودار بالا) و روش اتصال میانگین (خطوط ضخیم نمودار پایین)

سرمایه‌گذاری پیش‌بینی شده مربوط است. همان‌طور که ملاحظه می‌شود منحنی‌های مبتنی بر ماتریس همبستگی

خطوط ممتد در شکل (۵) به سبدهای سرمایه‌گذاری محقق شده و خطوط ناپیوسته مربوط به سبدهای

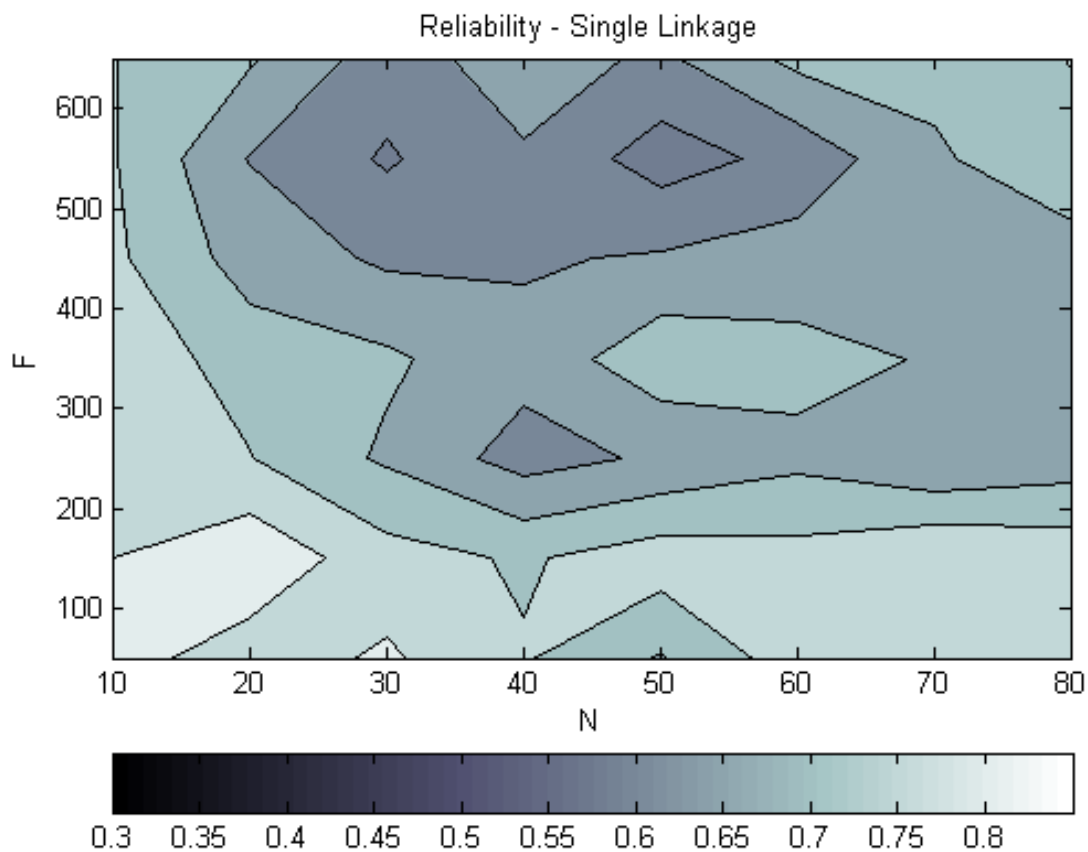
فیلترشده به یکدیگر نزدیک‌تر هستند و در ۱۰۰ درصد سبدهای سرمایه‌گذاری تشکیل شده قابل اعتمادتر از منحنی‌های عادی هستند. از لحاظ میزان ریسک پذیر بودن سبد سرمایه‌گذاری نیز در بازده‌های پایین به ترتیب، روش‌های اتصال واحد و اتصال میانگین، عملکرد بهتری داشته‌اند؛ اما در سطوح بالای بازده، هر سه روش تقریباً به نتایج مشابهی رسیده‌اند. در مجموع از این نظر، روش اتصال واحد در ۷۵/۶ درصد و روش اتصال میانگین در ۷۳/۱ درصد سبدهای سرمایه‌گذاری تشکیل شده، عملکرد بهتری داشته‌اند.

