
بسم الله الرحمن الرحيم

دانشکده علوم ریاضی

گروه ریاضی کاربردی

پایان‌نامه کارشناسی ارشد

یک کاربرد ازسیستمهای دینامیکی برای حل رده ای ازمسائل بهینه سازی ناهموار

عبدالسخی نظری

استاد راهنما

دکتر علیرضا ناظمی

۱۳۹۴

تقدیم به

روح پاک مادرم و

پدر بزرگوارم

سپاس گزاری

من لم یشکر المخلوق، لم یشکر الخالق حمد و سپاس یکتای بی همتا را که لطفش بر ما عیان است، ادای شکرش را هیچ زبان و دریای فضلش را هیچ کران نیست و اگر در این وادی هستیم، همه محبت اوست. الهی ای مهربانتر از ما به ما، از تو می خواهم همه کسانی را که حتی ذره ای در انجام این امر مرا یاری نموده اند، در سایه لطف و محبت بی کرانت، سلامت، شادکام و موفق بداری. بر خود لازم می دانم تا مراتب سپاس را از بزرگوارانی به جا آورم که اگر دست یاریگرشان نبود، هرگز این پایان نامه به انجام نمی رسید. ابتدا از استاد گرانقدرم جناب آقای دکتر علیرضا ناظمی که زحمت راهنمایی این پایان نامه را بر عهده داشتند، کمال سپاس را دارم. از جناب آقای دکتر جعفر فتحعلی و دکتر مهدی ایرانمنش که زحمت داوری این پایان نامه با ایشان بوده است، کمال سپاس را دارم. پدر، مادر،برادر،خواهر و دوستان عزیزم، از زحمات بی دریق و بی منت شما متشکرم. همیشه نیازمند محبت، لطف و دعای خیر شما بوده و هستم.

عبدالسخی نظری
۱۳۹۴

تعهدنامه

اینجانب **عبدالسخی نظری** دانشجوی کارشناسی ارشد رشته **ریاضی کاربردی** **دانشکده علوم ریاضی** دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان **یک کاربرد ازسیستمهای دینامیکی برای حل رده ای ازمسائل بهینه سازی ناهموار** ، 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، تا کنون 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.

عبدالسخی نظری
۱۳۹۴

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 ارتباط 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 بیش احساس می‌شود.

در چند دهه اخیر روش‌های بهینه‌سازی که بر پایه رویکرد هوش مصنوعی توسعه یافته‌اند، موفقیت‌های چشم‌گیری در حل مؤثر وکارای مسائل بهینه‌سازی به‌دست آورده‌اند.

روش‌هایی چون الگوریتم ژنتیک، جست‌وجوی 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 پایان‌نامه یک مدل شبکه عصبی بازگشتی برای حل رده‌ای از مسائل بهینه‌سازی مشتق ناپذیر ارائه می‌شود. تحلیل وجود یکتایی، پایداری و همگرای سراسری جواب‌ها مورد بررسی قرار می‌گیرند و عملکرد روش‌های ارائه شده با به کارگیری چند مثال از مسائل برنامه‌ریزی مشتق ناپذیر نشان داده می‌شود.

در انتها نتایج کار و پیشنهاداتی برای کارهای آتی ارائه می‌دهیم.

کلمات کلیدی: پایداری، همگرایی، برنامه ریزی مشتق ناپذیر، بهینه سازی، شبکه عصبی، مسائل مین ماکس، برنامه ریزی خطی و غیر خطی، آنتروپی

فهرست مطالب

فصل ۱

مقدمه

در این پایان نامه یک مدل شبکه عصبی[۱] بازگشتی برای حل مسائل بهینه سازی مشتق ناپذیر ارائه می کنیم. شبکه های عصبی ارائه شده در این پایان نامه در مقایسه با شبکه‌های عصبی قبلی پیچیدگی کمتری برای پیاده سازی دارد. علاوه بر این، شبکه های عصبی پیشنهادی همگرای سراسری به جواب بهینه مسأله هستند. نمونه های شبیه سازی بر اساس مشکلات پایه مورد بحث قرارگرفته اند تا عملکرد درست شبکه عصبی پیشنهادی را نشان دهند.

۱.۱ تاریخچه تحقیق

قدمت شبکه‌های عصبی مصنوعی به دهه ۵۰ میلادی بر می گردد. در سال ۱۹۵۶ بنیاد راکفلر برگزاری کنفرانسی را بر عهده داشت که چشم‌اندازش به صورت زیر بود:
امکان استفاده از کامپیوترها و شبیه سازی 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.

[۱] Neural network

۲.۱ کاربرد مسائل برنامه ریزی مشتق ناپذیر

مسأله برنامه ریزی مشتق ناپذیر 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 برنامه‌های کاربردی نیازمند پردازش در زمان واقعی داده‌ها است، اغلب نیازمند استفاده از شبکه‌های عصبی به عنوان یک ابزار کارا در یافتن جواب‌های زمان واقعی هستیم. چون زمان مورد نیاز برای حل مسأله بهینه سازی مشتق ناپذیر وابستگی زیادی به بعد و ساختار مسأله دارد لذا معمولاً روش‌های عددی مرسوم در کاربردهای واقعی چندان مؤثر نیستند. یک رهیافت امیدوار کننده عبارت است از استفاده از شبکه‌های عصبی بازگشتی مبنی بر پیاده سازی دوری با مدارهای الکترونیکی. در این فصل تعاریف مورد نیاز در فصل‌های بعدی به‌طور مختصر آورده شده است. ابتدا تعاریف مرتبط با تابع محدب [۲] را بیان نموده و در خصوص مسائل بهینه‌سازی مقید و نامقید به ذکر شرایط لازم و کافی بهینگی [۳] می‌پردازیم. در انتها مفاهیم پایداری [۴] و تابع انرژی در سیستم‌های دینامیکی را ارائه می‌دهیم.

۳.۱ تعاریف

تعریف ۱.۳.۱. تابع $f : \mathbb{R}^n \to \mathbb{R}^l$ 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 ۲.۳.۱. مجموعه$E \subset \mathbb{R}^n$ را محدب می‌گوییم اگر:

$$\forall x, y \in E, \lambda \in (۰, ۱) \ : \ \lambda x + (۱ - \lambda)y \in E.$$

تعریف ۳.۳.۱. تابع $f : \mathbb{R}^n \to \mathbb{R}$ بر روی مجموعه نقاط واقع در مجموعه محدب $E \subset \mathbb{R}^n$ یک تابع محدب نامیده می‌شود، اگر:

$$\forall x, y \in E, \ \lambda \in (۰, ۱) \ : \ f(\lambda x + (۱ - \lambda)y) \leq \lambda f(x) + (۱ - \lambda)f(y).$$

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

تعریف ۴.۳.۱. اگر $\bar{x} \in X$ و $\varepsilon-$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.

---

[۲] Convex

[۳] Optimality

[۴] Stability

تعریف ۵.۳.۱. تابع $\|.\| : \mathbb{R}^n \to \mathbb{R}$ را تابع نرم گوییم هرگاه به ازای هر $x, y \in \mathbb{R}$ و $\alpha \in \mathbb{R}$:

$$\begin{cases} \|x\| \geqslant ۰, \quad \forall\, x, \\ \|x\| = ۰ \Leftrightarrow x = ۰, \\ \|\alpha x\| = |\alpha|\|x\|, \\ \|x + y\| \leqslant \|x\| + \|y\|. \end{cases}$$

تعریف ۶.۳.۱. 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.

تعریف ۷.۳.۱. یک ماتریس $M(X)$ در اندازه $n \times n$ که عناصر آن m_{ij} $(i, j = ۱, ..., n)$ توابعی روی $E \subset \mathbb{R}^n$ هستند،

الف. نیمه معین مثبت [۵] روی E نامیده می‌شود، اگر
$$\forall\, V \in \mathbb{R}^n, V \neq ۰, X \in E \ : \ V^T M(X) V \geq ۰.$$

ب. معین مثبت [۶] روی E نامیده می‌شود، اگر
$$\forall\, V \in \mathbb{R}^n, V \neq ۰, X \in E \ : \ V^T M(X) V > ۰.$$

شرایط لازم و کافی برای مسائل نامقید

قضیه ۸.۳.۱. *(شرط لازم مرتبه اول [۳]): فرض کنید X یک مجموعه باز در $\mathbb{R}^n$ باشد. مسأله برنامه‌ریزی زیر را در نظر بگیرید:*

$$\min \qquad f(x) \tag{۱.۱}$$
$$\text{s.t} \qquad x \in X \tag{۲.۱}$$

که در آن $f : \mathbb{R}^n \to \mathbb{R}$ در $\hat{x}$ مشتق‌پذیر است. اگر $\hat{x}$ یک کمینه موضعی باشد، آنگاه $\nabla f(\hat{x}) = ۰$.

