



کاربرد الگوریتم‌های تبرید شبیه‌سازی شده و ژنتیک در تشکیل صندوق شاخصی

ابراهیم عباسی^۱
صمد اکبری^۲

تاریخ پذیرش: ۹۳/۲/۲۰

تاریخ دریافت: ۹۲/۶/۲۰

چکیده

هدف از این مطالعه تشکیل صندوق شاخصی با استفاده از الگوریتم‌های تبرید شبیه‌سازی شده و ژنتیک می‌باشد. صندوق‌های شاخصی پرتفوی‌هایی هستند که طوری طراحی می‌شوند که از طریق سرمایه‌گذاری در تعداد معدودی از اقلام تشکیل دهنده شاخص بتوانند ضمن کاهش هزینه‌های معاملاتی بازدهی نزدیک به بازده بازار را ایجاد نمایند. در این پژوهش به دنبال بررسی رابطه میان خطای ردیابی و تعداد سهام تشکیل دهنده صندوق هستیم. همچنین عملکرد الگوریتم تبرید شبیه‌سازی شده را در مقایسه با الگوریتم ژنتیک در تشکیل صندوق شاخصی می‌سنجیم. در این پژوهش برای انتخاب سهام‌هایی که بیش‌ترین تأثیر را بر شاخص می‌گذارند از یک تابع اولویت استفاده خواهد شد، سپس با استفاده از الگوریتم‌های فرا ابتکاری وزن‌های بهینه این سهام‌ها را به دست خواهیم آورد. نتایج حاصل نشان می‌دهد که هر چه تعداد سهام صندوق بیشتر باشد خطای ردیابی کاهش می‌یابد. بررسی‌ها دقت بالا و عملکرد برتر صندوق شاخصی ایجاد شده توسط الگوریتم ژنتیک را در مقایسه با الگوریتم تبرید شبیه‌سازی شده به اثبات رسانید. به منظور ایجاد این صندوق از سهام شرکت‌های موجود در بورس اوراق بهادار تهران استفاده خواهد شد.

واژه‌های کلیدی: صندوق شاخصی، سبد سهام، الگوریتم ژنتیک، الگوریتم تبرید شبیه‌سازی شده.

۱- دانشیار و عضو هیئت علمی دانشگاه الزهراء، دانشکده علوم اجتماعی و اقتصادی abbasiebrahim2000@yahoo.com
۲- کارشناس ارشد مهندسی صنایع گرایش مهندسی مالی samadakbari@yahoo.com

۱- مقدمه

ردیابی شاخص یکی از راهبردهای کم ریسک تخصیص منابع است که به اعتقاد بسیاری از صاحب‌نظران، بازدهی بالاتری را نسبت به رویکرد فعال در یک افق زمانی بلندمدت ایجاد می‌کند (۱۷). ایجاد یک پرتفوی با عملکردی مشابه شاخص و به دنبال آن دستیابی به ریسکی نزدیک به بازار، منطق زیربنایی این رویکرد سرمایه‌گذاری است. با توجه به اینکه شاخص کل معیاری از روند بازده واقعی سهام موجود در بازار است (۳)، در شرایطی که بازار از حرکت رو به رشد برخوردار باشد می‌توان با طراحی صندوقی که به وسیله تعداد کمی از سهام موجود در بازار که بیش‌ترین تأثیر را روی شاخص می‌گذارند، رفتار شاخص را تکرار کرد و به بازدهی نزدیک به بازده بازار دست یافت. از مزایای این روش کاهش ریسک و هزینه‌های معاملاتی است که در تعیین بازده نهایی تأثیر بسزایی دارند. در واقع صندوق‌های شاخصی در تلاش هستند که حرکت شاخص را با تعداد کمی سهام شبیه‌سازی کنند. بنابراین مسئله تعیین تعداد و نوع سهام موجود در صندوق‌های شاخصی همواره یکی از دغدغه‌های اصلی فعالان این حوزه بوده است به طوری که پژوهش‌های زیادی در زمینه انتخاب این گزینه‌ها صورت گرفته است و هر کدام از محققان از متغیرها و روش‌های مختلفی برای دستیابی به هدف استفاده کرده‌اند. استفاده از این رویکرد در بازار سرمایه ایران که صندوق‌های شاخصی در آن کمتر رشد کرده‌اند می‌تواند مورد استفاده قرار گیرد و موجب ترغیب سرمایه‌گذاران به ورود در بازار سرمایه شود.

در این راستا در پژوهش حاضر، مسئله تشکیل صندوق شاخصی را مبتنی بر الگوریتم ژنتیک و الگوریتم تبرید شبیه‌سازی شده مورد مطالعه قرار خواهیم داد و با تحلیل مقایسه‌ای میان الگوریتم‌های پیشنهادی، رویکرد برتر جهت سرمایه‌گذاری شاخص محور در بورس تهران را مشخص می‌کنیم. در این پژوهش برای انتخاب سهامی که می‌بایست در صندوق شاخصی قرار گیرند از یک تابع اولویت‌بندی استفاده می‌شود که سهام‌هایی را که بیش‌ترین تأثیر را بر روی شاخص می‌گذارند انتخاب می‌نماید. در ادامه ابتدا مبانی نظری و پیشینه پژوهش را بیان کرده و سپس به بیان روش پژوهش و نتایج تجربی خواهیم پرداخت. در پایان نیز از مباحث مطرح‌شده و یافته‌های پژوهش نتیجه‌گیری و پیشنهادهایی برای پژوهش‌های آتی ارائه خواهیم کرد.

۲- مبانی نظری و مروری بر پیشینه پژوهش

مارکوویتز (۱۳ و ۱۲) با ارائه نظریه پرتفوی مدرن روش مناسبی را برای ایجاد یک پرتفوی با بالاترین بازده در سطح ریسک معین ارائه داد اما پس از نظریه پرتفوی میانگین-واریانس، پیشرفت‌های زیادی در نظریه پرتفوی صورت گرفت و کارهای عملی متنوعی در فرموله کردن پرتفوی انجام گرفت. یافته‌های اخیر نشان می‌دهد که صندوق‌های شاخصی کم هزینه‌ای که توسط مدیریت منفعل اداره می‌شوند بازده بالاتری نسبت به دیگر صندوق‌های سرمایه‌گذاری مشترک دارند (۵). مدیریت صندوق شاخصی یک استراتژی انتخاب سهام است که تلاش می‌کند عملکردی مانند شاخص مدنظر بدست آورد. صندوق شاخصی معمولاً تمام سهم‌هایی را که شاخص را می‌سازند، شامل نمی‌شود. این صندوق‌ها طوری طراحی می‌شوند که عملکرد

