

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی شاهرود

دانشکده علوم ریاضی  
گروه ریاضی کاربردی

پایان نامه

برای دریافت درجه کارشناسی ارشد در رشته  
ریاضی کاربردی، گرایش تحقیق در عملیات

عنوان

# یک شبکه عصبی کارا برای حل ردهای از مسائل نامساوی‌های وردشی مخروطی

استاد راهنما

دکتر علیرضا ناظمی

دانشجو

عطیه سابتقی

۱۳۹۳

## تقدیم به پدر بزرگوار و مادر مهربانم

به پاس عاطفه سرشار و گرمای امید بخش وجودشان که در این سردترین روزگار ان بهترین پشتیبان من است

به پاس زحماتی که در تمام سال های تحصیل متقبل شده اند و پوزش تکل رنج دوری در این دوره تحصیلی در روزهایی که باید در کنارشان می بودم و نبودم

## تقدیم به همسر عزیزم

او که در سایه همیاری و همدلی اش به این منظور نائل شدم.

## خدایا...

به من زیستنی عطا کن که در لحظه مرگ، بر بی‌ثمری لحظه‌ای که برای زیستن گذشته است، حسرت نخورم و مردنی عطا کن که بر بیهودگیش، سوگوار نباشم. بگذار تا ان را، خود انتخاب کنم، اما آنچه‌ان که تو دوست می‌داری.

تو می‌دانی و همه می‌دانند شکنجه دیدن، بخاطر تو، زندانی کشیدن بخاطر تو و رنج بردن به پای تو تنها لذت زندگی من است، از شادی توست که من در دل می‌خندم، از امید رهایی توست که برق امید در چشمان خسته‌ام می‌درخشد و از خوشبختی توست که هوای پاک سعادت را در ریه‌هایم احساس می‌کنم. نمی‌توانم خوب حرف بزنم. نیروی شگفتی را که در زیر کلمات ساده و جمله‌های ضعیف و افتاده پنهان کرده‌ام دریاب، دریاب.

تو می‌دانی و همه می‌دانند که زندگی از تحمیل لبخندی بر لبان من، از آوردن برق امیدی در نگاه من، از برانگیختن موج شغفی در دل من، عاجز است.

تو، چگونه زیستن را به من بیاموز، چگونه مردن را خود خواهم آموخت.

به من توفیق تلاش در شکست، صبر در نومیدی، رفتن بی همراه، جهاد بی سلاح، کار بی پاداش، فداکاری در سکوت، دین بی دنیا، مذهب بی عوام، عظمت بی نام، خدمت بی نان، ایمان بی ریا، خوبی بی نمود، گستاخی بی خامی، قناعت بی غرور، عشق بی هوس، تنهایی در انبوه جمع‌آزیت، و دوست داشتن بی آنکه دوست بداند، روزی کن.<sup>۱</sup>

## سپاس گزارمی...

سپاس خدای را که سخنوران در ستودن او بمانند و شمارندگان، شمردن نعمت‌های او ندانند و کوشندگان، حق او را گزاردن نتوانند؛

از پدر و مادر مهربانم... این دو معلم بزرگوام... که همواره بر کوتاهی و درشتی من، قلم عفو کشیده و کریمانه از کنار غفلت‌هایم گذشته‌اند و در تمام عرصه‌های زندگی یار و یآوری بی چشم‌داشت برای من بوده‌اند؛

از همسر عزیزم... که سایه مهربانی‌اش سایه‌سار زندگیم است؛  
از خانواده محترم... که سهم بسزایی در پیشرفت‌های زندگیم داشته‌اند؛  
از استاد فرهیخته و فرزانه جناب آقای دکتر علیرضا ناظمی که با نکته‌های دل‌ویز و گفته‌های بلند، صحیفه‌های سخن را علم پرور نمود و همواره راهنما و راه‌گشای نگارنده در اتمام و اکمال پایان‌نامه بوده است؛

از اساتید فرزانه و دلسوز جناب آقای دکتر مهدی قوتمند و آقای دکتر محمد مهدی فاتح که زحمت داوری این رساله را متقبل شدند؛

هم‌چنین از دوستان عزیزم که در همه حال همراه و همیارم بوده‌اند از جمله خانوم‌ها: مهشید اخوان نژاد، مریم برزویی، فاطمه سلیمانی، مریم خان‌احمدی، سوسن عشوره، سمیرا سوخت‌سرای، سیمین پویا و مرضیه مرتضایی کمال تشکر و قدردانی را دارم. باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

عطه ساقی  
۱۳۹۳

## تعمدنامه

اینجانب عطیه سابقی دانشجوی کارشناسی ارشد رشته ریاضی کاربردی دانشکده علوم ریاضی دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان یک شبکه عصبی کارا برای حل رده‌ای از مسائل نامساوی‌های وردشی مخروطی، تحت راهنمایی دکتر علیرضا ناظمی متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، تاکنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه شاهرود متعلق دارد، و مقالات مستخرج با نام “ دانشگاه شاهرود “ یا “ Shahrood University “ به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده) شده است، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

عطیه سابقی  
۱۳۹۳

## مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

## چکیده

با پیشرفت فن‌آوری اطلاعات و ارتباطات و توسعه ارتباطات درون سازمانی و بین سازمانی نیاز به استفاده از مدل‌های بهینه‌سازی را برای استفاده منطقی از داده‌ها و اطلاعات فراهم شده گسترش داده است. این مطلب متضمن بزرگ شدن اندازه مسائل بهینه‌سازی که در عمل وجود دارند خواهد بود. در این شرایط لزوم به کارگیری روش‌های کارآمدی که بتوانند با سرعت بالا مسائل بسیار بزرگ را با کیفیت قابل قبول حل کنند بیش از پیش احساس می‌شود.

در چند دهه اخیر روش‌های بهینه‌سازی که بر پایه رویکرد هوش مصنوعی توسعه یافته‌اند، موفقیت‌های چشم‌گیری در حل مؤثر و کارایی مسائل بهینه‌سازی به دست آورده‌اند.

روش‌هایی چون الگوریتم ژنتیک، جستجوی ممنوع، شبیه‌سازی تبریدی، شبکه عصبی و... قابلیت‌های خود را در حل مسائل بزرگ عملی به خوبی نشان داده‌اند.

امتیازات ویژه‌ی موجود در شبکه‌های عصبی امکان کاربرد آنها را در حوزه وسیعی از تحقیقات فراهم ساخته است. از جمله آن امتیازات می‌توان به امکان یادگیری و بهبود عملکرد براساس داده‌های ورودی اشاره کرد. همچنین امکان انجام محاسبات به صورت موازی در شبکه‌های عصبی امتیاز دیگری است که با توجه به گسترش سخت‌افزارهای موازی، امکان حل مسائل بسیار بزرگ را توسط این رویکرد ممکن می‌سازد.

در این پایان‌نامه‌ی دو مدل شبکه عصبی بازگشتی برای حل رده‌ای از مسائل بهینه‌سازی ارائه می‌شود. تحلیل وجود یکتایی، پایداری و همگرایی سراسری جواب‌ها مورد بررسی قرار می‌گیرند و عملکرد روش‌های ارائه شده با به کارگیری چند مثال از مسائل نامساوی وردشی مخروطی نشان داده می‌شود. در انتها نتایج کار و پیشنهاداتی برای کارهای آتی ارائه می‌دهیم.

**واژه‌های کلیدی:** شبکه‌های عصبی، پایداری لیپانوف، همگرایی، نامساوی‌های وردشی،

برنامه‌ریزی مخروط مرتبه دوم.

# فهرست مطالب

۵	لیست تصاویر
۱	لیست جداول
۲	۱ مقدمات و مفاهیم اساسی
۲	۱.۱ مقدمه‌ای بر شبکه‌های عصبی
۲	۱.۱.۱ شبکه‌های عصبی بیولوژیکی
۳	۲.۱.۱ شبکه‌های عصبی مصنوعی
۷	۳.۱.۱ تاریخچه تکامل شبکه عصبی
۸	۴.۱.۱ تاریخچه حل مسائل بهینه‌سازی با استفاده از شبکه‌های عصبی
۱۰	۲.۱ مقدمه‌ای بر بهینه‌سازی غیرخطی
۱۲	۳.۱ سیستم‌های دینامیکی
۱۷	۴.۱ نامساوی‌های وردشی
۲۰	۲ کاربرد شبکه‌های عصبی در حل مسائل نامساوی‌های وردشی و مکمل
۲۰	۱.۲ مقدمه
۲۳	۲.۲ مدل شبکه عصبی
۲۵	۳.۲ تحلیل پایداری سیستم دینامیکی
۲۹	۴.۲ تعمیم
۳۰	۵.۲ مثال‌های عددی
۳۸	۳ یک شبکه عصبی لاگرانژی برای حل مسائل نامساوی وردشی مخروطی
۳۸	۱.۳ مقدمه
۳۸	۲.۳ بیان مسأله
۴۰	۳.۳ مدل شبکه عصبی
۴۱	۴.۳ تحلیل همگرایی و پایداری
۴۴	۵.۳ مثال‌های عددی



۵۰	یک شبکه عصبی جدید برای حل مسائل نامساوی‌های وردشی مخروطی و مکمل	۴
۵۰	..... مقدمه	۱.۴
۵۰	..... بیان مسأله	۲.۴
۵۲	..... مدل شبکه عصبی	۳.۴
۵۲	..... تحلیل پایداری و همگرایی	۴.۴
۵۶	..... مثال‌های عددی	۵.۴
۶۳	آ کد مثال‌های فصل ۳	
۶۳	..... مثال (۱.۵.۳)	۱.آ
۶۴	..... مثال (۲.۵.۳)	۲.آ
۶۴	..... مثال (۳.۵.۳)	۳.آ
۶۵	..... مثال (۴.۵.۳)	۴.آ
۶۶	ب کد مثال‌های فصل ۴	
۶۶	..... مثال (۱.۵.۴)	۱.ب
۶۷	..... مثال (۲.۵.۴)	۲.ب
۶۷	..... مثال (۳.۵.۴)	۳.ب
۶۸	..... مثال (۴.۵.۴)	۴.ب
۶۹	مراجع	
۷۵	واژه‌نامه فارسی به انگلیسی	
۷۷	واژه‌نامه انگلیسی به فارسی	

# 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<sup>۱</sup> 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<sup>۲</sup> نام دارند. دندریت‌ها و تنه سلولی، سیگنال‌ها را از نرون‌های مجاور دریافت می‌کنند و از طریق یک لوله باریک به نام آکسون<sup>۳</sup> 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<sup>۴</sup> می‌نامند. یک سیناپس مرکب از پایانه قبل از سیناپسی<sup>۵</sup>، شکاف سیناپسی<sup>۶</sup> و پایانه بعد از سیناپسی<sup>۷</sup> می‌باشد. همه سیگنال‌های جمع‌بندی شده، در تنه نرون ترکیب می‌شوند و اگر وسعت سیگنال‌های ترکیب شده به آستانه نرون برسد، مرحله تحریک شدن<sup>۸</sup> 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

---

<sup>۱</sup>Neuron

<sup>۲</sup>Dendrite

<sup>۳</sup>Axon

<sup>۴</sup>Synapse

<sup>۵</sup>Presynaptic Terminal

<sup>۶</sup>Synaptic Cleft

<sup>۷</sup>Postsynaptic Terminal

<sup>۸</sup>Firing

یک نرون به دیگری انتقال می‌یابد. تعداد بسیار زیادی آکسون از نرون‌های مختلف با دندریته‌های یک نرون به این صورت ارتباط برقرار می‌کنند.

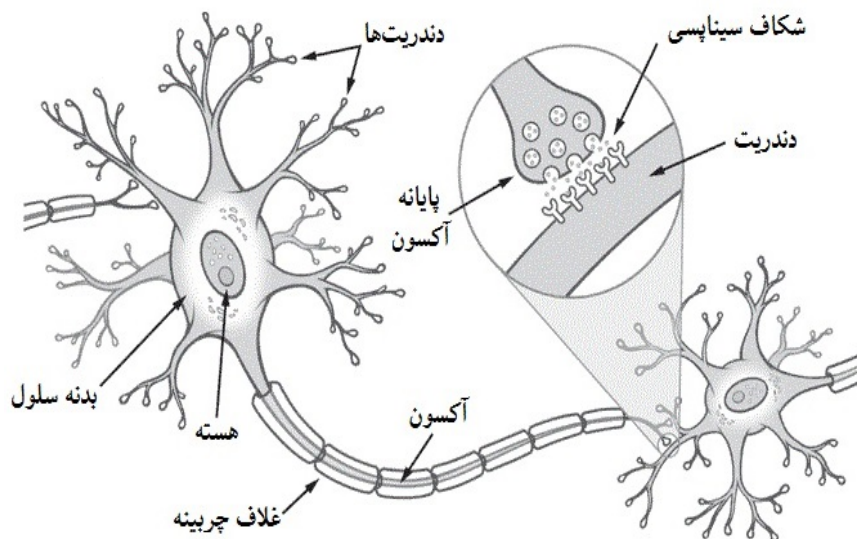
نرون‌ها با توجه به عملکردشان به سه دسته تقسیم می‌شوند:

(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 رابط دیگر 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.



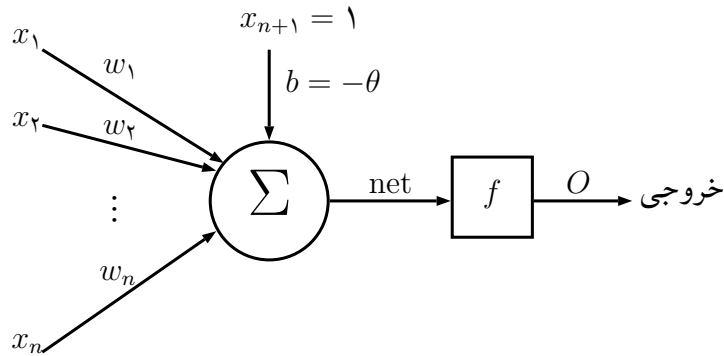
شکل ۱.۱: ساختار یک نرون.

## 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<sup>۹</sup> مدل‌های سخت افزاری یا نرم افزاری الهام گرفته از ساختار و رفتار نرون‌های بیولوژیکی و

<sup>۹</sup>Artificial Neural Networks

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 به‌جای استفاده از عبارت شبکه‌های عصبی مصنوعی به‌طور خلاصه عبارت شبکه‌های عصبی را به‌کار 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 آن را بایاس می‌نامند. سیگنال‌های عصبی  $x_1$  و  $x_2$  و  $\dots$  و  $x_{n+1}$  به ترتیب در  $w_1$  و  $w_2$  و  $\dots$  و  $w_{n+1}$  ضرب شده و مجموع این حاصل ضرب‌ها که در شکل آن را با  $\text{net}$  نشان می‌دهیم، وارد نرون می‌شوند. در این مرحله نرون همانند یک تابع روی  $\text{net}$  عمل می‌کند و حاصل کار خود (برد تابع) را از طریق آکسون‌های طرف راست به نرون‌های دیگر می‌فرستد. این تابع را تابع فعال‌سازی<sup>۱</sup> 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 پنهان می‌تواند از چندین لایه تشکیل گردد البته در شکل ۳.۱ لایه مخفی را تک لایه در نظر گرفته شده است.

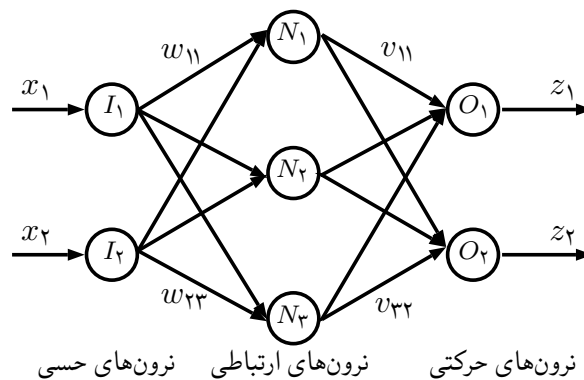
<sup>۱</sup> Activation Function

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: این‌که 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 می‌نامند. 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.



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

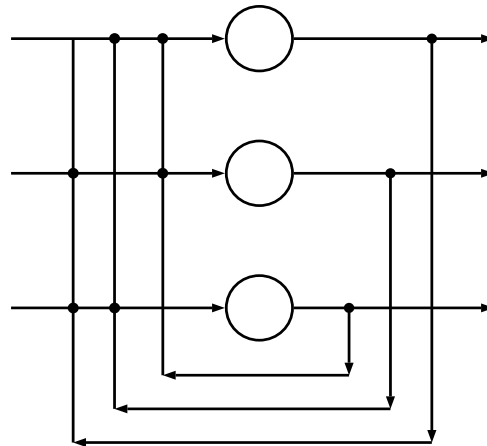
### شبکه‌های عصبی بازگشتی

در شبکه‌های عصبی بازگشتی<sup>۱۱</sup>، حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرونهای 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.

<sup>۱۱</sup>Feedback (Recurrent) Neural Network

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



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<sup>۱۲</sup> به همراه یک متخصص آمار به نام والتر پیتز<sup>۱۳</sup> اولین مدل ریاضی نرون را ارائه دادند [۶۸]. 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<sup>۱۴</sup> [۲۴] 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<sup>۱۵</sup> [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<sup>۱۶</sup> محقق شرکت IBM در سال ۱۹۵۸ با استفاده از کامپیوتر شبیه‌سازی شد. روزنبلات<sup>۱۷</sup> [۲۹، ۳۰] در سال ۱۹۵۸ نرون ساده مک‌کلاچ پیتز را اصلاح کرد و به آن قابلیت یادگیری و سازگاری را اضافه کرد و این نرون را پرسپترون نامید. قانون تعلیم پرسپترون<sup>۱۸</sup> اولین قانون رسمی برای شبکه‌های عصبی است.

برنارد ویدرو<sup>۱۹</sup> و مارسین هاف<sup>۲۰</sup> و همکارانشان یک مدل سه سطحی از نرون‌ها را ابداع کردند که

<sup>۱۲</sup>Warren McCulloch

<sup>۱۳</sup>Walter Pitts

<sup>۱۴</sup>Donald Hebb

<sup>۱۵</sup>Farley and Clark

<sup>۱۶</sup>Nathanial Rochester

<sup>۱۷</sup>Frank Rosenblatt

<sup>۱۸</sup>Perceptron

<sup>۱۹</sup>Bernard Widrow

<sup>۲۰</sup>Marcian Hoff

آدلاین<sup>۲۱</sup> نام‌گذاری شد و مختصر شده حروف ” عنصر خطی تطبیقی<sup>۲۲</sup>” بود [۱۵، ۱۶، ۱۷، ۱۸]. 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<sup>۲۳</sup> 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<sup>۲۴</sup> [۵۳] 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<sup>۲۵</sup> پال ورباس<sup>۲۶</sup> [۵۷] در سال ۱۹۷۴ ارائه شد و بعدها توسط رامل هارت<sup>۲۷</sup> [۲۱] به‌طور مستقل کشف گردید. این الگوریتم از زمان پیدایش به‌طور گسترده به‌عنوان یک الگوریتم آموزش در شبکه‌های عصبی پیش‌خورد<sup>۲۸</sup> 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<sup>۲۹</sup> [۴۰] بود که در سال ۱۹۸۲ منتشر شد. دومین اتفاق، حل مسأله فروشنده دوره‌گرد (TSP) توسط شبکه هاپفیلد بود [۳۸]. 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 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 امیدوارکننده برای حل مسائل بهینه‌سازی با ابعاد بالا استفاده از شبکه‌های عصبی مصنوعی می‌باشد.

در سال ۱۹۸۶ تانک و هاپفیلد [۲۶] 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

<sup>۲۱</sup> Adeline

<sup>۲۲</sup> Adaptive Linear Elements

<sup>۲۳</sup> Madaline

<sup>۲۴</sup> Papert and Minsky

<sup>۲۵</sup> Back Propagation

<sup>۲۶</sup> Paul Werbos

<sup>۲۷</sup> Rumelhart

<sup>۲۸</sup> Feed Forward

<sup>۲۹</sup> John Hopfield



با آن به طور 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 پیاده سازی کردند. هم‌چنین از این شبکه برای حل مسأله فروشنده دوره‌گرد با  $30$  شهر استفاده نمودند. این شبکه دارای نقص‌هایی بود به‌خصوص این‌که نقطه تعادل شبکه در شرایط  $KKT^{30}$  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 چا<sup>۳۱</sup> [۴۹] 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<sup>۳۲</sup> توسط رودریگز-وازکز<sup>۳۳</sup> و همکاران [۱۱] معرفی شد. هم‌چنین ما و شانبلات<sup>۳۴</sup> [۱۹] یک شبکه عصبی دو فازی ارائه کردند که فاز اول شبکه مشابه شبکه کندی و چا بود و در فاز دوم آن مسیر شبکه به جواب دقیق مسأله همگرا می‌شد بنابراین این روش جواب‌های دقیق‌تری نسبت به شبکه کندی و چا 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<sup>۳۵</sup> [۶۷] برمبنای 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. در سال ۱۹۹۳ بوزردوم و پتیسن<sup>۳۶</sup> [۵] شبکه‌ای را برمبنای گرادیان و روش تصویر و مستقل از پارامتر جریمه ابداع کردند که تنها قادر به حل مسائل درجه دوم با متغیرهای کران‌دار بود. این روش در عمل روش کارایی بود اما نمی‌توانست مسائل کلی برنامه‌ریزی خطی و درجه دوم را حل کند.

زیا<sup>۳۷</sup> و همکاران [۷۳، ۷۷، ۷۸، ۷۹] 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 و همکاران [۳، ۴، ۵۴، ۶۳، ۶۴، ۶۵] ارائه شده است.

<sup>۳۰</sup> Karush-Kuhn-Tucker

<sup>۳۱</sup> Kennedy and Chua

<sup>۳۲</sup> Switched-Capacitor

<sup>۳۳</sup> Rodriguez-Vazquez

<sup>۳۴</sup> Maa and Shanblatt

<sup>۳۵</sup> Zhang and Constantinides

<sup>۳۶</sup> Bouzardoum and Pattison

<sup>۳۷</sup> xia

## ۲.۱ مقدمه‌ای بر بهینه‌سازی غیرخطی

تعریف ۱.۲.۱ ([۵۱]). 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:

$$\begin{aligned} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \\ & h_j(x) = 0, \\ & x \in X \subset \mathbb{R}^n, \end{aligned} \quad (1.1)$$

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:

$$\begin{aligned} \min & f(x) \\ \text{s.t.} & g(x) \leq 0, \\ & h(x) = 0, \\ & x \in X \subset \mathbb{R}^n, \end{aligned} \quad (2.1)$$

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<sup>۳۸</sup> و در غیر این صورت یک مسأله بهینه‌سازی نامحدب<sup>۳۹</sup> نامیده می‌شود.

تعریف ۲.۲.۱ ([۵۱]). نقطه  $x$  جواب شدنی<sup>۴۰</sup> مسأله (۲.۱) نامیده می‌شود هرگاه در محدودیت‌های مسأله (۲.۱) صدق کند. نقطه شدنی  $x$  یک نقطه منظم<sup>۴۱</sup> نامیده می‌شود اگر بردارهای گرادیان  $\nabla g_i(x)$ ،  $i \in \{i | g_i(x) = 0\}$  و  $j = 1, \dots, p$ ، مستقل خطی باشند.

<sup>۳۸</sup>Convex Optimization

<sup>۳۹</sup>Nonconvex Optimization

<sup>۴۰</sup>Feasible Solution

<sup>۴۱</sup>Regular Point

قضیه ۳.۲.۱ (شرایط کافی مرتبه دوم<sup>۴۲</sup> [۲۲]). 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

$$\mu \geq 0, \quad (۳.۱)$$

$$\mu^T \mathbf{g}(x^*) = 0, \quad (۴.۱)$$

$$\nabla f(x^*) + \lambda^T \nabla \mathbf{h}(x^*) + \mu^T \nabla \mathbf{g}(x^*) = 0, \quad (۵.۱)$$

و ماتریس هسین<sup>۴۳</sup>

$$\nabla^2 L(x^*, \lambda, \mu) = \nabla^2 f(x^*) + \sum_{i=1}^m \mu_i \nabla^2 g_i(x^*) + \sum_{j=1}^p \lambda_j \nabla^2 h_j(x^*) \quad (۶.۱)$$

بر زیرفضای

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

که در آن

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

معین مثبت باشد آن گاه  $x^*$  یک نقطه مینیمم محلی اکید<sup>۴۴</sup> برای مسأله (۲.۱) است.

## مسأله دوگان لاگرانژ

مسأله برنامه‌ریزی غیرخطی (۱.۱) را در نظر بگیرید، این مسأله را مسأله اولیه (P) نیز می‌نامیم. مسأله دوگان لاگرانژ<sup>۴۵</sup> (D) به صورت زیر تعریف می‌شود:

$$\begin{aligned} \max \quad & \theta(u, v) \\ \text{s.t.} \quad & u \geq 0, \end{aligned} \quad (۷.۱)$$

که در آن

$$\begin{aligned} \theta(u, v) = \inf_x \left\{ f(x) + \sum_{i=1}^m u_i g_i(x) + \sum_{j=1}^p v_j h_j(x), x \in X \right\}, \quad (۸.۱) \\ u = (u_1, u_2, \dots, u_m)^t, \quad v = (v_1, v_2, \dots, v_p)^t, \end{aligned}$$

$u_i$  -امین مؤلفه از  $u$ ، متغیر دوگان<sup>۴۶</sup> یا ضربگر لاگرانژ<sup>۴۷</sup> مربوط به قید  $g_i(x) \leq 0$  می‌باشد و  $v_j$  -امین مؤلفه  $v$ ، متغیر دوگان یا ضربگر لاگرانژ مربوط به قید  $h_j(x) = 0$  می‌باشد.

<sup>۴۲</sup>Second-Order Sufficient Conditions

<sup>۴۳</sup>Hessian Matrix

<sup>۴۴</sup>Strict Local Minimum Point

<sup>۴۵</sup>Lagrangian Dual Problem

<sup>۴۶</sup>Dual Variable

<sup>۴۷</sup>Lagrangian Multiplier

مسئله دوگان لاگرانژ در شکل برداری به صورت زیر نوشته می‌شود:

$$\begin{aligned} \max \theta(u, v) \\ \text{s.t. } u \geq 0, \end{aligned} \quad (9.1)$$

که در آن

$$\theta(u, v) = \inf_x \{ f(x) + u^t g(x) + v^t h(x), x \in X \},$$

هم چنین  $h = (h_1, \dots, h_p) : \mathbb{R}^n \rightarrow \mathbb{R}^p$  و  $g = (g_1, \dots, g_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ،  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  هستند.

