

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده علوم ریاضی

رشته ریاضی کاربردی، گرایش ریاضی مالی

پایان نامه کارشناسی ارشد

بررسی رده ای از مسایل بهینه سازی سبد سهام با هزینه های معاملاتی

نگارنده: راضیه کی شمس وند

استادان راهنما

دکتر علیرضاناظمی
دکتر سید مجتبی میرلوحی

بهمن ۱۳۹۷



فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم راضیه کی شمس وند با شماره دانشجویی ۹۵۱۲۶۲۴ رشته ریاضی کاربردی گرایش ریاضی مالی تحت بررسی رده ای از مسایل بهینه سازی سبد سهام با هزینه های معاملاتی که در تاریخ ۱۳۹۷/۱۱/۰۲ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

<input type="checkbox"/> قبول (با درجه علمی خوب...)				<input type="checkbox"/> مردود
نوع تحقیق: <input checked="" type="checkbox"/> نظری <input type="checkbox"/> عملی				
عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء	
۱- استاد راهنمای اول	دکتر علیرضا ناظمی	دانشیار		
۲- استاد راهنمای دوم	دکتر سید مجتبی میرلوحی	استادیار		
۳- استاد مشاور				
۴- نماینده تحصیلات تکمیلی	دکتر سمیه مغاری	استادیار		
۵- استاد ممتحن اول	دکتر مهرداد غزنوی	استادیار		
۶- استاد ممتحن دوم	دکتر مجتبی غیائی	استادیار		

نام و نام خانوادگی رئیس دانشکده: دکتر ابراهیم هاشمی

تاریخ و امضاء و مهر دانشکده:



تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۴ ماه برگزار شود).

روردگارا...



نه می توانم مویشان را که در راه عزّت من
سفید شد، سیاه کنم و نه برای دست های
پینه بسته شان که ثمره تلاش برای افتخار
من است، مرهمی دارم. پس توفیقم ده
که هر لحظه شکرگزارشان باشم و ثانیه های
عمرم را در عصای دست بودنشان بگذارم.

تقدیم بہ

پدر و مادر عزیزم

سپاس‌گزاری...

نخستین سپاس و ستایش از آن خداوندی است که بنده کوچکش را در دریای بی‌کران اندیشه، قطره‌ای ساخت تا وسعت آن را از دریچه اندیشه‌های ناب آموزگارانی بزرگ به تماشا نشیند. لذا اکنون که در سایه‌سار بنده نوازی‌هایش پایان‌نامه حاضر به انجام رسیده است، بر خود لازم می‌دانم تا مراتب سپاس را از بزرگوارانی به‌جا آورم که اگر دست یاری‌گرشان نبود، هرگز این پایان‌نامه به انجام نمی‌رسید.

ابتدا از استادان گرانقدرم جناب آقای دکتر علی‌رضاناظمی و جناب آقای دکتر مجتبی میرلوحی که زحمت راهنمایی این پایان‌نامه را برعهده داشتند، کمال سپاس را دارم. سپاس آخر را به مهربانان‌ترین همراهان زندگی‌ام به خانواده‌ام و تمامی دوستانی که در طول انجام این پروژه کمک‌رسان من بوده‌اند تقدیم می‌کنم.

راضیه کی شمس‌وند
بهمن ۱۳۹۷

تعهد نامه

اینجانب راضیه کی شمس وند دانشجوی کارشناسی ارشد رشته ریاضی کاربردی علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان بررسی رده ای از مسایل بهینه سازی سبد سهام با هزینه های معاملاتی، تحت راهنمایی علیرضاناظمی متعهد می شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های دیگر پژوهش گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام “دانشگاه صنعتی شاهرود” یا “Shahrood University of Technology” به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده اند، در مقالات مستخرج از پایان نامه رعایت می گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

راضیه کی شمس وند

بهمن ۱۳۹۷

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی باشد.

چکیده

در این پایان‌نامه یک روش عددی بر مبنای شبکه‌های عصبی برای حل رده‌ای از مسائل بهینه‌سازی مالی با هزینه‌های معاملاتی ارائه می‌دهیم. برای این کار ابتدا شرایط بهینگی را برای مسأله بهینه‌سازی مورد نظر می‌نویسیم. سپس یک مدل شبکه عصبی متناظر با آن طراحی می‌کنیم. ثابت می‌کنیم که نقطه تعادل شبکه عصبی متناظر جواب بهینه مسأله اصلی است. با ارائه یک تابع لیاپانوف مناسب، پایه‌ای و همگرایی سراسری مدل ارائه شده را اثبات می‌کنیم. به این منظور تعاریفی از مفاهیم مالی در ریاضی و مقدماتی از بهینه‌سازی را ارائه می‌دهیم، سپس یک مدل شبکه عصبی گرادیانی جدید را برای حل مسائل بهینه‌سازی درجه دوم ارائه می‌دهیم، در ادامه یک مدل شبکه عصبی جدید برای حل مسائل انتخاب سبد سهام بهینه با فرض قرض گرفتن ارائه می‌دهیم و در پایان به حل مسأله انتخاب سبد سهام مجاز با هزینه‌های معاملاتی با استفاده از شبکه عصبی ارائه شده می‌پردازیم.

کلمات کلیدی: سبد سهام، بهینه‌سازی، هزینه‌های معاملاتی، شبکه‌های عصبی، ارزش در معرض ریسک، ارزش در معرض ریسک مشروط.

فهرست مطالب

ق	فهرست تصاویر
ش	فهرست جداول
۱	۱ مقدمات و مفاهیم اساسی
۱	۱.۱ مقدمه‌ای بر شبکه‌های عصبی
۱	۱.۱.۱ شبکه‌های عصبی بیولوژیکی
۳	۲.۱.۱ شبکه‌های عصبی مصنوعی
۶	۳.۱.۱ تاریخچه تکامل شبکه عصبی
۷	۴.۱.۱ تاریخچه حل مسائل بهینه‌سازی با استفاده از شبکه‌های عصبی
۹	۲.۱ مقدمه‌ای بر بهینه‌سازی غیرخطی
۱۰	۳.۱ سیستم‌های دینامیکی
۱۷	۲ یک مدل شبکه عصبی گرادیانی کارای جدید برای حل مسائل برنامه‌ریزی...
۱۸	۱.۲ مدل شبکه‌ی عصبی پیشنهادی
۲۲	۲.۲ مقایسه با برخی مدل‌های موجود
۲۴	۳.۲ تجزیه و تحلیل پایداری
۲۷	۳ انتخاب سبد سهام بهینه با فرض قرض گرفتن
۲۷	۱.۳ مقدمه
۲۸	۲.۳ مدل سبد سهام کارآمد قابل قبول با قرض گرفتن
۳۱	۳.۳ مثال عددی
۳۳	۱.۳.۳ ترسیم مثال ۳.۳ با شبکه‌ی عصبی
	۴ مسأله انتخاب سبد سهام مجاز با هزینه‌های معاملاتی با یک روش شبکه‌ی عصبی
۳۷	
۳۷	۱.۴ مقدمه

۳۸	مدل انتخاب سبد سهام با هزینه‌های معاملاتی	۲.۴
۴۱	مثال عددی	۳.۴
۴۷	ترسیم مثال ۳.۴ با شبکه‌ی عصبی	۱.۳.۴

فهرست تصاویر

۲	۱.۱	ساختار یک نرون.
۳	۲.۱	مدل ریاضی یک نرون.
۴	۳.۱	ساختار یک نرون با لایه ورودی، لایه پنهان و لایه خروجی.
۶	۴.۱	شبکه‌های عصبی دو لایه بازگشتی.
۱۱	۵.۱	پایداری لیاپانوف.
۱۲	۶.۱	پایداری مجانبی.
۲۲	۱.۲	یک نمودار بلوکی ساده برای شبکه عصبی (۱۴.۲) و (۱۵.۲)
۳۴	۱.۳	مثال مدل (۱۱.۳) با شبکه عصبی وقتی که $\mu = 1$
۳۵	۲.۳	مثال مدل (۱۲.۳) با شبکه عصبی وقتی که $\mu = 1$
		۱.۴	نمودار بالایی و پایینی مرزهای کارا با در نظر گرفتن خطای قابل قبول و
۴۵		هزینه‌های معاملاتی و محدودیت‌های کراندار
۴۶	۲.۴	مرزهای کارا بالایی تحت محدودیت‌های مختلف
۴۷	۳.۴	مرزهای کارا پایینی تحت محدودیت‌های مختلف
۴۸	۴.۴	مثال مدل (۹.۴) با شبکه عصبی وقتی که $r_0 = 0.05$
۴۹	۵.۴	مثال مدل (۱۰.۴) با شبکه عصبی وقتی که $r_0 = 0.05$
۵۰	۶.۴	مثال مدل (۹.۴) با شبکه عصبی وقتی که $r_0 = 0.1$
۵۱	۷.۴	مثال مدل (۱۰.۴) با شبکه عصبی وقتی که $r_0 = 0.1$
۵۲	۸.۴	مثال مدل (۹.۴) با شبکه عصبی وقتی که $r_0 = 0.15$
۵۳	۹.۴	مثال مدل (۱۰.۴) با شبکه عصبی وقتی که $r_0 = 0.15$

فهرست جداول

۵	توابع فعال سازی	۱.۱
۳۲	بالاترین سبد سهام کارایی مجاز	۱.۳
۴۳	بالاترین سبد سهام کارایی مجاز	۱.۴
۴۴	پایین ترین سبد سهام کارایی مجاز	۲.۴

فصل ۱

مقدمات و مفاهیم اساسی

۱.۱ مقدمه‌ای بر شبکه‌های عصبی

۱.۱.۱ شبکه‌های عصبی بیولوژیکی

واحد اصلی دستگاه عصبی، سلولی خاص به نام نرون^۱ می‌باشد و بدون شک نحوه کار مغز و راز شعور آدمی در آن نهفته است. نرون‌ها با وجود تفاوت‌های زیاد از نظر اندازه و شکل ظاهری، مشخصه‌های مشترکی دارند. همان‌طور که در شکل (۱.۱) نشان داده شده است از تنه سلولی نرون تعدادی شاخک کوتاه خارج می‌شود که دندریت^۲ نام دارند. دندریت‌ها و تنه سلولی، سیگنال‌ها را از نرون‌های مجاور دریافت می‌کنند و از طریق یک لوله باریک به نام آکسون^۳ به نرون‌های دیگر منتقل می‌گردد. آکسون در انتهای خود به تعدادی رشته جانبی باریک تقسیم می‌گردد که پایانه‌های آکسونی نام دارد و با دندریت‌های سایر نرون‌ها مرتبط است. ارتباط بین آکسون یک نرون و دندریت نرون دیگر را سیناپس^۴ می‌نامند. یک سیناپس

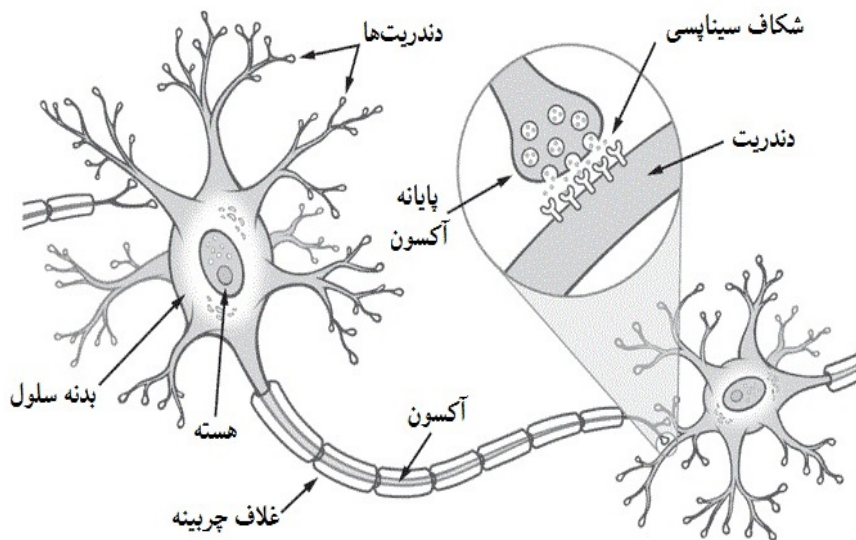
¹Neuron

²Dendrite

³Axon

⁴Synapse

مرکب از پایانه قبل از سیناپسی^۵، شکاف سیناپسی^۶ و پایانه بعد از سیناپسی^۷ می‌باشد. همه سیگنال‌های جمع‌بندی شده، در تنه نرون ترکیب می‌شوند و اگر وسعت سیگنال‌های ترکیب شده به آستانه نرون برسد، مرحله تحریک شدن^۸ فعال می‌شود و یک سیگنال خروجی تولید می‌شود. این سیگنال به صورت یک پالس منفرد یا بخشی از پالس‌ها در یک میزان خاص به موازات آکسون به پایانه‌های سیناپسی انتقال می‌یابد و موجب ترشح موادی به نام عصب-رسانه می‌شود. عصب-رسانه در داخل شکاف سیناپسی پخش می‌شود و نرون بعدی را تحریک می‌کند و به این ترتیب یک سیگنال از یک نرون به دیگری انتقال می‌یابد. تعداد بسیار زیادی آکسون از نرون‌های مختلف با دندریته‌های یک نرون به این صورت ارتباط برقرار می‌کنند.



شکل ۱.۱: ساختار یک نرون.

نرون‌ها با توجه به عملکردشان به سه دسته تقسیم می‌شوند:

(۱) **نرون‌های حسی:** این نرون‌ها پیام‌ها را از گیرنده‌ها به دستگاه عصبی مرکزی منتقل می‌کنند. گیرنده سلول خاصی است که در اندام‌های حسی، عضلات، پوست و ... قرار دارد و تغییرات فیزیکی یا شیمیایی را در می‌یابد، این تغییرات را به سیگنال‌های عصبی تبدیل می‌کند.

(۲) **نرون‌های حرکتی:** این نرون‌ها پیام‌ها را از مغز یا طناب نخاعی به اندام‌های عمل کننده می‌رسانند.

(۳) **نرون‌های رابط:** این نرون‌ها پیام‌ها را از نرون‌های حسی می‌گیرند و به یک نرون دیگر

⁵Presynaptic Terminal

⁶Synaptic Cleft

⁷Postsynaptic Terminal

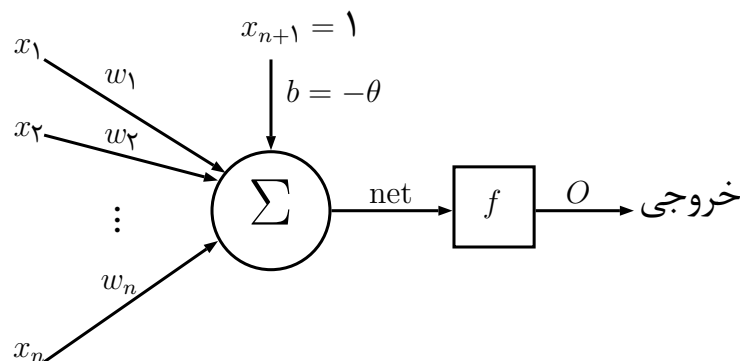
⁸Firing

انتقال می‌دهند و یا پیام‌ها را به یک نرون حرکتی می‌رسانند. نرون‌های رابط فقط در مغز، چشم و طناب نخاعی یافت می‌شوند.

حال منظور از یادگیری عبارت است از تغییری نسبتاً ساختاری در سیناپس (بعد از تحریک)، که این تغییر ساختاری باعث افزایش کارایی سیناپس می‌شود.

۲.۱.۱ شبکه‌های عصبی مصنوعی

بعد از اشاره‌ای گذرا بر شبکه عصبی بیولوژیکی و نشان دادن گوشه‌ای بسیار کوچک از این جهان مرموز و فوق‌العاده پیچیده به موضوع اصلی یعنی شبکه‌های عصبی مصنوعی می‌پردازیم. شبکه‌های عصبی مصنوعی^۹ مدل‌های سخت افزاری یا نرم افزاری الهام گرفته از ساختار و رفتار نرون‌های بیولوژیکی و سیستم‌های عصبی انسان هستند که شامل عناصر پردازشی (به نام نرون‌ها) و ارتباطات میان آنها با ضرایب (وزن‌ها) اختصاص یافته به ارتباطات است، که ساختار شبکه عصبی را تشکیل می‌دهند. در این پایان‌نامه ما شبکه‌های عصبی مصنوعی را یک مدل ریاضی فرض خواهیم کرد و به جای استفاده از عبارت شبکه‌های عصبی مصنوعی به‌طور خلاصه عبارت شبکه‌های عصبی را به کار خواهیم برد.



شکل ۲.۱: مدل ریاضی یک نرون.

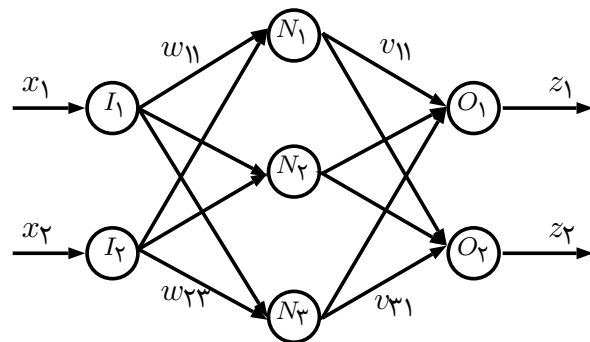
مدل شبیه‌سازی شده یک نرون، در شکل (۲.۱) نمایش داده شده است. یک نرون در حالت کلی یک مجموعه از $n+1$ ورودی x_j ($j = 1, \dots, n+1$) دارد. این ورودی‌ها همان سیگنال‌های عصبی می‌باشند که از محیط یا از یک نرون دیگر فرستاده می‌شوند. w_j ($j = 1, \dots, n+1$) که روی آکسون‌ها واقع هستند نقش ساختار سیناپسی را ایفا می‌کنند. این‌ها هستند که باید تغییر کنند تا فرآیند یادگیری اتفاق افتد حال این‌که این تغییر بر چه قاعده‌ای استوار باشد بستگی به الگوریتم آموزش یا تعلیم دارد. وزن آخرین ورودی یعنی $-\theta$ را با b نشان می‌دهند و آن‌را بایاس می‌نامند. سیگنال‌های عصبی x_1 و x_2 و ... و x_{n+1} به ترتیب در w_1 و w_2 و ... و

⁹Artificial Neural Networks

w_{n+1} ضرب شده و مجموع این حاصل ضربها که در شکل آن را با net نشان می‌دهیم، وارد نرون می‌شوند. در این مرحله نرون همانند یک تابع روی net عمل می‌کند و حاصل کار خود (برد تابع) را از طریق آکسون‌های طرف راست به نرون‌های دیگر می‌فرستد. این تابع را تابع فعال‌سازی^{۱۰} می‌نامند.

البته خاطر نشان می‌کنیم که اگر نرون به‌عنوان تابع فرض شود این تابع از \mathbb{R}^n به \mathbb{R} تعریف می‌گردد. یعنی سیگنال خروجی از نرون در تمام آکسون‌های طرف راست یکسان می‌باشد که روی شکل با O نمایش داده‌ایم. همان‌طور که قبلاً اشاره شد نرون‌ها می‌توانند حسی، ارتباطی و یا حرکتی باشند. اطلاعات از طریق نرون‌های حسی دریافت می‌شود و از طریق نرون‌های ارتباطی انتقال می‌یابد و بالاخره نتیجه کار شبکه عصبی و یا واکنش شبکه عصبی به اطلاعات دریافت شده از طریق نرون‌های حرکتی به خارج منتقل می‌گردد شکل (۳.۱).

نرون‌های حسی را لایه ورودی، نرون‌های حرکتی را لایه خروجی و نرون‌های ارتباطی را لایه پنهان نیز می‌گویند. لایه مخفی می‌تواند از چندین لایه تشکیل گردد البته در شکل (۳.۱) لایه مخفی را تک لایه در نظر گرفته است.



حرکتی نرون‌های ارتباطی نرون‌های حسی نرون‌های

شکل ۳.۱: ساختار یک نرون با لایه ورودی، لایه پنهان و لایه خروجی.

مشخصه‌های اصلی یک شبکه عصبی که در طراحی آن باید مدنظر قرار گیرند عبارت‌اند از:

معماری شبکه عصبی: نحوه اتصالات بین نرون‌ها، تعداد آن‌ها و تعداد لایه‌های تشکیل دهنده بخش نرون‌های ارتباطی را معماری شبکه عصبی می‌گوییم.

تابع‌های فعال‌سازی: این که چه تابعی روی نرون قرار بگیرد تا بر ورودی‌های نرون اثر کرده و خروجی نرون را تولید کند تابع فعال‌سازی آن نرون گویند. در جدول (۱.۱) تعدادی از مهم‌ترین توابع فعال‌سازی نشان داده شده است.

الگوریتم آموزش: روشی که وزن‌های روی آکسون‌ها بر اساس آن تغییر می‌کنند (w_{ij} ها و v_{jk} ها در شکل ۳.۱) تا خروجی شبکه عصبی به حالت مطلوب درآید را الگوریتم آموزش می‌گوییم.

¹⁰Activation Function

جدول ۱.۱: توابع فعال‌سازی

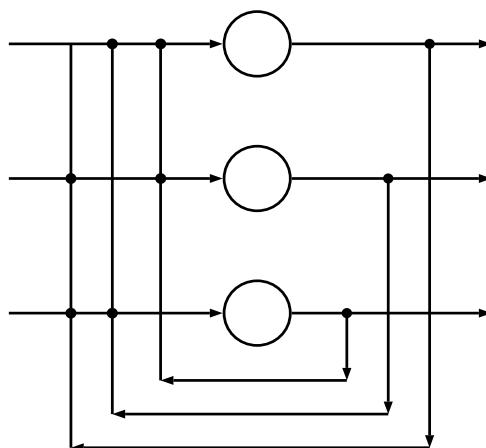
شکل تابع	تعریف تابع	نام	ردیف
	$f(x) = \begin{cases} 0 & x < 0, \\ 1 & x \geq 0. \end{cases}$	آستانه‌ای دو مقداره	۱
	$f(x) = \begin{cases} -1 & x < 0, \\ 1 & x \geq 0. \end{cases}$	آستانه‌ای دو مقداره متقارن	۲
	$f(x) = x$	خطی	۳
	$f(x) = \begin{cases} -1, & x < -1, \\ x, & -1 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$	آستانه‌ای خطی متقارن	۴
	$f(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$	آستانه‌ای خطی	۵
	$f(x) = \frac{1}{1+e^{-x}}$	سیگموئید	۶

با توجه به مطالب قبل می‌توان برای شبکه‌های عصبی مصنوعی خصوصیات عمومی زیر را نسبت داد:

۱. پردازش اطلاعات در عناصر بسیار ساده به نام نرون (سلول عصبی) انجام می‌گیرد.
۲. سیگنال‌های عصبی بین سلول‌های عصبی از طریق اتصالات بین آن‌ها مبادله می‌شوند.
۳. به هر خط اتصال (آکسون) یک وزن نسبت داده می‌شود که در شبکه‌های عصبی معمولی این وزن‌ها در سیگنال‌های منتقل شده ضرب می‌گردند.
۴. هر نرون دارای یک تابع فعال‌سازی می‌باشد که این تابع در اکثر موارد غیرخطی است.

شبکه‌های عصبی بازگشتی

در شبکه‌های عصبی بازگشتی^{۱۱}، حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرونهاي همان لایه و یا لایه قبل وجود دارد. چنین شبکه‌ای حافظه‌ای را نگه می‌دارد و حالت بعد نه تنها به سیگنالهای ورودی بلکه به حالات قبل شبکه نیز وابسته است. این شبکه‌های عصبی معمولاً در مدل ریاضی به صورت معادلات دیفرانسیل مرتبه اول ظاهر می‌شوند.