شاخص را با تعداد نسبتاً کمی از سهام تکرار کنند که به راحتی در بازار سرمایه قابل کنترل و مدیریت باشند. بنابراین عملکرد صندوق شاخصی به این بستگی دارد که ردیابی شاخص از طریق قرار دادن و جایگزینی مجموعه کوچکی از سهام داخل این پرتفوی، با چه مهارتی انجام شود. معیار سنجش ردیابی شاخص، خطای ردیابی است (۱۶) که توسط تغییرات خطای ردیابی، مجموعه انحراف معیارهای، تفاضل بین بازده صندوق ردیابی کننده از بازده شاخص مدنظر، اندازه‌گیری می‌شود. وقتی مدیر صندوق، یک صندوق شاخصی را طراحی می‌کند، سعی می‌کند تا سطح نوسان پذیری خطای ردیابی را کمینه نماید تا جایی که بازده صندوق به بازده شاخص نزدیک شود. در کل چندین روش برای اندازه‌گیری خطای ردیابی وجود دارد، برای مثال درجه دوم، خطی و مطلق که در این میان معیار سنجش درجه دوم به دلیل اینکه دارای تعدادی از خواص مطلوب آماری است از مطلوبیت بیشتری برخوردار است. در این مطالعه نیز از معیار سنجش درجه دوم استفاده شده است.

از آنجاییکه مساله ارایه شده در این پژوهش از نوع برنامه ریزی غیر خطی عدد صحیح و متعلق به مسائل دارای پیچیدگی بالا می باشد، می‌بایست از الگوریتم‌های فراابتکاری برای حل مدل استفاده نمود (۸). الگوریتم‌های مورد استفاده در این پژوهش الگوریتم ژنتیک و شبیه سازی تبرید می‌باشند که در ادامه توضیح مختصری در مورد آن‌ها آورده می‌شود.

الگوریتم ژنتیک

اصول بنیادی الگوریتم ژنتیک برای اولین بار توسط جان هالند در سال ۱۹۹۵ ابداع و مفاهیم الگوریتم ژنتیک در سال ۱۹۸۹ توسط گلبگ توسعه داده شد. الگوریتم ژنتیک یک روش جستجوی احتمالی است که از شبیه سازی تکاملی زیستی و طبیعی استفاده می کند. الگوریتم‌های ژنتیکی با بکارگیری اصل بقای بهترین‌ها برای تولید تخمین‌های هر چه بهتر یک جواب روی جمعیتی از جواب‌های بالقوه عمل می نمایند. در هر نسل، مجموعه ای از تخمین‌ها توسط فرآیند انتخاب افراد مطابق با سطح برازندگی‌شان در دامنه مساله و پرورش آنها با هم با استفاده از عملگرهای گرفته شده از ژنتیک طبیعی ایجاد می گردد. این فرایند ما را به سمت تکامل جمعیت‌هایی از افراد که با محیط مربوطشان بهتر از والدینشان وفق داده شده اند هدایت می کند. ساختار کلی الگوریتم ژنتیک به شرح زیر می‌باشد:

الف) ژن و کروموزم: ژن کوچکترین واحد سازنده الگوریتم ژنتیک می باشد. در حقیقت ژن‌ها برای نمایش شکل کد شده پارامترها می باشد. به رشته ای از ژن‌ها، کروموزوم می گویند.
ب) جمعیت: مجموعه ای از کروموزوم‌ها را جمعیت گویند. یکی از ویژگی‌های الگوریتم ژنتیک این است که بجای تمرکز بر روی یک نقطه از فضای جستجو یا یک کروموزوم بر روی جمعیتی از کروموزوم‌ها کار می کند.

ج) تابع برازندگی: تابع هدف جهت تعیین اینکه افراد چگونه در محدوده مساله ایفاء نقش می نمایند. در این تحقیق شایسته ترین افراد دارای کمترین مقدار عددی تابع هدف می‌باشند.

د) عملگرهای ژنتیک: برای تولید کروموزم های جدید (فرزندان) از طریق برخی از کروموزم های قدیمی نیاز به یکسری عملگرها داریم که در ادامه بحث به آن می پردازیم.

۱- عملگر تقاطع: عملگر تقاطع در یک لحظه بر روی دو کروموزم اعمال شده و دو فرزند به وسیله ترکیب ساختار دو کروموزوم ایجاد می کند. مفهوم مهمی که در ارتباط با این عملگر مطرح است نرخ تقاطع می باشد. نرخ تقاطع نسبت تعداد فرزندان تولید شده در هر نسل با اندازه جمعیت اصلی تعریف می شود.

۲- عملگر جهش: در سیر تکاملی طبیعی، جهش یک فرایند تصادفی است که در آن محتوای یک ژن با ژن دیگر جهت تولید یک ساختار ژنتیکی جدید جایگزین می گردد. نقش جهش اغلب به عنوان تضمینی است برای آنکه احتمال جستجو در رشته هرگز صفر نگردد.

ه) انتخاب کروموزوم و شرط توقف: عملگر انتخاب کروموزومهایی را از جمعیت جاری برای شرکت در مرحله بعدی با توجه به مقدار شایستگی شان انتخاب می کند. چندین روش انتخاب، نظیر چرخه رولت، تورنامنت، رتبه بندی و نخبه گرایی وجود دارد که در این پژوهش از چرخه رولت استفاده می شود. در GA بکارگیری شرط توقف آخرین گام بحساب می آید. به طور خلاصه معیارهای توقف در کل عبارتند از: الف) توقف روش با تعداد مشخصی از نسل، ب) به اتمام رسیدن با زمان محدود و ج) عدم بهبودی بهترین پاسخ در طی مدت زمان مشخص. در این پژوهش از تعداد نسل معین استفاده می شود.