قضیه ۹.۳.۱. *(شرط لازم مرتبه دوم [۳]) : فرض کنید مشتق دوم $f : \mathbb{R}^n \to \mathbb{R}$ در $\hat{x}$ موجود باشد. اگر $\hat{x}$ یک کمینه موضعی باشد، آنگاه $\nabla f(\hat{x}) = ۰$ و ماتریس هسین [۷] f در $\hat{x}$ نیمه معین مثبت است.*

قضیه ۱۰.۳.۱. *(شرط کافی [۳]): فرض کنید مشتق دوم $f : \mathbb{R}^n \to \mathbb{R}$ در $\hat{x}$ وجود باشد. اگر $\nabla f(\hat{x}) = ۰$ و ماتریس هسین f در $\hat{x}$ معین مثبت باشد، آنگاه $\hat{x}$ یک کمینه موضعی اکید است.*

قضیه ۱۱.۳.۱. *([۳]) : فرض کنید مجموعه E بصورت زیر باشد:*

$$E = \{x \in \mathbb{R}^n \mid g_k(x) \leq ۰,\ k = ۱, \cdots, m,\ h_p(x) = ۰,\ p = ۱, \cdots, l\}$$

اگر $g_k(x)$، $k = ۱, \cdots, m$ توابعی محدب و $h_p(x)$، $p = ۱, \cdots, l$ آفین باشند آنگاه، E یک مجموعه محدب است.

---

[۵] Positive semi-defnite

[۶] Positive defnite

[۷] Hessian

قضیه ۱.۲.۳.۱. *([۳]): اگر $f(x)$ بر مجموعه محدب E تابعی محدب باشد آنگاه $f(x)$ دارای یک کمینه موضعی[۸] است. اگر چنین کمینه‌ای یافت شود، یک کمینه سراسری[۹] است و بر روی مجموعه محدب به‌دست آورده می‌شود.*

شرایط لازم و کافی برای مسائل مقید

قضیه ۱.۳.۳.۱. *(شرایط لازم کروش-کان-تاکر[۱۰] $(K.K.T)$،[۳]): فرض کنید X یک مجموعه باز در $\mathbb{R}^n$ باشد. مسأله کمینه‌سازی مقید زیر را در نظر بگیرید:*

$$\min \qquad f(x) \tag{۳.۱}$$

$$\text{s.t.} \quad \begin{cases} g_k(x) \le ۰, & k = ۱, \cdots, m, \\ h_j(x) = ۰, & j = ۱, \cdots, l, \\ x = (x_۱, x_۲, ..., x_n)^T \in X, \end{cases} \tag{۴.۱}$$

که در آن $f : \mathbb{R}^n \to \mathbb{R}$ ، $g_k : \mathbb{R}^n \to \mathbb{R} \quad k = ۱, ..., m$ و $h_j : \mathbb{R}^n \to \mathbb{R} \quad j = ۱, ..., l$. فرض کنید $\hat{x}$ یک جواب شدنی این مسأله باشد و $K = \{k : g_k(\hat{x}) = ۰\}$. 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 $(j = ۱, ..., l)$ دارای مشتق پیوسته باشند. به‌علاوه فرض کنید $\nabla g_k(\hat{x})$ برای $k \in K$ و $\nabla h_j(\hat{x})$ برای $(j = ۱, ..., l)$ مجموعاً مستقل‌خطی باشند. اگر $\hat{x}$ کمینه موضعی برای مسأله (۳.۱) و (۴.۱) باشد، آنگاه اسکالرهای یکتای $\hat{u}_k \ge ۰$ برای $k \in K$ و $\hat{v}_j \ge ۰$ برای $(j = ۱, ..., l)$ موجود هستند به‌طوری که:

$$\begin{cases} \nabla f(\hat{x}) + \sum_{k \in K} \hat{u}_k \nabla g_k(\hat{x}) + \sum_{j=۱}^{l} \hat{v}_j \nabla h_j(\hat{x}) = ۰, \\ \hat{u_k} \ge ۰, \quad k \in K. \end{cases}$$

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

$$\begin{cases} \nabla f(\hat{x}) + \sum_{k=۱}^{m} \hat{u}_k \nabla g_k(\hat{x}) + \sum_{j=۱}^{l} \hat{v}_j \nabla h_j(\hat{x}) = ۰, \\ u_k g_k(\hat{x}) = ۰, \quad k = ۱, ..., m, \\ \hat{u_k} \ge ۰, \quad k = ۱, ..., m. \end{cases}$$

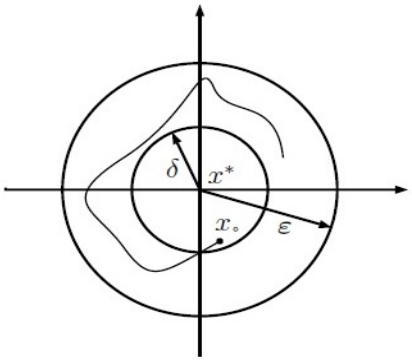
همچنین شرایط کروش-کان-تاکر می‌تواند به‌فرم ماتریسی زیر نوشته شود:

$$\begin{cases} \nabla f(\hat{x}) + \nabla g(\hat{x})^T u + \nabla h(\hat{x})^T v = ۰, \\ u^T g(\hat{x}) = ۰, \\ u \ge ۰, \end{cases}$$

[۸] Local minimum

[۹] Global minimum

[۱۰] Karush-Kuhn-Tucker



شکل ۱.۱: پایداری لیاپانوف

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 ۱۴.۳.۱. *(شرایط کافی کروش-کان-تاکر $(K.K.T)$،[۳]): فرض کنید X یک مجموعه باز در $\mathbb{R}^n$ باشد، مسأله (۳.۱) و (۴.۱) را در نظر بگیرید. فرض کنید $\hat{x}$ یک جواب شدنی باشد، همچنین فرض کنید f و g_k ها برای $(k = ۱, \cdots, m)$ توابعی محدب و h_j ها برای $(j = ۱, \cdots, l)$ توابعی آفین باشند و $K = \{k : g_k(\hat{x}) = ۰\}$ ، اگر شرایط کروش-کان-تاکر در $\hat{x}$ برقرار باشد، یعنی اسکالرهای $\hat{u}_k \geq ۰$ برای $k \in K$ و $\hat{v}_j \geq ۰$ برای $(j = ۱, ..., l)$ موجود باشند به‌طوری که:*

$$\nabla f(\hat{x}) + \sum_{k \in K} \hat{u}_k \nabla g_k(\hat{x}) + \sum_{j=۱}^{l} \hat{v}_j \nabla h_j(\hat{x}) = ۰,$$

آن‌گاه $\hat{x}$ یک کمینه موضعی برای (۳.۱) و (۴.۱) خواهد بود.

۴.۱ پایداری

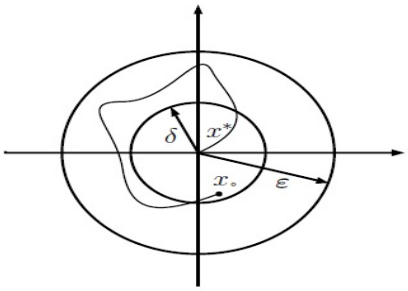
سیستم دینامیکي زیر را در نظر بگیرید:

$$\begin{cases} \dot{x} = f(x(t)), \\ x(t_۰) = x_۰ \in \mathbb{R}^n, \end{cases} \tag{۵.۱}$$

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

تعریف ۱.۴.۱. x^* یک نقطهٔ تعادل [۱۱] برای (۵.۱) نامیده می‌شود اگر $f(x^*) = ۰$.

[۱۱] Equilibrium point



شکل ۲.۱: پایداری مجانبی

تعریف ۲.۴.۱. فرض کنید که $x(t)$ یک جواب (۵.۱) باشد، نقطهٔ تعادل تنها x^*، پایدار به مفهوم لیاپانوف [۱۲] است (شکل (۱.۱) را ببینید)، اگر برای هر $x(t_\circ) = x_\circ$ و هر $\varepsilon > 0$ ، یک $\delta > 0$ وجود داشته باشد به‌طوری که اگر $\| x^* - x_\circ \| \leq \delta$ آنگاه

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

تعریف ۳.۴.۱. سیستم دینامیکی (۵.۱) همگرای سراسری[۱۳] به مجموعه

$\Omega^* = \{x \mid$ جواب سیستم (۵.۱) است. $\}$.

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

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

که در آن

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

تعریف ۴.۴.۱. سیستم دینامیکی در نقطهٔ تعادل یکتا x^* پایدار مجانبی سراسری [۱۴] نامیده می‌شود اگر x^* به مفهوم لیاپانوف پایدار باشد و در شرط زیر صدق کند:

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

شکل (۲.۱) را ببینید.

تعریف ۵.۴.۱. مجموعه $M \subseteq \mathbb{R}^n$، یک مجموعه پایدار[۱۵] نسبت به سیستم (۵.۱) گفته می‌شود اگر $x(t_\circ) \in M$ و $t_\circ \geq 0$، آنگاه به ازای هر $t \geq t_\circ$، $x(t) \in M$ باشد.

تعریف ۶.۴.۱. نگاشت F در شرط لیپ شیتز[۱۶] صدق می‌کند اگر عدد ثابت L وجود داشته باشد، به‌طوری که برای هر دو نقطه $x, y \in \mathbb{R}^n$ داشته باشیم:

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

[۱۲] Lyapunov
[۱۳] Global convergance
[۱۴] Globally asympotically stable
[۱۵] Invariant set
[۱۶] Lipschitz

تعریف ۷.۴.۱. F را لیپ شیتز محلی روی $\mathbb{R}^n$ نامیم اگر به ازای هر نقطه از $\mathbb{R}$، یک همسایگی مانند $D_\circ \subset \mathbb{R}^n$ وجود داشته باشد به‌طوری که نامساوی بالا برای هر دو نقطه $x, y \in D_\circ$ برقرار باشد.