شکل ۴.۱: شبکه‌های عصبی دو لایه بازگشتی.

۳.۱.۱ تاریخچه تکامل شبکه عصبی

در سال ۱۹۴۳ یک نرولوژیست به نام وارن مک کلاچ^{۱۲} به همراه یک متخصص آمار به نام والتر پیتز^{۱۳} اولین مدل ریاضی نرون را ارائه دادند [۷۷]. این نرون ساده یک عنصر محاسبه‌گر بود که عناصری به نام ورودی را در مقادیر ثابتی به نام وزن ضرب می‌کرد و آنها را از یک عملگر خطی عبور می‌داد و حاصل خروجی نرون را تشکیل می‌داد. در این مدل وزن‌ها مقادیر ثابتی بودند. در سال ۱۹۴۹ دونالد هب^{۱۴} [۱۸] کتابی درباره نحوه یادگیری در مغز انسان نوشت. بر این مبنا قانون یادگیری برای نرون را ارائه داد که به قانون هب معروف است و بعدها در تعلیم شبکه‌های عصبی از آن استفاده شد. در سال ۱۹۵۴ فارلی و کلارک^{۱۵} [۱] مدلی از نرون‌ها را به‌طور تصادفی به هم وصل کردند و قانون هب را پیاده‌سازی نمودند و نشان دادند که می‌توان دو الگوی ورودی را با این قانون تعلیم از هم تشخیص داد. با پیشرفت علم کامپیوتر اولین مدل از نرون عصبی مصنوعی توسط ناتانیال روچستر^{۱۶} محقق شرکت IBM در سال ۱۹۵۸ با

¹¹Feedback (Recurrent) Neural Network

¹²Warren McCulloch

¹³Walter Pitts

¹⁴Donald Hebb

¹⁵Farley and Clark

¹⁶Nathanial Rochester

استفاده از کامپیوتر شبیه‌سازی شد. روزنبلات^{۱۷} [۲۴، ۲۵] در سال ۱۹۵۸ نرون ساده مک کلاچ پیتز را اصلاح کرد و به آن قابلیت یادگیری و سازگاری را اضافه کرد و این نرون را پرسپترون نامید. قانون تعلیم پرسپترون^{۱۸} اولین قانون رسمی برای شبکه‌های عصبی است.

برنارد ویدرو^{۱۹} و مارسین هاف^{۲۰} و همکارانشان یک مدل سه سطحی از نرون‌ها را ابداع کردند که آدلاین^{۲۱} نام‌گذاری شد و مختصر شده حروف "عنصر خطی تطبیقی"^{۲۲} بود [۱۱، ۱۲، ۱۳، ۱۴]. و برای آدلاین یک قانون تعلیم بر مبنای مشتق‌گیری بیان نمودند. این قانون تعلیم به $\alpha - LMS$ معروف است. از متصل کردن چند آدلاین به یکدیگر مادالاین^{۲۳} ساخته شد. ویدرو و هاف از شبکه عصبی مادالاین برای حذف اکو از خطوط تلفن استفاده نمودند.

سال ۱۹۶۹ آغاز یک دوره رکود برای شبکه عصبی بود. پاپرت و مینسکی^{۲۴} [۵۵] کتابی به نام پرسپترون نوشتند و نشان دادند که اگرچه گیت‌های OR و AND را می‌توان با استفاده از پرسپترون پیاده‌سازی کرد ولی گیت XOR قابل پیاده‌سازی نمی‌باشد و لذا نمی‌توان از شبکه عصبی برای پردازش اطلاعات استفاده کرد. الگوریتم پس انتشار^{۲۵} پال ورباس^{۲۶} [۶۰] در سال ۱۹۷۴ ارائه شد و بعدها توسط رامل هارت^{۲۷} [۱۹] به‌طور مستقل کشف گردید. این الگوریتم از زمان پیدایش به‌طور گسترده به‌عنوان یک الگوریتم آموزش در شبکه‌های عصبی پیش‌خورد^{۲۸} مورد استفاده قرار گرفته است. سال ۱۹۸۲ تا ۱۹۸۶ را می‌توان تولد دوباره شبکه عصبی دانست. در این سال‌ها دو اتفاق مهم در این زمینه رخ داد. اول ارائه شبکه عصبی بازگشتی و مفهوم تابع انرژی توسط جان هاپفیلد^{۲۹} [۳۸] بود که در سال ۱۹۸۲ منتشر شد. دومین اتفاق، حل مسأله فروشنده دوره‌گرد (TSP) توسط شبکه هاپفیلد بود [۳۹]. هاپفیلد نشان داد که از این شبکه بازگشتی می‌توان برای حل مسائل بهینه‌سازی استفاده کرد و افق جدیدی را در شبکه‌های عصبی مصنوعی گشود.

۴.۱.۱ تاریخچه حل مسائل بهینه‌سازی با استفاده از شبکه‌های عصبی

طیف گسترده‌ای از مسائل مهم در رشته‌های علوم پایه و مهندسی از جمله کنترل بهینه، طراحی ساختمان، پردازش تصویر، تقریب توابع، تحلیل رگرسیون و... قابل تبدیل به مسائل

¹⁷Frank Rosenblatt

¹⁸Perceptron

¹⁹Bernard Widrow

²⁰Marcian Hoff

²¹Adeline

²²Adaptive Linear Elements

²³Madaline

²⁴Papert and Minsky

²⁵Back Propagation

²⁶Paul Werbos

²⁷Rumelhart

²⁸Feed Forward

²⁹John Hopfield

بهینه‌سازی خطی و غیرخطی هستند. در دهه‌های گذشته الگوریتم‌های بهینه‌سازی عددی فراوانی برای حل مسائل بهینه‌سازی خطی و غیرخطی ارائه شده‌اند. از آن‌جا که زمان مورد نیاز برای حل مسائل بهینه‌سازی و به‌خصوص مسائل بهینه‌سازی غیرخطی بسیار وابسته به بعد و ساختار مسأله است بنابراین الگوریتم‌های عددی کارایی کمی از خود نشان می‌دهند. یک رهیافت امیدوار کننده برای حل مسائل بهینه‌سازی با ابعاد بالا استفاده از شبکه‌های عصبی مصنوعی می‌باشد.

در سال ۱۹۸۶ تانک^{۳۰} و هاپفیلد [۷۰] اولین شبکه عصبی را برای حل مسائل برنامه‌ریزی خطی معرفی کردند و نشان دادند که وضعیت شبکه در هر تکرار به‌گونه‌ای تغییر می‌کند که تابع انرژی متناظر با آن به‌طور یکنواخت کاهش می‌یابد تا جایی که به نقطه مینیمم خود می‌رسد و این نقطه مینیمم متناظر با نقطه تعادل شبکه عصبی می‌باشد. آن‌ها این شبکه عصبی را توسط یک مدار الکتریکی پیاده‌سازی کردند. هم‌چنین از این شبکه برای حل مسأله فروشنده دوره‌گرد با ۳۰ شهر استفاده نمودند. این شبکه دارای نقص‌هایی بود به‌خصوص این که نقطه تعادل شبکه در شرایط KKT^{۳۱} صدق نمی‌کرد و لذا جواب مطلوبی از مسأله حاصل نمی‌شد. با این وجود کارهای هاپفیلد انگیزه بسیار خوبی را برای محققین به‌وجود آورد تا در این زمینه فعالیت کنند. کندی و چاو^{۳۲} [۴۵] با افزودن یک پارامتر جریمه متناهی به شبکه هاپفیلد آن را توسعه دادند و از آن برای حل مسأله برنامه‌ریزی غیرخطی محدب استفاده نمودند. با پارامترهای جریمه این شبکه در یافتن نقطه بهینه دقیق ناتوان است به‌خصوص اگر پارامتر جریمه بزرگ باشد این شبکه به‌سختی عمل می‌کند.

برای جلوگیری از بکار بردن پارامتر جریمه یک شبکه عصبی سوئیچ-خازن^{۳۳} توسط رودریگز-وازکز^{۳۴} و همکاران [۶۳] معرفی شد. هم‌چنین ما و شانبلات^{۳۵} [۱۶] یک شبکه عصبی دو فازیه ارائه کردند که فاز اول شبکه مشابه شبکه کندی و چاو بود و در فاز دوم آن مسیر شبکه به جواب دقیق مسأله همگرا می‌شد بنابراین این روش جواب‌های دقیق‌تری نسبت به شبکه کندی و چاو ارائه می‌داد. مشکل این شبکه در این بود که پایداری فاز دوم شبکه بستگی به انتخاب یک مقدار بزرگ از پارامتر جریمه داشت و لذا اگرچه تأثیر پارامتر جریمه در این روش کاهش یافته بود و جواب‌های حاصل دقیق‌تر بودند اما هنوز مستقل از پارامتر نبود. ژانگ و کنستانتینیدس^{۳۶} [۶۷] بر مبنای روش لاگرانژ یک شبکه عصبی ارائه کردند که کاملاً مستقل از پارامتر جریمه و قادر به حل مسائل غیرخطی بود. نقطه تعادل این شبکه عصبی در شرایط بهینگی مرتبه اول و دوم صدق می‌کرد و هم‌چنین شبکه حاصل همگرا بود.

³⁰Tank

³¹Karush-Kuhn-Tucker

³²Kennedy and Chua

³³Switched-Capacitor

³⁴Rodriguez-Vazquez

³⁵Maa and Shanblatt

³⁶Zhang and Constantinides

در سال ۱۹۹۳ بوزردوم و پتیسین^{۳۷} [۴] شبکه‌ای را برمبنای گرادیان و روش تصویر و مستقل از پارامتر جریمه ابداع کردند که تنها قادر به حل مسائل درجه دوم با متغیرهای کران دار بود. این روش در عمل روش کارایی بود اما نمی‌توانست مسائل کلی برنامه‌ریزی خطی و درجه دوم را حل کند.

ژیا^{۳۸} و همکاران [۸۲، ۸۳، ۸۶، ۸۷] چندین مدل را برای حل مسائل بهینه‌سازی غیرخطی با قیود خطی و غیرخطی ارائه کردند. همچنین چندین مدل شبکه عصبی برای حل مسائل بهینه‌سازی غیرخطی توسط عفتی^{۳۹} و همکاران [۲، ۳، ۲۱، ۵۶، ۶۸، ۶۹] ارائه شده است.

۲.۱ مقدمه‌ای بر بهینه‌سازی غیرخطی

تعریف ۱.۲.۱ ([۵۸]). تابع $f : \mathbb{R}^n \rightarrow \mathbb{R}$ روی مجموعه محدب و ناتهی $S \subseteq \mathbb{R}^n$ محدب نامیده می‌شود اگر برای هر $x, y \in S$ و $\alpha \in (0, 1)$ داشته باشیم

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

تابع f روی S اکیداً محدب نامیده می‌شود اگر نامساوی بالا زمانی که $x \neq y$ است، به‌طور اکید برقرار باشد.

یک مسأله برنامه‌ریزی غیرخطی در حالت کلی به‌صورت زیر در نظر گرفته می‌شود:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i = 1, \dots, m, \\ & h_j(x) = 0, \quad j = 1, \dots, p, \\ & x \in X \subset \mathbb{R}^n, \end{aligned} \tag{۱.۱}$$

که $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ، $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ($i = 1, \dots, m$) و $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$ ($j = 1, \dots, p$) توابع دو بار مشتق پذیر پیوسته هستند. برای تسهیل در نمادگذاری، توابع برداری $g = (g_1, g_2, \dots, g_m)$ و $h = (h_1, h_2, \dots, h_p)$ را معرفی می‌کنیم و (۱.۱) را به‌صورت زیر می‌نویسیم:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & \mathbf{g}(x) \leq 0, \\ & \mathbf{h}(x) = 0, \\ & x \in X \subset \mathbb{R}^n, \end{aligned} \tag{۲.۱}$$

³⁷Bouzerdoum and Pattison

³⁸Xia

³⁹Effati

که محدودیت‌های $g(x) \leq 0$ و $h(x) = 0$ ، محدودیت‌های تابعی و محدودیت $x \in X \subset \mathbb{R}^n$ ، محدودیت مجموعه‌ای نامیده می‌شود. هرگاه توابع f و g روی X محدب باشند و h آفین باشد یعنی $h = Ax - b$ ، $A \in \mathbb{R}^{p \times n}$ ، $rank(A) = p$ ($0 \leq p < n$) و $b \in \mathbb{R}^p$ باشد، آن‌گاه (۲.۱) یک مسأله بهینه‌سازی محدب^{۴۰} و در غیر این صورت یک مسأله بهینه‌سازی نامحدب^{۴۱} نامیده می‌شود.

قضیه ۱.۲.۱ (شرایط کافی مرتبه دوم^{۴۲} [۱۷]). فرض کنید x^* یک نقطه شدنی برای مسأله (۲.۱) باشد. اگر $\mu \in \mathbb{R}^m$ و $\lambda \in \mathbb{R}^p$ وجود داشته باشند به طوری که

$$\mu \geq 0, \quad (3.1)$$

$$\mu^T g(x^*) = 0, \quad (4.1)$$

$$\nabla f(x^*) + \lambda^T \nabla h(x^*) + \mu^T \nabla g(x^*) = 0, \quad (5.1)$$

و ماتریس هسین^{۴۳}

$$\nabla^2 L(x^*, \lambda, \mu) = \nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i \nabla^2 g_i(x^*) + \sum_{j=1}^p \mu_j \nabla^2 h_j(x^*) \quad (6.1)$$

بر زیرفضای

$$M(x^*) = \{d \mid \nabla h(x^*)d = 0, \nabla g_j(x^*)d = 0, \forall j \in J(x^*)\},$$

که در آن

$$J(x^*) = \{j \mid g_j(x^*) = 0, \mu_j > 0\},$$

معین مثبت باشد آن‌گاه x^* یک نقطه مینیمم محلی اکید^{۴۴} برای مسأله (۲.۱) است.

۳.۱ سیستم‌های دینامیکی

تعریف ۱.۳.۱ ([۲]). سیستم دینامیکی زیر را در نظر بگیرید:

$$\dot{x} = f(x(t)), \quad x(t_0) = x_0 \in \mathbb{R}^n, \quad (7.1)$$

که در آن f یک تابع از \mathbb{R}^n به \mathbb{R}^n می‌باشد. x^* یک نقطه تعادل^{۴۵} (۷.۱) نامیده می‌شود اگر $f(x^*) = 0$.

⁴⁰Convex Optimization

⁴¹Nonconvex Optimization

⁴²Second-Order Sufficient Conditions

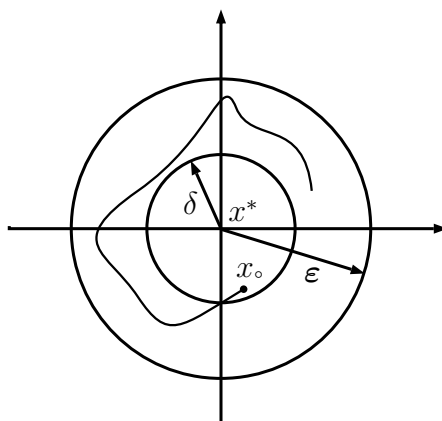
⁴³Hessian Matrix

⁴⁴Strict Local Minimum Point

⁴⁵Equilibrium Point

تعریف ۲.۳.۱ (پایداری به مفهوم لیاپانوف^{۴۶} [۲]). فرض می‌کنیم که $x(t)$ یک جواب (۷.۱) باشد، نقطه تعادل تنها x^* ، پایدار به مفهوم لیاپانوف است اگر برای هر $x_0 = x(t_0)$ و هر $\varepsilon > 0$ ، یک $\delta > 0$ وجود داشته باشد به طوری که اگر $\|x(t_0) - x^*\| < \delta$ آن گاه

$$\|x(t) - x^*\| < \varepsilon, \forall t \geq t_0.$$



شکل ۵.۱: پایداری لیاپانوف.

تعریف ۳.۳.۱ [۲]. سیستم دینامیکی (۷.۱) همگرای سراسری^{۴۷} به مجموعه جواب‌های (۷.۱) که با Ω^* نمایش داده می‌شود، گفته می‌شود اگر برای هر نقطه آغازین دلخواه، مسیر $x(t)$ در رابطه زیر صدق کند:

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), \Omega^*) = 0,$$

که در آن

$$\text{dist}(x, \Omega^*) = \inf_{y \in \Omega^*} \|x - y\|,$$

و $\|\cdot\|$ نرم l_2 را نشان می‌دهد.

تعریف ۴.۳.۱ [۲]. سیستم دینامیکی (۷.۱) در نقطه تعادل x^* که یکتا می‌باشد، پایدار مجانبی سراسری^{۴۸} نامیده می‌شود اگر x^* به مفهوم لیاپانوف پایدار باشد و در شرط زیر صدق کند:

$$\lim_{t \rightarrow \infty} x(t) = x^*.$$

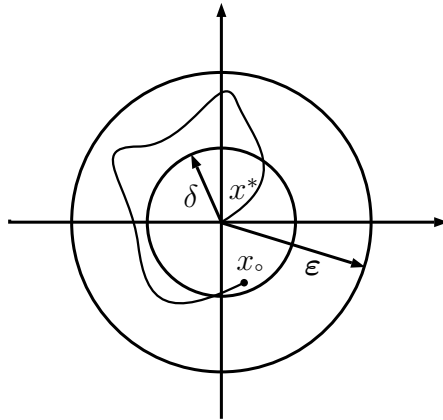
تعریف ۵.۳.۱ [۲]. مجموعه $M \subseteq \mathbb{R}^n$ یک مجموعه تغییر ناپذیر^{۴۹} نسبت به سیستم (۷.۱) گفته می‌شود اگر $x(t_0) \in M$ و $t_0 \geq 0$ ، به ازای هر $t \geq t_0$ $x(t) \in M$ باشد.

⁴⁶Stability in the Sense of Lyapunov

⁴⁷Globally Convergent

⁴⁸Globally Asymptotically Stable

⁴⁹Invariant Set



شکل ۶.۱: پایداری مجانبی.

تعریف ۶.۳.۱ ([۲]). فرض می‌کنیم که $\Omega \subset \mathbb{R}^l$ یک مجموعه بسته و محدب باشد که ممکن است بیکران باشد. $P_\Omega(x): \mathbb{R}^l \rightarrow \Omega$ یک عملگر تصویر^{۵۰} نامیده می‌شود و به صورت زیر تعریف می‌شود:

$$P_\Omega(x) = \arg \min_{v \in \Omega} \|x - v\|, \quad (۸.۱)$$

که $\|\cdot\|$ نرم l_2 را نشان می‌دهد.

تعریف ۷.۳.۱ ([۲]). تابع $F: E \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ روی E در شرط لیپ شیتز صدق می‌کند (پیوسته لیپ شیتز^{۵۱} است) اگر ثابت L وجود داشته باشد به طوری که برای هر دو نقطه $x, y \in E$ داشته باشیم:

$$\|F(x) - F(y)\| \leq L \|x - y\|.$$

تابع F پیوسته لیپ شیتز محلی^{۵۲} روی E نامیده می‌شود اگر برای نقطه $x_0 \in E$ ، یک $\varepsilon -$ همسایگی مانند $N_\varepsilon(x_0) \subset E$ و یک ثابت L_0 وجود داشته باشد به طوری که برای هر دو نقطه $x, y \in N_\varepsilon(x_0)$ داشته باشیم:

$$\|F(x) - F(y)\| \leq L_0 \|x - y\|.$$

قضیه ۱.۳.۱ ([۲]). تابع $F: E \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ را در نظر بگیرید. اگر F یک بار به طور پیوسته مشتق پذیر باشد یعنی $F \in C^1(E)$ ، آن گاه F روی E لیپ شیتز محلی است.

تعریف ۸.۳.۱ ([۲]). نگاشت $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

($\bar{\cdot}$) یکنوا نامیده می‌شود اگر برای هر دو نقطه $x, y \in \mathbb{R}^n$ داشته باشیم:

$$\langle x - y, F(x) - F(y) \rangle \geq 0.$$

⁵⁰Projection Operator

⁵¹Lipschitz Continuous

⁵²Locally Lipschitz Continuous

ب) اکیداً یکنوا نامیده می‌شود اگر برای هر دو نقطه متمایز $x, y \in \mathbb{R}^n$ داشته باشیم:

$$\langle x - y, F(x) - F(y) \rangle > 0.$$

که $\langle \cdot, \cdot \rangle$ ضرب داخلی را نشان می‌دهد.

تعریف ۹.۳.۱ ([۲]). اگر $g: \Omega_1 \subset \mathbb{R}^l \rightarrow \mathbb{R}$ ، آن‌گاه هر مجموعه غیر تهی به شکل زیر

$$L(r) = \{u \in \Omega_1 \mid g(u) \leq r, r \in \mathbb{R}\}$$

یک مجموعه سطح^{۵۳} از g نامیده می‌شود.

تعریف ۱۰.۳.۱ ([۲]). یک ماتریس $n \times n$ ، $M(x)$ که عناصر آن m_{ij} و $i = 1, \dots, n$ و $j = 1, \dots, n$ توابعی هستند که روی مجموعه $S \subset \mathbb{R}^n$ تعریف شده‌اند، روی S نیمه معین مثبت^{۵۴} نامیده می‌شود اگر

$$v^T M(x) v \geq 0, \forall v \in \mathbb{R}^n, \forall x \in S.$$

تعریف ۱۱.۳.۱ ([۲]). فرض می‌کنیم $V: \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع به‌طور پیوسته مشتق پذیر باشد. $V(x)$ به‌طور شعاعی بی‌کران گفته می‌شود اگر

$$\|x\| \rightarrow \infty \Rightarrow V(x) \rightarrow \infty.$$

لم ۱.۳.۱ ([۴۰]). فرض کنیم که $g: \Omega_1 \subset \mathbb{R}^l \rightarrow \mathbb{R}$ باشد که در آن Ω_1 بی‌کران است. آن‌گاه همه مجموعه‌های سطح g کراندار می‌باشند اگر و فقط اگر

$$\lim_{k \rightarrow \infty} g(u^k) = +\infty,$$

که در آن $\{u^k\} \subset \Omega_1$ و $\lim_{k \rightarrow \infty} \|u^k\| = +\infty$.

لم ۲.۳.۱ ([۴۰]). فرض کنیم که $g: \Omega_1 \subset \mathbb{R}^l \rightarrow \mathbb{R}$ روی مجموعه بسته Ω_1 پیوسته باشد آن‌گاه همه مجموعه‌های سطح g کراندار می‌باشند اگر و فقط اگر مجموعه مینیمم‌کننده‌های g غیرتهی و کراندار باشند.

قضیه ۲.۳.۱ ([۴۰]). فرض کنیم که تابع $F: K \subset \mathbb{R}^l \rightarrow \mathbb{R}^l$ روی K به‌طور پیوسته مشتق پذیر باشد، آن‌گاه F روی K یکنوا (اکیداً) است اگر و فقط اگر ماتریس ژاکوبین آن $\nabla F(x)$ برای هر $x \in K$ نیمه‌معین مثبت (معین مثبت) باشد.

