

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 مدل شبکه عصبی دوگانی برای حل رده
ای از مسائل برنامه ریزی درجه دوم

استاد راهنما

دکتر علیرضا ناظمی

دانشجو

معصومه عربشاهی مقدم

۱۳۹۳

تقدیم بہ روح پاک مادرم

او کہ آفتاب مهرش بر آستانہ قلم ہم چنان پابرجاست و ہرگز غروب نخواہد کرد.

تقدیم بہ پدر بزرگوارم

بہ پاس عاطفہ سرشار و کرمای امید بخش وجودش کہ در این سردترین روزگار ان بہترین پشتیبان من بودہ است.

تقدیم بہ ہمسر عزیزم

او کہ در سایہ ہمیاری و ہمدلی اش بہ این منظور نائل شدم.

تقدیم بہ دلبندم

او کہ وجودش نشانہ لطف الہی در زندگی من است.

کاش می شد...

کسی می آمد
باور تیره مارا می شست
و به مای فغاند
دل ما منزل تاریکی نیست...
"انخم" بر چهره بسی نازیباست...
بهترین واژه همان "لجنخند" است...
که ز لب های همه دور شدت...
کاش می شد...
که به انگشت نخ می بستیم...
تا فراموش نکردد...
که هنوز...
"انسایم"...

سپاس‌گزاری...

خدا را سپاس‌گزارم که هنوز راه‌هایی برای پیمودن و پله‌هایی برای صعود کردن برایم باقی است...

از پدر نازنینم... بابت تمام مهربانی‌ها و صبوری‌اش او که هم پدری فداکار بود و هم مادری دلسوز؛

از همسر مهربانم... اسطوره زندگی‌ام، امید بودنم او که با قلبی آکنده از عشق و معرفت محیطی سرشار از آرامش برایم فراهم آورده است؛

از خانواده محترم... به خصوص برادر بزرگوارم آقای عبدالهادی عربشاهی مقدم و خواهران بزرگوارم خانم‌ها زهرا و راضیه عربشاهی مقدم که سهم به‌سزایی در پیشرفت‌هایم داشته‌اند؛

از استاد با کمالات و فرزانه جناب آقای دکتر علیرضا ناظمی که در کمال سعه صدر و حسن خلق از هیچ کمکی در این عرصه دریغ ننمودند و همواره راهنما و راه‌گشای نگارنده در اتمام و اکمال پایان نامه بوده است؛

از اساتید فرزانه و دلسوز جناب آقای دکتر جعفر فتحعلی و جناب آقای دکتر مهرداد غزنوی که زحمت داوری این پایان نامه را متقبل شدند؛

از تمامی دوستان عزیزم که همیارم بودند از جمله خانم‌ها: شهلا خلیفی، عطیه سابقی، سمیرا سوخت‌سرایبی، مرضیه مرتضایی کمال تشکر و قدردانی را دارم. باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

تعمدنامه

اینجانب معصومه عربشاهی مقدم دانشجوی کارشناسی ارشد رشته ریاضی کاربردی دانشکده علوم ریاضی دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان یک مدل شبکه عصبی دوگانی برای حل رده‌ای از مسائل برنامه‌ریزی درجه دوم، 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.

معصومه عربشاهی مقدم

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 فن آوری 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 کارای مسائل بهینه‌سازی به دست آورده‌اند.

روش‌هایی چون الگوریتم ژنتیک، جست‌وجوی ممنوع، شبیه‌سازی تیریدی، 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 پایان‌نامه 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 برنامه‌ریزی درجه دوم محدب نشان داده می‌شود. در انتها نتایج کار و پیشنهاداتی برای کارهای آتی ارائه می‌دهیم.

فهرست مطالب

د	لیست تصاویر	
ذ	لیست جداول	
۱	مقدمه	۱
۱	تاریخچه تحقیق	۱.۱
۲	کاربرد مسائل برنامه ریزی درجه دوم	۲.۱
۳	مفاهیم اولیه و پیش‌نیازها	۲
۳	تعاریف	۱.۲
۳	توابع آفین و محدب	۲.۲
۷	پایداری	۳.۲
۹	تابع انرژی	۴.۲
۱۱	یک مدل شبکه عصبی دوگانی برای حل مسائل برنامه ریزی درجه دوم	۳
۱۱	مقدمه	۱.۳
۱۲	بررسی مدل های شبکه عصبی	۲.۳
۱۳	بیان مسأله	۳.۳
۱۴	مدل شبکه عصبی	۴.۳
۱۵	تحلیل پایداری مدل جدید شبکه عصبی	۵.۳
۱۶	محاسبات عددی	۶.۳
۲۱	یک مدل شبکه عصبی لاگرانژی برای حل مسائل برنامه ریزی درجه دوم کلی	۴
۲۱	مقدمه	۱.۴
۲۱	مدل شبکه عصبی	۲.۴
۲۷	مثال های محاسباتی	۳.۴
۳۹	مراجع	

خ

۴۲

واژه‌نامه فارسی به انگلیسی

۴۴

واژه‌نامه انگلیسی به فارسی

لیست تصاویر

۸	پایداری لیاپانوف	۱.۲
۸	پایداری مجانبی	۲.۲
۱۵	معماری شبکه های عصبی ارائه شده از (۱۵.۳)	۱.۳
۱۹	مسیرهای گذرای $y(t)$ از متغیرهای مسأله دوگان در مثال ۱.۶.۳	۲.۳
۲۰	مسیرهای گذرای $y(t)$ از متغیرهای مسأله دوگان در مثال ۲.۶.۳	۳.۳
۲۳	نمودار بلوکی ساده شده برای شبکه عصبی (۹.۴) (۱۰.۴)	۱.۴
	دیاگرام صوری شبکه عصبی (۹.۴) و (۱۰.۴) با پنج نقطه اولیه متفاوت در مثال	۲.۴
۲۸	رفتارهای گذرای $x(t)$ از شبکه عصبی (۹.۴) و (۱۰.۴) با ۲۰ نقطه اولیه متفاوت	۳.۴
۲۸	در مثال ۲.۳.۴	
۳۰	رفتار همگرایی $\ x(t) - x^*\ ^2$ در مثال ۲.۳.۴	۴.۴
۳۰	رفتارهای گذرا از شبکه عصبی (۹.۴) و (۱۰.۴) با ۲۰ نقطه اولیه متفاوت	۵.۴
	رفتار همگرایی $\ x(t) - x^*\ ^2$ در مثال ۳.۳.۴ با $y_0 = [-5, 4, -3, 2, -1, 0, 1, -2, 3, -4, 5]^T$	۶.۴
۳۱		
۳۲	نمودار شبکه عصبی با هشت نقطه اولیه در (۹.۴) و (۱۰.۴) مثال ۴.۳.۴	۷.۴
۳۲	رفتار گذرا از شبکه عصبی (۹.۴) و (۱۰.۴) با ۵۰ نقطه اولیه تصادفی در مثال ۵.۳.۴	۸.۴
۳۳	رفتار همگرایی $\ x(t) - x^*\ ^2$ در مثال ۵.۳.۴ با $y_0 = [-2, 2, -2, 2, -2, 2, -2, 2, -2, 2]^T$	۹.۴
	رفتار گذرای $x(t)$ از شبکه عصبی کندی چوآ (۲۲.۴) با ۵۰ نقطه اولیه متفاوت در	۱۰.۴
۳۳	مثال ۵.۳.۴	
	نتایج بردار پشتیبانی رگرسیون با استفاده از شبکه عصبی پیشنهادی (۹.۴) و (۱۰.۴)	۱۱.۴
۳۴	با یک کرنل RBF که $\xi = 100$ و $\delta = 1$ و سه پارامتر متفاوت ε	
	نتایج تقریب شبکه عصبی (۹.۴) و (۱۰.۴) در مثال ۵.۳.۴ که دایره معرف داده	۱۲.۴
۳۶	های تقریب و خم تابع تقریب است.	
	رفتار همگرایی $\ x(t) - x^*\ ^2$ با $x_0 = [1, -1, 1, -1, 1]^T$ و $\eta = 10$ در مثال	۱۳.۴
۳۷	مثال ۵.۳.۴	

لیست جداول

۱۸	$y_0 = [0, 0, 0, 0, 0, 0]^T$	نقطه اولیه	۱.۳
۱۸	$y_0 = [10, 10, 10, 10, 10, 10]^T$	نقطه اولیه	۲.۳
۲۰	$y_0 = [0, 0, 0, 0]^T$	نقطه اولیه	۳.۳
۳۴		داده های رگرسیونی	۱.۴
۳۵		داده های تقریبی	۲.۴

فصل ۱

مقدمه

در این پایان نامه دو مدل شبکه عصبی^۱ بازگشتی برای حل مسائل برنامه ریزی درجه دو^۲ ارائه می‌کنیم. شبکه‌های عصبی ارائه شده در این پایان نامه در مقایسه با شبکه‌های عصبی قبلی پیچیدگی کمتری برای پیاده‌سازی دارد. علاوه بر این، شبکه‌های عصبی پیشنهادی همگرایی سراسری به جواب بهینه مسأله هستند. نمونه‌های شبیه‌سازی بر اساس مشکلات پایه مورد بحث قرار گرفته اند تا عملکرد درست شبکه عصبی پیشنهادی را نشان دهند.

۱.۱ تاریخچه تحقیق

قدمت شبکه‌های عصبی مصنوعی به دهه ۵۰ میلادی بر می‌گردد. در سال ۱۹۵۶ بنیاد راکفلر برگزار کرد. کنفرانسی را بر عهده داشت که چشم‌اندازش به صورت زیر بود:

امکان استفاده از کامپیوترها و شبیه‌سازی در هر زمینه از یادگیری و سایر حوزه‌های آموزش. در همین کنفرانس بود که اصطلاح هوش مصنوعی مورد استفاده عمومی قرار گرفت. بطور کلی هوش مصنوعی را به صورت زیر می‌توان تعریف نمود:

فرایندهای کامپیوتری که سعی دارند تفکر انسان را تقلید نمایند، این فرایندها با فعالیت‌هایی که نیاز به استفاده از هوش دارند در ارتباطند.

عموماً این تعریف شامل زمینه‌های یادگیری خودکار، فهم زبان طبیعی، بینایی و تشخیص صدا، بازی کردن، حل مسائل ریاضی، رباتیک و سیستم‌های خبره است. در سال‌های اخیر برای محققان، شبکه‌های عصبی و فناوری‌های مرتبط با آن را به عنوان بخشی از هوش مصنوعی در نظر می‌گیرند. در حالی که بعضی دیگر با توجه به سررشته خود در علوم پزشکی، شبکه‌های عصبی را زیرمجموعه هوش مصنوعی نمی‌دانند.

^۱neural network

۲.۱ کاربرد مسائل برنامه ریزی درجه دوم

مسئله برنامه ریزی درجه دوم در زمینه های مختلف علوم و مهندسی مانند رگرسیون، طبقه بندی داده ها، پردازش تصویر، مسائل ریاضی مالی و... کاربرد فراوانی دارد. بالاخص در مهندسی و بسیاری از کاربردهای نظامی که برنامه های کاربردی نیازمند پردازش در زمان واقعی داده ها است، اغلب نیازمند استفاده از شبکه های عصبی به عنوان یک ابزار کارا در یافتن جواب های زمان واقعی هستیم. چون زمان مورد نیاز برای حل مسئله برنامه ریزی درجه دوم وابستگی زیادی به بعد و ساختار مسئله دارد لذا معمولاً روش های عددی مرسوم در کاربردهای واقعی چندان مؤثر نیستند. یک رهیافت امیدوار کننده عبارت است از استفاده از شبکه های عصبی بازگشتی مبنی بر پیاده سازی دوری با مدارهای الکترونیکی.

فصل ۲

مفاهیم اولیه و پیش‌نیازها

۱.۲ تعاریف

در این فصل تعاریف مورد نیاز در فصل‌های بعدی به‌طور مختصر آورده شده است. ابتدا تعاریف مرتبط با تابع محدب^۱ را بیان نموده و در خصوص مسائل بهینه‌سازی مقید و نامقید به ذکر شرایط لازم و کافی بهینگی^۲ می‌پردازیم. در انتها مفاهیم پایداری^۳ و تابع انرژی در سیستم‌های دینامیکی را ارائه می‌دهیم.

۲.۲ توابع آفین و محدب

تعریف ۱.۲.۲. تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ را تابع خطی می‌گوییم، هرگاه حافظ عملیات در فضای برداری \mathbb{R}^n باشد، یعنی:

$$f(x + y) = f(x) + f(y),$$

$$f(\alpha x) = \alpha f(x).$$

برای هر $x, y \in \mathbb{R}^n$ و هر $\alpha \in \mathbb{R}$.

تعریف ۲.۲.۲. تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}^l$ را تابع آفین می‌گوییم، هرگاه به‌صورت مجموع یک تابع خطی و یک تابع ثابت باشد. به عنوان مثال $f(x) = Ax + b$ که $A \in \mathbb{R}^{l \times n}$ و $b \in \mathbb{R}^l$ تابع آفین است.

تعریف ۳.۲.۲. مجموعه $E \subset \mathbb{R}^n$ را محدب می‌گوییم اگر:

$$\forall x, y \in E, \lambda \in (0, 1) : \lambda x + (1 - \lambda)y \in E.$$

^۱Convex

^۲Optimality

^۳Stability

تعریف ۴.۲.۲. تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ بر روی مجموعه نقاط واقع در مجموعه محدب $E \subset \mathbb{R}^n$ یک تابع محدب نامیده می‌شود، اگر:

$$\forall x, y \in E, \lambda \in (0, 1) : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

تابع f را روی E مقعر گوییم هرگاه $-f$ محدب باشد.

تعریف ۵.۲.۲. اگر $\bar{x} \in X$ و ε -همسایگی از \bar{x} مثل $N_\varepsilon(\bar{x})$ موجود باشد که به ازای هر $x \in X \cap N_\varepsilon(\bar{x})$ داشته باشیم $f(\bar{x}) \leq f(x)$ آنگاه، \bar{x} یک کمینه موضعی برای f است.

در توسعه برنامه‌ریزی غیرخطی، تابع محدب همواره به عنوان یک مفهوم اصلی ریاضی مورد نیاز است.

تعریف ۶.۲.۲. تابع $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$ را تابع نرم گوییم هرگاه به ازای هر $x, y \in \mathbb{R}^n$ و $\alpha \in \mathbb{R}$:

$$\begin{cases} \|x\| \geq 0, & \forall x, \\ \|x\| = 0 \Leftrightarrow x = 0, \\ \|\alpha x\| = |\alpha| \|x\|, \\ \|x + y\| \leq \|x\| + \|y\|. \end{cases}$$

برای بردار $x \in \mathbb{R}^n$ نرم‌های L_1, L_2 و L_r به صورت زیر تعریف می‌شوند،

$$\|x\|_1 = |x_1| + \dots + |x_n|,$$

$$\|x\|_2 = \sqrt{x^T x} = \sqrt{x_1^2 + \dots + x_n^2},$$

$$\|x\|_r = \sqrt[r]{|x_1^r| + \dots + |x_n^r|} = \sqrt[r]{\sum_{i=1}^n |x_i^r|}, \quad r \geq 1.$$

تعریف ۷.۲.۲. یک زیر مجموعه E از فضای متریک X را یک مجموعه فشرده می‌گوییم اگر هر پوشش باز E یک زیر پوشش باز متناهی داشته باشد.

شرایط لازم و کافی برای مسائل نامقید

قضیه ۸.۲.۲. (شرط لازم مرتبه اول [۲]): فرض کنید X یک مجموعه باز در \mathbb{R}^n باشد. مسأله برنامه‌ریزی زیر را در نظر بگیرید:

$$\text{minimize} \quad f(x) \tag{۱.۲}$$

subject to

$$x \in X \tag{۲.۲}$$

که در آن $f: \mathbb{R}^n \rightarrow \mathbb{R}$ در \hat{x} مشتق‌پذیر است. اگر \hat{x} یک کمینه موضعی باشد، آنگاه $\nabla f(\hat{x}) = 0$.