الگوریتم تبرید شبیه سازی شده

الگوریتم تبرید شبیه سازی شده، یک الگوریتم بهینه سازی فرآینتکاری ساده و اثربخش در حل مسائل بهینه سازی است. منشأ الگوریتم تبرید شبیه سازی شده، کارهای کریک پاتریک و کرنی و همکارانشان در سال های ۱۹۸۳ و ۱۹۸۵ است (۸). آنها برای حل مسائل پیچیده بهینه سازی، روشی مبتنی بر تکنیک تبرید تدریجی پیشنهاد نمودند. الگوریتم تبرید شبیه سازی شده برای حل یک مسئله بهینه سازی، ابتدا از یک جواب اولیه شروع می کند و سپس در یک حلقه تکرار به جواب های همسایه حرکت می کند. اگر جواب همسایه بهتر از جواب فعلی باشد، الگوریتم آن را به عنوان جواب فعلی قرار می دهد (به آن حرکت می کند)، در غیر این صورت، الگوریتم آن جواب را با احتمال $e^{-\Delta E/T}$ به عنوان جواب فعلی می پذیرد. در این رابطه ΔE تفاوت بین تابع هدف جواب فعلی و جواب همسایه است و T یک پارامتر به نام دما است. در هر دما، چندین تکرار اجرا می شود و سپس دما به آرامی کاهش داده می شود. در گام های اولیه دما خیلی بالا قرار داده می شود تا احتمال بیشتری برای پذیرش جواب های بدتر وجود داشته باشد. با کاهش تدریجی دما، در گام های پایانی احتمال کمتری برای پذیرش جواب های بدتر وجود خواهد داشت و بنابراین الگوریتم به سمت یک جواب خوب همگرا می شود.

در این الگوریتم، جواب های مسأله معادل اتم های جسمی است که می خواهیم تبرید کنیم و تابع هدف کمینه کردن میزان بی نظمی و تنش بین مولکول ها است. از این رو برای رسیدن به این هدف، باید دما را

کاهش داد تا میزان تنش کاهش یابد. دما در هر مرحله این الگوریتم، آستانه مجاز بی‌نظمی را نشان می‌دهد (۸). یکی دیگر از نکات اساسی در SA، استراتژی جستجوی همسایگی است. این امر می‌تواند با توجه به ماهیت مسأله، بسیار پیچیده باشد. چگونگی حرکت از نقطه‌ای به نقطه دیگر بستگی به ساختار عملگرهای طراحی شده دارد. دستیابی به نقاط کشف نشده فضای جواب و همچنین جستجوی دقیق همسایگی نقاط، منوط به طراحی عملگرهای کارا و مؤثر می‌باشد. این کار در بسیاری از مسایل پیچیده بهینه‌سازی با متغیرهای تصمیم متنوع و محدودیت‌های متعدد، امری دشوار است، تا آنجائیکه ممکن است بسیاری از رویکردهای فراابتکاری در عمل کارایی خود را از دست بدهند.

در پژوهش‌های مختلف روش‌های متنوعی برای ایجاد صندوق شاخص پیشنهاد شده است. بی‌زلی و همکاران (۴) از رویکرد تکاملی الگوریتم ژنتیک استفاده کرده و علاوه بر مسئله ایجاد پورتفوی شاخصی، تعدیل ترکیب آن و هزینه‌های معاملاتی را نیز مدنظر قرار دادند. رافائلی و همکاران (۱۵) یک رویکرد مقایسه‌ای مبتنی بر الگوریتم ژنتیک و برنامه‌ریزی کوادراتیک را به منظور ایجاد صندوق شاخصی بر مبنای شاخص بورس انگلیس توسعه دادند. نتایج ارائه‌شده مزیت روش ژنتیک را با در نظر گرفتن اندازه‌های مختلف و همچنین دوره‌های متعدد به‌روزرسانی ترکیب برای پورتفوی شاخصی به اثبات رسانید. کوریلی و مارسلینو (۷) در خصوص یک روش تشکیل صندوق شاخص مبتنی بر این فرض که قیمت‌های سهام متأثر از یک مدل عاملی هستند، پژوهش نمودند. در روش مورد استفاده آن‌ها شاخص و پورتفوی شاخصی آن دارای ساختار عاملی مشابه بودند. رویکرد بکار گرفته‌شده عبارت است از مرتب‌سازی عوامل و سپس اضافه کردن سهمی به پورتفوی که دارای بیش‌ترین همبستگی با عامل مورد نظر است. توروبیانو و سوارز (۱۸) یک راهبرد ترکیبی متشکل از یک الگوریتم تکاملی و برنامه‌ریزی کوادراتیک را به منظور حل مسئله صندوق شاخص طراحی کردند. یانسن و فن دیک (۹)، مسئله کمینه‌سازی خطای تشکیل صندوق را با در نظر گرفتن محدودیت برای تعداد سهام تشکیل‌دهنده پورتفوی، مدنظر قرار دادند. در روش مورد استفاده آن‌ها یک تابع هدف موزون متشکل از خطای تشکیل صندوق و تعداد سهام تشکیل‌دهنده پورتفوی کمینه می‌شود. زمانی که مجموعه سهام مشخص می‌شود، وزن‌های بهینه با استفاده از یک رویکرد استاندارد برنامه‌ریزی کوادراتیک محاسبه خواهد شد (کولمن و همکاران نیز چنین رویکردی را مورد استفاده قرار دادند (۶)). رولند و همکاران (۱۰) پرتفویی را برای ردیابی شاخص هلند ایجاد کردند که این پرتفوی شامل مجموعه از سهام‌های تشکیل‌دهنده این شاخص بود. وزن‌های بهینه این پرتفوی توسط کمینه کردن خطای ردیابی برای مجموعه‌ای از داده‌های تاریخی بازده و کوورایانس به دست آمد. برای بهینه کردن خطای ردیابی توسط الگوریتم ژنتیک، یک تابع برازندگی برای هر کروموزوم تعریف‌شده که برابر با حداقل خطای ردیابی دست‌یافتنی است. اوه و همکاران (۱۴) برای هر سهم یک تابع اولویت تعریف کردند. تابع مذکور عبارت بود از حاصل جمع موزون حجم معامله، ارزش بازار و بتا. یک هیوریستیک ساده که از این توابع رجحان بهره می‌گیرد، به منظور تعیین سهام تشکیل‌دهنده پورتفوی بکار گرفته شد، سپس از الگوریتم ژنتیک به منظور تعیین وزن بهینه هر سهم استفاده شد. حنیفی، بحرالعلوم و جوادی (۲) در پژوهش خود از الگوریتم‌های فرا ابتکاری جهت پیاده‌سازی

سرمایه‌گذاری شاخص محور در بورس تهران استفاده کردند. پژوهش آن‌ها مسئله انتخاب پورتفوی بهینه ردیابی کننده شاخص کل قیمت و بازده نقدی را با بهره‌گیری از رویکردهای فرا ابتکاری دنبال می‌کند. به منظور دستیابی به نمونه‌ای با این ویژگی، از روش نمونه‌گیری قضاوتی استفاده کردند. به عبارت دیگر آن دسته از شرکت‌هایی که در بازه زمانی پژوهش، دارای بیش از ۱۰۰ روز معاملاتی در سال باشند و لذا فعال‌ترند را انتخاب کردند

۳- فرضیه‌های پژوهش

فرضیه اول: میان خطای ردیابی و تعداد سهام تشکیل‌دهنده پرتفوی همبستگی وجود دارد.
فرضیه دوم: الگوریتم ژنتیک از دقت بالاتری در ردیابی شاخص نسبت به الگوریتم تبرید شبیه‌سازی شده برخوردار است.