قضیه ۳.۳.۱ ([۳۹]). در سیستم دینامیکی (۷.۱) فرض کنیم که f یک تابع پیوسته از \mathbb{R}^n به \mathbb{R}^n باشد، آن‌گاه برای هر $t_0 \geq 0$ و $x_0 \in \mathbb{R}^n$ ، یک جواب محلی $x(t)$ به ازای $t \in [t_0, \tau)$ که $\tau > t_0$ وجود دارد. علاوه بر این اگر f در x_0 پیوسته لیپ‌شیتز محلی باشد آن‌گاه جواب یکتاست و اگر f در \mathbb{R}^n پیوسته لیپ‌شیتز باشد آن‌گاه τ می‌تواند تا $+\infty$ توسعه داده شود.

⁵³Level Set

⁵⁴Positive Semidefinite

تعریف ۱۲.۳.۱ ([۲]). اگر یک جواب محلی تعریف شده در بازه $[t_0, \tau)$ نتواند به یک جواب محلی روی یک بازه بزرگتر $[t_0, \tau_1)$ که $\tau_1 > t_0$ می باشد گسترش یابد، آن گاه یک جواب ماکسیمال نامیده می شود و بازه $[t_0, \tau)$ ، بازه ماکسیمال وجود جواب نامیده می شود. بازه ماکسیمال وجود جواب مربوط به x_0 اغلب به صورت $[t_0, \tau(x_0))$ تعریف می شود.

قضیه ۴.۳.۱ ([۴۳]). فرض کنیم $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک تابع پیوسته باشد. اگر $t \in [t_0, \tau(x_0))$ ، $x(t)$ یک جواب ماکسیمال باشد و $\tau(x_0) < +\infty$ آن گاه

$$\lim_{t \rightarrow \tau(x_0)} \|x(t)\| = +\infty.$$

قضیه ۵.۳.۱ (قضیه پایداری لیاپانوف [۳۰]). فرض کنید که $x = 0$ یک نقطه تعادل سیستم (۷.۱) باشد و $V: \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع به طور پیوسته مشتق پذیر باشد به طوری که:

$$1. \quad V(0) = 0.$$

$$2. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad V(x) > 0.$$

$$3. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad \dot{V}(x) \leq 0.$$

آن گاه $x = 0$ نقطه پایداری سیستم خواهد بود.

قضیه ۶.۳.۱ (قضیه پایداری مجانبی [۳۰]). تحت شرایط قضیه ۵.۳.۱ اگر $V(\cdot)$ در شرایط زیر صدق کند:

$$1. \quad V(0) = 0.$$

$$2. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad V(x) > 0.$$

$$3. \quad \text{به ازای هر } x \in \mathbb{R}^n \setminus \{0\}, \quad \dot{V}(x) < 0.$$

آن گاه $x = 0$ پایدار مجانبی خواهد بود.

قضیه ۷.۳.۱ (قضیه پایداری مجانبی سراسری [۳۰]). تحت شرایط قضیه ۶.۳.۱ اگر $V(\cdot)$ به طور شعاعی بی کران باشد آن گاه نقطه $x = 0$ پایدار مجانبی سراسری است.

قضیه ۸.۳.۱ (اصل تغییرناپذیری لاسال^{۵۵} [۳۰]). فرض کنیم یک تابع به طور پیوسته مشتق پذیر باشد، همچنین فرض کنیم:

$$1. \quad M \subset \mathbb{R}^n \text{ یک مجموعه فشرده و پایدار نسبت به جواب سیستم (۷.۱) باشد.}$$

$$2. \quad \text{به ازای } x \in M, \quad \dot{V}(x) \leq 0.$$

⁵⁵LaSalle Principle of Invariance

۳. E مجموعه همه نقاط M باشد به طوری که $\dot{V}(x) = 0$.

۴. N بزرگترین مجموعه پایدار در E باشد.

آن‌گاه به ازای هر $x(t_0) \in M$ که $t_0 \geq 0$ می‌باشد داریم:

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), N) = 0.$$

فصل ۲

یک مدل شبکه عصبی گرادینانی کارای جدید برای حل مسائل برنامه‌ریزی درجه‌ی دوم

مسائل برنامه‌ریزی درجه دوم در طیف گسترده‌ای از کاربردهای علمی و مهندسی، تحلیل رگرسیون، پردازش سیگنال، بازسازی تصویر، کنترل ربات و غیره به وجود می‌آیند. همچنین، برخی مسائل بهینه‌سازی غیرخطی معمولاً با یک مسئله‌ی برنامه‌ریزی درجه دوم تقریب زده می‌شوند. به طور کلی، مسئله‌ی برنامه‌ریزی درجه دوم می‌تواند به شکل زیر نوشته شود

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^T W x + c^T x \\ \text{subject to} \quad & Ax = b, \\ & x \geq 0. \end{aligned}$$

بسیاری از شبکه‌های عصبی برای مسئله‌ی فوق ارائه شده‌اند. برای مثال، برخی شبکه‌های عصبی اولیه-دوگان ارائه شدند [۱۵]، [۲۰]، [۷۹] و [۸۰]. به منظور ساده‌سازی ساختار شبکه‌ی عصبی دوگان، شبکه‌ی عصبی دوگان ساده شده معرفی شد [۴۶] و [۳۷]. با استفاده از قضیه‌ی تصویر، چندین شبکه‌ی عصبی تصویری برای حل مسئله‌ی برنامه‌ریزی درجه دوم توسعه داده شدند [۲۳]، [۲۸]، [۲۷]، [۲۹]، [۳۶]، [۴۸]، [۴۷]، [۴۹]، [۸۴] و [۸۹] که به‌طور

کلی به جواب‌های بهینه‌ی دقیق همگرا بودند. در [۴۸] و [۴۷]، لیو و وانگ چند شبکه‌ی عصبی بازگشتی تک‌لایه را برای حل مسائل برنامه‌ریزی درجه دوم پیشنهاد دادند. شبکه‌های عصبی بازگشتی تک‌لایه پیچیدگی ساختاری ساده‌تری نسبت به سایر شبکه‌های عصبی مانند شبکه‌ی لاگرانژی و شبکه‌ی تصویری دارند. تا به حال، تنها چند شبکه‌ی عصبی برای حل مسائل برنامه‌ریزی درجه دوم زیر با محدودیت‌های تساوی و نامساوی توسعه یافته است.

$$\begin{aligned} & \text{minimize} && \frac{1}{2}x^T W x + c^T x \\ & \text{subject to} && Ax = b, \\ & && Bx \leq d. \end{aligned} \tag{۱.۲}$$

که در آن W یک ماتریس $n \times n$ نیمه معین مثبت متقارن حقیقی، $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ ، $d \in \mathbb{R}^p$ و $b \in \mathbb{R}^m$ ، $B \in \mathbb{R}^{p \times n}$ ، $\text{rank}(A) = m$ ($0 < m < n$)، $A \in \mathbb{R}^{m \times n}$ ، $c \in \mathbb{R}^n$ ژانگ و وانگ [۹۰] و ژیا و همکاران [۸۱] با محدودیت‌های تساوی و نامساوی یک طرفه همانند حالات خاص محدودیت‌های کران دوطرفه عمل کردند. نشان دادن کران پایین $-\infty$ یا تا کران $+\infty$ نیاز به اعداد به اندازه کافی بزرگ دارد که برای طراحی شبکه‌های عصبی نامساعد است. عفتی و ناظمی [۹] برای حل این مسئله با شرایط کاراش-کوهن-تاکر (KKT) یک شبکه‌ی عصبی پیشنهاد دادند. با این حال، این شبکه‌ی عصبی دارای نورون‌های بیشتری است و نیاز به شرایط همگرایی قوی‌تری دارد.

با انگیزه‌ی مباحث فوق، هدف این پایان نامه توسعه‌ی یک شبکه‌ی عصبی جدید برای حل (۱.۲) است که متغیرهای حالت کمتر، پیچیدگی ساختاری کمتر و شرایط همگرایی قوی‌تری دارد.

۱.۲ مدل شبکه‌ی عصبی پیشنهادی

مسئله‌ی برنامه‌ریزی درجه دوم محدب به صورت کلی زیر را در نظر بگیرید:

$$\begin{aligned} & \text{minimize} && \frac{1}{2}x^T W x + c^T x \\ & \text{subject to} && Ax = b \\ & && Bx \leq d, \end{aligned} \tag{۲.۲}$$

که در آن $W \in \mathbb{R}^{n \times n}$ ، $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ یک ماتریس نیمه معین مثبت متقارن، $d \in \mathbb{R}^p$ و $b \in \mathbb{R}^m$ ، $B \in \mathbb{R}^{p \times n}$ ، $\text{rank}(A) = m$ ($0 < m < n$)، $A \in \mathbb{R}^{m \times n}$ ، $c \in \mathbb{R}^n$ می‌توانیم بین جواب مسئله (۲.۲) و جواب یک مسئله مینیمم سازی نامقیدهم ارز توسط یک

تابع شایسته ارتباط ایجاد کنیم. تابع شایسته تابعی است که مینیمم سراسری اش با جوابهای مسأله مکمل غیرخطی (NCP) ^۱ مطابقت می کند. رده توابع NCP که در زیر تعریف می شود برای ساختن یک تابع شایسته استفاده می شود.

تعریف ۱.۱.۲. تابع $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$ یک تابع NCP نامیده می شود اگر در شرایط زیر صدق کند:

$$\phi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0.$$

یک تابع NCP متداول، تابع فیشر برمیستر ^۲ است که قویاً نیمه هموار است و به صورت زیر تعریف می شود:

$$\phi_{FB}(a, b) = (a + b) - \sqrt{a^2 + b^2}.$$

تابع شایسته $\psi_{FB}: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$ به وسیله مجذور ϕ_{FB} به دست می آید یعنی

$$\psi_{FB}(a, b) = \frac{1}{4} \|\phi_{FB}(a, b)\|^2.$$

تابع فیشر برمیستر آشفته نیز به صورت زیر داده شده است:

$$\phi_{FB}^\varepsilon(a, b) = (a + b) - \sqrt{a^2 + b^2 + \varepsilon}, \quad \varepsilon \rightarrow 0_+.$$

گزاره ۱.۱.۲. برای هر $\varepsilon \in \mathbb{R}$ داریم:

$$\phi_{FB}^\varepsilon(a, b) = 0 \iff a > 0, b > 0, ab = \frac{\varepsilon}{4}.$$

گزاره بالا بدیهی است و برهان حذف شده است. قابل توجه است که $\phi_{FB}^\varepsilon(a, b)$ با توجه به a, b برای هر $\varepsilon > 0$ هموار می باشد.

لم ۱.۱.۲. x^* جواب بهینه‌ی (۲.۲) است اگر و تنها اگر $u^* \geq 0$ ، P و Q موجود باشند به طوری که $(x^*, u^*)^T$ در شرایط زیر صدق کند:

$$(I - P)[Wx^* + c + B^T u^*] + Q(Ax^* - b) = 0 \quad (3.2)$$

$$\phi_{FB}^\varepsilon(u^*, d - Bx^*) = 0 \quad (4.2)$$

برهان. اگر x^* جواب بهینه‌ی (۲.۲) باشد، مطابق شرایط KKT برای مسائل بهینه‌سازی محدب (۱.۲.۱)، $(x^{*T}, v^{*T}, u^{*T})^T$ و $u^* \geq 0$ موجودند به طوری که در معادلات زیر صدق می کنند.

¹Nonlinear Complementarity Problem

²Fisher Burmister

$$Wx^* + c + A^T v^* + B^T u^* = 0 \quad (5.2)$$

$$Ax^* = b \quad (6.2)$$

$$d - Bx^* \geq 0, \quad u^* \geq 0, \quad (u^*)^T (d - Bx^*) = 0 \quad (7.2)$$

به آسانی دیده می‌شود که (۴.۲) و (۷.۲) معادل هستند. به این معنی که

$$d - Bx^* \geq 0, \quad u^* \geq 0, \quad (u^*)^T (d - Bx^*) = 0 \iff \phi_{FB}^{\epsilon}(u^*, d - Bx^*) = 0$$

در ادامه، نشان می‌دهیم که جواب‌های (۳.۲) و جواب‌های (۵.۲) و (۶.۲) معادلند. از (۵.۲) داریم

$$-A(Wx^* + c + A^T v^* + B^T u^*) = 0$$

با جمع کردن معادله‌ی بالا با معادله‌ی (۶.۲) داریم

$$Ax^* - A(Wx^* + c) - AA^T v^* - AB^T u^* = b$$

بنابراین

$$A^T v^* = A^T (AA^T)^{-1} (Ax^* - b) - A^T (AA^T)^{-1} A(Wx^* + c + B^T u^*) \quad (8.2)$$

با جایگذاری (۸.۲) در معادله‌ی (۵.۲) داریم

$$[I - A^T (AA^T)^{-1} A](Wx^* + c + B^T u^*) + A^T (AA^T)^{-1} (Ax^* - b) = 0$$

با قرار دادن $P = A^T (AA^T)^{-1} A$ و $Q = A^T (AA^T)^{-1}$ ، می‌توان معادله‌ی فوق را به صورت زیر نوشت

$$(I - P)[Wx^* + c + B^T (u^* + Bx^* - d)^+] + Q(Ax^* - b) = 0$$

برعکس، اگر u^* موجود باشد به طوری که $(x^*, u^*)^T$ در معادله‌ی (۳.۲) صدق کند با ضرب A در طرفین معادله‌ی (۳.۲) داریم

$$A(I - P)[Wx^* + c + B^T (u^* + Bx^* - d)^+] + AQ(Ax^* - b) = 0$$

توجه کنید که $AQ = I$ و $A(I - P) = 0$ پس داریم $Ax^* - b = 0$. در نتیجه از (۳.۲) داریم

$$(I - P)(Wx^* + c + B^T u^*) = 0. \text{ با قرار دادن } v^* = (AA^T)^{-1} A(Wx^* + c + B^T u^*) \text{ داریم}$$

$$Wx^* + c + A^T v^* + B^T u^* = Wx^* + c + B^T u^* - A^T (AA^T)^{-1} A(Wx^* + c + B^T u^*)$$

$$= (I - P)(Wx^* + c + B^T u^*)$$

$$= 0$$

□

پس حکم ثابت شده است.

لم ۲.۱.۲. $y^* = (x^{*T}, u^{*T})^T$ در معادله زیر صدق می کند

$$\eta(y) = \begin{bmatrix} (I - P)[Wx + c + B^T u] + Q(Ax - b) \\ \phi_{\text{FB}}^\varepsilon(u, d - Bx) \end{bmatrix} = \circ, \quad (9.2)$$

اگر و فقط اگر x^* برای هر $\varepsilon \rightarrow \circ_+$ یک نقطه KKT برای (۲.۲) باشد.

برهان. فرض کنید $\eta(y^*) = \circ$ باشد، آنگاه:

$$(I - P)[Wx^* + c + B^T u^*] + Q(Ax^* - b), \quad (10.2)$$

$$\phi_{\text{FB}}^\varepsilon(u^*, d - Bx^*) = \circ, \quad \text{for } \varepsilon \rightarrow \circ_+. \quad (11.2)$$

از گزاره ۱.۱.۲ به آسانی درمی یابیم که $\phi_{\text{FB}}^\varepsilon(u^*, d - Bx^*) = \circ$ اگر و فقط اگر

$$u^* \geq \circ, \quad (Bx^* - d) \leq \circ, \quad u^{*T}(Bx^* - d) = \circ, \quad \text{for } \varepsilon \rightarrow \circ_+. \quad (12.2)$$

از (۱۰.۲) و (۱۲.۲) درمی یابیم که $y^* = (x^{*T}, u^{*T})^T$ در شرایط KKT صدق می کند. عکس این قضیه بدیهی است.

□

اکنون با استفاده از لم ۲.۱.۲ به راحتی درمی یابیم که جواب سیستم (۹.۲) معادل مسئله مینیمم سازی هموار نامقید زیر است:

$$\min E(y) = \frac{1}{\gamma} \|\eta(y)\|_2^2. \quad (13.2)$$

روشن است که $E(y)$ یک تابع شایسته هموار برای سیستم (۹.۲) است. تابع شایسته E در (۱۳.۲) برای هر $y \in \mathbb{R}^{m+n+l}$ به طور پیوسته مشتق پذیر است.

اکنون فرض کنید $v(\cdot)$, $u(\cdot)$, $x(\cdot)$ متغیرهای وابسته به زمان باشند. هدف ما طراحی یک شبکه عصبی است که یک نقطه تعادل رامعین می کند که همچنین یک نقطه ایستای تابع انرژی $E(y)$ می باشد. از این رومی توانیم روش تندترین کاهش را برای ساخت مدل شبکه عصبی زیر برای حل مسئله (۹.۲) به صورت زیر استفاده کنیم:

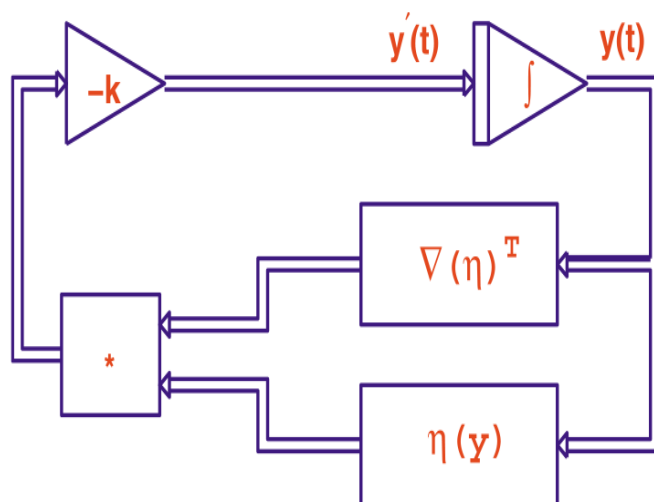
$$\frac{dy(t)}{dt} = -k \nabla E(y(t)), \quad k > \circ, \quad (14.2)$$

$$y(\circ) = y_\circ, \quad (15.2)$$

که k عامل مقیاس است و سرعت همگرایی شبکه عصبی (۱۴.۲) و (۱۵.۲) را نشان می دهد. برای دیدن اینکه چگونه شبکه‌های عصبی (۱۴.۲) و (۱۵.۲) می توانند برای حل (۱.۲) و (۲.۲) اعمال شوند. ما این را با تعدادی از مدل‌های شبکه‌های عصبی موجود مقایسه می کنیم.

۲.۲ مقایسه با برخی مدل‌های موجود

یک نوع از مدل شبکه‌های عصبی وجود دارد که آن را مدل گرادینان می‌نامند. برای استفاده از مدل شبکه عصبی گرادینان، یک مسئله‌ی بهینه‌سازی با محدودیت و یک مسئله‌ی بهینه‌سازی بدون محدودیت می‌تواند تقریب‌آمیز باشد. سپس تابع انرژی توسط روش تابع عملکرد ایجاد می‌شود. قابل توجه که مدل شبکه‌های عصبی گرادینانی است یک مزیت به‌عنوان مدل ممکن است، مستقیماً با استفاده از مشتقات تابع انرژی است وجود داشته باشد. اما این که همگرایی تضمین نشده، به ویژه در مورد مجموعه راه‌حل‌های بی‌کران کمبود است. [۸۸] علاوه بر این، شبکه عصبی گرادینان براساس تابع عملکرد مجاز به هر پارامتر قابل تنظیم عملکرد پارامتر می‌نامند. این که چطور شبکه عصبی (۱۴.۲) و (۱۵.۲) روی سخت‌افزار قابل اجراست در شکل ۱.۲ نشان داده شده است.



شکل ۱.۲: یک نمودار بلوکی ساده برای شبکه عصبی (۱۴.۲) و (۱۵.۲)

به‌عنوان مثال، روش عملکرد مجاز، مسئله‌ی (۱.۲) و (۲.۲) بهینه‌سازی محدودیت‌ها می‌تواند توسط به حداقل رساندن یک مسئله‌ی بهینه‌سازی بدون محدودیت تقریب زده شود.

$$\text{minimize } \bar{E}(x) = \frac{1}{\gamma} x^T W x + c^T x + \frac{\gamma}{\gamma} \left\{ \sum_{k=1}^p [(B_k x - d_k)^+]^2 + \sum_{q=1}^l (A_q x - b_q)^2 \right\},$$

که در آن γ یک پارامتر عملکرد است. مدل شبکه‌های عصبی گرادینانی است سپس داریم:

$$\frac{dx}{dt} = -\nabla \bar{E}(x) = -((Wx + c) + \gamma[B^T(Bx - d)^+ + A^T(Ax - b)]). \quad (۱۶.۲)$$

سیستم (۱۶.۲) به‌عنوان یک مدل شبکه عصبی کندی و چواس نامیده می‌شود [۵۷]. این شبکه به دلیل یک پارامتر عملکرد و یک راه‌حل بهینه برای اجرای وقتی که پارامترهای عملکرد بسیار زیاد هستند دقیق نیست ممنوع می‌باشد [۷۳].

بنابراین، این شبکه فقط برای هر پارامتر عملکرد داده شده به این راه‌حل تقریبی (۱۴.۲) و (۱۵.۲) می‌رسد. همچنین می‌توان نشان داد که کندی و چواس، شبکه عصبی (۱۶.۲) همگرای سراسری نیست، و به یک راه‌حل بهینه دقیق برخی از مسائل برنامه نویسی محدب باشد. برای مثال، می‌توان مثال ۴.۱ در [۷] مشاهده کرد. در [۸]، ناظمی مدل شبکه عصبی زیر را برای حل (۱.۲) و (۲.۲) به‌عنوان یک پیشنهاد ارائه داد:

$$\frac{dy}{dt} = \Phi(y), \quad (17.2)$$

$$y(t_0) = \circ, \quad y(t) = (x(t)^T, u(t)^T, v(t)^T)^T, \quad u(t_0) > \circ, \quad (18.2)$$

که در آن

$$\Phi(y) = \begin{bmatrix} -(Wx + c + \frac{1}{\gamma} B^T u^\gamma + A^T v) \\ \text{diag}(u_1, u_2, \dots, u_p)(Bx - d) \\ Ax - b \end{bmatrix}.$$

در [۷]، ناظمی همچنین شبکه عصبی زیر را برای حل مسائل (۱.۲) و (۲.۲) ارائه داد:

$$\frac{dx}{dt} = -(Wx + c + B^T(u + Bx - d)^+ + A^T v), \quad (19.2)$$

$$\frac{du}{dt} = (u + Bx - d)^+ - u, \quad (20.2)$$

$$\frac{dv}{dt} = Ax - b. \quad (21.2)$$

در [۹]، ناظمی و عفتی یک مدل گرادیان برای شبیه‌سازی با مسائل (۱.۲) و (۲.۲) ارائه داد:

$$\frac{d(y(t))}{dt} = -\nabla \hat{E}(y(t)), \quad (22.2)$$

$$y(\circ) = y_\circ, \quad y(t) = (x(t)^T, u(t)^T, v(t)^T)^T, \quad (23.2)$$

که در آن

$$\hat{E}(y) = \frac{1}{\gamma} \|\Psi(y)\|^\gamma, \quad (24.2)$$

$$\Phi(y) = \begin{bmatrix} -(Wx + c + A^T v + B^T u) \\ b - Ax \\ \phi_{\text{FB}}^\varepsilon(u, d - Bx) \end{bmatrix} = \circ, \quad (25.2)$$

در مدل‌های (۱۷.۲) - (۱۸.۲)، (۱۹.۲) - (۲۰.۲) و (۲۱.۲) - (۲۲.۲)، متغیرهای حالت در $x \in \mathbb{R}^n$ ، $u \in \mathbb{R}^p$ و $v \in \mathbb{R}^l$ وجود دارند. بنابراین، شبکه‌های عصبی $n + p + l$ نورون دارند. همچنین، شبکه عصبی ما در (۱۴.۲) فقط $n + p$ نورون دارند. علاوه بر این، شبکه عصبی ما به‌عنوان یک اصلاح از مدل و یک گسترش مدل پیشنهاد شده در [۹] برای حل مسائل برنامه نویسی محدب است.

