

پایدارسازی شاخه‌نگارهای مالی به‌عنوان روشی برای اندازه‌گیری تغییرات سیستمیک (مطالعه‌ای در شاخص‌های بورس اوراق بهادار تهران)

حجت‌الله صادقی*، مسلم نیلچی**

چکیده

تغییرات سیستمیک مالی، تأثیرات گوناگونی بر عناصر یک نظام مالی دارد. روش‌ها و نظریه‌های ریسک سیستمیک، ناظر به مدل‌بندی و تحلیل همین تغییرات هستند؛ از این‌رو اندازه‌گیری میزان ناپایداری یک سیستم مالی به‌عنوان متغیری پراهمیت در ریسک سیستمیک، هدف اصلی این پژوهش به‌شمار می‌رود. روش انتخابی برای این اندازه‌گیری به مفهوم دندروگرام‌ها (شاخه‌نگارها) متکی است. شاخه‌نگارها، اصلی‌ترین ابزار نگاره‌سازی در خوشه‌بندی‌های سلسله‌مراتبی هستند؛ بنابراین مقایسه شاخه‌نگارها متضمن مقایسه خوشه‌بندی‌های متناظر با آن‌ها است. در این پژوهش، داده‌های روزانه ۲۶ شاخص «بورس اوراق بهادار تهران» در ۲۴۱ روز از تاریخ ۱۳۹۶/۰۱/۰۵ تا ۱۳۹۷/۰۱/۰۵ به‌عنوان قلمرو پژوهش در نظر گرفته شده است. خوشه‌بندی‌های سلسله‌مراتبی ماهانه به روش وارد تدوین و با شاخص بیکر و شاخص ضریب همبستگی همسانی‌ها با یکدیگر مقایسه شد. از مقایسه شاخه‌نگارهای پیش از تحلیل PCA و پس از آن، میزان ناپایداری سیستم اندازه‌گیری شد. بر اساس داده‌های موجود و مطابق با نتایج، میزان درهم‌تنیدگی خوشه‌بندی‌ها قبل و بعد از PCA، ۴۳ درصد است. این نتیجه نشان می‌دهد که ۴۳ درصد از سیستم در قلمرو یادشده، ناپایدار و بنابراین میزان ثبات سیستم ۵۷ درصد بوده است.

کلیدواژه‌ها: خوشه‌بندی سلسله‌مراتبی؛ شاخه‌نگار (دندروگرام)؛ هم‌شاخه‌نگار (تانگلگرام)؛ تحلیل مؤلفه‌های اصلی.

تاریخ دریافت مقاله: ۱۳۹۷/۰۸/۱۹، تاریخ پذیرش مقاله: ۱۳۹۷/۱۱/۰۱.

* استادیار، دانشگاه یزد (نویسنده مسئول).

E-mail: sadeqi@yazd.ac.ir

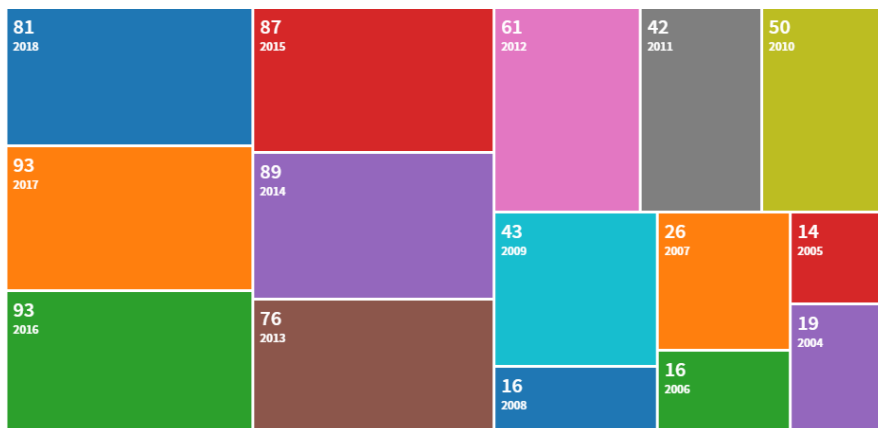
** دانشجوی دکتری، دانشگاه یزد.

۱. مقدمه

در پی بحران مالی سال ۲۰۰۷-۲۰۰۸، تلاش‌های بسیاری برای صورت‌بندی مفهوم تغییرات سیستمیک مالی صورت پذیرفته است. ریسک سیستمیک (نه سیستماتیک) به احتمال رخدادهایی اشاره دارد که به کنش - واکنش‌هایی در کل سیستم مالی منجر می‌شوند و به دنبال خود زیان‌هایی را به فعالان این سیستم تحمیل می‌کنند. کمی‌سازی میزان تغییرات سیستمیک از جمله مهم‌ترین دغدغه‌های نظریه‌های مالی در سال‌های اخیر بوده است [۳۲]. ریسک سیستمیک، بخشی از مفهوم کلی‌تری است که از آن با عنوان «ثبات مالی» یاد می‌شود. مفهوم ثبات مالی، در سال‌های اخیر مورد توجه پژوهشگران بوده است. نمودار ۱، تعداد مقالاتی که در مجلات نمایه WOS با عنوان «ثبات مالی» به چاپ رسیده است را نشان می‌دهد. با توجه به این نمودار، سال‌به‌سال بر بسامد پژوهش‌هایی مشتمل بر این موضوع افزوده شده است. یکی از چالش‌های این پژوهش‌ها، نحوه اندازه‌گیری ثبات مالی یک سیستم است. مسئله اساسی آن است که چگونه می‌توان، ثبات یک سیستم مالی را سنجید و از آن به‌عنوان یک متغیر وابسته یا متغیری مستقل در پژوهش‌های مالی استفاده کرد؛ به‌عبارت‌دیگر پس از کمی‌کردن مفهوم ثبات مالی، امکان طرح این پژوهش وجود دارد که چه عواملی بر ثبات مالی، مؤثر هستند و یا اینکه ثبات (یا بی‌ثباتی) مالی بر چه مسائل و متغیرهایی تأثیرگذار است. در این میان، تحلیل هر یک از زیرسیستم‌های مالی، مانند بورس اوراق بهادار یا ارزهای خارجی، بخشی از وضعیت کل سیستم را بازتاب می‌دهد. از تعداد تقریبی ۹۵۰ مقاله‌ای که با عنوان «ثبات مالی» در WOS چاپ شده‌اند، ۴۰۹ مقاله، ذیل دسته‌بندی مالی نمایه شده‌اند که نشان‌دهنده اهمیت این موضوع و رابطه آن به تصمیم‌گیری‌های مالی است. در همین بازه زمانی، نزدیک به ۸۰۰ مقاله با عنوان «ریسک سیستمیک» در نمایه یادشده به چاپ رسیده است که از نظر ترکیب و روند با مقالات ثبات مالی، مشابه هستند.

حدود ۳۶۰ مقاله موضوع ریسک سیستمیک را در یکی از حوزه‌های مالی بحث و بررسی کرده‌اند که نشان‌دهنده وزن این مفهوم در مبنی نظری امروز مالی است. در این پژوهش با تمرکز بر «بورس اوراق بهادار تهران»، موقعیت متریک شاخص‌های محاسبه‌شده آن به‌عنوان قلمرو مطالعه انتخاب شده‌اند. شاخص‌ها محاسبه‌شده، اندیکاتورهای صنایع مختلف هستند و دوری یا نزدیکی آن‌ها به یکدیگر نمایانگر فعل‌وانفعالات واقعی اقتصاد در آن صنایع است. مسئله اصلی آن است که آیا در طی زمان، موقعیت متریک شاخص‌های بورس تهران نسبت به یکدیگر ثابت بوده و بنابراین گونه‌ای پایداری بر سیستم حاکم بوده است یا برعکس، تغییرات این سیستم بر شکلی از ناپایداری دلالت دارد.

بر این اساس در بازه‌های مختلف زمانی با انجام خوشه‌بندی سلسله‌مراتبی شاخص‌ها، شاخه‌نگارهای متناظر با آن‌ها تدوین و با یکدیگر مقایسه می‌شوند. این مقایسه به روش‌های مختلفی صورت می‌پذیرد که در قسمت روش‌شناسی مطالعه، تشریح شده‌اند. در ادامه این نوشتار، ابتدا، ضمن تشریح مفهوم خوشه‌بندی سلسله‌مراتبی، برخی از کاربردهای متأخر آن در بازارهای مالی مرور می‌شود؛ سپس روش‌های مقایسه شاخه‌نگارها توضیح داده شده و نتایج عددی حاصل از آن‌ها در بورس تهران ارائه و تفسیر خواهد شد.



نمودار ۱. تعداد مقالات منتشرشده در WOS با موضوع ثابت مالی^۱ تا ژانویه ۲۰۱۹

۲. مبانی نظری و پیشینه پژوهش

در این بخش به مبانی نظری و مفاهیم ثبات مالی و ریسک سیستمیک اشاره و سپس به مفاهیم خوشه‌بندی و کاربردهای آن در مبانی نظری مالی پرداخته می‌شود. فصل مشترک تمام تعاریفی که برای ثبات مالی ارائه شده است، شرایطی است که احتمال رخداد بحران در سیستم مالی کم باشد. گاهی نیز از آن به «تاب‌آوری سیستم مالی در برابر تنیدگی» تعبیر می‌شود [۳]. یک سیستم مالی باثبات در تخصیص کارای منابع و مدیریت اثربخش ریسک‌ها، موفقیت بیشتری دارد. مفهوم مرتبط با ثبات مالی، ریسک سیستمیک است. دو مؤلفه بسیار مهم ریسک سیستمیک عبارت‌اند از: چگونگی اشاعه بحران‌ها و تنش‌های مالی از یک بخش اقتصاد به بخش دیگر و چگونگی فراگیری شوک ایجادشده در سیستم در طول زمان [۶].

۱. برای تدوین نمودار ۱، در پایگاه WOS، کلیدواژه‌های ثبات مالی (Financial Stability) جست‌وجو شد. این پایگاه این امکان را برای پژوهشگران فراهم می‌کند تا نمودارهای متعددی از جست‌وجوهای خود دریافت کنند. این نمودار نیز توسط پایگاه یادشده ارائه شده است.