تعریف ۹.۲.۲. [۹] یک ماتریس $M(X)$ در اندازه $n \times n$ که عناصر آن m_{ij} ($i, j = 1, \dots, n$) توابعی روی $E \subset \mathbb{R}^n$ هستند،

الف. نیمه معین مثبت^۴ روی E نامیده می‌شود، اگر

$$\forall V \in \mathbb{R}^n, V \neq 0, X \in E : V^T M(X) V \geq 0.$$

ب. معین مثبت^۵ روی E نامیده می‌شود، اگر

$$\forall V \in \mathbb{R}^n, V \neq 0, X \in E : V^T M(X) V > 0.$$

قضیه ۱۰.۲.۲. (شرط لازم مرتبه دوم [۴]) : فرض کنید مشتق دوم $f : \mathbb{R}^n \rightarrow \mathbb{R}$ در \hat{x} موجود باشد. اگر \hat{x} یک کمینه موضعی باشد، آنگاه $\nabla f(\hat{x}) = 0$ و ماتریس هسین^۶ f در \hat{x} نیمه معین مثبت است.

قضیه ۱۱.۲.۲. (شرط کافی [۴]) : فرض کنید مشتق دوم $f : \mathbb{R}^n \rightarrow \mathbb{R}$ در \hat{x} وجود باشد. اگر $\nabla f(\hat{x}) = 0$ و ماتریس هسین f در \hat{x} معین مثبت باشد، آنگاه \hat{x} یک کمینه موضعی اکید است.

قضیه ۱۲.۲.۲. [۴] : فرض کنید مجموعه E بصورت زیر باشد:

$$E = \{x \in \mathbb{R}^n \mid g_k(x) \leq 0, k = 1, \dots, m, h_p(x) = 0, p = 1, \dots, l\}$$

اگر $g_k(x)$ ، $k = 1, \dots, m$ ، توابعی محدب و $h_p(x)$ ، $p = 1, \dots, l$ ، آفین باشند آنگاه، E یک مجموعه محدب است.

قضیه ۱۳.۲.۲. [۴] : اگر $f(x)$ بر مجموعه محدب E تابعی محدب باشد آنگاه $f(x)$ دارای یک کمینه موضعی^۷ است. اگر چنین کمینه‌ای یافت شود، یک کمینه سراسری^۸ است و بر روی مجموعه محدب به دست آورده می‌شود.

قضیه ۱۴.۲.۲. [۹] اگر $\gamma(x)$ کوچکترین مقدار ویژه ماتریس متقارن $M(X)_{n \times n}$ باشد آنگاه:

الف. روی E نیمه معین مثبت اگر و فقط اگر به ازای هر $x \in E$ ، $\gamma(x) \geq 0$.

ب. روی E معین مثبت اگر و فقط اگر به ازای هر $x \in E$ ، $\gamma(x) > 0$.

شرایط لازم و کافی برای مسائل مقید

قضیه ۱۵.۲.۲. (شرایط لازم کروش-کان-تاکر^۹ (K.K.T))، [۴] : فرض کنید X یک مجموعه باز در \mathbb{R}^n باشد. مسأله کمینه‌سازی مقید زیر را در نظر بگیرید:

^۴Positive semi-definite

^۵Positive definite

^۶Hessian

^۷Local minimum

^۸Global minimum

^۹Karush-Kuhn-Tucker

$$\text{minimize} \quad f(x) \quad (۳.۲)$$

subject to

$$\begin{cases} g_k(x) \leq 0, & k = 1, \dots, m, \\ h_j(x) = 0, & j = 1, \dots, l, \\ x = (x_1, x_2, \dots, x_n)^T \in X, \end{cases} \quad (۴.۲)$$

که در آن $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ، $g_k: \mathbb{R}^n \rightarrow \mathbb{R}$ $k = 1, \dots, m$ و $h_j: \mathbb{R}^n \rightarrow \mathbb{R}$ $j = 1, \dots, l$ فرض کنید \hat{x} یک جواب شدنی این مسأله باشد و $K = \{k : g_k(\hat{x}) = 0\}$. همچنین فرض کنید f و g_k برای $k \in K$ در \hat{x} مشتق پذیر، g_k برای $k \notin K$ در \hat{x} پیوسته و h_j برای $(j = 1, \dots, l)$ دارای مشتق پیوسته باشند. به علاوه فرض کنید $\nabla g_k(\hat{x})$ برای $k \in K$ و $\nabla h_j(\hat{x})$ برای $(j = 1, \dots, l)$ مجموعاً مستقل خطی باشند. اگر \hat{x} کمینه موضعی برای مسأله (۳.۲) و (۴.۲) باشد، آن‌گاه اسکالرهایی یکتای $\hat{u}_k \geq 0$ برای $k \in K$ و $\hat{v}_j \geq 0$ برای $(j = 1, \dots, l)$ موجود هستند به طوری که:

$$\begin{cases} \nabla f(\hat{x}) + \sum_{k \in K} \hat{u}_k \nabla g_k(\hat{x}) + \sum_{j=1}^l \hat{v}_j \nabla h_j(\hat{x}) = 0, \\ \hat{u}_k \geq 0, \quad k \in K. \end{cases}$$

علاوه بر فرض‌های بالا اگر g_k برای $k \notin K$ در \hat{x} مشتق پذیر باشد، آن‌گاه شرایط کروش-کان-تاکر می‌تواند به صورت زیر نوشته شود:

$$\begin{cases} \nabla f(\hat{x}) + \sum_{k=1}^m \hat{u}_k \nabla g_k(\hat{x}) + \sum_{j=1}^l \hat{v}_j \nabla h_j(\hat{x}) = 0, \\ \hat{u}_k g_k(\hat{x}) = 0, \quad k = 1, \dots, m, \\ \hat{u}_k \geq 0, \quad k = 1, \dots, m. \end{cases}$$

همچنین شرایط کروش-کان-تاکر می‌تواند به فرم ماتریسی زیر نوشته شود:

$$\begin{cases} \nabla f(\hat{x}) + \nabla g(\hat{x})^T u + \nabla h(\hat{x})^T v = 0, \\ u^T g(\hat{x}) = 0, \\ u \geq 0, \end{cases}$$

که در آن $\nabla g(\hat{x})$ ماتریس ژاکوبی $m \times n$ و $\nabla h(\hat{x})$ ماتریس ژاکوبی $l \times n$ است و \hat{u} یک بردار m تایی و \hat{v} یک بردار l تایی است. $(\hat{u}^T, \hat{v}^T)^T$ بردار ضرایب لاگرانژ نامیده شده است.

قضیه ۱۶.۲.۲. (شرایط کافی کروش-کان-تاکر (K.K.T)، [۴]): فرض کنید X یک مجموعه باز در \mathbb{R}^n باشد، مسأله (۳.۲) و (۴.۲) را در نظر بگیرید. فرض کنید \hat{x} یک جواب شدنی باشد، همچنین فرض کنید f و g_k ها برای $(k = 1, \dots, m)$ توابعی محدب و h_j ها برای $(j = 1, \dots, l)$ توابعی آفین باشند و $K = \{k : g_k(\hat{x}) = 0\}$ ، اگر شرایط کروش-کان-تاکر در \hat{x} برقرار باشد، یعنی اسکالرهایی $\hat{u}_k \geq 0$ برای $k \in K$ و $\hat{v}_j \geq 0$ برای $(j = 1, \dots, l)$ موجود باشند به طوری که،

$$\nabla f(\hat{x}) + \sum_{k \in K} \hat{u}_k \nabla g_k(\hat{x}) + \sum_{j=1}^l \hat{v}_j \nabla h_j(\hat{x}) = 0,$$

آن‌گاه \hat{x} یک کمینه موضعی برای (۳.۲) و (۴.۲) خواهد بود.

قضیه ۱۷.۲.۲ ([۴]): اگر در مسأله (۳.۲) و (۴.۲)، g_k ها ($k = 1, \dots, m$) محدب و h_j ها ($j = 1, \dots, l$) آفینی باشند، شرایط کروش-کان-تاکر لازم و کافی هستند. مسأله برنامه‌ریزی درجه دوم^{۱۰} QP در حالت کلی به صورت زیر بیان می‌شود:

$$\text{minimize} \quad q(x) = \frac{1}{2}x^T Qx + x^T c$$

subject to

$$a_i^T x_i = b_i, \quad i \in \varepsilon,$$

$$a_i^T x_i \geq b_i, \quad i \in \eta.$$

که در آن Q یک ماتریس متقارن $n \times n$ است. ε و η مجموعه متناهی از اندیس‌ها هستند و c و x و a_i بردارهایی در \mathbb{R}^n هستند. اگر ماتریس Q نیمه‌معین مثبت باشد آن‌گاه مسأله برنامه‌ریزی درجه دوم محدب است و اگر ماتریس Q نامعین باشد آن‌گاه مسأله کمی چالش برانگیز خواهد شد، چون مسأله می‌تواند چندین نقطه مینیم موضعی داشته باشد.

۳.۲ پایداری

سیستم دینامیکی زیر را در نظر بگیرید:

$$\begin{cases} \dot{x} = f(x(t)), \\ x(t_0) = x_0 \in \mathbb{R}^n. \end{cases} \quad (5.2)$$

که در آن $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ است.

تعریف ۱.۳.۲. x^* یک نقطه تعادل^{۱۱} برای (۵.۲) نامیده می‌شود اگر $f(x^*) = 0$.

تعریف ۲.۳.۲. فرض کنید که $x(t)$ یک جواب (۵.۲) باشد، نقطه تعادل تنها x^* ، پایدار به مفهوم لیاپانوف^{۱۲} است (شکل (۱.۲) را ببینید)، اگر برای هر $x_0 = x(t_0)$ و هر $\varepsilon > 0$ ، یک $\delta > 0$ وجود داشته باشد به طوری که اگر $\|x(t) - x_0\| \leq \delta$ آن‌گاه

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

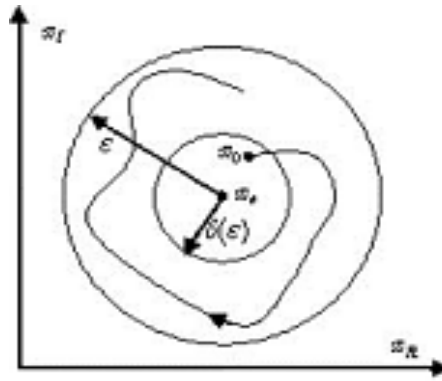
تعریف ۳.۳.۲. سیستم دینامیکی (۵.۲) همگرای سراسری^{۱۳} به مجموعه

^{۱۰}Quadratic programming

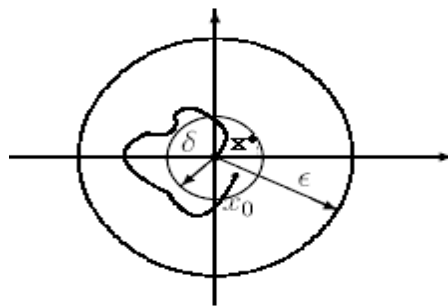
^{۱۱}Equilibrium point

^{۱۲}Lyapunov

^{۱۳}Global convergance



شکل ۱.۲: پایداری لیپانوف



شکل ۲.۲: پایداری مجانبی

$\Omega^* = \{x \mid \text{سیستم (۵.۲) جواب دارد}\}$.

گفته می‌شود اگر هر جواب دلخواه $x(t)$ از سیستم در رابطه زیر صدق کند:

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

که در آن

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

تعریف ۴.۳.۲. سیستم دینامیکی در نقطه تعادل یکتا x^* پایدار مجانبی سراسری^{۱۴} نامیده می‌شود اگر x^* به مفهوم لیپانوف پایدار باشد و در شرط زیر صدق کند:

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

شکل (۲.۲) را ببینید.

تعریف ۵.۳.۲. مجموعه $M \subseteq \mathbb{R}^n$ ، یک مجموعه پایدار^{۱۵} نسبت به سیستم (۵.۲) گفته می‌شود اگر $x(t) \in M$ و $t \geq 0$ ، آن‌گاه به ازای هر $t_0 \geq 0$ ، $x(t) \in M$ باشد.

تعریف ۶.۳.۲. نگاهت F در شرط لیپ شیتز^{۱۶} صدق می‌کند اگر عدد ثابت L وجود داشته باشد،

^{۱۴}Globally asymptotically stable

^{۱۵}Invariant set

^{۱۶}Lipschitz

به طوری که برای هر دو نقطه $x, y \in \mathbb{R}^n$ داشته باشیم:

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

تعریف ۷.۳.۲. F را لیپ شیتز محلی روی \mathbb{R}^n نامیم اگر به ازای هر نقطه از \mathbb{R}^n ، یک همسایگی مانند $D \subset \mathbb{R}^n$ وجود داشته باشد به طوری که نامساوی بالا برای هر دو نقطه $x, y \in D$ برقرار باشد.

تعریف ۸.۳.۲. نگاشت $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یکنوا^{۱۷} گفته می‌شود، اگر برای هر زوج از نقاط $x, y \in \mathbb{R}^n$ داشته باشیم:

$$(x - y)^T (F(x) - F(y)) \geq 0.$$

همچنین روی \mathbb{R}^n اکیداً یکنوا^{۱۸} است، اگر نامساوی فوق به صورت اکید برای هر $x \neq y$ برقرار باشد. و F روی \mathbb{R}^n یکنوای قوی است اگر برای هر زوج از نقاط $x, y \in \mathbb{R}^n$ ثابت $\beta > 0$ وجود داشته باشد، به طوری که:

$$(x - y)^T (F(x) - F(y)) \geq \beta \|x - y\|^2.$$

قضیه ۹.۳.۲. ([۱]): فرض می‌کنیم که $F: K \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ ، یک تابع به طور پیوسته مشتق پذیر روی K باشد و ماتریس ژاکوبین F که لزوماً متقارن نیست، نیمه معین مثبت (معین مثبت) باشد، آنگاه $F(x)$ یکنوا (یکنوای قوی) است.

قضیه ۱۰.۳.۲. ([۱]): در سیستم دینامیکی (۵.۲) فرض می‌کنیم که $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ یک تابع پیوسته باشد، آنگاه برای هر $t_0 > 0$ و $x_0 \in \mathbb{R}^n$ ، یک جواب محلی^{۱۹} $x(t)$ به ازای $t \in [t_0, \tau)$ که $\tau > t_0$ وجود دارد. علاوه بر این اگر f در x در شرط لیپ شیتز محلی صدق کند آنگاه جواب یکتا خواهد بود و اگر f در \mathbb{R}^n در شرط لیپ شیتز صدق کند آنگاه τ می‌تواند تا $+\infty$ توسعه داده شود.

۴.۲ تابع انرژی

قضیه ۱۰.۴.۲. ([۱]): فرض کنید که $x = 0$ یک نقطه تعادل برای سیستم (۵.۲) و تابع $E: \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ به طور پیوسته مشتق پذیر باشد اگر

$$E(0) = 0. \quad ۱.$$

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

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

آنگاه $x = 0$ نقطه پایداری لیاپانوف سیستم خواهد بود و $E(x)$ را «تابع لیاپانوف» یا «تابع انرژی» برای سیستم (۵.۲) نامیم.

^{۱۷}Monotone

^{۱۸}Strictly monotonic

^{۱۹}Local solution

قضیه ۲.۴.۲. [۱] اصل تغییرناپذیری (ناوردای) لازال^{۲۰}: فرض می‌کنیم تابع $V : \mathbb{R}^n \rightarrow \mathbb{R}$ به‌طور پیوسته مشتق پذیر باشد، همچنین فرض می‌کنیم

۱. $M \subseteq \mathbb{R}^n$ یک مجموعه فشرده و پایدار نسبت به جواب سیستم (۵.۲) باشد،

۲. به ازای هر $x \in M$ ، $\dot{V}(x) \leq 0$ ،

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

۴. E بزرگ‌ترین مجموعه پایدار در \mathbb{R}^n ، $N \subseteq \mathbb{R}^n$ ، باشد

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

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

^{۲۰}Lasalle's invariance principle

فصل ۳

یک مدل شبکه عصبی دوگانی برای حل مسائل برنامه ریزی درجه دوم

۱.۳ مقدمه

مسائل بهینه سازی با تابع هدف غیر خطی اغلب با یک سیستم مرتبه دوم تقریب زده می شود و جواب به طور تقریبی با روش برنامه ریزی درجه دوم استاندارد به دست می آید. برنامه ریزی درجه دوم به یک رده خاصی از برنامه ریزی غیر خطی اشاره می کند که در آن تابع هدف درجه دوم است و قیدها خطی هستند. روش مرسوم برای حل مسائل برنامه ریزی درجه دوم به طور معمول شامل یک فرایند تکرار شونده است، اما زمان محاسباتی طولانی استفاده از آن ها را محدود می کند، زیرا الگوریتم های معمول برای کامپیوترهای دیجیتال مؤثر نیستند. همچنین زمان محاسبات مورد نیاز برای یک راه حل، تا حد زیادی به ابعاد و ساختار مسأله وابسته است. یک راه حل امیدوارکننده برای حل مسائل بهینه سازی زمان واقعی^۱ به کار بردن شبکه های عصبی مصنوعی است. برای فرموله کردن یک مسأله بهینه سازی در نظریه شبکه عصبی دو نوع روش وجود دارد. روش اول مسأله بهینه سازی مقید را به یک مسأله بهینه سازی نامقید تبدیل می کند و آن گاه یک شبکه عصبی را با روش گرادیان پیشنهاد می کنند. روش دیگر [۹]، [۱۲]، [۳۰] شامل یک مجموعه از معادلات دیفرانسیلی بر مبنای نظریه دوگان است به طوری که نقاط تعادل مدل ارائه شده جواب بهینه مسأله اصلی است. آن گاه یک تابع لیاپانف مناسب به طوری که تمام مسیرهای سیستم همگرا به نقاط تعادل باشد، پیدا می کنیم. در حال حاضر شبکه های عصبی مصنوعی در رده بسیار وسیعی از مسائل بهینه سازی مقید استفاده می شود.