۳.۲ تجزیه و تحلیل پایداری

در این بخش همگرایی مسی‌ر شبکه عصبی (۱۴.۲) و (۱۵.۲) را بررسی می‌کنیم.

قضیه ۱.۳.۲. فرض کنید ماتریس ژاکوبین $\nabla\eta(y)$ نامنفرد و x^* نقطه KKT (۲.۲) باشد آنگاه u^* نقطه تعادل شبکه عصبی (۱۴.۲) و (۱۵.۲) است. از سوی دیگر اگر $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ نقطه تعادل شبکه عصبی (۱۴.۲) و (۱۵.۲) باشد و ماتریس ژاکوبین $\eta(y)$ در (۹.۲) غیرمنفرد باشد آنگاه x^* نقطه KKT مسئله (۲.۲) می‌باشد.

برهان. فرض کنید ماتریس ژاکوبین x^* جواب بهینه (۲.۲) باشد. ازلم ۲.۱.۲ به روشنی درمی‌یابیم که $\eta(y^*) = 0$. بایک محاسبه ساده به وضوح داریم:

$$\nabla E(y) = (\nabla\eta(y))^T \eta(y), \quad (26.2)$$

که $\nabla\eta(y)$ ماتریس ژاکوبین $\eta(y)$ است. با استفاده از (۲۶.۲) داریم $\nabla E(y^*) = 0$ یعنی y^* نقطه تعادل مدل دینامیکی (۱۴.۲) و (۱۵.۲) است، عکس قضیه بدیهی است. \square

لم ۱.۳.۲. نقطه تعادل مدل شبکه عصبی ارائه شده (۱۴.۲) و (۱۵.۲) منحصر به فرد است.

برهان. چون مسئله (۲.۲) دارای جواب بهینه منحصر به فرد x^* ، است شرایط KKT برای داشتن جواب منحصر به فرد $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ لازم و کافی است. علاوه بر آن از قضیه ۱.۳.۲ درمی‌یابیم که جواب سیستم (۹.۲)، نقطه تعادل مدل شبکه عصبی ارائه شده (۱۴.۲) و (۱۵.۲) است. بنابراین نقطه تعادل شبکه (۱۴.۲) و (۱۵.۲) منحصر به فرد است. \square

قضیه ۲.۳.۲. فرض کنید y^* یک نقطه تعادل تنه‌برای (۱۴.۲) و (۱۵.۲) باشد آنگاه y^* برای (۱۴.۲) و (۱۵.۲) پایدار مجانبی است.

برهان. ابتدا ملاحظه می‌کنیم که $E(y) \geq 0$ و $E(y^*) = 0$. علاوه بر این چون y^* یک نقطه تعادل تنه‌برای (۱۴.۲) و (۱۵.۲) است پس یک همسایگی $\Omega_* \subseteq \mathbb{R}^{n+m+l}$ برای y^* وجود دارد به طوری که:

$$\nabla E(y^*) = 0, \text{ and } \nabla E(y) \neq 0, \forall y \in \Omega_* \setminus \{y^*\}.$$

ادعای می‌کنیم که برای هر $y \in \Omega_* \setminus \{y^*\}$ داریم $E(y) > 0$. در غیر این صورت اگر $y \in \Omega_* \setminus \{y^*\}$ وجود داشته باشد به طوری که $E(y) = 0$ شود آنگاه با استفاده از (۱۳.۲) و (۲۶.۲) داریم $\nabla E(y) = 0$. یعنی y یک نقطه تعادل (۱۵.۲) می‌باشد که به وضوح با این فرض که y یک نقطه تعادل تنه‌دار Ω_* است در تناقض است. علاوه بر این

$$\frac{dE(y(t))}{dt} = [\nabla E(y(t))]^T \frac{dy(t)}{dt} = -k \|\nabla E(y(t))\|^2 \leq 0, \quad (27.2)$$

$$\frac{dE(y(t))}{dt} < \circ, \quad \forall y(t) \in \Omega_* \text{ and } y(t) \neq y^*.$$

با استفاده از قضیه (۶.۳.۱) درمی یابیم که y^* پایدارمجبانی است.

□

قضیه ۳.۳.۲. فرض کنید $y = y(t, y_0)$ با نقطه شروع $y_0 = y(0, y_0)$ یک مسیر برای (۱۴.۲) و (۱۵.۲) باشد و مجموعه سطح $L(y_0) = \{y \in \mathbb{R}^{n+m+l} : E(y) \leq E(y_0)\}$ کراندار باشد آن گاه:

$$(a) \quad \gamma^+(y_0) = \{y(t, y_0) | t \geq 0\} \text{ کراندار است.}$$

$$(b) \quad \text{وجود دارد } \bar{y} \text{ به طوریکه } \lim_{t \rightarrow \infty} y(t, y_0) = \bar{y}.$$

برهان. (a) فرض کنید y^* یک نقطه تعادل شبکه (۱۴.۲) و (۱۵.۲) باشد آن گاه می گوییم $\nabla E(y^*) = 0$. با محاسبه مشتق $E(y)$ در طول مسیر $y(t, y_0), (t \geq 0)$ داریم:

$$\frac{dE(y(t))}{dt} = [\nabla E(y(t))]^T \frac{dy}{dt} = -k \|\nabla E(y(t))\|^2 \leq 0. \quad (28.2)$$

بنابراین در طول مسیر $y = y(t, y_0), (t \geq 0)$ ، $E(y)$ به طور یکنواخت غیرصعودی است.

بنابراین $\gamma^+(y_0) \subseteq L(y_0)$ است یعنی $\gamma^+(y_0) = \{y(t, y_0) | t \geq 0\}$ کراندار است.

(b) با استفاده از (a)، $\gamma^+(y_0) = \{y(t, y_0) | t \geq 0\}$ مجموعه ای کراندار از نقاط است. در مسیر مستقیم

دنباله صعودی یکنواخت $\{\bar{t}_n\}$ ، $\bar{t}_n \rightarrow +\infty, 0 \leq \bar{t}_1 \leq \dots \leq \bar{t}_n \leq \dots$ یک زیر دنباله

کراندار ساخته شده از تعداد زیادی نقاط نامتناهی می باشد. بنابراین نقطه حدی \bar{y} وجود دارد

یعنی اینکه یک زیر دنباله $\{t_n\} \subseteq \{\bar{t}_n\}, t_n \rightarrow +\infty$ وجود دارد به طوریکه:

$$\lim_{n \rightarrow +\infty} y(t_n, y_0) = \bar{y},$$

که \bar{y} در رابطه زیر صدق می کند

$$\frac{dE(y(t))}{dt} = 0,$$

که نشان می دهد \bar{y} نقطه ω -limit است. با استفاده از قضیه مجموعه تغییر ناپذیر

لسال داریم:

$$y(t, y_0) \rightarrow \bar{y} \in M \quad \text{as } t \rightarrow \infty,$$

که M بزرگترین مجموعه تغییر ناپذیر در $K = \{y(t, y_0) | \frac{dE(y(t, y_0))}{dt} = 0\}$ است. از (۱۴.۲)

، (۱۵.۲) و (۲۸.۲) نتیجه می شود که $\frac{dx}{dt} = 0, \frac{du}{dt} = 0$ و $\frac{dE(y(t))}{dt} = 0 \Leftrightarrow \frac{dv}{dt} = 0$. بنابراین

با استفاده از $M \subseteq K \subseteq D^*$ داریم $\bar{y} \in D^*$. بنابراین از هر y_0 حالت شروع، مسیر $y(t, y_0)$ از

(۱۴.۲) و (۱۵.۲) به \bar{y} میل می کند و این برهان را کامل می کند.

□

۲۶ یک مدل شبکه عصبی گرادیانی کارای جدید برای حل مسائل برنامه‌ریزی...

نتیجه ۱.۳.۲. اگر $D^* = \{(x^{*T}, u^{*T}, v^{*T})^T\}$ آن گاه شبکه عصبی (۱۴.۲) و (۱۵.۲) برای حل (۲.۲) به نقطه تعادل منحصر به فرد $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ پایدارمجانبی سراسری است که D^* به عنوان مجموعه نقاط بهینه (۲.۲) تعریف می‌شود.

فصل ۳

انتخاب سبد سهام بهینه با فرض قرض گرفتن

۱.۳ مقدمه

روش میانگین-واریانس برای مسئله‌ی انتخاب سبد سهام، ابتدا توسط مارکوویتز مطرح شده است، که نقش مهمی در توسعه نظریه انتخاب سبد سهام مدرن داشته است. مدل میانگین-واریانس مارکوویتز در انتخاب سبد سهام شامل ترسیم یک مرز کارا است [۳۳، ۳۴]. برای فرموله کردن این مدل، لازم است توزیع احتمال تخمین زده شود، همچنین یک بردار میانگین و یک ماتریس کوواریانس مورد بحث قرار می‌گیرد. در مسئله‌ی انتخاب سبد سهام میانگین-واریانس، مطالعات پیشین شامل [۷۸]، شارپ [۶۱]، مرتون [۲۶]، سزگو [۵]، پرولد [۴۲]، ووروس [۴۱]، پنگ [۵۲]، بست و گراور [۵۳]، بست و جاروسلاوا [۱۰]، یوشیموتو و [۵۹] کاوادی و کونو هستند. در تحقیقات آن‌ها معمولاً فرض می‌شود که بردار میانگین و ماتریس کوواریانس معلوم این موضوع به این معنی است که همه‌ی بازده‌های میانگین، واریانس‌ها و کوواریانس‌ها از دارایی‌های ریسکی را می‌توان توسط یک سرمایه‌گذار به‌طور دقیق برآورد شود. علاوه بر این، فرض اصلی برای استفاده کردن از مدل مارکوویتز این است که وضعیت دارایی در آینده می‌تواند به‌درستی توسط داده‌های دارایی در گذشته منعکس شود، یعنی هرچند، این نوع فرض برای هر تغییر در بازارهای واقعی سخت است. بست و گراور [۵۲]، هر دو تغییراتی در میانگین‌ها

در مورد حساسیت سبد اوراق بهادار کارا از دیدگاه تحلیلی و محاسباتی را مورد بحث قرار دادند. چوپرا و زیمبا [۷۲] در مورد اثرات تخمین خطاها در میانگین و کوواریانس‌ها، روی انتخاب سبد سهام بهینه بحث کردند. تعدادی از مطالعات درباره‌ی شبکه‌های عصبی کارا برای پیدا کردن سبد سهام‌های کارا از حل مدل میانگین-واریانس وقتی که همیشه نقش بنیادی در بهینه سازی سبد سهام است وجود دارد، اما این برای به دست آوردن سبد سهام‌ها در قالب بسته تحت قیدهای عمومی مانند قیدهای نامنفی روی دارایی‌های همبسته خیلی دشوار است (جلوگیری از فروش کوتاه مدت).

تاناکا و همکاران [۳۵] مدل انتخاب سبد سهام را براساس احتمال فازی پیشنهاد کردند، که می‌تواند به‌عنوان یک توسعه طبیعی از مدل مارکویتز در نظر گرفته شود زیرا این مدل احتمال را به احتمال فازی توسعه می‌دهد. این روش اجازه مشارکت دانش تخصصی را به‌وسیله استفاده از درجه امکان، برای منعکس کردن درجه تشابه بین وضعیت آینده از دارایی‌ها و وضعیت دوره‌های قبلی را می‌دهد. هرچند، بازده مورد انتظار و ریسک از دارایی نمی‌تواند به‌طور دقیق پیشگویی شود زیرا عوامل نامعلومی است که روی بازارهای مالی تأثیر می‌گذارد. ژانگ و نی [۷۶] در مورد انتخاب سبد سهام کارای مجاز برای دارایی‌های ریسکی تحت فرض بر این که بازده مورد انتظار و ریسک از دارایی خطاهای مجاز برای نشان دادن عدم اطمینان واقعی سهام‌های سرمایه گذار دارد بحث می‌کنند آن‌ها یک استخراج تحلیلی از مرز کارای مجاز وقتی که فروش کوتاه مدت وجود ندارد را مورد بحث قرار می‌دهند. در بازار کامل، سرمایه گذاران نه تنها دارایی‌های ریسکی را خریداری می‌کنند، بلکه آن‌ها نیز می‌توانند برای خرید دارایی‌های ریسکی پول قرض بگیرند. چگونه سرمایه گذاران تصمیم درستی برای قرض گرفتن دارند؟ این برای یک مسئله‌ی انتخاب سبد سهام بسیار مهم است.

۲.۳ مدل سبد سهام کارآمد قابل قبول با قرض گرفتن

فرض کنید w سرمایه‌ی اولیه سرمایه‌گذار و v حداکثر مقدار پولی که سرمایه‌گذار قادر است امانت بگیرد باشند. v را یک کران بالا در قرض گرفتن در نظر می‌گیریم، در غیر این صورت قرض گرفتن بی‌کران است. به‌منظور تشریح بهتر، ما از نشانه‌های زیر استفاده می‌کنیم:

r_j ، بازده مورد انتظار دارایی ریسکی j ، $j = 1, 2, \dots, n$ ؛
 r_j^* ، نرخ بهره قرض گرفتن؛

x_j ، مقدار سرمایه‌گذاری شده در دارایی ریسکی j ، $j = 1, 2, \dots, n$ ؛

$x = (x_1, x_2, \dots, x_n)^T$ ، بردار مقدار سرمایه‌گذاری شده از دارایی‌های ریسکی؛

$r = (r_1, r_2, \dots, r_n)^T$ ، بردار بازده مورد انتظار؛

$V = (\sigma_{ij})_{n \times n}$ ، ماتریس کوواریانس بازده‌ها؛

$l = (1, 1, \dots, 1)^T$

اگر قرض گرفتن با نرخ بدون ریسک ممکن باشد، آن‌گاه بازده مورد انتظار و واریانس مربوط

به پرتفوی x به ترتیب برابر است با $x^T V x$ و $r^T x + r^*(w - l^T x)$. بنابراین، مسئله‌ی انتخاب سبد سهام‌ها ممکن است توسط برنامه‌ریزی درجه‌ی دوم زیر تشریح شود:

$$\begin{aligned} & \text{minimize} && x^T V x \\ & \text{subject to} && r^T x + r^*(w - l^T x) = \mu, \\ & && w \leq l^T x \leq w + v, \\ & && x \geq 0, \end{aligned} \quad (1.3)$$

که در آن μ یک بازده مورد انتظار از پرتفوی‌ها است. فرض کنید داده‌های مشاهده شده روی بازده‌ها از دارایی‌ها روی m دوره داده شود. در زمان گسسته $k = (1, 2, \dots, m)$ ، نوع از نرخ‌های مورد انتظار به صورت یک بردار $r_k = (r_{k1}, r_{k2}, \dots, r_{kn})^T$ نشان داده می‌شود. در [۳۵]، میانگین وزنی فازی از برآورد r_j و کوواریانس فازی از برآورد σ_{ij} برای داده‌های داده شده وجود دارد، که به صورت زیر تعریف می‌شوند:

$$\bar{r}_j = \frac{\sum_{k=1}^m (h_k r_{kj})}{\sum_{k=1}^m h_k}, \quad j = 1, 2, \dots, n, \quad (2.3)$$

و

$$\bar{\sigma}_{ij} = \frac{\sum_{k=1}^m (r_{ki} - \bar{r}_i)(r_{kj} - \bar{r}_j)h_k}{\sum_{k=1}^m h_k}, \quad i, j = 1, 2, \dots, n, \quad (3.3)$$

که در آن h_k یک درجه‌ی امکان^۱ است که منعکس کننده میزان شباهت بین وضعیت آینده از بازارهای دارایی و k امین نمونه‌ی ارائه شده توسط کارشناسان است. چون $j = 1, 2, \dots, n$ ، r_j در یک محیط اقتصادی نامعلوم فازی وجود دارد، و از یک زمان به زمان دیگر تغییر می‌کند، در آینده تعیین کردن بازده‌ها و ریسک از n دارایی را نمی‌توان به طور دقیق پیش بینی کرد. در [۷۶] بازده‌های متوسط شدنی و کوواریانس‌ها به ترتیب به صورت زیر تعریف می‌شوند:

$$r_j^* = \bar{r}_j + \phi_j, \quad \phi_{jl} \leq \phi_j \leq \phi_{jh}, \quad j = 1, 2, \dots, n, \quad (4.3)$$

و

$$\sigma_{ij}^* = \bar{\sigma}_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ijl} \leq \varepsilon_{ij} \leq \varepsilon_{ijh}, \quad j = 1, 2, \dots, n, \quad (5.3)$$

که در آن ϕ_j خطای قابل قبول^۲ برای \bar{r}_j را نشان می‌دهد، ϕ_{jl} کران پایین از ϕ_j را نشان می‌دهد، ϕ_{jh} کران بالا از ϕ_j را نشان می‌دهد، ε_{ij} خطای قابل قبول برای $\bar{\sigma}_{ij}$ را نشان می‌دهد، ε_{ijl} کران پایین از ε_{ij} را نشان می‌دهد و ε_{ijh} کران بالا از ε_{ij} را نشان می‌دهد.

با ترکیب اطلاعات آینده درباره بازده دارایی‌ها با نظر کارشناسان، ϕ_{jh} ، ϕ_{jl} ، ε_{ijh} و ε_{ijl} را می‌توان

¹Possibility grade

²Admissible error

۳۰ انتخاب سبد سهام بهینه با فرض قرض گرفتن

برآورد کرد. به ترتیب، فاصله‌های $[\bar{r}_j + \phi_{jl}, \bar{r}_j + \phi_{jh}]$ و $[\bar{\sigma}_{ij} + \varepsilon_{ijl}, \bar{\sigma}_{ij} + \varepsilon_{ijh}]$ نیز تعیین می‌شوند. ε_{ij} و ϕ_j می‌توان براساس نگرش خود روی بازده و ریسک توسط سرمایه‌گذار انتخاب شود. به‌طور متوسط بردار بازده متوسط شدنی را می‌توان با

$$\bar{r}^* = a + \Phi, \quad \Phi_l \leq \Phi \leq \Phi_h, \quad (۶.۳)$$

نشان داد که در آن

$$a = (\bar{r}_1, \bar{r}_2, \dots, \bar{r}_n)^T,$$

$$\Phi = (\phi_1, \phi_2, \dots, \phi_n)^T,$$

$$\Phi_l = (\phi_{1l}, \phi_{2l}, \dots, \phi_{nl})^T,$$

و

$$\Phi_h = (\phi_{1h}, \phi_{2h}, \dots, \phi_{nh})^T.$$

به‌طور مشابه، ماتریس کوواریانس قابل قبول را می‌توان به‌صورت زیر تعریف کرد:

$$V^* = \Sigma + V_\varepsilon, \quad V_{\varepsilon l} \leq V_\varepsilon \leq V_{\varepsilon h}, \quad (۷.۳)$$

که در آن

$$\Sigma = (\bar{\sigma}_{ij})_{n \times n},$$

$$V_\varepsilon = (\varepsilon_{ij})_{n \times n},$$

$$V_{\varepsilon l} = (\varepsilon_{ijl})_{n \times n},$$

و

$$V_{\varepsilon h} = (\varepsilon_{ijh})_{n \times n}.$$

بنابراین، بازده‌شدنی و کوواریانس وابسته به پرتفوی x به‌صورت زیر داده می‌شود:

$$\hat{r} = (a + \Phi)^T x + r^*(w - l^T x), \quad (۸.۳)$$

و

$$\hat{\sigma}^2 = x^T \Sigma x + x^T V_\varepsilon x. \quad (۹.۳)$$

اکنون برای هر $x \geq 0$ داریم:

$$(a + \Phi_l)^T x + r^*(w - l^T x) \leq \hat{r} \leq (a + \Phi_h)^T x + r^*(w - l^T x),$$

و

$$x^T \Sigma x + x^T V_{\varepsilon l} x \leq \hat{\sigma}^2 \leq x^T \Sigma x + x^T V_{\varepsilon h} x.$$

بنابراین مدل سبد سهام قابل قبول زیر را ایجاد می‌کنیم:

$$\begin{aligned} & \text{minimize} && x^T(\Sigma + V_\varepsilon)x \\ & \text{subject to} && (a + \Phi)^T x + r^*(w - l^T x) \geq \mu, \\ & && w \leq l^T x \leq w + v, \\ & && x \geq 0 \end{aligned} \tag{۱۰.۳}$$

اگر $(\Phi, V_\varepsilon) = (\Phi_h, V_{\varepsilon l})$ باشد آن‌گاه (۱۰.۳) را می‌توان به صورت زیر نوشت:

$$\begin{aligned} & \text{minimize} && x^T(\Sigma + V_{\varepsilon l})x \\ & \text{subject to} && (a + \Phi_h)^T x + r^*(w - l^T x) \geq \mu, \\ & && w \leq l^T x \leq w + v, \\ & && x \geq 0 \end{aligned} \tag{۱۱.۳}$$

اگر $(\Phi, V_\varepsilon) = (\Phi_l, V_{\varepsilon h})$ باشد آن‌گاه (۱۰.۳) را می‌توان به صورت زیر نوشت:

$$\begin{aligned} & \text{minimize} && x^T(\Sigma + V_{\varepsilon h})x \\ & \text{subject to} && (a + \Phi_l)^T x + r^*(w - l^T x) \geq \mu, \\ & && w \leq l^T x \leq w + v, \\ & && x \geq 0 \end{aligned} \tag{۱۲.۳}$$

مدل (۱۱.۳) به این معنی است که سرمایه‌گذار بازده و ریسک را خوش‌بینانه برآورد می‌کند. مدل (۱۲.۳) به این معنی است که سرمایه‌گذار بازده و ریسک را بدبینانه برآورد می‌کند. مدل (۱۰.۳) طرحی است که سرمایه‌گذار پرتفویی را انتخاب می‌کند که نه خیلی خوش‌بینانه و نه خیلی بدبینانه است. در مدل‌های سبد سهام کارای مجاز از دارایی‌های ریسکی در [۷۶] می‌تواند توسط مدل‌های (۱۰.۳ – ۱۲.۳) وقتی که $w = 1$ و $v = 0$ به دست آید. واضح است که مدل (۱۰.۳) توسعه‌هایی برای مدل‌های قبلی در [۳۵، ۷۶] برای مسئله‌ی انتخاب سبد سهام‌ها است.