۴- روش شناسی پژوهش

ردیابی شاخص عبارت است از ایجاد و طراحی یک پرتفوی سرمایه‌گذاری به منظور دستیابی به عملکردی مشابه با شاخص مینا. به دلیل کاهش هزینه‌های معاملاتی تنها زیر مجموعه‌ای از سهام تشکیل‌دهنده در صندوق شاخصی لحاظ خواهد شد (۱۰). به عبارت دیگر در جستجوی مجموعه‌ای مناسب از L سهم هستیم که به خوبی شاخص را در یک بازه زمانی در گذشته $(0, T)$ ردیابی کرده و بتواند به گونه‌ای اثربخش عملکردی مشابه با شاخص را در زمان $[T, T + \epsilon]$ ایجاد نماید. تابع هدف این مسئله حداقل کردن خطای ردیابی این زیرمجموعه از سهام است که در این پژوهش تابعی از اختلاف میان بازده پورتفوی ردیابی کننده و شاخص در نظر گرفته شده است. پژوهش حاضر از یک فرآیند دو مرحله‌ای تشکیل شده است:

- انتخاب مجموعه‌ای از سهام ردیابی کننده شاخص توسط محاسبه تابع اولویت
- بهینه‌سازی سهام انتخاب‌شده توسط الگوریتم‌های فرا ابتکاری

۴-۱- انتخاب مجموعه‌ای از سهام ردیابی کننده شاخص توسط محاسبه تابع اولویت

برای انتخاب مجموعه‌ای از سهام که بیش‌ترین تأثیر را بر روش شاخص دارند، فرض می‌کنیم n تعداد سهام موجود در شاخص و L تعداد سهام انتخاب‌شده برای صندوق شاخصی باشد ($L < n$). همچنین c_k ، $(k=1, 2, \dots, L)$ ، k امین سهمی باشد که در صندوق قرار می‌گیرد. به عبارت دیگر $\Phi_p = \{c_1, c_2, \dots, c_L\}$ مجموعه پرتفوی صندوق شاخصی است که از میان n سهم موجود در شاخص انتخاب‌شده است. s نمایانگر تعداد بخش‌های صنعت است که در شاخص قرار دارند و d_i تعداد سهمی است که در i امین بخش صنعت وجود دارد ($\sum_{i=1}^s d_i = n$). همچنین بتا برای i امین سهم از i امین بخش صنعت ($i=1, 2, \dots, s$ و $j=1, 2, \dots, d$) با β_{ij} نشان داده می‌شود. برای یک سهم خاص $x_{i(j)}(t)$ و $A_{i(j)}(t)$ و $M_{i(j)}(t)$ نشان‌دهنده نرخ بازده، حجم معاملات و ارزش بازار در زمان t است. $\bar{r}_m(t)$ نرخ بازده شاخص در زمان t است.

در اینجا یک تابع اولویت، $P_{i(j)}$ ، برای هر سهم تعریف می‌کنیم ($j=1,2,d_i$ و $i=1,2,\dots,s$):

$$P_{i(j)} = v_1 \{B_{i(j)}\}^{-1} + v_2 \bar{A}_{i(j)} + v_3 \bar{M}_{i(j)} \quad (1)$$

که v_1 ، v_2 و v_3 مقادیر مثبت از پیش تعیین‌شده هستند و $B_{i(j)}$ انحراف

معیار استاندارد $\hat{\beta}_{i(j)}$ ($\hat{\beta}_{i(j)}$ تخمینی برای $\beta_{i(j)}$ است) است. $\bar{A}_{i(j)}$ و $\bar{M}_{i(j)}$ دارای مقادیر از صفر تا یک هستند. توجه داشته باشید که $P_{i(j)}$ در نهایت نشان‌دهنده اهمیت نسبی یک سهم خاص در پرتفوی است.

روند این الگوریتم برای انتخاب سهام برای صندوق شاخصی به صورت زیر است:

ابتدا ارزش بازار را برای هر بخش از صنعت محاسبه می‌کنیم ($mc_1(1), mc_2(1), \dots, mc_s(1)$) سپس فرایند زیر را برای $k=1,2,\dots,L$ تکرار می‌کنیم تا این گام L سهم را برای پرتفوی انتخاب کند. $\bar{A}_{i(k)}$ امین بخش صنعت را که دارای بزرگ‌ترین ارزش بازار است را انتخاب می‌کنیم. $mc_{i(k)} = \max_{i=1,2,\dots,s} mc_i(k)$ برای i_k امین بخش صنعت انتخاب‌شده $P_{i_k(j)}$ را برای $j=1,2,\dots,d_{i_k}$ محاسبه کرده و سهم $i_k(j_k)$ را که دارای بیشتر مقدار اولویت ($P_{i_k(j_k)} = \max_{j=1,2,\dots,d_{i_k}} P_{i_k(j)}$) است را به پرتفوی اضافه کرده و آن سهم را از بخش صنعت i_k که انتخاب‌شده بود حذف می‌نماییم، سپس بدون این سهم ارزش بازار صنعت‌ها را به روز می‌کنیم ($mc_{1(k+1)}, \dots, mc_{s(k+1)}$) و به ابتدای این گام باز می‌گردیم.

۴-۲- بهینه‌سازی سهام انتخاب‌شده توسط الگوریتم‌های ابتکاری

مدل پیشنهادی این پژوهش برای بهینه‌سازی وزن سهم‌های تشکیل‌دهنده صندوق شاخصی به شکل زیر می‌باشد:

$$\text{Minimize} \quad \frac{(\sum_{t=1}^T [(\sum_{i=1}^n z_i r_{it} w_i) - R_t]^2)^{\frac{1}{2}}}{T} \quad (2)$$

$$\text{Subject to:} \quad \sum_{i=1}^n z_i w_i = 1 \quad (3)$$

$$\sum_{i=1}^n z_i = L \quad (4)$$

$$0 < w_i < 1, i = 1, 2, \dots, n \quad (5)$$

$$z_i \in \{0,1\} \quad (6)$$

که در آن:

T : دوره زمانی که در آن اقدام به ردیابی شاخص می‌نماییم.