قضیه ۴.۲.۱ (دوگانی موضعی<sup>۴۸</sup> [۲۲]). فرض کنید مسئله

$$\begin{aligned} \min f(x) \\ \text{s.t. } h(x) = 0, \end{aligned} \quad (10.1)$$

$$x \in X \subset \mathbb{R}^n,$$

که دارای یک جواب موضعی در  $x^*$  و ضریب لاگرانژ  $v^*$  است. هم چنین فرض کنید که  $x^*$  یک نقطه منظم برای قیود است و هسین لاگرانژ  $\nabla^2 L(x^*, v^*)$  معین مثبت است. در این صورت مسئله دوگان

$$\max \theta(v) \quad (11.1)$$

که در آن

$$\theta(v) = \inf_x \{ f(x) + v^t h(x), x \in X \},$$

دارای یک جواب موضعی در  $v^*$  است و  $x^*$  جواب متناظر با  $v^*$  در تعریف  $\theta$  است.

به سادگی می‌توان نتایج فوق را به مسائلی که علاوه بر قیود تساوی قیود نامساوی هم دارند تعمیم داد

[۲۲].

## ۳.۱ سیستم‌های دینامیکی

تعریف ۱.۳.۱ ([۳]). سیستم دینامیکی زیر را در نظر بگیرید:

$$\dot{x} = f(x(t)), \quad x(t_0) = x_0 \in \mathbb{R}^n, \quad (12.1)$$

که در آن  $f$  یک تابع از  $\mathbb{R}^n$  به  $\mathbb{R}^n$  می‌باشد.  $x^*$  یک نقطه تعادل<sup>۴۹</sup> (۱۲.۱) نامیده می‌شود اگر  $f(x^*) = 0$ . اگر همسایگی  $\Omega^* \subseteq \mathbb{R}^n$  از  $x^*$  وجود داشته باشد که  $f(x^*) = 0$  و  $\forall x \in \Omega^* \setminus \{x^*\}$ ،  $f(x) \neq 0$ ، آن‌گاه  $x^*$  یک نقطه تعادل تنها<sup>۵۰</sup> نامیده می‌شود.

تعریف ۲.۳.۱ (پایداری به مفهوم لیاپانوف<sup>۵۱</sup> [۳]). فرض می‌کنیم که  $x(t)$  یک جواب (۱۲.۱) باشد،

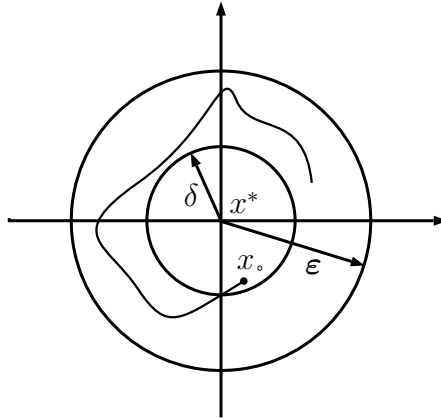
<sup>۴۸</sup>Local Duality

<sup>۴۹</sup>Equilibrium Point

<sup>۵۰</sup>Isolated Equilibrium Point

<sup>۵۱</sup>Stability in the Sense of Lyapunov

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  $\|x(t_0) - x^*\| < \delta$  آنگاه

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


شکل ۵.۱: پایداری لیاپانوف.

تعریف ۳.۳.۱ ([۳]). سیستم دینامیکی (۱۲.۱) همگرای سراسری<sup>۵۲</sup> به مجموعه جواب‌های (۱۲.۱) که با  $\Omega^*$  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 ([۳]). سیستم دینامیکی (۱۲.۱) در نقطه تعادل  $x^*$  که یکتا می‌باشد، پایدار مجانبی سراسری<sup>۵۳</sup> 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 ([۳]). مجموعه  $M \subseteq \mathbb{R}^n$  یک مجموعه تغییر ناپذیر<sup>۵۴</sup> نسبت به سیستم (۱۲.۱) 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.

تعریف ۶.۳.۱ ([۳]). 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<sup>۵۵</sup> نامیده می‌شود و به صورت زیر تعریف می‌شود:

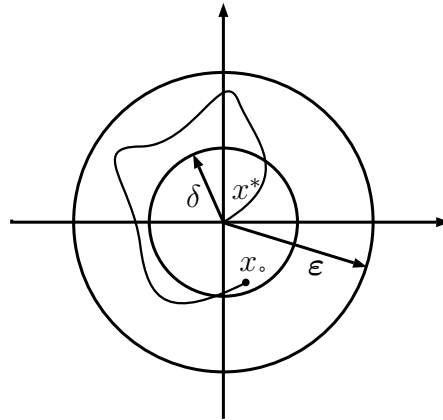
$$P_\Omega(x) = \arg \min_{v \in \Omega} \|x - v\|, \quad (۱۳.۱)$$

<sup>۵۲</sup>Globally Convergent

<sup>۵۳</sup>Globally Asymptotically Stable

<sup>۵۴</sup>Invariant Set

<sup>۵۵</sup>Projection Operator



شکل ۶.۱: پایداری مجانبی.

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 لیب شیتز صدق می‌کند (پیوسته لیب شیتز<sup>۵۶</sup> 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 لیب شیتز محلی<sup>۵۷</sup> روی  $E$  نامیده می‌شود اگر برای نقطه  $x_0 \in E$ ، یک  $\varepsilon$ -همسایگی مانند  $N_\varepsilon(x_0) \subset E$  و یک ثابت  $L$  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 ۸.۳.۱**. تابع  $F : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  را در نظر بگیرید. اگر  $F \in C^1(E)$ ، آنگاه  $F$  روی  $E$  لیب شیتز محلی است.

□

برهان. به [۴۷] مراجعه شود.

**تعریف ۹.۳.۱** ([۳]). نگاشت  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$

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

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

(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.

<sup>۵۶</sup>Lipschitz Continuous

<sup>۵۷</sup>Locally Lipschitz Continuous

تعریف ۱۰.۳.۱ ([۳]). اگر  $g : \Omega \subset \mathbb{R}^l \rightarrow \mathbb{R}$ ، آنگاه هر مجموعه غیر تهی به شکل زیر

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

یک مجموعه سطح<sup>۵۸</sup> از  $g$  نامیده می‌شود.

تعریف ۱۱.۳.۱ ([۳]). یک ماتریس  $n \times n$ ،  $M(x)$  که عناصر آن  $m_{ij}$ ،  $i = 1, \dots, n$  و  $j = 1, \dots, n$  هستند که روی مجموعه  $S \subset \mathbb{R}^n$  تعریف شده‌اند، روی  $S$  نیمه معین مثبت<sup>۵۹</sup> نامیده می‌شود اگر

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

اگر  $\gamma(x)$  کوچکترین مقدار ویژه ماتریس  $M(x)$  باشد سه حالت زیر را می‌توان نتیجه گرفت:

۱.  $M(x)$  روی  $S$  نیمه معین مثبت است اگر و فقط اگر به ازای هر  $x \in S$ ،  $\gamma(x) \geq 0$ .

۲.  $M(x)$  روی  $S$  معین مثبت<sup>۶۰</sup> است اگر و فقط اگر به ازای هر  $x \in S$ ،  $\gamma(x) > 0$ .

تعریف ۱۲.۳.۱ ([۳]). 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 ۱۳.۳.۱ ([۴۲]). فرض کنیم که  $g : \Omega \subset \mathbb{R}^l \rightarrow \mathbb{R}$  باشد که در آن  $\Omega$  بی‌کران است. آنگاه همه مجموعه‌های سطح  $g$  کراندار می‌باشند اگر و فقط اگر

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

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

قضیه ۱۴.۳.۱ (قضیه نقطه ثابت بروئر<sup>۶۱</sup> [۴۲]). تابع پیوسته  $f : \Omega \rightarrow \Omega$ ، که  $\Omega$  یک مجموعه فشرده و محدب می‌باشد حداقل یک نقطه ثابت دارد.

قضیه ۱۵.۳.۱ ([۴۲]). 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  $x \in K$  نیمه معین مثبت (معین مثبت) باشد.

قضیه ۱۶.۳.۱ ([۴۷]). در سیستم دینامیکی (۱۲.۱) 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 براین اگر  $f$  در  $x_0$  پیوسته لیپ شیتز محلی باشد آنگاه جواب یکتاست و اگر  $f$  در  $\mathbb{R}^n$  پیوسته لیپ شیتز باشد آنگاه  $\tau$  می‌تواند تا  $+\infty$  توسعه داده شود.

<sup>۵۸</sup>Level Set

<sup>۵۹</sup>Positive Semidefinite

<sup>۶۰</sup>Positive Definite

<sup>۶۱</sup>Brouwer's Fixed Point Theorem

تعریف ۱۷.۳.۱ ([۲۳]). 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، آنگاه یک جواب ماکسیمال نامیده می‌شود و بازه  $[t_0, \tau)$ ، بازه ماکسیمال وجود جواب نامیده می‌شود. بازه ماکسیمال وجود جواب مربوط به  $x_0$  اغلب به صورت  $[t_0, \tau(x_0))$  تعریف می‌شود.

قضیه ۱۸.۳.۱ ([۲۴]). فرض کنیم  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  یک تابع پیوسته باشد. اگر  $x(t) \in [t_0, \tau(x_0))$ ، یک جواب ماکسیمال باشد و  $\tau(x_0) < +\infty$  آنگاه

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

قضیه ۱۹.۳.۱ (قضیه پایداری لیاپانوف [۳۳]). فرض کنید که  $x = 0$  یک نقطه تعادل سیستم (۱۲.۱) باشد و  $V: \mathbb{R}^n \rightarrow \mathbb{R}$  یک تابع به طور پیوسته مشتق پذیر باشد به طوری که:

$$V(0) = 0. \quad ۱.$$

$$V(x) > 0, \quad x \in \mathbb{R}^n \setminus \{0\} \quad ۲.$$

$$\dot{V}(x) \leq 0, \quad x \in \mathbb{R}^n \setminus \{0\} \quad ۳.$$

آنگاه  $x = 0$  نقطه پایداری سیستم خواهد بود.

قضیه ۲۰.۳.۱ (قضیه پایداری مجانبی [۳۳]). تحت شرایط قضیه (۱۹.۳.۱) اگر  $V(\cdot)$  در شرایط زیر صدق کند:

$$V(0) = 0. \quad ۱.$$

$$V(x) > 0, \quad x \in \mathbb{R}^n \setminus \{0\} \quad ۲.$$

$$\dot{V}(x) < 0, \quad x \in \mathbb{R}^n \setminus \{0\} \quad ۳.$$

آنگاه  $x = 0$  پایدار مجانبی خواهد بود.

قضیه ۲۱.۳.۱ (قضیه پایداری مجانبی سراسری [۳۳]). تحت شرایط قضیه ۲۰.۳.۱ اگر  $V(\cdot)$  به طور شعاعی بی‌کران باشد آنگاه نقطه  $x = 0$  پایدار مجانبی سراسری است.

قضیه ۲۲.۳.۱ (اصل تغییرناپذیری لاسال<sup>۶۲</sup> [۳۳]). فرض کنیم  $V$  یک تابع به طور پیوسته مشتق پذیر باشد، همچنین فرض کنیم:

$$۱. \quad M \subset \mathbb{R}^n \quad \text{یک مجموعه فشرده و پایدار نسبت به جواب سیستم (۱۲.۱) باشد.}$$

$$۲. \quad \dot{V}(x) \leq 0, \quad x \in M \quad \text{به ازای}$$

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

<sup>۶۲</sup>LaSalle Principle of Invariance



۴.  $N$  بزرگترین مجموعه پایدار در  $E$  باشد.

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

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

## ۴.۱ نامساوی‌های وردشی

### مقدمه‌ای بر نامساوی وردشی

نظریه نامساوی وردشی<sup>۶۳</sup> اولین بار در سال ۱۹۶۶ به وسیله هارتمن<sup>۶۴</sup> و استمپکیا<sup>۶۵</sup> به عنوان ابزاری برای مطالعه معادلات دیفرانسیل جزئی با کاربردهایی در مکانیک معرفی شد [۲۸، ۳۱، ۳۲، ۳۹، ۵۶، ۵۹]. این نامساوی‌ها با بعد نامتناهی بودند. نظریه نامساوی وردشی با بعد متناهی در سال ۱۹۸۰ که به وسیله دافرmos<sup>۶۶</sup> معرفی شد، دافرmos متوجه شد که شرایط تعادل شبکه ترافیک که به وسیله اسمیت<sup>۶۷</sup> در سال ۱۹۷۹ بیان شد ساختار نامساوی وردشی دارد. نامساوی‌های وردشی در اقتصاد، علم مدیریت، تحقیق در عملیات و همچنین در مهندسی کاربردهای فراوانی دارد.

**تعریف ۱.۴.۱ ([۳]).** فرض کنید  $F : K \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  تابعی پیوسته روی مجموعه بسته و محدب  $K$  باشد، مسأله نامساوی وردشی که آن را با نماد  $VI(F, K)$  نشان می‌دهند یافتن یک بردار  $x^* \in K$  است به طوری که

$$F(x^*)^T(x - x^*) \geq 0, \quad \forall x \in K, \quad (14.1)$$

یا به طور معادل

$$\langle F(x^*), x - x^* \rangle \geq 0, \quad \forall x \in K, \quad (15.1)$$

که در آن  $\langle \cdot, \cdot \rangle$  ضرب داخلی در فضای اقلیدسی  $n$  بعدی است.

از دیدگاه هندسی، نامساوی وردشی (۱۵.۱) بیان می‌کند که  $F(x^*)^T$  بر ناحیه شدنی  $K$  در نقطه  $x^*$  عمود است (شکل ۷.۱).

**قضیه ۲.۴.۱.** مسأله بهینه‌سازی زیر را در نظر بگیرید:

$$\begin{aligned} \min & f(x) \\ \text{s.t.} & x \in K, \end{aligned}$$

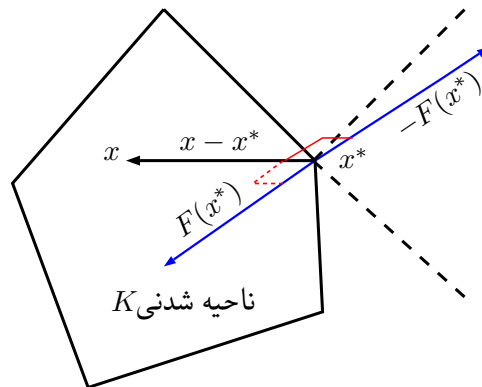
<sup>۶۳</sup>Variational Inequalities

<sup>۶۴</sup>Hartman

<sup>۶۵</sup>Stampacchia

<sup>۶۶</sup>Dafermos

<sup>۶۷</sup>Smith



شکل ۷.۱: تعبیر هندسی مسأله نامساوی وردشی.

که در آن  $f$  یک تابع به‌طور پیوسته مشتق‌پذیر و محدب است و  $K$  یک مجموعه محدب و غیرتهی در  $\mathbb{R}^n$  است آن‌گاه  $x^*$  جواب بهینه مسأله فوق است اگر و تنها اگر  $x^*$  جواب مسأله نامساوی وردشی

$$\nabla f(x^*)^T (x - x^*) \geq 0, \quad \forall x \in K$$

باشد.

□ برهان. با توجه به [۵۱] (نتیجه ۲، ص ۱۲۸) اثبات روشن است.

قضیه ۳.۴.۱ ([۲]). فرض کنید تابع  $F : K \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  روی  $K$  به‌طور پیوسته مشتق‌پذیر باشد و ماتریس ژاکوبین  $F$  متقارن و نیمه معین مثبت باشد، آن‌گاه تابع محدب حقیقی مقدار  $f : K \rightarrow \mathbb{R}$  وجود دارد که در رابطه

$$\nabla f(x) = F(x),$$

صدق می‌کند و  $x^*$  جواب  $VI(F, K)$ ، جواب مسأله بهینه‌سازی

$$\min f(x)$$

$$\text{s.t. } x \in K, \quad (۱۶.۱)$$

نیز است.

تعریف ۴.۴.۱ ([۲]). فرض می‌کنیم  $\mathbb{R}_+^n$  ناحیه نامنفی روی  $\mathbb{R}^n$  را نشان دهد و فرض کنیم  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  باشد. مسأله مکمل غیرخطی<sup>۶۸</sup> روی  $\mathbb{R}_+^n$  یک سیستم از معادلات و نامعادلات است که به صورت زیر است:

یافتن  $x^* \geq 0$  به‌طوری که

$$F(x^*) \geq 0, \quad F(x^*)^T x^* = 0. \quad (۱۷.۱)$$

هرگاه نگاشت  $F$  آفین باشد یعنی  $F(x) = Mx + b$  که در آن  $M$  یک ماتریس  $n \times n$  و  $b$  یک بردار  $n \times ۱$  می‌باشد مسأله (۱۷.۱) یک مسأله مکمل خطی<sup>۶۹</sup> نامیده می‌شود.

<sup>۶۸</sup>Nonlinear Complementarity Problem (NCP)

<sup>۶۹</sup>Linear Complementarity Problem (LCP)

مسئله نامساوی وردشی، مسئله مکمل را به عنوان یک حالت خاص در برمی‌گیرد. قضیه زیر رابطه بین مسئله مکمل و مسئله نامساوی وردشی را نشان می‌دهد.

قضیه ۵.۴.۱ ([۲]). مسئله نامساوی وردشی  $VI(F, \mathbb{R}_+^n)$  و مسئله مکمل (۱۷.۱) جواب‌های یکسان دارند.

قضیه ۶.۴.۱ ([۲]). فرض می‌کنیم که  $K$  یک مجموعه بسته محدب باشد آنگاه  $y = P_K(x)$  اگر و فقط اگر

$$y^T(z - y) \geq x^T(z - y), \quad \forall z \in K,$$

یا

$$(y - x)^T(z - y) \geq 0, \quad \forall z \in K.$$

برهان. توجه کنید که  $y = P_K(x)$  مینیمم کننده  $g(z) = \|x - z\|^2$  روی همه  $z \in K$  است. از آنجا که  $\nabla g(z) = 2(x - z)$  است، با تعریف مسئله زیر

$$\begin{aligned} \min g(z) &= \|x - z\|^2 \\ \text{s.t. } z &\in K, \end{aligned}$$

با توجه به قضیه (۲.۴.۱)،  $y = P_K(x)$  جواب بهینه مسئله فوق است اگر و فقط اگر

$$\nabla g(z)^T(z - y) \geq 0, \quad \forall z \in K,$$

بنابراین

$$(y - x)^T(z - y) \geq 0, \quad \forall z \in K.$$

□

نتیجه ۷.۴.۱ ([۲]). فرض می‌کنیم که  $K$  یک مجموعه بسته و محدب باشد، آنگاه عملگر تصویر  $P_K(\cdot)$  در شرط لیپ شیتز صدق می‌کند یعنی:

$$\|P_K(u) - P_K(v)\| \leq \|u - v\|, \quad \forall u, v \in \mathbb{R}^n.$$

قضیه ۸.۴.۱ ([۲]). فرض کنیم  $K$  مجموعه‌ای بسته و محدب باشد، آنگاه  $x^* \in K$  یک جواب مسئله نامساوی وردشی  $VI(F, K)$  است اگر و فقط اگر برای هر  $\gamma > 0$ ، یک نقطه ثابت نگاشت

$$P_K(I - \gamma F) : K \rightarrow K,$$

باشد یعنی

$$x^* = P_K(x^* - \gamma F(x^*)).$$

# فصل ۲

## کاربرد شبکه‌های عصبی در حل مسائل نامساوی‌های وردشی و مکمل

### ۱.۲ مقدمه

در این فصل از پایان‌نامه ابتدا مسائل نامساوی‌های وردشی و مکمل<sup>۱</sup> را معرفی کرده و مروری بر کارهای انجام شده در این زمینه داریم. در ادامه یک مدل شبکه عصبی را معرفی می‌کنیم، و به اثبات پایداری و همگرایی آن می‌پردازیم. در پایان کارایی این مدل شبکه عصبی را با ارائه چندین مثال عددی از مسائل نامساوی وردشی و مکمل نشان می‌دهیم.

### بیان مسأله

مسأله پیدا کردن نقطه  $x^* \in S \subseteq \mathbb{R}^n$  را در نظر می‌گیریم به طوری که:

$$VI(F, S) \quad \langle F(x^*), x - x^* \rangle \geq 0, \quad \forall x \in S, \quad (1.2)$$

که در آن  $F$  نگاشت پیوسته‌ای از  $\mathbb{R}^n$  به  $\mathbb{R}^n$  است و  $S = \{x \in \mathbb{R}^n \mid Ax - b \geq 0, Bx = c, x \geq 0\}$  که  $c \in \mathbb{R}^r$ ,  $b \in \mathbb{R}^m$ ,  $r \leq n$ ,  $0 \leq m$ ,  $\text{rank}(B) = r$ ,  $B \in \mathbb{R}^{r \times n}$ ,  $\text{rank}(A) = m$ ,  $A \in \mathbb{R}^{m \times n}$  و  $S$  زیرمجموعه محدب بسته غیرتهی در  $\mathbb{R}^n$  است، که  $\langle \cdot, \cdot \rangle$  حاصل ضرب داخلی در  $\mathbb{R}^n$  را نشان می‌دهد. این مسأله را نامساوی وردشی می‌نامیم.

در حالت خاص برای  $S = \mathbb{R}_+^n$  مسأله (۱.۲) می‌تواند به صورت مسأله مکمل غیرخطی<sup>۲</sup> زیر نوشته شود:

$$x^* \geq 0, \quad F(x^*) \geq 0, \quad \langle x^*, F(x^*) \rangle = 0. \quad (2.2)$$

<sup>۱</sup>Complementarity and Variational Inequalities

<sup>۲</sup>Nonlinear Complementarity Problem (NCP)

برای  $S = \mathbb{R}^n$  به راحتی مسأله (۱.۲) به حل سیستم معادلات غیرخطی  $F(x) = 0$  ساده می‌شود [۷۱]. مخصوصاً برای تابع به‌طور پیوسته مشتق‌پذیر  $f$  اگر  $x^*$  جواب مسأله

$$\min \{f(x) | x \in S\} \quad (۳.۲)$$

باشد، سپس  $x^*$  جواب مسأله (۱.۲) با  $F(x) = \nabla f(x)$  است، که  $S$  همان تابع تعریف شده در (۱.۲) است و بردار گرادین  $f(x)$  در نقطه  $x$  به صورت زیر تعریف می‌شود:

$$\nabla f(x) = \left( \frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T \in \mathbb{R}^n.$$

بنابراین تعدادی از مسائل بهینه‌سازی مانند برنامه ریزی‌های خطی و درجه دو [۱۲، ۱۴، ۶۲]، مین ماکس و برنامه‌ریزی غیرخطی ممکن است با استفاده از (۳.۲) حل شود.

در سال‌های اخیر، چون مسائل نامساوی‌های وردشی کاربرد مهمی در طیف گسترده‌ای از رشته‌های علمی و مهندسی از جمله اقتصاد شبکه، علوم حمل و نقل، تئوری بازی‌ها، برنامه‌ریزی نظامی، کنترل اتوماتیک، پردازش سیگنال، تجزیه و تحلیل رگرسیون، طراحی ساختار، طراحی مکانیکی، شبکه‌های برنامه‌ریزی و غیره دارند، تعدادی الگوریتم بهینه‌سازی و شبکه‌های عصبی مصنوعی<sup>۳</sup> برای حل این مسائل کشف شدند [۱۳، ۴۱، ۵۰، ۵۱، ۵۲، ۵۸].

با این حال، الگوریتم‌های بهینه‌سازی سنتی [۱۲، ۱۳، ۱۴، ۵۰، ۵۱، ۵۲]، برای مسائل زمان واقعی<sup>۴</sup> برای پیاده‌سازی روی کامپیوتر مناسب نیستند. یکی از روش‌های امیدوارکننده برای برطرف کردن این مشکلات به کار بردن شبکه عصبی مصنوعی مبتنی بر پیاده‌سازی مداری است. مهم‌ترین مزیت شبکه‌های عصبی، اجرای موازی و همگرایی سریع آن است. بسیاری از شبکه‌های عصبی زمان پیوسته با استفاده از پارامترهای جریمه، برای حل مسائل بهینه‌سازی مقید، توسعه داده شده است [۷، ۴۶، ۶۰، ۷۶، ۸۱]. در سال‌های اخیر، برای جلوگیری از به کار بردن پارامترهای جریمه کارهای قابل توجهی انجام شده است [۶]. تعدادی از شبکه‌های عصبی اولیه و دوگان با ساختار دو لایه و یک لایه کشف شده‌اند [۴۶، ۶۰، ۷۶]. این شبکه‌های عصبی همگرایی سراسری به یک جواب دقیق هستند [۸۰، ۸۲].

در این قسمت به منظور حل (۱.۲) و (۲.۲) و مسائل بهینه‌سازی متناظر، رده‌ای از شبکه‌های عصبی با ساختار لایه پنهان را پیشنهاد می‌دهیم. این مدل شبکه‌های عصبی پارامترهای قابل تنظیم ندارد بنابراین پیچیدگی زمانی آن کمتر است.

**تعریف ۱.۱.۲ ([۲۰]).** نگاشت  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  روی  $S$  یکنوا است، اگر برای هر  $x, x' \in S$  داشته باشیم:

$$\langle F(x) - F(x'), x - x' \rangle \geq 0 \quad (۴.۲)$$

$F$  روی  $S$  اکیداً یکنوا است، اگر برای هر  $x \neq x'$  نامساوی (۴.۲) اکید برقرار باشد.

لم ۲.۱.۲ ([۱۳، ۴۱، ۵۲، ۵۸]). اگر  $F$  به‌طور پیوسته مشتق‌پذیر و ماتریس ژاکوبین  $\nabla F$  برای هر  $x \in S$  معین مثبت باشد، در نتیجه  $F$  روی  $S$  اکیداً یکنوا است.

<sup>۳</sup>Artificial Neural Networks

<sup>۴</sup>Real-Time

لم ۳.۱.۲. بردار  $x^*$  جواب مسأله (۱.۲) است، اگر و تنها اگر  $y^* \in \mathbb{R}^m$  و  $z^* \in \mathbb{R}^r$  که  $y^* \geq 0$  وجود داشته باشد به طوری که  $u^* = (x^*, y^*, z^*)$  جواب مسأله زیر باشد:

$$\langle G(u^*), u - u^* \rangle \geq 0 \quad \forall u \in S_0.$$

که مجموعه  $S_0 = \{u = (x, y, z) | x \geq 0, y \geq 0, z\}$  است و

$$G(u) = \begin{pmatrix} F(x) - A^T y - B^T z \\ Ax - b \\ Bx - c \end{pmatrix}.$$

برهان. لم با استفاده از شرایط کاروش-کان تاکر<sup>۵</sup>، اثبات می‌شود [۲۰]. □

لم ۴.۱.۲ ([۲۳، ۵۲]). تصویر نقطه  $u$  روی مجموعه محدب بسته  $S_0$  با  $P_{S_0}(u)$  نشان داده می‌شود و به عنوان جواب یکتای مسأله زیر تعریف می‌شود:

$$\begin{aligned} \min \quad & \| \eta - u \| \\ \text{s.t.} \quad & \eta \in S_0. \end{aligned}$$

عملگر تصویر به صورت  $P_{S_0}(\cdot) : \mathbb{R}^l \rightarrow S_0$  در نظر گرفته می‌شود که  $(l = n + m + r)$  است. همچنین برای هر  $u, u' \in \mathbb{R}^l$  داریم:

$$\| P_{S_0}(u) - P_{S_0}(u') \| \leq \| u - u' \|.$$

علاوه بر این، برای هر  $u \in \mathbb{R}^l$ ،  $P_{S_0}(u)$  در رابطه زیر صدق می‌کند:

$$\langle u - P_{S_0}(u), u' - P_{S_0}(u') \rangle \leq 0, \quad \forall u' \in S_0, u \in \mathbb{R}^l,$$

که در آن  $\| \cdot \|$  نرم اقلیدسی را نشان می‌دهد.