۳.۳ مثال عددی

در این فصل به منظور نشان دادن روش شبکه عصبی پیشنهاد شده و فرمول‌های تحلیلی از سبدهای کارآمد قابل قبول، ما یک مثال کاربردی نشان داده شده در جدول (۱.۳) که توسط مارکوینتز [۳۴]، تاناکا و همکاران [۳۵] در نظر گرفته شده را ارائه می‌دهیم. که در این مثال، ستون‌های (۱۰ – ۲) به ترتیب نشان دهنده‌ی تأمین داده‌های تنباکو آمریکایی، *A.T.&T.*، فولاد

جدول ۱.۳: بالاترین سبد سهام کارایی مجاز

h_t	Year	Am.T.	A.T.&T.	U.S.S.	G.M.	A.T.&S.	C.C.	Bdn.	Frstn.	S.S.
۰/۱	۱۹۳۷(۱)	-۰/۳۰۵	-۰/۱۷۳	-۰/۳۱۸	-۰/۴۷۷	-۰/۴۵۷	-۰/۰۶۵	-۰/۳۱۹	-۰/۴	-۰/۴۳۵
۰/۱۱۸	۱۹۳۸(۲)	۰/۵۱۳	۰/۰۹۸	۰/۲۵۸	۰/۷۱۴	۰/۱۰۷	۰/۳۳۸	۰/۰۷۶	۰/۳۳۶	۰/۲۳۸
۰/۱۳۵	۱۹۳۹(۳)	۰/۰۵۵	۰/۲	-۰/۰۴۷	۰/۱۶۵	-۰/۴۲۴	-۰/۰۷۸	۰/۳۸۱	-۰/۰۹۳	-۰/۲۹۵
۰/۱۵۳	۱۹۴۰(۴)	-۰/۱۲۶	۰/۰۳	۰/۱۰۴	-۰/۰۴۳	-۰/۱۸۹	-۰/۰۷۷	-۰/۰۵۱	-۰/۰۹	-۰/۰۳۶
۰/۱۷۱	۱۹۴۱(۵)	-۰/۲۸	-۰/۱۸۳	-۰/۱۷۱	-۰/۲۷۷	۰/۶۳۷	-۰/۱۸۷	۰/۰۸۷	-۰/۱۹۴	-۰/۲۴
۰/۱۸۸	۱۹۴۲(۶)	-۰/۰۰۳	۰/۰۶۷	-۰/۰۳۹	۰/۴۷۶	۰/۸۶۵	۰/۱۵۶	۰/۲۶۲	۱/۱۱۳	۰/۱۲۶
۰/۲۰۶	۱۹۴۳(۷)	۰/۴۲۸	۰/۳	۰/۱۴۹	۰/۲۲۵	۰/۳۱۳	۰/۳۵۱	۰/۳۴۱	۰/۵۸	۰/۶۳۹
۰/۲۲۴	۱۹۴۴(۸)	۰/۱۹۲	۰/۱۰۳	۰/۲۶	۰/۲۹	۰/۶۳۷	۰/۲۳۳	۰/۲۲۷	۰/۴۷۳	۰/۲۸۲
۰/۲۴۱	۱۹۴۵(۹)	۰/۴۴۶	۰/۲۱۶	۰/۴۱۹	۰/۲۱۶	۰/۳۷۳	۰/۳۴۹	۰/۳۵۲	۰/۲۲۹	۰/۵۷۸
۰/۲۵۹	۱۹۴۶(۱۰)	-۰/۰۸۸	-۰/۰۴۶	-۰/۰۷۸	-۰/۲۷۲	-۰/۰۳۷	-۰/۲۰۹	۰/۱۵۳	-۰/۱۲۶	۰/۲۸۹
۰/۲۷۶	۱۹۴۷(۱۱)	-۰/۱۲۷	-۰/۰۷۱	۰/۱۶۹	۰/۱۴۴	۰/۰۲۶	۰/۳۵۵	-۰/۰۹۹	۰/۰۰۹	۰/۱۸۴
۰/۲۹۴	۱۹۴۸(۱۲)	-۰/۰۱۵	۰/۰۵۶	-۰/۰۳۵	۰/۱۰۷	۰/۱۵۳	-۰/۲۳۱	۰/۰۳۸	۰	۰/۱۱۴
۰/۳۱۲	۱۹۴۹(۱۲)	۰/۳۰۵	۰/۰۳۸	۰/۱۳۳	۰/۳۲۱	۰/۰۶۷	۰/۲۴۶	۰/۲۷۳	۰/۲۲۳	-۰/۲۲۲
۰/۳۲۹	۱۹۵۰(۱۴)	-۰/۰۹۶	۰/۰۸۹	۰/۷۳۲	۰/۳۰۵	۰/۵۷۹	-۰/۲۴۸	۰/۰۹۱	۰/۶۵	۰/۳۲۷
۰/۳۴۷	۱۹۵۱(۱۵)	۰/۰۱۶	۰/۰۹	۰/۰۲۱	۰/۱۹۵	۰/۰۴	-۰/۰۶۴	۰/۰۵۴	-۰/۱۳۱	۰/۳۳۳
۰/۳۶۵	۱۹۵۲(۱۶)	۰/۱۲۸	۰/۰۸۳	۰/۱۳۱	۰/۳۹	۰/۴۳۴	۰/۰۷۹	۰/۱۰۹	۰/۱۷۵	۰/۰۶۲
۰/۳۸۲	۱۹۵۳(۱۷)	-۰/۰۰۱	۰/۰۳۵	۰/۰۰۶	-۰/۰۷۲	-۰/۰۲۷	۰/۰۶۷	۰/۲۱	-۰/۰۸۴	-۰/۰۴۸
۰/۰۴	۱۹۴۵(۱۸)	۰/۱۵۴	۰/۱۷۶	۰/۹۰۸	۰/۷۱۵	۰/۴۶۹	۰/۰۷۷	۰/۱۱۲	۰/۷۵۶	۰/۱۸۵

ایالات متحده، جنرال موتورز، اتچسون & توپکا & سانتا فا، کوکا-کولا، بوردن، فیرستون و شارون استیل را مورد بررسی قرار داده‌ایم، Φ_l و Φ_h به صورت زیر هستند

$$\Phi_h = (0/014, 0/014, 0/039, 0/041, 0/047, 0/01, 0/027, 0/043, 0/028)^T, \quad \Phi_l = -\Phi_h$$

به منظور مقایسه با تاناکا [۳۵]، ژانگ و نی [۷۶]، فرض کنیم $V_\varepsilon = 0$ و درجه‌ی امکان h_t به صورت زیر تعریف می‌شود

$$h_{kij} = h_{ki} = h_k = 0/1 + 0/3 \frac{k-1}{17}, \quad k = 1, 2, \dots, 18.$$

با توجه به (۳.۳) \hat{a} ، $\hat{a} + \Phi_h$ و $\hat{a} + \Phi_l$ به شرح زیر به دست می‌آیند:

$$\hat{a} = (0/07099, 0/07012, 0/19645, 0/20610, 0/23390, 0/05317, 0/13619, 0/21603, 0/14386)^T,$$

$$\hat{a} + \Phi_h = (0/08499, 0/08412, 0/23545, 0/24710, 0/28090, 0/06317, 0/16319, 0/25903, 0/17186)^T,$$

$$\hat{a} + \Phi_l = (0/05699, 0/05612, 0/15745, 0/16510, 0/18690, 0/04317, 0/10919, 0/17303, 0/11586)^T.$$

با توجه به (۴.۳) ماتریس کوواریانس قابل قبول V^* به شرح زیر بیان می‌شود:

$$V^* = \hat{\Sigma} = \begin{pmatrix} 0/04042 & 0/01593 & 0/01933 & 0/03277 & 0/01046 & 0/02665 & 0/0183 & 0/02816 & 0/02077 \\ 0/01593 & 0/01119 & 0/01701 & 0/01875 & 0/00771 & 0/00705 & 0/00912 & 0/02106 & 0/01518 \\ 0/01933 & 0/01701 & 0/09852 & 0/06287 & 0/04545 & 0/00694 & 0/0035 & 0/07752 & 0/03114 \\ 0/03277 & 0/01875 & 0/06287 & 0/07954 & 0/04465 & 0/02072 & 0/00904 & 0/07894 & 0/01964 \\ 0/01046 & 0/00771 & 0/04545 & 0/04465 & 0/09713 & 0/00585 & 0/01096 & 0/0086 & 0/0285 \\ 0/02665 & 0/00705 & 0/00694 & 0/02072 & 0/00585 & 0/0405 & 0/00953 & 0/01952 & 0/01095 \\ 0/0183 & 0/00912 & 0/0035 & 0/00904 & 0/01096 & 0/00953 & 0/01986 & 0/01784 & 0/00079 \\ 0/02816 & 0/02106 & 0/07752 & 0/07894 & 0/0086 & 0/01952 & 0/01784 & 0/13343 & 0/03455 \\ 0/02077 & 0/01518 & 0/03114 & 0/01964 & 0/0285 & 0/01095 & 0/00079 & 0/03455 & 0/06211 \end{pmatrix}$$

فرض کنید که سرمایه اولیه سرمایه گذار صد هزار دلار، $w = 100,000$ باشد. فرض می‌کنیم که نرخ بهره وام ۴٪ و حداکثر مبلغی است که سرمایه گذار قادر به قرض گرفتن صد هزار دلار است، یعنی $r^* = 4\%$ و $v = 100,000$. از روش شبکه‌های عصبی برای یافتن سبد سهام‌های کارای قابل قبول استفاده می‌کنیم، ما مقدار سرمایه گذاری از بالاترین و پایین ترین سبد سهام‌ها را نشان می‌دهیم.

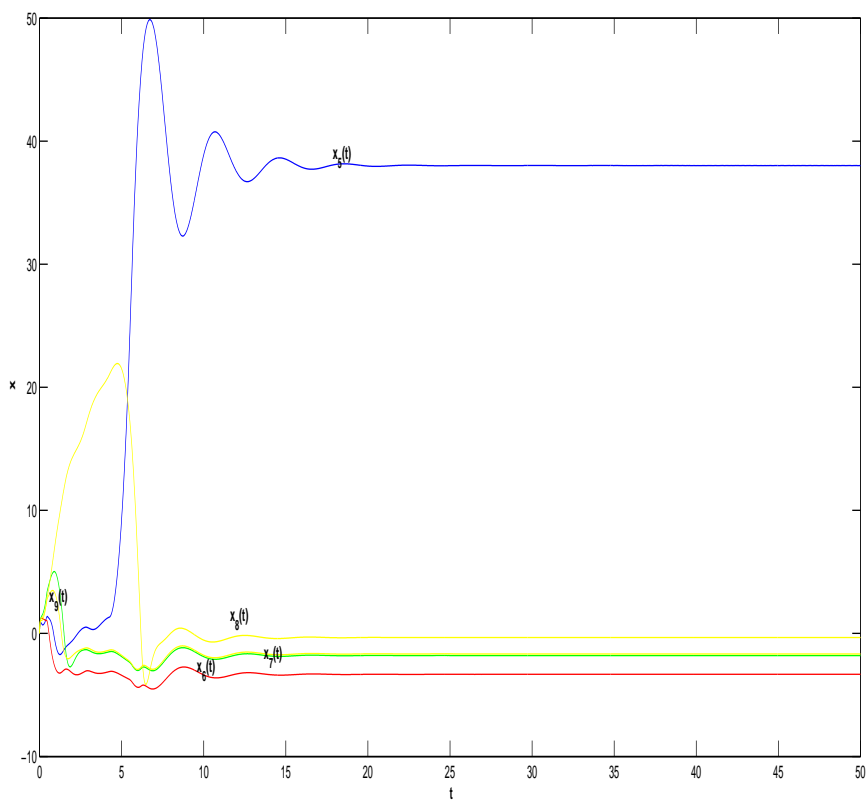
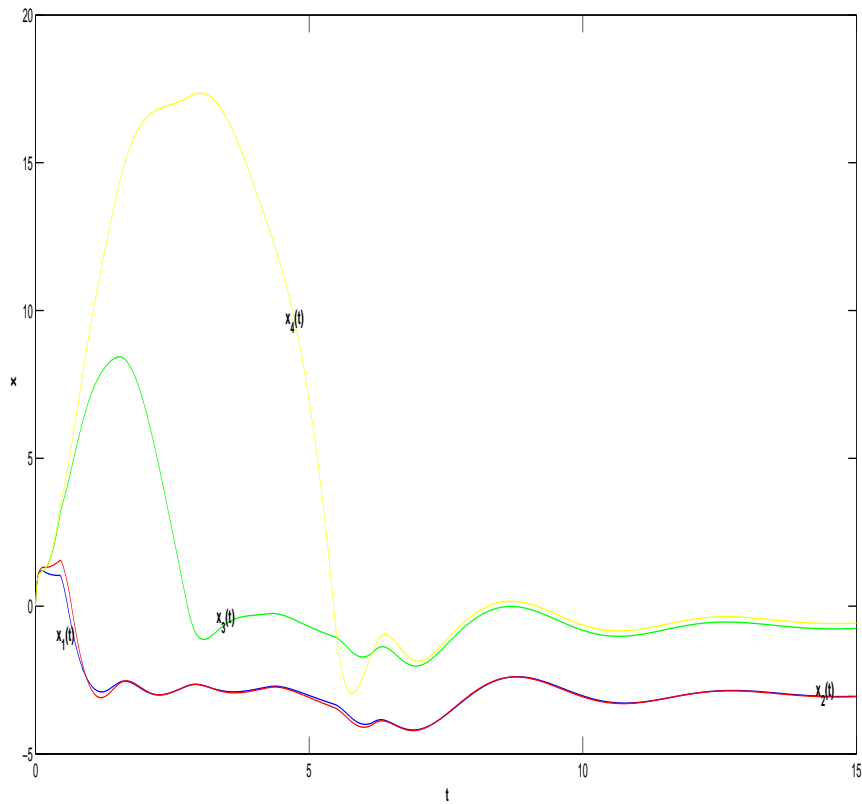
۱.۳.۳ ترسیم مثال ۳.۳ با شبکه‌ی عصبی

حال مثال (۳.۳) وقتی که $\mu = 1$ است را برای مسئله‌های (۱۱.۳) و (۱۲.۳) با استفاده از شبکه‌های عصبی ترسیم می‌کنیم. در این صورت جواب بهینه برای مسأله (۱۱.۳) عبارت است از:

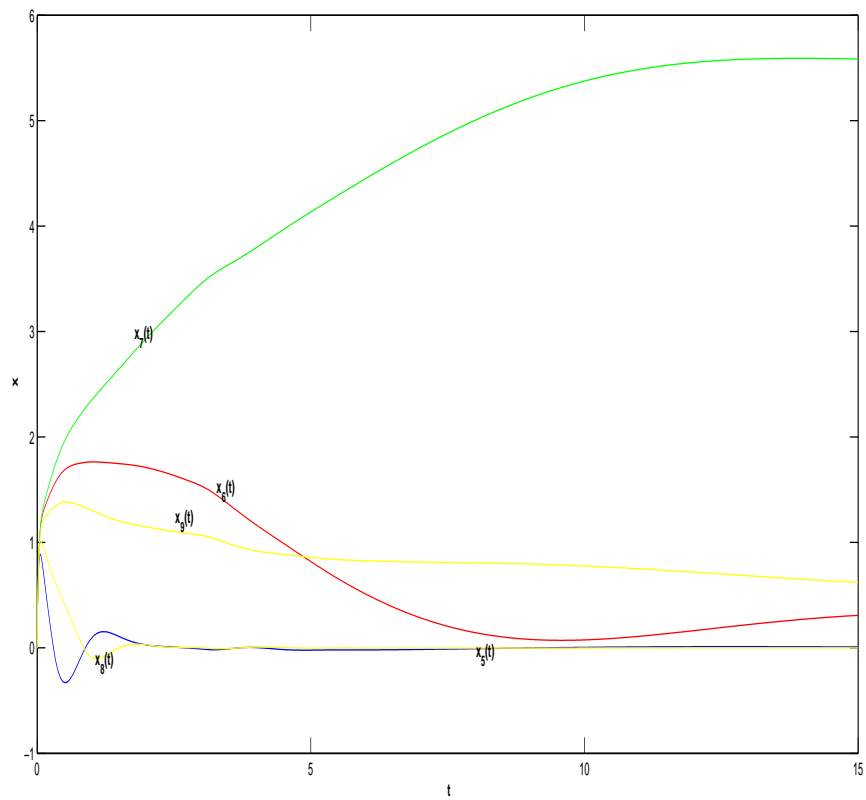
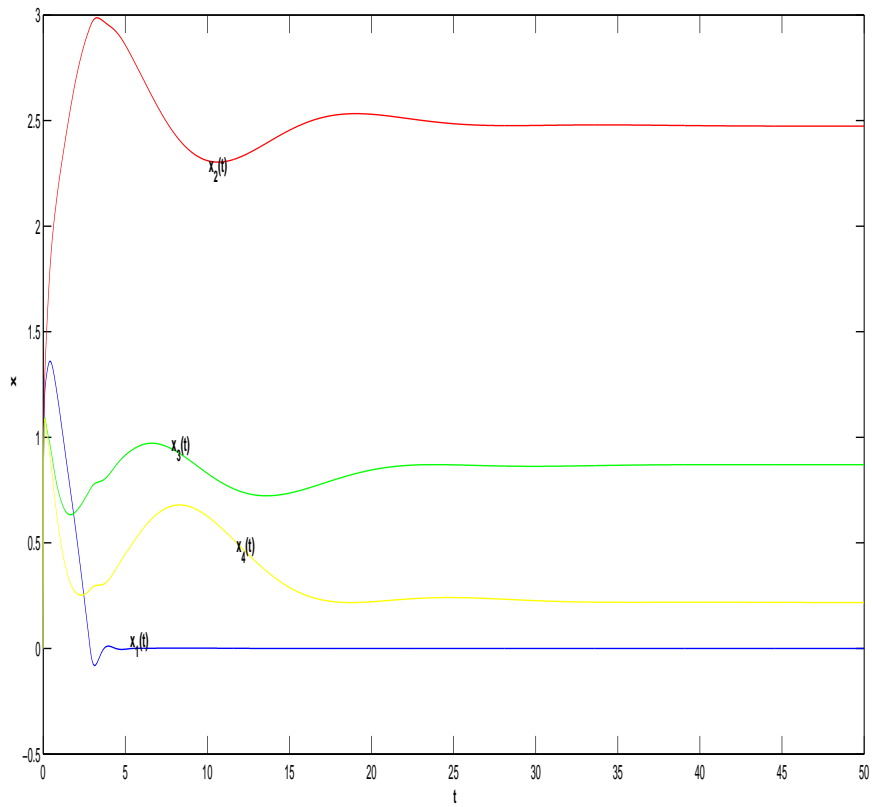
$$x^* = (-2/993, -3/006, -0/695, -0/516, 38/014, -3/326, -1/798, -0/335, -1/665)^T$$

و جواب بهینه برای مسأله (۱۲.۳) عبارت است از:

$$x^* = (-0/0001, 2/5298, 0/8453, 0/2210, -0/0004, 0/3074, 5/5416, 0/0001, 0/5557)^T$$



شکل ۱.۳: مثال مدل (۱۱.۳) با شبکه عصبی وقتی که $\mu = 1$



شکل ۲.۳: مثال مدل (۱۲.۳) با شبکه عصبی وقتی که $\mu = 1$

فصل ۴

مسأله انتخاب سبد سهام مجاز با هزینه‌های معاملاتی با یک روش شبکه‌ی عصبی

۱.۴ مقدمه

در سال ۱۹۵۲، مارکوویتز [۳۲] کار پیش‌گام خود که مبنای تحلیلی از سبد سهام‌های مدرن بود را منتشر کرد. مدل میانگین-واریانس مارکوویتز به‌عنوان مبنایی برای توسعه نظریه مالی مدرن در پنج دهه گذشته بوده است. با تلاش مستمر محققان مختلف توسعه کار اولیه مارکوویتز ادامه یافت [۵۴، ۴۱، ۵]. تحقیقات بیش‌تر در مورد انتخاب سبد سهام‌ها در مراجع [۴۲، ۶۶، ۶۴] یافت می‌شود. اخیراً تعدادی از محققان مسائل مربوط به انتخاب سبد انتزاعی فازی را مورد بررسی قرار داده‌اند. تاناکا و همکاران [۳۵] مدل انتخاب سبد سهام‌ها براساس احتمال فازی را پیشنهاد کردند که می‌تواند به‌عنوان یک توسعه مدل مارکوویتز در نظر گرفته شود. به دلیل گسترش احتمال به احتمال فازی با استفاده از رویکردهای فازی، دانش کارشناسان و سرمایه‌گذاران می‌تواند به یک مدل انتخاب سبد سهام بهتر منتج شود. ایدا [۵۱]، لای و همکاران [۴۴] و گیووی و همکاران در [۶۵] مدل‌های برنامه‌ریزی بازه‌ای از انتخاب سبد سهام‌ها را معرفی کردند. ژانگ و نی [۷۶]، ژانگ و همکاران [۷۴] ژانگ و وانگ [۷۵] در مورد انتخاب

سبد کارآمد قابل قبول، که بادر نظرگرفتن بازده مورد انتظار و ریسک مورد انتظار است بحث کردند. به‌خوبی شناخته شده است که هزینه‌های معاملاتی یکی از نگرانی‌های اصلی مدیران بورس است. این تأثیر مهمی در مورد بهینه‌سازی سبد سهام‌ها و فرکانس زمان تعادل سبدها دارد. آرنوت و واگنر [۶۲] این را یافتند که نادیده گرفتن هزینه‌های معاملاتی منجر به سبد نا کارآمد می‌شود. تجزیه و تحلیل تجربی انجام شده توسط یاشیموتو [۱۰] این واقعیت را تأیید کرد. علاوه بر این هنگامی که برخی از محدودیت‌های واقعی‌تر مانند هزینه‌های معاملاتی، محدودیت‌های کران‌دار، نقدینگی محدودیت‌ها، محدودیت‌های معاملاتی و محدودیت‌های قدرت‌مندی در نظر گرفته شده است، مشکل یک مسئله برنامه ریزی انتخاب سبد سهام پیچیده‌تر می‌شود و الگوریتم‌های بهینه‌سازی سنتی موفق به یافتن راه حل مؤثر مطلوب نمی‌شود. بنابراین بسیاری از محققان با استفاده از الگوریتم‌های تکاملی، مشکلات پیچیده محدودیت‌های مجموعه را حل می‌کنند. به‌عنوان مثال ژیا و همکاران [۸۵] یک الگوریتم ژنتیک برای مشکل انتخاب سبدها با نظم بازده مورد انتظار طراحی کرده‌اند. چانگ و همکاران [۷۱] از الگوریتم‌های تکاملی براساس الگوریتم‌های ژنتیکی، جستجوی تابو و برای مسأله میانگین واریانس استفاده کرده‌اند. اسپرانزا [۵۰] از الگوریتم‌های تکاملی مبتنی بر برنامه نویسی خطی برای یک مدل بهینه‌سازی سبدها با هزینه‌های معاملاتی، حداقل واحدهای معامله و محدودیت‌های حداقل موجودی استفاده کرد. فرناندز و گومز [۶] الگوریتم‌های تکاملی را براساس شبکه‌های عصبی برای مدل استاندارد واریانس مارکوویتز با توانایی و محدودیت‌های محدود ارائه دادند.