بر اساس همین دو مؤلفه، سیاست‌گذاران و ناظران سیستم، تمهیدات مربوط به ریسک سیستمیک را تدارک می‌بینند. در مفهوم ثبات مالی، برخی به‌دنبال ساختن شاخص‌های کمی و برخی به‌دنبال نگاره‌سازی از آن بوده‌اند تا بتوان بر اساس این نگاره به دریافتی از وضعیت ثبات سیستم مالی دست یافت [۳۳]. در مبانی نظری ریسک سیستمیک نیز به همین ترتیب تلاش‌هایی برای کمی‌سازی و سنجش آن صورت گرفته است. به‌صورت مشخص برخی از این سنجها بر ویژگی‌های توپولوژیک شبکه‌های مالی استوار است و برخی رهیافت‌های دیگری مانند رهیافت مبتنی بر بازار را دارند [۱۶، ۲۱، ۲۵، ۲۶]. در رویکرد انتخابی این پژوهش نیز ویژگی‌های شبکه‌ای شاخص‌ها نسبت به یکدیگر اهمیت دارد؛ بنابراین در ادامه به مفاهیم خوشه‌بندی، ثبات خوشه‌بندی‌ها و شاخه‌نگار پرداخته می‌شود.

در مبانی نظری داده‌کاوی، خوشه‌بندی سلسله‌مراتبی یا همان تحلیل سلسله‌مراتبی خوشه‌ها، روشی از خوشه‌بندی است که به‌دنبال پیدا کردن سلسله‌مراتبی از خوشه‌هایی از مشاهدات، اشیا یا اعیان موردنظر است. نحوه جست‌وجوی این خوشه از یکی از این دو منطق پیروی می‌کند. در منطق پایین به بالا یا تجمعی، هر مشاهده‌ای در خوشه‌ای تک‌عضوی عضویت دارد؛ یعنی خوشه‌ای با حضور خودش. در گام بعد، شبیه‌ترین خوشه‌ها با یکدیگر ادغام می‌شوند تا به خوشه‌هایی بزرگ‌تر از مشاهدات منجر شود [۳۹]. معیارهای گوناگونی برای این شباهت‌یابی وجود دارد؛ بنابراین سلسله‌مراتب خوشه‌ها از طریق ترکیب خوشه‌های کوچک به بزرگ حاصل می‌شود. در منطق دوم، ابتدا همه مشاهدات در یک خوشه بزرگ قرار گرفته و با تجزیه آن خوشه، خوشه‌های کوچک‌تر، اما همگن‌تر، شکل می‌گیرند. سلسله‌مراتب خوشه‌ها در این روش، منطقی بالا به پایین یا تجزیه‌ای دارد. هم ادغام‌ها (در روش اول) و هم تجزیه‌ها (در روش دوم) بر اساس الگوریتم‌های خاصی صورت می‌پذیرند که در داده‌کاوی به آن‌ها پرداخته می‌شود [۱۵، ۲۹]. نکته دیگر به انتخاب معیاری برمی‌گردد که شباهت هر مشاهده با مشاهده دیگر را نشان می‌دهد. متریک‌های مختلفی برای اندازه‌گیری شباهت مشاهدات یا فاصله آن‌ها از یکدیگر وجود دارد که از جمله آن‌ها می‌توان به «فاصله اقلیدسی»، «فاصله منهن» و «فاصله ماکسیمم» اشاره کرد [۲۴]. در تقسیم یا ادغام خوشه منطق پیوند آن‌ها با یکدیگر نیز مهم است که روش‌های خاص خود را دارد. حاصل یک خوشه‌بندی سلسله‌مراتبی به‌صورت نگاره‌ای که دندروگرام یا شاخه‌نگار نامیده می‌شود، نشان داده می‌شود.

استفاده از خوشه‌بندی به‌عنوان رویکردی در تحلیل بازارهای مالی، کاربردهای روزافزونی یافته است. در این رویکرد هر یک از سهم‌های موجود در بازارهای مالی به‌عنوان اشیا یا مشاهداتی در نظر گرفته می‌شوند که باید ابتدا خوشه‌بندی شوند. خوشه‌بندی سهام، ممکن است بر اساس بازده آن‌ها، عوامل بنیادین متناظر و یا سایر مختصات مالی مهم صورت پذیرد. در

بخش قابل توجهی از مطالعات، ابتدا شبکه‌ای از سهام تدوین می‌شود. منظور از شبکه، نموداری است که رئوس آن، سهم‌های منتخب و یال‌های آن، فاصله یا شباهت سهام از یکدیگر است. تحلیل این شبکه یا نمودار به روش‌های گوناگونی امکان‌پذیر است که خوشه‌بندی سلسله‌مراتبی یک از آن‌ها است. ویژگی‌های یک نمودار، مانند معیارهای مرکزیت رئوس یا پیروی درجه رئوس از قانون توانی، استخراج و تحلیل درخت کمینه پوشا و سایر تحلیل‌های کمی تحلیل شبکه در مطالعات متعدد مالی صورت گرفته است [۱۸]. در پژوهش بریدا و همکاران (۲۰۱۰) سهام موجود در شاخص DAX30 با همین منطق بررسی و تحلیل شده است [۴]. این شبکه، البته در طول زمان، پویایی‌های خود را دارد که تغییرپذیری تاریخی آن‌ها موضوع برخی از پژوهش‌ها بوده است. دینامیسم تغییرات سیستم سهام، اطلاعات مفیدی برای پیش‌بینی آن‌ها در زمان‌های بعدی را به دست می‌دهد [۴۲]. تغییرات سیستم سهام در مواقع بحران نظیر بحران مالی ۲۰۰۷-۲۰۰۸ نیز در ذیل همین عنوان جای می‌گیرد [۴۱]. برای مثال در پژوهش متأخر آراجو و همکاران (۲۰۱۸) که در بورس برزیل صورت گرفته است، ابتدا شاخصی کمی برای حال و هوای فعالان بازار تکوین شده و سپس تأثیرات نمودار سهام بورس یادشده از این شاخص کمی مطالعه شده است. شاخصی که حال و هوای فعالان بازار را نشان بدهد از تحلیل متن‌ها و بسامد واژگان با بار معنایی مثبت و منفی استخراج می‌شود که روش‌های تحلیل متن را به استخدام می‌گیرد [۱]. در مطالعه‌ای دیگر که در بورس کازابلانکا، صورت گرفته است، سهام مختلف بر اساس «نمای هرست» خوشه‌بندی شده‌اند. نمای هرست که از سری زمانی بازده سهام محاسبه می‌شود، نشان‌دهنده تصادفی بودن قیمت یا رونددار بودن آن و به‌طور کلی حافظه بلندمدت سری‌های زمانی است. موضعی که می‌تواند به فرضیه بازار کارا یا شکل بدیل آن، فرضیه بازار فراکتال، مربوط باشد. این مطالعه به چند روش، نمای هرست سهام منتخب محاسبه می‌شود و خوشه‌بندی بر اساس آن‌ها صورت می‌پذیرد [۲۲]. در برخی از مطالعات، چند بازار مالی به عنوان قلمرو پژوهش انتخاب شده‌اند. برای مثال در مطالعه ی کای و همکاران (۲۰۱۷)، ۵۲ بازار مالی مختلف از کشورهای پراکنده انتخاب و خوشه‌بندی شده‌اند. در این پژوهش حاصل خوشه‌بندی‌ها و موقعیت جغرافیایی این ۵۲ شاخص بررسی و تحلیل شده است [۵].

در پژوهش اسمالیفلک و همکاران (۲۰۱۵) ۴۱ شاخص از ۴۰ کشور برای مدت ۱۲ فصل، انتخاب و میزان یکپارچگی یا عدم یکپارچگی آن‌ها به عنوان معیاری از روند جهانی شدن تحلیل شده است. در این پژوهش نیز شاخه‌نگار این شاخص‌ها یکی از ابزارهای منتخب بوده است [۱۰]. از خوشه‌بندی سلسله‌مراتبی به عنوان روشی در تکوین پرتفوی سهام نیز استفاده می‌شود. پس از تکوین خوشه‌ها، می‌توان پرتفوی‌های مرکزی و پرتفوی‌های پیرامونی را از یکدیگر تمیز داد. آنگاه با استفاده از روش‌های بهینه‌سازی سبد سهام در هر یک از خوشه‌ها، وزن بهینه هر یک از

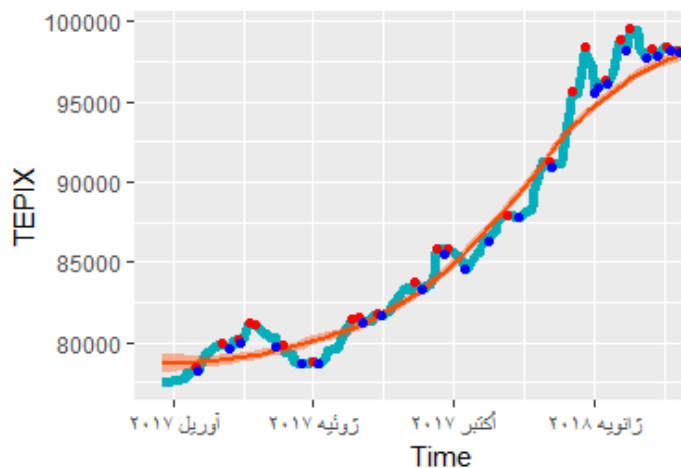
سهم‌ها در سبد منتخب را تعیین کرد. مطالعه‌ای در «بورس چین» با همین روش صورت گرفته است [۲۷]. علاوه بر محاسبات عددی، شیوه نمایش خوشه‌بندی‌ها و نمودار نیز اهمیت دارد. پژوهش‌های مستقل، اما ناکافی‌ای که نگاره‌سازی داده‌های مالی را دنبال کرده‌اند، روش‌های متعددی برای تدوین نگاره‌ها، فهم و تفسیر واقعیات مالی پیشنهاد و نقاط ضعف و قوت آن‌ها را بررسی کرده‌اند [۳۱]. رویکردهای مبتنی بر شبکه و ارتباط متقابل شاخص‌های مختلف در اندازه‌گیری و تحلیل ریسک سیستمیک نیز پیگیری شده است. در پژوهش لی و همکاران (۲۰۱۸) ویژگی‌های توپولوژیک شبکه‌های مالی به‌عنوان شاخصی از ریسک سیستمیک مطرح و در «بورس اروپایی استاکس» محاسبه شده است [۲۵].