لم ۵.۱.۲ ([۲۵]). فرض کنید  $G(u) : \mathbb{R}^l \rightarrow \mathbb{R}^l$  تابع پیوسته‌ای باشد، داریم:

$$\langle G(u^*), u - u^* \rangle \geq 0 \quad \forall u \in \mathbb{R}^l \iff u^* = P_{S_0}(u^* - G(u^*)).$$

تعریف ۶.۱.۲ ([۶۱]). بردار  $u^* = u(t^*)$  یک نقطه تعادل یا حالت پایدار سیستم دینامیکی

$\frac{du(t)}{dt} = H(u(t))$  در زمان  $t^* \in \mathbb{R}^+$  است اگر برای هر  $t \geq t^*$  داشته باشیم:

$$H(u^*) = 0.$$

تعریف ۷.۱.۲ ([۷۴]). فرض کنید  $S_1 \subset \mathbb{R}^l$  یک همسایگی باز از  $u^*$  باشد، تابع لیاپانوف<sup>۶</sup>

$E : \mathbb{R}^l \rightarrow \mathbb{R}$  یک تابع به طور پیوسته مشتق پذیر است که در شرایط زیر صدق می‌کند:

(۱)  $E(u(t))$  نامنفی است، یعنی  $E(u(t)) \geq 0$ . مخصوصاً  $E(u^*) = 0$  و برای هر  $u \in S_1$  داریم

<sup>۵</sup>Karush-Kuhn-Tucker conditions

<sup>۶</sup>Lyapunov Function

$$.E(u(t)) > 0$$

(۲) مشتق  $E(u(t))$  نسبت به زمان نامثبت است، یعنی

$$\frac{dE(u(t))}{dt} = [\nabla_{u(t)} E(u(t))]^T \frac{du}{dt} = [\nabla_{u(t)} E(u(t))]^T H(u(t)) \leq 0, \quad \forall u \in S_1.$$

$$. \frac{dE(u(t))}{dt} = 0 \text{ برای } u(t) = u^*$$

قضیه ۸.۱.۲ ([۷۴]). نقطه تعادل  $u^*$  از سیستم دینامیکی  $\frac{du(t)}{dt} = H(u(t))$  یک نقطه ثابت لیاپانوف است اگر تابع انرژی متناظری وجود داشته باشد که در تعریف (۷.۱.۲) صدق کند.

تعریف ۹.۱.۲ ([۶۱]). یک سیستم دینامیکی زمان پیوسته همگرای سراسری نامیده می‌شود، اگر برای هر نقطه اولیه داده شده، مسیر سیستم دینامیکی، همگرا به نقطه تعادل مربوطه باشد.

تعریف ۱۰.۱.۲ ([۷۰]). فرض کنید  $u(t)$  جواب سیستم  $\frac{du(t)}{dt} = H(u(t))$  باشد، در نتیجه  $u(t)$  همگرا به مجموعه  $\tilde{S}$  است اگر

$$\lim_{t \rightarrow +\infty} \text{dist}(u(t), \tilde{S}) = 0,$$

$$\text{که در آن } \text{dist}(u(t), \tilde{S}) = \inf_{\eta \in \tilde{S}} \|u(t) - \eta\|$$

قبل از بیان مدل شبکه عصبی، فرض‌های زیر را در نظر می‌گیریم:

فرض اول: نگاشت  $F$  روی مجموعه محدب باز شامل  $S$  یک بار مشتق‌پذیر است.

فرض دوم: نگاشت  $F$  روی  $S$  یکنوا است.

فرض سوم:  $\nabla F$  روی  $\mathbb{R}_+^n = \{x \in \mathbb{R}^n | x \geq 0\}$  معین مثبت است.

## ۲.۲ مدل شبکه عصبی

از لم‌های (۳.۱.۲) و (۵.۱.۲) می‌توانیم قضیه زیر را به‌دست آوریم.

قضیه ۱.۲.۲. فرض کنید  $P_S(u) = [(x_1)^+, \dots, (x_n)^+, (y_1)^+, \dots, (y_m)^+, z_1, \dots, z_r]$

$$. (x_i)^+ = \max\{0, x_i\} \text{ و } (y_j)^+ = \max\{0, y_j\} \text{ برای } i = 1, \dots, n \text{ و } j = 1, \dots, m$$

بردار  $x^*$  جواب مسأله نامساوی وردشی (۱.۲) است، اگر و تنها اگر

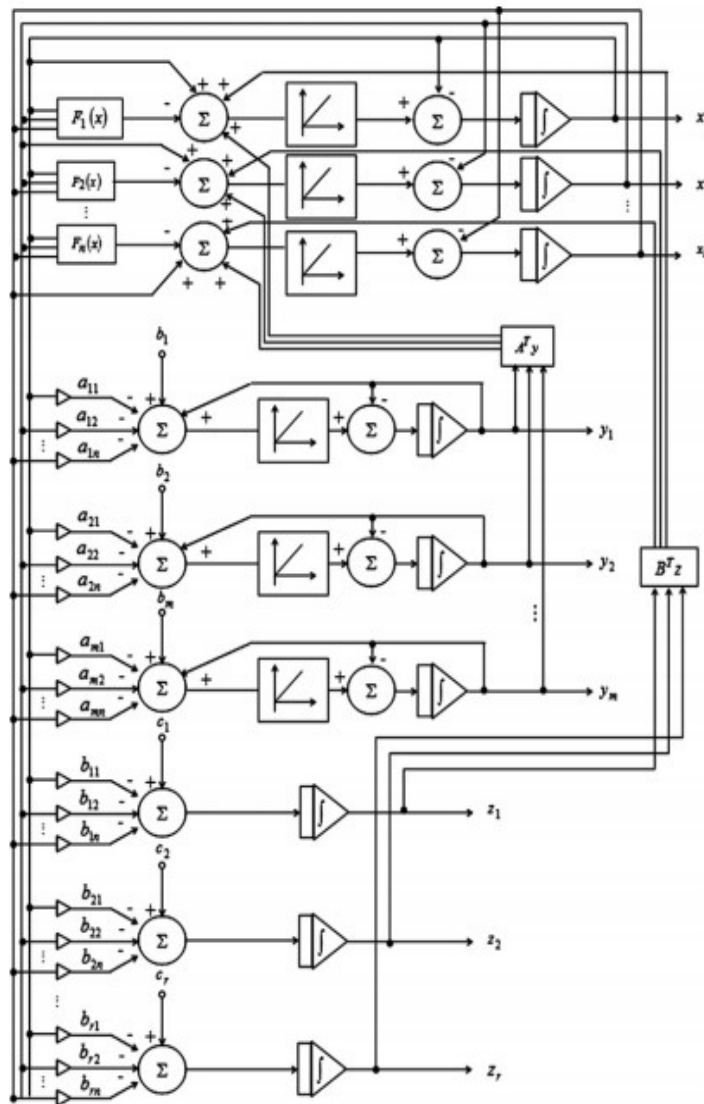
$$\begin{cases} x^* = (x^* - F(x^*) + A^T y^* + B^T z^*)^+, \\ y^* = (y^* - Ax^* + b)^+, \\ Bx^* = c, \end{cases} \quad (5.2)$$

که در آن  $x(\cdot)$  و  $y(\cdot)$  و  $z(\cdot)$  وابسته به زمان‌اند.

مدل شبکه عصبی مربوط به قضیه (۱.۲.۲) به شکل سیستم دینامیکی غیرخطی زیر تعریف می‌شود:

$$\frac{du}{dt} = \frac{d}{dt} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = H(u) = \begin{pmatrix} (x - F(x) + A^T y + B^T z)^+ - x \\ (y - Ax + b)^+ - y \\ -Bx + c \end{pmatrix}. \quad (6.2)$$

این مدل شبکه عصبی چون پارامترهای قابل تنظیم ندارد، به محاسبات کمتری نیاز دارد. شکل (۱.۲)، نمودار بلوکی از این شبکه عصبی را نشان می‌دهد که  $x$  و  $y$  و  $z$  خروجی‌های شبکه‌اند. بردارهای  $b = (b_1, b_2, \dots, b_m)^T$  و  $c = (c_1, c_2, \dots, c_r)^T$  داده‌های خروجی‌اند. پارامترهای دیگر  $F(x) = (F_1(x), \dots, F_n(x))^T$  و  $A = (a_{ij})_{m \times n}$  و  $B = (b_{pj})_{r \times n}$  هستند.



شکل ۱.۲: نمودار بلوکی شبکه عصبی (۶.۲)

۲.۲.۲. فرض کنید  $u^*$  جواب (۵.۲) است  $S^* = \{u^* \in \mathbb{R}^l \mid u^* \in S^* \text{ است، اگر و تنها اگر } u^* \text{ نقطه تعادل شبکه (۶.۲) باشد.}$

□

برهان. نتیجه مستقیم قضیه (۱.۲.۲) است.



## ۳.۲ تحلیل پایداری سیستم دینامیکی

در این بخش نشان می‌دهیم که شبکه عصبی (۶.۲) پایدار لیاپانوف است و همگرایی سراسری به مجموعه جواب مسأله (۱.۲) است.

قضیه ۱.۳.۲. اگر فرض‌های اول تا سوم برقرار باشد، سپس برای هر نقطه اولیه  $u(t_0) = u_0 \in S$  داریم:

(الف) جواب منحصر به فرد پیوسته  $u(t) = (x(t), y(t), z(t)) \in S$  برای (۶.۲) در بازه  $[t_0, T(u_0)]$  وجود دارد.

(ب) اگر  $x(t_0) \geq 0$  و  $y(t_0) \geq 0$  باشد داریم  $x(t) \geq 0$  و  $y(t) \geq 0$ .

برهان. (الف) چون  $(x - F(x) + A^T y + B^T z)^+ - x$ ،  $(y - Ax + b)^+ - y$  و  $-Bx + c$  پیوسته لیپ شیتس محلی هستند، با استفاده از قضیه (۱۶.۳.۱)، جواب یکتای  $u(t)$  برای سیستم دینامیکی (۶.۲) در بازه  $[t_0, T(u_0)]$  وجود دارد به طوری که  $T(u_0)$  به  $u_0$  وابسته است.

(ب) برای اثبات (ب) ابتدا نشان می‌دهیم که  $x(t) \geq 0$  داریم:

$$\frac{dx}{dt} = (x - F(x) + A^T y + B^T z)^+ - x.$$

همچنین

$$\frac{dx}{dt} + x(t) = (x - F(x) + A^T y + B^T z)^+.$$

با ضرب هر دو طرف معادلات بالا در  $e^t$  و انتگرال گیری بر حسب  $t$  در بازه  $[t_0, T(u_0)]$  داریم:

$$\frac{dx}{dt} e^t + x(t) e^t = e^t (x - F(x) + A^T y + B^T z)^+,$$

$$\frac{d}{dt} [e^t x(t)] = e^t (x - F(x) + A^T y + B^T z)^+,$$

$$\int_{t_0}^t \frac{d}{ds} [e^s x(s)] ds = \int_{t_0}^t e^s (x - F(x) + A^T y + B^T z)^+ ds,$$

$$e^s x(s) \Big|_{t_0}^t = \int_{t_0}^t e^s (x - F(x) + A^T y + B^T z)^+ ds,$$

$$e^t x(t) - e^{t_0} x(t_0) = \int_{t_0}^t e^s (x - F(x) + A^T y + B^T z)^+ ds,$$

$$x(t) = e^{-(t-t_0)} x(t_0) + e^{-t} \int_{t_0}^t e^s (x - F(x) + A^T y + B^T z)^+ ds \geq 0.$$

در نتیجه چون دو جمله سمت راست مقادیر نامنفی‌اند،  $x(t) \geq 0$ . همین نتیجه برای  $y(t) \geq 0$  صادق است. □

لم ۲.۳.۲. اگر فرض دوم برقرار باشد سپس  $G$  یکنواخت (یکنوا) است.

برهان. نشان می‌دهیم که برای هر  $\hat{u}, \tilde{u} \in S$  داریم  $\langle (\hat{u} - \tilde{u}), G(\hat{u}) - G(\tilde{u}) \rangle \geq 0$ . داریم:

$$\begin{aligned} \langle (\hat{u} - \tilde{u}), G(\hat{u}) - G(\tilde{u}) \rangle &= \langle \hat{x} - \tilde{x}, (F(\hat{x}) - A^T \hat{y} - B^T \hat{z}) - (F(\tilde{x}) - A^T \tilde{y} - B^T \tilde{z}) \rangle \\ &+ \langle \hat{y} - \tilde{y}, (A\hat{x} - b) - (A\tilde{x} - b) \rangle + \langle \hat{z} - \tilde{z}, (B\hat{x} - c) - (B\tilde{x} - c) \rangle \\ &= \langle \hat{x} - \tilde{x}, F(\hat{x}) - F(\tilde{x}) \rangle - \langle (\hat{x} - \tilde{x}), A^T(\hat{y} - \tilde{y}) \rangle - \langle \hat{x} - \tilde{x}, B^T(\hat{z} - \tilde{z}) \rangle \\ &+ \langle \hat{y} - \tilde{y}, A(\hat{x} - \tilde{x}) \rangle + \langle \hat{z} - \tilde{z}, B(\hat{x} - \tilde{x}) \rangle \\ &= \langle \hat{x} - \tilde{x}, F(\hat{x}) - F(\tilde{x}) \rangle \geq 0. \end{aligned}$$

طبق فرض، چون  $F(x)$  روی  $S$  یکنواخت است، داریم:

$$\langle (\hat{u} - \tilde{u}), G(\hat{u}) - G(\tilde{u}) \rangle \geq 0.$$

□

قضیه ۳.۳.۲. فرض کنید  $u^* = (x^{*T}, y^{*T}, z^{*T})^T$  نقطه تعادل (۶.۲) باشد، که  $x^*$  جواب بهینه (۱.۲) است. در نتیجه شبکه عصبی پیشنهادی (۶.۲) با نقطه اولیه  $u_0 = (x_0, y_0, z_0) \in S$  پایدار به مفهوم لیاپانوف است.

برهان. تابع  $E: \mathbb{R}^l \rightarrow \mathbb{R}$  را به صورت زیر تعریف می‌کنیم:

$$E(u) = -\langle G(u), H(u) \rangle - \frac{1}{\rho} \|H(u)\|^2 + \frac{1}{\rho} \|u - u^*\|^2,$$

که در آن  $G(u)$  و  $H(u)$  به ترتیب در لم (۳.۱.۲) و (۶.۲) داده شده‌اند. فرض کنید که  $S_1 \subseteq \mathbb{R}^l$  همسایگی برای  $u^*$  باشد. نشان می‌دهیم که  $E(u)$  تابع لیاپانوف مناسبی برای سیستم دینامیکی (۶.۲) است. با استفاده از نتیجه داده شده در [۴۳] واضح است که:

$$-\langle G(u), H(u) \rangle \geq \|H(u)\|^2, \quad (۷.۲)$$

$$\langle H(u) + u - u^*, -H(u) - G(u) \rangle \geq 0. \quad (۸.۲)$$

از (۷.۲) واضح است که  $E(u)$  روی  $S_1$  نامنفی است و  $E(u^*) = 0$ . می‌بینیم که برای همه

$$u \in S_1 \setminus \{u^*\}, E(u) > 0 \text{ داریم}$$

حال نشان می‌دهیم که  $\frac{dE(u)}{dt} \leq 0$ . چون

$$\frac{dE}{dt} = \left\langle \frac{\partial E}{\partial u}, \frac{du}{dt} \right\rangle = \langle \nabla E(u), \frac{du}{dt} \rangle,$$

در نتیجه از قضیه ۳.۲ از [۷۰] داریم:

$$\nabla E(u) = G(u) - (\nabla G(u) - I)H(u) + u - u^*,$$

که در آن  $\nabla G(u)$  ماتریس ژاکوبین  $G$  را نشان می‌دهد و

$$\nabla G(u) = \begin{pmatrix} \nabla F(x) & -A^T & -B^T \\ A & 0 & 0 \\ B & 0 & 0 \end{pmatrix}.$$

داریم:

$$\begin{aligned} \frac{dE(u)}{dt} &= \langle G(u) - (\nabla G(u) - I)H(u) + u - u^*, H(u) \rangle \\ &= \langle G(u) + u - u^*, H(u) \rangle + \|H(u)\|^2 - \langle H(u), \nabla G(u)H(u) \rangle. \end{aligned}$$

با استفاده از رابطه (۸.۲) می‌توان نوشت:

$$\begin{aligned} \langle H(u) + u - u^*, -H(u) - G(u) \rangle &\geq 0 \\ \Rightarrow -\langle H(u), H(u) \rangle - \langle H(u), u - u^* \rangle - \langle G(u), H(u) \rangle - \langle G(u), u - u^* \rangle &\geq 0 \\ \Rightarrow -\|H(u)\|^2 - \langle G(u) + u - u^*, H(u) \rangle - \langle G(u), u - u^* \rangle &\geq 0 \\ \Rightarrow -\langle G(u) + u - u^*, H(u) \rangle \geq \langle G(u), u - u^* \rangle + \|H(u)\|^2 \\ \Rightarrow \langle G(u) + u - u^*, H(u) \rangle \leq -\langle G(u), u - u^* \rangle - \|H(u)\|^2 \end{aligned}$$

بنابراین

$$\begin{aligned} \frac{dE(u)}{dt} &\leq -\langle u - u^*, G(u) \rangle - \|H(u)\|^2 + \|H(u)\|^2 - \langle H(u), \nabla G(u)H(u) \rangle \\ &= -\langle u - u^*, G(u) \rangle - \langle H(u), \nabla G(u)H(u) \rangle. \end{aligned} \quad (9.2)$$

حال نشان می‌دهیم که  $\langle u - u^*, G(u) \rangle$  و  $\langle H(u), \nabla G(u)H(u) \rangle$  مثبت‌اند. عبارت  $\langle H(u), \nabla G(u)H(u) \rangle$  مثبت است چون  $\nabla F$  معین مثبت است. در نتیجه  $\nabla G(u)$  نیمه‌معین مثبت می‌شود. از لم (۲.۳.۲) داریم:

$$\langle u - u^*, G(u) - G(u^*) \rangle \geq 0.$$

همچنین

$$\langle u - u^*, G(u) - G(u^*) \rangle = \langle u - u^*, G(u) \rangle - \langle u - u^*, G(u^*) \rangle \geq 0.$$

بنابراین

$$\langle u - u^*, G(u) \rangle \geq \langle u - u^*, G(u^*) \rangle$$

با استفاده از رابطه

$$\langle G(u^*), u - u^* \rangle \geq 0 \quad \forall u \in S..$$

که در لم (۳.۱.۲) بیان شد داریم:

$$\langle u - u^*, G(u^*) \rangle \geq 0 \quad \forall u \in S..$$

بنابراین

$$\langle u - u^*, G(u) \rangle \geq 0 \quad \forall u \in S.. \quad (10.2)$$

□

اثبات کامل شد.

قضیه ۴.۳.۲. مسیر جواب شبکه عصبی (۶.۲) همگرای سراسری به  $S^*$  است.

برهان. قضیه (۳.۳.۲) نشان می‌دهد که تابع  $E(u(t))$  برای همه  $t \geq t_0$  یکنواخت کاهشی<sup>۷</sup> است. مجموعه  $\Phi = \{u \in S_0 | E(u) \leq E(u_0)\}$  کراندار است [۴۱]. علاوه بر این تابع  $E(u)$  به طور پیوسته مشتق پذیر است. بنابراین از اصل تغییرناپذیری لسال (۲۲.۳.۱) مسیر  $u(t)$  سیستم دینامیکی (۶.۲) زمانی که  $t \rightarrow +\infty$  همگرا به  $\nu$  است، که  $\nu$  بزرگترین مجموعه تغییرناپذیر در  $\Pi = \{u \in \Phi | \frac{dE(u)}{dt} = 0\}$  است.

حال نشان می‌دهیم که  $\frac{du}{dt} = 0 \iff \frac{dE}{dt} = 0$ . واضح است که اگر  $\frac{du}{dt} = 0$  در نتیجه  $\frac{dE}{dt} = \langle \frac{\partial E}{\partial u}, \frac{du}{dt} \rangle = 0$ . برای اثبات برعکس، فرض کنید  $\hat{u} = (\hat{x}^T, \hat{y}^T, \hat{z}^T)^T \in \Pi$ . کفایت نشان دهیم که  $\frac{d\hat{x}}{dt} = 0$ ،  $\frac{d\hat{y}}{dt} = 0$  و  $\frac{d\hat{z}}{dt} = 0$ . می‌توان دید که  $\frac{dE(\hat{u})}{dt} = 0$  از (۹.۲) داریم:

$$\langle \hat{u} - u^*, G(\hat{u}) \rangle + \langle H(\hat{u}), \nabla G(\hat{u}) H(\hat{u}) \rangle = 0. \quad (۱۱.۲)$$

چون  $\nabla G(\hat{u})$  نیمه معین مثبت است،  $\langle G(\hat{u}), \hat{u} - u^* \rangle \geq 0$ ، که در آن

$$\nabla G(\hat{u}) = \begin{pmatrix} \nabla F(\hat{x}) & -A^T & -B^T \\ A & 0 & 0 \\ B & 0 & 0 \end{pmatrix}.$$

در نتیجه با توجه به (۱۱.۲) داریم:

$$\langle H(\hat{u}), \nabla G(\hat{u}) H(\hat{u}) \rangle = 0, \quad \langle G(\hat{u}), \hat{u} - u^* \rangle = 0, \quad \langle G(\hat{u}) - G(u^*), \hat{u} - u^* \rangle = 0,$$

داریم:

$$\langle H(\hat{u}), \nabla G(\hat{u}) H(\hat{u}) \rangle = \langle ((\hat{x} - F(\hat{x}) + A^T \hat{y} + B^T \hat{z})^+ - \hat{x}), \nabla F(\hat{x})(\hat{x} - F(\hat{x}) + A^T \hat{y} + B^T \hat{z}) - \hat{x} \rangle = 0.$$

معین مثبت بودن  $\nabla F(\hat{x})$  نشان می‌دهد که

$$(\hat{x} - F(\hat{x}) + A^T \hat{y} + B^T \hat{z})^+ - \hat{x} = 0, \quad \frac{d\hat{x}}{dt} = 0.$$

از طرف دیگر،  $\langle G(\hat{u}) - G(u^*), \hat{u} - u^* \rangle = 0$ ، لذا

$$\langle F(\hat{x}) - F(x^*), \hat{x} - x^* \rangle = \langle \hat{x} - x^*, \nabla F(x_\mu)(\hat{x} - x^*) \rangle = 0,$$

که در آن برای همه  $0 \leq \mu \leq 1$  داریم  $x_\mu = (1 - \mu)\hat{x} + \mu x^*$

چون  $\hat{x} = x^*$  در نتیجه

$$B\hat{x} - c = 0, \quad \frac{d\hat{z}}{dt} = 0.$$

حال  $\langle G(\hat{u}), \hat{u} - u^* \rangle = 0$  را در نظر بگیرید، که می‌تواند به صورت زیر نوشته شود:

$$\langle F(\hat{x}) - A^T \hat{y} - B^T \hat{z}, \hat{x} - x^* \rangle + \langle A\hat{x} - b, \hat{y} - y^* \rangle + \langle B\hat{x} - c, \hat{z} - z^* \rangle = 0.$$

چون  $\hat{x} = x^*$  می‌توان نوشت:

$$\langle A\hat{x} - b, \hat{y} - y^* \rangle = 0 \quad (۱۲.۲)$$

<sup>۷</sup>Monotonically Nonincreasing

با توجه به شرایط کاروش-کان-تاگر [۲۳]،  $x^*$  جواب (۱.۲) است اگر و تنها اگر  
وجود داشته باشد به طوری که  $(y^{*T}, z^{*T})^T \in \mathbb{R}^{m+r}$  در شرایط زیر صدق کند:

$$Ax^* - b \geq 0, \quad Bx^* - c = 0, \quad \langle Ax^* - b, y^* \rangle = 0,$$

$$F(x^*) - A^T y^* - B^T z^* \geq 0, \quad \langle F(x^*) - A^T y^* - B^T z^*, x^* \rangle = 0, \quad y^* \geq 0.$$

رابطه (۱۲.۲) معادل با رابطه زیر است:

$$\langle A\hat{x} - b, \hat{y} \rangle = \langle A\hat{x} - b, y^* \rangle = \langle Ax^* - b, y^* \rangle = 0.$$

علاوه بر این،  $\hat{y} \geq 0$  و  $A\hat{x} - b \geq 0$  و  $\langle A\hat{x} - b, \hat{y} \rangle = 0$  اگر و تنها اگر

$$(\hat{y} - A\hat{x} + b)^+ - \hat{y} = 0, \quad \frac{d\hat{y}}{dt} = 0.$$

در نتیجه،  $\frac{du}{dt} = 0$  اگر و تنها اگر  $\frac{dE}{dt} = 0$ . با توجه به لم (۲.۲.۲) می‌دانیم  $\Pi \subseteq S^*$ . در نتیجه فاصله  $(u(t), S^*)$  زمانی که  $t \rightarrow +\infty$  به صفر میل می‌کند.  $\square$

## ۴.۲ تعمیم

مسئله نامساوی وردشی زیر را در نظر بگیرید:

$$VI(F, \tilde{S}) \quad x^* \in \tilde{S}, \quad \langle F(x^*), x - x^* \rangle \geq 0, \quad \forall x \in \tilde{S}, \quad (۱۳.۲)$$

که در آن  $x, A, B, c, d$  و  $F$  در بخش ۱ همین فصل تعریف شده‌اند و برای  $\Omega = \{l \leq x \leq h\}$  داریم:

$$\tilde{S} = \{x \in \mathbb{R}^n \mid Ax \geq b, Bx = c, x \in \Omega\}.$$

نتیجه ۱.۴.۲. سیستم (۵.۲) می‌تواند به صورت زیر تغییر کند:

$$\begin{cases} x^* = P_{\Omega}(x^* - F(x^*) + A^T y^* + B^T z^*), \\ y^* = (y^* - Ax^* + b)^+, \\ Bx^* = c, \end{cases}$$

که در آن  $P_{\Omega}: \mathbb{R}^n \rightarrow \Omega$  عملگر تصویر<sup>۸</sup> است و به صورت زیر تعریف می‌شود:

$$P_{\Omega}(x) = [P_{\Omega}(x_1), \dots, P_{\Omega}(x_n)]^T, \quad \forall i = 1, 2, \dots, n,$$

$$P_{\Omega}(x_i) = \begin{cases} l_i & x_i < l_i, \\ x_i & l_i \leq x_i \leq h_i, \\ h_i & x_i > h_i. \end{cases}$$

<sup>۸</sup>Projection Operator

با توجه به (۶.۲)، مدل شبکه عصبی برای حل (۱۳.۲) به صورت زیر به دست می‌آید:

$$\frac{d}{dt} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} P_{\Omega}(x - F(x) + A^T y + B^T z) - x \\ (y - Ax + b)^+ - y \\ -Bx + c \end{pmatrix}. \quad (14.2)$$

با توجه به قضیه‌های (۱.۳.۲)، (۳.۳.۲) و (۴.۳.۲) می‌توانیم نتایج پایداری مدل شبکه عصبی (۱۴.۲) را به صورت زیر بیان کنیم:

قضیه ۲.۴.۲ ([۴۸]). فرض کنید  $F$  مشتق‌پذیر و  $\nabla F$  برای هر  $x \in \Omega$  معین مثبت باشد، در نتیجه مدل شبکه عصبی (۱۴.۲) به مفهوم لیاپانوف پایدار و همگرا به نقطه تعادل  $u^* = (x^*, y^*, z^*)$  است، که  $x^*$  جواب بهینه (۱۳.۲) است.