برای تحلیل حساسیت این نتایج در حالات و شرایط مختلف، مطالعه بوت استرپینگ با نمونه‌برداری داخلی انجام می‌شود. برای هر زوج مرتب (N, F) که نشان‌دهنده تعداد دارایی‌ها و F نشان‌دهنده افق سرمایه‌گذاری است، ۵۰ مجموعه متشکل از N دارایی و ۵۰ زمان t_0 انتخاب می‌شود. برای هر مجموعه، ۱۰ مقدار برای بازده مورد انتظار سبد سرمایه‌گذاری بهینه، r_p انتخاب می‌شود. بدین منظور بالاترین بازده دارایی‌های مجموعه و بازده کم ریسک‌ترین دارایی مجموعه پیدا می‌شود و با فواصل مساوی، از میان این دو حد بالا و پایین، ۱۰ مقدار استخراج می‌شود. با توجه به روند توضیح داده شده، در مجموع برای هر زوج مرتب (N, F) ، حالت و در کل با احتساب مقادیر ۵۰، ۱۵۰، ۲۵۰، ۳۵۰، ۴۵۰، ۵۵۰ و ۶۵۰ برای F و ۱۰، ۲۰، ۳۰، ۴۰، ۵۰، ۶۰، ۷۰ و ۸۰ برای N ، ۲۸۰۰ مجموعه و ۲۸ هزار سبد سرمایه‌گذاری بهینه بررسی می‌شود.

برای مطالعه میزان قابل اعتماد بودن سبدهای سرمایه‌گذاری تشکیل شده، در هر کدام از مجموعه‌های گزینش شده، روش فیلتر کردن اتصال واحد و اتصال میانگین، پیاده می‌شود و ماتریس همبستگی فیلترشده محاسبه می‌شود. سبد سرمایه‌گذاری بهینه با ماتریس

همبستگی فیلترشده حاصل از این دو روش و نیز با ماتریس همبستگی فیلتر نشده، برای هر کدام از بازده‌های مورد انتظار، محاسبه و مقدار α برای آن‌ها تعیین می‌شود؛ سپس تعداد دفعاتی که در آن $\alpha_{SL} < \alpha_M$ یعنی شاخص قابل اعتماد بودن روش عادی کمتر از روش اتصال واحد بوده است و نیز به‌طور مشابه تعداد دفعاتی که $\alpha_{AL} < \alpha_M$ یعنی شاخص قابل اعتماد بودن روش عادی کمتر از روش اتصال میانگین بوده است، شمرده می‌شود. شکل (۳) یک نمودار تراکمی^۱ است که براساس درصد پیروزی روش اتصال واحد بر روش عادی از نظر معیار قابل اعتماد بودن در حالات مختلف ترسیم شده است. نمودار تراکمی، یک نوع نمودار سه بعدی است که به هر زوج مرتب حاصل از عناصر محورهای افقی و عمودی، یک رنگ منتسب می‌کند. هر رنگ در یک نمودار میله‌ای جداگانه به عددی مشخص نسبت داده شده است. برای مثال، شکل (۶) نشان می‌دهد در سبدهای سرمایه‌گذاری بهینه با ۲۰ دارایی و افق زمانی ۱۰۰ روز، روش اتصال واحد بیش از ۸۰ درصد از روش عادی قابل اعتمادتر است. شکل (۴) نیز همین موارد را برای روش اتصال میانگین به تصویر کشیده است.