در پژوهش حاضر نیز پس از انجام خوشه‌بندی سلسله‌مراتبی در زمان‌های مختلف، از مقایسه این خوشه‌بندی‌ها، تغییرات سیستمیک بورس تهران محاسبه شده است. پس از کمی‌کردن تغییرات سیستمی که در قالب ضریب همبستگی شاخه‌نگارها بیان شده است، استفاده از روش تجزیه مؤلفه‌های اصلی به‌عنوان روشی برای پیراستن بازده‌ها و رسیدن به داده‌هایی که حاصل بازسازی مؤلفه‌های اصلی است، پیگیری شده است. این روش از تحلیل در پژوهش‌های پیشین مالی کاربردهای فراوانی داشته است. برای مثال، در پژوهش نوبی و لی (۲۰۱۶) از PCA برای پویایی‌ها و تعاملات بین شاخص‌های بورس‌های گوناگون در فاصله سال‌ها ۱۹۹۸ تا ۲۰۱۲ استفاده شده است [۳۰]. از این روش اخیراً در تحلیل ماتریس واریانس - کواریانس حاصل از داده‌های با بسامد بالا نیز استفاده شده است [۲۰]. پس از بحران مالی ۲۰۱۷-۲۰۱۸ در چند پژوهش دیگر از این روش برای اندازه‌گیری ریسک سیستمیک در نهاد مالی و بانک‌های چین استفاده شده است [۱۲، ۱۷، ۲۳]. در بخش بعدی به روش پژوهش، شامل قلمروهای پژوهش، خوشه‌بندی سلسله‌مراتبی، روش مقایسه شاخه‌نگارها و نیز روش مؤلفه‌های اصلی پرداخته می‌شود.

۳. روش‌شناسی

این پژوهش در پارادایم پژوهش‌های کمی، شکل گرفته است. در این‌گونه پژوهش‌ها داده‌های کمی مبنای تصمیم‌گیری خواهند بود. با توجه به اینکه نوع مسئله مقتضی فرضیه خاصی نیست، پژوهش حاضر مبتنی بر پرسشی در باب اندازه‌گیری میزان ثبات سیستم مالی و کارآمدی روش ارائه شده است؛ بنابراین، هدف از این مطالعه، اندازه‌گیری رابطه دو کمیت یا متغیر است که ماهیتی تبیینی و فرضیه‌پذیر داشته باشد و ارائه فرضیه، بلامحل است؛ به‌بیان‌دیگر این پژوهش ماهیتی اکتشافی دارد و از نظر هدف، توصیفی است؛ زیرا توصیف و اندازه‌گیری میزان ثبات سیستم، مدنظر بوده است. بدیهی است، سایر پژوهشگران می‌توانند از روش ارائه‌شده

برای سنجش ثبات مالی و ریسک سیستمیک در پژوهش‌هایی که ماهیت تبیینی و فرضیه‌پذیر دارند، استفاده کنند. داده‌های روزانه ۲۶ شاخص «بورس اوراق بهادار تهران» در ۲۴۱ روز کاری از «سایت بورس اوراق بهادار تهران» از تاریخ ۱۳۹۶/۰۱/۰۵ تا ۱۳۹۷/۰۱/۰۵ دریافت شد. این داده‌ها متناظر با یک سال شمسی است که به عنوان نمونه انتخاب شده‌اند تا بر اساس آن‌ها، شاخه‌نگارها و مقایسه آن‌ها صورت پذیرد. در صورتی که نتایج این مقطع زمانی با مقاطع دیگر مقایسه شود، تفسیرپذیر بهتری داشته و تصویر روشن‌تری از ثبات سیستم مالی را به دست خواهد داد؛ موضوعی که می‌تواند دست‌مایه پژوهش‌های آتی قرار بگیرد. اگر مقدار شاخص i در روز t و $P_i(t-1)$ مقدار آن در پایان روز $t-1$ را نشان دهد، آنگاه بازده لگاریتمی یا پیوسته شاخص در فاصله زمانی Δt با استفاده از رابطه ۱، به دست می‌آید.



نمودار ۲. رفتار شاخص کل در قلمرو پژوهش

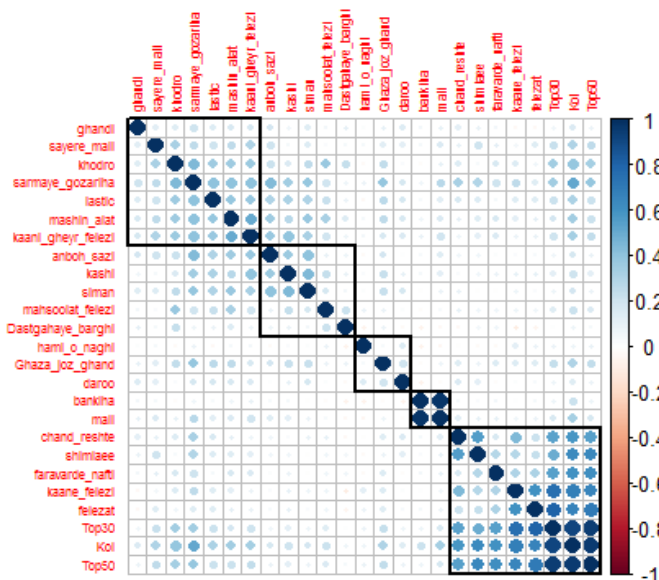
$$R_i(t) = \ln \frac{P_i(t)}{P_i(t-1)} \quad \text{رابطه (۱)}$$

در بازه زمانی انتخاب شده شاخص کل بورس اوراق بهادار به صورت نمودار ۲، تغییر داشته است. در این نمودار مشخص است که شاخص کل از عدد ۷۷۴۸۰ آغاز و در انتهای قلمرو زمانی پژوهش به ۹۶۴۲۰ می‌رسد. علاوه بر رفتار شاخص، روند کلی و نقاط قله و دره آن در نمودار مشخص است. در همین بازه، ماتریس همبستگی ۲۶ شاخص نیز محاسبه شده است که در نمودار ۳، مشاهده می‌شود. عناصر قطر اصلی ماتریس همبستگی که اساس ماتریس فاصله شاخص‌ها از یکدیگر است، مقدار یک را به خود می‌گیرند (چرا که میزان همبستگی هر متغیر

با خودش، عدد یک است) و سایر عناصر میزان ضریب همبستگی آن‌ها را نشان می‌دهد. بخش عمده‌ای از شبکه‌های مالی بر اساس همبستگی بازده‌ها به‌دست می‌آید [۴۰]. محاسبه ماتریس همبستگی نخستین گام در تدوین شبکه سهام و به‌طور کلی خوشه‌بندی‌های مبتنی بر همبستگی است. در این پژوهش از همبستگی پیرسون برای به‌دست‌آوردن ماتریس همبستگی استفاده شده است.

$$\rho_{ij} = \frac{\langle Y_i Y_j \rangle - \langle Y_i \rangle \langle Y_j \rangle}{\sqrt{(\langle Y_i^2 \rangle - \langle Y_i \rangle^2)(\langle Y_j^2 \rangle - \langle Y_j \rangle^2)}} \quad \text{رابطه (۲)}$$

در رابطه ۲، i و j نمادهای عددی از سهم‌ها هستند و (...) نشان‌دهنده متوسط آماری در طول دوره مورد مطالعه است. ضریب همبستگی ρ_{ij} می‌تواند در بازه -1 تا $+1$ قرار بگیرد که در آن -1 به معنای همبستگی کاملاً مخالف دو قیمت سهام، است و $+1$ یعنی دو قیمت سهام کاملاً همبسته هستند. زمانی که $\rho_{ij} = 0$ باشد، آنگاه قیمت‌های دو سهم غیرهمبسته است. ماتریس ضریب همبستگی، یک ماتریس متقارن $N \times N$ با $\rho_{ii} = 1$ در قطر اصلی است [۱۹، ۳۷].

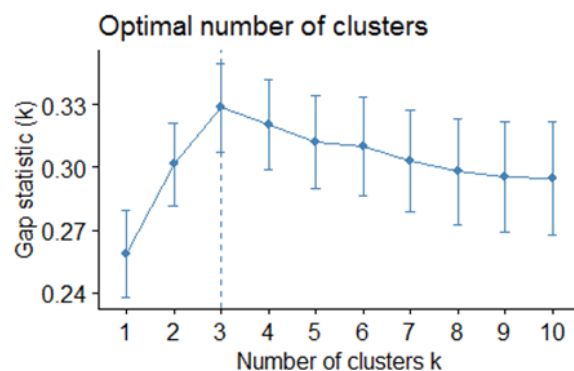


نمودار ۳. همبستگی شاخص‌ها

پس از محاسبه همبستگی، با استفاده از رابطه ۳، همبستگی به متریک فاصله تبدیل می‌شود.

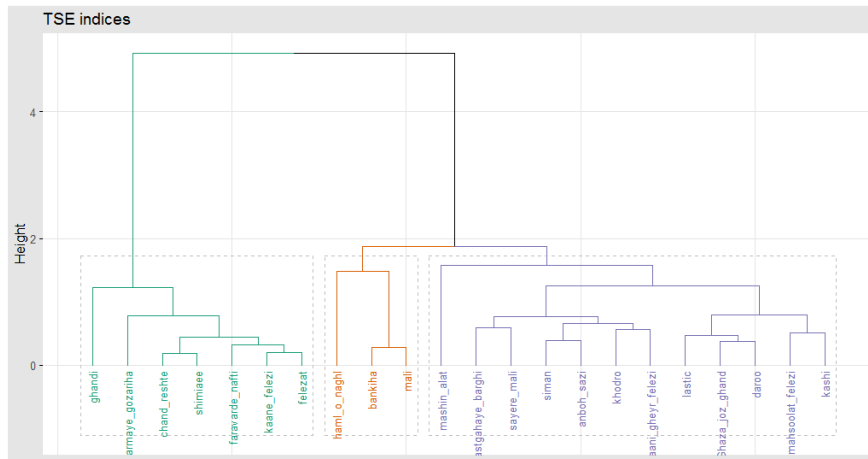
$$d(i, j) = \sqrt{2(1 - \rho_{ij})} \quad \text{رابطه (۳)}$$

بر اساس این متریک، بیشترین فاصله بین دو سهم، زمانی رخ می‌دهد که همبستگی بازده آن‌ها، ۱- کمترین فاصله، در همبستگی کامل +۱ اتفاق می‌افتد. استفاده از این متریک به عنوان مبنایی برای خوشه‌بندی به دلیل خواص ریاضی، همچون متقارن بودن و نامساوی مثلثی، توجیه‌پذیر است [۸، ۹]. در گام بعد ماتریس فاصله‌ها، به عنوان ورودی روشی که وارد در سال ۱۹۶۳ ارائه کرد، استفاده می‌شود. بر اساس «روش وارد»، تابع هدف، کمینه‌سازی مجموع واریانس‌های درون خوشه‌ای است. در خوشه‌بندی هر یک از خوشه‌ها واریانس درون خوشه‌ای دارند که در حالت ایدئال صفر است. واریانس درون خوشه‌ای کمتر، نشان‌دهنده همگنی بیشتر مشاهداتی است که در یک خوشه جای گرفته‌اند. در این روش در هر مرحله خوشه‌هایی با یکدیگر ادغام می‌شوند که ترکیب آن‌ها کمترین میزان افزایش در مجموع واریانس‌های درون خوشه‌ای را به دنبال داشته باشد؛ به عبارت دیگر همگن‌ترین خوشه‌ها تا حد ممکن با یکدیگر ادغام می‌شوند و خوشه‌های بزرگ‌تر و در پایان، سلسله‌مراتبی از خوشه‌ها شکل می‌گیرد؛ بنابراین روش وارد از جمله الگوریتم‌های پایین به بالا در خوشه‌بندی سلسله‌مراتبی است [۷]. در گام بعد خوشه‌بندی حاصل از روش وارد به صورت شاخه‌نگار نمایش داده می‌شود. پیش از نمایش خوشه‌ها می‌توان تعداد خوشه‌ها را تعیین کرد. تعیین تعداد بهینه خوشه‌ها به روش‌های گوناگونی از جمله «روش البو»، «روش میانگین سیلووت» و «روش آماره شکاف» صورت می‌گیرد. در این پژوهش از «روش آماره شکاف» که در سال ۲۰۰۱ معرفی شد، استفاده شده است [۳۶]. بر اساس داده‌های قلمرو این پژوهش، مقدار آماره شکاف محاسبه شده در ۳ خوشه، بهینه است.