^۱Real-time optimization

۲.۳ بررسی مدل های شبکه عصبی

در این بخش برخی مدل های شبکه عصبی پیشین را بررسی می کنیم [۱۰]. در سال ۱۹۸۵ تنک و هاپ فیلد مقاله ای منتشر کردند [۲۸] که در آن یک روش جدید برای حل مسائل برنامه نویسی غیر خطی با استفاده از شبکه عصبی مطرح شده است. برخلاف روش های سنتی، تنک و هاپ فیلد^۲ توانستند از قطعات الکترونیکی آنالوگ که به صورت موازی عمل می کنند، استفاده کنند. در سال ۱۹۸۷ کندی و چوا [۱۷]^۳ مدل بهتری را ارائه کردند که همیشه همگرایی را تضمین می کند. مسأله برنامه ریزی غیر خطی زیر را بررسی می کنیم:

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && g(x) = [g_1(x), \dots, g_m(x)] \leq 0 \\ & && f(x), g_j(x) : \mathbb{R}^n \rightarrow \mathbb{R}^1, \quad j = 1, 2, \dots, m, \\ & && f(x), g(x) \in C^2. \end{aligned}$$

مدل کندی و چوا [۱۷] به شکل زیر تشریح می شود:

$$\dot{x} = C^{-1} \{-\nabla f(x) - s \nabla g(x) g^+(x)\}, \quad (1.3)$$

که در آن C یک ماتریس $n \times n$ است، s یک پارامتر جریمه است و

$$g^+(x) = [g_1^+(x), g_2^+(x), \dots, g_m^+(x)],$$

که

$$g_j^+(x) = \max\{0, g_j(x)\},$$

تابع انرژی برای مدل (۱.۳) به صورت زیر تعریف می شود:

$$E(x) = f(x) + \frac{s}{2} \sum_{j=1}^m (g_j^+(x))^2. \quad (2.3)$$

این مدل تنها به یک تقریب از جواب بهینه همگراست. علاوه بر این در سال ۱۹۹۶ زیا و همکاران [۳۲] مدلی جدید معرفی و این اشکال را برطرف کردند. مسأله QP را به شکل زیر بررسی می کنیم:

$$\begin{aligned} & \text{minimize} && f(x) = \frac{1}{2} x^T A x + a^T x \\ & \text{subject to} && g(x) = D x - b = 0, \quad x \geq 0, \end{aligned} \quad (3.3)$$

که در آن A ماتریس $m \times n$ نیمه معین مثبت، D ماتریس $n \times m$ و $rank(D) = m$ و $b \in \mathbb{R}^n$ و $a \in \mathbb{R}^m$. دوگان مسأله (۳.۳)، DQP را به صورت زیر تعریف می کنیم:

$$\begin{aligned} & \text{maximize} && f(x) = -\frac{1}{2} x^T A x + b^T y \\ & \text{subject to} && g(x) = D^T y - A x - a \leq 0. \end{aligned} \quad (4.3)$$

^۲Hopfield and Tank

^۳Kennedy and Chua

طبق شرایط کان تاکر، مدل زیا و همکاران به شکل زیر توصیف می‌شود:

$$\frac{dx}{dt} = - \left\{ D^T(Dx - b) + \gamma(-D^T y + Ax + a) + \gamma A \left[x - (x + D^T y - Ax - a)^+ \right] \right\}, \quad (5.3)$$

$$\frac{dy}{dt} = -\gamma \{ Dx - b + D [(x + D^T y - Ax - a)^+ - x] \}, \quad (6.3)$$

که در آن

$$\gamma = \|(x + D^T y - Ax - a)^+ - x\|^{-1}, \quad (x, y)^T \in G = \{(x, y)^T | x \geq 0\}.$$

مدل زیا و همکاران هر دو مسئله اولیه و دوگان مسئله برنامه ریزی درجه دوم را با هم حل می‌کند و نیاز به پارامتر میزان سازی^۴ ندارد. در سال ۲۰۰۵ عفتی و بایمانی^۵ یک مدل دینامیکی تعریف کردند که هر دو مسئله اولیه و دوگان را با همگرایی سریع‌تر حل می‌کند. مدل عفتی و بایمانی [۸] با سیستم دینامیکی غیر خطی زیر تشریح می‌شود:

$$\frac{dx}{dt} = -a - A \left(x + k \frac{dx}{dt} \right) + D^T \left(y + k \frac{dy}{dt} \right), \quad x \geq 0, \quad (7.3)$$

$$\frac{dy}{dt} = b - D \left(x + k \frac{dx}{dt} \right), \quad (8.3)$$

که در آن ضریب k یک ثابت مثبت است. در سال ۲۰۰۷ اسکویی^۶ و همکاران ([۱۲] را ببینید) مدل دیگری با یک سیستم غیر خطی دینامیکی معرفی کردند که از شرایط اولیه اختیاری استفاده می‌کند. این مدل هم‌زمان به جواب‌های دقیق اولیه و دوگان همگراست.

۳.۳ بیان مسأله

مسأله برنامه ریزی درجه دوم را با فرض داشتن جواب یکتا به شکل زیر بررسی می‌کنیم:

$$\begin{aligned} \text{minimize} \quad & f(x) = \frac{1}{\gamma} x^T H x + c^T x \\ \text{subject to} \quad & Ax - b \leq 0, \end{aligned} \quad (9.3)$$

که در آن H یک ماتریس $n \times n$ متقارن معین مثبت است. بنابراین تابع هدف اکیدا محدب است. A یک ماتریس $m \times n$ و $rank(A) = m$ و $b \in \mathbb{R}^m$ و $c \in \mathbb{R}^n$. مسئله دوگان لاگرانژ عبارت است از ماکزیم کردن $\theta(y)$ روی $y \geq 0$ که در آن:

$$\theta(y) = \inf \left\{ \frac{1}{\gamma} x^T H x + c^T x + y^T (Ax - b) : x \in \mathbb{R}^n \right\}. \quad (10.3)$$

^۴Tunning

^۵Effati and Baymani

^۶Oskoei

توجه کنید برای یک y مفروض، تابع $\frac{1}{2}x^T Hx + c^T x + y^T (Ax - b)$ اکیدا محدب است و مینیم آن در یک نقطه مناسب به دست می‌آید که در رابطه زیر صادق است:

$$Hx + A^T y + c = 0. \quad (11.3)$$

بنابراین مسأله دوگان می‌تواند به شکل زیر نوشته شود:

$$\begin{aligned} \text{maximize} \quad & \frac{1}{2}x^T Hx + c^T x + y^T (Ax - b) \\ \text{subject to} \quad & Hx + A^T y = -c, \\ & y \geq 0. \end{aligned} \quad (12.3)$$

حال شکل دیگری از مسأله دوگان را بررسی می‌کنیم. چون H ماتریس مثبت معین است پس H^{-1} وجود دارد، بنابراین:

$$x = -H^{-1}(c + A^T y). \quad (13.3)$$

لم ۱۰.۳.۳. مسأله (۱۲.۳) هم‌ارز با مسئله زیر است:

$$\begin{aligned} \text{maximize} \quad & \theta(y) = \frac{1}{2}y^T D y + y^T d - \frac{1}{2}c^T H^{-1}c \\ \text{subject to} \quad & y \geq 0. \end{aligned} \quad (14.3)$$

که در آن $D = -AH^{-1}A^T$ و $d = -b - AH^{-1}c$

برهان. با جای‌گذاری (۱۳.۳) در (۱۰.۳) داریم:

$$\begin{aligned} \theta(y) &= \frac{1}{2} [(-H^{-1}(c + A^T y))^T H (-H^{-1}(c + A^T y))] \\ &\quad + c^T (-H^{-1}(c + A^T y)) + y^T [A(-H^{-1}(c + A^T y)) - b] \\ &= -\frac{1}{2}y^T AH^{-1}A^T y - y^T b - y^T AH^{-1}c - \frac{1}{2}c^T H^{-1}c. \end{aligned}$$

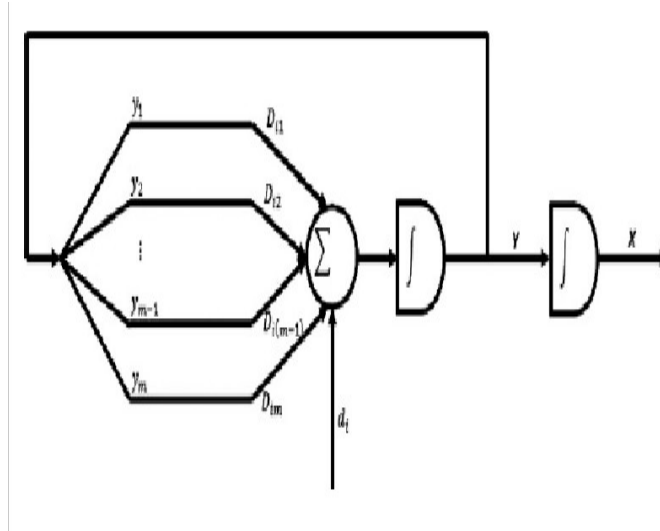
با در نظر گرفتن $D = -AH^{-1}A^T$ و $d = -b - AH^{-1}c$ ، مسأله (۱۴.۳) به دست می‌آید و لم ثابت می‌شود. \square

۴.۳ مدل شبکه عصبی

مدل جدید شبکه عصبی بر طبق مسئله (۱۴.۳) با یک سیستم دینامیکی غیر خطی ساده به شکل زیر تشریح می‌شود:

$$\frac{dy}{dt} = Dy + d, \quad y \geq 0. \quad (15.3)$$

سیستم تشریح شده با (۱۵.۳) می‌تواند به آسانی با یک شبکه عصبی بازگشتی تک لایه ای در شکل (۱.۳) نشان داده شود.



شکل ۱.۳: معماری شبکه های عصبی ارائه شده از (۱۵.۳)

۵.۳ تحلیل پایداری مدل جدید شبکه عصبی

در این بخش پایداری نقطه تعادل و همگرایی شبکه عصبی (۱۵.۳) را مطالعه می‌کنیم.

قضیه ۱.۵.۳. الف. y^* نقطه تعادل شبکه عصبی (۱۵.۳) است اگر و تنها اگر $y^* \in Y^* = \{y \mid \text{است (۱۴.۳)}\}$

ب. برای هر $y_0 = y(t_0) \in \mathbb{R}_m^+$ یک جواب پیوسته یکتا برای معادله (۱۵.۳) روی $[t_0, +\infty)$ وجود دارد.

برهان. الف. فرض کنید $y \in \mathbb{R}_m^+$ و $G(y) = Dy + d$. بنابراین $G(y^*) = 0$ برای $y^* \in \mathbb{R}_m^+$ اگر و تنها اگر $y^* \in Y^*$.

ب. برای $u, v \in \mathbb{R}_m^+$ داریم:

$$\|G(u) - G(v)\| = \|Du + d - Dv - d\| = \|Du - Dv\| = \|D\| \|u - v\|$$

بنابراین G پیوسته لیپ شیتز با ثابت $\|D\|$ است طبق قضیه (۱۰.۳.۲) از فصل ۲ سیستم (۱۵.۳) جواب یکتا دارد.

□

در قضیه زیر نشان می‌دهیم معادله دیفرانسیل (۱۵.۳) به طور مجانبی پایدار است.

قضیه ۲.۵.۳. شبکه عصبی توصیف شده با معادله دیفرانسیل (۱۵.۳) پایدار مجانبی به $y^* \in Y^*$ است.

برهان. با استفاده از قضیه (۱.۵.۳) می‌دانیم یک جواب پیوسته و یکتا $y(t)$ از (۱۵.۳) روی $[t_0, +\infty)$ با نقطه اولیه $y(t_0) \in \mathbb{R}_m^+$ وجود دارد. تابع انرژی زیر را بررسی می‌کنیم:

$$E(y) = e^{-(\frac{1}{\gamma} y^T D y + y^T d)} - e^{-(\frac{1}{\gamma} y^{*T} D y^* + y^{*T} d)}. \quad (۱۶.۳)$$

نشان می‌دهیم $E(y)$ یک تابع لیاپانف سیستم (۱۵.۳) است. در نقطه تعادل y^* ، $E(y^*) = 0$. همچنین $E(y)$ مشتق پذیر و معین مثبت در همسایگی y^* است زیرا $E(y) > 0$. کافی است نشان دهیم $\frac{d(E(y(t)))}{dt} \leq 0$. مشتق $E(y)$ را نسبت به زمان t پیدا می‌کنیم، داریم:

$$\frac{dE(y)}{dt} = \frac{dE(y)}{dy} \frac{dy}{dt} = -(Dy + d) e^{-(\frac{1}{\gamma} y^T D y + y^T d)} (Dy + d) = -\|Dy + d\|^2 e^{-(\frac{1}{\gamma} y^T D y + y^T d)} \leq 0.$$

لذا $\frac{dE(y)}{dt} < 0$ به جز برای $y = y^*$. بنابراین $y(t)$ همگرای مجانبی به y^* است. \square

برای حل شبکه عصبی (۱۵.۳) از روش اویلر استفاده می‌شود. کد مطلب زیر پیاده سازی اجرای شبکه عصبی را شرح می‌دهد.

```
for i = 1 : n;
    dy = dt(Dy + d);
    dy = max(y + dy, 0) - y;
    y = y + dy;
    x = -H^-1(c + A^T y);
    z = theta(y) = \frac{1}{\gamma} y^T D y + y^T d - \frac{1}{\gamma} c^T H^-1 c;
    Y(:, i) = y;
end.
```

۶.۳ محاسبات عددی

مثال ۱.۶.۳. مسأله زیر را در نظر بگیرید:

$$\begin{aligned} &\text{minimize} && x_1^2 + x_2^2 + x_1 x_2 - 30x_1 - 30x_2 \\ &\text{subject to} && \frac{5}{12}x_1 - x_2 \leq \frac{35}{12}, \\ &&& \frac{5}{2}x_1 + x_2 \leq \frac{35}{2} \\ &&& -x_1 \leq 5, \\ &&& x_2 \leq 5, \\ &&& x_1, x_2 \geq 0, \end{aligned}$$

مسأله بالا می‌تواند به مسأله زیر تبدیل شود:

$$\begin{aligned}
 \text{minimize} \quad & f(x) = x_1^2 + x_2^2 + x_1x_2 - 30x_1 - 30x_2 \\
 \text{subject to} \quad & g_1(x) = \frac{5}{12}x_1 - x_2 \leq \frac{35}{12}, \\
 & g_2(x) = \frac{5}{4}x_1 + x_2 \leq \frac{35}{2} \\
 & g_3(x) = -x_1 \leq 5, \\
 & g_4(x) = x_2 \leq 5, \\
 & g_5(x) = -x_1 \leq 0, \\
 & g_6(x) = -x_2 \leq 0,
 \end{aligned}$$

مسأله QP می‌تواند به صورت زیر نوشته شود:

$$\begin{aligned}
 \text{minimize} \quad & f(x) = \frac{1}{2}x^T Hx + c^T x \\
 \text{subject to} \quad & Ax - b \leq 0,
 \end{aligned}$$

که در آن

$$c^T = [-30, -30], \quad A = \begin{bmatrix} \frac{5}{12} & -1 \\ \frac{5}{4} & 1 \\ -1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} \frac{35}{12} \\ \frac{35}{2} \\ 5 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad H = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