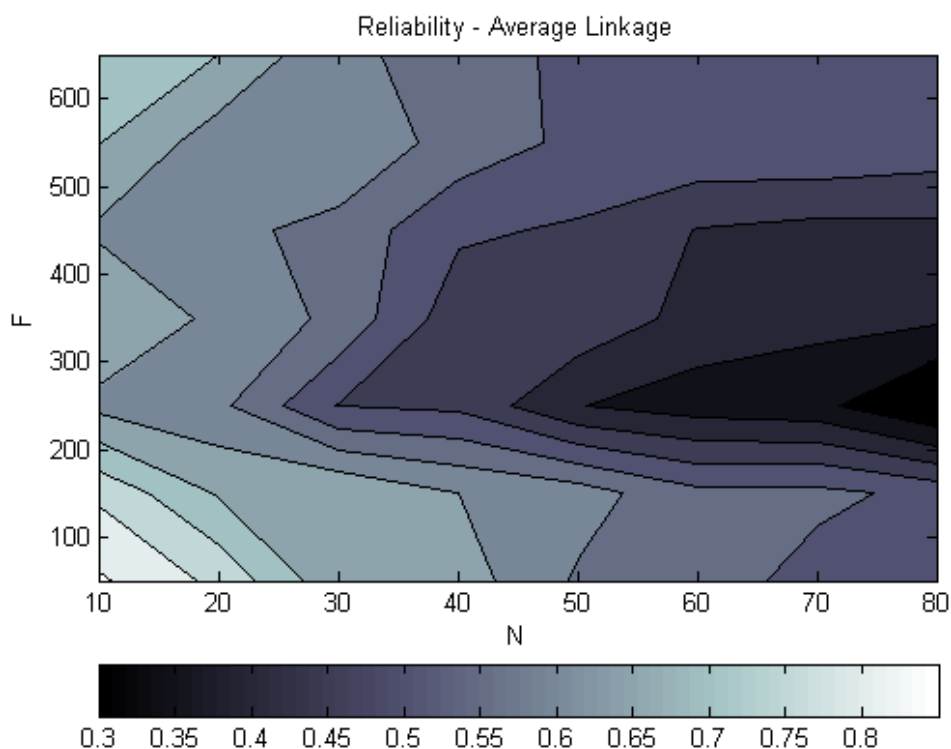
۱. Density Plot



شکل (۶) درصد پیروزی روش اتصال واحد بر روش عادی به صورت تابعی از N و F ، از لحاظ قابل اعتماد بودن

حالت، ۸۳ درصد است و در ۱۰۰ درصد کل حالات، عملکرد بهتری را نسبت به روش بدون فیلتر کردن نشان می‌دهد. نتایج شکل (۴) نشان می‌دهد عملکرد روش اتصال میانگین از لحاظ قابل اعتماد بودن، ضعیف‌تر است و بیشتر برای سبدهای سرمایه‌گذاری با اندازه کوچک‌تر از ۳۰ کارایی دارد. درصد پیروزی این روش در بهترین حالت، ۸۵ درصد است و در ۷۳ درصد کل حالات، عملکرد بهتری را نسبت به روش بدون فیلتر کردن نشان می‌دهد.

نرخ پیروزی بیش از ۵۰ درصد، نشان‌دهنده عملکرد بهتر هر کدام از روش‌ها نسبت به روش عادی (که از ماتریس فیلتر نشده استفاده می‌کند) است. همان‌طور که در شکل (۷) مشخص است، در روش اتصال میانگین، در تمامی حالات مختلف (N, F) ، این روش از لحاظ قابل اعتماد بودن بر روش عادی برتر است. این موضوع برای آن دسته از سبدهای سرمایه‌گذاری که افق سرمایه‌گذاری در آن‌ها کمتر از ۴۰۰ روز باشد، در خور توجه است. درصد پیروزی این روش در بهترین



شکل (۷) درصد پیروزی روش اتصال میانگین بر روش عادی به صورت تابعی از N و F از لحاظ قابل اعتماد بودن

کمتری دارند و سبدهای سرمایه‌گذاری با تعداد دارایی کمتر از ۳۰ نیز بدون توجه به افق سرمایه‌گذاری، ریسک تحقق‌یافته کمتری نشان می‌دهند. به طور کلی روش اتصال واحد در ۶۹ درصد و روش اتصال میانگین در ۶۷ درصد موارد عملکرد بهتری را نشان می‌دهد.