تعریف ۸.۴.۱. نگاشت $F : \mathbb{R}^n \to \mathbb{R}^n$ یکنوا[۱۷] گفته می‌شود، اگر برای هر زوج از نقاط $x, y \in \mathbb{R}^n$ داشته باشیم:

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

همچنین روی $\mathbb{R}^n$ اکیداً یکنوا [۱۸] است، اگر نامساوی فوق به‌صورت 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 $\beta > \circ$ وجود داشته باشد، به‌طوری که:

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

قضیه ۹.۴.۱. *([۱]): فرض می‌کنیم که $F : K \subset \mathbb{R}^n \to \mathbb{R}^n$، یک تابع به‌طور 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 ۱۰.۴.۱. *([۱]): در سیستم دینامیکی (۵.۱) فرض می‌کنیم که $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ یک تابع پیوسته باشد، آنگاه برای هر $t_\circ > \circ$ و $x_\circ \in \mathbb{R}^n$، یک جواب محلی[۱۹] $x(t)$ به ازای $t \in [t_\circ, \tau)$ که $\tau > t_\circ$ وجود دارد. علاوه بر این اگر f در $x_\circ$ 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.*

۵.۱ تابع انرژی

قضیه ۱.۵.۱. *([۱]): فرض کنید که $x = \circ$ یک نقطهٔ تعادل برای سیستم (۵.۱) و تابع $E : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ به‌طور پیوسته مشتق پذیر باشد اگر*

۱. $E(\circ) = \circ$،

۲. *به ازای هر* $x \in \Omega \subseteq \mathbb{R}^n \setminus \{\circ\}$ ، $E(x) > \circ$،

۳. *به ازای هر* $x \in \Omega \subseteq \mathbb{R}^n \setminus \{\circ\}$ ، $\dot{E}(x) \leq \circ$،

آنگاه $x = \circ$ نقطهٔ پایداری لیاپانوف سیستم خواهد بود و $E(x)$ را« تابع لیاپانوف» یا «تابع انرژی» برای سیستم (۵.۱) نامیم.

قضیه ۲.۵.۱. *: [۱] اصل تغییرناپذیری (ناوردای) لازال [۲۰] : فرض می‌کنیم تابع $V : \mathbb{R}^n \to \mathbb{R}$ به‌طور پیوسته مشتق پذیر باشد، همچنین فرض می‌کنیم*

[۱۷] Monotone
[۱۸] Strictly monotonic
[۱۹] Local solution
[۲۰] Lasalle's invariance principle

۱. $M \subseteq \mathbb{R}^n$ یک مجموعه فشرده و پایدار نسبت به جواب سیستم (۵.۱) باشد،

۲. به ازای هر $x \in M$ ، $\dot{V}(x) \leq ۰$،

۳. E مجموعه همه نقاط M باشد به‌طوری‌که $\dot{V}(x) = ۰$،

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

آن‌گاه به ازای هر $x(t_۰) \in M$ که $t_۰ \geq ۰$ باشد، داریم:

$$\lim_{t\to\infty} dist(x(t), N) = ۰.$$

فصل ۲

تاریخچه مختصر از کاربرد شبکه های عصبی در بهینه‌سازی

۱.۲ مقدمه

دهه‌های آغازین سده بیستم میلادی و دوران پیشرفت شگرف صنعتی، همراه با تولید خودرو بود که انقلاب همه جانبه این در ترابری، افزایش شتاب جابجایی و صدها کار و پیشه جدید در رشته‌های بازرگانی بوجود آورده است. به نظر می‌رسد که سمبل دوران فراصنعتی و نماد فرآورده‌های بی‌همتای قرن آینده «هوش مصنوعی» است. امروزه موضوع هوش مصنوعی داغ ترین بحث میان کارشناسان دانش رایانه و اطلاعات و دیگر دانشمندان و تصمیم‌گیرندگان است.در سراسر تاریخ تا به امروز انسان از جنبه تن و روان، مرکز و محور بحث‌ها و پژوهش‌ها بوده است.ولی اکنون موجودی با رتبه‌ای پائین‌تر، بی‌جان و ساختگی مي‌خواهد جانشین او شود، امری که بدون شك می‌توان ادعا نمود بیشتر انسان‌ها با آن مخالفند. هوش مصنوعی چنانچه به هدف هاي والای خود برسد،جهش بزرگی در راه دستیابی بشر به رفاه بیشتر و حتی ثروت افزون‌تر خواهد بود.هم اکنون نمونه‌هاي خوبی از هوش مصنوعی در دنیای واقعی ما به کار افتاده است.چنین دستاوردهایی صرف منابع لازم در آینده را همچنان توجیه خواهد کرد. از سوی دیگر، منتقدین هوش مصنوعی چنین استدلال می‌کنند که صرف زمان و منابع ارزشمند دیگر در راه ساخت فراورده‌ای که پر از نقص و کاستی و دست‌آوردهای مثبت اندکی است، مایه بدنام کردن و زیر پا گذاشتن توانمندی‌ها و هوشمندی‌های انسان می‌باشد. شبکه عصبی مصنوعی روشی عملی برای یادگیری توابع گوناگون نظیر توابع با مقادیر حقیقی، توابع با مقادیر گسسته و توابع با مقادیر برداری می‌باشد. یادگیری شبکه‌ی عصبی در برابر خطاهای داده های آموزشی مصون بوده و شبکه با موفقیت به مسائلی نظیر شناسائی گفتار، شناسائی و تعبیر تصاویر، و یادگیری روبات اعمال شده‌اند. لذا در این فصل تاریخچه‌ای از شبکه‌ی عصبی و کاربرد آن در زمینه‌های مختلف علوم اشاره کرده و در انتها مدل‌هایی از شبکه‌های عصبی که برای حل مسائل گوناگون بهینه‌سازی (مسائل خطی یا غیر خطی) به دست آمده است را معرفی می‌کنیم.

۱.۱.۲ تاریخچه مختصر شبکه‌های عصبی

دیدگاه جدید شبکه‌های عصبی در دهه‌ی چهل قرن بیستم شروع شد، زمانیکه وارن مک کلوث[۱] و والترپیتز[۲] نشان دادند که شبکه‌های عصبی در اصل می‌توانند هر تابع حسابی[۳] و منطقی[۴] را محاسبه نمایند. کار این افراد را می‌توان نقطه‌ی شروع حوزه‌ی علمی شبکه‌های عصبی مصنوعی نامید و این موضوع با دونالدهب[۵] ادامه یافت. نخستین کاربرد عملی شبکه‌های عصبی در اواخر دهه‌ی پنجاه قرن بیستم مطرح شد، زمانیکه فرانک روزنبلات[۶] در سال ۱۹۵۸ شبکه‌ی پرسپترون[۷] را معرفی نمود. روزنبلات و همکارانش شبکه‌ای ساختند که قادر بود الگوها را از هم شناسایی نماید. در همین زمان بود که برناردویدرو[۸] در سال ۱۹۶۰ شبکه‌ی عصبی تطبیقی خطی آدلاین[۹] را با قانون یادگیری جدید مطرح نمود که از لحاظ ساختار شبیه شبکه‌ی پرسپترون می‌باشد. پیشرفت شبکه‌های عصبی تا دهه هفتاد قرن بیستم ادامه یافت. در ۱۹۷۲ تئوکوهونن[۱۰] و جیمزاندرسون[۱۱] به طور مستقل و بدون اطلاع از هم شبکه‌های عصبی جدید را معرفی نمودند که قادر بودند به عنوان عناصر ذخیره‌ساز عمل نمایند. استفان گروسبرگ[۱۲] در این دهه روی شبکه‌های خود سازمانده[۱۳] فعالیت می‌کرد. علاقه به کارکردن روی شبکه‌های عصبی در دهه شصت قرن بیستم در قیاس با دهه هشتاد به علت عدم بروز ایده‌های جدید و نبود کامپیوترهای سریع جهت پیاده‌سازی کمرنگ می‌نمود. لیکن در خلال دهه هشتاد، تحقیقات روی شبکه‌های عصبی دوباره افزایش یافت، در این زایش دوباره شبکه‌های عصبی دو نگرش جدید قابل تأمل می‌باشند. استفاده از مکانیسم تصادفی جهت توضیح عملکرد یک طبقه وسیع از شبکه‌های برگشتی[۱۴] که می‌توان آن‌ها را جهت ذخیره‌سازی اطلاعات استفاده نمود. این ایده توسط جان هاپفیلد[۱۵] در سال ۱۹۸۲ مطرح شد. دومین ایده‌ی مهم که کلید توسعه‌ی شبکه‌های عصبی در دهه هشتاد شد، الگوریتم پس انتشار خطا[۱۶] می‌باشد که توسط دیویدراملهارت[۱۷] و جیمزمکلند[۱۸] مطرح گردید. با بروز این دو ایده شبکه‌های

---

[۱] Warren McCulloch
[۲] Walter Pitts
[۳] Logical function
[۴] Arithmetic
[۵] Donald Hebb
[۶] Frank Rosenblalt
[۷] Pereceptron
[۸] Bernurd Widrow
[۹] Adaptive linear element
[۱۰] Teo Kohonen
[۱۱] Jemes Anderson
[۱۲] Stefan Grossberg
[۱۳] Self-organizing
[۱۴] Feedback(Recurrent)
[۱۵] John Hopfield
[۱۶] Error back-propagation
[۱۷] Daivid Rummelhurt
[۱۸] James Mcland

عصبی متحول شدند و هزاران مقاله نوشته شد و شبکه‌های عصبی کاربردهای زیادی در رشته‌های مختلف علوم پیدا کردند. بیشتر پیشرفت‌ها در شبکه‌های عصبی به ساختارهای نوین و روش‌های یادگیری جدید مربوط می‌شوند.