نتیجه ۳.۴.۲. برای مجموعه  $\hat{S} = \{x \in \mathbb{R}^n | x \in \Omega\}$  مسأله (۱۳.۲) به مسأله مکمل زیر تبدیل می‌شود:

$$x \in \Omega, \quad F(x) \geq 0, \quad \langle x, F(x) \rangle = 0, \quad (15.2)$$

و می‌تواند مدل شبکه عصبی برای حل (۱۵.۲) به شکل زیر بیان شود:

$$\frac{dx}{dt} = P_{\Omega}(x - F(x)) - x. \quad (16.2)$$

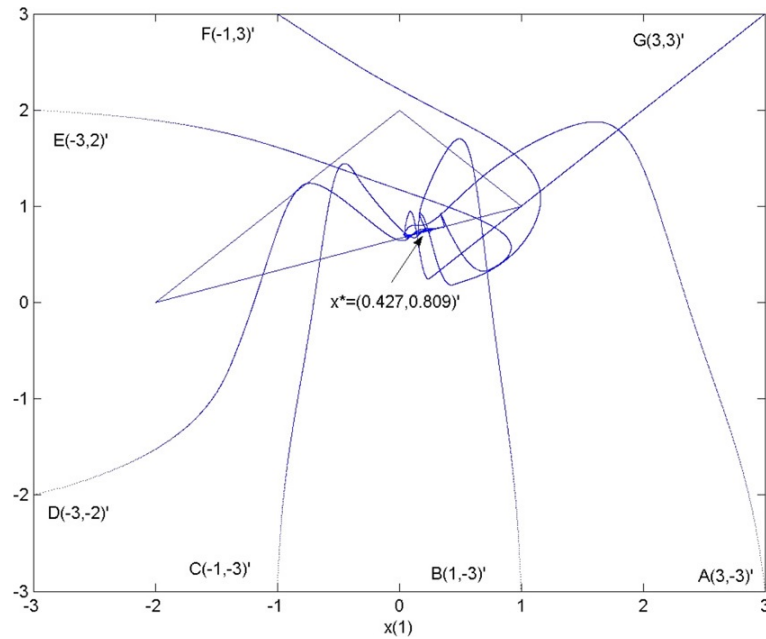
## ۵.۲ مثال‌های عددی

در این بخش چند مثال عددی به منظور نشان دادن کارایی شبکه‌های عصبی (۶.۲)، (۱۴.۲) و (۱۶.۲) ارائه شده است. تمام این مثال‌ها به وسیله نرم افزار Matlab حل شده است.

مثال ۱.۵.۲ ([۴۸]). مسأله برنامه‌ریزی غیرخطی زیر را در نظر بگیرید:

$$\begin{aligned} \min \quad & f(x) = \frac{1}{4}x_1^4 + 0.5x_1^2 + \frac{1}{4}x_2^4 + 0.5x_2^2 - 0.9x_1x_2 \\ \text{s.t} \quad & -x_1 - x_2 \geq -2, \\ & x_1 - x_2 \geq -2, \\ & x_1 - 3x_2 = -2, \\ & x \geq 0. \end{aligned}$$

جواب بهینه این مسأله  $x^* = (0.427, 0.809)^T$  است. توجه کنید که  $\nabla^2 f(x) = \nabla F(x)$  روی  $\mathbb{R}_+^2$  معین مثبت است. با توجه به قضیه (۴.۳.۲) می‌دانیم که مدل شبکه عصبی (۶.۲) به  $x^*$  همگراست. شکل (۲.۲) رفتار گذرا  $x(t)$  با هفت نقطه اولیه  $A(3, -3)$ ،  $B(1, -3)$ ،  $C(-1, -3)$ ،  $D(-3, -2)$ ،  $E(-3, 2)$  و  $F(-1, 3)$ ،  $G(3, 3)$  در نظر گرفته شده است.



شکل ۲.۲: مثال (۱.۵.۲): رفتار گذرا  $x(t) = (x_1(t), x_2(t))^T$  از مدل شبکه عصبی (۶.۲).

مثال ۲.۵.۲ ([۴۸]). مسأله برنامه‌ریزی غیرخطی محدب زیر را در نظر بگیرید:

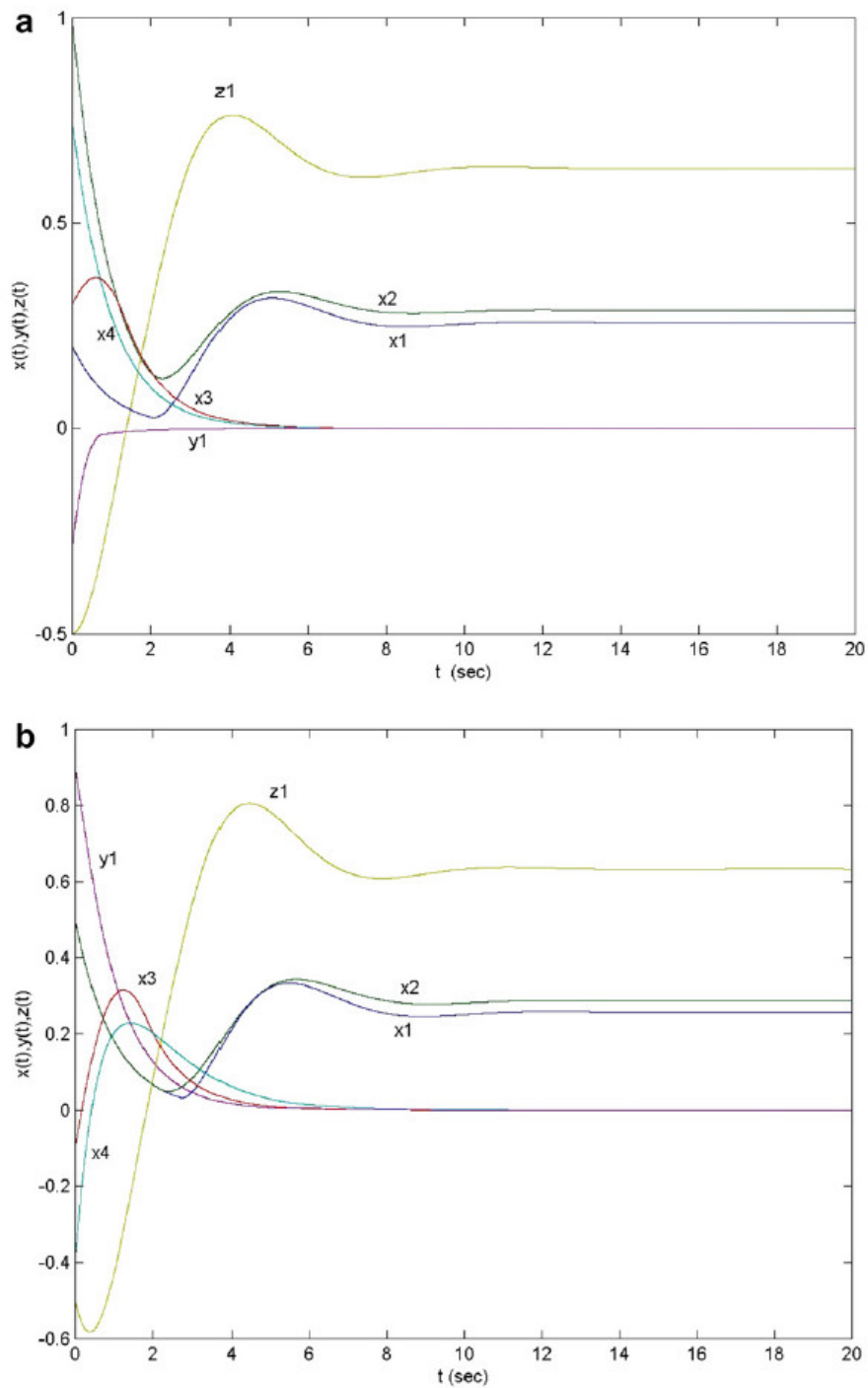
$$\begin{aligned} \min \quad & f(x) = 0.4x_1 + x_2^2 + x_3^2 - x_1x_2 + 0.5x_2^2 + 0.5x_3^2 + \frac{x_1^2}{3} \\ \text{s.t.} \quad & -0.5x_1 - x_2 + x_3 \geq -0.5, \\ & x_1 + 0.5x_2 - x_3 = 0.4, \\ & x \geq 0. \end{aligned}$$

جواب بهینه این مسأله با استفاده از  $F(x) = \nabla f(x)$  و  $S = \{x \in \mathbb{R}^3 \mid -0.5x_1 - x_2 + x_3 \geq -0.5, x_1 + 0.5x_2 - x_3 = 0.4, x \geq 0\}$  به طوری

که

$$F(x) = \begin{bmatrix} 0.4 + 2x_1 - x_2 + 0.1x_1^2 \\ 2x_2 - x_1 \\ x_3 \\ x_3 \end{bmatrix}, \quad \nabla F(x) = \begin{bmatrix} 2 + 0.2x_1 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

توجه کنید که  $\nabla F(x)$  روی  $\mathbb{R}_+^n$  معین مثبت است. با توجه به قضیه (۴.۳.۲) می‌بینیم که مدل شبکه عصبی (۶.۲) به  $x^*$  همگراست. شکل (۳.۲) مسیرهای شبکه عصبی (۶.۲) را با نقاط اولیه شدنی  $(0.2, 1, 0.3, 0.75, -0.3, -0.5)^T$  و نقاط اولیه نشدنی  $(0.5, 0.5, -0.1, -0.4, 0.9, -0.5)^T$  نشان می‌دهد، که همگرای به جواب دقیق  $x^*$  است.



شکل ۳.۲: مثال (۲.۵.۲): رفتار گذرا مدل شبکه عصبی (۶.۲) برای دو حالت مختلف: نقاط (a) اولیه شدنی (b) نقاط اولیه نشدنی.



مثال ۳.۵.۲ ([۴۸]). مسأله نامساوی وردشی غیرخطی  $VI(F, \tilde{S})$  را در نظر بگیرید. نگاشت  $F$  و مجموعه محدودیت‌های  $\tilde{S}$  را به صورت زیر در نظر بگیرید:

$$F(x) = \begin{bmatrix} 3x_1 - \frac{1}{x_1} + 3x_2 - 2 \\ 3x_1 + 3x_2 \\ 4x_3 + 4x_4 \\ 4x_3 + 4x_4 - \frac{1}{x_4} - 3 \end{bmatrix}, \quad \tilde{S} = \{x \in \mathbb{R}^n | x_1 + x_2 = 1, x_3 + x_4 \geq 0, l \leq x \leq h\},$$

که در آن  $h = (1, 0, 1, 0)^T$  و  $l = (0, 1, 0, 0)^T$ . این مسأله جواب یکتای  $x^* = (1, 0, 0, 1)^T$  دارد. توجه کنید که

$$\nabla F = \begin{bmatrix} 3 + \frac{1}{x_1^2} & 3 & 0 & 0 \\ 3 & 3 & 0 & 0 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 4 & 4 + \frac{1}{x_4^2} \end{bmatrix},$$

برای  $l \leq x \leq h$  معین مثبت است. بر طبق قضیه (۲.۴.۲) می‌دانیم که مدل شبکه عصبی (۱۴.۲) به  $x^*$  همگراست. شکل (۴.۲) مسیرهای شبکه عصبی (۱۴.۲) را با نقاط اولیه شدنی  $(0, 1, 0, 0)^T$  و نقاط اولیه نشدنی  $(-2, 5, -1, -0.5, 2, 1)^T$  را نشان می‌دهد، که هر دو همگرا به جواب دقیق  $x^*$  است.

مثال ۴.۵.۲ ([۴۸]). مسأله نامساوی وردشی خطی زیر را در نظر بگیرید. نگاشت  $F$  و مجموعه محدودیت‌های  $S$  را به صورت زیر در نظر بگیرید:

$$F(x) = \begin{bmatrix} 22x_1 - 2x_2 + 6x_3 - 4 \\ 2x_2 - 2x_1 \\ 2x_3 + 6x_1 \end{bmatrix},$$

$$S = \{x \in \mathbb{R}^3 | x_1 - x_2 \geq 1, -3x_1 - x_3 \geq -4, 2x_1 + 2x_2 + x_3 = 0, l \leq x \leq h\},$$

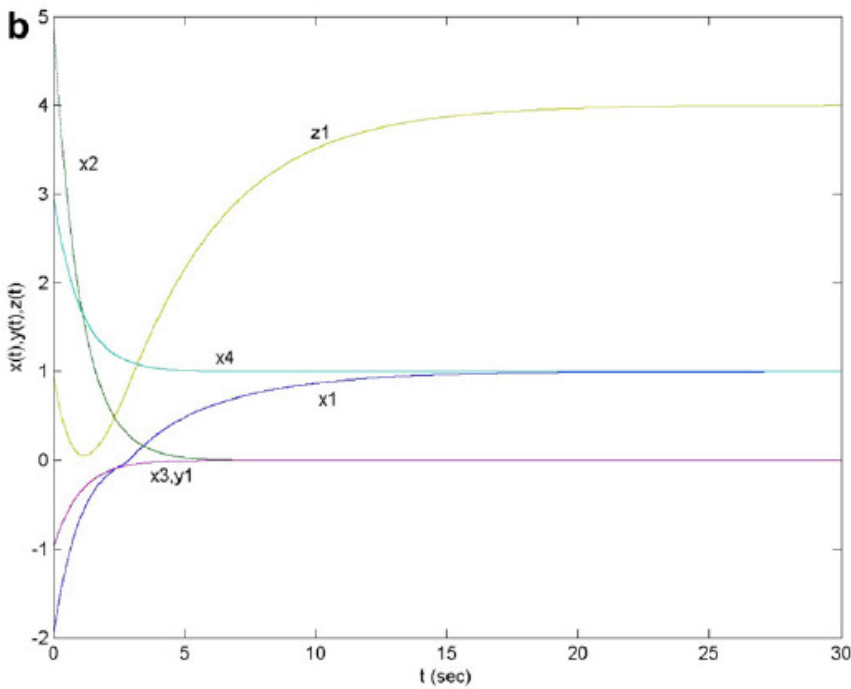
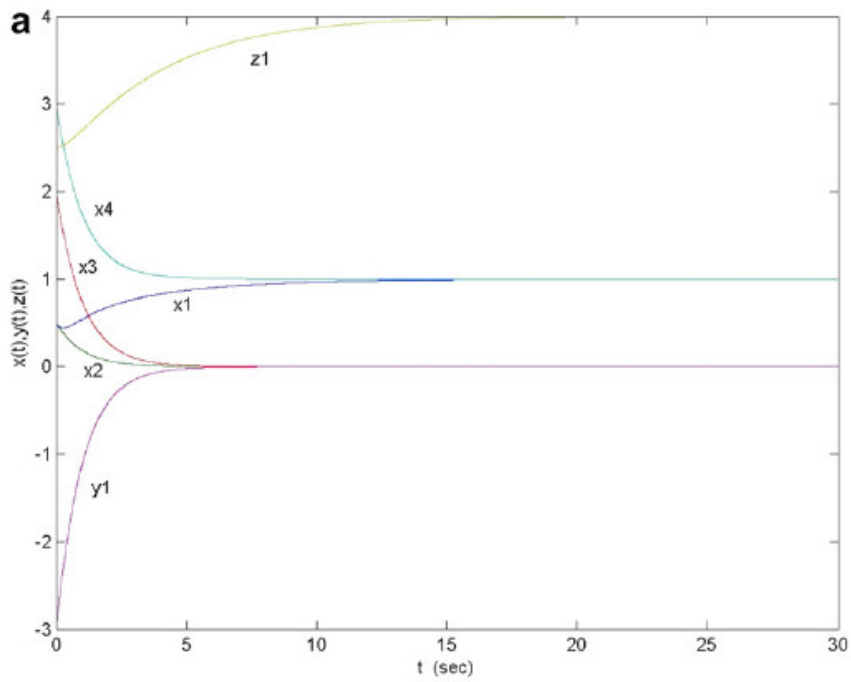
که در آن  $h = (6, 6, 6)^T$  و  $l = (-6, -6, -6)^T$ . این مسأله جواب یکتای  $x^* = (2, 1, -6)^T$  را دارد. توجه کنید که

$$\nabla F = \begin{bmatrix} 22 & -2 & 6 \\ -2 & 2 & 0 \\ 6 & 0 & 2 \end{bmatrix},$$

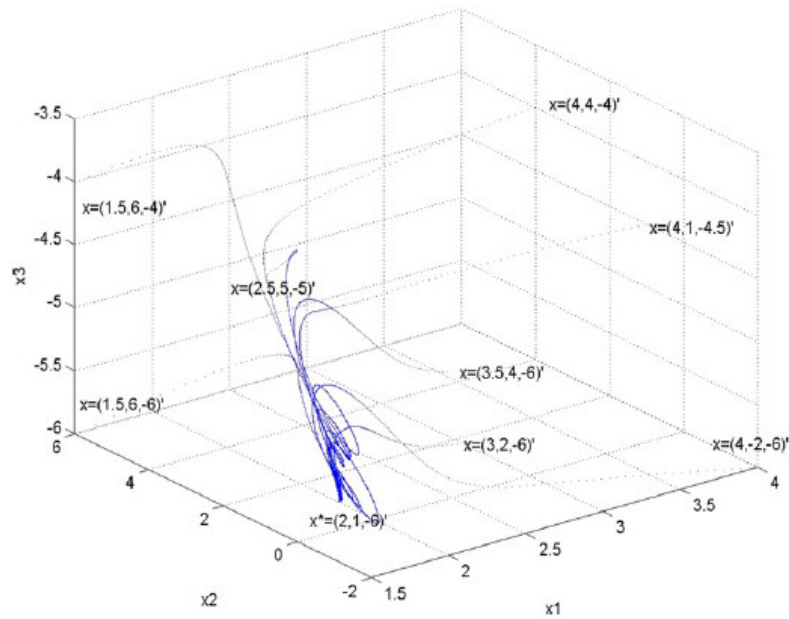
برای  $l \leq x \leq h$  معین مثبت است. بر طبق قضیه (۲.۴.۲) می‌دانیم که مدل شبکه عصبی (۱۴.۲) به  $x^*$  همگراست. شکل (۵.۲) هشت مسیر شروع شده با نقاط شروع مختلف دلخواه را نشان می‌دهد، که این مسیرها همگرا به جواب دقیق  $x^*$  است.

مثال ۵.۵.۲ ([۷۰]). مسأله مکمل خطی زیر را در نظر بگیرید:

$$x \geq 0, \quad Qx + \theta \geq 0, \quad x^T(Qx + \theta) = 0 \quad (17.2)$$



شکل ۴.۲: مثال (۳.۵.۲): رفتار گذرا مدل شبکه عصبی (۱۴.۲) برای دو حالت مختلف: (a) نقاط اولیه شدنی (b) نقاط اولیه نشدنی.



شکل ۵.۲: مثال (۴.۵.۲): نتیجه شبیه‌سازی مدل شبکه عصبی (۱۴.۲) با هشت نقطه اولیه مختلف.

که در آن

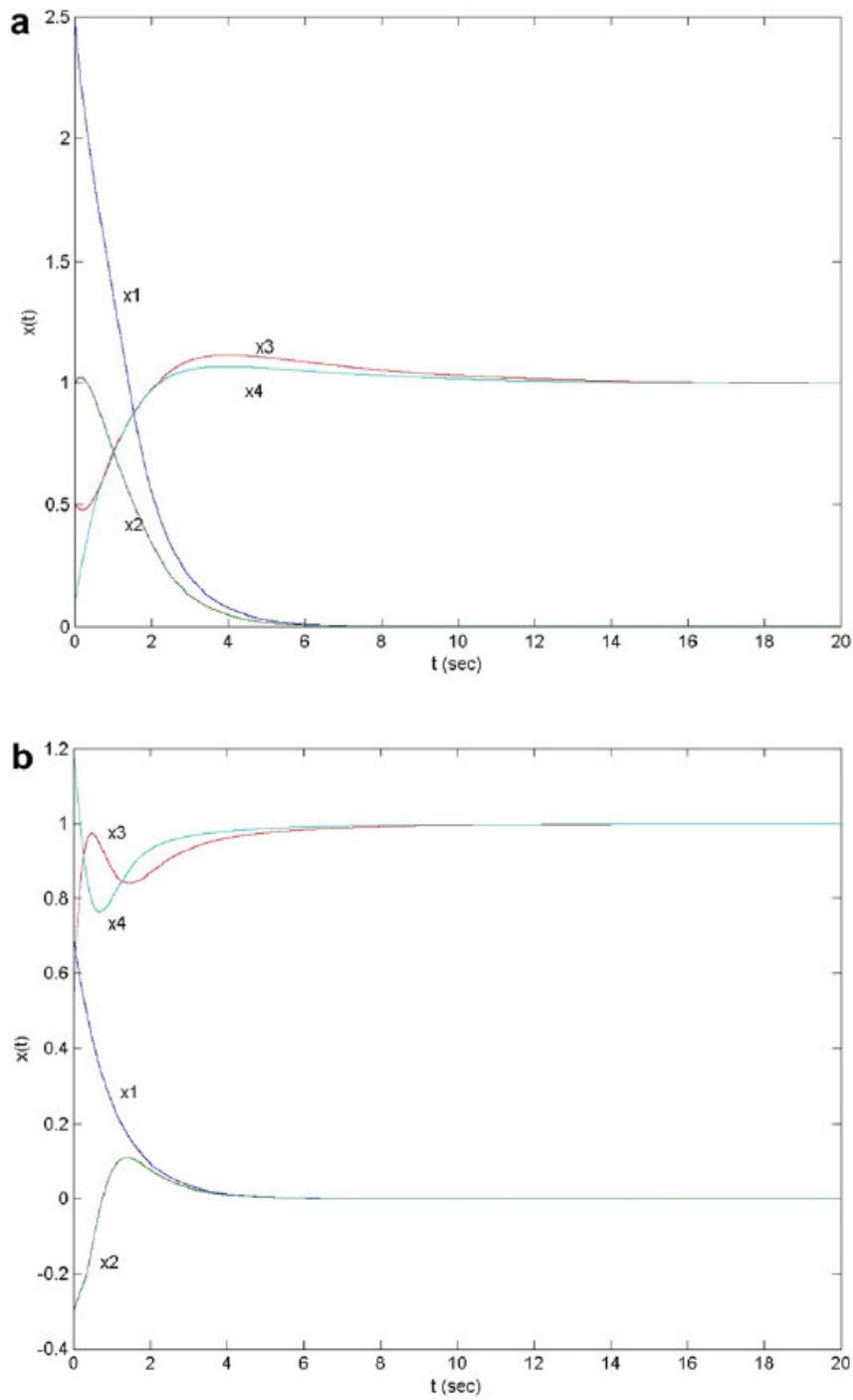
$$Q = \begin{bmatrix} 1 & -2 & 1 & -1 \\ 0 & 1 & -1 & 2 \\ -1 & 1 & 2 & -3 \\ 1 & -2 & -1 & 2 \end{bmatrix}, \quad \theta = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}.$$

این مسأله جواب یکتای  $x^* = (0, 0, 1, 1)^T$  را دارد. شکل (۶.۲) مسیرهای شبکه عصبی (۱۴.۲) را با نقاط اولیه  $(2/5, 1, 0/5, 0)^T$  و  $(0/7, -0/3, 0/6, 1/2)^T$  نشان می‌دهد.

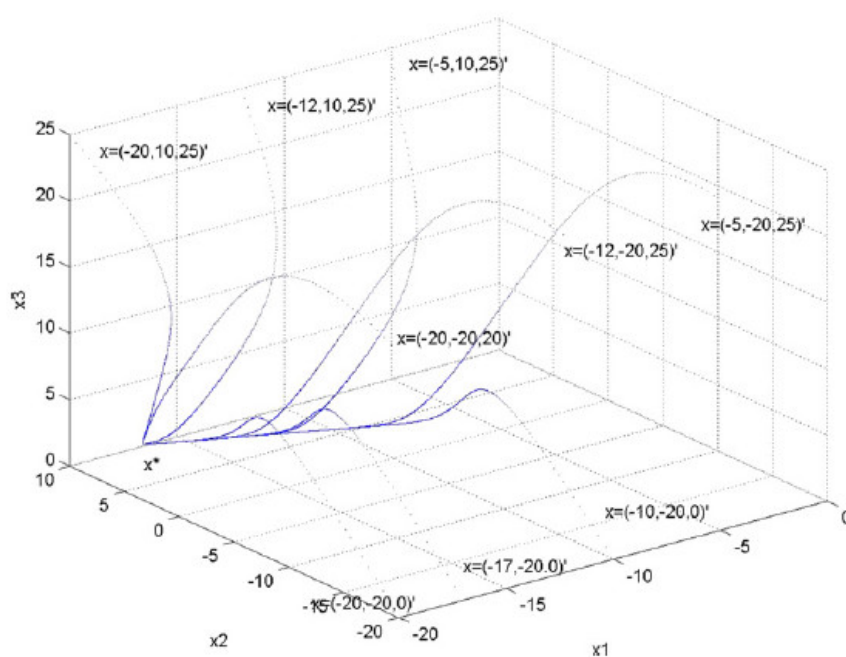
مثال ۶.۵.۲ ([۷۲]). مسأله مکمل خطی (۱۵.۲) را به طوری که  $F(x) = Qx + \theta$

$$Q = \begin{bmatrix} 0/180 & 0/648 & 0/288 \\ 0/648 & 0/880 & 0/720 \\ 0/288 & 0/720 & 0/720 \end{bmatrix}, \quad \theta = \begin{bmatrix} 0/4 \\ 0/2 \\ 0/3 \end{bmatrix},$$

برای  $l = (-20, -20, -20)^T$  و  $h = (20, 20, 20)^T$  در نظر بگیرید. این مسأله جواب یکتای  $x^* = (-20/0000, 3/3793, 4/2043)^T$  را دارد. شکل (۷.۲) نه مسیر شروع شده با نقاط شروع مختلف دلخواه را نشان می‌دهد، که این مسیرها همه همگرا به جواب دقیق  $x^*$  هستند.