جواب دقیق مسائل QP و DQP به شکل زیر به دست آمده است:

$$x = [5, 5]^T, \quad y = [0, -6, 0, 9, 0, 0]^T.$$

حال از مدل جدید شبکه عصبی برای حل مسأله QP به شکل زیر استفاده می‌کنیم:

$$\frac{dy}{dt} = Dy + d, \quad y \geq 0,$$

که در آن $D = -AH^{-1}A^T$ و $d = -b - AH^{-1}c$. با استفاده از روش اوایلر برای حل شبکه عصبی فوق با انتخاب $n = 100$ ، $dt = 0.2$ ، جواب بهینه مسأله DQP به شکل (۲.۳) به دست می‌آید:

$$y^* = [0, 6, 0, 9, 0, 0]^T$$

با جای‌گذاری y^* در (۱۱.۳) جواب بهینه مسأله اصلی به دست می‌آید:

$$x^* = [5, 5]^T.$$

مدل ارئه شده	مدل عفتی بايمني	مدل زيا	تعداد تکرارها
۱۰۰	۱۵۰۰	۲۰۰,۰۰۰	
$x^* = [۵, ۵]^T$	$x^* = [۵, ۵]^T$	$x^* = [۵,۰۰۴۵, ۴,۹۹۰۶]^T$	جواب QP
$y^* = [۰, ۶, ۰, ۹, ۰, ۰]^T$	$y^* = [۰, ۶, ۰, ۹]^T$	$y^* = [۰,۱۱۸, ۵,۹۳۸, ۰,۰۵۱, ۹,۳۶۲]^T$	جواب DQP
۰,۰۱۱۲۵	۱,۶۵	۹۲۵,۰۳	زمان (s)

جدول ۱.۳: نقطه اوليه $y_0 = [۰, ۰, ۰, ۰, ۰, ۰]^T$

مدل ارئه شده	مدل عفتی بايمني	مدل زيا	تعداد تکرارها
۱۰۰	۱۵۰۰	۲۰۰,۰۰۰	
$x^* = [۵, ۵]^T$	$x^* = [۵, ۵]^T$	$x^* = [۴,۹۹۸, ۵,۰۰۴]^T$	جواب QP
$y^* = [۰, ۶, ۰, ۹, ۰, ۰]^T$	$y^* = [۰, ۶, ۰, ۹]^T$	$y^* = [۰,۰۳۳, ۶,۰۴۰, ۰,۰۵۱, ۸,۸۳۴]^T$	جواب DQP
۰,۰۰۰۹۲	۱,۵۹	۹۲۰,۷۴	زمان (s)

جدول ۲.۳: نقطه اوليه $y_0 = [۱۰, ۱۰, ۱۰, ۱۰, ۱۰, ۱۰]^T$

برای مسأله QP با دو متغیر و شش ثابت، مدل جدید پس از حدود ۱۰۰ تکرار همگراست. در مقایسه، بیشتر از ۱۵۰۰ و ۲۰۰۰۰۰ تکرار برای حل همان مسأله از مدل عفتی-بایماني طول می‌کشد. [۸] را مشاهده کنید). این مدل نسبت به مدل‌های قبلی شکل ساده‌تر، سرعت همگرایی بیشتر، تعداد عملیات محاسباتی و متغیرهای کمتری دارد. همچنین با شروع از هر نقطه اولیه، سیستم پویای غیرخطی سرعت همگرایی خوبی دارد (جدول (۱.۳) و (۲.۳) را مشاهده کنید).

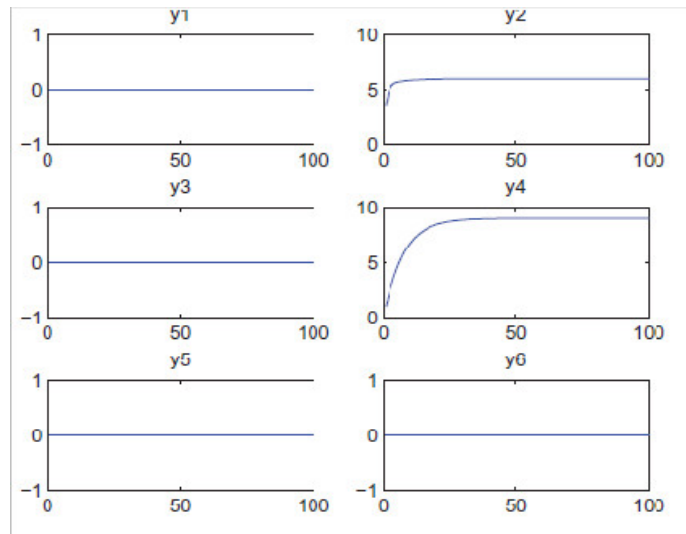
مثال ۲.۶.۳. مسأله زیر مفروض است:

$$\begin{aligned} &\text{minimize} && -2x_1 - 6x_2 + x_1^2 - 2x_1x_2 + 2x_2^2 \\ &\text{subject to} && x_1 + x_2 \leq 2, \\ &&& -x_1 + 2x_2 \leq 2, \\ &&& x_1, x_2 \geq 0. \end{aligned}$$

جواب دقیق مسأله $x^* = [۰,۸, ۱,۲]^T$ است.

شکل ماتریسی مسأله QP را به صورت زیر بررسی می‌کنیم:

$$\begin{aligned} &\text{minimize} && f(x) = \frac{1}{2}x^T Hx + c^T x \\ &\text{subject to} && Ax - b \leq 0, \end{aligned}$$



شکل ۲.۳: مسیرهای گذرای $y(t)$ از متغیرهای مسأله دوگان در مثال ۱.۶.۳

که در آن

$$c^T = [-2, -6], \quad A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad H = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}.$$

حال از مدل جدید شبکه عصبی برای حل مسئله QP به شکل زیر استفاده می‌کنیم:

$$\frac{dy}{dt} = Dy + d, \quad y \geq 0,$$

که در آن $D = -AH^{-1}A^T$ و $d = -b - AH^{-1}c$. و با استفاده از روش اویلر^۷ برای حل شبکه عصبی فوق با انتخاب $n = 20$ ، $dt = 0.2$ و جواب بهینه مسأله DQP به شکل (۳.۳) به دست می‌آید:

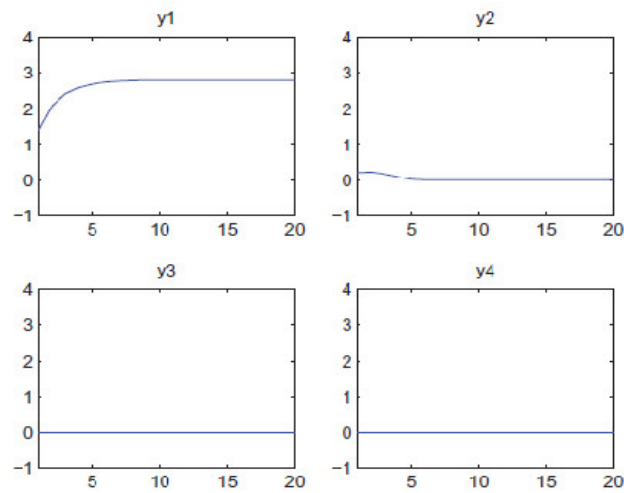
$$y^* = [2.8, 0, 0, 0]^T$$

با جای‌گذاری y^* در (۱۱.۳) جواب بهینه مسأله اصلی به صورت زیر به دست می‌آید:

$$x^* = [0.8, 1.2]^T.$$

برای مسأله QP با دو متغیر و چهار ثابت، مدل جدید پس از حدود ۲۰ تکرار همگراست. در مقایسه، بیشتر از ۲۰۰ تکرار برای حل همان مسأله از مدل عفتی بایمانی طول می‌کشد (جدول (۳.۳) را مشاهده کنید).

^۷Euler method



شکل ۳.۳: مسیره‌های گذرای $y(t)$ از متغیرهای مسأله دوگان در مثال ۲.۶.۳

مدل عفتی بایمانی	مدل ارائه شده	
۲۰۰	۲۰	تعداد تکرارها
$x^* = [0.8, 1.2]^T$	$x^* = [0.8, 1.2, 0, 0.0022]^T$	جواب QP
$y^* = [2.8, 0]^T$	$y^* = [2.8, 0, 0, 0]^T$	جواب DQP
۰.۰۰۴۰	۰.۰۰۰۶	زمان (s)

جدول ۳.۳: نقطه اولیه $y_0 = [0, 0, 0, 0]^T$

فصل ۴

یک مدل شبکه عصبی لاگرانژی برای حل مسائل برنامه ریزی درجه دوم کلی

۱.۴ مقدمه

در این فصل توانایی یک شبکه عصبی برای حل مسائل برنامه‌نویسی درجه دوم محدب با محدودیتهای خطی کلی بیان می‌شود. مدل شبکه عصبی ارائه شده به مفهوم لیاپانف پایدار است و به جواب بهینه دقیق مساله‌ی اصلی همگرا می‌شود. نمودار بلوکی مدل پیشنهادی نیز داده می‌شود. بسیاری از مسائل بهینه‌سازی، ماهیتی متغیر با زمان دارند، بنابراین آنها باید در زمان واقعی حل شوند. یک روش امیدوارکننده برای بررسی مسائل بهینه‌سازی با ابعاد بالا و ساختار پیچیده این است که شبکه عصبی مصنوعی را بر اساس اجرای مداری به‌کار بگیریم. هدف اصلی شبکه عصبی برای بهینه‌سازی این است که تابع انرژی نامنفی بسازد و سیستمی پویا ایجاد کند که شبکه عصبی مصنوعی را ارائه می‌دهد. سیستم پویا معمولاً در شکل معادلات دیفرانسیل معمولی درجه اول است، علاوه بر این انتظار می‌رود که سیستم پویا به حالت تعادل همگرا شود. شبکه عصبی برای حل مسائل برنامه‌نویسی ریاضی اولین بار توسط تنک و هاپ فیلد^۱ [۲۸] مطرح شد. از آن به بعد شبکه‌های عصبی برای حل انواع متفاوتی از مسائل بهینه‌سازی درجه دوم به‌طور گسترده بررسی شده است و برخی از نتایج مهم نیز به‌دست آمده است. ([۱۲] - [۷], [۲۰] - [۱۴], [۲۳] - [۲۲], [۲۷] - [۲۵])

۲.۴ مدل شبکه عصبی

مسأله زیر را در نظر بگیرید:

$$\text{minimize } \frac{1}{2} x^T Q x + D^T x \quad (1.4)$$

^۱Tank and Hopfield

$$\text{subject to} \quad Ax - b \leq 0 \quad (2.4)$$

$$Ex - f = 0 \quad (3.4)$$

که $Q \in \mathbb{R}^{n \times n}$ ماتریس متقارن معین مثبت، $A \in \mathbb{R}^{m \times n}$ ، $b \in \mathbb{R}^m$ ، $E \in \mathbb{R}^{l \times m}$ ، $f \in \mathbb{R}^l$ ، $rank(A, E) = m + l$ و $x \in \mathbb{R}^n$ ، $D \in \mathbb{R}^n$ تابع لاگرانژ زیر را در نظر بگیرید:

$$L(x, u, v) = \frac{1}{2} x^T Q x + D^T x + \frac{1}{2} \sum_{k=1}^m u_k (a_k x - b_k) + \sum_{p=1}^l v_p (e_p x - f_p) \quad (4.4)$$

بر طبق شرایط KKT برای مسئله (۱.۴)-(۳.۴) در [۴]، $x^* \in \mathbb{R}^n$ جواب بهینه (۱.۴)-(۳.۴) است اگر و تنها اگر $u^* \in \mathbb{R}^m$ و $v^* \in \mathbb{R}^l$ وجود داشته باشد به طوری که $(x^{*T}, u^{*T}, v^{*T})^T$ در زیر صدق کند:

$$\begin{cases} u^* \geq 0, Ax^* - b \leq 0, u^{*T}(Ax^* - b) = 0, \\ Qx^* + D + A^T u^* + E^T v^* = 0, \\ Ex^* - f = 0. \end{cases} \quad (5.4)$$

x^* نقطه KKT (۱.۴)-(۳.۴) نامیده می‌شود و $(u^{*T}, v^{*T})^T$ بردار ضریب لاگرانژ مرتبط به x^* است. با توجه به [۴] x^* جواب بهینه (۱.۴)-(۳.۴) است اگر و تنها اگر x^* نقطه KKT (۱.۴)-(۳.۴) باشد. حال فرض می‌کنیم $x(\cdot)$ ، $u(\cdot)$ و $v(\cdot)$ برخی متغیرهای وابسته به زمان باشند. به منظور استفاده از یک شبکه عصبی که مسأله بهینه‌سازی درجه دوم (۱.۴)-(۳.۴) را حل می‌کند، یک سیستم دینامیکی به صورت زیر در نظر بگیرید:

$$\frac{dx}{dt} = -\nabla_x L(x, u, v) = -(Qx + D + \frac{1}{2} \sum_{k=1}^m u_k a_k^T + \sum_{p=1}^l v_p e_p^T), \quad (6.4)$$

$$\frac{du}{dt} = -\nabla_u L(x, u, v) = \text{diag}(u_1, \dots, u_m)(Ax - b), \quad (7.4)$$

$$\frac{dv}{dt} = -\nabla_v L(x, u, v) = Ex - f, \quad (8.4)$$

با نقطه اولیه $(x(t_0)^T, u(t_0)^T, v(t_0)^T)^T$ و $u(t_0) > 0$ فرض کنید $y = (x^T, u^T, v^T)^T \in \mathbb{R}^{m+n+l}$ و D^* به عنوان مجموعه نقاط بهینه (۱.۴)-(۳.۴) و دوگان آن است و فرض کنید

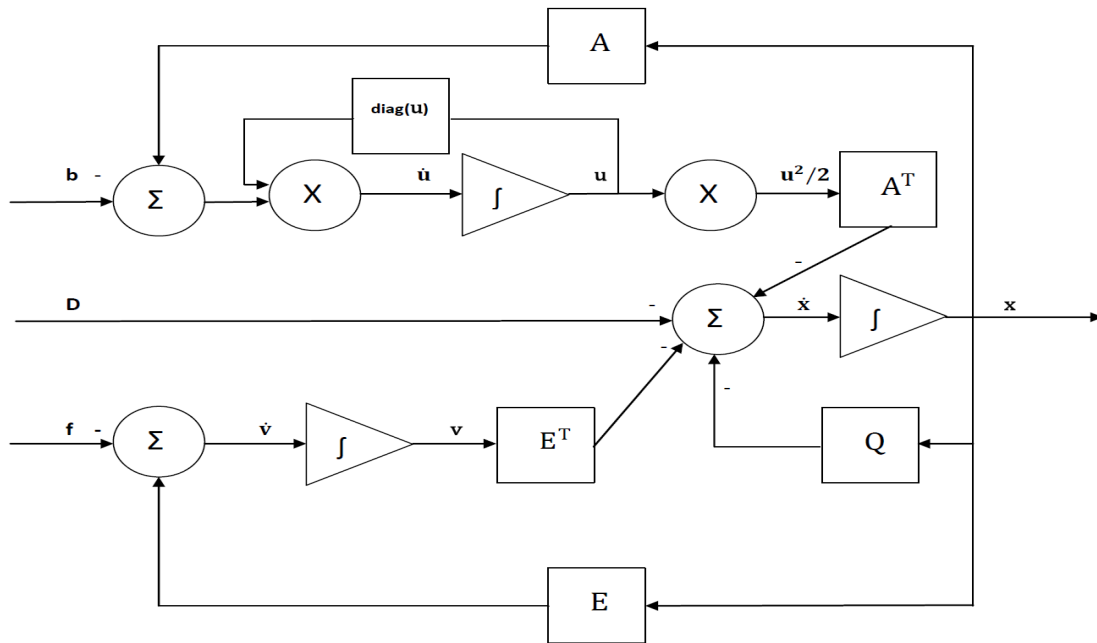
$$F(y) = \begin{bmatrix} -(Qx + D + \frac{1}{2} A^T u + E^T v) \\ \text{diag}(u_1, \dots, u_m)(Ax - b) \\ Ex - f \end{bmatrix}.$$

بنابراین شبکه عصبی (۶.۴)-(۸.۴) به شکل زیر می‌تواند نوشته شود:

$$\frac{dy}{dt} = \eta F(y), \quad (9.4)$$

$$y(t_0) = y_0, \quad u(t_0) > 0, \quad (10.4)$$

که η پارامتر مقیاس است و سرعت همگرایی شبکه عصبی (۹.۴) و (۱۰.۴) را نشان می‌دهد. برای سادگی فرض می‌کنیم $\eta = 1$. اجرای سخت افزاری شبکه عصبی (۹.۴) و (۱۰.۴) در شکل (۱.۴) نشان داده می‌شود.