برای بررسی دقیق‌تر گستره بازده‌های مورد انتظار، دو گروه برای بازده زیاد و بازده‌های کم در نظر گرفته می‌شود. برای این کار بیشترین بازده دارایی‌های مجموعه و بازده کم ریسک‌ترین دارایی مجموعه پیدا می‌شود و با فواصل مساوی از میان این دو حد بالا و پایین، ۱۰ مقدار استخراج می‌شود. ۵ مقدار بزرگ‌تر در گروه بازده‌های زیاد و ۵ مقدار کوچک‌تر در گروه بازده‌های کم قرار داده می‌شوند و سپس رفتار روش‌های اتصال واحد و اتصال میانگین در هر یک از

برای مطالعه ریسک تحقق‌یافته در سبدهای سرمایه‌گذاری تشکیل شده برای هر مجموعه در سطوح بازده مختلف، میان روش‌های اتصال واحد، اتصال میانگین و روش بدون فیلتر کردن مقدار ریسک تحقق‌یافته سبد سرمایه‌گذاری، با یکدیگر مقایسه می‌شود. تعداد دفعاتی که ریسک تحقق‌یافته روش اتصال واحد، کمتر از روش عادی بوده است و نیز تعداد دفعاتی که ریسک تحقق‌یافته روش اتصال میانگین، کمتر از روش عادی بوده است، شمرده می‌شود. از نظر ریسک‌پذیر بودن، هر دو روش اتصال واحد و اتصال میانگین، عملکردی به نسبت مشابه دارند. آن دسته از سبدهای سرمایه‌گذاری که با افق سرمایه‌گذاری بیش از ۴۰۰ روز تشکیل می‌شوند، بدون توجه به تعداد دارایی‌های داخل سبد سرمایه‌گذاری نسبت به روش بدون فیلتر کردن، ریسک تحقق‌یافته

و روش‌های ذکر شده در آن‌ها پیاده شد. نتایج به دست آمده نشان می‌دهد روش‌های فیلتر کردن ارائه شده از نظر قابل اعتماد بودن و از نظر ریسک تحقق یافته، در مجموع عملکرد بهتری نسبت به روش‌های بدون فیلتر کردن از خود نشان می‌دهند. این نتایج که بر اساس اطلاعات برداشت شده از بورس اوراق بهادار تهران به دست آمده است، با نتایج پژوهش‌های مشابه در بورس نیویورک مطابقت دارد [۱۴].

با توجه به نتایج حاصل از این پژوهش پیشنهاد می‌شود از تلفیق هر دو روش اتصال میانگین و اتصال واحد، یک سیستم بهینه‌سازی پویا طراحی شود. همان‌طور که در نتایج مشاهده می‌شود در هر ترکیب مختلف از تعداد سهام و افق سرمایه‌گذاری، یکی از این روش‌ها بر دیگری برتری دارد. با در نظر گرفتن ملاحظات ریسک‌گریزی سرمایه‌گذار نیز معیارهای انتخاب روش فیلتر کردن گسترده‌تر می‌شود و انتظار می‌رود این سیستم با در نظر گرفتن همه آن‌ها، بهترین روش را در شرایط مختلف انتخاب کند. سیستم بهینه‌سازی پویا باید بتواند در طول زمان با رصد کردن بازار و ماتریس‌های همبستگی فیلتر شده، به‌طور مستمر سبد سرمایه‌گذاری مناسب سرمایه‌گذار را انتخاب کند و آن را در معرض کمترین ریسک ممکن قرار دهد.

منابع

- [1] Bellman, R. (1961). Adaptive control processes—A guided tour. *Naval Research Logistics Quarterly*, 8(3): 315–316.
- [2] Bonanno, G., Caldarelli, G., Lillo, F., & Mantegna, R. N. (2003). Topology of correlation based minimal spanning trees in real and model markets. *Physical Review E*, 68(4), 4.
- [3] Bonanno, G., Lillo, F., & Mantegna, R. N. (2000). High-frequency Cross-correlation

این گروه‌ها بررسی می‌شود. در بازده‌های مورد انتظار پایین، عملکرد دو روش از نظر ریسک تحقق یافته بسیار بهبود می‌یابد، به گونه‌ای که هر دو روش در ۷۳ درصد موارد، عملکرد بهتری نسبت به روش بدون فیلتر کردن نشان می‌دهند. این در حالی است که در همین نرخ بازده‌ها، از نظر قابل اعتماد بودن، عملکرد روش اتصال واحد همچنان با ۷۵ درصد پیروزی قابل قبول است؛ اما عملکرد روش اتصال میانگین، افت شدیدی کرده است و عملکرد بهتری از روش بدون فیلتر کردن نمی‌تواند نشان دهد. در بازده‌های مورد انتظار بالا، عملکرد دو روش از نظر ریسک تحقق یافته، ضعیف‌تر از روش بدون فیلتر کردن است؛ به بیان دیگر، در سبدهای سرمایه‌گذاری با بازده مورد انتظار زیاد، استفاده از روش‌های فیلترینگ ماتریس ضرایب همبستگی به ریسک کمتری نمی‌انجامد؛ اما از نظر قابل اعتماد بودن سبدهای سرمایه‌گذاری، هر دو روش، عملکرد مناسبی دارند، به گونه‌ای که روش اتصال واحد در ۸۹ درصد و روش اتصال میانگین در ۱۰۰ درصد موارد بر روش عادی غلبه دارد.