شکل ۶.۲: مثال (۵.۵.۲): رفتار گذرا مدل شبکه عصبی (۱۴.۲) برای دو مورد مختلف: (a) نقاط اولیه شدنی (b) نقاط اولیه نشدنی.



شکل ۷.۲: مثال (۶.۵.۲): نتیجه شبیه‌سازی مدل شبکه عصبی (۱۴.۲) با نه نقطه اولیه مختلف.

## فصل ۳

# یک شبکه عصبی لاگرانژی برای حل مسائل نامساوی وردشی مخروطی

### ۱.۳ مقدمه

در این فصل از پایان‌نامه ابتدا مسائل نامساوی‌های وردشی مخروطی و مکمل<sup>۱</sup> را معرفی کرده و سپس ساختار یک مدل از شبکه‌های عصبی، بر پایه قضایای دوگانی، بهینه‌سازی، تجزیه و تحلیل توابع محدب، پایداری لیاپانوف و اصل تغییرناپذیری لسلال برای حل مسائل نامساوی وردشی مخروطی و مکمل ارائه می‌شود. ثابت می‌شود که نقطه‌ی تعادل شبکه‌ی عصبی پیشنهاد شده همان جواب بهینه‌ی مسأله نامساوی وردشی مخروطی و مکمل است و برعکس. همچنین نشان داده می‌شود که شبکه‌ی عصبی پیشنهادی پایدار به مفهوم لیاپانوف و همگرای سراسری به جواب بهینه مسأله اصلی است. کارآمدی مدل پیشنهادی با ارائه چندین مثال نشان داده می‌شود.

قابل ذکر است که این مدل شبکه عصبی فقط برای حل مسائل بهینه‌سازی غیرخطی محدب کاربرد دارد و از آن جهتی که مسائل نامساوی وردشی نوع خاصی از مسائل بهینه‌سازی غیرخطی محدب هستند، می‌توان از این مدل شبکه عصبی برای حل این مسأله استفاده کرد.

### ۲.۳ بیان مسأله

نظریه نامساوی وردشی اولین بار در سال ۱۹۶۶ به وسیله هارتمن<sup>۲</sup> و استمپکیا<sup>۳</sup> به عنوان ابزاری برای مطالعه معادلات دیفرانسیل جزئی با کاربردهایی در مکانیک معرفی شد [۲۸، ۳۱، ۳۲، ۳۹، ۵۶، ۵۹]. نامساوی‌های وردشی در اقتصاد، علم مدیریت، تحقیق در عملیات و همچنین در مهندسی کاربردهای

<sup>۱</sup>Complementarity and Variational Inequalities

<sup>۲</sup>Hartman

<sup>۳</sup>Stampacchia

فراوانی دارند. مسائل نامساوی وردشی مخروطی مسائلی هستند که محدودیت‌های آن‌ها شامل ضرب دکارتی مخروط‌های مرتبه دوم است.

پیدا کردن نقطه  $x \in C$  را در نظر می‌گیریم به طوری که:

$$\langle F(x), y - x \rangle \geq 0, \quad \forall y \in C, \quad (1.3)$$

که در آن  $C$  مجموعه متناهی است و به صورت

$$C = \{x \in \mathbb{R}^n : h(x) = 0, \quad -g(x) \in \mathcal{K}\}, \quad (2.3)$$

تعریف می‌شود، که  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n, h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  و  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  توابع به‌طور پیوسته مشتق‌پذیر هستند. همچنین فرض کنید  $\nabla F$  یک ماتریس نیمه معین مثبت است.

$\mathcal{K}$  ضرب دکارتی مخروط‌های مرتبه دوم است و به صورت زیر تعریف می‌شود:

$$\mathcal{K} = \mathcal{K}^{m_1} \times \mathcal{K}^{m_2} \times \dots \times \mathcal{K}^{m_p}, \quad (3.3)$$

که در آن

$$m_1 + m_2 + \dots + m_p = m, \quad m_1, m_2, \dots, m_p \geq 1, \quad l \geq 0,$$

و

$$\mathcal{K}^{m_i} := \{(x_{i1}, x_{i2}, \dots, x_{im_i})^T \in \mathbb{R}^{m_i} \mid \|(x_{i2}, \dots, x_{im_i})\| \leq x_{i1}\},$$

که در آن  $\|\cdot\|$  نشان دهنده نرم اقلیدسی است. حالت خاصی از (۳.۳) حالتی است که  $\mathcal{K} = \mathbb{R}_+^n$  (یعنی مقادیر مثبت در  $\mathbb{R}^n$ )، که مربوط به  $p = n$  و  $m_1 = \dots = m_p = 1$  است. زمانی که  $h$  آفین است، مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (۱.۳)-(۳.۳) معادل مسأله برنامه‌ریزی محدب مخروط مرتبه دوم زیر است [۴۴]:

$$\min f(x) \quad (4.3)$$

$$\text{s.t. } Ax = b, \quad (5.3)$$

$$-g(x) \in \mathcal{K}, \quad (6.3)$$

که در آن  $A \in \mathbb{R}^{l \times n}$  رتبه سطری کامل دارد،  $b \in \mathbb{R}^l$ ،  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ،  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  و  $F = \nabla f$  مسأله (۴.۳)-(۶.۳) معادل با مسأله زیر است:

$$\min f(x) \quad (7.3)$$

$$\text{s.t. } Ax = b, \quad (8.3)$$

$$\sqrt{g_{i2}^2 + \dots + g_{im}^2} + \epsilon \leq g_{i1}, \quad i = 1, \dots, p. \quad (9.3)$$

از طرفی مسأله (۴.۳)-(۶.۳) معادل مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (SOCCVI) زیر است:

پیدا کردن  $x \in C$  را چنان در نظر بگیرید که در رابطه زیر صدق کند:

$$\langle \nabla f(x), y - x \rangle \geq 0 \quad \forall y \in C, \quad (10.3)$$

که در آن

$$C = \{x \in \mathbb{R}^n : Ax - b = 0, \quad -g(x) \in \mathcal{K}\}. \quad (11.3)$$

در نتیجه حل نامساوی وردشی (۱۰.۳) معادل با حل مسأله بهینه‌سازی محدب (۷.۳)–(۹.۳) است. بنابراین مسأله برنامه‌ریزی محدب را به صورت کلی زیر در نظر می‌گیریم:

$$\min f(x) \quad (12.3)$$

$$s.t. \quad h(x) = 0, \quad (13.3)$$

$$g(x) \leq 0, \quad (14.3)$$

که در آن  $h(x) = Ax - b$  و  $g(x) \leq 0$  قیود مخروطی مرتبه دوم مشتق‌پذیر (۹.۳) را نشان می‌دهند و  $g(x) \in \mathbb{R}^m$ .

برای حل مسائل نامساوی وردشی مقید و مکمل روش‌های محاسباتی زیادی ارائه شده‌اند [۲۷، ۳۵، ۳۶، ۳۷، ۶۶، ۶۹]. این روش‌ها شامل روش‌هایی بر پایه تابع ایده‌آل<sup>۴</sup>، روش نقطه داخلی<sup>۵</sup>، روش نیوتن<sup>۶</sup>، روش معادله غیرخطی<sup>۷</sup>، روش تصویر<sup>۸</sup> و نسخه‌های مختلف آن است. علاوه‌براین مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (SOCCVI) می‌تواند به وسیله یافتن جواب شرایط کاروش-کان-تاگر متناظر این مسأله حل شود.

### ۳.۳ مدل شبکه عصبی

تابع لاگرانژ متناظر با (۱۲.۳)–(۱۴.۳) را مشابه با [۷۵] به صورت زیر در نظر می‌گیریم:

$$L(x, u, v) = f(x) + \frac{1}{\nu} \sum_{k=1}^m u_k^* g_k(x) + \sum_{p=1}^l v_p h_p(x). \quad (15.3)$$

از [۵۱] می‌دانیم که  $x^* \in \mathbb{R}^n$  یک جواب بهینه برای (۱۲.۳)–(۱۴.۳) است اگر و تنها اگر  $u^* \in \mathbb{R}^m$  و  $v^* \in \mathbb{R}^l$  وجود داشته باشند به قسمی که  $(x^{*T}, u^{*T}, v^{*T})^T$  در شرایط KKT زیر صدق کند:

$$\begin{cases} u^* \geq 0, \quad g(x^*) \leq 0, \quad u^{*T} g(x^*) = 0, \\ \nabla f(x^*) + \nabla g(x^*)^T u^* + \nabla h(x^*)^T v^* = 0, \\ h(x^*) = 0. \end{cases} \quad (16.3)$$

$x^*$  یک نقطه KKT مسأله (۱۲.۳)–(۱۴.۳) نامیده می‌شود و  $(u^{*T}, v^{*T})^T$  بردار ضرب‌گر لاگرانژ متناظر با  $x^*$  نامیده می‌شود.

<sup>۴</sup>Merit Function

<sup>۵</sup>Interior Method

<sup>۶</sup>Newton Method

<sup>۷</sup>Nonlinear Equation Method

<sup>۸</sup>Projection Method



لم ۱.۳.۳ ([۵۱]). اگر  $f$  و  $g_k$ ،  $k = 1, \dots, m$  محدب باشند، آنگاه  $x^*$  جواب بهینه مسأله (۱۲.۳) - (۱۴.۳) است اگر و تنها اگر  $x^*$  نقطه KKT مسأله (۱۲.۳) - (۱۴.۳) باشد.

اکنون یک شبکه عصبی برای حل (۱۲.۳) - (۱۴.۳) و دوگان آن با سیستم دینامیکی زیر را معرفی می‌کنیم:

$$\frac{dx_i}{dt} = -\nabla_{x_i} L(x, u, v) = -\left( \frac{\partial f}{\partial x_i} + \sum_{k=1}^m u_k \frac{\partial g_k}{\partial x_i} + \sum_{p=1}^l v_p \frac{\partial h_p}{\partial x_i} \right), \quad i = 1, \dots, n, \quad (17.3)$$

$$\frac{du_k}{dt} = \nabla_{u_k} L(x, u, v) = u_k g_k(x), \quad u_k(t_0) > 0 \quad k = 1, \dots, m, \quad (18.3)$$

$$\frac{dv_p}{dt} = \nabla_{v_p} L(x, u, v) = h_p(x), \quad p = 1, \dots, l, \quad (19.3)$$

با نقطه آغازین  $(x_0^T, u_0^T, v_0^T)^T$ . توجه کنید در شبکه عصبی (۱۷.۳) - (۱۹.۳) ضرب‌گرهای متناظر با محدودیت‌های نامساوی به صورت  $u_k$  ( $k = 1, \dots, m$ ) تعریف شده‌اند. حال با در نظر گرفتن  $W^*$ ،  $w = (x^T, u^T, v^T)^T \in \mathbb{R}^{n+m+l}$  به‌عنوان مجموعه نقاط بهینه (۱۲.۳) - (۱۴.۳) و دوگان آن در نظر گرفته شده، و

$$\Phi(x, u, v) = \begin{bmatrix} -\left( \nabla f(x) + \sum_{k=1}^m \nabla g_k(x)^T u_k + \nabla h(x)^T v \right) \\ \text{diag}(u_1, \dots, u_m) g(x) \\ h(x) \end{bmatrix}.$$

شبکه عصبی (۱۷.۳) - (۱۹.۳) را می‌توان به صورت زیر نوشت:

$$\frac{dw}{dt} = \Phi(w), \quad (20.3)$$

$$w(t_0) = w_0, \quad u(t_0) > 0. \quad (21.3)$$

### ۴.۳ تحلیل همگرایی و پایداری

در این بخش پایداری و همگرایی شبکه عصبی (۲۰.۳) و (۲۱.۳) را مورد بررسی قرار می‌دهیم.

قضیه ۱.۴.۳. فرض کنید  $w^* = (x^{*T}, u^{*T}, v^{*T})^T$  نقطه تعادل شبکه عصبی (۲۰.۳) و (۲۱.۳) باشد. آنگاه  $x^*$  نقطه KKT مسأله (۱۲.۳) - (۱۴.۳) است.

برهان. با مقایسه شرایط KKT (۱۶.۳) با سیستم دینامیکی (۲۰.۳) و (۲۱.۳) می‌بینیم آن‌چه باید اثبات شود این است که  $g_k(x^*) \leq 0$  و  $u_k^* g_k(x^*) = 0$  ( $k = 1, \dots, m$ ). جواب معادله دیفرانسیل (۱۸.۳) به‌صورت زیر است:

$$u_k(t) = u_k(t_0) e^{\int_{t_0}^t g_k(x) dt}, \quad k = 1, \dots, m,$$

که  $u_k(t_0) > 0$  یک مقدار آغازین مثبت برای  $u_k(t)$  است. اگر  $x$  متعلق به ناحیه شدنی نباشد یعنی حداقل یک  $k$  وجود داشته باشد که  $g_k(x) > 0$ ، ضرب‌گرهای متناظر  $u_k(t)$  به صورت نمایی با زمان افزایش می‌یابد، و در نتیجه عبارت  $\frac{1}{\rho} u_k^* g_k(x)$  در تابع لاگرانژ (۱۵.۳) با زمان افزایش می‌یابد. بنابراین بعد از یک زمان متناهی منحنی  $x(t)$  به سمت مجموعه شدنی حرکت می‌کند. این روند به طور پیوسته تکرار خواهد شد مگر آن‌که تمام محدودیت‌ها برقرار شوند. بنابراین  $(x^T, u^T, v^T)^T$  سرانجام به نقطه KKT  $(x^{*T}, u^{*T}, v^{*T})^T$  میل می‌کند، در حالی که ضرب‌گرهای  $u_k(t)$  متناظر با محدودیت‌های غیرفعال ( $g(x) < 0$ ) به سمت صفر میل می‌کند، و باقیمانده ضرب‌گرها متناظر با محدودیت‌های فعال به اعداد ثابت میل می‌کنند. بنابراین برهان کامل است.  $\square$

لم ۲.۴.۳. نقطه تعادل شبکه عصبی معرفی شده (۲۰.۳) و (۲۱.۳) یکتاست.

برهان. از آن‌جا که مسأله (۱۲.۳)–(۱۴.۳) دارای جواب بهینه منحصر به فرد  $x^*$  است، شرایط KKT (۱۶.۳) دارای جواب یکتا  $w^* = (x^{*T}, u^{*T}, v^{*T})^T$  است [۵۱]. از طرفی بنا به قضیه (۱.۴.۳) نقطه تعادل شبکه عصبی (۲۰.۳) و (۲۱.۳) در شرایط KKT (۱۶.۳) صدق می‌کند. بنابراین نقطه تعادل شبکه عصبی (۲۰.۳) و (۲۱.۳) یکتاست.  $\square$

لم ۳.۴.۳. برای هر نقطه آغازین  $w(t_0) = (x(t_0)^T, u(t_0)^T, v(t_0)^T)^T$ ، یک جواب یکتا و پیوسته  $w(t) = (x(t)^T, u(t)^T, v(t)^T)^T$  برای سیستم دینامیکی (۲۰.۳)–(۲۱.۳) وجود دارد.

برهان. چون  $\nabla f(x)$ ،  $\nabla g_k(x)$  ( $k = 1, \dots, m$ ) و  $\nabla h_p(x)$  ( $p = 1, \dots, l$ ) روی مجموعه محدب و باز  $D \subseteq \mathbb{R}^{n+m+l}$  به طور پیوسته مشتق‌پذیر هستند، بنابراین  $\nabla f(x) + \frac{1}{\rho} \nabla g(x)^T u^2 + \sum_{p=1}^l v_p \nabla h_k(x)$ ،  $h(x)$  و  $u_k g_k(x)$  ( $k = 1, \dots, m$ ) پیوسته لیپ‌شیتز محلی هستند و طبق قضیه (۱۶.۳.۱)، شبکه عصبی (۲۰.۳) و (۲۱.۳) دارای جواب پیوسته یکتا  $w(t)$ ، برای  $t \in [t_0, \tau)$ ، به طوری که  $\tau \rightarrow \infty$  می‌باشد.  $\square$

قضیه ۴.۴.۳. اگر  $\nabla^2 f(x)$  معین مثبت و  $\nabla^2 g_k(x)$  ( $k = 1, \dots, m$ ) نیمه‌معین مثبت باشد یا اگر  $\nabla^2 f(x)$  نیمه‌معین مثبت و  $\nabla^2 g_k(x)$  ( $k = 1, \dots, m$ ) معین مثبت باشد، آنگاه شبکه عصبی (۲۰.۳) و (۲۱.۳) پایدار سراسری به مفهوم لیاپانوف و همگرایی سراسری به یک نقطه KKT  $w^* = (x^{*T}, u^{*T}, v^{*T})^T$  می‌باشد به طوری که  $x^*$  جواب بهینه مسأله (۱۲.۳)–(۱۴.۳) و  $(u^{*T}, v^{*T})^T$  جواب بهینه دوگان آن است.

برهان. تابع لیاپانوف زیر را در نظر می‌گیریم:

$$E(w) = \|\Phi(w)\|^2 + \frac{1}{\rho} \|w - w^*\|^2. \quad (22.3)$$

بنابراین

$$\frac{d\Phi}{dt} = \frac{\partial \Phi}{\partial w} \frac{dw}{dt} = \nabla \Phi(w) \Phi(w),$$

که ماتریس ژاکوبین نگاشت  $\Phi$  به صورت زیر به دست می آید  $[\Lambda]$ :

$$\nabla\Phi(w) = \begin{bmatrix} -\nabla^T f(x) - \frac{1}{\nu} \sum_{k=1}^m u_k^T \nabla^T g_k(x) & -\nabla g(x)^T A & -\nabla h(x)^T \\ A^T \nabla g(x) & B & O_{m \times l} \\ \nabla h(x) & O_{l \times m} & O_{l \times l} \end{bmatrix}, \quad (23.3)$$

که  $O$  ماتریس صفر،  $A = \text{diag}(u_1, \dots, u_m)$  و  $B = \text{diag}(g_1(x), \dots, g_m(m))$ . حال با استفاده از (۲۳.۳) داریم:

$$\begin{aligned} \frac{dE(w(t))}{dt} &= \left( \frac{d\Phi}{dt} \right)^T \Phi + \Phi^T \left( \frac{d\Phi}{dt} \right) + (w - w^*)^T \frac{dw(t)}{dt} \\ &= \Phi^T \left( \nabla\Phi(w)^T + \nabla\Phi(w) \right) \Phi + (w - w^*)^T \Phi(w) \\ &= -\nu \left[ \left( \nabla f(x) + \frac{1}{\nu} \nabla g(x)^T u^T + \nabla h(x)^T v \right)^T \left( \nabla^T f(x) + \frac{1}{\nu} \sum_{k=1}^m u_k^T \nabla^T g_k(x) \right) \right. \\ &\quad \left. \left( \nabla f(x) + \frac{1}{\nu} \nabla g(x)^T u^T + \nabla h(x)^T v \right) \right] + \nu \sum_{k=1}^m (u_k g_k)^T g_k + (w - w^*)^T \Phi(w) \\ &\leq \nu \sum_{k=1}^m (u_k g_k)^T g_k + (w - w^*)^T \Phi(w). \end{aligned}$$

طبق فرض قضیه  $\left( \nabla^T f(x) + \frac{1}{\nu} \sum_{k=1}^m u_k^T \nabla^T g_k(x) \right) -$  معین منفی است.

حال اگر  $g_k(x) \leq 0$ ،  $(k = 1, \dots, m)$ ، در نتیجه  $\sum_{k=1}^m (u_k g_k)^T g_k \leq 0$ . به علاوه در این حالت  $\nabla\Phi(w)$  نیمه معین منفی است و بنا به قضیه (۱۵.۳.۱)  $-\Phi(w)$  یکنواست یعنی

$$(w - w^*)^T (\Phi(w) - \Phi(w^*)) = (w - w^*)^T \Phi(w) \leq 0,$$

که  $\Phi(w^*) = 0$ . بنابراین

$$\frac{dE(w(t))}{dt} \leq 0. \quad (24.3)$$

از طرف دیگر اگر  $g_k(x(t)) > 0$  برای بعضی  $k \in \{1, \dots, m\}$ ، آن گاه طبق اثبات قضیه (۱.۴.۳) می بینیم که  $x(t)$  بعد از یک زمان متناهی به درون ناحیه شدنی میل می کند، درحالی که ضرب گره های متناظر  $u_k$  تا زمانی که  $u_k \geq 0$  افزایش می یابد. در نتیجه بعد از یک زمان متناهی و با روندی مشابه حالت قبل به دست می آید  $\frac{dE(w(t))}{dt} \leq 0$ . بنابراین شبکه عصبی (۲۰.۳) و (۲۱.۳) پایدار سراسری به مفهوم لیاپانوف است. حال چون

$$E(w) \geq \frac{1}{\nu} \|w - w^*\|^2,$$

است. برای هر نقطه آغازین  $w_0 \in \mathbb{R}^{n+m+l}$ ، منحنی جواب  $\{w(t)\}$  کراندار است.

فرض کنید  $\Lambda = \{w \in \mathbb{R}^{n+m+l} | E(w) \leq E(w_0)\}$ ،  $\Gamma = \{w(t) | \frac{dE(w(t))}{dt} = 0\}$  و  $\Pi$  بزرگترین مجموعه تغییرناپذیر در  $\Gamma$  باشد بنابراین با استفاده از اصل تغییرناپذیری لسال (۲۲.۳.۱)، منحنی های  $w(t)$  به  $\Pi$  همگرا خواهند شد، یعنی

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

اما از (۲۰.۳) و (۲۱.۳) و (۲۴.۳)،  $\frac{dw}{dt} = 0$  و تنها اگر  $\frac{dE}{dt} = 0$ ، بنابراین  $\Gamma = \{w \in \Lambda | \frac{dw}{dt} = 0\} = W^*$ ، در نتیجه شبکه عصبی (۲۰.۳) و (۲۱.۳) همگرای سراسری به جواب های بهینه مسأله (۱۲.۳) - (۱۴.۳) است.  $\square$

نتیجه ۵.۴.۳. فرض کنید شرایط قضیه (۴.۴.۳) برقرار باشد. اگر  $W^* = \{(x^{*T}, u^{*T}, v^{*T})^T\}$  شبکه عصبی (۲۰.۳) و (۲۱.۳) برای حل (۱۲.۳)-(۱۴.۳) پایدار مجانبی سراسری به نقطه تعادل  $w^* = (x^{*T}, u^{*T}, v^{*T})^T$  است.

## ۵.۳ مثال‌های عددی

در این بخش مثال‌هایی ارائه خواهد شد تا نتایجی را که در بخش‌های قبل حاصل شد نشان دهند. شبیه‌سازی این مثال‌ها با استفاده از نرم افزار Matlab، و حل کننده معادلات دیفرانسیل معمولی ode45 انتخاب شده است.

مثال ۱.۵.۳ ([۴۴]). مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (SOCCVI) زیر را در نظر بگیرید:

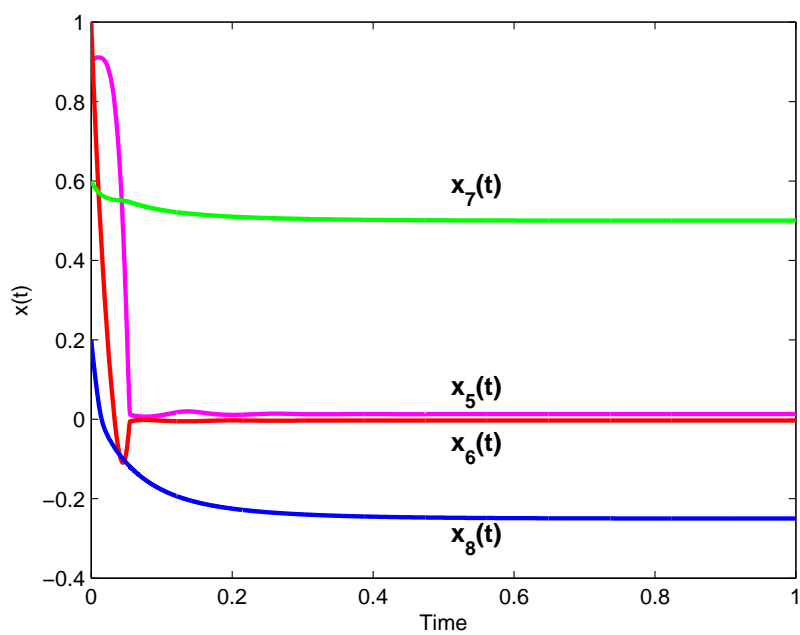
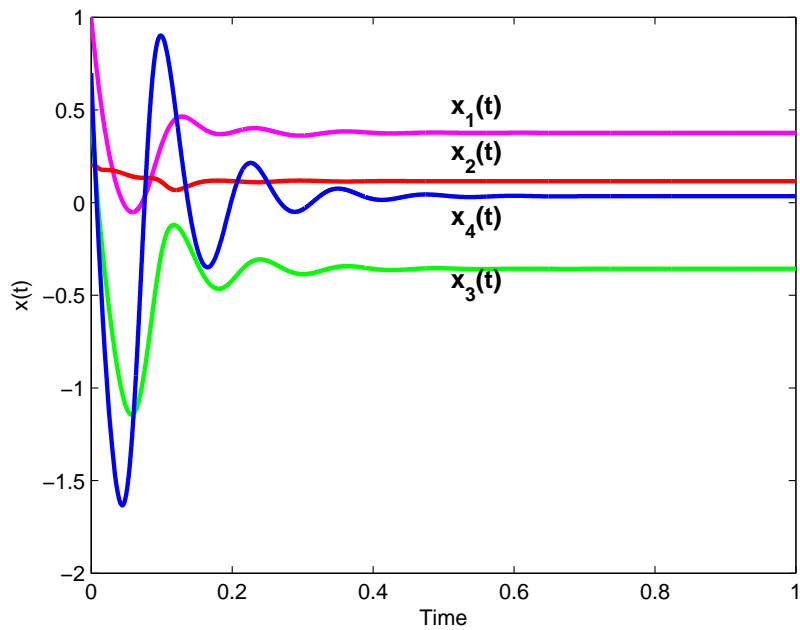
$$F(x) = \begin{pmatrix} 2x_1 + x_2 + 1 \\ x_1 + 6x_2 - x_3 - 2 \\ -x_2 + 3x_3 - \frac{\epsilon}{\delta}x_4 + 3 \\ -\frac{\epsilon}{\delta}x_3 + 2x_4 + \frac{1}{\gamma} \sin x_4 \cos x_5 \sin x_6 + 6 \\ \frac{1}{\gamma} \cos x_4 \sin x_5 \sin x_6 + 2x_5 - \frac{\delta}{\gamma} \\ -\frac{1}{\gamma} \cos x_4 \cos x_5 \cos x_6 + 2x_6 + \frac{1}{\gamma} \cos x_6 \sin x_7 \cos x_8 + 1 \\ \frac{1}{\gamma} \sin x_6 \cos x_7 \cos x_8 + 4x_7 - 2 \\ -\frac{1}{\gamma} \sin x_6 \sin x_7 \sin x_8 + 2x_8 + \frac{1}{\gamma} \end{pmatrix},$$

$$C = \{x \in \mathbb{R}^8 : -g(x) = x \in \mathcal{K}^3 \times \mathcal{K}^3 \times \mathcal{K}^2\}.$$

جواب تقریبی این مسأله

$$x^* = (0.3867, 0.1155, -0.3671, 0.0044, 0.0031, 0.0032, 0.5002, -0.2500)^T$$

است. شکل (۱.۳) رفتار گذرا از مدل شبکه عصبی (۲۰.۳) و (۲۱.۳) را برای مثال (۱.۵.۳) نشان می‌دهد.