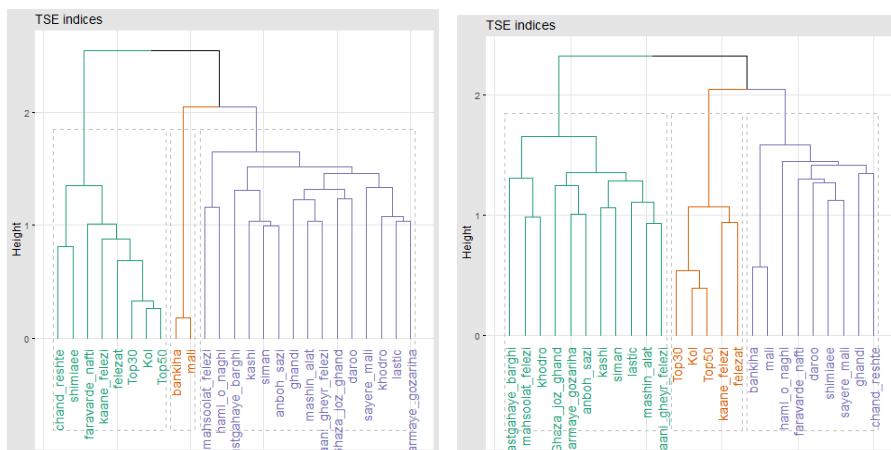
نمودار ۴. مقدار بهینه تعداد خوشه‌ها بر اساس آماره شکاف

اصلی‌ترین ابزار نگاره‌سازی در خوشه‌بندی‌های سلسله‌مراتبی، شاخه‌نگار است. در ارائه شاخه‌نگار به تعداد، خوشه‌ها و سلسله‌مراتب آن‌ها و نیز موقعیت سهم‌ها از یکدیگر توجه شده است.



نمودار ۵. شاخه‌نگار شاخص‌ها پیش از انجام PCA

بر اساس این خوشه‌بندی، شاخص‌های قندی، سرمایه‌گذاری، چندرشته‌ای، شیمیایی فراورده‌های نفتی، کانه فلز و فلزات در خوشه اول، بانک‌ها و مالی و حمل‌ونقل در خوشه دوم و سایر شاخص‌ها در خوشه سوم جای گرفته‌اند. موقعیت متریک هر یک از شاخص‌ها نسبت به یکدیگر در نمودار ۵، نشان داده شده است. طبیعی است که با تغییر قلمرو زمانی، خوشه‌بندی‌های متفاوتی بروز پیدا می‌کند. برای مثال، اگر قلمرو زمانی این پژوهش به دو دوره شش‌ماهه تقسیم شده و در هر بازه پس از خوشه‌بندی، شاخه‌نگار مربوطه تدوین شود، می‌توان میزان تغییرات این دو دوره را تعیین کرد. در نمودار ۶ مشخص است که ترکیب خوشه‌ها در دو بازه موردنظر تا حدودی متفاوت بوده و موقعیت شاخص‌ها، نسبت به یکدیگر، شاهد تغییراتی بوده است.



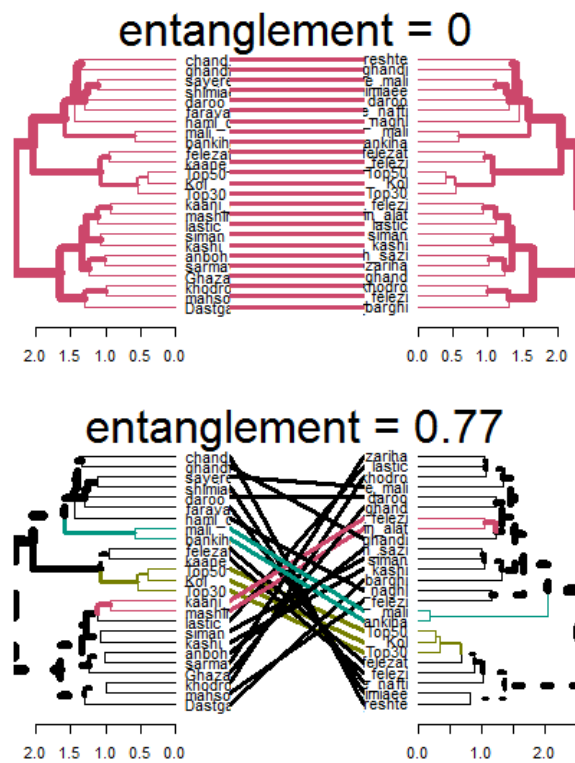
نمودار ۶. شاخه‌نگارهای شش‌ماهه

تغییرات شاخه‌نگارها در طول زمان می‌تواند نشان‌دهنده تغییرات سیستم باشد. مسئله اصلی این پژوهش، مقایسه شاخه‌نگارها به دو روش است: ۱. تدوین هم‌شاخه‌نگارها و ۲. محاسبه شاخص‌های کمی برای مقایسه شاخه‌نگارها. نتیجه چنین مقایسه‌ای، رسیدن به پاسخ این پرسش است که آیا شاخه‌نگارها که نشان‌دهنده موقعیت متریک شاخص‌های گوناگون «بورس اوراق بهادار تهران» هستند، در طول زمان تغییر کرده‌اند و بنابراین دلالت بر گونه‌ای ناپایداری در سیستم را دارند؟ یا برعکس شاخه‌نگارهای یادشده در طول زمان، پایدار بوده‌اند و تغییرات زیادی نداشته‌اند. برای پاسخ به این پرسش‌ها در ادامه کار، ابتدا مفهوم هم‌شاخه‌نگار، شاخص گامای بیکر و ضریب همبستگی همسانی توضیح داده می‌شود. در قسمت بعد، نتایج محاسبه آن در قلمرو زمانی این پژوهش تشریح خواهد شد.

هم‌شاخه‌نگارها که میزان شباهت دو یا چند شاخه‌نگار را نشان می‌دهند، یکی از روش‌های مقایسه خوشه‌بندی‌های سلسله‌مراتبی به‌شمار می‌روند. در برخی از پژوهش‌ها، هم‌شاخه‌نگارهایی که دو یا بیش از دو خوشه‌بندی را مقایسه می‌کند، معرفی و روش تدوین آن‌ها ارائه شده است [۳۸]. بیشتر کاربردهای این روش در آمار زیستی و مسائلی از قبیل شباهت و خوشه‌بندی مشاهدات زیست‌شناختی بوده است [۲۸، ۳۴].

به نظر می‌رسد تعمیم و به‌کارگیری آن در مطالعات مالی، فارغ از نوآوری این پژوهش، می‌تواند در آشکارسازی تغییرات سیستم مالی سهم داشته باشد. هم‌شاخه‌نگار، دو شاخه‌نگار را به‌عنوان ورودی می‌گیرد و آن‌ها را مقابل یکدیگر قرار می‌دهد و مشاهدات مشابه را پیدا و با خط به یکدیگر متصل می‌کند. در صورتی که شاخه‌نگارهای ورودی عین هم باشند، هیچ‌گونه درهم‌تنیدگی اتفاق نمی‌افتد و خطوطی که مشاهدات مشابه را به یکدیگر وصل می‌کنند موازی هستند؛ اما اگر شاخه‌نگارها، متفاوت باشند، برخی از این خطوط یکدیگر را قطع می‌کنند. هر چه میزان درهم‌تنیدگی خطوط بیشتر باشد، شاخه‌نگارها تفاوت بیشتری دارند؛ بر همین اساس میزان درهم‌تنیدگی محاسبه می‌شود که مقدار صفر یعنی شباهت کامل شاخه‌نگارهای ورودی (بدون هیچ خط متقاطع) و مقدار یک یعنی عدم‌شباهت کامل آن‌ها (مقاطع بودن همه خطوط) است [۱۴].

در نمودار ۷، هم‌شاخه‌نگار خوشه‌بندی شش‌ماهه نخست با خودش (به‌عنوان مثالی از شباهت کامل با درهم‌تنیدگی صفر) و با شش‌ماهه دوم (به‌عنوان مثالی از شباهت جزئی با درهم‌تنیدگی ۷۷ درصد) نشان داده شده است.



نمودار ۷. هم‌شاخه‌نگار شش‌ماهه اول و دوم

بیکر (۱۹۷۴)، روشی برای مقایسه خوشه‌بندی‌های سلسله‌مراتبی ارائه کرد. بر اساس روش وی، دو مشاهده (در پژوهش حاضر دو شاخص) انتخاب می‌شود، سپس تعیین می‌کنند در خوشه‌بندی اول در چه صورت آن دو شاخص هم‌خوشه می‌شوند و تعداد خوشه‌ها در این حالت چند است؟ حالا این دو شاخص در خوشه‌بندی دوم کنترل شده و مجدداً تعداد خوشه‌ها شمارش می‌شود. برای این زوج از شاخص‌ها دو عدد (تعداد خوشه‌ها در حالت هم‌خوشه‌گی اول و تعداد خوشه‌ها در حالت هم‌خوشه‌گی دوم) به دست می‌آید. برای همه زوج‌ها (همه انتخاب‌های دوتایی از شاخص‌ها) این دو عدد محاسبه می‌شود؛ سپس همبستگی رتبه‌ای (اسپیرمن) این دو ستون عدد محاسبه می‌شود که عددی بین -1 و $+1$ است و مقادیر نزدیک به صفر، عدم شباهت آماری خوشه‌بندی‌های موردنظر را نشان می‌دهد. [۲].