R_t : بازده شاخص در زمان t

r_{it} : بازده سهم i ام (در زمان t) از مجموعه سهام در اختیار جهت تشکیل صندوق شاخصی

L : تعداد سهام تشکیل‌دهنده صندوق شاخصی

Z_i : متغیری است که در صورت وجود سهم i ام در صندوق شاخصی معادل یک و در غیر این صورت برابر صفر در نظر گرفته می‌شود.

W_i : وزن سهم i ام در پرتفوی ردیابی کننده.

رابطه ۲ نشان‌دهنده خطای ردیابی و تابع هدف در مسئله مورد نظر است. رابطه ۳ و ۵ به محدودیت وزن ارقام تشکیل‌دهنده پرتفوی و معادله ۴ و ۶ به محدودیت عدد صحیح اشاره دارند. بر اساس محدودیت ۴، اگر در دارایی i سرمایه‌گذاری شود، مقدار Z_i برابر یک و در غیر این صورت برابر صفر خواهد بود. پارامتر L نشانگر تعداد سهامی است که سرمایه‌گذار مایل به سرمایه‌گذاری در آن می‌باشد، بنابراین این محدودیت، سرمایه‌گذاری در L سهم از n سهم را تضمین می‌نماید این محدودیت، فضای جستجو را به یک فضای گسسته و غیرخطی بدل می‌نماید. این امر موجب پدید آمدن ترکیبی پیچیده از برنامه‌ریزی عدد صحیح غیرخطی شده که یک مسئله سخت برای حل است و برای حل آن از الگوریتم‌های فرا ابتکاری استفاده می‌نماییم (۱).

برای جمع‌آوری داده‌ها ابتدا با استفاده از سایت بورس اوراق بهادار و نرم‌افزار ره آورد نوین داده‌های مورد نیاز جمع‌آوری شده و سپس فیلترینگ آن‌ها در محیط نرم‌افزار Excel صورت پذیرفت و به منظور اجرا و ارزیابی الگوریتم پیشنهادی از نرم‌افزار MATLA استفاده شد. برای آزمون فرضیه‌ها نیز از نرم‌افزار SPSS استفاده شده است. همچنین در این پژوهش از اطلاعات روزانه سهام شرکت‌ها از تاریخ ۹۱/۱/۵ الی ۹۱/۶/۳۱، به منظور انتخاب سهم‌هایی که بیش‌ترین تأثیر را بر روی شاخص می‌گذارند و یا به عبارتی به عنوان داده‌های آموزش و اطلاعات مربوط به بازه زمانی ۹۱/۷/۱ الی ۹۱/۱۲/۲۸ به عنوان داده‌های آزمایش و به عبارتی جهت ارزیابی عملکرد صندوق شاخصی در نظر گرفته شده است. برای بدست آوردن بهترین نرخ‌های تقاطع و جهش از روش تیونینگ استفاده شده است. ابتدا به صورت تصادفی چند عدد در بازه $0,9-0,7$ برای نرخ تقاطع انتخاب گردید و سپس با مشاهده بهترین جواب برای مسئله سعی شد تا عدد تصادفی بعدی به نرخ تقاطع بهتر نزدیک باشد. همین روش در بازه $0,4-0,2$ برای نرخ جهش نیز تکرار شد. در نتیجه پارامترهای بهینه به این ترتیب بدست آمد: تعداد تکرارهای الگوریتم: ۲۰۰ مرتبه، نرخ تقاطع: $0,8$ و نرخ جهش: $0,35$. در مورد الگوریتم تبرید نیز همین فرآیند تکرار شد و پارامترهای بهینه به این ترتیب تعیین گردید: دمای اولیه: 100 ، تابع کاهش دما: $0,99$ و تعداد تکرار الگوریتم ۲۰۰ مرتبه. مروری بر ادبیات موضوع نشان می‌دهد که پارامترها در پژوهش‌های مشابه به مقادیر این پژوهش نزدیک است (۳، ۱۱ و ۱۲).

۵- تشکیل صندوق شاخصی با استفاده از داده‌های تاریخی

پس از تعیین پارامترهای ورودی، الگوریتم‌های حل مسئله پژوهش اجرا و نتایج آن در جدول ۱ آورده شده است. به دلیل ماهیت تصادفی الگوریتم‌های ابتکاری، برای حل این مسئله الگوریتم‌ها را صد بار تکرار کرده و نتایج پس از محاسبه میانگین در جدول فوق درج گردیده است.

جدول ۱- خروجی الگوریتم‌ها با توجه به محدودیت عدد صحیح

زمان اجرا	بهترین خطای ردیابی (دوره آموزش)	میانگین خطای ردیابی (دوره آموزش)	تعداد سهام صندوق
الگوریتم ژنتیک			
۲۵,۷۴۳۵	1.06×10^{-7}	1.62×10^{-7}	۵
۲۹,۴۸۱۷	4.10×10^{-8}	5.60×10^{-8}	۱۰
۳۳,۴۱۸۲	2.40×10^{-8}	4.40×10^{-8}	۱۵
۳۵,۶۵۸۸	1.70×10^{-8}	1.90×10^{-8}	۲۰
الگوریتم تبرید شبیه‌سازی شده			
۲۹,۹۱۳۴	3.14×10^{-7}	4.37×10^{-7}	۵
۳۱,۲۶۶	1.64×10^{-7}	2.39×10^{-7}	۱۰
۳۴,۴۲۴	1.16×10^{-7}	1.71×10^{-7}	۱۵
۳۷,۶۲۴۳	3.9×10^{-8}	6.2×10^{-8}	۲۰