۲.۱.۲ چرا از شبکه عصبی استفاده می‌کنیم؟

شبکه‌های عصبی با توانایی قابل توجه خود در استنتاج نتایج از داده‌های پیچیده می‌توانند در استخراج الگوها و شناسایی گرایش‌های مختلفی که برای انسان‌ها و کامپیوتر شناسایی آنها بسیار دشوار است استفاده شوند. از مزایای شبکه‌های عصبی می‌توان موارد زیر را نام برد:

۱. یادگیری تطبیقی: توانایی یادگیری اینکه چگونه وظایف خود را بر اساس اطلاعات داده شده به آن و یا تجارب اولیه انجام دهد در واقع اصلاح شبکه را گویند.

۲. خود سازماندهی: یک شبکه عصبی مصنوعی به صورت خودکار سازماندهی و ارائه داده‌هایی که در طول آموزش دریافت کرده را انجام دهد. نرون‌ها با قاعده‌ی یادگیری سازگار شده و پاسخ به ورودی تغییر می‌یابد.

۳. عملگرهای بی‌درنگ: محاسبات در شبکه‌ی عصبی مصنوعی می‌تواند به صورت موازی و به وسیله سخت‌افزارهای مخصوصی که طراحی و ساخت آن برای دریافت نتایج بهینه قابلیت‌های شبکه عصبی مصنوعی است انجام شود.

۴. تحمل خطا: با ایجاد خرابی در شبکه مقداری از کارایی کاهش می‌یابد ولی برخی امکانات آن با وجود مشکلات بزرگ همچنان حفظ می‌شود.

۵. دسته بندی: شبکه‌های عصبی قادر به دسته بندی ورودی‌ها برای دریافت خروجی مناسب می‌باشند.

۶. تعمیم‌دهی: این خاصیت شبکه را قادر می‌سازد تا تنها با برخورد با تعداد محدودی نمونه، یک قانون کلی از آن را به دست آورده، نتایج آموخته‌ها را به موارد مشابه از قبل تعمیم دهد. توانایی که در صورت نبود آن سامانه باید بی نهایت واقعیت‌ها و روابط را به خاطر سپارد.

۷. پایداری-انعطاف پذیری: یک شبکه عصبی هم به حد کافی پایدار است تا اطلاعات فراگرفته خود را حفظ کند و هم قابلیت انعطاف و تطبیق را دارد و بدون از دست دادن اطلاعات قبلی می‌تواند موارد جدید را بپذیرد.

۳.۱.۲ معایب شبکه‌ی عصبی

با وجود برتری‌هایی که شبکه‌های عصبی نسبت به سامانه‌های مرسوم دارند، معایبی نیز دارند که پژوهشگران این رشته تلاش دارند که آنها را به حداقل برسانند، از جمله:

۱. قواعد یا دستورات مشخصی برای طراحی شبکه جهت یک کاربرد اختیاری وجود ندارد.

۲. در مورد مسائل مدل سازی، صرفاً نمی‌توان با استفاده از شبکه عصبی به فیزیک مساله پی برد. به عبارت دیگر مرتبط ساختن پارامترها یا ساختار شبکه به پارامترهای فرآیند معمولاً غیر ممکن است.

۳. دقت نتایج بستگی زیادی به اندازه مجموعه آموزش دارد.

۴. آموزش شبکه ممکن است مشکل و یا حتی غیر ممکن باشد.

۵. پیش بینی عملکرد آینده شبکه (عمومیت یافتن) آن به سادگی امکان پذیر نیست.

۲.۲ شبکه‌های عصبی در بهینه‌سازی

همانطور که گفته شد یکی از کاربردهای هوش مصنوعی در مسائل بهینه سازی می‌باشد، بهینه نمودن تابع هدف (ماکزیمم یا مینیمم کردن آن) با استفاده از روش‌های تحلیلی سر راست بسیار پیچیده می باشد. با وجود این امکان رسیدن به جواب با بکارگیری روش‌های عددی رایانه‌ای میسر می‌گردد. در این روش‌ها با تغییر متغیرهای طراحی در دوره‌های متوالی و با در نظر گرفتن قیود به جواب بهینه نزدیک می‌شویم. استفاده از روش‌های بهینه‌سازی متداول و تحلیل مستقیم سازه‌ها در دفعات مکرر موجب افزایش قابل ملاحظه‌ای زمان محاسبه می‌گردد. بحث و بررسی در مورد کاربرد شبکه‌های عصبی (شبکه‌های عصبی مصنوعی) در بهینه‌سازی از اوایل سال ۱۹۸۰ میلادی آغاز شده است. نتایج پژوهش‌های انجام شده از آن زمان تاکنون، مواردی چون برنامه‌ریزی خطی، برنامه‌ریزی درجه دوم، برنامه‌ریزی هندسی و برنامه‌ریزی غیرخطی را در بر می‌گیرد. ایده اصلی در استفاده از شبکه‌های عصبی مصنوعی برای مسائل بهینه سازی استفاده از یک تابع انرژی (نامنفی) و یک سیستم دینامیکی است که این دو بیان کننده‌ی مدل‌های شبکه‌های عصبی مصنوعی متناظر مسائل بهینه‌سازی هستند. سیستم دینامیکی بیان شده معمولاً یک دستگاه معادلات دیفرانسیل غیرخطی مرتبه اول است. انتظار می‌رود که برای یک نقطه‌ی آغازین نقطه‌ی پایداری دستگاه معادلات دیفرانسیل غیرخطی به‌دست آمده، جواب بهینه‌ی مسأله بهینه‌سازی اصلی باشد. یک اصل اساسی در استفاده از تابع انرژی این است که برای همگرایی دستگاه معادلات دیفرانسیل به نقطه‌ی پایداری آن، تابع انرژی متناظر با آن حتماً باید نامنفی باشد. البته می‌توان مدل شبکه‌های عصبی نظیر مسائل بهینه سازی را حتی بدون در نظر گرفتن تابع انرژی نیز بیان کرد. بنابراین برای یک مدل متناظر با مسائل بهینه‌سازی، اصل اساسی استفاده از شبکه‌های عصبی در اینگونه مسائل به‌صورت زیر بیان می‌شود:

برای یک نقطه‌ی آغازین دلخواه، نقطه‌ی پایداری دستگاه معادلات دیفرانسیل غیرخطی بدست آمده جواب بهینه مسأله اصلی است و برعکس.

مدل‌های مطرح شده متناظر مسائل مختلف بهینه‌سازی را می‌توان به دو قسمت مدل‌های دوگانی و مدل‌های جریمه‌ای تقسیم نمود. نظریه دوگانی و توابع جریمه‌ای دو نظریه بسیار مهم در مسائل بهینه‌سازی هستند که اکثر این مسائل بر مبنای این دو روش کلاسیک قابل حل هستند. در نظریه‌ی توابع جریمه‌ای

معمولاً از مدل‌های گرادیانی برای معرفی مدل مورد نظر استفاده می‌شود، ولی در نظریه دوگانی از مدل‌های اولیه –دوگان برای حل این مسائل استفاده می‌شود. اگر بتوان متناظر با هر مسأله بهینه‌سازی، روش مشخصی را ارائه نمود که آن روش برای حل آن مسأله شرایط لازم وکافی را برآورده سازد، آنگاه می‌توان متناظر با آن روش، یک مدل شبکه عصبی برای مسأله مورد نظر بسازیم. در زیر به بیان چند مدل که تاکنون برای مسائل برنامه‌ریزی خطی و درجه دوم و غیرخطی ارائه شده‌اند می‌پردازیم.

۳.۲ برخی مدل های پیشین ارائه شده شبکه عصبی در بهینه سازی

مدل اول

مسأله برنامه‌ریزی خطی زیر را در نظر بگیرید:

$$\min \quad f(x) = a^T x \tag{۱.۲}$$

$$\text{s.t.} \begin{cases} g(x) = Dx - b = ۰, \\ x \geq ۰, \end{cases} \tag{۲.۲}$$

که در آن $D \in \mathbb{R}^{m\times n}$ ، $\text{rank}(D) = m$، $x, a \in \mathbb{R}^n$ و $b \in \mathbb{R}^m$ است. دوگان (۱.۲) و (۲.۲) به‌صورت زیر بیان می‌شود:

$$\max \quad \bar{f}(y) = b^T y \tag{۳.۲}$$

$$\text{s.t.} \begin{cases} \bar{g}(y) = D^T y - a \leq ۰, \\ y \in \mathbb{R}^m. \end{cases} \tag{۴.۲}$$

در [۲۵] مدل متناظر برای حل (۱.۲) و (۲.۲) به‌صورت زیر نمایش داده شده است:

$$\begin{cases} \dfrac{dx}{dt} = -[D^T(Dx - b) - \beta(D^T y - a)], \\ \dfrac{dy}{dt} = -\beta[(Dx - D^T y - a)^+ - b], \end{cases}$$

که در آن $\beta = \| (x + D^T y - a)^+ - x \|_۲^۲$ ، $(x, y)^T \in \{(x, y)^T \mid x \geq ۰\}$ و $(Dx - D^T y - a)^+ = max\{(Dx - D^T y - a), ۰\}$.