روش دیگر مقایسه خوشه‌بندی‌های سلسله‌مراتبی، استفاده از مفهوم ناهمسانی مشاهداتی است که خوشه‌بندی شده‌اند. این فاصله، ناهمسانی خوشه‌های مختلف از یکدیگر را اندازه‌گیری می‌کند. در خوشه‌بندی، هدف اولیه افزایش ناهمسانی بین خوشه‌ای و کاهش ناهمسانی

درون خوشه‌ای است. همبستگی ناهمسانی‌ها عبارت است از: ضریب همبستگی دو ماتریس که هر یک فاصله‌های ناهمسانی دو خوشه‌بندی را نشان می‌دهند. تفسیر همبستگی ناهمسانی‌ها نیز مشابه «شاخص بیکر» است [۳۵]. «شاخص فاولکس - مالوز» نیز شباهت دو خوشه‌بندی را نشان می‌دهد. این شاخص، عددی بین صفر تا ۱ را می‌گیرد و هر چه شباهت خوشه‌بندی‌ها بیشتر باشد، به ۱ نزدیک‌تر می‌شود [۱۳].

در ادامه، شاخص‌های انتخاب‌شده در قلمرو زمانی پژوهش به صورت ماهانه، خوشه‌بندی شده و سپس این خوشه‌بندی‌ها با یکدیگر مقایسه می‌شوند. با توجه به اینکه رسم 12×12 هم‌شاخه‌نگار ممکن نیست؛ بنابراین در مقایسه آن‌ها به «شاخص بیکر» و ضریب همبستگی ناهمسانی‌ها اکتفا شده است. در این پژوهش از بسته *dendextend* از نرم‌افزار *R* استفاده شده است [۱۴]. پس از مشخص شدن میزان ناپایداری خوشه‌بندی‌ها، با استفاده از روش تجزیه مؤلفه‌های اصلی (PCA)، سری‌های زمانی بازده شاخص‌های منتخب به صورت ترکیبی از عناصر پایدار سیستم تحلیل می‌شود و خوشه‌بندی نهایی بر اساس مؤلفه‌های اصلی سیستم صورت می‌گیرد. در انجام تجزیه مؤلفه‌های اصلی به صورت کلی مراحل زیر طی شده است:

الف) ابتدا بازده لگاریتمی شاخص‌های قلمرو مطالعه، محاسبه و به صورت ماتریس $T \times n$ که در آن n تعداد شاخص‌ها و T تعداد بازده‌ها یا عمق سری زمانی است، نشان داده می‌شود (ماتریس X).

ب) ماتریس V یا ماتریس همبستگی بازده‌ها یک ماتریس $n \times n$ است که از آن با محاسبه بردارها و مقدارهای ویژه آن می‌توان به ماتریسی $n \times n$ رسید که W یا «ماتریس بارعاملی نامیده» می‌شود.

ج) با استفاده از تبدیل $P = XW$ ، ماتریس بازده‌ها به ماتریس مؤلفه‌های اصلی تبدیل می‌شود که یک ماتریس $T \times n$ است.

د) در آخرین گام k مؤلفه اصلی که بیشترین بار عاملی را دارند و ماتریس بازده‌ها بر اساس آن‌ها، به این شکل تبدیل می‌شوند: $X \approx P^*W^*$. ماتریس X ، همان ماتریس بازده‌هایی است که مؤلفه‌های اصلی آن‌ها را ساخته‌اند و بنابراین مؤلفه‌های غیراصولی که نوفه یا تصادفی هستند از آن حذف شده‌اند [۱۱]. با اجرای این روش و رسیدن به این داده‌های پیراسته و انجام خوشه‌بندی بر اساس آن‌ها می‌توان به خوشه‌بندی سلسله‌مراتبی‌ای دست یافت که بر اساس تحلیل مؤلفه‌های اصلی شکل گرفته است و بنابراین نشان‌دهنده موقعیت متریک ثبت‌شده آن‌ها نسبت به یکدیگر است. جدول ۱، مراحل انجام کار و توابع استفاده‌شده را نشان می‌دهد.

۴. تحلیل داده‌ها

پس از تقسیم بازه زمانی به صورت ماهانه، شاخه‌نگار مربوط به هر ماه تدوین شد؛ سپس شاخص بیکر، شاخص همبستگی همسانی‌ها و شاخص *FIM*، به صورت زوجی محاسبه و نتایج حاصله نگاره‌سازی شد.

جدول ۱. توابع و بسته‌های استفاده‌شده

مرحله	شرح	تابع استفاده‌شده	بسته استفاده‌شده
۱	ورود داده‌ها شاخص‌ها به R	read.csv()	base
۲	تبدیل داده‌ها به سری زمانی	timeDate() as.timeSeries()	timeSeries
۳	محاسبه بازده لگاریتمی شاخص‌ها	returns	quantmod
۴	خوشه‌بندی سلسله‌مراتبی شاخص‌ها و تدوین شاخه‌نگار یک‌ساله	hclust() as.dendrogram()	stats
۵	تدوین هم‌شاخه‌نگار بازه‌های شش‌ماهه	tanglegram()	dendextend
۶	نگاره‌سازی شاخه‌نگارها	fviz_dend()	factextra
۷	محاسبه شاخص بیکر برای خوشه‌بندی‌های ماهانه	cor_bakers_gamma()	dendextend
۸	محاسبه شاخص همبستگی ناهمسانی‌ها	cor_cophenetic()	dendextend
۹	نگاره‌سازی نتایج	ggplot()	ggplot2
۱۰	تجزیه مؤلفه‌های اصلی	princom()	stats

جدول ۲، مقادیر «شاخص بیکر» را نشان می‌دهد. قطر اصلی ماتریس نشان‌دهنده شباهت هر خوشه‌بندی به خودش، یعنی یک است؛ اما همان‌طور که مشاهده می‌شود، خوشه‌بندی‌ها در بیشتر موارد به یکدیگر شبیه نیستند و مقدار «شاخص بیکر» در آن‌ها به صفر نزدیک است. روند قابل ملاحظه‌ای در «شاخص بیکر» و ضریب همبستگی همسانی‌ها وجود ندارد. این واقعیت در مقایسه با حالتی فرضی که در آن همبستگی‌ها، نزدیک به یک و نمایانگر این حقیقت باشند که پایداری قابل ملاحظه‌ای در سیستم وجود دارد، قابل تفسیر است؛ از طرف دیگر ضریب همبستگی همسانی‌ها نیز در بازه موردنظر رفتاری مشابه «شاخص بیکر» دارد. جدول ۳، نشان می‌دهد که بر اساس ضریب همبستگی همسانی‌ها، شباهت بین خوشه‌بندی‌ها در ماه‌های مختلف سال ناچیز است و به ندرت در برخی از مواقع این ضریب قابل ملاحظه به نظر می‌رسد.

به منظور درک روشن تری از پایداری «بورس اوراق بهادار تهران»، ماتریس‌های فوق به ماتریس‌هایی بالامثلی تبدیل می‌شود تا مقادیر تکراری حذف شوند. روشن است که هر دو ماتریس، متقارن هستند و با در نظر گرفتن ماتریس بالامثلی یا پایین‌مثلثی، اطلاعات خاصی از بین نخواهد رفت. آمار توصیفی مقادیر باقی مانده و چگالی احتمال آن‌ها در ادامه بررسی می‌شود.

جدول ۲. مقادیر «شاخص بیکر» قبل از PCA

	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
۱	۱											
۲	-۰.۴۲	۱										
۳	-۰.۲۸	-۰.۱۳	۱									
۴	-۰.۳۶	-۰.۱۳	-۰.۲۶	۱								
۵	-۰.۵۴	-۰.۵۷	-۰.۱۳	-۰.۲۹	۱							
۶	-۰.۴۵	-۰.۲۹	-۰.۷۱	-۰.۰۳	-۰.۳۱	۱						
۷	-۰.۱۴	-۰.۰۹	-۰.۱۹	-۰.۱۷	-۰.۳۱	-۰.۰۲	۱					
۸	-۰.۵۱	-۰.۰۲	-۰.۱۶	-۰.۰۱	-۰.۳۹	-۰.۴۳	-۰.۵۱	۱				
۹	-۰.۱۱۶	-۰.۵۵	-۰.۲۴	-۰.۳۱	-۰.۰۲	-۰.۲۰۶	-۰.۲۳	-۰.۴۲	۱			
۱۰	-۰.۵۴	-۰.۰۶	-۰.۱۵	-۰.۰۶۱	-۰.۱۰۷	-۰.۱۷	-۰.۳۷	-۰.۴۶	-۰.۰۱	۱		
۱۱	-۰.۵۲	-۰.۰۱	-۰.۲۴	-۰.۵۵	-۰.۳۶	-۰.۱۸	-۰.۰۲	-۰.۵۵	-۰.۰۶	-۰.۸۸	۱	
۱۲	-۰.۲۴	-۰.۲۱	-۰.۱۹	-۰.۲۸	-۰.۴۱	-۰.۰۹	-۰.۰۹	-۰.۱۴۶	-۰.۲۹	-۰.۰۵	-۰.۲۳	۱

جدول ۳. شاخص همبستگی همسانی‌ها پیش از PCA

	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
۱	۱											
۲	-۰.۵۸	۱										
۳	-۰.۳۴	-۰.۰۱	۱									
۴	-۰.۵۶	-۰.۲۵	-۰.۱۸	۱								
۵	-۰.۷۱	-۰.۶۴	-۰.۵۸	-۰.۲۱	۱							
۶	-۰.۱۳	-۰.۰۵	-۰.۹۳	-۰.۳۱	-۰.۰۷	۱						
۷	-۰.۶۵	-۰.۱۳	-۰.۴۳	-۰.۱۳	-۰.۲۳	-۰.۱۵	۱					
۸	-۰.۶۲	-۰.۳۵	-۰.۲۶	-۰.۲۲	-۰.۰۶	-۰.۴۲	-۰.۰۷	۱				
۹	-۰.۷۸	-۰.۷۶	-۰.۱۶	-۰.۲۳	-۰.۰۶	-۰.۲۲	-۰.۰۴	-۰.۴۲	۱			
۱۰	-۰.۰۴	-۰.۰۴	-۰.۲۴	-۰.۴۶	-۰.۴۵	-۰.۱۹	-۰.۴۳	-۰.۲۱	-۰.۳۲	۱		
۱۱	-۰.۳۲	-۰.۰۱	-۰.۱۷	-۰.۵۱	-۰.۰۵	-۰.۳۲	-۰.۰۸	-۰.۰۱	-۰.۱۷	-۰.۵۶	۱	
۱۲	-۰.۴۳	-۰.۳۹	-۰.۶۵	-۰.۲۹	-۰.۰۴	-۰.۱۸	-۰.۱۱	-۰.۱۷۲	-۰.۱۶	-۰.۰۸	-۰.۴۲	۱