نتیجه‌گیری و پیشنهادها

پژوهش حاضر با استفاده از ماتریس ضرایب همبستگی فیلتر شده، بهینه‌سازی سبد سرمایه‌گذاری را بررسی می‌کند. بدین منظور، دو روش مبتنی بر خوشه‌بندی اتصال واحد و اتصال میانگین برای فیلتر کردن استفاده شد و عملکرد سبدهای سرمایه‌گذاری بهینه‌شده حاصل از اعمال روش‌های مذکور در مقایسه با روش بدون فیلتر کردن تحلیل و بررسی شد. برای تحلیل حساسیت نتایج با استفاده از تکنیک بوت استرپینگ و باز نمونه‌گیری، تعداد زیادی سبد سرمایه‌گذاری با ابعاد و ویژگی‌های متنوع تشکیل

- Physical Journal B - Condensed Matter*, 30(3), 285–288.
- [12] Pafka, S., & Kondor, I. (2003). Noisy Covariance Matrices and Portfolio Optimization. *Physica A: Statistical Mechanics and Its Applications*, 319, 487–494. Statistical Mechanics; Portfolio Management.
- [13] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T., & Stanley, H. E. (2001). A Random Matrix Approach to Cross-Correlations in Financial Data. *Physical Review E*, 65(6): 20-38.
- [14] Raei, R., Jafari, G., Namaki, A. (2010). Analysis of Tehran Stock Exchange using complex network based on threshold method. *Journal of the Accounting and Auditing Review*, 17(62): 33-48. (in Persian)
- [15] Sandoval Jr., L., Bortoluzzo, A., & Venezuela, M. (2014). Not all that glitters is RMT in the forecasting of risk of portfolios in the Brazilian stock market. *Physica A: Statistical Mechanics*, 410: 94-109.
- [16] Sharifi, S., Crane, M., Shamaie, A., & Ruskin, H. (2004). Random matrix theory for portfolio optimization: a stability approach. *Physica A: Statistical Mechanics*, 335: 629–643.
- [17] Tola, V., Lillo, F., Gallegati, M., & Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1): 235–258.
- in a Set of Stocks. *Quantitative Finance*, 1(1), 96–104.
- [4] Drozd, S., Kwapien, J., Gruemmer, F., Ruf, F., & Speth, J. (2001). Quantifying dynamics of the financial correlations, *Physica A: Statistical Mechanics and its Applications*, 299(1): 144-153.
- [5] Galluccio, S., Bouchaud, J., & Potters, M. (1998). Rational Decisions, Random Matrices and Spin Glasses, *Physica A: Statistical Mechanics and its Applications*, 259(3): 449-456.
- [6] Gower, J., & Ross, G. (1969). Minimum spanning trees and single linkage cluster analysis. *Applied Statistics*, 18(1), 54–64.
- [7] Laloux, L., Cizeau, P., Bouchaud, J., & Potters, M. (1998). Noise Dressing of Financial Correlation Matrices, *Physical Review Letters*, 83(7), 1467-1470.
- [8] LaLoux, L., Cizeau, P., Potters, M., & Bouchaud, J. (2000). Random Matrix Theory and Financial Correlations. *International Journal of Theoretical and Applied Finance*, 3(03): 391–397.
- [9] Mantegna, R. N., Stanley, H. E., & Chriss, N. A. (2000). *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press.
- [10] Mohammadi, S., Khojasteh, M. (2009). Portfolio optimization based on filtered covariance matrix. *Economics Letters*, 72(5): 33-48. (in Persian)
- [11] Onnela, J.-P., Chakraborti, a., Kaski, K., & Kertiész, J. (2002). Dynamic asset trees and portfolio analysis. *The European*