مدل دوم

مسأله برنامه‌ریزی درجه دوم زیر را در نظر بگیرید:

$$\min \quad f(x) = \frac{۱}{۲} x^T A x + a^T x \tag{۵.۲}$$

$$\text{s.t.} \begin{cases} Dx - b = ۰, \\ x \geq ۰, \end{cases} \tag{۶.۲}$$

که در آن A یک ماتریس متقارن معین مثبت است. دوگان (۵.۲) و (۶.۲) به‌صورت زیر است:

$$\max \quad \bar{f}(y) = b^T y - \frac{۱}{۲} x^T A x \tag{۷.۲}$$

$$\text{s.t.} \quad \bar{D} y - A x - a \leq ۰. \tag{۸.۲}$$

در [۲۵] مدل متناظر برای حل (۵.۲) و (۶.۲) به‌صورت زیر نمایش داده شده است:

$$\begin{cases} \frac{dx}{dt} = -\{D^T(x-b) + \gamma(-D^T y + Ax + a) + \gamma A[x - (x + D^T y - Ax - a)^+]\}, \\ \\ \frac{dy}{dt} = -\gamma\{Dx - b + D[(x + D^T y - Ax - a)^+ - x]\}, \end{cases}$$

که در آن $\gamma = \|(x + D^T y - Ax - a)^+ - x\|_۲^۲$ و $(x,y)^T \in \{(x,y)^T \mid x \geq ۰\}$.

مدل سوم

مسأله برنامه‌ریزی غیرخطی زیر را در نظر بگیرید:

$$\min \quad f(x) \tag{۹.۲}$$

$$\text{s.t.} \quad g(x) = [g_۱(x), g_۲(x), ..., g_m(x)]^T \leq ۰, \tag{۱۰.۲}$$

که در آن $f(x), g(x) \in C^۲$ و $f(x), g_i(x) : \mathbb{R}^n \to \mathbb{R}^۱$. برای حل (۹.۲) و (۱۰.۲) هاپفیلد و تنک[۱۹] در [۲۳]، یک مدل شبکه عصبی به‌صورت زیر ارائه کردند:

$$\dot{x} = C^{-۱}\{-\nabla f(x) - \nabla g(x) g^+(x) - \frac{۱}{s} Q^{-۱} x\} s,$$

که در آن C یک ماتریس قطری $n \times n$ بیان کننده ظرفیت هر نرون و Q یک ماتریس قطری $n \times n$ است که اعضای آن بصورت زیر تعریف می‌شوند:

$$q_{ij} = \frac{۱}{\frac{۱}{\rho_i} + \sum_{j=۱}^{m} -d_{ji}},$$

که در آن $\frac{۱}{\rho_i}$ ضریب هدایت گرمایی هر نرون، d_{ji} وزن‌هاي تخصیص داده شده به نرونهای داخلی و s پارامتر جریمه است. تابع انرژی متناظر به‌صورت زیر انتخاب شده است:

$$E_۱(x) = f(x) + \sum_{j=۱}^{m} [g_i^+]^۲ + \sum_{i=۱}^{n} \frac{x_i^۲}{۲ s r_{ii}}.$$

مدل چهارم

برای حل (۹.۲) و (۱۰.۲) ، کندی و چوآ[۲۰] در [۲۴] مدلی دیگر به‌صورت زیر ارائه داده‌اند:

$$\dot{x} = C^{-۱}[-\nabla f(x) - \nabla g(x) g^+(x)],$$

[۱۹] Hopfield and Tank

[۲۰] Kennedy and Chua

که در آن C و s شرایط مدل سوم را دارند. تابع انرژی متناظر این مدل به‌صورت زیر انتخاب شده است:

$$E_۲(x) = f(x) + \frac{s}{۲}\sum_{j=۱}^{m}[g_i^+]^۲.$$

مدل پنجم

مجدداً برای حل (۹.۲) و (۱۰.۲) توسط رودریگز-واسکوئز[۲۱] در [۲۵] مدلی به‌صورت زیر ارائه شده است:

$$\dot{x} = -u_x \nabla f(x) - s\nabla g(x) g^+(x),$$

که در آن u_x متناظر اندیس‌های شدنی x است چنان‌که

$$u_x = \begin{cases} ۱ & g(x) \leq ۰, \\ ۰ & \text{در غیر اینصورت} \end{cases}$$

تابع انرژی متناظر این مدل نیز به‌صورت زیر انتخاب شده است:

$$E_۳(x) = -u_x f(x) + \frac{s}{۲}\sum_{j=۱}^{m}[g_j^+(x)]^۲.$$

همان طور که ملاحظه می‌کنید، مدل‌های بیان شده به دلیل مشکلاتی که هر کدام داشته‌اند به طور متوالی اصلاح شده‌اند.

مدل ششم

برنامه‌ریزی محدب درجه دوم[۲۲] زیر را در نظر بگیرید:

$$\min \quad f(x) = \frac{۱}{۲}x^T Q x + D^T x \tag{۱۱.۲}$$

$$\text{s.t.} \begin{cases} g(x) = Ax - b \leq ۰, \\ h(x) = Ex - f = ۰, \end{cases} \tag{۱۲.۲}$$

که در آن $Q \in \mathbb{R}^{n\times n}$ یک ماتریس متقارن معین مثبت، $A \in \mathbb{R}^{m\times n}$، $b \in \mathbb{R}^m$ ، $E \in \mathbb{R}^{l\times n}$، $f \in \mathbb{R}^l$ و $x \in \mathbb{R}^n$ است. بر طبق شرایط کروش-کان-تاکر ، $x^* \in \mathbb{R}^n$ جواب بهینه (۱۱.۲) و (۱۲.۲) است اگر وتنها اگر بردارهای $u^* \in \mathbb{R}^m$ و $v^* \in \mathbb{R}^l$ وجود داشته باشد که در رابطه زیر صدق کنند:

$$\begin{cases} u^* \geq ۰, Ax^* - b \leq ۰, {u^*}^T(Ax^* - b) = ۰, \\ Qx^* + D + A^T u^* + E^T v^* = ۰, \\ Ex^* - f = ۰. \end{cases} \tag{۱۳.۲}$$

[۲۱] Rodríguez Vazquez

[۲۲] Convex quadratic programming

برای حل مسأله (۱۳.۲) تابع زیر را که معروف به تابع فیشر-برمیستر[۲۳] است، معرفی می‌کنیم:

$$\begin{cases} \phi : \mathbb{R}^۲ \to \mathbb{R} \\ \phi(a,b) = \sqrt{a^۲ + b^۲} - a - b. \end{cases}$$

قضیه ۱.۳.۲. *[۲۵]:*

۱. $\phi(a,b) = ۰$ *اگر وفقط اگر* $a \geq ۰, b \geq ۰, ab = ۰$.

۲. *مربع* $\phi(a,b)$ *به‌طور پیوسته مشتق پذیر باشد.*

۳. $\phi(a,b)$ *همه جا به جز در مبدأ دو بار به‌طور پیوسته مشتق پذیر است، ولی در مبدأ قویاً*[۲۴] *هموار است.*

تابع ϕ *که در شرط ۱ قضیه‌ی (۱.۳.۲) صدق کند، مساله مکمل غیر خطی*[۲۵] *نامیده می‌شود.*

متناظر با دستگاه (۱۳.۲) دستگاه زیر را در نظر می‌گیریم:

$$\phi(x,u,v) = \begin{bmatrix} \phi(\alpha, Qx + D + A^T u + E^T v) \\ \phi(u, Ax - b) \\ \phi(\beta, Ex - f) \end{bmatrix} = ۰, \qquad (۱۴.۲)$$

که در آن $\alpha = (\alpha_۱, \ldots, \alpha_m)^T > ۰$، $u = (u_۱, \ldots, u_m)$ و $\beta = (\beta_۱, \ldots, \beta_l)^T > ۰$ است. مدل شبکه عصبی متناظر با مساله‌ی (۱۴.۲) به‌صورت زیر بیان شده است:

$$\begin{cases} \dfrac{dy}{dt} = -\tau \bigtriangledown E(y(t)), \quad \tau > ۰ \\ y(t_۰) = y_۰ \in \mathbb{R}^{n+m+l}. \end{cases} \qquad (۱۵.۲)$$