جدول ۴. آمار توصیفی شاخص بیکر و شاخص همبستگی همسانی

	mean	sd	median	min	max	range
شاخص بیکر	۰/۲۲۷	۰/۱۶۹	۰/۱۹۶	-۰/۰۱۰	۰/۵۷۹	۰/۵۹۰
شاخص همبستگی همسانی	۰/۲۸۷	۰/۱۳۵	۰/۲۵۰	۰/۰۵۲	۰/۶۲۳	۰/۵۷۰

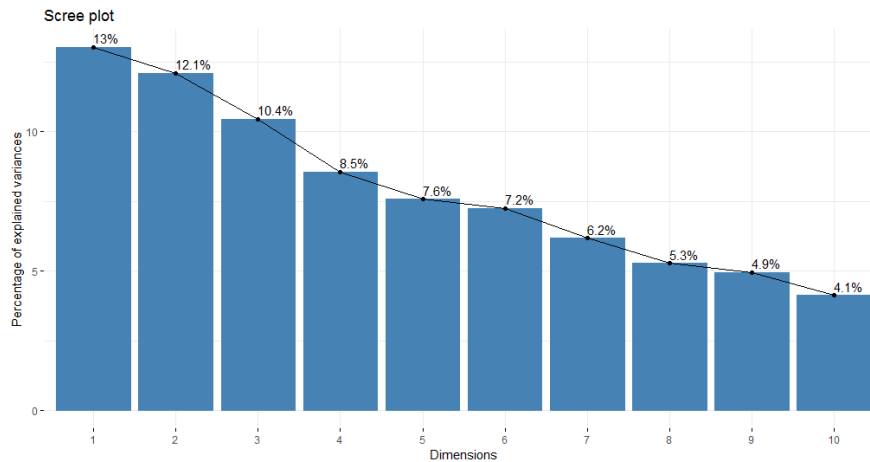
با این توصیف، میانگین شاخص بیکر و شاخص همبستگی همسانی، به ترتیب ۰/۲۲۷ و ۰/۲۸۷ است که هر دو گویای پایداری کم سیستم مورد مطالعه هستند. پایداری کم سیستم نشان‌دهنده تغییرپذیری مداوم شاخه‌نگارها و خوشه‌بندی‌های سلسله‌مراتبی است؛ بنابراین نتیجه عملی، آن است که در هر زمان موقعیت متریک شاخص‌ها نسبت به یکدیگر تغییر می‌کند؛ باوجود این روش‌هایی وجود دارند که امکان اندازه‌گیری جزء پایدار سیستم را به دست می‌دهند. یکی از این روش‌ها، تجزیه مؤلفه‌های اصلی (PCA) است. در روش PCA، ماتریس سری‌های زمانی‌ای که مبنای خوشه‌بندی در یک سال هستند، بر اساس مؤلفه‌های اصلی (حاصل از مقادیر ویژه و بردارهای ویژه ماتریس همبستگی آن‌ها) بازسازی می‌شوند. در این بازسازی بر اساس مبنای روش PCA، تأثیر عوامل اصلی یا همان مؤلفه‌های باثبات سیستم که بیشترین بار عاملی در شکل‌گیری نوسان کل سیستم را داشته‌اند در نظر گرفته می‌شود و از تأثیر عواملی که اهمیت کمتری دارند و به سبب ناپایداری این سیستم بوده‌اند، چشم‌پوشی می‌شود. در گام نخست، مقادیر ویژه ماتریس همبستگی بازده‌ها محاسبه می‌شود. برای انتخاب تعداد ابعاد به مقادیر ویژه ی این ماتریس، توجه می‌شود. ملاک دیگر درصد تجمعی واریانس است. در پژوهش‌های مالی معمولاً تعداد ابعادی که ۷۰ درصد تغییرات را نشان دهد، به عنوان ملاک انتخاب تعداد ابعاد در نظر گرفته می‌شود. این تعداد در پژوهش حاضر، ۸ بُعد است.

جدول ۵. مقادیر ویژه

	مقادیر ویژه	درصد واریانس	درصد تجمعی واریانس
Dim.1	۳۱/۲۲۸۹۲۷۱۷	۱۳/۰۱۲۰۵۲۹۹	۱۳/۰۱۲۰۵۲۹۹
Dim.2	۲۹/۰۱۸۲۰۵۷۶	۱۲/۰۹۰۹۱۹۰۷	۲۵/۰۲۹۷۲۰۵
Dim.3	۲۵/۰۷۶۲۸۰۸۵	۱۰/۴۴۸۴۵۰۳۶	۳۵/۵۵۱۴۲۲۴۱
Dim.4	۲۰/۵۱۶۰۴۷۱۹	۸/۵۴۸۳۵۲۹۹۷	۴۴/۰۹۹۷۷۵۴
Dim.5	۱۸/۱۵۲۸۱۷۹۱	۷/۵۶۳۶۹۹۶۲۴	۵۱/۶۶۳۴۷۵۰۳
Dim.6	۱۷/۳۶۲۲۲۰۴۱	۷/۲۳۴۳۴۱۸۳۶	۵۸/۸۹۷۸۱۶۸۶
Dim.7	۱۴/۸۱۱۲۷۷۲۲	۶/۱۷۱۳۶۵۵۰۷	۶۵/۰۶۹۱۸۳۳۷
Dim.8	۱۲/۶۴۱۴۸۸۸۸	۵/۲۶۷۲۸۷۰۳۳	۷۰/۳۳۶۴۶۹۴
Dim.9	۱۱/۸۵۱۵۰۷۹۹	۴/۹۳۸۱۲۸۳۳۱	۷۵/۲۷۴۵۹۷۷۴
Dim.10	۹/۸۸۷۶۶۸۶۴۱	۴/۱۱۹۸۶۱۹۳۴	۷۹/۳۹۴۴۵۹۶۷

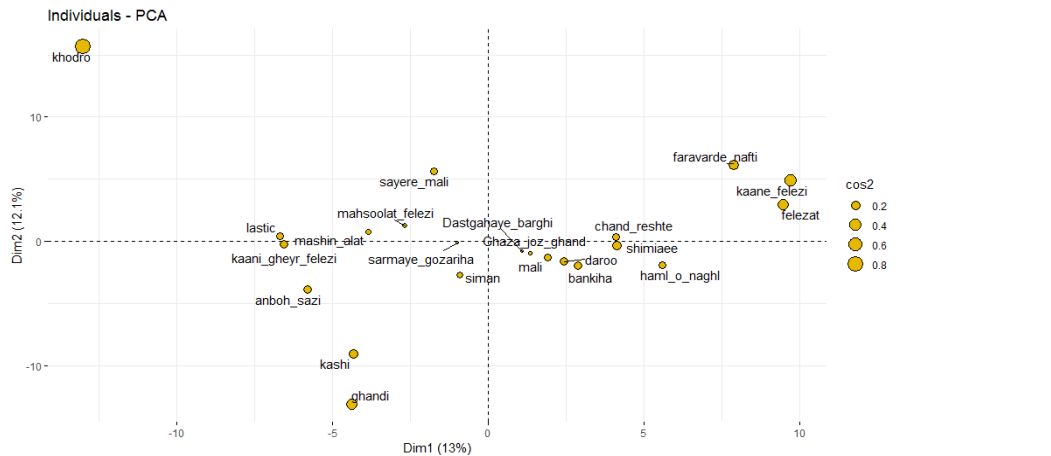
Dim.11	۹/۲۲۷۴۴۲۱۵۹	۳/۸۴۸۹۳۴۲۲۳	۸۳/۲۴۳۳۹۳۹
Dim.12	۷/۸۲۷۶۱۶۹۴۲	۳/۲۶۱۵۰۷۰۵۹	۸۶/۵۰۴۹۰۰۹۶
Dim.13	۷/۲۱۹۴۹۲۸۶۴	۳/۰۰۸۱۲۲۰۲۷	۸۹/۵۱۳۰۲۲۹۹
Dim.14	۵/۵۰۵۱۰۹۰۳۹	۲/۲۹۳۷۹۵۴۳۳	۹۱/۸۰۶۸۱۸۴۲
Dim.15	۵/۲۹۰۹۰۵۶۸۸	۲/۲۰۴۵۴۴۰۳۷	۹۴/۰۱۱۳۶۲۴۶
Dim.16	۴/۳۱۹۵۰۸۳۴۶	۱/۷۹۹۷۹۵۱۴۴	۹۵/۸۱۱۱۵۷۶
Dim.17	۳/۵۹۳۴۹۸۱۷	۱/۴۹۷۲۹۰۹۰۴	۹۷/۳۰۸۴۴۸۵۱
Dim.18	۳/۰۲۹۴۴۶۳۳	۱/۲۶۲۲۶۹۳۰۴	۹۸/۵۷۰۷۱۷۸۱
Dim.19	۱/۹۹۲۴۴۶۹۳۴	-۰/۸۳۰۱۸۶۲۲۳	۹۹/۴۰۰۹۰۴۰۳
Dim.20	۱/۴۲۵۶۶۱۸۳۵	-۰/۵۹۴۰۲۵۷۶۵	۹۹/۹۹۴۹۲۹۸
Dim.21	-۰/۰۱۲۱۶۸۴۸۶	-۰/۰۰۵۰۷۰۲۰۲	۱۰۰

مقادیر ویژه و نقش آن‌ها در تشکیل مؤلفه‌های اصلی این پژوهش به صورت نمودار ۸، است. همان‌گونه که مشاهده می‌شود این نمودار در تناظر با مقادیر جدول ۵، است.