شکل ۱.۳: رفتار گذرا شبکه عصبی در مثال (۱.۵.۳)

مثال ۲.۵.۳ ([۴۴]). مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (SOCCVI) زیر را در نظر بگیرید:

$$\langle \frac{1}{\sqrt{3}} Dx, y - x \rangle \geq 0, \quad \forall y \in C,$$

که در آن

$$C = \{x \in \mathbb{R}^n : Ax - a = 0, Bx - b \preceq 0\}.$$

$D$  ماتریس متقارن  $n \times n$  است،  $A$  و  $B$  به ترتیب ماتریس‌های  $l \times n$  و  $m \times n$  هستند،  $d$  یک بردار  $n \times 1$  است،  $a$  و  $b$  بردارهای  $l \times 1$  و  $m \times 1$  با  $l + m \leq n$  هستند. می‌توانیم داده‌های  $A, B, a, b$  را تصادفی انتخاب کنیم، اما با مراجعه به [۴۵]، مجموعه داده‌ها را به صورت زیر تعریف می‌کنیم:

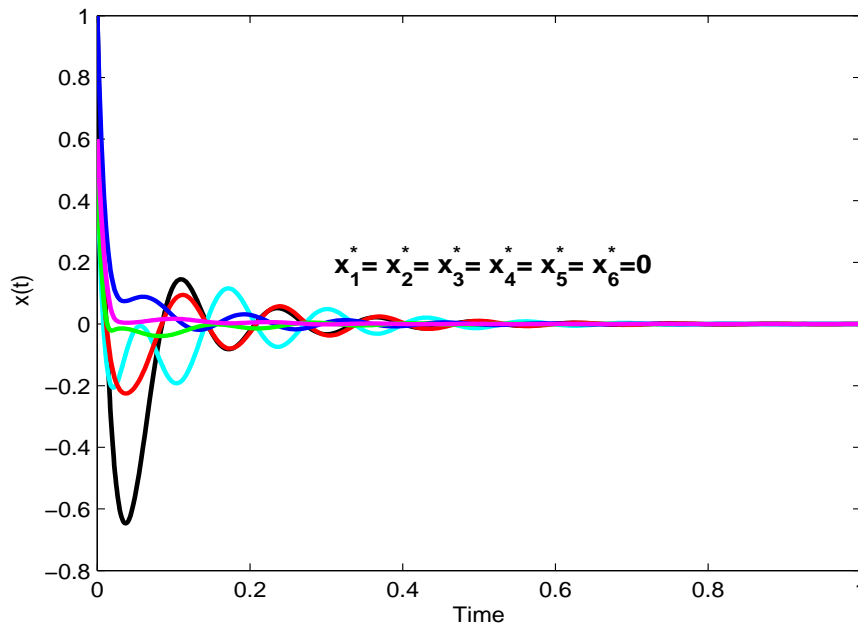
$$D = [D_{ij}]_{n \times n},$$

که در آن

$$D_{ij} = \begin{cases} 2, & i = j \\ 1, & |i - j| = 1, \\ 0, & \text{otherwise,} \end{cases}$$

$b = (e_{m_1}, e_{m_2}, \dots, e_{m_p})^T$ ,  $a = O_{l \times 1}$ ,  $B = \begin{bmatrix} O_{m \times (n-m)} & I_{m \times m} \end{bmatrix}_{m \times n}$ ,  $A = \begin{bmatrix} I_{l \times 1} & O_{l \times (n-1)} \end{bmatrix}_{l \times n}$  که  $e_{m_i} = (1, 0, \dots, 0)^T \in \mathbb{R}^{m_i}$  و  $l + m \leq n$ . رتبه  $A$  و  $B$  کامل است و  $\text{rank}([A^T \ B^T]) = l + m$  است.

پارامترهای  $l = 3$ ،  $m = 3$  و  $n = 6$  است. این مسأله جواب تقریبی  $x^* = (0, 0, 0, 0, 0, 0)^T$  را دارد. شکل (۲.۳) رفتار گذرا از مدل شبکه عصبی (۲۰.۳) و (۲۱.۳) را برای مثال (۲.۵.۳) نشان می‌دهد.



شکل ۲.۳: رفتار گذرا شبکه عصبی در مثال (۲.۵.۳)

مثال ۳.۵.۳ ([۴۴]). مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (SOCCVI) زیر را در نظر بگیرید:

$$F(x) = \begin{pmatrix} x_3 \exp(x_1 x_2) + \phi(x_1 + x_2) \\ \phi(x_1 + x_2) + \frac{2(x_2 - x_1)}{\sqrt{1 + (x_2 - x_1)^2}} \\ x_1 \exp(x_1 x_2) - \frac{2x_2 - x_1}{\sqrt{1 + (x_2 - x_1)^2}} \\ x_4 \\ x_5 \end{pmatrix},$$

و

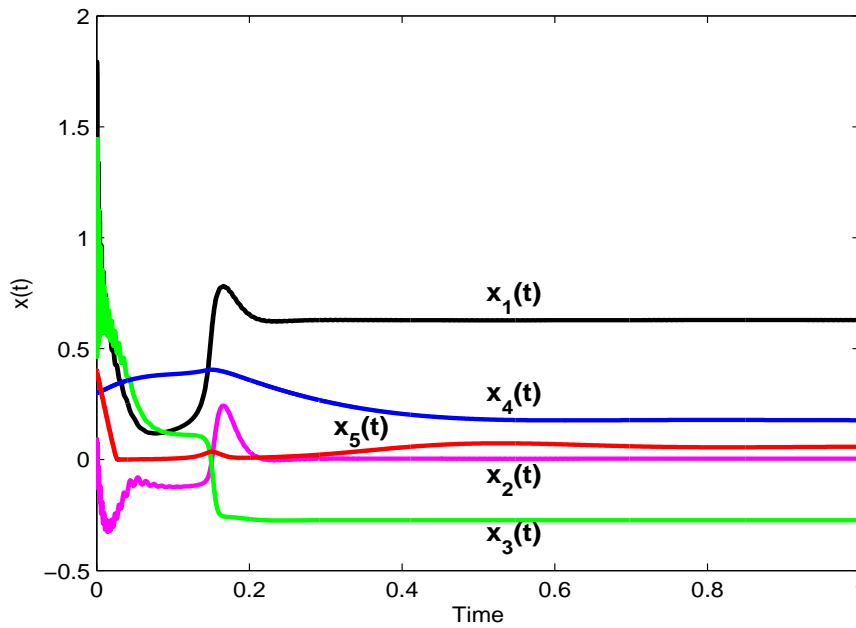
$$C = \{x \in \mathbb{R}^5 : h(x) = 0, -g(x) \in \mathcal{K}^3 \times \mathcal{K}^2\},$$

با

$$h(x) = -62x_1^3 + 58x_2 + 167x_3^3 - 29x_4 - x_5 - 3x_5 + 11,$$

$$g(x) = \begin{pmatrix} -3x_1^3 - 2x_2 + x_3 - 5x_3^3 \\ 5x_1^3 - 4x_2 + 2x_3 - 10x_3^3 \\ -x_4 \\ -x_5 \\ -3x_5 \end{pmatrix}.$$

جواب تقریبی این مسأله  $x^* = (0.6287, 0.0042, -0.2733, 0.1773, 0.0582)^T$  است. شکل (۳.۳) رفتار گذرا از مدل شبکه عصبی (۲.۵.۳) و (۲۱.۲) را برای مثال (۳.۵.۳) نشان می‌دهد.



شکل ۳.۳: رفتار گذرا شبکه عصبی در مثال (۳.۵.۳)

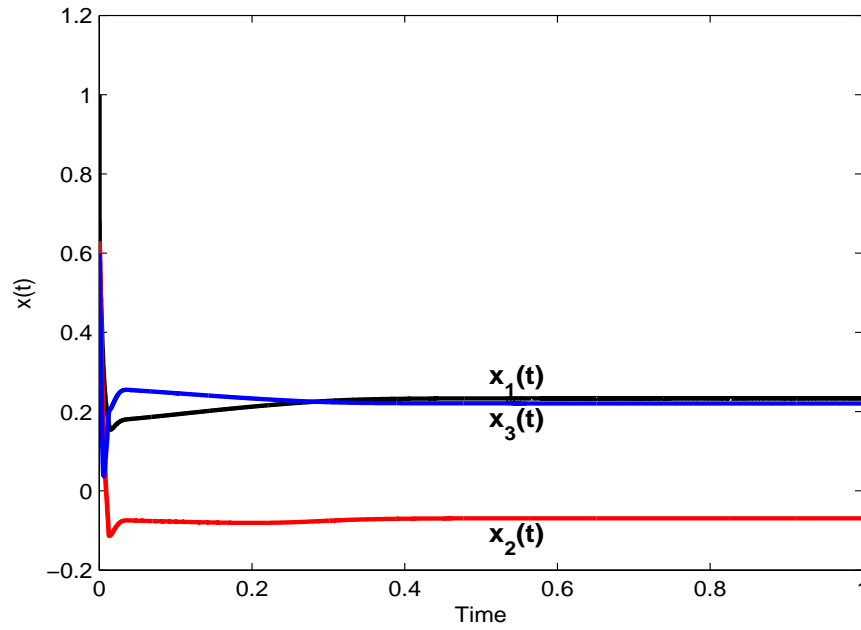
مثال ۴.۵.۳ ([۴۴]). مسأله محدب غیر خطی زیر را در نظر بگیرید:

$$\min \quad \exp(x_1 - x_3) + 3(2x_1 - x_2)^4 + \sqrt{1 + (3x_2 + 5x_3)^2}$$

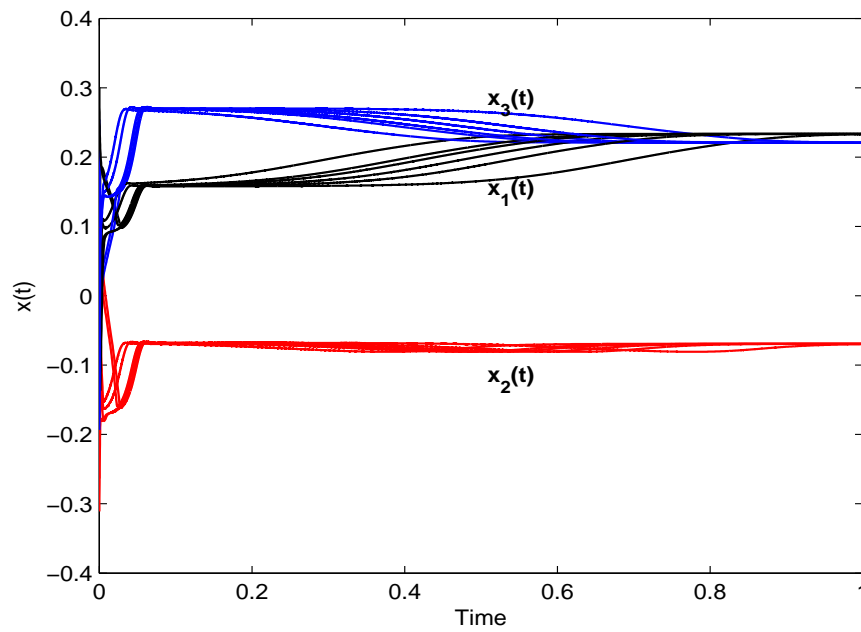
$$\text{s.t.} \quad -g(x) = \begin{pmatrix} 4x_1 + 6x_2 + 3x_3 - 1 \\ -x_1 + 7x_2 - 5x_3 + 2 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathcal{K}^2 \times \mathcal{K}^3.$$

جواب تقریبی این مسأله  $x^* = (0.2335, -0.0710, 0.2207)^T$  است. شکل (۴.۳) رفتار گذرا از مدل شبکه عصبی (۲۰.۳) و (۲۱.۳) را برای مثال (۴.۵.۳) نشان می‌دهد.





شکل ۴.۳: رفتار گذرا شبکه عصبی در مثال (۴.۵.۳)



شکل ۵.۳: رفتار گذرا شبکه عصبی با پنج نقطه شروع در مثال (۴.۵.۳)

# فصل ۴

## یک شبکه عصبی جدید برای حل مسائل نامساوی‌های وردشی مخروطی و مکمل

### ۱.۴ مقدمه

در این فصل از پایان‌نامه مدل شبکه عصبی جدیدی را برای حل مسائل نامساوی وردشی مخروطی ارائه می‌دهیم و ثابت می‌کنیم که نقطه‌ی تعادل شبکه‌ی عصبی پیشنهاد شده همان جواب بهینه‌ی مسأله اصلی است و برعکس. همچنین نشان داده می‌شود که شبکه‌ی عصبی پیشنهادی پایدار به مفهوم لیاپانوف و همگرایی سراسری به جواب بهینه مسأله اصلی است. کارآمدی مدل پیشنهادی با ارائه چندین مثال نشان داده می‌شود.

در این فصل در بخش دوم تعریفی از مسائل نامساوی وردشی و مکمل داریم، در بخش سوم یک شبکه عصبی برای حل مسائل نامساوی وردشی مخروطی معرفی شده است. در بخش چهارم پایداری و همگرایی این شبکه عصبی را مورد بررسی قرار داده شده است. در بخش پنجم مثال‌هایی برای نشان دادن کارایی و عملکرد شبکه عصبی ارائه شده است.

### ۲.۴ بیان مسأله

مسأله نامساوی وردشی زیر را در نظر بگیرید:

$$\langle F(x), y - x \rangle \geq 0, \quad \forall y \in C, \quad (1.4)$$

که در آن  $C$  مجموعه متناهی است و به صورت

$$C = \{x \in \mathbb{R}^n : h(x) = 0, \quad -g(x) \in \mathcal{K}\}, \quad (2.4)$$

تعریف می‌شود.  $\langle \cdot, \cdot \rangle$  نشان دهنده ضرب داخلی است،  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^l$  و  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  توابع به‌طور پیوسته مشتق‌پذیر هستند. فرض کنید  $\nabla F$  یک ماتریس نیمه معین مثبت

است.  $\mathcal{K}$  ضرب دکارتی مخروط‌های مرتبه دوم است و به صورت زیر بیان می‌شود:

$$\mathcal{K} = \mathcal{K}^{m_1} \times \mathcal{K}^{m_2} \times \dots \times \mathcal{K}^{m_p}, \quad (3.4)$$

که در آن

$$m_1 + m_2 + \dots + m_p = m, \quad m_1, m_2, \dots, m_p \geq 1, \quad l \geq 0,$$

و

$$\mathcal{K}^{m_i} := \{(x_{i1}, x_{i2}, \dots, x_{im_i})^T \in \mathbb{R}^{m_i} \mid \|(x_{i2}, \dots, x_{im_i})\| \leq x_{i1}\},$$

که در آن  $\|\cdot\|$  نشان دهنده نرم اقلیدسی و  $\mathcal{K}^1$  مجموعه نامنفی حقیقی  $\mathbb{R}_+$  است. حالت خاصی از (۳.۴) حالتی است که  $\mathcal{K} = \mathbb{R}_+^n$  (یعنی مقادیر مثبت در  $\mathbb{R}^n$ )، که مربوط به  $p = n$  و  $m_1 = \dots = m_p = 1$  است. همانطور که در فصل سوم گفته شد، زمانی که  $h$  آفین است، مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (۱.۴)-(۳.۴) معادل مسأله برنامه‌ریزی محدب مخروط مرتبه دوم زیر است:

$$\min f(x) \quad (4.4)$$

$$s.t. \quad Ax = b, \quad (5.4)$$

$$\sqrt{g_{i2}^2 + \dots + g_{im}^2} + \epsilon \leq g_{i1} \quad i = 1, \dots, p, \quad (6.4)$$

که در آن  $A \in \mathbb{R}^{l \times n}$  رتبه سطری کامل دارد،  $b \in \mathbb{R}^l$ ،  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ،  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  و  $f = \nabla F$ . بنابراین یک مسأله برنامه‌ریزی محدب را به صورت کلی زیر در نظر می‌گیریم:

$$\min f(x) \quad (7.4)$$

$$s.t. \quad h(x) = 0, \quad (8.4)$$

$$g(x) \leq 0 \quad (9.4)$$

که در آن  $h(x) = Ax - b$  و  $g(x) \leq 0$  قيود مخروطی مرتبه دوم مشتق‌پذیر (۹.۴) را نشان می‌دهند و  $g(x) \in \mathbb{R}^m$ .

مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم (۷.۴)-(۹.۴) می‌تواند به وسیله یافتن جواب شرایط کاروش-کان-تاگر زیر حل شود:

$$u^* \geq 0, g(x^*) \leq 0, u^{*T} g(x^*) = 0,$$

$$\nabla f(x^*) + \nabla g(x^*)^T u^* + \nabla h(x^*)^T v^* = 0, \quad (10.4)$$

$$h(x^*) = 0.$$

$x^*$  نقطه‌ی کاروش-کان-تاگر مسأله (۴.۴)-(۶.۴) است و زوج  $(u^{*T}, v^{*T})^T$  بردار ضرایب لاگرانژین مربوط  $x^*$  نامیده می‌شود.

### ۳.۴ مدل شبکه عصبی

در این بخش برای حل مسأله نامساوی وردشی مخروطی تعریف شده در (۱.۴)-(۳.۴) سیستم دینامیکی زیر را بیان می‌کنیم.

فرض کنید  $x(\cdot)$ ،  $u(\cdot)$  و  $v(\cdot)$  متغیرهای وابسته به زمان باشند. هدف، طراحی یک سیستم دینامیکی زمان پیوسته است آن چنان که منجر به یافتن نقطه کاروش-کان-تاگر مسأله برنامه‌ریزی غیرخطی محدب (۷.۴)-(۹.۴) شود. مدل شبکه عصبی را به صورت زیر در نظر می‌گیریم:

$$\frac{dx}{dt} = -(\nabla f(x) + \nabla g(x)^T(u + g(x))^+ + \nabla h(x)^T v), \quad (11.4)$$

$$\frac{du}{dt} = (u + g(x))^+ - u, \quad (12.4)$$

$$\frac{dv}{dt} = h(x), \quad (13.4)$$

با نقطه اولیه  $y_0 = (x_0^T, u_0^T, v_0^T)^T$  که در آن

$$(u + g(x))^+ = ([u_1 + g_1(x)]^+, \dots, [u_m + g_m(x)]^+)^T,$$

$$[u_k + g_k(x)]^+ = \max\{u_k + g_k(x), 0\}, \quad k = 1, 2, \dots, m.$$

حال برای سادگی فرض کنید  $y = (x^T, u^T, v^T)^T \in \mathbb{R}^{n+m+l}$  را مجموعه نقاط بهینه (۱.۴)-(۳.۴) در نظر بگیریم تعریف می‌کنیم:

$$\Psi(y) = \begin{bmatrix} -(\nabla f(x) + \nabla g(x)^T(u + g(x))^+ + \nabla h(x)^T v) \\ (u + g(x))^+ - u \\ h(x) \end{bmatrix}. \quad (14.4)$$

در نتیجه شبکه عصبی (۱۱.۴)-(۱۳.۴) را می‌توان به صورت زیر نوشت:

$$\frac{dy}{dt} = \eta \Psi(y), \quad (15.4)$$

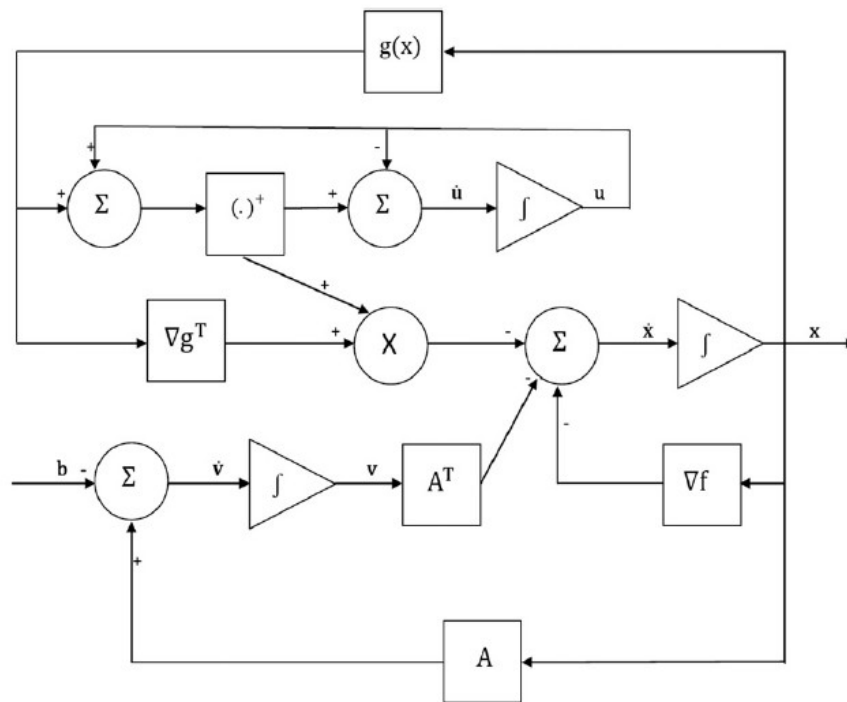
$$y(t_0) = y_0, \quad (16.4)$$

که در آن  $\eta$  پارامتر اسکالر و نشان‌دهنده نرخ همگرایی شبکه عصبی (۱۵.۴) و (۱۶.۴) است. برای سادگی بحث  $\eta = 1$  در نظر می‌گیریم. شکل (۱.۴) بیانگر این است که مدل فوق روی سخت افزار چگونه اجرا می‌شود.

### ۴.۴ تحلیل پایداری و همگرایی

در این بخش به اثبات پایداری و همگرایی شبکه عصبی پیشنهاد شده (۱۵.۴) و (۱۶.۴) می‌پردازیم.

قضیه ۱.۴.۴. فرض کنید  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  نقطه تعادل شبکه عصبی (۱۵.۴) و (۱۶.۴) باشد. آنگاه  $x^*$  نقطه کاروش-کان-تاگر از مسأله (۴.۴)-(۶.۴) است. برعکس اگر  $x^* \in \mathbb{R}^n$  جواب



شکل ۱۰۴: دیاگرام بلوکی برای شبکه عصبی (۱۵.۴) و (۱۶.۴)

بهینه مسأله (۴.۴)-(۶.۴) باشد، آنگاه بردارهای  $u^* \in \mathbb{R}^m$  و  $v^* \in \mathbb{R}^l$  وجود دارند به طوری که  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  نقطه تعادل شبکه عصبی پیشنهاد شده (۱۵.۴) و (۱۶.۴) است.

برهان. فرض کنید  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  نقطه تعادل شبکه عصبی (۱۵.۴) و (۱۶.۴) باشد، آنگاه  $\frac{dx^*}{dt} = 0$ ،  $\frac{du^*}{dt} = 0$  و  $\frac{dv^*}{dt} = 0$  در نتیجه

$$\nabla f(x^*) + \nabla g(x^*)^T (u^* + g(x^*))^+ + \nabla h(x^*)^T v = 0, \quad (17.4)$$

$$(u^* + g(x^*))^+ = u^*, \quad (18.4)$$

$$h(x^*) = 0. \quad (19.4)$$

همچنین، داریم  $(u^* + g(x^*))^+ = u^*$  اگر و تنها اگر

$$u^* \geq 0, \quad g(x^*) \leq 0, \quad u^{*T} g(x^*) = 0. \quad (20.4)$$

بنابراین با جایگذاری (۱۸.۴) در (۱۷.۴) داریم:

$$\nabla f(x^*) + \nabla g(x^*)^T u^* + \nabla h(x^*) v = 0 \quad (21.4)$$

از (۱۹.۴) و (۲۰.۴) نتیجه می‌شود که  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  در شرایط کاروش-کان-تاکر (K.K.T) صدق می‌کند و این اثبات را کامل می‌کند. □

قضیه ۲.۴.۴. نقطه تعادل مدل شبکه عصبی پیشنهاد شده (۱۵.۴) و (۱۶.۴) یکتاست.

برهان. چون مسأله (۴.۴)-(۶.۴) دارای جواب بهینه یکتای  $x^*$  است، شرایط لازم و کافی کاروش-کان-تاکر دارای جواب یکتای  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  است، بنابراین از قضیه (۱.۴.۴) می‌بینیم که نقطه تعادل مدل پیشنهاد شده (۱۵.۴) و (۱۶.۴) در سیستم (۱۰.۴) صدق می‌کند. بنابراین نقطه تعادل شبکه عصبی (۱۵.۴) و (۱۶.۴) یکتاست. □

قضیه ۳.۴.۴. برای هر نقطه آغازین  $y(t_0) = (x(t_0)^T, u(t_0)^T, v(t_0)^T)^T$  سیستم (۱۵.۴) و (۱۶.۴) دارای جواب یکتای پیوسته  $y(t) = (x(t)^T, u(t)^T, v(t)^T)^T$  است.

برهان. چون  $\nabla f(x)$ ،  $\nabla g_k(x)$ ،  $(k = 1, \dots, m)$  و  $\nabla h_l(x)$ ،  $(l = 1, \dots, p)$  توابع به طور پیوسته مشتق‌پذیر روی مجموعه باز  $D \subseteq \mathbb{R}^{n+m+l}$  هستند، بنابراین  $(u + \nabla f(x) + \nabla g(x)^T(u + g(x)))^+$  و  $g(x)^+ - u$  لیپ‌شیتز محلی پیوسته روی مجموعه باز  $D \subseteq \mathbb{R}^{n+m+l}$  هستند. با توجه به قضیه (۲۰.۳.۱)، شبکه عصبی (۱۵.۴) و (۱۶.۴) دارای جواب یکتای پیوسته  $y(t)$  برای  $t \in [t_0, \tau)$  وقتی  $\tau \rightarrow \infty$  است. □

لم ۴.۴.۴. ماتریس ژاکوبی  $\nabla \Psi(y)$  از نگاشت  $\Psi$  تعریف شده در (۱۴.۴) نیمه معین منفی است.