۲.۴ مدل انتخاب سبد سهام با هزینه‌های معاملاتی

انتخاب یک مسئله‌ی پرتفوی با n دارایی را بررسی می‌کنیم. فرض کنید r_j بازده مورد انتظار از دارایی j ام و x_j مقدار سرمایه گذاری شده از دارایی j ، $j = 1, 2, \dots, n$ باشد. برای توصیف بهتر، مجموعه‌ی $x = (x_1, x_2, \dots, x_n)^T$ ، $r = (r_1, r_2, \dots, r_n)^T$ و $e = (1, 1, \dots, 1)^T$ داریم. آن‌گاه بازده مورد انتظار و واریانس مربوط به پرتفوی $x = (x_1, x_2, \dots, x_n)^T$ وجود دارد، که به‌ترتیب عبارت‌اند را:

$$E(r) = r^T x, \quad D(r) = x^T V x.$$

که در آن $V = (\sigma_{ij})_{n \times n}$ ماتریس کوواریانس بازده مورد انتظار است. مدل میانگین-واریانس مارکوویتز از مسئله‌ی انتخاب سبد سهام با برنامه ریزی درجه‌ی دوم زیر بیان می‌شود:

$$\begin{aligned} & \text{minimize} && x^T V x \\ & \text{subject to} && r^T x = r_0, \\ & && e^T x = 1, \\ & && x \geq 0. \end{aligned} \quad (1.4)$$

به منظور استفاده از مدل (۱.۴) در یک مسئله‌ی سرمایه‌گذاری کاربردی، به تخمین r و $V = (\sigma_{ij})_{n \times n}$ نیاز داریم. فرض کنید داده‌های مشاهده شده روی بازده‌ها از دارایی‌های روی m دوره داده شود. در زمان k ($k = 1, 2, \dots, m$)، نوع n از بازده‌ها که به صورت یک بردار $r_k = (r_{k1}, r_{k2}, \dots, r_{kn})^T$ نشان داده می‌شود، وجود دارد. بازده مورد انتظار $\bar{r} = (\bar{r}_1, \bar{r}_2, \dots, \bar{r}_n)^T$ و کوواریانس $\bar{V} = (\bar{\sigma}_{ij})_{n \times n}$ از [۷۴، ۷۶] به صورت زیر به دست می‌آیند:

$$\bar{r}_j = \frac{\sum_{k=1}^m (h_{kj} r_{kj})}{\sum_{k=1}^m h_{kj}}, \quad j = 1, 2, \dots, n, \quad (2.4)$$

و

$$\bar{\sigma}_{ij} = \frac{\sum_{k=1}^m (r_{ki} - \bar{r}_i)(r_{kj} - \bar{r}_j)h_{kij}}{\sum_{k=1}^m h_{kij}}, \quad i, j = 1, 2, \dots, n, \quad (3.4)$$

که در آن h_{kj} یک درجه‌ی امکان منعکس کننده میزان شباهت بین وضعیت آینده از دارایی j و k امین نمونه‌ی ارائه شده توسط کارشناسان است، h_{kij} یک درجه‌ی امکان^۱ منعکس کننده میزان شباهت بین رابطه وضعیت آینده از دارایی i و دارایی j و k امین نمونه‌ی ارائه شده توسط کارشناسان است.

چون r_j ، $j = 1, 2, \dots, n$ به وسیله‌ی عوامل نامشخص تحت تأثیر قرار می‌گیرد، به درستی بازده مورد انتظار و دارایی‌های ریسکی قابل پیش بینی نیست. در [۷۴]، [۷۶] و [۷۵]، بازده و کوواریانس متوسط شدنی به ترتیب به صورت زیر تعریف می‌شوند:

$$r_j^* = \bar{r}_j + \phi_j, \quad \phi_{jl} \leq \phi_j \leq \phi_{jh}, \quad j = 1, 2, \dots, n, \quad (4.4)$$

و

$$\sigma_{ij}^* = \bar{\sigma}_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ijl} \leq \varepsilon_{ij} \leq \varepsilon_{ijh}, \quad j = 1, 2, \dots, n, \quad (5.4)$$

که در آن ϕ_j خطای قابل قبول برای \bar{r}_j است، ϕ_{jl} و ϕ_{jh} کران پایین و بالا از ϕ_j هستند، ε_{ij} خطای قابل قبول برای $\bar{\sigma}_{ij}$ است، ε_{ijl} و ε_{ijh} کران پایین و بالا از ε_{ij} هستند. برای توصیف بهتر از نمادهای زیر استفاده می‌کنیم:

$$a = (\bar{r}_1, \bar{r}_2, \dots, \bar{r}_n)^T, \quad \Sigma = (\bar{\sigma}_{ij})_{n \times n},$$

^۱Possibility grade

$$\Phi = (\phi_1, \phi_2, \dots, \phi_n)^T, \quad V_\varepsilon = (\varepsilon_{ij})_{n \times n},$$

$$\Phi_l = (\phi_{1l}, \phi_{2l}, \dots, \phi_{nl})^T, \quad \Phi_h = (\phi_{1h}, \phi_{2h}, \dots, \phi_{nh})^T,$$

$$V_{\varepsilon l} = (\varepsilon_{ijl})_{n \times n}, \quad V_{\varepsilon h} = (\varepsilon_{ijh})_{n \times n},$$

$a + \Phi$ را بردار بازده‌شدنی می‌نامیم. $\Sigma + V_\varepsilon$ ماتریس کوواریانس‌شدنی می‌نامیم. بنابراین، بازده‌شدنی و کوواریانس وابسته به پرتفوی x با روابط زیر تعریف می‌شوند:

$$\hat{r} = (a + \Phi)^T x, \quad (6.4)$$

و

$$\hat{\sigma}^2 = x^T (\Sigma + V_\varepsilon) x. \quad (7.4)$$

برای هر $x \geq 0$ بازده‌شدنی و کوواریانس وابسته به پرتفوی x با روابط زیر بیان می‌شوند:

$$(a + \Phi_l)^T x \leq \hat{r} \leq (a + \Phi_h)^T x,$$

و

$$x^T (\Sigma + V_{\varepsilon l}) x \leq \hat{\sigma}^2 \leq x^T (\Sigma + V_{\varepsilon h}) x.$$

حال مسئله‌ی انتخاب پرتفوی قابل قبول را با هزینه‌های معاملاتی در نظر می‌گیریم. در این‌جا فرض می‌کنیم هزینه‌های معاملاتی یک تابع V شکل است که از تفاوت بین پرتفوی جدید $x = (x_1, x_2, \dots, x_n)$ و پرتفوی موجود $x_0 = (x_1^0, x_2^0, \dots, x_n^0)$ به‌دست آید. منظور این است که هزینه‌های معاملاتی برای i دارایی i دارای $c_i = k_i |x_i - x_i^0|$ است. هزینه‌های معاملاتی کل برابر با $\sum_{i=1}^n c_i = \sum_{i=1}^n k_i |x_i - x_i^0|$ است. بنابراین، بازده‌شدنی پس از پرداخت هزینه‌های معاملاتی به‌صورت زیر است:

$$(a + \Phi)^T x - \sum_{i=1}^n k_i |x_i - x_i^0|.$$

علاوه‌براین، می‌دانیم که هر وزن کوچکی از یک دارایی اثر محسوسی روی بازده‌ی پرتفوی نخواهد داشت، اما هزینه‌های اجرایی و نظارت افزایش خواهد یافت. به‌طور مشابه، وزن خیلی بالا در هر دارایی باعث خواهد شد سرمایه‌گذارها از ریسک زیاد آسیب ببینند. بنابراین قیدهای کراندار باید وقتی که به‌طور عملی سبد اوراق بهادار تشکیل می‌دهیم در نظر گرفته شود. از این‌رو، مسئله‌ی انتخاب پرتفوی قابل قبول با هزینه‌های معاملاتی^۲ و قیدهای کراندار^۳

²Transaction costs

³Bounded constraints

می تواند به شرح زیر بیان شود:

$$\begin{aligned} & \text{minimize} && x^T(\Sigma + V_\varepsilon)x \\ & \text{subject to} && (a + \Phi)^T x - \sum_{i=1}^n k_i |x_i - x_i^\circ| \geq r_\circ, \\ & && l \leq x \leq u, \\ & && e^T x = 1, \end{aligned} \tag{۸.۴}$$

که در آن $u = (u_1, u_2, \dots, u_n)^T$ ، $l = (l_1, l_2, \dots, l_n)^T$ ، و u_i و l_i مینیمم و ماکسیمم سهم سرمایه‌گذاری در دارایی i ، $i = 1, 2, \dots, n$ ، هستند.

اگر $(\Phi, V_\varepsilon) = (\Phi_h, V_{\varepsilon l})$ باشد آن‌گاه (۸.۴) را می‌توان به صورت زیر نوشت:

$$\begin{aligned} & \text{minimize} && x^T(\Sigma + V_{\varepsilon l})x \\ & \text{subject to} && (a + \Phi_h)^T x - \sum_{i=1}^n k_i |x_i - x_i^\circ| \geq r_\circ, \\ & && l \leq x \leq u, \\ & && e^T x = 1, \end{aligned} \tag{۹.۴}$$

اگر $(\Phi, V_\varepsilon) = (\Phi_l, V_{\varepsilon h})$ باشد آن‌گاه (۸.۴) را می‌توان به صورت زیر نوشت:

$$\begin{aligned} & \text{minimize} && x^T(\Sigma + V_{\varepsilon h})x \\ & \text{subject to} && (a + \Phi_l)^T x - \sum_{i=1}^n k_i |x_i - x_i^\circ| \geq r_\circ, \\ & && l \leq x \leq u, \\ & && e^T x = 1, \end{aligned} \tag{۱۰.۴}$$

جواب بهینه مسائل (۹.۴) و (۱۰.۴) بیش‌ترین و کم‌ترین پرتفوی قابل قبول نامیده می‌شود. مسئله‌ی (۹.۴) به این معنی است که سرمایه‌گذار بازده و ریسک را خوش‌بینانه برآورد می‌کند. مسئله‌ی (۱۰.۴) به این معنی است که سرمایه‌گذار بازده و ریسک را بدبینانه برآورد می‌کند. مسئله‌ی (۸.۴) طرحی است که سرمایه‌گذار پرتفویی را انتخاب می‌کند که نه خیلی خوش‌بینانه و نه خیلی بدبینانه است.

۳.۴ مثال عددی

برای نشان دادن ابزار و رویکردهای پیشنهاد شده، ما یک مثال عددی با مارکوویتز [۳۱] را معرفی می‌کنیم. در جدول‌های (۱.۴) و (۲.۴) ستون (۱۰-۲) به ترتیب نشان دهنده‌ی تأمین داده‌های تنباکو آمریکایی، A.T.&T.، فولاد ایالات متحده، ماشین عمومی، اتچسون & توپکا

۴۲ مسأله انتخاب سبد سهام مجاز با هزینه‌های معاملاتی با یک روش شبکه‌ی عصبی

& سانتا فا، کوکا-کولا، بوردن، فیرستون و شارون استیل است. در این مثال، فرض می‌کنیم x°, l, u به ترتیب برابر باشند با: $x^\circ = (0/1, 0/4, 0, 0, 0, 0, 0/3, 0/2, 0)$ ، $l = (0, 0, 0, 0, 0, 0, 0, 0, 0)$ ، $u = (0/2, 0/2, 0/3, 0/3, 0/35, 0/40, 0/2, 0/25, 0/3)$ ، باشند و درجه‌ی امکان h_k تعریف می‌شود با:

$$h_{kij} = h_{ki} = h_k = 0/1 + 0/3 \frac{k-1}{17}, \quad k = 1, 2, \dots, 18.$$

$\Phi_l, \Phi_h, V_\epsilon = 0$ معین وجود دارند با:

$$\Phi_h = (0/014, 0/014, 0/039, 0/041, 0/047, 0/01, 0/027, 0/043, 0/028)^T, \quad \Phi_l = -\Phi_h$$

علاوه بر این، برای روش شبکه‌های عصبی از مجموعه پارامترهای زیر تنظیم می‌شوند: حداکثر تکرار ۱۰۰۰ و مقدار جمعیت ۲۰ نفر است، حداکثر سرعت ۰/۴ و حداقل سرعت ۰/۴- است. با استفاده از (۲.۴) $a, a + \Phi_l, a + \Phi_h$ به شرح زیر تعریف می‌شوند:

$$a = (0/07099, 0/07012, 0/19645, 0/20610, 0/23390, 0/05317, 0/13619, 0/21603, 0/14386)^T,$$

$$a + \Phi_h = (0/08499, 0/08412, 0/23545, 0/24710, 0/28090, 0/06317, 0/16319, 0/25903, 0/17186)^T,$$

$$a + \Phi_l = (0/05699, 0/05612, 0/15745, 0/16510, 0/18690, 0/04317, 0/10919, 0/17303, 0/11586)^T.$$

با استفاده از (۳.۴) ماتریس کوواریانس \bar{V} به شرح زیر بیان شود:

$$\bar{V} = \begin{pmatrix} 0/04042 & 0/01593 & 0/01933 & 0/03277 & 0/01046 & 0/02665 & 0/0183 & 0/02816 & 0/02077 \\ 0/01593 & 0/01119 & 0/01701 & 0/01875 & 0/00771 & 0/00705 & 0/00912 & 0/02106 & 0/01518 \\ 0/01933 & 0/01701 & 0/09852 & 0/06287 & 0/04545 & 0/00694 & 0/0035 & 0/07752 & 0/03114 \\ 0/03277 & 0/01875 & 0/06287 & 0/07954 & 0/04465 & 0/02072 & 0/00904 & 0/07894 & 0/01964 \\ 0/01046 & 0/00771 & 0/04545 & 0/04465 & 0/09713 & 0/00585 & 0/01096 & 0/086 & 0/0285 \\ 0/02665 & 0/00705 & 0/00694 & 0/02072 & 0/00585 & 0/0405 & 0/00953 & 0/01952 & 0/01095 \\ 0/0183 & 0/00912 & 0/0035 & 0/00904 & 0/01096 & 0/00953 & 0/01986 & 0/01784 & 0/00079 \\ 0/02816 & 0/02106 & 0/07752 & 0/07894 & 0/086 & 0/01952 & 0/01784 & 0/13343 & 0/03455 \\ 0/02077 & 0/01518 & 0/03114 & 0/01964 & 0/0285 & 0/01095 & 0/00079 & 0/03455 & 0/06211 \end{pmatrix}$$

ما برای مسئله‌ی انتخاب سبد سهام پیشنهادی از روش شبکه‌های عصبی استفاده می‌کنیم. با توجه به نرخ بازده‌های مختلف r_0 نتایج در جداول (۱.۴) و (۲.۴) زیر بیان شده‌اند. در جداول (۱.۴) و (۲.۴) نتیجه‌های زیر را می‌توانیم به‌دست آوریم. هرچه بازده بالاتر باشد ریسک بالاتر است که این نتیجه با تئوری سرمایه‌گذاری سازگار است. علاوه بر این، برای بازده‌های مشابه، سبد سهام کارا بالاتر با خطای مجاز^۴ ریسک کم‌تری از سبد سهام کارا پایینی مجاز دارد.

با استفاده از بالاترین و پایین‌ترین سبد سهام مدل‌های (۹.۴) و (۱۰.۴)، ما دو مرز سبد سهام کارا با خطای مجاز در شکل (۱.۴) را به‌دست آوردیم.

⁴Admissible error

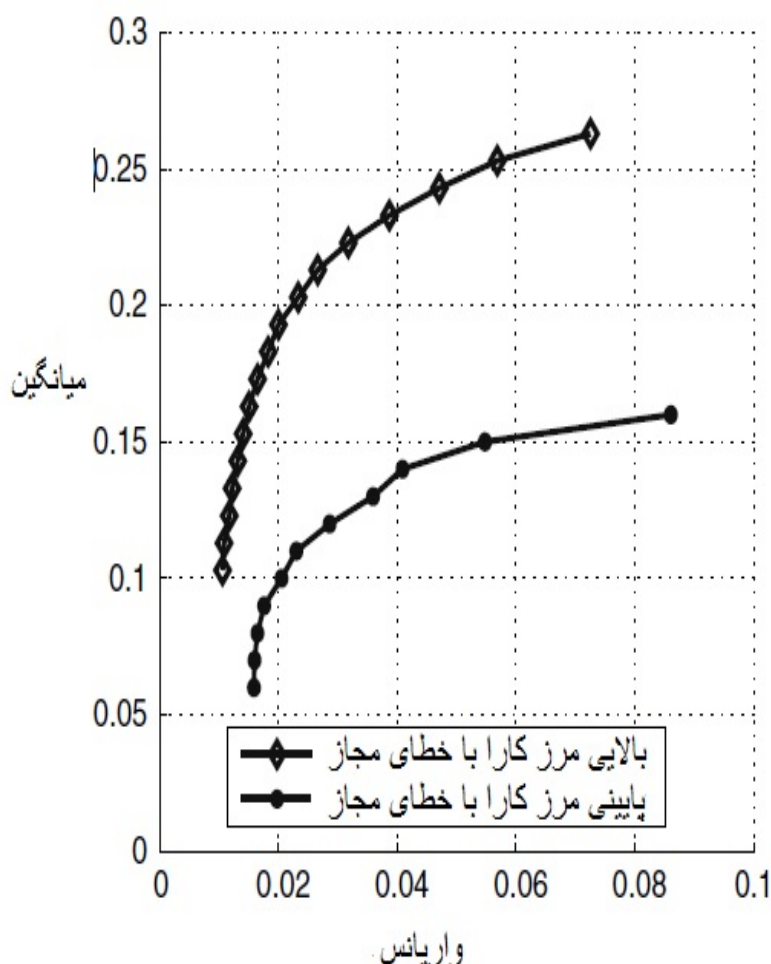
جدول ۱.۴: بالاترین سبد سهام کارایی مجاز

<i>Return</i>	<i>Am.T.</i>	<i>A.T.&T.</i>	<i>U.S.S.</i>	<i>G.M.</i>	<i>A.T.&S.</i>	<i>C.C.</i>	<i>Bdn.</i>	<i>Frstn.</i>	<i>S.S.</i>	<i>Risk</i>
۰/۰۵	۰/۰۵۶۳	۰/۲۰۰۰	۰/۰۳۸۳	۰	۰/۱۰۰۴	۰/۲۹۰۰	۰/۲۰۰۰	۰	۰/۱۱۵۰	۰/۰۱۶۱
۰/۱۰	۰/۰۶۹۸	۰/۲۰۰۰	۰/۰۵۴۵	۰	۰/۰۹۲۶	۰/۲۸۸۱	۰/۲۰۰۰	۰	۰/۰۹۴۸	۰/۰۱۶۱
۰/۱۵	۰/۰۳۶۸	۰/۲۰۰۰	۰/۱۱۱۶	۰/۰۱۶۸	۰/۱۵۳۴	۰/۱۲۲۲	۰/۲۰۰۰	۰	۰/۱۵۹۰	۰/۰۱۹۰

جدول ۲.۴: پایین‌ترین سبد سهام کارایی مجاز

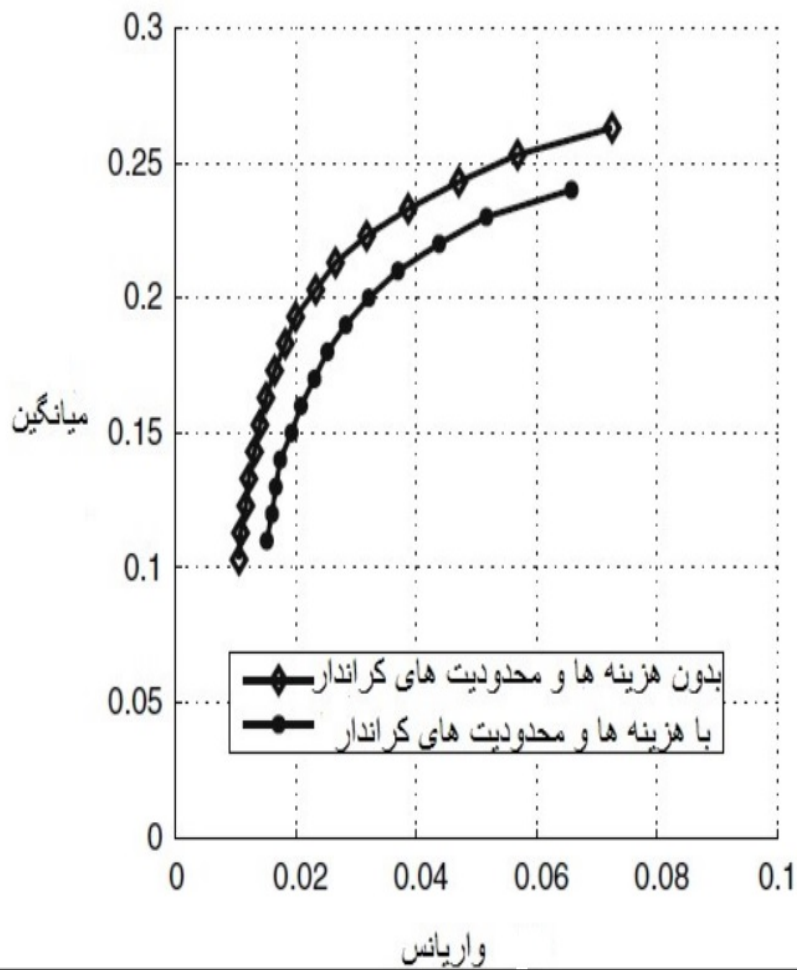
<i>Return</i>	<i>Am.T.</i>	<i>A.T.&T.</i>	<i>U.S.S.</i>	<i>G.M.</i>	<i>A.T.&S.</i>	<i>C.G.</i>	<i>Bdm.</i>	<i>Frstm.</i>	<i>S.S.</i>	<i>Risk</i>
۰/۰۵	۰/۰۴۱۱	۰/۲۰۰۰	۰/۰۵۲۵	۰	۰/۰۶۲۲	۰/۳۰۱۶	۰/۲۰۰۰	۰	۰/۱۴۲۵	۰/۰۱۶۱
۰/۱۰	۰	۰/۲۰۰۰	۰/۰۶۷۵	۰/۰۸۳۵	۰/۱۷۷۲	۰/۱۴۶۸	۰/۲۰۰۰	۰/۰۰۶۵	۰/۱۱۸۴	۰/۰۱۹۷
۰/۱۵	۰	۰	۰/۲۲۴۳	۰/۰۷۲۳	۰/۳۵۰۰	۰	۰/۱۷۲۳	۰/۱۸۱۱	۰	۰/۰۵۵۹

در شکل (۱.۴) مشاهده می‌شود که گستره‌ی بازده‌ی بالاترین سبد سهام با خطای مجاز همیشه بزرگ‌تر از پایین‌ترین سبد سهام با خطای مجاز است. این واقعیت ناشی از مفهوم بالاترین و پایین‌ترین سبد اوراق بهادار با خطای مجاز است. پایین‌ترین و بالاترین سبد اوراق بهادار با خطای مجاز را می‌توان به‌عنوان دو نظر افراطی که یک نقش مرجع برای یک سرمایه‌گذار ایفا می‌کنند در نظر گرفت. برای نشان دادن اینکه هزینه‌های معاملاتی



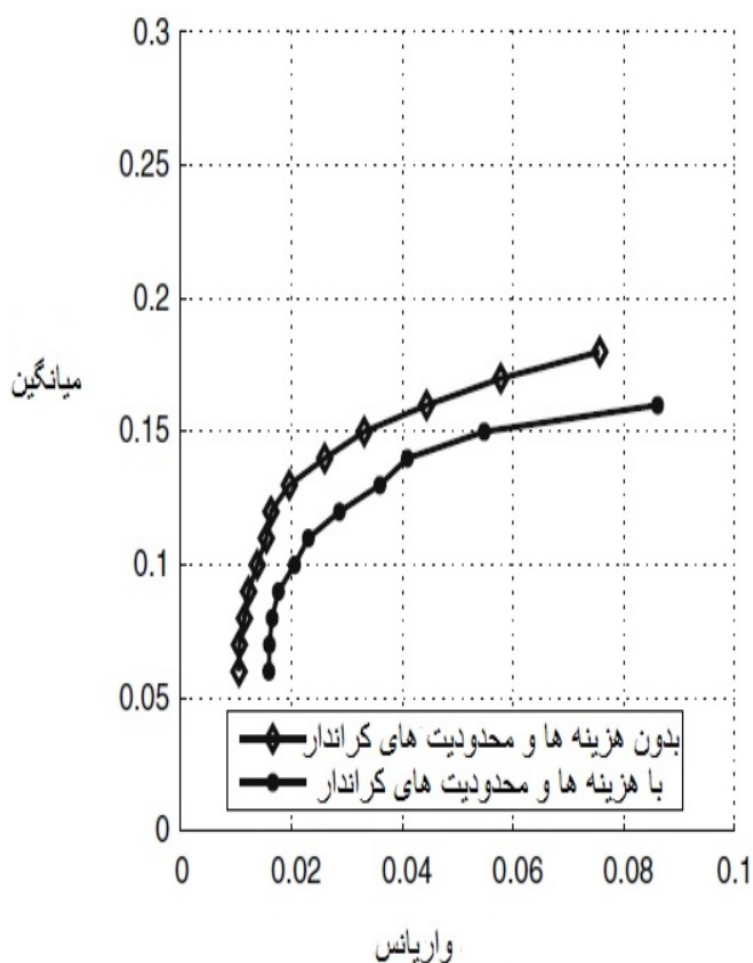
شکل ۱.۴: نمودار بالایی و پایینی مرزهای کارا با در نظر گرفتن خطای قابل قبول و هزینه‌های معاملاتی و محدودیت‌های کراندار

و محدودیت‌های کراندار تأثیر زیادی بر انتخاب سبد سهام کارا دارند، بالاترین و پایین‌ترین مرزهای کارایی سبد سهام با و بدون هزینه‌های معاملاتی و محدودیت‌های کراندار که به ترتیب در شکل (۲.۴) و (۳.۴) نشان داده شده است. می‌توانیم از شکل (۲.۴) و (۳.۴) دریابیم که برای بالاترین سبد سهام کارایی مجاز یا برای پایین‌ترین سبد سهام کارا با خطای مجاز، منحنی با هزینه‌های معاملاتی و محدودیت‌های کراندار پایینی از منحنی بدون این محدودیت‌ها را است. گفتیم که، گستره‌ی بازده سبد سهام کارا با خطای مجاز بدون هزینه‌های معاملاتی و محدودیت‌های کراندار بزرگ‌تر از با هزینه‌های معاملاتی و محدودیت‌های کراندار است.