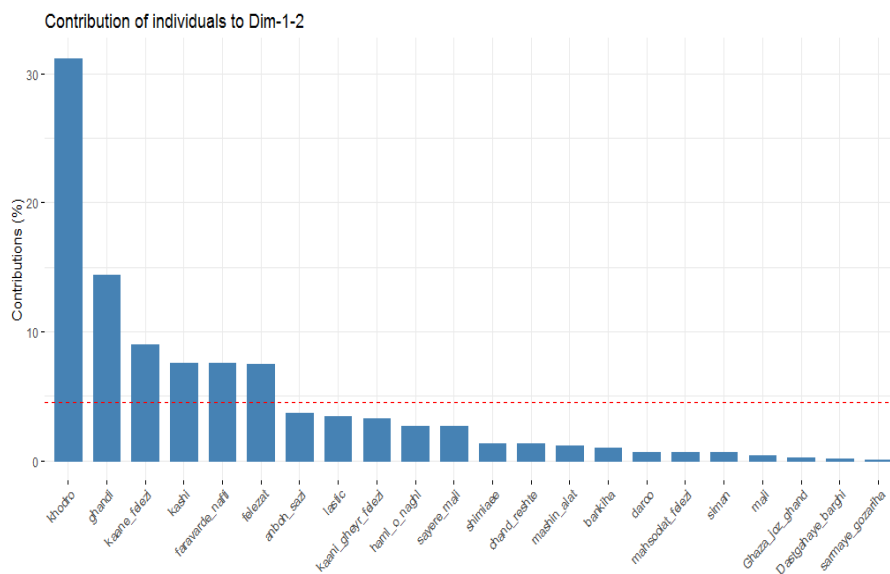
نمودار ۸. میزان اهمیت مؤلفه‌های ویژه

پس از آنکه تعداد مؤلفه‌های اصلی مشخص شد، می‌توان ماتریس بازده‌ها را به ماتریس بازده‌های تعدیل‌شده بر مبنای PCA تبدیل کرد و خوشه‌بندی را بر اساس آن انجام داد. در نمودار ۹، موقعیت هر یک از شاخص‌ها نسبت به دو عامل نخست که دارای بیشترین اهمیت هستند و زاویه‌ای که با مؤلفه دوم دارند، مشخص شده است.



نمودار ۹. موقعیت هر یک از شاخص‌ها نسبت به مؤلفه‌های اول و دوم

نمودار ۱۰، نقشی که هر یک از صنایع در دو مؤلفه اصلی (یکم و دوم) داشته‌اند را نشان می‌دهد. بر اساس نتایج، صنایعی که بیشترین نقش را در مؤلفه‌های اول و دوم داشته‌اند و به تعبیری بیشترین تأثیرگذاری سیستمی را به خود اختصاص داده‌اند عبارت‌اند از: خودرو؛ صنایع قندی؛ کانی فلزی؛ کاشی؛ فرآورده‌ها نفتی؛ فلزات و انبوه‌سازی.

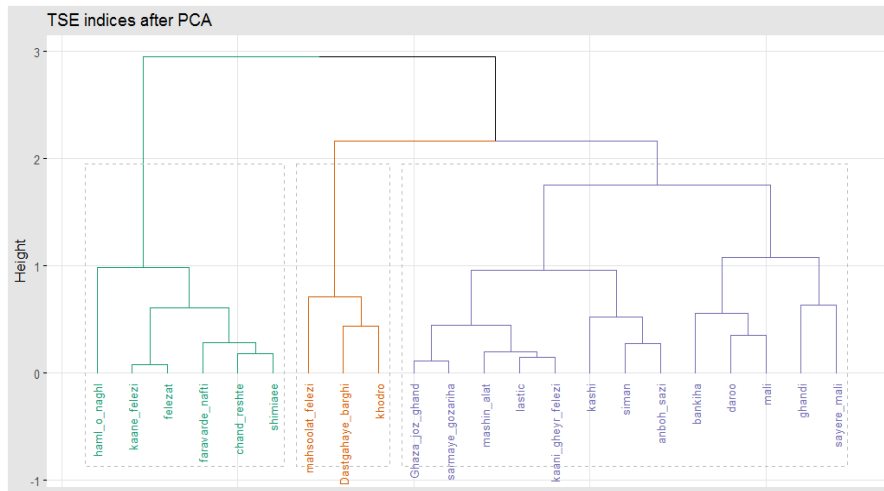


نمودار ۱۰. سهم هر یک از شاخص‌ها در مؤلفه‌های اول و دوم

پس از تولید سری‌های زمانی جدید بر مبنای PCA، برای هر ماه به صورت جداگانه، فرایند خوشه‌بندی سلسله‌مراتبی، صورت پذیرفت و «شاخص بیکر» برای آن‌ها مجدداً محاسبه شده و در جدول ۶، گزارش شده است. با توجه به جدول ۶، همبستگی تعدادی از خوشه‌بندی‌ها در ماه‌های مختلف به مقادیر نزدیکی به عدد یک رسیده است و این بدان معنا است که شاخه‌نگارهای متناظر با آن‌ها تقریباً مثل هم بوده‌اند یا شباهت زیادی به هم داشته‌اند. در بسیاری از حالت‌ها نیز همبستگی از مقادیری که نزدیک به صفر بوده‌اند به مقادیر فاصله‌دار از صفر رسیده‌اند و به مفهوم آن است که سیستم پس از تحلیل PCA نسبت به حالت پیشین، پایدارتر است؛ بنابراین می‌توان نتیجه گرفت در صورتی که خوشه‌بندی سلسله‌مراتبی بر اساس PCA صورت پذیرد، نتایج حاصل از آن اتکاپذیری بالاتری خواهد داشت.

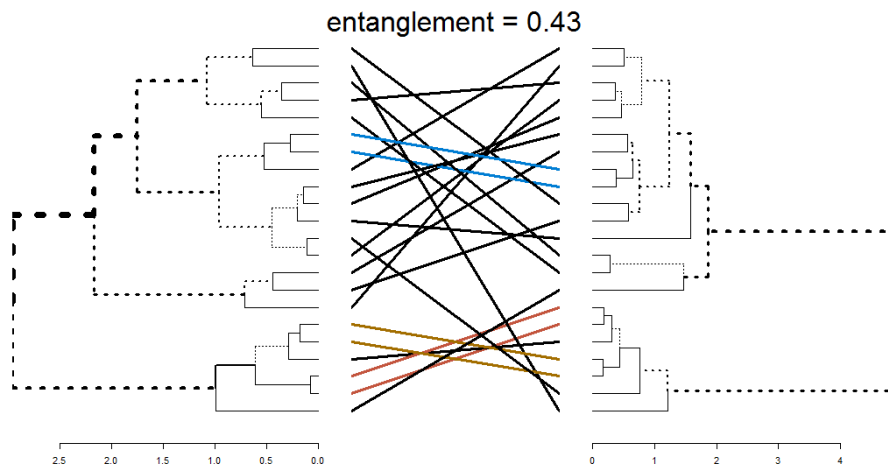
جدول ۶ «شاخص بیکر» بعد از PCA

	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
۱	۱											
۲	۰/۳۶۷	۱										
۳	۰/۵۲۲	۰/۵۵۹	۱									
۴	۰/۴۷۸	۰/۷۰۳	۰/۹۳۹	۱								
۵	۰/۵۰۱	۰/۶۶۳	۰/۸۴۴	۰/۸۹۹	۱							
۶	۰/۵۰۸	۰/۶۷۷	۰/۸۱۵	۰/۸۶۸	۰/۹۷۱	۱						
۷	۰/۵۴۸	۰/۶۷۶	۰/۸۶۳	۰/۹۰۳	۰/۹۸۱	۰/۹۸۳	۱					
۸	۰/۷۲۱	۰/۲۷۲	۰/۴۹۳	۰/۴۱۶	۰/۴۴۸	۰/۴۵۲	۰/۵۰۳	۱				
۹	۰/۳۵	۰/۲۳۹	۰/۳۲۸	۰/۳۹۹	۰/۵۱۹	۰/۴۸۱	۰/۴۶۴	۰/۳۲۸	۱			
۱۰	۰/۴۷	۰/۶۶۸	۰/۸۱۷	۰/۸۷۲	۰/۹۷۵	۰/۹۷۹	۰/۹۷۸	۰/۴۱۷	۰/۴۸۱	۱		
۱۱	۰/۳۸۸	۰/۳	۰/۳۲۲	۰/۳۷۶	۰/۴۴۶	۰/۵۵۳	۰/۴۷۴	۰/۳۷۵	۰/۳۴۱	۰/۴۶۳	۱	
۱۲	۰/۵۹۲	۰/۵۹۹	۰/۸۷۷	۰/۸۶۸	۰/۹۳۶	۰/۹۳۸	۰/۹۵۴	۰/۵۱۶	۰/۴۴۳	۰/۹۲۳	۰/۴۵۳	۱



نمودار ۱۱. خوشه‌بندی سلسله‌مراتبی بعد از PCA

در پایان خوشه‌بندی پایداری که حاصل از تحلیل PCA است، ارائه می‌شود. در این خوشه‌بندی، گروه‌های سه‌گانه‌ای مشخص شده است که ترکیب آن با خوشه‌بندی اولیه متفاوت است. این تفاوت، به دلیل نوفه‌هایی است که سیستم را تحت تأثیر قرار می‌دهند و در اثر تحلیل PCA بخش زیادی از آن‌ها، حذف شده‌اند. اگر این شاخه‌نگار و شاخه‌نگار قبل از PCA با یکدیگر مقایسه شوند، میزان شباهت آن‌ها نشان‌دهنده ثبات سیستم و عدم شباهت آن‌ها نشان‌دهنده ناپایداری سیستم است.



نمودار ۱۲. هم‌شاخه‌نگار خوشه‌بندی‌های سلسله‌مراتبی بعد و قبل از PCA

میزان درهم‌تنیدگی این دو خوشه‌بندی ۴۳ درصد است. اگر این عدد صفر باشد، به معنای شباهت کامل خوشه‌ها و اگر ۱۰۰ درصد باشد به معنای عدم شباهت کامل خوشه‌ها است؛ بنابراین در صورتی که دو خوشه‌بندی (یعنی قبل از تحلیل PCA و پایدارسازی خوشه‌ها یا بعد از تحلیل PCA) مقایسه شود، این عدد می‌تواند کمیتی برای اندازه‌گیری پایداری سیستم باشد. هر چه این عدد به صفر نزدیک باشد، ناپایداری آن کمتر و باثبات‌تر است.