که در آن τ ضریبی برای افزایش سرعت همگرایی و کاهش تعداد تکرارها در روش عددی انتخاب شده برای حل دستگاه دینامیکی (۱۵.۲) است. تابع انرژی متناظر این مدل به‌صورت زیر انتخاب شده است:

$$E(y) = \frac{۱}{۲} \| \phi(y) \|^۲.$$

مدل هفتم

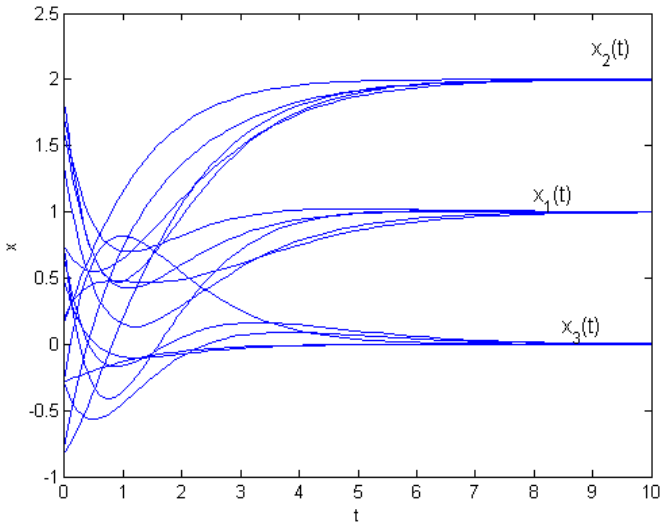
برای حل (۱۱.۲) و (۱۲.۲) تعریف کنید:

$$U(x,u,v) = \begin{bmatrix} -(Qx + D + A^T u + E^T v) \\ (u + Ax - b)^+ - u \\ (Ex - f) \end{bmatrix}. \qquad (۱۶.۲)$$

[۲۳] Ficher-Burmeister

[۲۴] Strongly smooth

[۲۵] Nonlinear complementarity problem



شکل ۱.۲: مسیر جواب در مثال ۲.۳.۲

مدل شبکه عصبی برای حل (۱۱.۲) و (۱۲.۲) در [۲۵] به‌صورت زیر بیان شده است:

$$\begin{cases} \dfrac{dy}{dt} = \kappa U(y), \\ y(t_\circ) = y_\circ, \end{cases} \tag{۱۷.۲}$$

که در آن κ نرخ همگرایی (نرخ همگرایی پارامتری است که سرعت همگرایی مدل شبکه عصبی با آن قابل تنظیم است) می‌باشد.

مثال ۲.۳.۲.

$$\min \ \ x_۱^۲ + x_۲^۲ + x_۳^۲$$

$$\text{s.t.} \begin{cases} ۲x_۱ + x_۲ - ۵ \leq ۰, \\ x_۱ + x_۳ - ۲ \leq ۰, \\ -x_۱ + ۱ \leq ۰, \\ -x_۲ + ۲ \leq ۰, \\ -x_۳ \leq ۰, \end{cases}$$

جواب بهینه به ازای $\kappa = ۱$ است. $x = (۱, ۲, ۰)^T$ با استفاده از شبکه عصبی (۱۷.۲) نقطه تعادل بدست آمده جواب بهینه مساله می‌باشد که مسیرهای جواب در شکل (۱.۳) رسم شده است.

مدل هشتم

مسأله‌ی برنامه ریزی درجه دوم (۱۱.۲) و (۱۲.۲) را در نظر بگیرید که در آن $Q \in \mathbb{R}^{n\times n}$ یک ماتریس نیمه معین مثبت است. برای حل مساله (۱۱.۲) و (۱۲.۲) در [۲۹] یک مدل شبکه عصبی به‌صورت

زیر ارائه می‌شود:

$$\frac{dy}{dt} = \kappa(I + M^T) \begin{bmatrix} -(Qx + D + A^T u + E^T v) \\ (u + Ax - b)^+ - u \\ (Ex - f) \end{bmatrix}, \tag{۱۸.۲}$$

که در آن κ نرخ همگرایی است و

$$M = \begin{bmatrix} -Q & A^T & E^T \\ A & \circ_{m\times m} & \circ_{m\times l} \\ E & \circ_{l\times m} & \circ_{l\times l} \end{bmatrix}.$$

مدل نهم

مسأله برنامه‌ریزی درجه دوم زیر را در نظر بگیرید:

$$\min \quad f(x) = \frac{۱}{۲} x^T Ax + a^T x \tag{۱۹.۲}$$

$$\text{s.t.} \begin{cases} Dx \le b, \\ x \ge \circ. \end{cases} \tag{۲۰.۲}$$

دوگان آن عبارت است از:

$$\max \quad \bar{f}(w) = b^T w - \frac{۱}{۲} x^T Ax \tag{۲۱.۲}$$

$$\text{s.t.} \begin{cases} \bar{D}^T w - Ax \le a, \\ w \ge \circ. \end{cases} \tag{۲۲.۲}$$

که در آن $A \in \mathbb{R}^{n\times n}$ متقارن و نیمه معین مثبت است، $a \in \mathbb{R}^n$ و $b \in \mathbb{R}^m$ و $D \in \mathbb{R}^{m\times n}$ است. با توجه به شرایط شرایط کروش-کان-تاکر برای مسائل برنامه‌ریزی محدب، $(x^{*T}, w^{*T})^T$ به‌ترتیب یک جواب بهینه برای مسائل (۱۹.۲) - (۲۲.۲) است اگر و فقط اگر $(x^{*T}, w^{*T})^T$ در شرایط کروش-کان-تاکر زیر صدق کند:

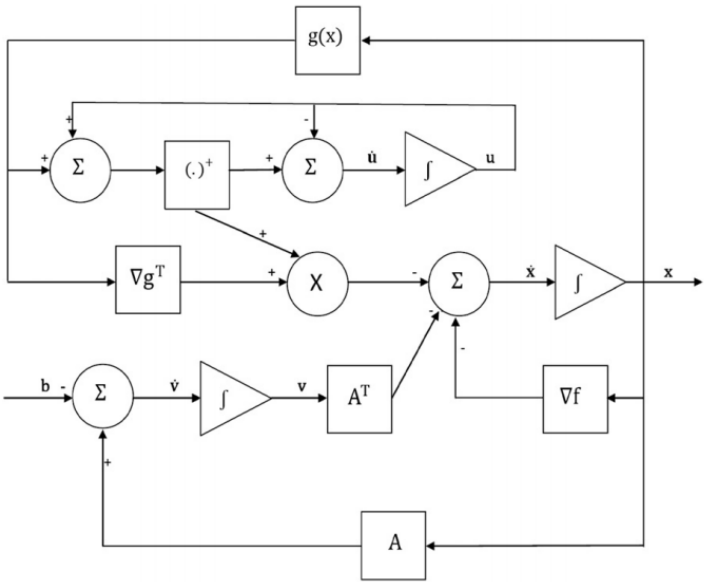
$$\begin{cases} w^* \ge \circ, Dx^* - b \le \circ, w^{*T}(Dx^* - b) = \circ, \\ Ax^* + a + D^T w^* \ge \circ, x \ge \circ, \\ x^{*T}(Ax^* + a + D^T w^*) = \circ. \end{cases}$$

مدل متناظر با (۱۹.۲) و (۲۰.۲) در [۲۹] به‌صورت زیر بیان شده است:

$$\dot{u} = B(I + M^T)\{(u - (Mu + q)^+ - u)\},$$

که در آن

$$u = \begin{bmatrix} x \\ w \end{bmatrix}, q = \begin{bmatrix} a \\ b \end{bmatrix}, M = \begin{bmatrix} A & D^T \\ -D & \circ \end{bmatrix}.$$



شکل ۲.۲: دیاگرام مدل شبکه عصبی ۲۵.۲

M یک ماتریس نیمه معین مثبت است زیرا $u^TMu = x^TAx \geq 0$. همچنین $B = \lambda I$ که $\lambda > 0$ و با افزایش λ سرعت همگرایی افزایش می‌یابد.

مدل دهم [۲]

فرم کلی یک مساله‌ی برنامه‌ریزی غیر خطی محدب[۲۶] را به صورت زیر در نظر بگیرید:

$$\min \quad f(x) \tag{۲۳.۲}$$

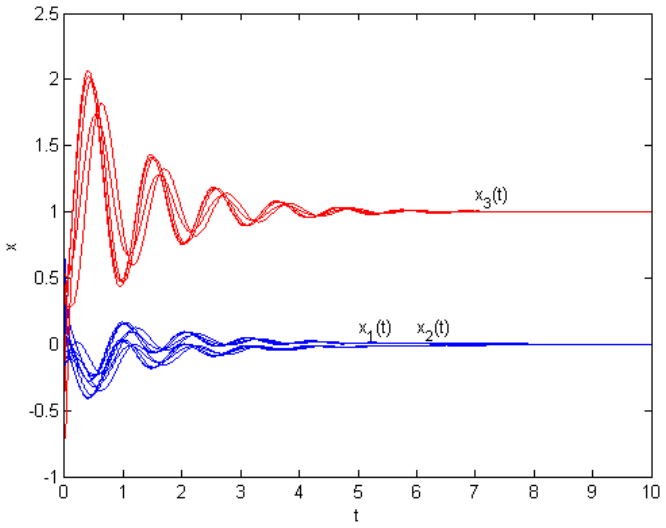
$$\text{s.t.} \begin{cases} g(x) \leq 0, \\ h(x) = 0, \end{cases} \tag{۲۴.۲}$$

که در آن $x \in \mathbb{R}^n$ ، $f : \mathbb{R}^n \to \mathbb{R}^n$ ، $g(x) = (g_1(x), g_2(x), \ldots, g_m(x))^T$ یک تابع برداری m-بعدی است. توابع $f, g_1, g_2, \ldots, g_m$ محدب و دو بار مشتق‌پذیر فرض شده‌اند. $h(x) = Ax - b$، $A \in \mathbb{R}^{l \times n}$، $\text{rank}(A) = l, 0 \leq l < n$ و $b \in \mathbb{R}^l$. در اینجا فرض می‌کنیم که مساله (۲۳.۲) و (۲۴.۲) یک جواب بهینه دارد. متناظر با مساله (۲۳.۲) و (۲۴.۲) شبکه عصبی زیر را معرفی می‌کنیم:

$$\begin{cases} \frac{dx}{dt} = -(\nabla f(x) + \nabla g(x)^T(u + g(x))^+ + \nabla h(x)^T v), \\ \frac{du}{dt} = (u + g(x))^+ - u, \\ \frac{dv}{dt} = h(x), \end{cases} \tag{۲۵.۲}$$

با نقطه‌ی آغازی $y_0 = (x_0^T, u_0^T, v_0^T)^T$ شکل (۲.۲) معماری این مدل را نشان می دهد.