شکل ۲.۴: مرزهای کارا بالایی تحت محدودیت‌های مختلف

دلیل این موضوع این است که این محدودیت‌ها تأثیر عمیقی روی وزن‌دهی بهینه یک دارایی هنگام حداقل کردن ریسک دارند، لذا پرتفویهای محدود شده کاملاً تحت غلبه پرتفویهای محدود نشده قرار می‌گیرند.

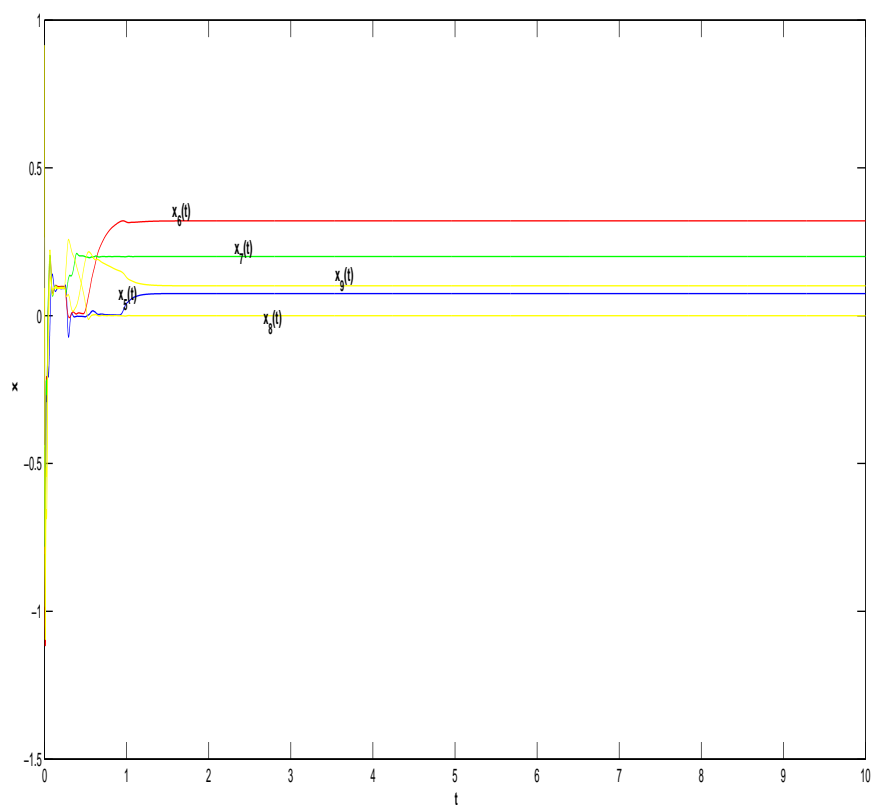
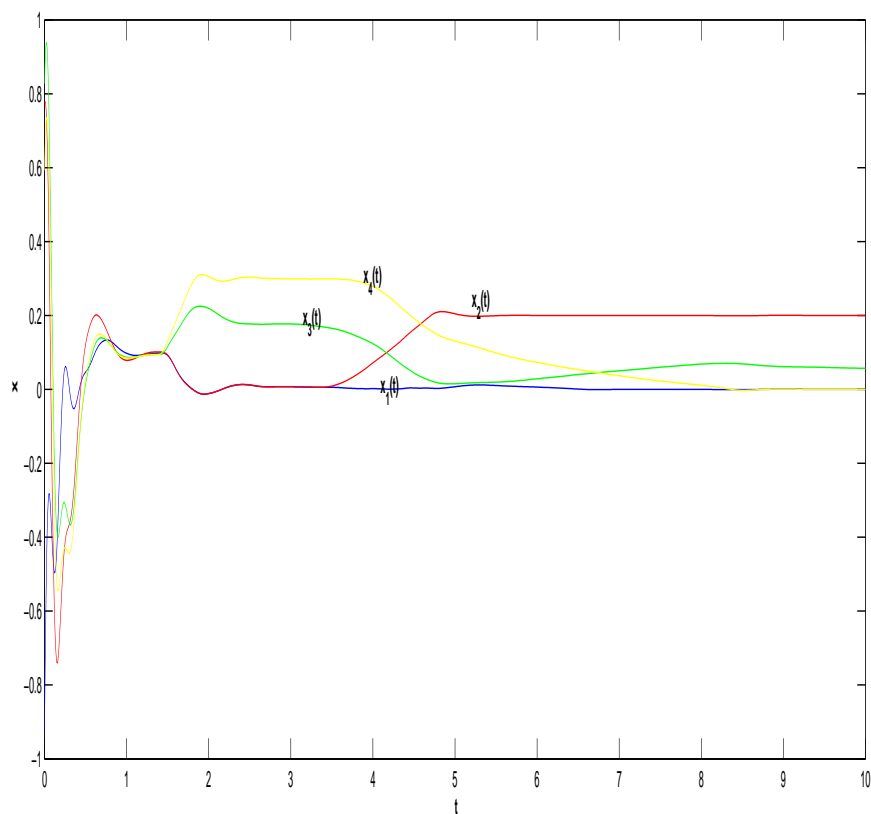


شکل ۳.۴: مرزهای کارا پایینی تحت محدودیت‌های مختلف

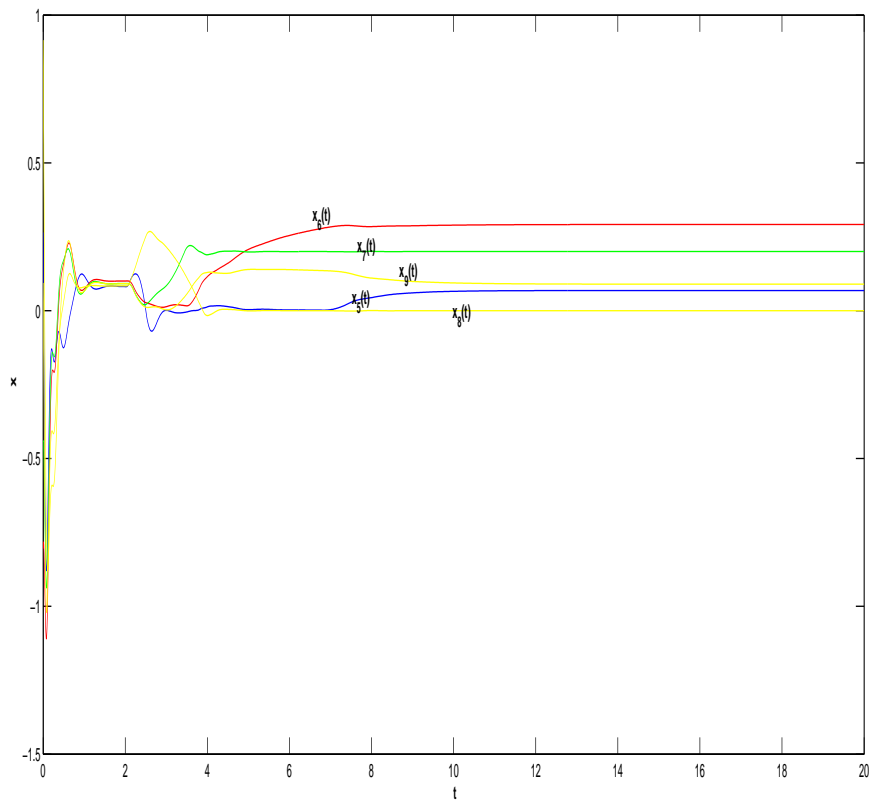
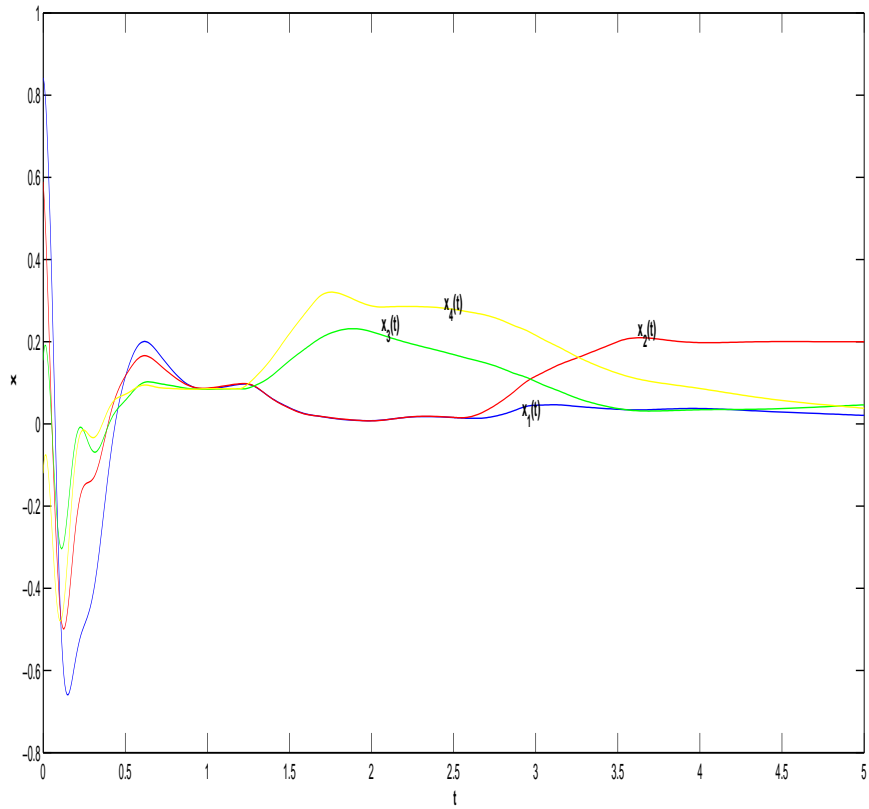
۱.۳.۴ ترسیم مثال ۳.۴ با شبکه‌ی عصبی

حل مثال (۳.۴) وقتی که r_0 به ترتیب برابر با $0.05, 0.1, 0.15$ است را برای مسئله‌های (۹.۴) و (۱۰.۴) با استفاده از شبکه‌های عصبی ترسیم می‌کنیم. جواب‌ها در جدول (۱.۴) و (۲.۴) نشان داده شده است.

۴۸ مسأله انتخاب سبد سهام مجاز با هزینه‌های معاملاتی با یک روش شبکه‌ی عصبی

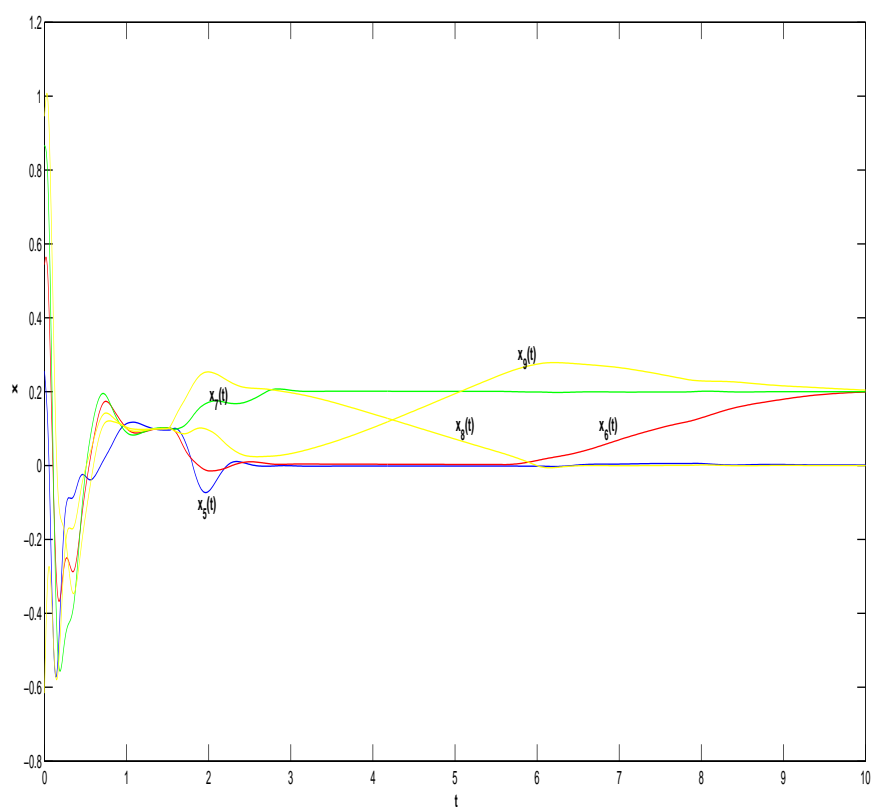
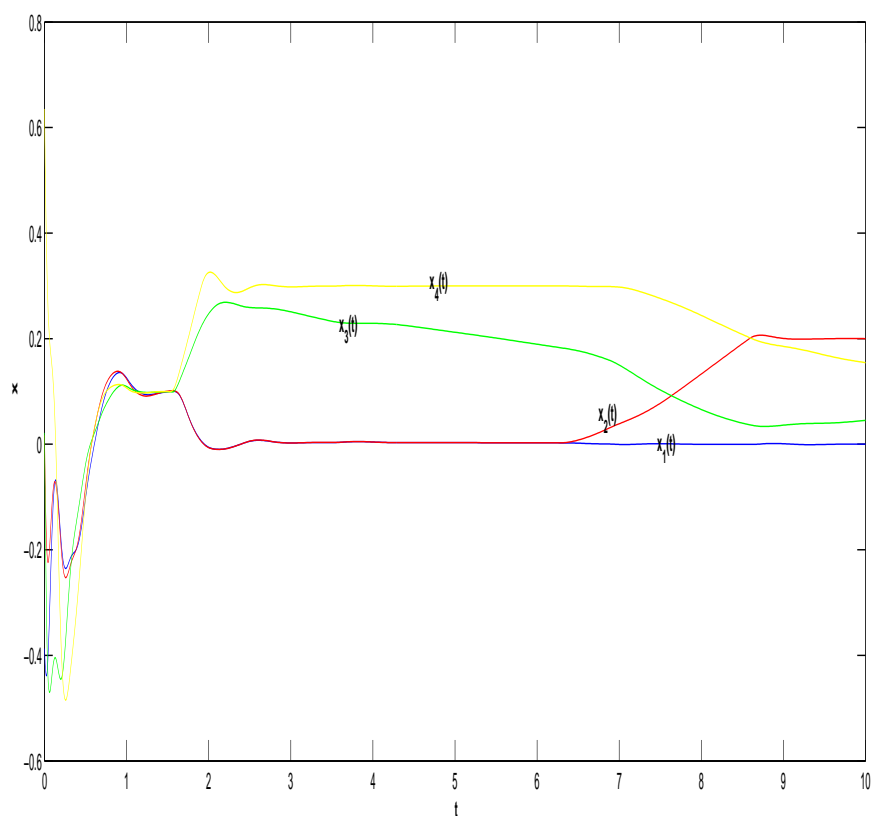


شکل ۴.۴: مثال مدل (۹.۴) با شبکه عصبی وقتی که $r_0 = 0.05$

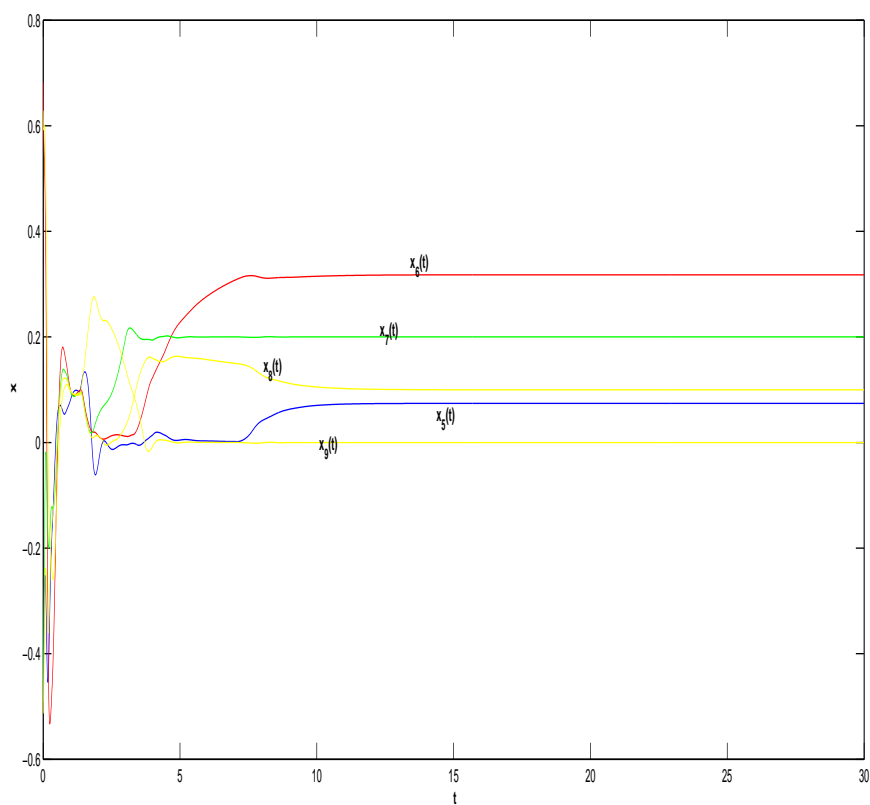
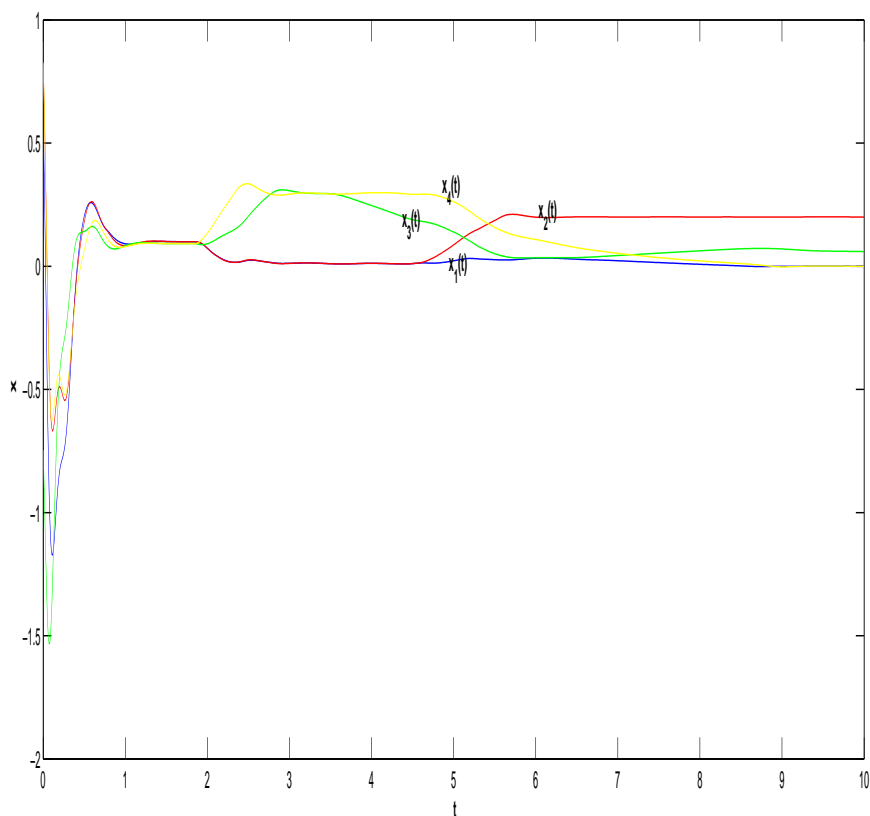


شکل ۵.۴: مثال مدل (۱۰.۴) با شبکه عصبی وقتی که $r_0 = 0.05$

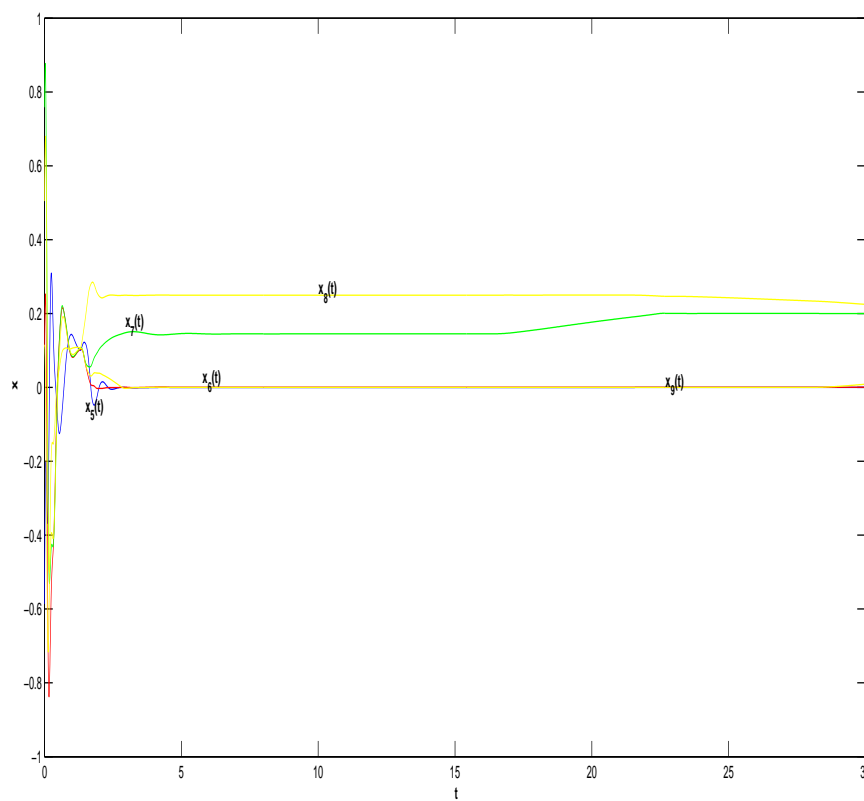
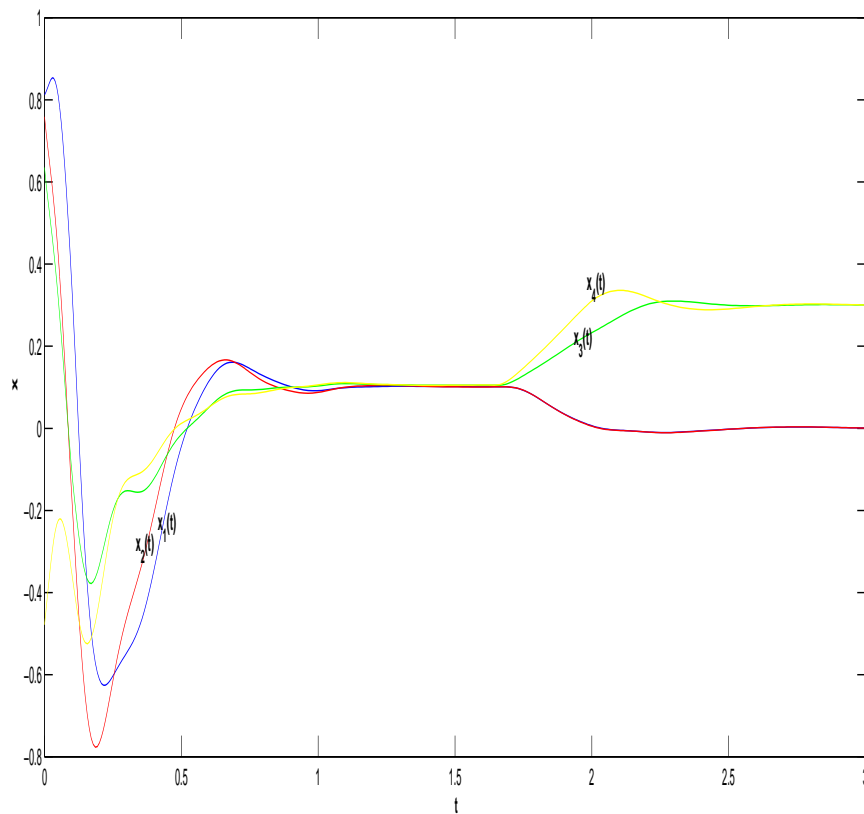
۵۰ مسأله انتخاب سبد سهام مجاز با هزینه‌های معاملاتی با یک روش شبکه‌ی عصبی