شکل ۱.۴: نمودار بلوکی ساده شده برای شبکه عصبی (۹.۴) (۱۰.۴)

به منظور بررسی کارایی مدل (۹.۴) و (۱۰.۴) آن را با برخی مدل های موجود مقایسه می‌کنیم. ابتدا مسأله زیر را در نظر می‌گیریم:

$$\text{minimize } \frac{1}{2} x^T Q x + D^T x \quad (11.4)$$

$$\text{subject to } Ax - b \leq 0 \quad (12.4)$$

$$x \geq 0 \quad (13.4)$$

که $Q \in \mathbb{R}^{n \times n}$ ماتریس متقارن معین مثبت است، $A \in \mathbb{R}^{m \times n}$ و $b \in \mathbb{R}^m$ یک مدل شبکه عصبی اولیه-دوگان برای (۱۱.۴)-(۱۳.۴) به صورت زیر در نظر بگیرید:

$$\frac{d}{dt} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} Q + I_n & A^T \\ -A & I_m \end{bmatrix} \times \begin{bmatrix} -x + (x - (Qx + D) + A^T u)^+ \\ u - (u - Ax + b)^+ \end{bmatrix}, \quad (14.4)$$

که I_m و I_n به ترتیب $m \times m$ و $n \times n$ هستند. مدل پیشنهادی شبکه عصبی برای حل (۱۱.۴)-(۱۳.۴) به صورت زیر داده می‌شود [۳۲]:

$$\frac{dx}{dt} = - \left(Qx + D + \frac{1}{2} \left(\sum_{k=1}^m u_k^+ a_k^T - \sum_{k=m+1}^{m+n} u_k^+ e_{(k-m)}^T \right) \right), \quad (15.4)$$

$$\frac{du}{dt} = \text{diag}(u_1, \dots, u_m, u_{m+1}, \dots, u_{m+n}) (\hat{A}x - \hat{b}), \quad (16.4)$$

که $\hat{A} = \begin{bmatrix} A \\ -I_n \end{bmatrix}$ و $\hat{b} = \begin{bmatrix} b \\ 0 \end{bmatrix}$ و $e_i \in \mathbb{R}^{1 \times n}$ بردار واحد استاندارد (مولفه i ام یک است، همه مولفه‌های دیگر صفر است) معادله حالت در (۱۵.۴) و (۱۶.۴) به $3m + 3n + 3mn + n^2$ ضرب و $2n + m + mn + n^2$ جمع/تقسیم نیاز دارد، در حالی که معادله حالت در (۱۴.۴) به $6mn + n^2$ ضرب و $4mn + 2n^2$ جمع/تقسیم در هر تکرار نیاز دارد. بنابراین شبکه عصبی ارائه شده در (۱۵.۴) و (۱۶.۴) پیچیدگی کمتری را برای اجرا نسبت به شبکه موجود در (۱۴.۴) دارد. مسأله زیر را در نظر می‌گیریم:

$$\text{minimize} \quad \frac{1}{2} x^T A_0 x + a^T x \quad (17.4)$$

$$\text{subject to} \quad x \in \Omega. \quad (18.4)$$

$$Dx = b \quad (19.4)$$

که $A_0 \in \mathbb{R}^{m \times m}$ ماتریس متقارن نیمه معین مثبت است، D ماتریس $n \times m$ و $\Omega \subset \mathbb{R}^m$ مجموعه بسته محدب و $x = (x_1, \dots, x_m)^T$ است. در [۲۷] شبکه عصبی برای حل (۱۷.۴)-(۱۹.۴) به صورت زیر است:

$$\frac{dx}{dt} = P. (x - A_0 x + D^T y - a) - x \quad (20.4)$$

$$\frac{dx}{dt} = -DP. (x - A_0 x + D^T y - a) + b \quad (21.4)$$

که در آن P عملگر تصویر در Ω است، مدل شبکه عصبی (۲۰.۴)-(۲۱.۴) در زمان متناهی همگرا نیست. کندی و چوآ [۱۷] یک مدل شبکه عصبی را برای حل مسائل برنامه‌نویسی غیر خطی محدب پیشنهاد کردند. این شبکه برای حل (۱.۴)-(۳.۴) به شکل زیر استفاده می‌شود:

$$\frac{dx}{dt} = - (Qx + D + s (A^T (Ax - b)^+ + E^T (Ex - f))), \quad (22.4)$$

که

$$(Ax - b)^+ = ([a_1 x - b_1]^+, [a_2 x - b_2]^+, \dots, [a_m x - b_m]^+),$$

$$[a_k x - b_k] = \max \{ (a_k x - b_k), 0 \}, \quad k = 1, \dots, m,$$

که در آن s پارامتر جریمه است. این شبکه به دلیل پارامتر جریمه مشخص قادر نیست جواب بهینه دقیق را بیابد و زمانی که پارامتر جریمه خیلی بزرگ باشد اجرای آن سخت است. بنابراین شبکه (۲۲.۴) تنها به جواب تقریبی (۱.۴)-(۳.۴) برای هر پارامتر جریمه متناهی همگراست. در مقایسه، شبکه عصبی ارائه شده در (۹.۴)-(۱۰.۴) به‌طور سراسری به جواب بهینه (۱.۴)-(۳.۴) همگراست.

قضیه ۱.۲.۴. الف. فرض کنید $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ با $u^* \geq 0$ یک نقطه تعادل شبکه عصبی (۹.۴) و (۱۰.۴) باشد آن‌گاه x^* نقطه KKT مسأله (۱.۴)-(۳.۴) است.

ب. فرض کنید $y(t) = (x(t)^T, u(t)^T, v(t)^T)^T$ منحنی مسیر (۹.۴) و (۱۰.۴) با نقطه اولیه $y(t_0) = (x(t_0)^T, u(t_0)^T, v(t_0)^T)^T$ اگر $u(t_0) > 0$ باشد. اگر $u(t) \geq 0$.

برهان. الف. با مقایسه سیستم (۵.۴) با معادله پویای عصبی (۹.۴) و (۱۰.۴) باقی می ماند ثابت کنیم

$$u_k^*(a_k x^* - b_k) = 0, \quad a_k x^* - b_k \leq 0, \quad (k = 1, \dots, m).$$

جواب تحلیلی معادله (۷.۴) به صورت زیر است:

$$u_k(t) = u_k(t_0) e^{\int_{t_0}^t (a_k x - b) dt}, \quad (k = 1, \dots, m). \quad (۲۳.۴)$$

که $u_k(t_0)$ مقدار اولیه $u_k(t)$ است. این عبارت نشان می دهد که اگر قیدهای نامساوی برآورده نشوند (یعنی حداقل یک k وجود دارد که $a_k x - b_k > 0$) عبارت $\frac{1}{\tau} u_k^*(a_k x - b_k)$ در تابع لاگرانژ (۴.۴) با زمان افزایش می یابد چون هدف مینیم سازی $L(x, u, v)$ است، پس از مدت زمان محدود $x(t)$ به سمت ناحیه شدنی حرکت می کند. بنابراین به ازای $t \geq t_0$ ، $t \rightarrow +\infty$ ، $u_k(t) \rightarrow 0$. در نتیجه $u_k^*(a_k x^* - b_k) = 0$.

ب. برای نقطه اولیه داده شده $y(t_0)$ با $u(t_0) > 0$ و از (۲۳.۴)، چون

$$e^{\int_{t_0}^t (a_k x - b_k) dt} \geq 0, \quad (k = 1, \dots, m)$$

در نتیجه برای هر $t \geq t_0$ ، $u_k(t) \geq 0$.

لم ۲.۲.۴. برای هر نقطه اولیه $y(t_0) = (x(t_0)^T, u(t_0)^T, v(t_0)^T)^T$ جواب پیوسته یکتا $y(t)$ برای سیستم (۹.۴) و (۱۰.۴) وجود دارد.

برهان. $Qx + D + \frac{1}{\tau} A^T u^* + E^T v$ و $\text{diag}(u_1, \dots, u_m)(Ax - b)$ و $Ex - f$ پیوسته لیپ شیتز موضعی است. با توجه به قضیه وجود و یکتایی جواب معادلات دیفرانسیل معمولی، شبکه عصبی (۹.۴) و (۱۰.۴) جواب پیوسته منحصر بفرد $y(t)$ ، $t \in [t_0, \tau)$ را برای برخی از $\tau > t_0$ دارد.

قضیه ۳.۲.۴. اگر Q معین مثبت باشد آن‌گاه شبکه عصبی پیشنهادی در (۹.۴) و (۱۰.۴) به مفهوم لیاپانف پایدار است و به طور سراسری به $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ همگراست که در آن y^* نقطه تعادل (۹.۴) و (۱۰.۴) است.

برهان. تابع لیاپانف زیر را تعریف کنید:

$$E(y) = \|F(y)\|^2 + \frac{1}{\tau} \|y - y^*\|^2 \quad (۲۴.۴)$$

آن‌گاه

$$\frac{dF}{dt} = \frac{\partial F}{\partial y} \frac{dy}{dt} = \nabla F(y) F(y)$$

که ماتریس ژاکوبین نگاشت F به صورت زیر است:

$$\nabla F(y) = \begin{bmatrix} -Q & -A^T S & -E^T \\ S^T A & K & Q_{m \times 1} \\ E & Q_{l \times m} & Q_{l \times l} \end{bmatrix}, \quad (25.4)$$

$$K = \text{diag}(a_1 x - b_1, \dots, a_m x - b_m), \quad S = \text{diag}(u_1, \dots, u_m)$$

داریم:

$$\begin{aligned} \frac{dE(y(t))}{dt} &= \left(\frac{dF}{dt}\right)^T F + F^T \left(\frac{dF}{dt}\right) + (y - y^*) \frac{dy(t)}{dt} \\ &= F^T (\nabla F(y))^T + \nabla F(y) F + (y - y^*)^T F(y) \\ &= -\gamma (Qx + D + \frac{1}{\gamma} A^T u^\gamma + E^T v)^T Q (Qx + D + \frac{1}{\gamma} A^T u^\gamma + E^T v) \\ &\quad + \gamma \sum_{k=1}^m (u_k (a_k x - b_k))^\gamma (a_k x - b_k) + (y - y^*)^T F(y) \\ &\leq \gamma \sum_{k=1}^m (u_k (a_k x - b_k))^\gamma (a_k x - b_k) + (y - y^*)^T F(y) \end{aligned}$$

حال اگر $(a_k x - b_k) \leq 0, (k = 1, \dots, m)$ آن‌گاه $\sum_{k=1}^m (u_k (a_k x - b_k))^\gamma (a_k x - b_k) \leq 0$ علاوه بر این $\nabla F(y)$ معین منفی است و $F(y)$ اکیدا یکنواخت است. بنابراین $(y - y^*)^T (F(y) - F(y^*)) = (y - y^*)^T F(y) \leq 0$,

در نتیجه

$$\left(\frac{dE(y(t))}{dt}\right) \leq 0. \quad (26.4)$$

از طرفی، اگر $(a_k x - b_k) > 0$ برای برخی از $k \in \{1, \dots, m\}$. آن‌گاه بعد از مدت زمان متناهی $x(t)$ نهایتاً در ناحیه شدنی قرار می‌گیرد. در نتیجه بعد از یک زمان متناهی $\frac{dE(y(t))}{dt} \leq 0$ را به دست می‌آوریم. بنابراین شبکه عصبی (۹.۴) و (۱۰.۴) به مفهوم لیاپانف پایدار است. چون

$$E(y) \geq \frac{1}{\gamma} \|y - y^*\|^\gamma, \quad (27.4)$$

برای هر نقطه اولیه $y_0 = (x(t_0)^T, u(t_0)^T, v(t_0)^T) \in R^{n+m+l}$ با $y_0 > 0$ جواب $y(t)$ در $[t_0, \tau)$ کراندار می‌شود بنابراین $\tau \rightarrow +\infty$ و زیر دنباله‌ای همگرا به صورت زیر وجود دارد.

$$\left\{ (x(t_k)^T, u(t_k)^T, v(t_k)^T) \Big| t_0 < t_1 < \dots < t_k < t_{k+1}, \quad t_k \rightarrow \infty \text{ و } k \rightarrow \infty \right\}$$

آن چنان‌که:

$$\lim_{k \rightarrow \infty} (x(t_k)^T, u(t_k)^T, v(t_k)^T)^T = (\bar{x}^T, \bar{u}^T, \bar{v}^T)^T$$

که در آن $(\bar{x}^T, \bar{u}^T, \bar{v}^T)^T$ در \circ صدق می‌کند. بنابراین $(\bar{x}^T, \bar{u}^T, \bar{v}^T)^T$ یک نقطه ω حدی از مجموعه $\left\{ (x(t)^T, u(t)^T, v(t)^T)^T \mid t \geq t_0 \right\}$ است. با استفاده از قضیه مجموعه تغییرناپذیری لازال $\{ (x(t)^T, u(t)^T, v(t)^T)^T \} \rightarrow M$ که M بزرگترین مجموعه تغییرناپذیر در

$$K = \left\{ (x(t)^T, u(t)^T, v(t)^T)^T \mid \frac{dE(y(t))}{dt} = \circ \right\}$$

است. از (۹.۴)، (۱۰.۴) و (۲۶.۴) نتیجه می‌گیریم که

$$\frac{dx}{dt} = \frac{du}{dt} = \frac{dv}{dt} = \circ \iff \frac{dE(y(t))}{dt} = \circ$$

بنابراین $(\bar{x}^T, \bar{u}^T, \bar{v}^T)^T \in D^*$ با $M \subseteq K \subseteq D^*$ لذا شبکه عصبی ارائه شده در (۹.۴) و (۱۰.۴) به‌طور سراسری به نقطه تعادل \bar{y} همگراست. \square

نتیجه ۴.۲.۴. فرض کنید که شرایط قضیه (۳.۲.۴) برقرار باشند، اگر $D^* = \{ (x^{*T}, u^{*T}, v^{*T})^T \}$ آن‌گاه شبکه عصبی (۹.۴) و (۱۰.۴) برای حل (۱.۴)–(۳.۴) به‌طور پایدار سراسری مجانبی به نقطه تعادل منحصر به فرد $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ است.

۳.۴ مثال‌های محاسباتی

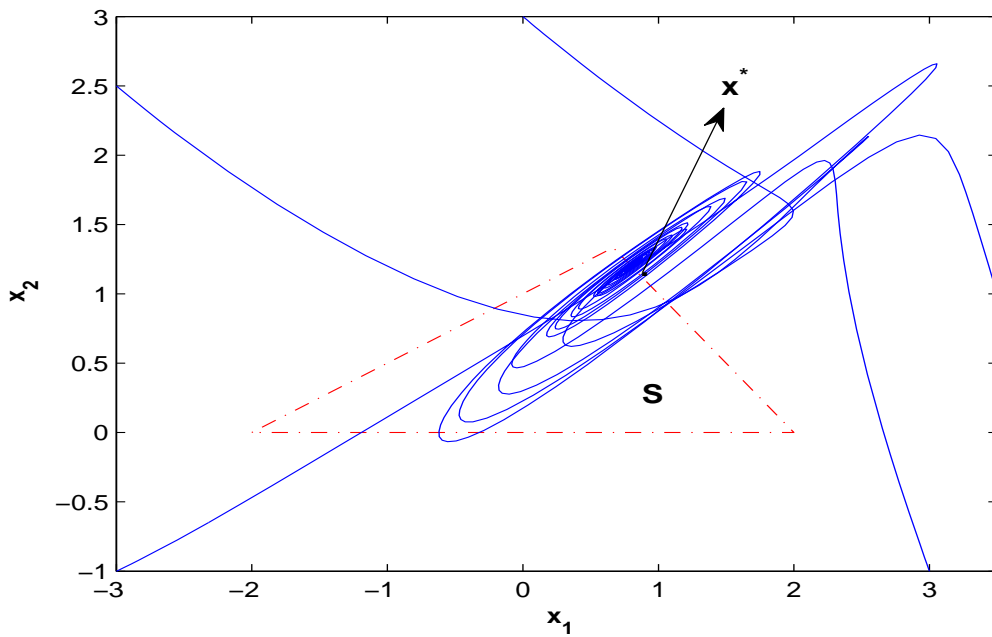
برای نشان دادن قابلیت شبکه عصبی ارائه شده (۹.۴) و (۱۰.۴) چندین مثال را آزمایش می‌کنیم. برای آزمایش برخی از مسائل، عملکرد شبکه عصبی ارائه شده (۹.۴) و (۱۰.۴) را با مقادیر مختلف η و حالت‌های اولیه مختلف $x(0)$ مقایسه می‌کنیم. همچنین دو مثال کاربردی در مهندسی را به‌عنوان رگرسیون و مسأله حداقل مربعات تقریبی محدود شده را ارائه می‌دهیم.