۵. بحث و نتیجه‌گیری

از آنجاکه اندازه‌گیری پایداری یک سیستم مالی، نخستین گام در تخمین ریسک سیستمیک است، در این پژوهش با استفاده از خوشه‌بندی سلسله‌مراتبی شاخص‌های «بورس اوراق بهادار تهران» و پایدارسازی آن با استفاده از تحلیل PCA، کمیت جدیدی برای اندازه‌گیری میزان پایداری سیستم معرفی شد. نتایج پژوهش نشان می‌دهد که شاخه‌نگارهای شاخص‌ها در طول زمان متغیر هستند و بنابراین موقعیت متریک شاخص‌ها نسبت به یکدیگر مدام تغییر می‌کند. در صورتی که بتوان این تغییر را به مؤلفه‌های اصلی تجزیه کرد، می‌توان سهم مؤلفه‌های مهم و مؤلفه کمتر مهم را از یکدیگر تفکیک کرد؛ بنابراین پس از حذف تأثیر مؤلفه‌های کم‌اهمیت که در تحلیل PCA صورت می‌گیرد، می‌توان به خوشه‌بندی‌هایی دست یافت که از ثبات بیشتری برخوردارند. البته این پایداری کامل نیست؛ چراکه مؤلفه‌های اصلی انتخاب‌شده، خود دارای تغییراتی هستند. تغییرات یادشده را می‌توان تغییرات سیستمیک تلقی کرد.

برای مقایسه خوشه‌بندی‌ها و تشخیص شباهت موجود بین آن‌ها از شاخص‌های گوناگونی چون «شاخص بیکر» و ضریب همبستگی همسانی‌ها استفاده می‌شود. در این پژوهش مشخص شد که خوشه‌بندی‌ها، پس از تحلیل PCA از شباهت بیشتری برخوردارند و بنابراین پایداری بالاتری دارند. در نهایت خوشه‌بندی سلسله‌مراتبی حاصل از PCA با خوشه‌بندی اولیه مقایسه شد که میزان ناپایداری (۴۳ درصد) سیستم در قلمرو مطالعه به دست آمد. پژوهش حاضر برای محدوده زمانی یک‌ساله صورت گرفته و طبیعی است که با توسعه قلمرو پژوهش به چند سال و به‌ویژه سال‌هایی که تغییرات سیستمیک آشکاری در بازار بورس رخ داده است و نیز توسعه قلمرو آن به شاخص‌های صنایع گوناگون و سایر بازارهای مالی می‌توان نتایج حاصل از آن را تعمیم داد.

در انتها به سرمایه‌گذاران، تحلیلگران، سبذگردانان و نیز مدیران ریسک توصیه می‌شود در استفاده از تکنیک‌های خوشه‌بندی به عنوان ابزاری در متنوع‌سازی یا مدیریت ریسک به این مهم توجه داشته باشند در زمان‌های مختلف موقعیت متریک مشاهداتی که خوشه‌بندی می‌شوند بر اساس میزان ناپایداری سیستم تغییر می‌کند و بنابراین اتکا به آن‌ها ممکن است قابل تعمیم به

زمان‌های دیگر نباشد؛ مگر آنکه با روش‌هایی از جمله روش‌هایی که در این پژوهش به آن اشاره شد، خوشه‌بندی‌ها پایدار شود و شکل پایدارشده آن‌ها برای تصمیم‌گیری‌های مالی به‌کار رود. در پژوهش‌های آتی، پژوهشگران می‌توانند سایر روش‌های پایدارسازی را مدنظر قرار دهند و شکاف پژوهشی موجود را از طریق بسط قلمروهای این پژوهش یا اصلاح و تکمیل روش‌های آن ترمیم کنند.

منابع

1. Araújo, Ta., Eleutério, S., & Louçã, F. (2018). Do sentiments influence market dynamics? A reconstruction of the Brazilian stock market and its mood. *Physica A: Statistical Mechanics and its Applications*, 505, 1139-1149.
2. Baker, F. B. (1974). Stability of Two Hierarchical Grouping Techniques Case I: Sensitivity to Data Errors. *Journal of the American Statistical Association*, 69(346), 440-445.
3. Bank, W. (2013). *Global financial development report 2014: Financial inclusion*: World Bank Publications.
4. Brida, J. Gabriel, & Risso, W. Adrián. (2010). Hierarchical structure of the German stock market. *Expert Systems with Applications*, 37(5), 3846-3852.
5. Cai, Yumei, C., Xiaomei, H., Q., & Sun, J. (2017). Hierarchy, cluster, and time-stable information structure of correlations between international financial markets. *International Review of Economics & Finance*, 51, 562-573
6. Caruana, J. (2010). Systemic risk: how to deal with it? *BIS Publications*.
7. de Amorim, Renato Cordeiro. (2015). Feature relevance in ward's hierarchical clustering using the L p norm. *Journal of Classification*, 32(1), 46-62.
8. Deza, Michel-Marie, & Deza, Elena. (2006). *Dictionary of distances*: Elsevier.
9. Deza, Michel Marie, & Deza, Elena. (2009). Encyclopedia of distances *Encyclopedia of Distances* (pp. 1-583): Springer.
10. Esmalifalak, H., Ajirlou, A. Irannezhad, B., Pordeli, S. & Esmalifalak, M. (2015). (Dis)integration levels across global stock markets: A multidimensional scaling and cluster analysis. *Expert Systems with Applications*, 42(22), 8393-8402.
11. Everitt, Brian S, & Dunn, Graham. (2001). Principal components analysis. *Applied multivariate data analysis*, 48-73.
12. Fang, L., Xiao, B., Yu, H., & You, Q. (2018). A stable systemic risk ranking in China's banking sector: Based on principal component analysis. *Physica A: Statistical Mechanics and its Applications*, 492, 1997-2009.
13. Fowlkes, E. B., & Mallows, C. L. (1983). A Method for Comparing Two Hierarchical Clusterings. *Journal of the American Statistical Association*, 78(383), 553-569.
14. Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, 31(22), 3718-3720.
15. Goldberger, J., & Tassa, T. (2008). A hierarchical clustering algorithm based on the Hungarian method. *Pattern Recognition Letters*, 29(11), 1632-1638
16. Guerra, S, Maria, S, Thiago C, Tabak, B Miranda P, Rodrigo A, Miranda, R (2016). Systemic risk measures. *Physica A: Statistical Mechanics and its Applications*, 442, 329-342.
17. Huang, W. Q., & Wang, D. (2018). Systemic importance analysis of chinese financial institutions based on volatility spillover network. *Chaos, Solitons & Fractals*, 114, 19-30.
18. Huang, W.Q., Yao, S., Zhuang, X. T., & Yuan, Y. (2017). Dynamic asset trees in the US stock market: Structure variation and market phenomena. *Chaos, Solitons & Fractals*, 94, 44-53.
19. Iori, G., & Mantegna, R. N. (2018). Chapter 11 - Empirical Analyses of Networks in Finance. In C. Hommes & B. LeBaron (Eds.). Handbook of Computational Economics (Vol. 4, pp. 637-685): Elsevier.

20. Jian, Z., Deng, P., & Zhu, Z. (2018). High-dimensional covariance forecasting based on principal component analysis of high-frequency data. *Economic Modelling*, 75, 422-431.
21. Kleinow, J., Moreira, F., Strobl, S., & Vähämaa, S. (2017). Measuring systemic risk: A comparison of alternative market-based approaches. *Finance Research Letters*, 21, 40-46.
22. Lahmiri, S. (2016). Clustering of Casablanca stock market based on hurst exponent estimates. *Physica A: Statistical Mechanics and its Applications*, 456, 310-318.
23. Lee, Byung-Joo. (2018). Asian financial market integration and the role of Chinese financial market. *International Review of Economics & Finance*, 59, 490-499
24. Li, C.Z., Xu, Z.B., & Luo, T. (2013). A heuristic hierarchical clustering based on multiple similarity measurements. *Pattern Recognition Letters*, 34(2), 155-162.
25. Li, We., Hommel, U., & Paterlini, S. (2018). Network topology and systemic risk: Evidence from the Euro Stoxx market. *Finance Research Letters*, 27, 105-112
26. Li, Y., Hao, A., Zhang, X., & Xiong, X. (2018). Network topology and systemic risk in Peer-to-Peer lending market. *Physica A: Statistical Mechanics and its Applications*, 508, 118-130.
27. Lu, Ya-Nan, Li, Sai-Ping, Zhong, Li-Xin, Jiang, Xiong-Fei, & Ren, Fei. (2018). A clustering-based portfolio strategy incorporating momentum effect and market trend prediction. *Chaos, Solitons & Fractals*, 117, 1-15.
28. Matsen, F. A., Billey, S. C., Kas, A., & Konvalinka, M. (2018). Tanglegrams: A Reduction Tool for Mathematical Phylogenetics. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(1), 343-349.
29. Mondal, Sakib A. (2018). An improved approximation algorithm for hierarchical clustering. *Pattern Recognition Letters*, 104, 23-28.
30. Nobi, A., & Lee, J. W. (2016). State and group dynamics of world stock market by principal component analysis. *Physica A: Statistical Mechanics and its Applications*, 450, 85-94.
31. Rea, A., & Rea, W. (2014). Visualization of a stock market correlation matrix. *Physica A: Statistical Mechanics and its Applications*, 400, 109-123.
32. Rodríguez-Moreno, M., & Peña, J. I. (2013). Systemic risk measures: The simpler the better? *Journal of Banking & Finance*, 37(6), 1817-1831.
33. Sarlin, Peter, & Peltonen, Tuomas A. (2013). Mapping the state of financial stability. *Journal of International Financial Markets, Institutions and Money*, 26, 46-76.
34. Scornavacca, C., Zickmann, F., & Huson, D. H. (2011). Tanglegrams for rooted phylogenetic trees and networks. *Bioinformatics*, 27(13), i248-i256.
35. Sokal, R. R., & Rohlf, F. Ja. (1962). The comparison of dendrograms by objective methods. *Taxon*, 11(2) 33-40.
36. Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411-423.
37. Tumminello, M., Lillo, F., & Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization*, 75(1), 40-58.

38. Venkatachalam, B., Apple, J., John, K. St., & Gusfield, D. (2010). Untangling Tanglegrams: Comparing Trees by Their Drawings. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 7(4), 588-597.
39. Vinothkumar, K., & Selvan, M. P. (2014). Hierarchical Agglomerative Clustering Algorithm method for distributed generation planning. *International Journal of Electrical Power & Energy Systems*, 56, 259-269.
40. Wang, Y., Li, H., Guan, J., & Liu, N. (2019). Similarities between stock price correlation networks and co-main product networks: Threshold scenarios. *Physica A: Statistical Mechanics and its Applications*, 516, 66-77.
41. Zhao, L., Li, W., & Cai, X. (2016). Structure and dynamics of stock market in times of crisis. *Physics Letters A*, 380(5), 654-666.
42. Zhao, L., Wang, G.J., Wang, M., Bao, W., Li, W., & Stanley, H. E. (2018). Stock market as temporal network. *Physica A: Statistical Mechanics and its Applications*, 506, 1104-1112.