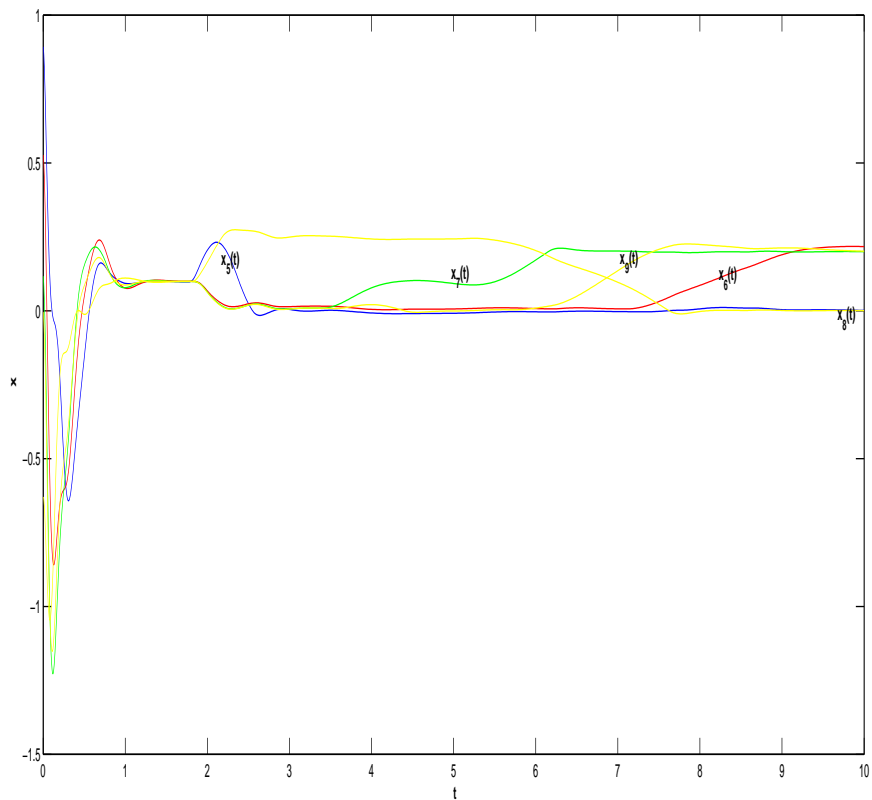
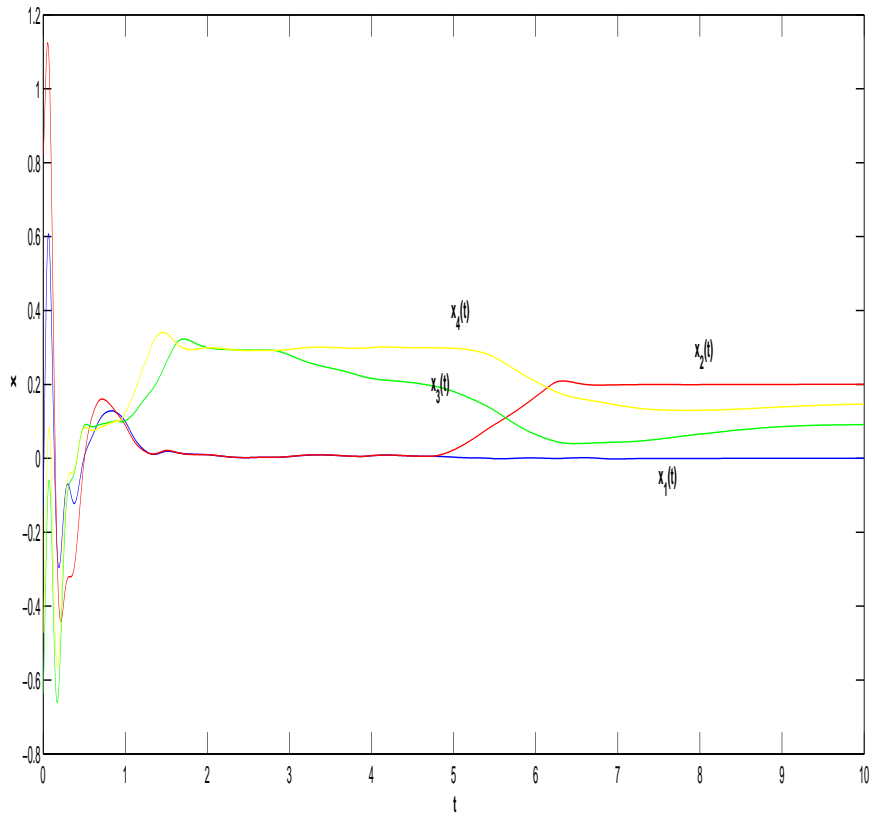
شکل ۶.۴: مثال مدل (۹.۴) با شبکه عصبی وقتی که $r_0 = 0/1$



شکل ۷.۴: مثال مدل (۱۰.۴) با شبکه عصبی وقتی که $r_0 = 1/10$



شکل ۸.۴: مثال مدل (۹.۴) با شبکه عصبی وقتی که $r_0 = 0.15$



شکل ۹.۴: مثال مدل (۱۰.۴) با شبکه عصبی وقتی که $r_0 = 0.15$

مراجع

- [۱] آذر. ضمیری. شبکه‌های عصبی مصنوعی و بهینه‌سازی. پایان‌نامه کارشناسی‌ارشد، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، آبان ۱۳۸۷.
- [۲] محمد. مقدس. حل رده‌ای از مسائل بهینه‌سازی غیرخطی با استفاده از شبکه‌های عصبی. پایان‌نامه کارشناسی‌ارشد، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، تابستان ۱۳۹۲.
- [۳] محمد. مقدس، سهراب. عفتی و محمد. اسحاق‌نژاد. یک شبکه عصبی برای حل مسائل بهینه‌سازی غیرخطی نامحدب. در ششمین کنفرانس بین‌المللی انجمن ایرانی تحقیق در عملیات، تهران، ایران، اردیبهشت ۱۳۹۲.
- [4] A. Bouzerdoum and T. R. Pattison. Neural network for quadratic optimization with bound constraints. *IEEE Transactions on Neural Networks*, 4(2): 293–304, (1993).
- [5] A.F. Perold, Large-scale portfolio optimization, *Management Science*, 30(1984) 1143-1160.
- [6] A. Fernández, S. Gómez, Portfolio selection using neural networks, *Computers & Operations Research*, 34 (2007) 1177-1191.
- [7] A. R. Nazemi, A dynamical model for solving degenerate quadratic minimax problems with constraints, *Journal of Computational and Applied Mathematics* 236 (2011) 1282-1295.
- [8] A. R. Nazemi, A dynamic system model for solving convex nonlinear optimization problems, *Communications in Nonlinear Science and Numerical Simulation* 17 (2012) 1696-1705.

- [9] A. R. Nazemi, S. Effati, An application of a merit function for solving convex programming problems, *Computers and Industrial Engineering*, 66 (2013) 212-221.
- [10] A. Yoshimoto, The mean-variance approach to portfolio optimization subject to transaction costs, *Journal of the Operational Research Society, Japan* 39 (1996) 99-117.
- [11] B. Widrow. Generalization and information storage in networks of adaline 'neurons'. in *Self Organizing Systems*, pp. 435–461. Spartan Books, (1959).
- [12] B. Widrow and M. E. Hoff. Adapting switching circuits. in 1960 ire wescon Convention Record, no. 4, pp. 96–104, New York, (1960). ire.
- [13] B. Widrow and M. A. Lehr. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. *Proceedings of the IEEE. Institute of Electrical and Electronics Engineers*, 78(9): 1415–1442, (1990).
- [14] B. Widrow and R. Winter. Neural nets for adaptive filtering and adaptive pattern recognition. *Computer*, 21(3): 25–39, (1988).
- [15] Chen Y. and Jiao Y. and Gao X. and Leung K. (2001), "A new gradient-based neural network for solving linear and quadratic programming problems", *IEEE Transactions on Neural Networks*, No. 12(5), 1074–1083.
- [16] C.Y. Maa and M. A. Schanblatt. A two-phase optimization neural network. *IEEE Transactions on Neural Networks*, 3(6): 1003–1009, (1992).
- [17] D. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. Springer, 3rd ed., (2008).
- [18] D. O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. John Wiley Sons, Inc., New York, (1949).
- [19] D. E. Rumelhart, G. E. Hinton and R. J. Williams. Learning representations by backpropagating errors. *Nature*, 323: 533–536, (1986).

- [20] Effati S. and Baymani M. (2005), "A new nonlinear neural network for solving convex nonlinear programming problems", *Applied Mathematics and Computation*, 168, 1370 -1379.
- [21] Effati S, Ghomashi A, Nazemi AR. Application of projection neural network in solving convex programming problems. *Applied Mathematics and Computation*, 2007;188:1103–14.
- [22] Effati S. and Nazemi A.R. (2006), Neural network models and its application for solving linear and quadratic programming problems, *Applied Mathematics and Computation*, 172(1), 305-331.
- [23] Effati S. and Ranjbar M. (2011), "A novel neural network for solving quadratic programming problems", *Applied Mathematics and Computation*, No. 35(4), 1688–1695.
- [24] F. Rosenblatt. *Neurocomputing: foundations of research*. chap. The perception: a probabilistic model for information storage and organization in the brain, pp. 89–114. MIT Press, Cambridge, MA, USA, (1988).
- [25] F. Rosenblatt. *Principles of neurodynamics: perceptrons and the theory of brain mechanisms*. Report (Cornell Aeronautical Laboratory). Spartan Books, (1962).
- [26] G.P. Szegö , *Portfolio Theory*, Academic Press, New York, 1980.
- [27] Gao X. and Liao L.Z. (2010), A new one-layer neural network for linear and quadratic programming. *IEEE Transactions on Neural Networks*, 21(6), 918-929.
- [28] Ghasabi-Oskoei H. and Mahdavi-Amiri N. (2006), An efficient simplified neural network for solving linear and quadratic programming problems. *Applied Mathematics and Computation*, 175(1), 452-464.
- [29] Ghasabi-Oskoei H. and Malek A. and Ahmadi A. (2007), Novel artificial neural network with simulation aspects for solving linear and quadratic programming problems. *Computers and Mathematics with Applications*, 53(9), 1439-1454
- [30] H. K. Khalil. *Nonlinear Systems*. Prentice Hall PTR, 3rd ed., (2002).

- [31] H.M. Markowitz, Portfolio Selection: Efficient Diversification of Investments, Wiley, New York, 1959.
- [32] H. Markowitz, Portfolio selection, The Journal of Finance, 7 (1952) 77-91.
- [33] H. Markowitz, Portfolio Selection: Efficient Diversification of Investments, Wiley, New York, 1959.
- [34] H. Markowitz, Analysis in Portfolio Choice and Capital Markets, Basil Blackwell, Oxford, 1987.
- [35] H. Tanaka, P.J. Guo, I.B. Türksen, Portfolio selection based on fuzzy probabilities and possibility distributions, Fuzzy Sets and Systems 111 (2000) 387–397.
- [36] Hu X.L. (2009), "Applications of the general projection neural network in solving extended linear-quadratic programming problems with linear constraints", Neuro-computing, No. 72, 1131–1137.
- [37] Hu X. and Wang J. (2008), An improved dual neural network for solving a class of quadratic programming problems and its k -winners-take-all application. IEEE Transactions on Neural Networks, 19(12), 2022-2031.
- [38] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79(8): 2554–2558, (1982).
- [39] J. J. Hopfield and D. W. Tank. "neural" computation of decisions in optimization problems. Biological Cybernetics, 52(3): 141–152, (1985).
- [40] J. M. Ortega and W. C. Rheinboldt. Iterative Solution of Nonlinear Equations in Several Variables. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, (2000).
- [41] J.S. Pang, A new efficient algorithm for a class of portfolio selection problems, Operational Research, 28 (1980) 754-767.
- [42] J. Vörös, Portfolio analysis-An analytic derivation of the efficient portfolio frontier, European Journal of Operational Research, 203 (1986) 294-300.

- [43] J. Zabczyk. Mathematical control theory: an introduction. Systems control. Birkhäuser, (1992).
- [44] K.K. Lai, S.Y. Wang, J.P. Xu, S.S. Zhu, Y. Fang, A class of linear interval programming problems and its application to portfolio selection, IEEE Transactions on Fuzzy Systems, 10 (2002) 698-704.
- [45] Kennedy, M. P., Chua, L. O.: Neural Networks for Nonlinear Programming, IEEE Transactions on Circuits and Systems, 35, (1988), 554-562.
- [46] Liu S. and Wang J. (2006), A simplified dual neural network for quadratic programming with its KWTA application. IEEE Transactions on Neural Networks, 17(6), 1500-1510.
- [47] Liu Q. and Wang J. (2008, November), A one-layer recurrent neural network for non-smooth convex optimization subject to linear equality constraints. In International Conference on Neural Information Processing (1003-1010). Springer Berlin Heidelberg.
- [48] Liu Q. and Wang J. (2008), A one-layer recurrent neural network with a discontinuous hard-limiting activation function for quadratic programming. IEEE Transactions on Neural Networks, 19(4), 558-570.
- [49] Malek A. and Alipour M. (2007), Numerical solution for linear and quadratic programming problems using a recurrent neural network. Applied Mathematics and Computation, 192(1), 27-39.
- [50] M.G. Speranza, A heuristic algorithm for a portfolio optimization model applied to the Milan stock market, Computer Operational Research, 23 (1996) 433-441.
- [51] M. Ida, Portfolio selection problem with interval coefficients, Applied Mathematics Letters, 16 (2003) 709-713.
- [52] M.J. Best, R.R. Grauer, The efficient set mathematics when mean-variance problems are subject to general linear constraints, Journal of Economics and Business 42 (1990) 105-120.

-
-
- [53] M.J. Best, H. Jaroslava, The efficient frontier for bounded assets, *Mathematical Methods of Operations Research* 52 (2000) 195–212.
- [54] M.J. Best, J. Hlouskova, The efficient frontier for bounded assets, *Mathematical Methods of Operations Research*, 52 (2000) 195-212.
- [55] M. L. Minsky and S. Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, (1972).
- [56] M. Moghaddas and S. Effati. A novel recurrent neural network based on ncp function for solving convex nonlinear optimization problems. in *The 5th Iranian Conference on Applied Mathematics*, Hamadan, Iran, September (2013). Bu-Ali Sina University.
- [57] M. P. Kennedy and L. O. Chua, Neural networks for nonlinear programming, *IEEE Transactions on Circuits and Systems*, 35 (1988) 554-562.
- [58] M.S. Bazaraa, H.D. Sherali and C. M. Shetty. *Nonlinear Programming*. John Wiley Sons, Inc., 3rd ed. , (2005).
- [59] N. Kawadai, H. Konno, Solving large scale mean-variance models with dense non-factorable covariance matrices, *Journal of Operations Research Society of Japan* 44 (2001) 251–260.
- [60] P. Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Ph.D. thesis, Harvard University, Cambridge, MA, (1974).
- [61] R.C. Merton, An analytic derivation of the efficient frontier, *Journal of Finance and Quantitative Analysis* 9 (1972) 1851–1872.
- [62] R.D. Arnott, W.H. Wanger, The measurement and control of trading costs, *Financial Analysts Journal*, 46 (1990) 73-80.
- [63] Rodriguez-Vazquez, A., Dominguez-Castro, R., Rueda, A., Huertas, J.L., Sanchez-Sinencio, E., 1990. Nonlinear switched-capacitor neural networks for optimization problems. *IEEE Transactions on Biomedical Circuits and Systems*, 37, 384–397.

- [64] S. Fang, A mean-variance analysis of arbitrage portfolios, *Physica A* 375 (2007) 625-632.
- [65] S. Giove, S. Funari, C. Nardelli, An interval portfolio selection problems based on regret function, *European Journal of Operational Research*, 170 (2006) 253-264.
- [66] S. Pafka, I. Kondor, Estimated correlation matrices and portfolio optimization, *Physica A* 343 (2004) 623-634.
- [67] S. Zhang and A.G. Constantinides. Lagrange programming neural networks. *IEEE Trans. Circuits Syst. II: Analog Digital. Signal Process*, 39(7): 441–452, (1992).
- [68] S. Effati, and M. Baymani. A new nonlinear neural network for solving convex nonlinear programming problems. *Applied Mathematics and Computation*, 168(2): 1370 – 1379, (2005).
- [69] S. Effati, A. Ghomashi and M. Abbasi. A novel recurrent neural network for solving mlcps and its application to linear and quadratic programming. *Asia-Pacific Journal of Operational Research*, 28(4): 523–541, November (2011).
- [70] Tank, D. W., Hopfield, J. J.: Simple ‘Neural’ Optimization Network: An A/D Converter, Signal Decision Circuit and A Linear Programming Circuit, *IEEE Transactions on Biomedical Circuits and Systems*, 33, (1986), 533-541.
- [71] T.J. Chang, N. Meade, J.E. Beasley, Y.M. Sharaiha, Heuristics for cardinality constrained portfolio optimization, *Computer & Operation Research*, 27 (2000) 1271-1302.
- [72] V.K. Chopra, W.T. Ziemba, The effect of errors in means, variances and covariances on optimal portfolio choice, *The Journal of Portfolio Management* (Winter) (1991) 6–11.
- [73] W. E. Lillo, M. H. Loh, S. Hui and S. H. Zak, On solving constrained optimization problems with neural networks: A penalty method approach, *IEEE Transactions on Neural Networks*, 4 (1993) 931-939.

-
-
- [74] W.G. Zhang, W.A. Liu, Y.L. Wang, On admissible efficient portfolio selection problem: Models and algorithms, *Applied Mathematics Computation*, 176 (2006) 208-218.
- [75] W.G. Zhang, Y.L. Wang, An analytic derivation of admissible efficient frontier with borrowing, *European Journal of Operational Research*, 184 (2008) 229-243.
- [76] W.G. Zhang, Z.K. Nie, On admissible efficient portfolio selection problem, *Applied Mathematics and Computation* 159 (2004) 357–371.
- [77] W. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4): 115–133, (1943).
- [78] W. Sharpe, *Portfolio Theory and Capital Markets*, McGraw-Hill, New York, 1970.
- [79] Wu X. and Xia Y. and Li J. and Chen W. (1996), "A high-performance neural network for solving linear and quadratic programming problems", *IEEE Transactions on Neural Networks*, No. 7(3), 643–651.
- [80] Xia Y. (1996), A new neural network for solving linear and quadratic programming problems. *IEEE Transactions on Neural Networks*, 7(6), 1544-1548.
- [81] Xia Y. and Feng G. and Wang J. (2004), A recurrent neural network with exponential convergence for solving convex quadratic program and related linear piecewise equations. *Neural Networks*, 17(7), 1003-1015.
- [82] Xia Y, Leng H, Wang J. A projection neural network and its application to constrained optimization problems. *IEEE Transactions on Biomedical Circuits and Systems*, 2002;49:447–58.
- [83] Xia Y, Wang J. A recurrent neural network for nonlinear convex optimization subject to nonlinear inequality constraints. *IEEE Transactions on Biomedical Circuits and Systems*, 2004;51:447–58.
- [84] Yang Y. and Cao J. (2006), Solving quadratic programming problems by delayed projection neural network. *IEEE Transactions on Neural Networks*, 17(6), 1630-1634.

- [85] Y.S. Xia, B. Liu, S.Y. Wang, K.K. Lai, A model for portfolio selection with order of expected returns, *Computers & Operations Research*, 27 (2000) 409-422.
- [86] Y. Xia, G. Feng and J. Wang. A novel recurrent neural network for solving nonlinear optimization problems with inequality constraints. *IEEE Transactions on Neural Networks*, 19(8): 1340–1353, (2008).
- [87] Y. Xia and J. Wang. A recurrent neural network for solving nonlinear convex programs subject to linear constraints. *IEEE Transactions on Neural Networks*, 16(2): 379–386, (2005).
- [88] Y. Yang, J. Cao, A feedback neural network for solving convex constraint optimization problems, *Applied Mathematics and Computation* 201 (2008) 340-350.
- [89] Zhang Y. and Li Z. (2009), Zhang neural network for online solution of time-varying convex quadratic program subject to time-varying linear-equality constraints. *Physics Letters A*, 373(18), 1639-1643.
- [90] Zhang Y. and Wang J. (2002), A dual neural network for convex quadratic programming subject to linear equality and inequality constraints. *Physics Letters A*, 298(4), 271-278.

Abstract

In this thesis, a numerical method based on neural networks is presented to solve a class of the financial optimization problems with transaction costs. To do this, we first write the optimality conditions for the optimization problem. Then we design a neural network model related to it. We prove that the equilibrium point of the neural network is the main optimal solution. By presenting an appropriate Lyapunov function, stability and convergence properties of the proposed neural network are stated. For this purpose some definitions of financial concepts in mathematics and some concepts of optimization are presented, then a novel gradient based neural network model is presented, in continuation an analytic derivation of admissible efficient frontier with borrowing by a neural network model is studied and in end the admissible portfolio selection problem with transaction costs with a neural network model is considered.

Keywords: Portfolio, Optimization, Transaction Costs, Neural Networks, Value at Risk, Conditional Value at Risk.



Shahrood University of Technology

Faculty of Mathematical Sciences

MSc Thesis in: Financial Mathematics

**The study of a class of portfolio
optimization problems with transaction costs**

By: Raziye Keyshamsvand

Supervisors

Dr. Alireza Nazemi

Dr. Seyed Mojtaba Mirlohi

Feb. 2019