مثال ۱.۳.۴. [۲۶]

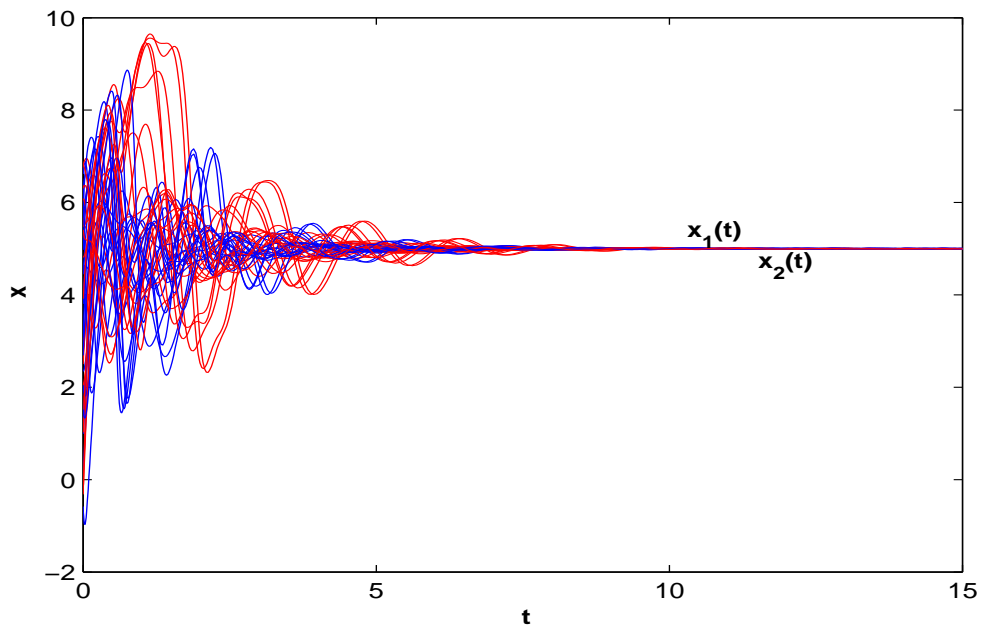
$$\text{minimize } x_1 + 2x_2 - 2x_1x_2 - 2x_1 - 6x_2$$

$$\text{subject to } \begin{cases} x_1 + x_2 \leq 2, \\ -x_1 + 2x_2 \leq 2, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

این مسأله تنها یک جواب $x^* = [0.8, 1.2]^T$ دارد. قضیه (۳.۲.۴) و نتیجه (۴.۲.۴) بیان می‌کند که مدل بیان شده در (۹.۴) و (۱۰.۴) به‌طور سراسری به x^* همگراست. شکل (۲.۴) نمودار متغیرهای حالت $(x_1(t), x_2(t))^T$ را با پنج نقطه اولیه مختلف به تصویر می‌کشد، که S ناحیه شدنی است. این بردارها به‌طور سراسری به جواب دقیق x^* همگراست.



شکل ۲.۴: دیاگرام صوری شبکه عصبی (۹.۴) و (۱۰.۴) با پنج نقطه اولیه متفاوت در مثال ۱.۳.۴



شکل ۳.۴: رفتارهای گذرای $x(t)$ از شبکه عصبی (۹.۴) و (۱۰.۴) با ۲۰ نقطه اولیه متفاوت در مثال ۲.۳.۴

مثال ۲.۳.۴. [۳۴]

$$\begin{aligned} &\text{minimize } x_1^2 + x_2^2 + x_1 x_2 - 3x_1 - 3x_2 \\ &\text{subject to } \begin{cases} \frac{5}{12}x_1 - x_2 \leq \frac{35}{12}, \\ \frac{5}{4}x_1 + x_2 \leq \frac{35}{4}, \\ -5 < x_1 \leq 0, \quad x_2 \geq 5. \end{cases} \end{aligned}$$

جواب بهینه برای این مسأله $x^* = [5, 5]^T$ است. شبکه عصبی ارائه شده در (۹.۴) و (۱۰.۴) را برای حل این برنامه ریزی ریاضی به کار می‌بریم. نتایج مثال‌های محاسباتی نشان می‌دهد که خروجی (۹.۴) و (۱۰.۴) با هر نقطه اولیه همیشه به $y^* = [x^{*T}, u^{*T}]^T$ همگراست. به طور مثال شکل (۳.۴) رفتار $x(t)$ را بر اساس (۹.۴) و (۱۰.۴) با 20 نقطه تصادفی نشان می‌دهد. رفتار همگرایی l_2 نرم خطای x و x^* نیز در شکل (۴.۴) نشان داده می‌شود. برای مقایسه، مسأله بالا در ما ۲ و شنبلیت ۳ [۲۲]، با استفاده از شبکه کندی-چوآ حل می‌شود. این شبکه با $s = 50$ به نقطه تعادل $\tilde{x} = [4, 97778, 5, 1745]^T$ همگراست، که می‌تواند به عنوان جواب تقریبی مسأله بالا دیده شود و این نقطه تعادل شدنی نیست.

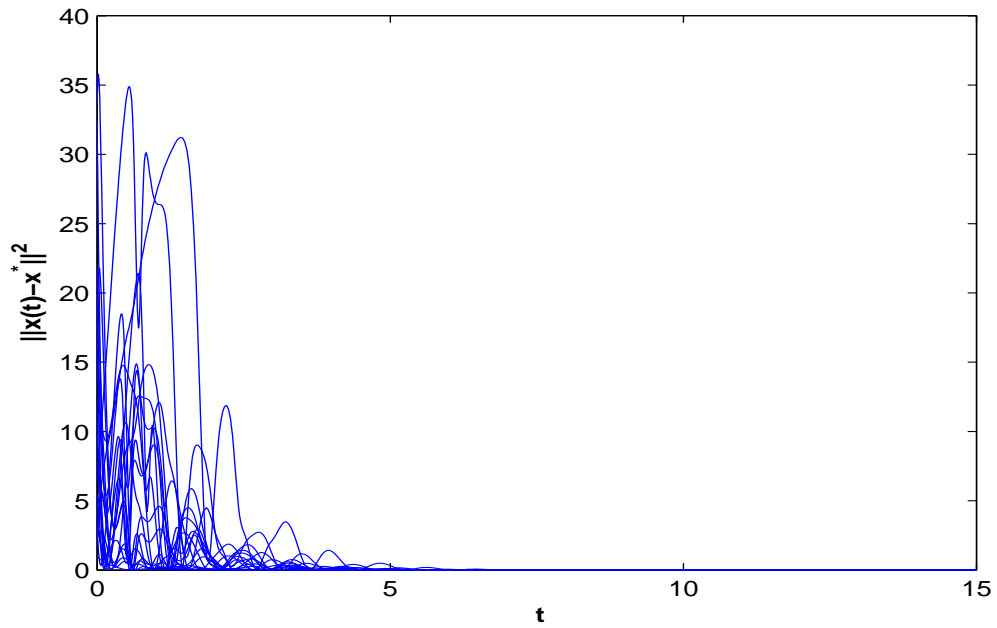
مثال ۳.۳.۴. [۲۶]

$$\begin{aligned} &\text{minimize } x_1^2 + x_2^2 + 0.5x_3^2 + x_1 x_2 - 4x_1 - 3x_2 - 2x_3 \\ &\text{subject to } \begin{cases} x_1 + x_2 + 2x_3 \leq 3, \\ 3x_1 - 9x_2 + 9x_3 = 1, \\ 0 \leq x_i \leq \frac{4}{3} \quad (i = 1, 2, 3). \end{cases} \end{aligned}$$

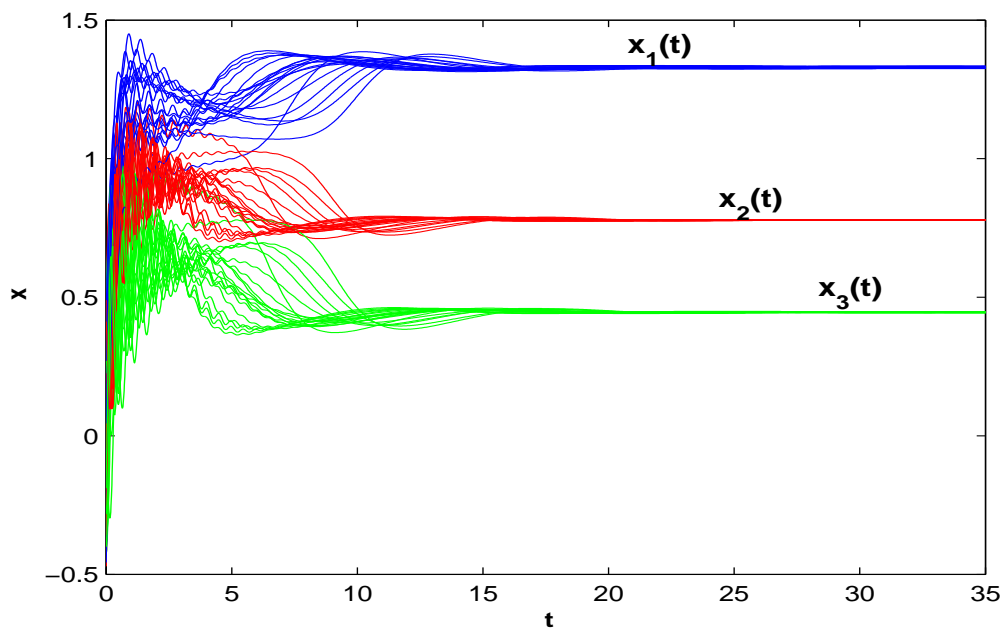
جواب بهینه برای این مسأله $x^* = [\frac{4}{3}, \frac{7}{9}, \frac{4}{9}]^T$ است. شکل (۵.۴) نشان می‌دهد که مسیرهای شبکه عصبی ارائه شده در (۹.۴) و (۱۰.۴) که مسأله بالا را با 20 نقطه تصادفی اولیه حل کرد، به جواب بهینه این مسأله همگراست. تاثیر پارامتر η را در شبکه عصبی (۹.۴) و (۱۰.۴) بر روی مقدار $\|x(t) - x^*\|^2$ آزمایش می‌کنیم. از شکل (۶.۴) واضح است که با $\eta = 0.1$ ، سرعت همگرایی $\|x(t) - x^*\|^2$ کمتر از زمانی است که $\eta = 5$ اختیار می‌شود.

^۲Maa

^۳Shanblatt



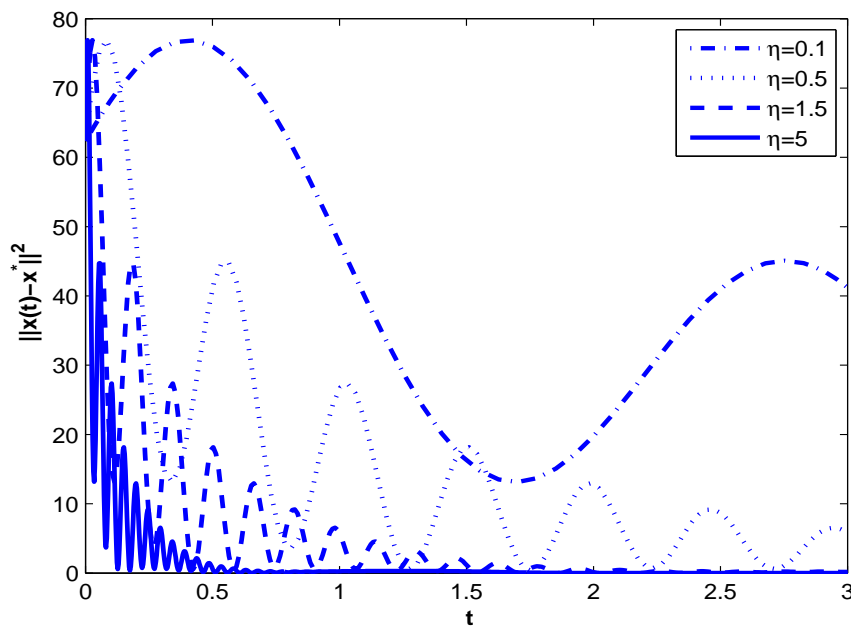
شکل ۴.۴: رفتار همگرایی $\|x(t) - x^*\|^2$ در مثال ۲.۳.۴



شکل ۵.۴: رفتارهای گذرا از شبکه عصبی (۹.۴) و (۱۰.۴) با ۲۰ نقطه اولیه متفاوت

مثال ۴.۳.۴. [۳۰]

$$\text{minimize } 1 \cdot x_1^2 + 2x_2^2 + 2x_3^2 - 2(x_1x_2 + 3x_1x_3 - x_2x_3)$$



شکل ۶.۴: رفتار همگرایی $\|x(t) - x^*\|^2$ در مثال ۳.۳.۴ با $y_0 = [-5, 4, -3, 2, -1, 0, 1, -2, 3, -4, 5]^T$

$$\text{subject to } \begin{cases} -1 \leq x_2 - x_1 \leq 0, \\ -1 \leq x_3 - 3x_1 \leq 1, \\ 1 \leq x_2 + x_3 \leq 2. \end{cases}$$

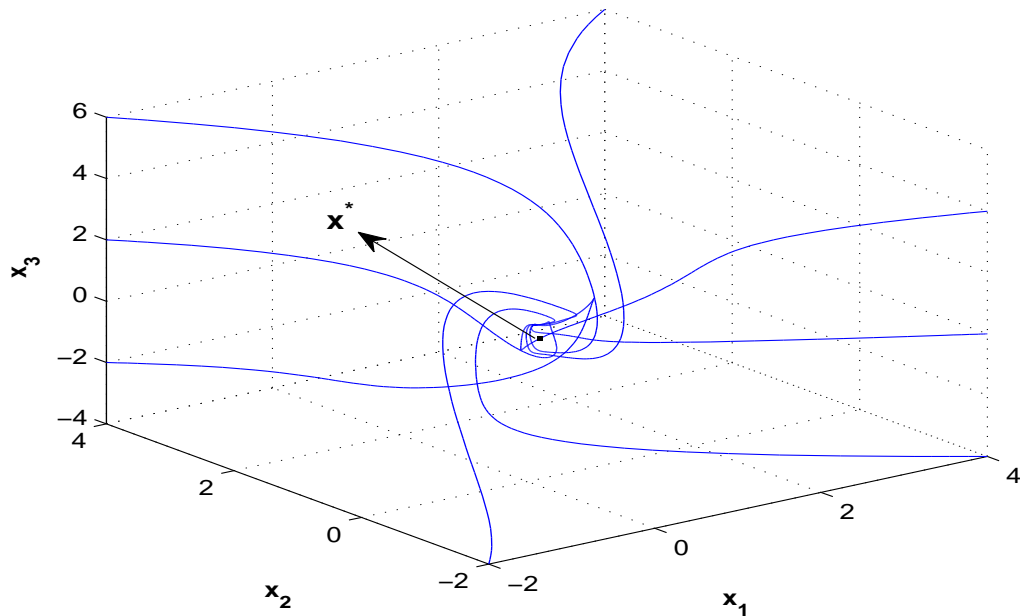
نتایج مثال‌های محاسباتی نشان می‌دهد که خط مسیر خروجی $x(t)$ مدل ارائه شده به جواب بهینه $x^* = [\frac{1}{4}, \frac{1}{4}, \frac{3}{4}]^T$ (مثلاً شکل (۷.۴) همگرایی با هشت نقطه اولیه تصادفی را نشان می‌دهد. علاوه بر این، وقتی نقطه اولیه به‌عنوان نقطه نشدنی انتخاب شود، مسیرهای شبکه هم‌چنان به x^* همگراست.