البته در ستون سوم جدول بهترین مقدار تابع هدف و یا به عبارتی کم‌ترین خطای ردیابی از میان این صد بار تکرار درج شده است. با توجه به نتایج می‌توان نتیجه گرفت که با افزایش تعداد سهام تشکیل‌دهنده پرتفوی، میانگین خطای ردیابی و بهترین مقدار تابع هدف کاهش می‌یابد. به منظور بررسی این رابطه از ضریب همبستگی پیرسون استفاده شد که مقدار آن برابر ۹۶,۸۷- درصد و سطح معناداری دو دامنه آن ۰,۰۰۰ محاسبه شد بنابراین همبستگی میان تعداد سهام پرتفوی و خطای ردیابی با اطمینان ۹۵ درصد تأیید شد و فرضیه اول پژوهش مورد تأیید قرار گرفت. علت این امر نزدیک شدن ترکیب صندوق به ترکیب شاخص، افزایش تنوع‌پذیری و افزایش قابلیت آن در دستیابی به عملکردی مشابه شاخص است. همان‌طور که انتظار می‌رفت زمان مورد نیاز برای حل مسئله با افزایش تعداد سهام تشکیل‌دهنده صندوق به دلیل افزایش میزان محاسبات افزایش می‌یابد. نتایج حاصل از تجزیه و تحلیل داده‌ها نشان می‌دهد که الگوریتم ژنتیک پیشنهادی نسبت به الگوریتم تبرید شبیه‌سازی شده از دقت بالاتری در ردیابی شاخص برخوردار است چرا که خطای ردیابی کمتری را به ازای مقادیر مختلف تعداد سهام تشکیل‌دهنده صندوق به خود اختصاص داده است. به منظور بررسی آماری این موضوع از آزمون مقایسه زوج‌ها استفاده شد. آماره میانگین در این آزمون که نشانگر تفاوت میان خطای ردیابی حاصل از الگوریتم ژنتیک و الگوریتم SA است برابر با ۳,۲۲- محاسبه گردید. بدین ترتیب پایین‌تر بودن دقت ردیابی الگوریتم SA نسبت به الگوریتم ژنتیک قابل استنتاج است. از آنجایی که سطح معناداری محاسبه‌شده (۰,۰۲۴) از سطح خطای ۰,۰۵ کمتر است فرض صفر رد شده و فرضیه دوم پژوهش نیز با اطمینان ۹۵ درصد مورد تأیید قرار می‌گیرد. با توجه به اینکه الگوریتم ژنتیک از توانمندی بالاتری در ردیابی شاخص برخوردار است، بهترین پرتفوی حاصل از این روش را که دارای ۲۰ سهام است مورد بررسی بیشتر قرار می‌دهیم. همان‌گونه که در جدول ۲ مشاهده می‌شود،

پرتفوی ردیابی کننده به خوبی و با دقت بالایی شاخص را ردیابی می‌کند به گونه‌ای که خطای ردیابی برابر $10^{-8} * 1,70$ و ضریب همبستگی

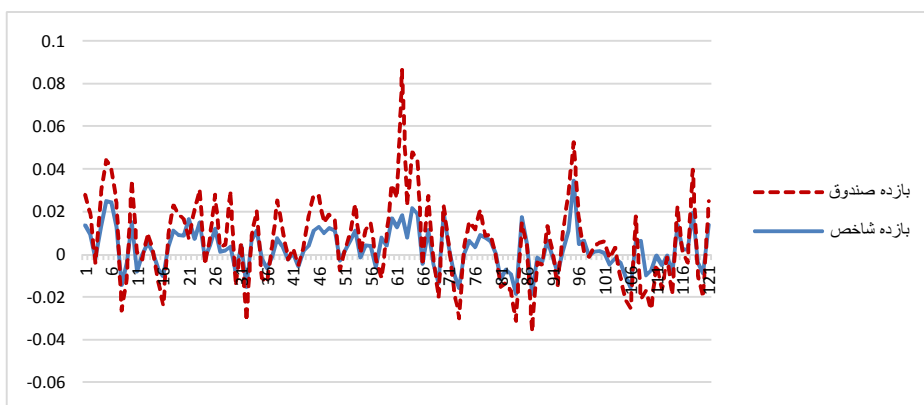
جدول ۲- نتایج اجرای الگوریتم‌ها با محدودیت عدد صحیح در دوره آموزش و آزمایش

خطای ردیابی (آموزش)	بازده اضافی (آموزش)	ضریب همبستگی (آموزش)	خطای ردیابی (آزمایش)	بازده اضافی (آزمایش)	ضریب همبستگی (آزمایش)
۱,۷۰*۱۰ ^{-۸}	۳,۷۳٪	۸۷٪	۱,۹۹*۱۰ ^{-۷}	۷,۰۰٪	۷۶٪
۱,۵۰*۱۰ ^{-۹}	۲,۸۵٪	۹۰٪	۱,۸۱*۱۰ ^{-۷}	۷,۷۹٪	۹۲٪

میان سری‌های زمانی بازده پرتفوی و شاخص برابر با ۸۷ درصد است. بعد از مشخص شدن توانمندی الگوریتم طراحی شده در ردیابی شاخص، هدف بهبود و توسعه الگوریتم پیشنهادی و ارزیابی عملکرد آن در خصوص داده‌های آتی (آزمایش) را فراتر از اهداف اولیه تعیین شده پژوهش و به عنوان نوآوری‌های این پژوهش دنبال می‌کنیم.

۵-۱- تشکیل صندوق شاخصی با استفاده از داده‌های آزمایش

بر اساس نتایج حاصل از تجزیه و تحلیل داده‌ها صندوق متشکل از ۲۰ سهم که مبنی بر الگوریتم ژنتیک ایجاد شده بود، کم‌ترین خطای ردیابی را حاصل نمود، بنابراین در این بخش نیز عملکرد این پرتفوی را در خصوص داده‌های آزمایش (بازه زمانی ۹۱/۰۷/۰۱ الی ۹۱/۱۲/۲۹) مورد بررسی و تحلیل قرار می‌دهیم.



نمودار ۱- بازده صندوق شاخصی در مقایسه با شاخص در دوره آزمایش

همان‌گونه که از نمودار ۱ قابل استنتاج است خطای ردیابی در خصوص داده‌های آزمایش افزایش یافته به گونه‌ای که مقدار آن به $10^{-7} * 1,99$ و ضریب همبستگی میان سری‌های زمانی بازده صندوق و شاخص به ۷۰ درصد رسیده است.

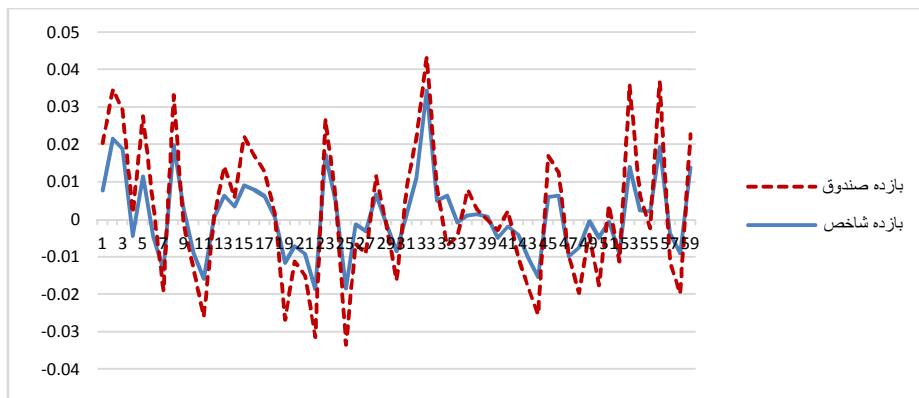
عملکرد صندوق شاخصی در خصوص داده‌های آزمایش از دو منظر قابل بررسی است:

۱. ترکیب صندوق شاخصی بایستی به صورت منظم و در فواصل زمانی نسبتاً مشخص مورد بازبینی قرار گیرد. علت این موضوع تغییر در ترکیب شاخص در نتیجه ورود و خروج شرکت‌ها به بورس و همچنین تعدیل وزن ارقام تشکیل‌دهنده شاخص می‌باشد. تعیین فواصل زمانی مورد نظر امری غیرقطعی بوده و بایستی توسط تحلیلگر و با توجه به رویدادهای بازار مشخص شود. بر اساس دیدگاه بیزلی (۴) این فاصله زمانی بایستی کمتر از ۶ ماه باشد. البته این زمان تعیین شده، سازگار با بورس‌های پیشرفته‌ای است که از تلاطم کمتری برخوردارند، لذا این فاصله زمانی برای بورس تهران بسیار کوتاه‌تر پیشنهاد می‌گردد. با توجه به نمودار ۱ تا حدود ۳ ماه پس از انتخاب پرتفوی، ردیابی شاخص با دقت بالایی صورت پذیرفته و ضریب همبستگی در این بازه زمانی برابر ۷۸ درصد محاسبه شده است. بنابراین می‌توان استدلال نمود که ترکیب پرتفوی پس از گذشت یک بازه زمانی نیازمند بروز رسانی است و این امر بخش جدایی‌ناپذیر از فرآیند ردیابی شاخص می‌باشد.
۲. افزایش خطای ردیابی و کاهش ضریب همبستگی در خصوص داده‌های آزمایش در جهت مثبت بوده است، یعنی عدم تطابق با شاخص به دلیل عملکرد بهتر پرتفوی مورد نظر نسبت به شاخص است. این موضوع از طریق محاسبه تفاوت میانگین بازده پرتفوی و شاخص در بازه زمانی اشاره‌شده قابل بررسی است که نتیجه آن در حدود ۴,۵۲ درصد بازده اضافی را نشان می‌دهد.

۵-۲- به روز رسانی ترکیب صندوق شاخصی

بررسی نوسانات بازده صندوق شاخصی در دوره آزمایش و مقایسه آن با شاخص که در نمودار ۱ قابل مشاهده است، قابلیت صندوق شاخصی تشکیل‌شده در ردیابی مناسب شاخص را بعد از گذشت ۶۰ روز معاملاتی به اثبات می‌رساند. علاوه بر بررسی شماتیک، ضریب همبستگی ۹۲ درصدی میان سری زمانی بازده پرتفوی و شاخص در این بازه زمانی ۳ ماهه این موضوع را مورد تأیید قرار می‌دهد. همان‌طور که در نمودار ۱ مشخص است، خطای ردیابی بعد از گذشت دوره ۳ ماهه افزایش می‌یابد. لذا بروز رسانی ترکیب صندوق و مطالعه اثر آن بر خطای ردیابی در بازه زمانی باقی‌مانده را مد نظر قرار می‌دهیم. بدین منظور پایان دوره آموزش با ۳ ماه افزایش از تاریخ ۹۱/۶/۳۱ به ۹۱/۹/۳۰ منتقل شد. پس از بهینه‌سازی وزن سهام تشکیل‌دهنده پرتفوی با استفاده از داده‌های جدید آموزش، پرتفوی مورد نظر نگهداری، شاخص‌های ردیابی محاسبه و مقایسه شماتیک میان سری زمانی بازده صندوق و شاخص در دوره آزمایش انجام شد که نتایج حاصل در ادامه آورده شده است. خطای ردیابی صندوق شاخصی بروز شده برابر با $10^{-7} * 1,16$ و بازده اضافی

آن نسبت به شاخص برابر با ۱,۱۹٪ محاسبه شد. همبستگی میان سری‌های زمانی بازده صندوق و شاخص برابر با ۸۳ درصد می‌باشد.



نمودار ۲- بازده صندوق شاخصی بروز شده در مقایسه با شاخص در دوره آزمایش

به منظور بررسی بهبود حاصل شده در فرآیند ردیابی، خطای ردیابی و بازده اضافی پرتفوی بهینه قبل از بروز رسانی، در بازه زمانی جدید داده‌های آزمایش محاسبه شد که نتایج آن عبارت‌اند از: خطای ردیابی برابر با $10^{-7} * 1,94$ و بازده اضافی نسبت به شاخص -۵,۴٪ تعیین گردید. همان طور که مشخص است در نتیجه بروز رسانی ترکیب صندوق شاخصی خطای ردیابی کاهش و بازده اضافی نسبت به شاخص از -۵,۴٪ به ۱,۱۹٪ رسید. به عبارت دیگر نوسانات منفی بازده پرتفوی نسبت به شاخص با نوسانات مثبت جایگزین و بدین ترتیب عملکردی بهتر حاصل شد.

جدول ۳- نتایج اجرای الگوریتم‌ها قبل و پس از به روز رسانی داده‌ها

خطای ردیابی	ضریب همبستگی	بازده اضافی
قبل از به روز رسانی	۷۱٪	-۵,۴٪
پس از به روز رسانی	۸۳٪	۱,۱۹٪

۶- نتیجه گیری و بحث

در این پژوهش از یک الگوریتم دو مرحله‌ای برای تشکیل صندوق شاخصی استفاده کردیم و با بررسی رویکردهای مختلف و تحلیل مقایسه‌ای آن‌ها رویکرد برتر را که در بورس اوراق بهادار تهران به نتیجه‌ای مطلوب‌تر دست یافته بود، مشخص نمودیم. در این راستا نتایج حاصل از تجزیه و تحلیل داده‌ها نشان داد که