[۲۶] General convex nonlinear programming



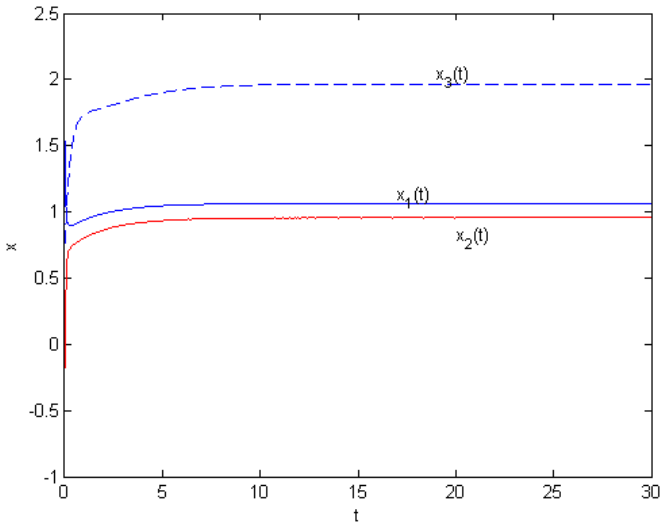
شکل ۳.۲: مسیر جواب در مثال ۳.۳.۲

مثال ۳.۳.۲.

$$\min \ (x_۱ + ۳x_۲ + x_۳)^۲ + ۴(x_۱ - x_۲)^۲$$

$$\text{s.t.} \begin{cases} x_۱^۳ - ۶x_۲ - ۴x_۳ \leq ۰, \\ x_۱ + x_۲ + x_۳ - ۱ = ۰, \\ x_۱, x_۲, x_۳ \geq ۰. \end{cases}$$

جواب بهینه $x^* = (۰, ۰, ۱)^T$ است. مسیرهای جواب در شکل (۳.۲) نشان داده شده‌اند.



شکل ۴.۲: مسیر جواب در مثال ۴.۳.۲

مثال ۴.۳.۲.

$$\min \ \frac{۱}{۲}((x_۱ - x_۲)^۴ + (x_۲ + x_۳)^۲ + (x_۱ + x_۳)^۲)$$

$$\text{s.t.} \begin{cases} x_۱^۲ + x_۲^۴ - x_۳ \leq ۰, \\ (۲ - x_۱)^۲ + (۲ - x_۲)^۲ - x_۳ \leq ۰, \\ ۲e^{-x_۱ + x_۲} - x_۳ \leq ۰, \\ x_۱^۲ + x_۲^۲ - ۲x_۱ + x_۲ - ۴ \leq ۰, \\ |x_۱| \leq ۲, |x_۲| \leq ۲, x_۳ \geq ۰, \end{cases}$$

جواب بهینه $x^* = (۱/۱۰۹, ۰/۹۲۲, ۱/۹۵۴)^T$ است. مسیرهای جواب در شکل (۴.۲) نشان داده شده‌اند.

مدل یازدهم [۱]

مساله برنامه‌ریزی غیر خطی محدب[۲۷] زیر را در نظر بگیرید:

$$\min \ f(x) \tag{۲۶.۲}$$

$$\text{s.t.} \begin{cases} g(x) \leq ۰, \\ h(x) = ۰. \end{cases} \tag{۲۷.۲}$$

[۲۷] Convex nonlinear programming

که در آن $x \in \mathbb{R}^n$ ، $f : \mathbb{R}^n \to \mathbb{R}^n$، $g(x) = (g_1(x), g_2(x), \ldots, g_m(x))^T$ یک تابع برداری m-بعدی است. توابع $f, g_1, g_2, \ldots, g_m$ محدب و دو بار مشتق‌پذیر فرض شده‌اند. $h(x) = Ax - b$، $A \in \mathbb{R}^{l \times n}$، $\text{rank}(A) = l, 0 \leq l < n$ و $b \in \mathbb{R}^l$. ابتدا تابع لاگرانژ زیر را در نظر می‌گیریم:

$$L(x, u, v) = f(x) + \frac{1}{2} \sum_{k=1}^{m} u_k^2 g_k(x) + \sum_{p=1}^{l} v_p h_p(x)$$

که در آن $u \in \mathbb{R}^m$ و $v \in \mathbb{R}^l$ را ضرایب لاگرانژ می‌نامند. $x^* \in \mathbb{R}^n$ یک جواب بهینه (۲۶.۲) و (۲۷.۲) اگر و فقط اگر $u^* \in \mathbb{R}^m$ و $v^* \in \mathbb{R}^l$ وجود داشته باشد به‌طوری‌که (x^{*T}, u^{*T}, v^{*T}) در شرایط کروش-کان-تاکر زیر صدق کند:

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

برای حل مساله‌ی (۲۶.۲) و (۲۷.۲) از سیستم دینامیکی زیر استفاده می‌کنیم:

$$\begin{cases} \frac{dx_i}{dt} = -\nabla_{x_i} L(x, u, v) = -(\frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{k=1}^{m} u_k^2 \frac{\partial g_k}{\partial x_i} + \sum_{p=1}^{l} v_p \frac{\partial h_p}{\partial x_i}), \ i = 1, \ldots, n, \\ \frac{du}{dt} = \nabla_u L(x, u, v) = diag(u_1, \ldots, u_m) g(x), \\ \frac{dv_p}{dt} = \nabla_{v_p} L(x, u, v) = h_p(x), \ i = 1, \ldots, l, \end{cases} \quad (۲۸.۲)$$

با یک نقطه آغازین $(x_0^T, u_0^T, v_0^T)^T$ و $u_k(t_0) > 0 (k = 1, 2, \ldots, m)$.

مثال ۵.۳.۲.

$$\min \ x_1^2 + x_2^2 + 2x_1 x_2 + (x_1 - 1)^4 + (x_2 - 3)^4$$

$$\text{s.t.} \begin{cases} x_1^2 + x_2^2 - 64 \leq 0, \\ (x_1 + 3)^2 + (x_2 + 4)^2 - 36 \leq 0, \\ (x_1 - 3)^2 + (x_2 + 4)^2 - 36 \leq 0, \\ x_1 \geq 0, \ x_2 \geq 0, \end{cases}$$

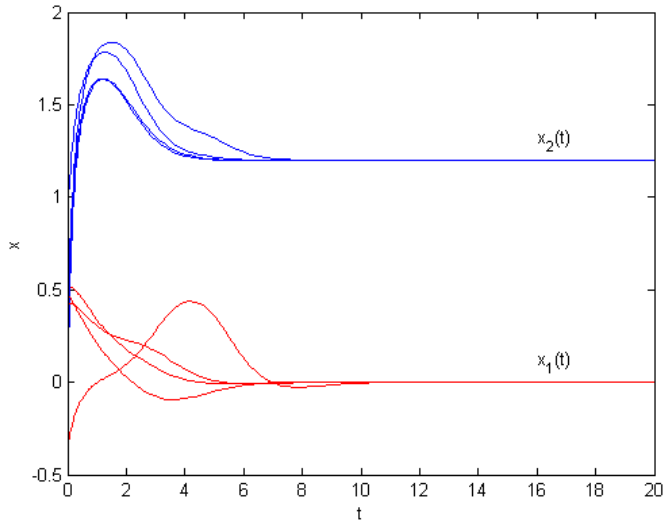
جواب بهینه $x^* = (0, 1/96)^T$ است و مسیرهای خروجی $x(t)$ در شکل (۵.۲) نشان داده شده است.

مثال ۶.۳.۲.