برهان. بدون کاستن از کلیت مسأله، فرض می‌کنیم  $0 < p < m$  وجود دارد به طوری که

$$(u + g)^+ = \underbrace{((u_1 + g_1(x)), (u_2 + g_2(x)), \dots, (u_p + g_p(x)), \circ, \circ, \dots, \circ)^T}_{>_0} \geq \circ.$$

با مراجعه به [۹] می‌توان نشان داد که

$$\nabla \Psi(y) = \begin{bmatrix} -(\nabla^T f(x) + \sum_{k=1}^p ((u_k + g_k) \nabla^T g_k(x)) + \nabla g^p(x)^T \nabla g^p(x)) & -\nabla g^p(x)^T & -\nabla h(x)^T \\ \nabla g^p(x) & S_{m \times m} & O_{m \times l} \\ \nabla h(x) & O_{l \times m} & O_{l \times l} \end{bmatrix},$$

که در آن

$$\nabla g^p(x) = \begin{bmatrix} U_{p \times n} \\ O_{(m-p) \times n} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_{p-1}} & \frac{\partial g_1}{\partial x_p} & \frac{\partial g_1}{\partial x_{p+1}} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \dots & \frac{\partial g_2}{\partial x_{p-1}} & \frac{\partial g_2}{\partial x_p} & \frac{\partial g_2}{\partial x_{p+1}} & \dots & \frac{\partial g_2}{\partial x_n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial g_p}{\partial x_1} & \dots & \frac{\partial g_p}{\partial x_{p-1}} & \frac{\partial g_p}{\partial x_p} & \frac{\partial g_p}{\partial x_{p+1}} & \dots & \frac{\partial g_p}{\partial x_n} \\ \circ & \dots & \circ & \circ & \circ & \dots & \circ \\ \circ & \dots & \circ & \circ & \circ & \dots & \circ \\ \circ & \dots & \circ & \circ & \circ & \dots & \circ \end{bmatrix},$$

و  $\nabla^T g_k(x)$  ماتریس هسین از  $g_k$ ،  $k = 1, \dots, p$  است، و

$$S_{m \times m} = \begin{bmatrix} O_{p \times p} & O_{p \times (m-p)} \\ O_{(m-p) \times p} & -I_{(m-p) \times (m-p)} \end{bmatrix},$$

که در آن  $O$  ماتریس صفر است. ماتریس  $\nabla g^p(x)^T \nabla g^p(x)$ ، ماتریس نیمه معین مثبت است [۱۰]. چون توابع  $f, g_1, \dots, g_m$  محدب و دوبار مشتق‌پذیر فرض شده، بنابراین ماتریس‌های هسین  $\nabla^2 f(x)$  و  $\nabla^2 g_k(x)$  برای  $(k = 1, 2, \dots, p)$ ، ماتریس‌های نیمه معین مثبت هستند. همچنین واضح است که ماتریس  $S_{m \times m}$ ، نیمه معین منفی است. طبق توضیحات گفته شده، ما می‌توانیم نتیجه بگیریم ماتریس ژاکوبی  $\nabla \Psi(y)$ ، نیمه معین منفی است.

اگر  $p = m$  یعنی  $(u_1 + g_1(x), u_2 + g_2(x), \dots, u_m + g_m(x))^T$ ، آنگاه داریم:

$$\nabla \Psi(y) = \begin{bmatrix} -(\nabla^2 f(x) + \sum_{k=1}^m ((u_k + g_k) \nabla^2 g_k(x)) + \nabla g(x)^T \nabla g(x)) & -\nabla g(x)^T & -\nabla h(x)^T \\ \nabla g(x) & O_{m \times m} & O_{m \times l} \\ \nabla h(x) & O_{l \times m} & O_{l \times l} \end{bmatrix}.$$

که در آن  $\nabla g(x)$  ماتریس ژاکوبی از توابع برداری  $g(x)$  و  $\nabla^2 g_k(x)$  ماتریس هسین از  $g_k(x)$  برای  $(k = 1, 2, \dots, m)$  هستند. مشابه حالت قبلی، می‌توان بسادگی نشان داد که  $\nabla \Psi(y)$  ماتریس نیمه معین منفی است.

سرانجام، اگر  $p = 0$  یعنی  $(0, 0, \dots, 0)^T$ ، آنگاه داریم:

$$\nabla \Psi(y) = \begin{bmatrix} -\nabla^2 f(x) & O_{n \times m} & -\nabla h(x)^T \\ O_{m \times n} & -I_{m \times m} & O_{m \times l} \\ \nabla h(x) & O_{l \times m} & O_{l \times l} \end{bmatrix}.$$

در این حالت هم، می‌توان نشان داد که  $\nabla \Psi(y)$  ماتریس نیمه معین منفی است، این اثبات را کامل می‌کند.  $\square$

قضیه ۵.۴.۴. مدل شبکه عصبی پیشنهاد شده در (۱۵.۴) و (۱۶.۴)، پایدار به مفهوم لیاپانوف و همگرایی سراسری به  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  است که در آن  $x^*$  جواب بهینه از (۴.۴)–(۶.۴) است.

تابع لیاپانوف  $E: \mathbb{R}^{m+n+l} \rightarrow \mathbb{R}$  زیر را در نظر بگیرید:

$$E(y) = \|\Psi(y)\|^2 + \frac{1}{\gamma} \|y - y^*\|^2. \quad (22.4)$$

مشتق  $\Psi(y)$  را به صورت زیر در نظر می‌گیریم:

$$\frac{d\Psi}{dt} = \frac{\partial \Psi}{\partial y} \frac{dy}{dt} = \nabla \Psi(y) \Psi(y).$$

با محاسبه مشتق  $E(y(t))$  از شبکه عصبی (۱۵.۴) و (۱۶.۴) داریم:

$$\begin{aligned} \frac{dE(y(t))}{dt} &= \left(\frac{d\Psi}{dt}\right)^T \Psi + \Psi^T \left(\frac{d\Psi}{dt}\right) + (y - y^*)^T \frac{dy}{dt} = \\ &= \Psi^T (\nabla \Psi(y)^T + \nabla \Psi(y)) \Psi + (y - y^*)^T \Psi(y). \end{aligned}$$

با به کار بردن لم (۴.۴.۴) داریم:

$$\Psi^T(y)(\nabla\Psi(y)^T + \nabla\Psi(y))\Psi(y) \leq 0, \quad (23.4)$$

با استفاده از تعریف (۱۰.۳.۱) و قضیه (۱۵.۳.۱) داریم:

$$(y - y^*)^T(\Psi(y) - \Psi(y^*)) = (y - y^*)^T\Psi(y) \leq 0,$$

بنابراین

$$\frac{dE(y(t))}{dt} \leq 0. \quad (24.4)$$

این بدان معنی است که شبکه عصبی (۱۵.۴) و (۱۶.۴)، پایدار به مفهوم لیاپانوف است. در ادامه چون

$$E(y_0) \geq E(y) \geq \frac{1}{\rho}\|y - y^*\|^2, \quad (25.4)$$

پس  $y(t)$  کراندار است. فرض کنید  $\Lambda = \{y \in \mathbb{R}^{m+n+l} | E(y) \leq E(y_0)\}$ . با استفاده از قضیه لاسال داریم  $\{(x(t)^T, u(t)^T, v(t)^T)^T \rightarrow M, t \rightarrow \infty\}$  که  $M$  بزرگترین مجموعه تغییرناپذیر در  $K = \{(x(t)^T, u(t)^T, v(t)^T)^T | \frac{dE(y(t))}{dt} = 0\}$  است. از (۱۵.۴) و (۱۶.۴) و (۲۳.۴) طبق [۹] داریم:

$$\frac{dx}{dt} = 0, \quad \frac{du}{dt} = 0, \quad \frac{dv}{dt} = 0 \Leftrightarrow \frac{dE(y(t))}{dt} = 0.$$

بنابراین  $(\bar{x}^T, \bar{u}^T, \bar{v}^T)^T \in D^*$  با  $M \subseteq K \subseteq D^*$ .

بنابراین، شبکه عصبی پیشنهاد شده در (۱۵.۴) و (۱۶.۴) همگرای سراسری به نقطه تعادل  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  است که در آن جواب بهینه از (۴.۴)–(۶.۴) است. از قضیه (۵.۴.۴) می‌توان گزاره زیر را نتیجه گرفت:

گزاره ۶.۴.۴. اگر  $D^* = \{(x^{*T}, u^{*T}, v^{*T})^T\}$  باشد، آنگاه شبکه عصبی (۱۵.۴) و (۱۶.۴) برای حل مسأله (۴.۴)–(۶.۴) پایدار مجانبی سراسری به نقطه تعادل یکتای  $y^* = (x^{*T}, u^{*T}, v^{*T})^T$  است.

## ۵.۴ مثال‌های عددی

به منظور نشان دادن کارایی و تأثیر شبکه عصبی پیشنهاد شده (۱۵.۴) و (۱۶.۴)، در این بخش، به ذکر چندین مثال عددی می‌پردازیم. نتایج شبیه‌سازی در ode45، از نرم افزار Matlab انجام شده است.



مثال ۱.۵.۴ ([۴۴]). مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم زیر را در نظر بگیرید:

$$F(x) = \begin{pmatrix} 2x_1 + x_2 + 1 \\ x_1 + 6x_2 - x_3 - 2 \\ -x_2 + 3x_3 - \frac{1}{5}x_4 + 3 \\ -\frac{1}{5}x_3 + 2x_4 + \frac{1}{4}\sin x_4 \cos x_5 \sin x_6 + 6 \\ \frac{1}{4}\cos x_4 \sin x_5 \sin x_6 + 2x_5 - \frac{1}{4} \\ -\frac{1}{4}\cos x_4 \cos x_5 \cos x_6 + 2x_6 + \frac{1}{4}\cos x_6 \sin x_7 \cos x_8 + 1 \\ \frac{1}{4}\sin x_6 \cos x_7 \cos x_8 + 4x_7 - 2 \\ -\frac{1}{4}\sin x_6 \sin x_7 \sin x_8 + 2x_8 + \frac{1}{4} \end{pmatrix},$$

$$C = \{x \in \mathbb{R}^8 : -g(x) = x \in \mathcal{K}^3 \times \mathcal{K}^3 \times \mathcal{K}^2\}.$$

جواب تقریبی این مسأله

$$x^* = (0.3820, 0.1148, -0.3644, 0.0000, 0.0000, 0.0000, 0.5000, -0.2500)^T$$

است. شکل (۲.۴) رفتار گذرا از مدل شبکه عصبی (۱۶.۴) و (۱۵.۴) را برای مثال (۱.۵.۴) نشان می‌دهد.

مثال ۲.۵.۴ ([۴۴]). مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم زیر را در نظر بگیرید:

$$\langle \frac{1}{4}Dx, y - x \rangle \geq 0, \quad \forall y \in C,$$

که در آن

$$C = \{x \in \mathbb{R}^n : Ax - a = 0, Bx - b \leq 0\}.$$

$D$  ماتریس متقارن  $n \times n$  است،  $A$  و  $B$  به ترتیب ماتریس‌های  $l \times n$  و  $m \times n$  هستند،  $d$  یک بردار  $n \times 1$  است،  $a$  و  $b$  بردارهای  $l \times 1$  و  $m \times 1$  با  $l + m \leq n$  هستند. می‌توانیم داده‌های  $a, B, A, b$  و  $D$  را تصادفی انتخاب کنیم، اما با مراجعه به [۴۵]، مجموعه داده‌ها را به صورت زیر تعریف می‌کنیم:

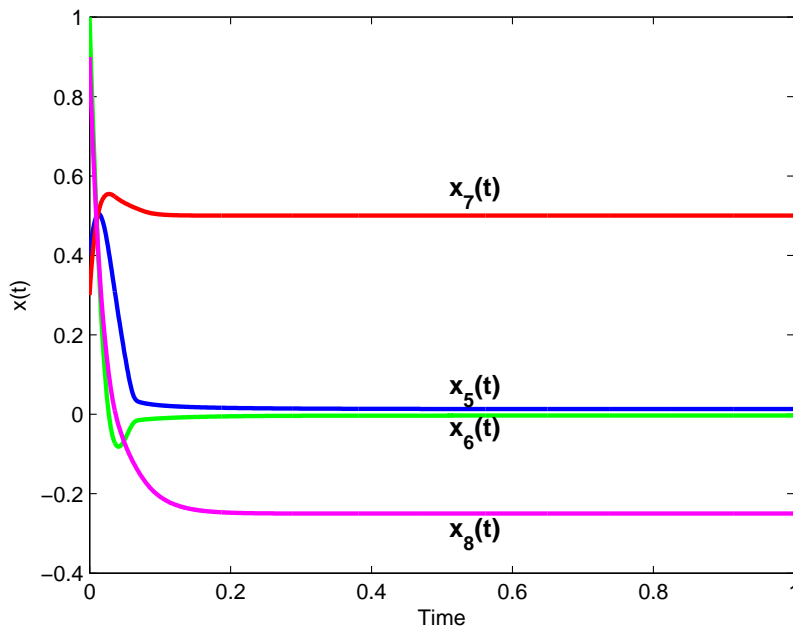
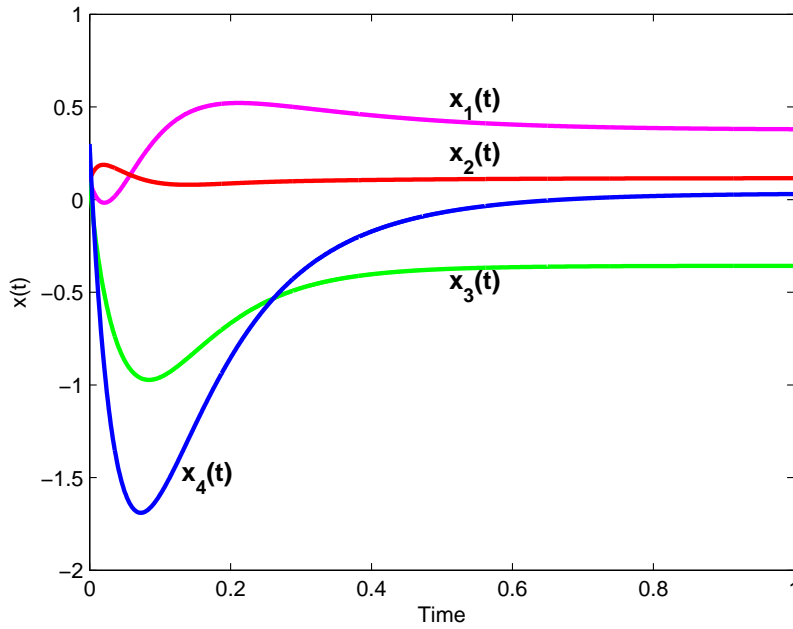
$$D = [D_{ij}]_{n \times n},$$

که در آن

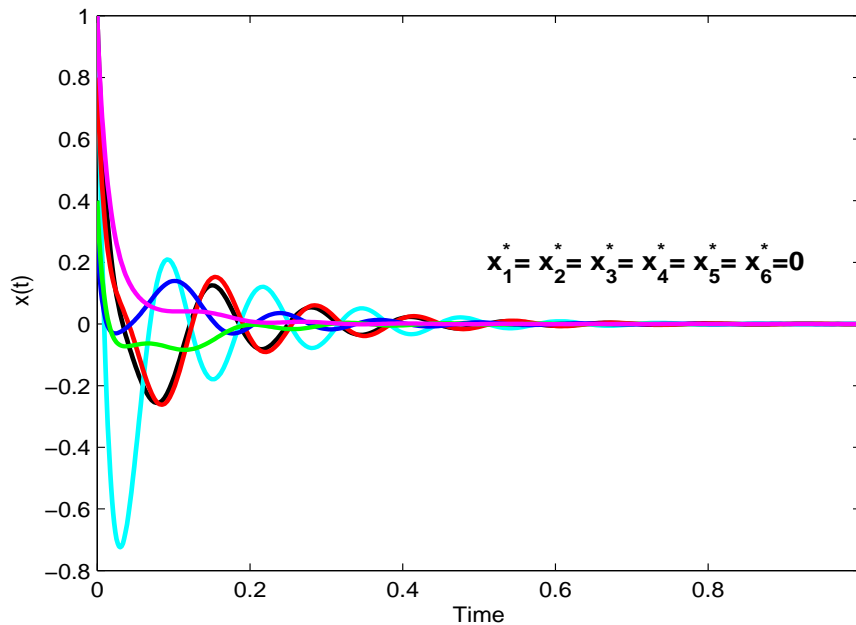
$$D_{ij} = \begin{cases} 2, & i = j \\ 1, & |i - j| = 1, \\ 0, & \text{otherwise,} \end{cases}$$

$b = (e_{m_1}, e_{m_2}, \dots, e_{m_p})^T$ ,  $a = O_{l \times 1}$ ,  $B = [O_{m \times (n-m)} \quad I_{m \times m}]_{m \times n}$ ,  $A = [I_{l \times 1} \quad O_{l \times (n-1)}]_{l \times n}$  که  $e_{m_i} = (1, 0, \dots, 0)^T \in \mathbb{R}^{m_i}$  و  $l + m \leq n$  رتبه  $A$  و  $B$  کامل است و رتبه آن  $([A^T \ B^T]) = l + m$  است.

پارامترهای  $l = 3$  و  $m = 3$  و  $n = 6$  است. این مسأله جواب تقریبی  $x^* = (0, 0, 0, 0, 0, 0)^T$  را دارد. شکل (۳.۴) رفتار گذرا از مدل شبکه عصبی (۱۶.۴) و (۱۵.۴) را برای مثال (۲.۵.۴) نشان می‌دهد.



شکل ۲.۴: رفتار گذرا شبکه عصبی در مثال (۱.۵.۴)



شکل ۳.۴: رفتار گذرا شبکه عصبی در مثال (۲.۵.۴)

مثال ۳.۵.۴ ([۴۴]). مسأله نامساوی وردشی با محدودیت مخروطی مرتبه دوم زیر را در نظر بگیرید:

$$F(x) = \begin{pmatrix} x_3 \exp(x_1 x_3) + \wp(x_1 + x_2) \\ \wp(x_1 + x_2) + \frac{2(x_2 - x_3)}{\sqrt{1 + (2x_2 - x_3)^2}} \\ x_1 \exp(x_1 x_3) - \frac{2x_2 - x_3}{\sqrt{1 + (2x_2 - x_3)^2}} \\ x_4 \\ x_5 \end{pmatrix},$$

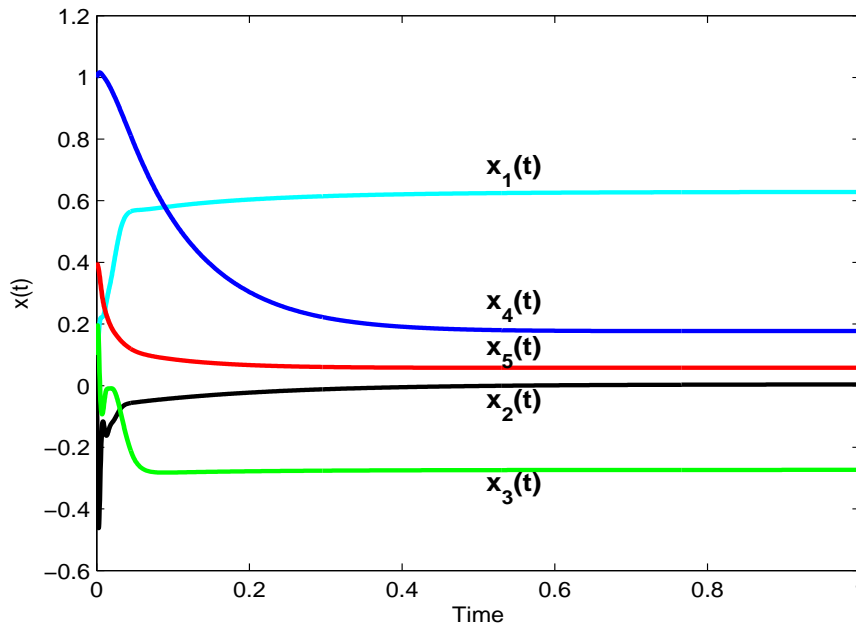
$$C = \{x \in \mathbb{R}^5 : h(x) = 0, -g(x) \in \mathcal{K}^3 \times \mathcal{K}^2\},$$

که در آن

$$h(x) = -62x_1^3 + 58x_2 + 167x_3^2 - 29x_3 - x_4 - 3x_5 + 11,$$

$$g(x) = \begin{pmatrix} -3x_1^3 - 2x_2 + x_3 - 5x_3^2 \\ 5x_1^3 - 4x_2 + 2x_3 - 10x_3^2 \\ -x_3 \\ -x_4 \\ -3x_5 \end{pmatrix}$$

جواب تقریبی این مسأله  $x^* = (0.6287, 0.0042, -0.2733, 0.1773, 0.587)^T$  است. شکل (۴.۴) رفتار گذرا از مدل شبکه عصبی (۱۶.۴) و (۱۵.۴) را برای مثال (۳.۵.۴) نشان می‌دهد.



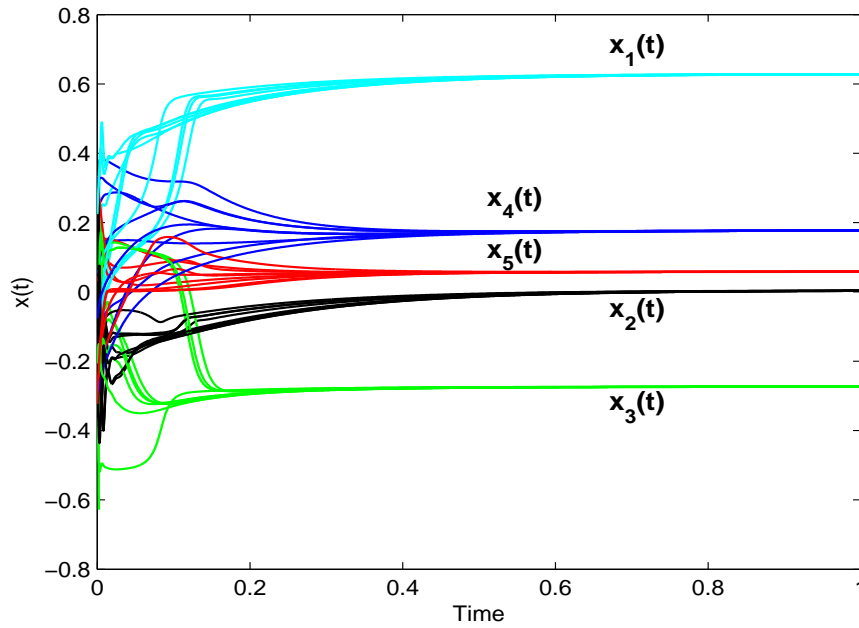
شکل ۴.۴: رفتار گذرا شبکه عصبی در مثال (۳.۵.۴)

مثال ۴.۵.۴ ([۴۴]). مسأله محدب غیر خطی زیر را در نظر بگیرید:

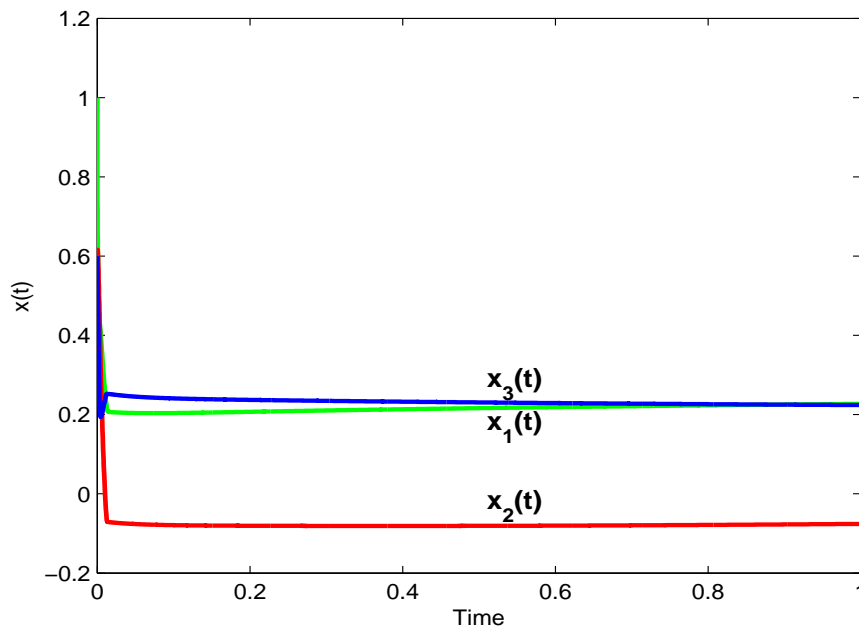
$$\min \quad \exp(x_1 - x_3) + 3(2x_1 - x_2)^4 + \sqrt{1 + (3x_2 + 5x_3)^2}$$

$$\text{s.t} \quad -g(x) = \begin{pmatrix} 4x_1 + 6x_2 + 3x_3 - 1 \\ -x_1 + 7x_2 - 5x_3 + 2 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathcal{K}^2 \times \mathcal{K}^3.$$

جواب تقریبی این مسأله  $x^* = (0.2336, -0.0700, 0.2206)^T$  است. شکل (۶.۴) رفتار گذرا از مدل شبکه عصبی (۱۶.۴) و (۱۵.۴) را برای مثال (۴.۵.۴) نشان می‌دهد.



شکل ۵.۴: رفتار گذرا شبکه عصبی با ده نقطه شروع در مثال (۳.۵.۴)



شکل ۶.۴: رفتار گذرا شبکه عصبی در مثال (۴.۵.۴)

# نتایج و پیشنهادهایی برای کارهای آتی

در این پایان‌نامه ابتدا مروری بر شبکه‌های عصبی داشته و با استفاده از دو مدل شبکه عصبی کارا به حل مسأله‌ی نامساوی وردشی مخروطی پرداختیم. همچنین پایداری و همگرایی سراسری به جواب بهینه را مورد بررسی قرار دادیم. از جمله مزیت‌های مدل‌های ارائه شده می‌توان به موارد زیر اشاره کرد:

- حل مسائل بهینه‌سازی غیرخطی محدب
- حل مسأله تخصیص
- حل مسأله ربات چند انگشتی
- حل مسأله پردازش تصویر
- حل مسائل بهینه‌سازی سبد سهام
- استفاده از مدل پیشنهادی برای حل مساله طبقه بندی.
- تعمیم مدل پیشنهادی برای داده‌های رگرسیون با ابعاد بالاتر.
- حل مساله رگرسیون ريج<sup>۱</sup> با مدل‌های ارائه شده.
- مدل‌های ذکر شده همگرایی تضمین شده دارند.
- این مدل‌ها نیاز به هیچگونه پارامتر جریمه ندارد.
- با وجود سادگی مدل‌ها، در رده بسیاری از مسائل بهینه‌سازی محدب قابل استفاده‌اند.
- این مدل‌ها هیچگونه وابستگی به نقطه شروع نداشته و این نقطه می‌تواند خارج از ناحیه شدنی و نقطه‌ای تصادفی باشد.
- این مدل‌ها قادر به حل مسائل وابسته به زمان و مسائل با ابعاد بزرگ هستند.
- مدل فصل سوم مسائل برنامه‌ریزی خطی را حل نمی‌کند اما مدل فصل چهارم تمام مسائل بهینه‌سازی محدب را حل می‌کند.