با افزایش تعداد سهام تشکیل‌دهنده پرتفوی، میانگین خطای ردیابی و بهترین مقدار تابع هدف کاهش می‌یابد. بررسی‌ها عملکرد بهتر الگوریتم ژنتیک در مقایسه با الگوریتم تبرید شبیه‌سازی شده را به دلیل دستیابی به صندوقهایی با خطای ردیابی کمتر به اثبات رسانید. نکته قابل توجه در این میان عملکرد مشابه و فراتر از شاخص توسط صندوق‌های حاصل از این روش در دفعات مختلف تکرار و در خصوص داده‌های آزمایش است. این در حالی است که در تحقیقات انجام شده توسط بیزلی (۴)، یورسین (۱۰) و رافائلی (۱۵) عملکرد مشابه و پایین‌تر نسبت به شاخص، نتیجه به کارگیری الگوریتم‌های پیشنهادی آن‌ها بوده است. از طرف دیگر دقت ردیابی این مدل نیز مشابه و یا بهتر از روش‌های ارائه شده در تحقیقات دیگر است. دقت ردیابی این الگوریتم در حد ۷-۱۰ می‌باشد و این در حالی است که در تحقیقات بیزلی (۴)، یورسین (۱۰)، لارسن (۱۰) و تورویانو (۱۸) این مقدار بین ۳-۱۰ تا ۵-۱۰ متغیر است. علاوه بر موارد ذکر شده نتایج حاصل از ارزیابی عملکرد صندوقی در دوره آزمایش، اهمیت بروز رسانی ترکیب آن را منعکس کرد و آن را جزئی جدایی‌ناپذیر از رویکرد سرمایه‌گذاری شاخص محور نشان داد. بررسی صندوق شاخصی با مشخص نمود که گروه‌های محصولات شیمیایی، شرکت‌های چند رشته‌ای، گروه کانه‌های فلزی و فلزات اساسی بیش‌ترین سهم از صندوق را به خود اختصاص داده‌اند، همچنین پنج سهم صندوق، شپدیس، کگل، وامید و کرمانشاه به ترتیب با ۱۷٫۱۸٪، ۱۰٫۶۵٪، ۱۰٫۵۸٪، ۹٫۹۹٪ و ۹٫۳۴٪ بیش‌ترین وزن را در صندوق شاخصی به خود اختصاص داده‌اند. الگوی پیشنهادی در این پژوهش هر دو طیف سرمایه‌گذاران متخصص و مبتدی را مخاطب قرار می‌دهد. الگوریتم فراابتکاری طراحی‌شده به عنوان یک ابزار پشتیبان تصمیم، راهنمای سرمایه‌گذاران متخصص در ایجاد صندوق شاخصی بوده و همچنین با توجه به اینکه منجر به پیاده‌سازی یک استراتژی کم ریسک و با هزینه‌های پایین معاملاتی خواهد شد، بهترین ابزار سرمایه‌گذاری برای افرادی که قصد ورود به بازار سرمایه را داشته و یا از تجربه کافی برخوردار نمی‌باشند، قلمداد می‌شود. پیشنهادهای زیر برای پژوهش‌های آتی ارائه می‌گردد:

- (۱) لحاظ نمودن هزینه‌های معاملاتی و تلاش در جهت حداقل سازی آن و همچنین تعیین دوره بروز رسانی ترکیب صندوق شاخصی به عنوان پیشنهاد مطرح می‌گردد.
- (۲) جهت حل مسئله پژوهش از الگوریتم‌های فرا ابتکاری ژنتیک و تبرید شبیه‌سازی شده بهره گرفته شد، در حالی که مطالعه و بررسی تکنیک‌های متنوعی چون بهینه‌سازی فوجی ذرات و منطق فازی، می‌تواند در پژوهش‌های آتی مدنظر قرار گیرد.
- (۳) رویکرد شاخص محور بنیادی که سرمایه‌گذاری «محوری-جانبی» یکی از نمونه‌های آن است، در کنار رویکرد شاخص محور مورد مطالعه قرار گیرد.

فهرست منابع

- * تقوی فرد، محمدتقی، منصوری، طاها؛ خوش طینت، محسن. ارائه یک الگوریتم فرا ابتکاری جهت انتخاب سبد سهام با در نظر گرفتن محدودیت‌های عدد صحیح، فصلنامه پژوهش‌های اقتصادی، زمستان ۱۳۸۶، سال هفتم، شماره چهارم، صفحات ۴۹ - ۶۹.
- * حنیفی، فرهاد، بحرالعلوم، محمد مهدی، جوادی، بابک. طراحی و تحلیل مقایسه‌ای الگوریتم‌های فرا ابتکاری جهت پیاده‌سازی سرمایه‌گذاری شاخص محور در بورس تهران، چشم‌انداز مدیریت، پاییز ۸۸، شماره ۳۲، صفحات ۸۹-۱۰۸.
- * راعی، رضا، تلنگی، احمد، مدیریت سرمایه گذاری پیشرفته، تهران، انتشارات سمت، ۱۳۸۳. چاپ اول.
- * Beasley, J. E. Meade, N. AND Chang, T. J. (2003), "An Evolutionary Heuristic for the Index Tracking Problem", *European Journal of Operational Research*, 148, 621-643.
- * Bogle, J. C. (1998), "The implications of style analysis for mutual fund performance". *Journal of Portfolio Management*, 24(4), 34-42.
- * Coleman, T. F, Y. Li and J. Henniger. (2006), "Minimizing Tracking Error while Restricting the Number of Assets", *Journal of Risk*, 8, 33-56.
- * Corielli, F. and Marcellino, M. (2006), "Factor Based Index Tracking", *Journal of Banking and Finance*, 30, 2215-2233.
- * E.-g. Talbi (2009), "Metaheuristics : from design to implementation", John Wiley & Sons, New York, NY, USA.
- * Jansen, R. and Dijk, R. Van. (2002), "Optimal Benchmark Tracking with Small Portfolios", *Journal of Portfolio Management*, 28, 33-39.
- * Jeurissen, R. (2005), "A Hybrid Genetic Algorithm to Track the Dutch AEX-index", Bachelor Thesis, Erasmus University Rotterdam.
- * Larsen Jr., G. and Resnick, B. (1998), "Empirical Insights on Indexing", *the Journal of Portfolio Management*, 25 (1), 51-60.
- * Markowitz, H. (1952), Portfolio selection. *Journal of Finance*, 7, 77-91.
- * Markowitz, H. (1959), Portfolio selection: efficient diversification of investments. New York: Wiley.
- * Oh, K.J. Kim, T.Y. and Min, S. (2005), "Using Genetic Algorithm to Support Portfolio Optimization for Index fund Management", *Expert Systems with Applications*, 28, 371-379.
- * Rafaely, B. and Bennell, J. (2006), "Optimisation of FTSE 100 Tracker Funds: A Comparison of Genetic Algorithms and Quadratic Programming", *Managerial Finance*, 32 (6), 477-492.
- * Rohweder, H. (1998), "Implementing Stock Selection Ideas: Does Tracking Error Optimization do any Good", *Portfolio Management*, 24 (3), 49-59.
- * Sharpe, W. F. (1991), "The Arithmetic of Active Management", *Financial Analyst Journal*, 47 (1), 7-9.
- * Torrubiano, R. and Suarez, A. (2008), "A Hybrid Optimization Approach to Index Tracking", *Operational Research*, 166 (1), 57-71.