مثال ۵.۳.۴. [۳۵]

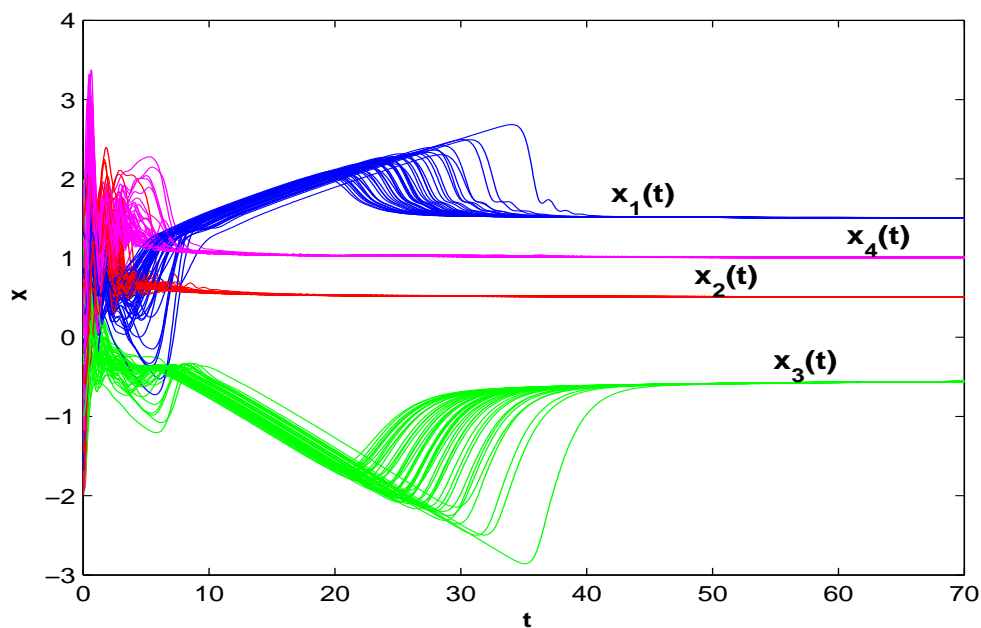
$$\text{minimize } 10x_1^2 - x_2^2 + x_3^2 + x_4^2 - 2x_1 - x_4$$

$$\text{subject to } \begin{cases} 2x_1 + x_2 + x_3 + 4x_4 = 7, \\ 0.2x_1 + 2x_2 + 2x_4 = 6, \\ x_1 \leq 1.5, \quad +x_2 \leq 0.5, \\ x_3 \leq 1.5, \quad x_4 \leq 1. \end{cases}$$

این مسأله یک جواب بهینه $x^* = [1/5, 0/5, -0/5, 1]^T$ دارد. شکل (۸.۴) رفتار همگرایی $x(t)$ را بر اساس شبکه عصبی ارائه شده با ۵۰ نقطه اولیه تصادفی نشان می‌دهد. تأکید می‌کنیم که در هر صورت



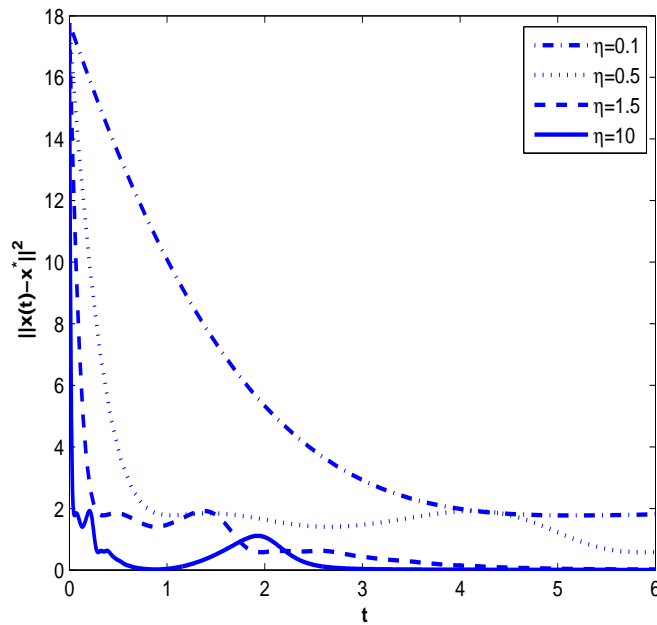
شکل ۷.۴: نمودار شبکه عصبی با هشت نقطه اولیه در (۹.۴) و (۱۰.۴) مثال ۴.۳.۴.



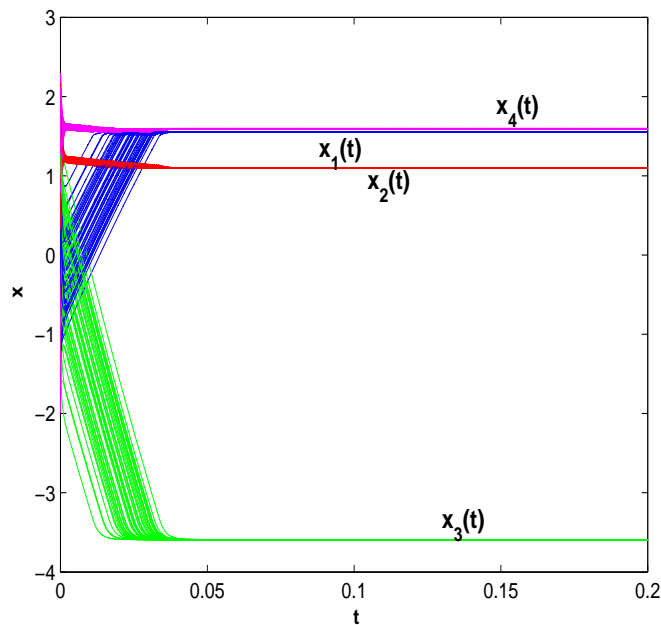
شکل ۸.۴: رفتار گذرا از شبکه عصبی (۹.۴) و (۱۰.۴) با ۵۰ نقطه تصادفی در مثال ۵.۳.۴.

اگر نقطه اولیه داخل یا خارج فضای شدنی گرفته می‌شود، شبکه ارائه شده همیشه به جواب بهینه x^* همگراست. شکل (۹.۴) نیز رفتار همگرای خطای $\|x(t) - x^*\|^2$ را با η های مختلف و نقطه اولیه

$$y_0 = [-2, 2, -2, 2, -2, 2, -2, 2, -2, 2, -2, 2, -2, 2]^T$$



شکل ۹.۴: رفتار همگرایی $\|x(t) - x^*\|^2$ در مثال ۵.۳.۴ با $y = [-2, 2, -2, 2, -2, 2, -2, 2, -2, 2]^T$



شکل ۱۰.۴: رفتار گذرای $x(t)$ از شبکه عصبی کندی چوآ (۲۲.۴) با ۵۰ نقطه اولیه متفاوت در مثال ۵.۳.۴

۱۲٫۷	۱۱٫۵	۱۱٫۰	۱۰٫۲	۷٫۸	۵٫۶	۴	۳٫۰	۱٫۰	x
-۲٫۰	۴٫۰	۱۲٫۸	۱۷٫۶	۲۰	۸٫۰	-۲٫۴	-۲٫۸	-۳٫۲	y

جدول ۱.۴: داده های رگرسیونی

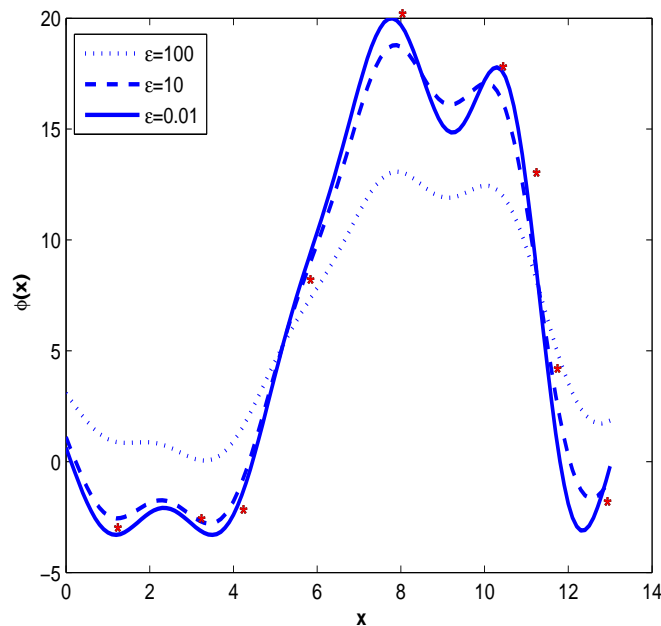
نشان می دهد. بدیهی است که η بزرگتر منجر به میزان همگرایی بهتر می شود. برای مقایسه، شبکه کندی-چوا در (۲۲.۴) را برای حل این مثال به کار می بریم. شکل (۱۰.۴) رفتار گذرا را که بر اساس مدل کندی-چوا در (۲۲.۴) است را با ۵۰ نقطه اولیه تصادفی و $s = 150000$ نشان می دهد. از این شکل واضح است که مسیرهای مدل کندی-چوا (۲۲.۴) به x^* همگرا نیستند.

مثال ۶.۳.۴. مسأله رگرسیون تقریبی از مجموعه داده ها را در نظر بگیرید:

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}, \quad (28.4)$$

فرض کنید تابع رگرسیونی به شکل زیر باشد:

$$\phi(x) = \sum_{i=1}^N \alpha_i \phi_i(x) + \beta \quad (29.4)$$



شکل ۱۱.۴: نتایج بردار پشتیبانی رگرسیون با استفاده از شبکه عصبی پیشنهادی (۹.۴) و (۱۰.۴) با یک کرنل RBF که $\xi = 100$ و $\delta = 1$ و سه پارامتر متفاوت ε .

و α_i و β پارامترهای این مدل هستند که تخمین زده می شوند. در اینجا می توانیم شبکه عصبی در (۹.۴) و (۱۰.۴) برای تخمین این پارامترها به کار ببریم با استفاده از تابع کاهش هابر^۴ تابع رگرسیونی

^۴Huber

۴/۵	۴	۳/۵	۳	۲/۵	۲	۱/۵	۱	۰/۵	۰	x
۵/۸	۱/۲	۰/۸	۳	۶/۲	۶/۲	۸	۷/۹	۷/۲	۷/۶	y

جدول ۲.۴: داده‌های تقریبی

که در (۲۹.۴) تعریف شد، به شکل زیر تبدیل می‌شود:

$$\phi(x) = \sum_{i=1}^N \theta_i K(x, x_i) + \beta,$$

که $K(x, y)$ تابع کرنل است و

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j).$$

برطبق [۲۶] $\theta_i (i = 1, \dots, N)$ ، جواب بهینه مسأله بهینه‌سازی درجه دوم به شکل زیر است:

$$\begin{aligned} \text{minimize } & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \theta_i \theta_j K(x_i, x_j) - \sum_{i=1}^N \theta_i \gamma_i + \frac{\varepsilon}{2\xi} \sum_{i=1}^N \theta_i^2 \\ \text{subject to } & \begin{cases} \sum_{i=1}^N \theta_i = 0, \\ -\xi \leq \theta_i \leq \xi, \quad (i = 1, \dots, N), \end{cases} \end{aligned}$$

که در آن $\varepsilon > 0$ پارامتر دقیق مورد نیاز برای تقریب و $\xi > 0$ یک پارامتر مشخص است. همچنین نشان داده می‌شود که $\beta = v^*$ که v^* نقطه تعادل مدل شبکه عصبی ارائه شده برای حل مسأله بهینه‌سازی درجه دوم است. به عنوان مثال داده‌های رگرسیون در جدول (۱۰.۴) را در نظر می‌گیریم. شبکه عصبی ارائه شده (۹.۴) و (۱۰.۴) با یک کرنل RBF گوسین زیر در نظر می‌گیریم:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\delta^2}\right)$$

شکل (۱۱.۴) نتیجه رگرسیون را برای $\xi = 100$ و $\delta = 1$ و سه پارامتر مختلف ε نشان می‌دهد واضح است که وقتی ε به صفر میل می‌کند تقریب دقیق‌تر است.

مثال ۷.۳.۴. مسأله تقریب حداقل-مربعات محدود را در نظر بگیرید. پارامترهای $(a_0, a_1, a_2, a_3, a_4)^T$ را برای تابع:

$$y(x) = a_4 e^x + a_3 x^3 + a_2 x^2 + a_1 x + a_0$$

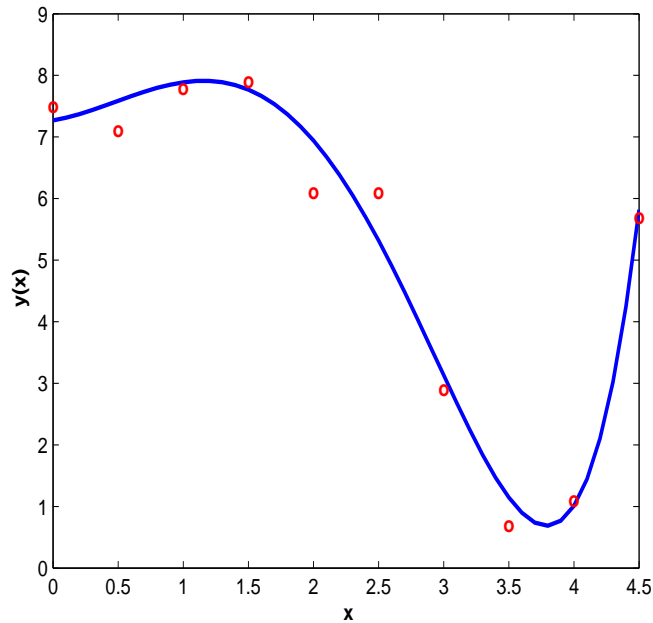
که متناسب با داده‌های مفروض در جدول (۲.۴) است و با توجه به محدودیت‌های زیر بیابید:

$$8.1 \leq y(1/3) \leq 8.3, \quad 2.25 \leq y(4/2) \leq 2.26, \quad 3.4 \leq y(2/8) \leq 3.5$$

این مسأله می‌تواند به شکل زیر فرموله شود:

$$\text{minimize } \|Cx - d\|^2$$

$$\text{subject to } Ax \in \Omega,$$



شکل ۱۲.۴: نتایج تقریب شبکه عصبی (۹.۴) و (۱۰.۴) در مثال ۵.۳.۴. که دایره معرف داده های تقریب و خم تابع تقریب است.

که در آن

$$x = (x_1, x_2, x_3, x_4, x_5)^T = (a_4, a_3, a_2, a_1, a_0)^T,$$

$$C = \begin{bmatrix} 1 & 1/649 & 2/718 & 4/482 & 7/389 & 12/183 & 33/116 & 54/598 & 90/017 \\ 0 & 0/125 & 1 & 3/375 & 8 & 15/625 & 42/875 & 64 & 91/125 \\ 0 & 0/25 & 1 & 2/25 & 4 & 6/25 & 12/25 & 16 & 20/25 \\ 0 & 0/5 & 1 & 1/5 & 2 & 2/5 & 3/5 & 4 & 4/5 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T,$$

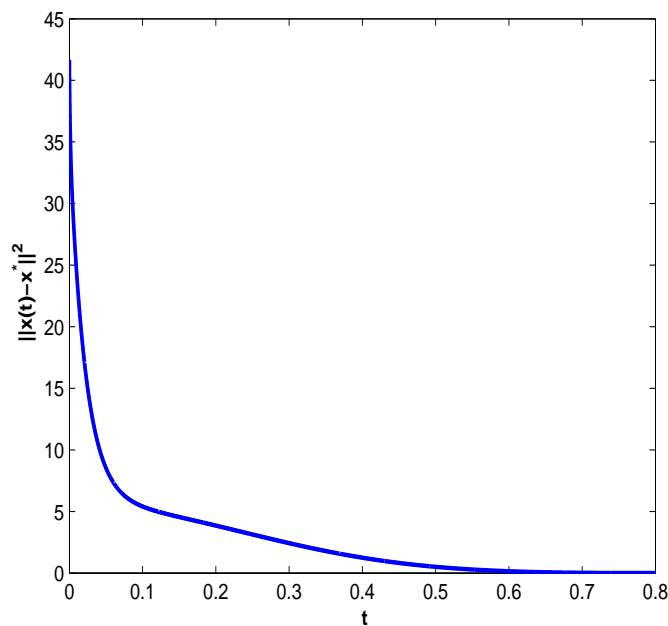
$$d = [7/2, 7/2, 7/9, 8, 6/2, 0/8, 1/2, 5/8]^T,$$

$$A = \begin{bmatrix} 3/669 & 2/192 & 1/69 & 1/3 & 1 \\ 16/449 & 21/952 & 7/84 & 2/8 & 1 \\ 96/686 & 74/088 & 17/64 & 4/2 & 1 \end{bmatrix},$$

و

$$\Omega = \{x \in \mathbb{R}^3 : 1/1 \leq x_1 \leq 1/3, 3/4 \leq x_2 \leq 3/5, 2/25 \leq x_3 \leq 2/26\}$$

شکل (۱۲.۴) نتایج تقریب را با استفاده از شبکه عصبی ارائه شده نمایش می‌دهد. خطای نرم l_2 بین x و $x^* = [1, -1, 1, -1, 1]^T$ نیز در شکل (۱۳.۴) نشان داده می‌شود.