---

<sup>۱</sup>Ridge regression

# پیوست آ

## کد مثال‌های فصل ۳

### آ.۱ مثال (۱.۵.۳)

```
[t,x]=ode45('s',[0 1],[1 0.2 0.5 0.7 0.9 1 0.6 0.2 1 0.9 0.8]);
plot(t,x(:,1),'m','linewidth',2);
hold on;
text(0.5,0.5,'\bf x_1(t)','fontsize',12)
plot(t,x(:,2),'r','linewidth',2);
hold on;
text(0.5,0.25,'\bf x_2(t)','fontsize',12)
plot(t,x(:,3),'g','linewidth',2);
hold on;
text(0.5,-0.44,'\bf x_3(t)','fontsize',12)
plot(t,x(:,4),'b','linewidth',2);
hold on;
text(0.5,-0.1,'\bf x_4(t)','fontsize',12)
plot(t,x(:,5),'m','LineWidth',2);
hold on;
text(0.5,0.07,'\bf x_5(t)','FontSize',12)
plot(t,x(:,6),'r','LineWidth',2);
hold on;
text(0.5,-0.07,'\bf x_6(t)','FontSize',12)
plot(t,x(:,7),'g','LineWidth',2);
hold on;
```

```

text(0.5,0.58,'\bf x_7(t)','FontSize',12)
plot(t,x(:,8),'b','LineWidth',2);
hold on;
text(0.5,-0.30,'\bf x_8(t)','FontSize',12)
ylabel('x(t)');
xlabel('Time');

```

## ٢.٥.٣. مثال (٢.٥.٣)

```

[t,x]=ode45('s4',[0 1],[1 0.5 0.6 1 0.5 0.6 0.8 1 0.5 0.3]);
plot(t,x(:,1),'k','LineWidth',2);
hold on;
text(0.3,0.19,' x_1(t)= x_2(t)= x_3(t)= x_4(t)= x_5(t)= x_6(t)=0')
plot(t,x(:,2),'c','LineWidth',2);
hold on;
plot(t,x(:,3),'r','LineWidth',2);
hold on;
plot(t,x(:,4),'b','LineWidth',2);
hold on;
plot(t,x(:,5),'g','LineWidth',2);
hold on;
plot(t,x(:,6),'m','LineWidth',2);
hold on;
ylabel(' x(t)');
xlabel('Time');

```

## ٣.٥.٣. مثال (٣.٥.٣)

```

[t,x]=ode45('s5',[0 1],[1 0.1 0.8 0.3 0.4 0.5 0.9 0.8]);
plot(t,x(:,1),'k','LineWidth',2);
hold on;
plot(t,x(:,2),'m','LineWidth',2);
hold on;
plot(t,x(:,3),'g','LineWidth',2);

```



```

hold on;
plot(t,x(:,4),'b','LineWidth',2);
hold on;
plot(t,x(:,5),'r','LineWidth',2);
hold on;
text(0.5,0.73,'\bf x_1(t)','FontSize',12)
text(0.5,-0.09,'\bf x_2(t)','FontSize',12)
text(0.5,-0.35,'\bf x_3(t)','FontSize',12)
text(0.5,0.28,'\bf x_4(t)','FontSize',12)
text(0.3,0.13,'\bf x_5(t)','FontSize',12)
ylabel('x(t)');
xlabel('Time');

```

### آ.۴ مثال (۴.۵.۳)

```

[t,x]=ode45('s6',[0 1],[1 0.5 0.6 0.8 1]);
plot(t,x(:,1),'k','LineWidth',2);
hold on;
text(0.5,0.28,'\bf x_1(t)','FontSize',12)
plot(t,x(:,2),'r','LineWidth',2);
hold on;
text(0.5,-0.12,'\bf x_2(t)','FontSize',12)
plot(t,x(:,3),'b','LineWidth',2);
hold on;
text(0.5,0.17,'\bf x_3(t)','FontSize',12)
ylabel('x(t)');
xlabel('Time');

```

# پیوست ب

## کد مثال‌های فصل ۴

### ب.۱ مثال (۱.۵.۴)

```
[t,x]=ode45('s4',[0 3],[0.1 0.1 0.1 0.3 0.4 0 0 0 0 0]);
plot(t,x(:,1),'k');
hold on;
text(1.5,0.45,'x(1)')
plot(t,x(:,2),'y');
hold on;
text(1.5,0.20,'x(2)')
plot(t,x(:,3),'g');
hold on;
text(0.7,-0.46,'x(3)')
plot(t,x(:,4),'b');
hold on;
text(0.2,-1.5,'x(4)')
plot(t,x(:,5),'r');
hold on;
text(1.5,-0.07,'x(5)')
plot(t,x(:,6),'g');
hold on;
text(0.1,-0.1,'x(6)')
plot(t,x(:,7),'r');
hold on;
text(1.5,0.61,'x(7)')
```

```

plot(t,x(:,8),'m');
hold on;
text(1.5,-0.30,'x(8)')
ylabel('x(t)');
xlabel('Time');

```

## ب.۲ مثال (۲.۵.۴)

```

[t,x]=ode45('s500',[0 1],[1 1 0.8 0.3 0.4 1 0 0 1 0]);
plot(t,x(:,1),'k','LineWidth',2);
hold on;
text(0.29,0.2,' x_1(t)= x_2(t)= x_3(t)= x_4(t)= x_5(t)= x_6(t)=0')
plot(t,x(:,2),'c','LineWidth',2);
hold on;
plot(t,x(:,3),'r','LineWidth',2);
hold on;
plot(t,x(:,4),'b','LineWidth',2);
hold on;
plot(t,x(:,5),'g','LineWidth',2);
hold on;
plot(t,x(:,6),'m','LineWidth',2);
hold on;
ylabel(' x(t)');
xlabel('Time');

```

## ب.۳ مثال (۳.۵.۴)

```

[t,x]=ode45('s600',[0 1],[0.1 0.1 0.1 1 0.4 0.9 0.7 0.8]);
plot(t,x(:,1),'c','LineWidth',2);
hold on;
text(0.5,0.7,'\bf x_1(t)','FontSize',12)
plot(t,x(:,2),'k','LineWidth',2);
hold on;
text(0.5,-0.06,'\bf x_2(t)','FontSize',12)

```

```

plot(t,x(:,3),'g','LineWidth',2);
hold on;
text(0.5,-0.33,'\bf x_3(t)','FontSize',12)
plot(t,x(:,4),'b','LineWidth',2);
hold on;
text(0.5,0.26,'\bf x_4(t)','FontSize',12)
plot(t,x(:,5),'r','LineWidth',2);
hold on;
text(0.5,0.11,'\bf x_5(t)','FontSize',12)
ylabel('x(t)');
xlabel('Time');

```

### ب.٤. مثال (٤.٥.٤)

```

[t,x]=ode45('s800',[0 1],[1 0.5 0.6 0.8 1]);
plot(t,x(:,1),'g','LineWidth',2);
hold on;
plot(t,x(:,2),'r','LineWidth',2);
hold on;
plot(t,x(:,3),'b','LineWidth',2);
hold on;
text(0.5,-0.03,'\bf x_2(t)','FontSize',12)
text(0.5,0.17,'\bf x_1(t)','FontSize',12)
text(0.5,0.28,'\bf x_3(t)','FontSize',12)
ylabel('x(t)');
xlabel('Time');

```

## مراجع

- [۱] آذر. ضمیری. شبکه‌های عصبی مصنوعی و بهینه‌سازی. پایان‌نامه کارشناسی‌ارشد، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، آبان ۱۳۸۷.
- [۲] عباس. قماش لنگرودی. حل مسائل بهینه‌سازی با استفاده از شبکه‌های عصبی تصویر. پایان‌نامه کارشناسی‌ارشد، دانشکده علوم پایه، دانشگاه تربیت معلم سبزوار، آبان ۱۳۸۳.
- [۳] محمد. مقدس. حل رده‌ای از مسائل بهینه‌سازی غیرخطی با استفاده از شبکه‌های عصبی. پایان‌نامه کارشناسی‌ارشد، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، تابستان ۱۳۹۲.
- [۴] 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.
- [5] A. Bouzerdoum and T. R. Pattison. Neural network for quadratic optimization with bound constraints. *IEEE Trans. Neural Netw.*, 4(2): 293–304, (1993).
- [6] A. Malek and H. G. Oskoei. Primal–dual solution for the linear programming problems using neural networks, *Appl. Math. Comput.* 169, 451–471, (2005).
- [7] A. Malek, A. Yari. Primal–dual solution for the linear programming problem using neural networks, *Appl. Math. Comput.* 169, 198–211, (2005).
- [8] A. R. Nazemi. A dynamic system model for solving convex nonlinear optimization problems, *Commun Nonlinear Sci Numer Simulat* 17, 1696–1705, April (2012).
- [9] A. R. Nazemi. Solving general convex nonlinear optimization problems by an efficient neurodynamic model, *Engineering Applications of Artificial Intelligence* 26, 685–696, (2013).
- [10] A. Quarteroni, R. Sacco and F. Saleri. *Numerical mathematics. Texts in Applied Mathematics*, vol. 37, second ed, (2007). Springer-Verlag, Berlin.
- [11] A. Rodriguez-Vazquez. Dominguez-Castro, R., Rueda, A., Huertas, J.L., and Sanchez Sinencio, E. Nonlinear switched capacitor ‘neural’ networks for optimization problems. *IEEE Trans. Circuits Syst.*, 37(3): 384–398, (1990).

- 
- [12] B.S. He. A class of projection and contraction methods for monotone variational inequalities, *Appl. Math. Optim. Ser. A* 35, 69–76, (1997).
- [13] B.S. He, L.Z. Liao. Improvements of some projection methods for monotone nonlinear variational inequalities, *J. Optim. Theory Appl.* 112 (1), 111–128, (2002) .
- [14] B.S. He. Solution and applications of a class of general linear variational inequalities, *Sci. China Ser. A* 39, 395–404, (1996).
- [15] B. Widrow. Generalization and information storage in networks of adaline ‘neurons’ in *Self Organizing Systems*, pp. 435–461. Spartan Books, (1959).
- [16] B. Widrow and M. E. Hoff. Adapting switching circuits. in 1960 ire wescon Convention Record, no. 4, pp. 96–104, New York, NY, (1960). ire.
- [17] B. Widrow and M. A. Lehr. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. *Proc. IEEE*, 78(9): 1415–1442, (1990).
- [18] B. Widrow and R. Winter. Neural nets for adaptive filtering and adaptive pattern recognition. *Computer*, 21(3): 25–39, (1988).
- [19] C.Y. Maa and M. A. Schanblatt. A two-phase optimization neural network. *IEEE Trans. Neural Netw.*, 3(6): 1003–1009, (1992).
- [20] D.G. Luenberger. *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, (1973).
- [21] D. E. Rumelhart, G. E. Hinton and R. J. Williams. Learning representations by backpropagating errors. *Nature*, 323: 533–536, (1986).
- [22] D. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. Springer, 3rd ed., (2008).
- [23] D.P. Bertsekas, J.N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, (1989).
- [24] D. O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. John Wiley & Sons, Inc., New York, (1949).
- [25] D. Kinderlehrer and G. Stampcchia. *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, (1980).
- [26] D. Tank and J. J. Hopfield. Simple ‘neural’ optimization networks: An a/d converter, signal decision circuit, and a linear programming circuit. *IEEE Trans. Circuits Syst.*, 33(5): 533–541, (1986).

- [27] F. Facchinei, J. Pang. *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer, New York (2003).
- [28] F. Giannessi and A. Maugeri. *Variational Analysis And Applications. Nonconvex Optimization and Its Applications*. Springer, (2005).
- [29] 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).
- [30] F. Rosenblatt. *Principles of neurodynamics: perceptrons and the theory of brain mechanisms*. Report (Cornell Aeronautical Laboratory). Spartan Books, (1962).
- [31] G. Stampacchia. Formes bilineaires coercives sur les ensembles convexes. *C. R. Acad. Sci. Paris* 258, 4413–4416 (1964).
- [32] G. Stampacchia. *Variational inequalities. Theory and applications of monotone operators*. In: *Proceedings of the NATO Advanced Study Institute*, pp. 101–192, Venice (1968).
- [33] H. K. Khalil. *Nonlinear Systems*. Prentice Hall PTR, 3rd ed., (2002).
- [34] J. Zabczyk. *Mathematical control theory: an introduction*. Systems & control. Birkhäuser, (1992).
- [35] J. S. Chen, H. T. Gao, S. H. Pan. An R-linearly convergent derivative-free algorithm for nonlinear complementarity problems based on the generalized Fischer-Burmeister merit function. *J. Comput. Appl. Math.* 232, 455–471 (2009).
- [36] J. S. Chen, S.H. Pan. A family of NCP functions and a descent method for the nonlinear complementarity problem. *Comput. Optim. Appl.* 40, 389–404 (2008).
- [37] J. S. Chen. The semismooth-related properties of a merit function and a decent method for the nonlinear complementarity problem. *J. Glob. Optim.* 36, 565–580 (2006).
- [38] J. J. Hopfield and D. W. Tank. “neural” computation of decisions in optimization problems. *Biological Cybernetics*, 52(3): 141–152, (1985).
- [39] J. Lions and G. Stampacchia. *Variational inequalities*. *Commun. Pure Appl. Math.* 20, 493–519, (1967).
- [40] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci.*, 79(8): 2554–2558, (1982).
- [41] J.M. Ortega and W.G. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, (1970).

- [42] 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).
- [43] J.S. Pang. A posteriori error bounds for the linearly-constrained variational inequality problem, *Math. Oper. Res.* 12, 474–484, (1987).
- [44] J. Sun, J.S. Chen and C.H. Ko. Neural networks for solving second-order con constrained variational inequality problem. *Appl. Optim. Comput.* 51, 623–648 (2010).
- [45] J. H. Sun and L.W. Zhang. A globally convergent method based on Fischer-Burmeister operators for solving second-order-cone constrained variational inequality problems. *Comput. Math. Appl.* 58, 1936–1946, (2009).
- [46] J. Wang, Q. Hu and D. Jiang. A Lagrangian neural network for kinematics control of redundant robot manipulators, *IEEE Trans. Neural Networks* 10 (5), 1123–1132, (1999).
- [47] L. Perko. *Differential Equations and Dynamical Systems*. Springer, 3rd ed., (2001).
- [48] M. Yashtini and A. Malek. Solving complementarity and variational inequalities problems using neural networks, *Applied Mathematics and Computation* 190, 216–230, (2007).
- [49] M. P. Kennedy and L. O. Chua. Neural networks for nonlinear programming. *IEEE Trans. Circuits Syst.*, 35(5): 554–562, (1988).
- [50] M. Avriel. *Nonlinear Programming: Analysis and Methods*, Prentice-Hall, Englewood Cliffs, NJ, (1976).
- [51] M.S. Bazaraa, H.D. Sherali and C. M. Shetty. *Nonlinear Programming*. John Wiley & Sons, Inc., 3rd ed. , (2005).
- [52] M. Fukushima. Equivalent differentiable optimization problems and descent method for asymmetric variational inequality problems, *Math. Program.* 53, 99–110, (1992).
- [53] M. L. Minsky and S. Papert. *Perceptrons: An Introduction to Computational Geometry*. MitPress, (1972).
- [54] 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.
- [55] M. Moghaddas and M. Eshaghnejad. A recurrent neural network for solving non-convex non linear optimization problem. in *The 44th Annual Iranian Mathematics Conference*, Mashhad, Iran, August (2013). Ferdowsi University of Mashhad.



- [56] O. Mancino and G. Stampacchia. Convex programming and variational inequalities. *J. Optim. Theory Appl.* 9, 3–23, (1972).
- [57] P. Werbos. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. thesis, Harvard University, Cambridge, MA, (1974).
- [58] P.T. Harker and J.S. Pang. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms, and applications, *Math. Program.* 48B, 161–220, (1990).
- [59] P. Hartman, G. Stampacchia. On some nonlinear elliptic differential functional equations. *ActaMath.* 115, 271–310 (1966).
- [60] Q. Tao, J. D. Cao, M. S. Xue and H. Qiao. A high performance neural network for solving nonlinear programming problems with hybrid constraints, *Phys. Lett. A* 288 (2), 88–94, (2001).
- [61] R.K. Miller and A.N. Michel. *Ordinary Differential Equations*, Academic Press, (1982).
- [62] R.T. Rockafellar. Linear-quadratic programming and optimal control, *SIAM J. Contr. Optim.* 25 (3), 781–814, (1987).
- [63] S. Effati, and M. Baymani. A new nonlinear neural network for solving convex nonlinear programming problems. *Appl. Math. and Comput.*, 168(2): 1370 – 1379, (2005).
- [64] 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).
- [65] S. Effati, A. Ghomashi and A. R. Nazemi. Application of projection neural network in solving convex programming problems. *Appl. Math. Comput.*, 188(2): 1103 – 1114, (2007).
- [66] S.J. Wright. An infeasible-interior-point algorithm for linear complementarity problems. *Math. Program.* 67, 29–51 (1994).
- [67] S. Zhang and A.G. Constantinides. Lagrange programming neural networks. *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, 39(7): 441–452, (1992).
- [68] W. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, 5(4): 115–133, (1943).
- [69] X. Chen, L. Qi and D. Sun. Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities. *Math. Comput.* 67(222), 519–540, (1998).

- [70] X.B. Gao, L.Zh. Liao. A neural network for monotone variational inequalities with linear constraints, *Phys. Lett. A* 307, 118–128, (2003).
- [71] X.B. Gao, L.Zh. Liao and L. Qi. A novel neural network for variational inequalities with linear and nonlinear constraints, *IEEE Trans. Neural Networks* 16 (6), 1305–1317, (2005).
- [72] X.B. Liang and J. Wang. A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bound constraints 11 (6), 1251–1262, (2000).
- [73] Y. Xia, H. Leung and J. Wang. A projection neural network and its application to constrained optimization problems. *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl*, 49(4), 447–458, (2002).
- [74] Y. H. Chen, S.C. Fang. Solving convex programming problems with equality constraints neural networks, *Comput. Math. Appl.* 36 (7), 41–68, (1998).
- [75] Y. C. Huang . A novel method to handle inequality constraints for convex programming neural network. *Neural Process Lett*; 16: 17–27, (2002).
- [76] Y.S. Xia. A new neural network for solving linear and quadratic programming problems, *IEEE Trans. Neural Networks* 7 (6), 1544–1547, (1996).
- [77] Y. Xia, G. Feng and J. Wang. A novel recurrent neural network for solving nonlinear optimization problems with inequality constraints. *IEEE Trans. Neural Netw.*, 19(8): 1340–1353, (2008).
- [78] Y. Xia and J. Wang. A recurrent neural network for nonlinear convex optimization subject to nonlinear inequality constraints. *IEEE Trans. Circuits Syst. I, Reg. Papers*, 51(7): 1385–1394, (2004).
- [79] Y. Xia and J. Wang. A recurrent neural network for solving nonlinear convex programs subject to linear constraints. *IEEE Trans. Neural Netw.*, 16(2): 379–386, (2005).
- [80] Y.S. Xia, H. Leung and J. Wang. A projection neural network and its application to constrained optimization problems, *IEEE Trans. Circ. Syst. I, Fundam. Theory Appl.* 49 (4), 447–458, (2002).
- [81] Y.S. Xia, J. Wang. A general methodology for designing globally convergent optimization neural networks, *IEEE Trans. Neural Networks* 9 (6), 1331–1343, (1998).
- [82] Y. S. Xia and J. Wang. A general projection neural network for solving monotone variational inequality and related optimization problems, *IEEE Trans. Neural Networks* 15 (2), 318–328, (2004) .

# واژه‌نامه فارسی به انگلیسی

LaSalle Principle of Invariance	اصل تغییرناپذیری لسال
Axon	آکسون
Convex Optimization	بهینه‌سازی محدب
Nonconvex Optimization	بهینه‌سازی نامحدب
Postsynaptic Terminal	پایانه بعد از سیناپسی
Presynaptic Terminal	پایانه قبل از سیناپسی
Globally Asymptotically Stable	پایدار مجانبی سراسری
Stability in the Sense of Lyapunov	پایداری به مفهوم لیاپانوف
Globally Exponential Stability	پایداری نمایی سراسری
Back Propagation	پس انتشار
Feed Forward	پیش‌خورد
Lipschitz Continuous	پیوسته لیپ‌شیتز
Locally Lipschitz Continuous	پیوسته لیپ‌شیتز محلی
Activation Function	تابع فعال‌سازی
Firing	تحریک شدن
Feasible Solution	جواب شدنی
Dendrite	دندریت
Local Duality	دوگانی موضعی
Switched-Capacitor	سوئیچ-خازن
Synapse	سیناپس
Neural Networks	شبکه‌های عصبی
Feedback Neural Network	شبکه‌های عصبی بازگشتی
Artificial Neural Networks	شبکه‌های عصبی مصنوعی
Second-Order Sufficient Conditions	شرایط کافی مرتبه دوم
Synaptic Cleft	شکاف سیناپسی

Lagrangian Multiplier	ضربگر لاگرانژ
Projection Operator	عملگر تصویر
Brouwer's Fixed Point Theorem	قضیه نقطه ثابت بروئر
Hessian Matrix	ماتریس هسین
Dual Variable	متغیر دوگان
Invariant Set	مجموعه تغییرناپذیر
Level Set	مجموعه سطح
Lagrangian Dual Problem	مسأله دوگان لاگرانژ
Linear Complementarity Problem	مسأله مکمل خطی
Nonlinear Complementarity Problem	مسأله مکمل غیرخطی
Positive Definite	معین مثبت
Strongly Positive Definite	معین مثبت قوی
Variational Inequalities	نامساوی‌های وردشی
Neuron	نرون
Equilibrium Point	نقطه تعادل
Isolated Equilibrium Point	نقطه تعادل تنها
Regular Point	نقطه منظم
Strict Local Minimum Point	نقطه مینیمم محلی اکید
Strictly Monotone Map	نگاشت اکیداً یکنوا
Strongly Monotone Map	نگاشت قویاً یکنوا
Monotone Map	نگاشت یکنوا
Positive Semidefinite	نیمه معین مثبت
Globally Convergent	همگرای سراسری

# واژه‌نامه انگلیسی به فارسی

Activation Function	تابع فعال‌سازی
Artificial Neural Networks	شبکه‌های عصبی مصنوعی
Axon	آکسون
Back Propagation	پس انتشار
Brouwer's Fixed Point Theorem	قضیه نقطه ثابت بروئر
Convex Optimization	بهینه‌سازی محدب
Dendrite	دندریت
Dual Variable	متغیر دوگان
Equilibrium Point	نقطه تعادل
Feasible Solution	جواب شدنی
Feedback Neural Network	شبکه‌های عصبی بازگشتی
Feed Forward	پیش‌خورد
Firing	تحریک شدن
Globally Asymptotically Stable	پایدار مجانبی سراسری
Globally Convergent	همگرای سراسری
Globally Exponential Stability	پایداری نمایی سراسری
Hessian Matrix	ماتریس هسین
Invariant Set	مجموعه تغییرناپذیر
Isolated Equilibrium Point	نقطه تعادل تنها
Lagrangian Dual Problem	مسأله دوگان لاگرانژ
Lagrangian Multiplier	ضرب‌گر لاگرانژ
LaSalle Principle of Invariance	اصل تغییرناپذیری لسال
Level Set	مجموعه سطح
Linear Complementarity Problem	مسأله مکمل خطی
Lipschitz Continuous	پیوسته لیپ شیتز

Local Duality	دوگانی موضعی
Locally Lipschitz Continuous	پیوسته لیپ‌شیتز محلی
Monotone Map	نگاشت یکنوا
Neural Networks	شبکه‌های عصبی
Neuron	نرون
Nonconvex Optimization	بهبینه‌سازی نامحدب
Nonlinear Complementarity Problem	مسئله مکمل غیرخطی
Positive Definite	معین مثبت
Positive Semidefinite	نیمه معین مثبت
Postsynaptic Terminal	پایانه بعد از سیناپسی
Presynaptic Terminal	پایانه قبل از سیناپسی
Projection Operator	عملگر تصویر
Regular Point	نقطه منظم
Second-Order Sufficient Conditions	شرایط کافی مرتبه دوم
Stability in the Sense of Lyapunov	پایداری به مفهوم لیاپانوف
Strict Local Minimum Point	نقطه مینیمم محلی اکید
Strictly Monotone Map	نگاشت اکیداً یکنوا
Strongly Monotone Map	نگاشت قویاً یکنوا
Strongly Positive Definite	معین مثبت قوی
Switched-Capacitor	سوئیچ-خازن
Synapse	سیناپس
Synaptic Cleft	شکاف سیناپسی
Variational Inequalities	نامساوی‌های وردشی

## **Aabstract**

This involves enlarging the size of the optimization problems that exist in practice. The necessary conditions of efficiency in the use of techniques that enable high-speed, very large problems solved with acceptable quality can be felt more than.

Recently methods of optimization based on artificial intelligence approaches have been developed remarkable success in solving optimization problems efficiently acquired. Methods such as Genetic Algorithms, Tabu Search, refrigeration simulation and neural networks, their ability to solve large problems have good action. Special rates available on the possible application of neural networks in a wide range of research has provided. It points to the possibility of learning and performance improvement based on the input data point. It also allows parallel computations in a neural network is another advantage of the parallel hardware, enabling very large problems by this approach possible.

In this thesis, we tried two different models of recursive neural network is presented to solve optimization problems in the traces. Analysis of uniqueness, stability and convergence of global solutions are examined and the performance of the proposed methods using several examples of second-order cone constrained variational inequality problems is shown.

Finally, we provide conclusions and recommendations for future work.

**Keywords:** Neural networks, Stability of Lyapunov, Convergent, Variational Inequalities, second order cone programming.



Shahrood University  
Faculty Of Mathematical Sciences

Dissertation Submitted in Partial  
Fulfillment of The Requirements For The  
Degree of Master of Science in  
Applied Mathematics

**An efficient neural network for solving  
second-order con constrained variational  
inequality problem**

Supervisor  
**Dr. Alireza Nazemi**

by  
**Atiye Sabeghi**

2015