شکل ۳.۴: رفتار همگرایی $\|x(t) - x^*\|^2$ با $x_0 = [1, -1, 1, -1]^T$ و $\eta = 10^{-5}$ در مثال ۵.۳.۴.

برای پایان دادن به این بخش، به سوال زیر پاسخ می‌دهیم.

آیا مزایای شبکه عصبی ارائه شده در مقایسه با مدل‌های موجود وجود دارد؟ برای جواب دادن به این سوال نتایج به دست آمده از آزمایش‌های عددی و مباحث نظری ارائه شده را خلاصه می‌کنیم: در مقایسه با الگوریتم‌های بهینه‌سازی عددی سنتی، روش شبکه عصبی چند مزیت بالقوه در مسائل زمان واقعی دارد:

- ساختار شبکه عصبی می‌تواند به‌طور موثر با استفاده از یکپارچه‌سازی در مقیاس بسیار بزرگ و فن‌آوری‌های نوین اجرا شود. دوم، شبکه‌های عصبی می‌توانند بسیاری از مسائل بهینه‌سازی را با پارامترهای زمان-متغیر حل کنند. سوم، روش‌های پویا و روش‌های ODE می‌توانند مستقیماً شبکه عصبی زمان-پیوسته را برای حل مسائل به‌طور موثر پیاده‌سازی کنند.
- مدل شبکه عصبی ارائه شده را با برخی مدل‌های موجود مقایسه می‌کنیم که برای مسایل بهینه‌سازی درجه دوم نیز کار می‌کند، مثلاً مدل‌هایی که در (۲۰.۴)، (۲۱.۴) و (۲۲.۴) استفاده می‌شود. در نگاه اول، به نظر می‌رسد که این مدل‌های شبکه عصبی پیچیدگی‌های کمتری داشته باشند. با این حال می‌بینیم که تفاوت عملکرد مدل ارائه شده با آزمایش برخی مسایل بهینه‌سازی درجه دوم بسیار واضح است.
- تغییر نقاط اولیه اثر زیادی بر مدل شبکه عصبی ارائه شده ندارد، در حالی که برای برخی مدل‌های موجود این تأثیر وجود دارد. به این دلیل است که مدل ارائه شده به‌طور سراسری به جواب بهینه مسأله همگراست.

• دو مثال برای رگرسیون و مسأله تقریب حداقل مربعات محدود برای نشان دادن شبکه عصبی ارائه شده بیان شده است.

در این مقاله تاثیر پارامتر η را در مدل پویای (۹.۴) و (۱۰.۴) درباره میزان همگرایی مسیرها و رفتار همگرایی $\|x(t) - x^*\|^2$ نیز تجزیه و تحلیل کردیم. واضح است که η بزرگتر منجر به میزان همگرایی بهتری می‌شود. نتایج شبیه‌سازی در چندین مثال، رفتار همگرایی و جذابیت شبکه عصبی ارائه شده را به وضوح نشان می‌دهد.

نتیجه‌گیری و پیشنهادات برای کارهای آتی

در این پایان نامه ابتدا مروری بر شبکه‌های عصبی داشته و با استفاده از دو مدل شبکه عصبی کارا به حل مسائل برنامه ریزی درجه دوم محدب پرداختیم. هم‌چنین پایداری و همگرایی و سراسری به جواب بهینه را مورد بررسی قرار دادیم. از جمله مزیت‌های مدل ارائه شده می‌توان به موارد زیر اشاره کرد:

- حل مسأله ربات چند انگشتی
- ارائه یک مدل دوگان برای حل مسائل QP با استفاده از مدل فصل ۴
- کاربرد مدل ارائه شده از فصل ۴ در مسائل پردازش تصویر
- حل مسأله رگرسیون و طبقه‌بندی داده‌ها با توسعه دادن مدل فصل ۳
- حل مسائل بهینه‌سازی سبد سهام
- این مدل نیاز به پارامتر جریمه ندارد
- با وجود سادگی این مدل، در رده بسیاری از مسائل بهینه‌سازی محدب قابل استفاده است
- این مدل هیچ‌گونه وابستگی به نقطه شروع نداشته و این نقطه می‌تواند خارج از ناحیه شدنی و نقطه‌ای تصادفی باشد
- این مدل می‌تواند قادر به حل مسائل وابسته به زمان و مسائل با ابعاد بزرگ باشد

مراجع

- [1] Agraalw. S. K. and B. C. Fabien., 1999. Optimization of Dynamic Systems, Kluwer Academic Publishers, Netherlands.
- [2] Angurita. D. Boni, A., 2002. Improved neural network for SVM learning IEEE Trance, Neural Network. 13, 1243-1244.
- [3] Avriel. M., 1976. Nonlinear Programming: Analysis and Methods, Prentice-Hall, Englewood Cliffs, NJ.
- [4] Bazaraa.M. S. Sherali, C.M. and Shetty, C. M., 1979. Nonlinear Programming, Theory and Algorithms, John Wiley Sons, New York.
- [5] Bazaraa. M. S, H. D. Sherali and C. M. Shetty., 1993. Nonlinear Programming- Theory and Algorithms, 2nd ed. New York: Wiley.
- [6] Burges. C., 1998. A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery 2 (2) 121–167.
- [7] Ding. K. and N.-J. Huang., 2008. A new class of interval projection neural networks for solving interval quadratic program, Chaos, Solitons and Fractals, 35 718–725.
- [8] Effati.S, M. Baymani., 2005. A new nonlinear neural network for solving quadratic programming problems, Appl. Math. Comput. 165. 719-729.
- [9] Effati.S, Nazemi. A. R., 2006. Neural network models and its application for solving linear and quadratic programming problems, Appl. Math. Comput. 172 305 -331.
- [10] Effati.S, Ranjbar, M., 2011. A novel reccurent nonlinear neural network for solving quadratic programming problems, Appl. Mach. 35, 1688-1695.
- [11] Gao. X.B., Liao L.Z., 2010. A new one-layer neural network for linear quadratic programming, IEEE Trans. Neural Netw 6.918-929.
- [12] Ghasabi-Oskoei.H, a. Maleki, A. Ahmadi., 2007. Novel artificial neural network with simulation aspects for solving quadratic programming problems, comput. Math.App. 53 1493-1454.

- [13] Hale. J.K., 1969. Ordinary Differential Equations, Wiley-Interscience, New York .
- [14] Hu. X, Wang, J., 2007 . Design of general projection neural networks for solving monotone linear variational inequalities and linear and quadratic optimization problems. IEEE Transe . Syst, Man Cybern, Part B 37 , 1414-1421.
- [15] Huang. Y.C., 2002. A novel method to handle inequality constraints for convex programming neural network. Neural Process. Lett. 16, 17-27.
- [16] Kennedy. M.P, L.O. Chua., 1998. Neural network for nonlinear programming, IEEE Trans. Curcuits Syst. 35. 554-562.
- [17] Kennedy. M.P, L.O. Chua, Neural network for nonlinear programming, IEEE Trans. Curcuits Syst. 35 (1998) 554-562.
- [18] Liao. L, H. Qi, L. Qi., 2001. Solving nonlinear complementarity problems with neural networks: a reformulation method approach, J. Comput Appl. Math . 131. 343-359.
- [19] Liu,Q., Cao, J., 2011. Global exponented stability of discreted-time recurrent neural network for solving quadratic programming problems subject to linear constraints. Neuro Comput. 74. 3494-3501.
- [20] Lillo, W.E., Loh, M.H., Hui, S., Za[˘] k, S.H., 1993. On solving constrained optimization problems with neural networks: a penalty method approach. IEEE Trans. Neural Netw. 4, 931–939.
- [21] Luenberger. D.G., 1973. Introduction to Linear and Nonlinear programming, Addison Wesley, Reading, MA.
- [22] Maa, C.-Y. and Shanblatt, M. A., 1992. A two-phase optimization neural network. IEEE Trans. Neural Netw., 3(6):1003–1009.
- [23] Maleki. A, N. Hosseini-pour-Mahani, S. Ezazipour., 2010. Efficient recurrent neural network model for the solution of general nonlinear optimization problems, Optim Meth.Soft. 25. 1–18.
- [24] Miller. R. K, Michel. A. N., 1982. Ordinary Differential Equations, Academic Press, NewYork.
- [25] Nazemi, A.R., 2011. Adynamical model for solving quadratic minimax problems with constraint . J. Comput Appl. Math . 236. 1282-1295.
- [26] Nazemi, A.R., 2014. A neural network model for solving convex quadratic programming problems with some applications . J. Comput Appl. Math., 32. (54-62).

-
- [27] Tao. Q, J.D. Cao, M.S. Xue, H. Qiao., 2001. A high performance neural network for solving nonlinear programming problems with hybrid constraints, *Phys. Lett. A.*, 288. (2) 88–94.
- [28] Tank. D.W, J.J. Hopfield., 1989. Simple neural optimization networks: an A/D converter, singal descision network and a linear programming circuit, *IEEE Trans. Curcuits Syst.* 33. 533-541.
- [29] Wu. X, Y. Xia, J. Li, W.K. Chen., 1996. A high-performance neural network for solving linear and quadratic programming problems, *IEEE Trans. Neural Networks.*, 7. 643-651.
- [30] Xia. Y, J.Wang., 2000. Global exponential stability of recurrent neural network for solving optimization and related problems, *IEEE Trans. Neural Network.*, 11. 1017-1022.
- [31] Xia. Y, J.Wang., 1995. Neural Network for solving linear programming problems with bounded variables, *IEEE Trans. Neural Networks* 6. 1544-1547.
- [32] Xia. Y., 1996. A new neural network for solving nonlinear programming problems and its applications, *IEEE Trans. Neural Network* 7. 525-529.
- [33] Xia. Y., 1997. Neural Network for solving extended linear programming problems. *IEEE Trans. Neural Network* 8. 519-525.
- [34] Xia, Y., Feng, G., and Wang, J., 2004. A recurrent neural network for nonlinear convex opti mization subject to nonlinear inequality constraints. *IEEE Trans. Circuits Syst.*, 51. 1385–1394. 2008.
- [35] Xia, Y., Feng, G., 2003. Solving convex quadratic progarmming problems by an modified neural network with exponential convergence, In: *IEEE International Conference Neural Networks Signal Processing*, Nanjing, China, PP. 14-17.
- [36] Zhang, S. and Constantinides, A.G., 1992. Lagrange programming neural networks. *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, 39(7):441–452.

واژه‌نامه فارسی به انگلیسی

LaSalle Principle of Invariance	اصل تغییرناپذیری لسال
Strictly convex	اکیداً محدب
Strictly Monotone	اکیداً یکنوا
Convex Optimization	بهینه‌سازی محدب
Nonconvex Optimization	بهینه‌سازی نامحدب
Globally Asymptotically Stable	پایدار مجانبی سراسری
Stability in the Sense of Lyapunov	پایداری به مفهوم لیاپانوف
Globally Exponential Stability	پایداری نمایی سراسری
Lipschitz Continuous	پیوسته لپ شیتز
Locally Lipschitz Continuous	پیوسته لپ‌شیتز محلی
Energy Function	تابع انرژی
Regression Function	تابع رگرسیون
Hubber Function	تابع هوپر
Transpose Jacobian	ترانپوز ژاکوبین
Activation Function	تابع فعال‌سازی
Firing	تحریک شدن
Feasible Solution	جواب شدنی
Least Squares	حداقل مربعات
State Trajectory	خط سیر
Local Duality	دوگانی موضعی
Transient Behavior	رفتار ناپایدار
Euler Method	روش اویلر
Discrete time	زمان گسسته
Neural Networks	شبکه‌های عصبی
Feedback Neural Network	شبکه‌های عصبی بازگشتی

Artificial Neural Networks	شبکه‌های عصبی مصنوعی
Lagrangian Multiplier	ضربگر لاگرانژ
Projection Operator	عملگر تصویر
Global Minimum	کمینه سراسری
Local Minimum	کمینه موضعی
Hessian Matrix	ماتریس هسین
Jacobian Matrix	ماتریس ژاکوبین
Support Vector Machine	ماشین بردار پشتیبانی
Dual Variable	متغیر دوگان
Invariant Set	مجموعه تغییرناپذیر
Level Set	مجموعه سطح
Quadratic Programming	مسئله برنامه ریزی درجه دوم
Lagrangian Dual Problem	مسئله حداقل مربعات
Dual Program	مسئله دوگان
Linear Complementarity Problem	مسئله مکمل خطی
Nonlinear Complementarity Problem	مسئله مکمل غیرخطی
Positive Definite	معین مثبت
Strongly Positive Definite	معین مثبت قوی
Negative Definite	معین منفی
Nonnegative Region	ناحیه نامنفی
Initial Point	نقطه اولیه
Equilibrium Point	نقطه تعادل
Differentiable Mapping	نگاشت مشتق پذیر
Positive Semidefinite	نیمه معین مثبت
Globally Convergent	همگرای سراسری

واژه‌نامه انگلیسی به فارسی

Activation Function	تابع فعال سازی
Artificial Neural Networks	شبکه های عصبی مصنوعی
Convex Optimization	بهینه سازی محدب
Differentiable Mapping	نگاشت مشتق پذیر
Discrete time	زمان گسسته
Dual Program	مسأله دوگان
Dual Variable	متغیر دوگان
Equilibrium Point	نقطه تعادل
Energy Function	تابع انرژی
Euler Method	روش اویلر
Feasible Solution	جواب شدنی
Feedback Neural Network	شبکه های عصبی بازگشتی
Firing	تحریک شدن
Globally Asymptotically Stable	پایدار مجانبی سراسری
Globally Convergent	همگرای سراسری
Globally Exponential Stability	پایدار نمایی سراسری
Global Minimum	کمینه سراسری
Hessian Matrix	ماتریس هسین
Hubber Function	تابع هوبر
Initial Point	نقطه اولیه
Invariant Set	مجموعه تغییر ناپذیر
Jacobian Matrix	ماتریس ژاکوبین
Lagrangian Dual Problem	مسأله دوگان لاگرانژ
Lagrangian Multiplier	ضربگر لاگرانژ
Least Squares	حداقل مربعات

Level Set.....	مجموعه سطح
Linear Complementarity Problem.....	مسئله مکمل خطی
Lipschitz Continuous.....	پیوسته لیپ شیتز
Local Duality.....	دوگانی موضعی
Local Minimum.....	کمینه موضعی
Locally Lipschitz Continuous.....	پیوسته لیپ شیتز محلی
Negative Definite.....	معین منفی
Neural Networks.....	شبکه های عصبی
Nonconvex Optimization.....	بهینه سازی نا محدب
Nonlinear Complementarity Problem.....	مسئله مکمل غیرخطی
Nonnegative Region.....	ناحیه نامنفی
Positive Definite.....	معین مثبت
Projection Operator.....	عملگر تصویر
Positive Semidefinite.....	نیمه معین مثبت
Quadratic Programming.....	مسئله برنامه ریزی درجه دوم
Regression Function.....	تابع رگرسیون
Stability in the Sense of Lyapunov.....	پایداری به مفهوم لیاپانوف
State Trajectory.....	خط سیر
Strictly convex.....	اکیداً محدب
Strictly Monotone.....	اکیداً یکنوا
Support Vector Machine.....	ماشین بردار پشتیبانی
Strongly Positive Definite.....	معین مثبت قوی
Transpose Jacobian.....	ترانپوزه ژاکوبین
Transient Behavior.....	رفتار ناپایدار

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 a model 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 method using several examples of convex quadratic programming problems is shown.

Finally, we provide conclusions and recommendations for future work.



Shahrood University
Faculty Of Mathematical Sciences

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

**A duality model of neural network for solving
a class of nonlinear quadratic programming
problems**

Supervisor
Dr. Alireza Nazemi

by
Masoume Arabshahi Moghadam

2015