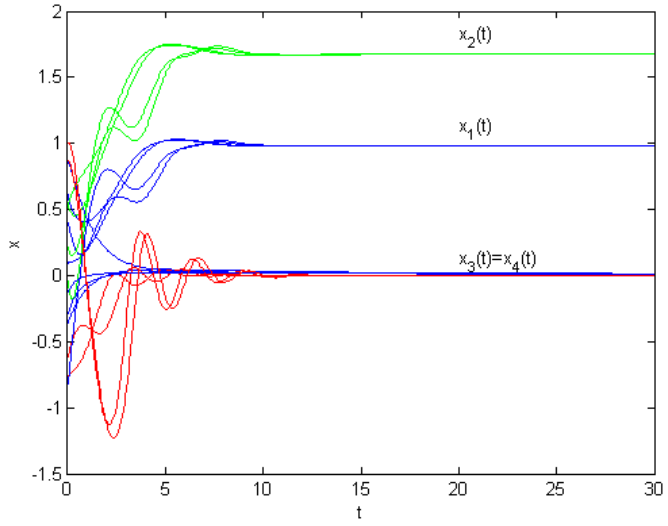
$$\min \ 0/4x_1 + x_1^2 + x_2^2 - x_1 x_2 + 0/5x_3^2 + 0/5x_4^2 + \frac{1}{30} x_1^3$$

$$\text{s.t.} \begin{cases} -x_1 + x_2 - x_3 \leq 2, \\ 3x_1 + x_2 - x_3 - x_4 \leq 18, \\ \frac{1}{3}x_1 + x_2 - x_4 = 2, \\ x \geq 0, \end{cases}$$

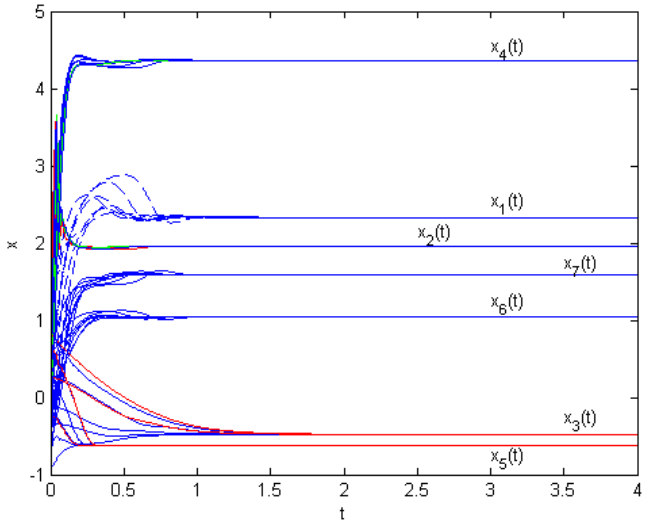
جواب بهینه $x^* = (0/982, 1/672, 0, 0)^T$ است که همگرایی مسیرها در شکل (۶.۲) نشان داده شده است.



شکل ۵.۲: مسیر جواب در مثال ۵.۳.۲



شکل ۶.۲: مسیر جواب در مثال ۶.۳.۲



شکل ۷.۲: مسیر جواب در مثال ۷.۳.۲

مثال ۷.۳.۲.

$$\min\ (x_۱-۱۰)^۲+۵(x_۲-۱۲)^۲+x_۳^۴+۳(x_۴-۱۱)^۲+۱۰x_۵^۶+۷x_۶^۲+x_۷^۴-۴x_۶x_۷-۱۰x_۶-۸x_۷$$

$$\text{s.t.}\begin{cases} ۲x_۱^۲+۳x_۲^۴+x_۳+۴x_۴^۲+۵x_۵-۱۲۷\le ۰,\\ ۷x_۱+۳x_۲+۱۰x_۳^۲+x_۴-x_۵-۲۸۲\le ۰,\\ ۲۳x_۱+x_۲^۲+۶x_۶^۲-۸x_۷-۱۹۶\le ۰,\\ ۴x_۱^۲+x_۲^۲+۲x_۳^۲-۳x_۱x_۲+۵x_۶-۱۱x_۷\le ۰, \end{cases}$$

نقطه بهینه به صورت $x^* = (۲/۳۳۰, ۱/۹۵۱, -۰/۴۷۷۵, ۴/۳۶۶, -۰/۶۲۵, ۱/۰۳۸, ۱/۵۹۴)^T$ می‌باشد.شکل (۷.۲) مسیرهای جواب را نمایش می دهد.

فصل ۳

یک روش منظم سازی آنتروپی برای حل مسائل مین-ماکس با برخی کاربردها

۱.۳ مقدمه

برخی مسائل علمی و مهندسی می توانندبا مساله مین - ماکس [۱] بصورت زیر فرموله شوند[۱۰]:

$$\min_{x\in X} \max_{1\le i\le m}\{f_i(x)\}, \tag{۱.۳}$$

که در آن Xیک دامنه مشترک از مولفه توابع $f_i(x)$ $(i = 1,\dots,m)$، است وf_i ها دوبار مشتق پذیر هستند. یک مشکل مهم مشتق پذیری تابع زیر است:

$$F(x) \equiv \max_{1\le i\le m}\{f_i(x)\}, \qquad x \in X. \tag{۲.۳}$$

توجه داشته باشید که مساله مین- ماکس (۱.۳) دارای تابع لاگرانژ[۲]زیر است:

$$L(x,\lambda) = \sum_{i=1}^{m} \lambda_i f_i(x)\ , \tag{۳.۳}$$

برای هر$x \in X$ که در آن m بردار و λ نشان دهنده ضرایب لاگرانژ هستند که در شرایط نامنفی و نرمال بودن صدق می کنند، یعنی این ضرایب در مجموعه زیر قرار می گیرند:

$$\Lambda \equiv \{\lambda | \lambda \ge 0;\ \sum_{i=1}^{m} \lambda_i = 1\}. \tag{۴.۳}$$

با در نظر گرفتن برخی مفروضات منظم سازی، تابع ماکزیمم $F(x)$ را می توان به صورت زیربه دست آورد:

$$F(x) = \max_{\lambda\in\Lambda} L(x,\lambda), \qquad x \in X. \tag{۵.۳}$$

---

[۱] Min-max

[۲] Lagrange

متاسفانه، $\max \sum_{i=۱}^{m} \lambda_i f_i(x)$ روی $\lambda \in \Lambda$، به ندرت یک راه حل صریح و روشن برای بدست آوردن $\lambda(x)$ ارائه می دهد. در این فصل ما یک روش منظم سازی با جایگزین کردن تابع $L_p(x,\lambda)$ به جای $L(x,\lambda)$ به صورت زیر در نظر میگیریم:

$$L_p(x,\lambda) = \sum_{i=۱}^{m} \lambda_i f_i(x) + \frac{۱}{p} R(\lambda;\beta), \qquad (۶.۳)$$

که در آن $p > ۰$ پارامتر کنترل است،R یک تابع تنظیم است، و β بردار پارامتر اختیاری از R است. با انتخاب دقیق تابع R، ماکزیمم تابع (۶.۳) را در نظر می گیریم:

$$F_p(x) \equiv \max_{\lambda \in \Lambda} L_p(x,\lambda). \qquad (۷.۳)$$

مساله (۷.۳) یک تقریب برای مساله (۵.۳) است.

برخی از روش های منظم سازی شناخته شده از توابع زیر استفاده می کنند [۶]،[۱۸]:

$$R(\lambda;\beta) = -\frac{۱}{۲} \sum_{i=۱}^{m} (\lambda_i - \beta_i)^۲.$$

$$R(\lambda;\beta) = \sum_{i=۱}^{m} l_n \lambda_i.$$

در قسمت بعد، یک روش منظم سازی با استفاده از توابع آنتروپی[۳]معرفی شده است. با این روش منظم سازی یک تقریب خوب و هموار $F_p(x)$ برای تابع $F(x)$ در (۲.۳) بدست می آید.

۲.۳ روش منظم سازی آنتروپی

۱.۲.۳ تابع آنتروپی شانون

فرض کنید $x \in X$ داده شده است، هنگامی که تابع آنتروپی شانون[۴] [۸] به عنوان تابع تنظیم به صورت $R(\lambda) = -\sum_{i=۱}^{m} (\ln \lambda_i)\lambda_i$، در نظر گرفته شود، تعریف می کنیم:

$$L_p(x,\lambda) \equiv \sum_{i=۱}^{m} \lambda_i f_i(x) - \frac{۱}{p} \sum_{i=۱}^{m} (\ln \lambda_i)\lambda_i, \qquad (۸.۳)$$

که در آن $\lambda \in \Lambda$. چون $L_p(x,\lambda)$ در λاکیدا مقعر است، با ماکزیمم کردن آن روی Λ جواب منحصر به فرد بهینه زیر حاصل می شود:

$$\lambda_i^*(x,p) = \frac{\exp[p f_i(x)]}{Z}\ , \quad ۱ \le i \le m\ , \qquad (۹.۳)$$

که در آن

$$Z = \sum_{i=۱}^{m} \exp[p f_i(x)]\ . \qquad (۱۰.۳)$$

---

[۳] Entropic

[۴] Shannon's

با جایگزین کردن λ_i در (۸.۳) با $\lambda_i^*(x,p)$ داریم:

$$F_p(x) = \frac{1}{p}\ln\{\sum_{i=1}^{m}\exp[pf_i(x)]\}, \tag{۱۱.۳}$$

که یک تابع هموار [۵] است. لازم به تاکید است که تابع $F_p(x)$ بدست آمده در اینجا معادل با ماکزیمم کردن $L_p(x,\lambda)$ روی Λ برای هر $x \in X$ است.

۲.۲.۳ خواص $F_p(x)$

لم ۱.۲.۳. *$F_p(x)$ در X به صورت نقطه به نقطه به $F(x)$ همگرا [۶] است، زمانی‌که $p \to +\infty$.*

برهان. فرض کنید $x \in X$ و $V(x)$ را به صورت یک تابع برداری با مولفه

$$v_i(x) = \exp[f_i(x)],\ 1 \le i \le m,$$

تعریف کنید در این صورت نرم- l_p از $V(x)$ به صورت زیر بدست می آید:

$$\|V(x)\|_p \equiv \left\{\sum_{i=1}^{m}[v_i(x)]^p\right\}^{\frac{1}{p}} = \left\{\sum_{i=1}^{m}\exp[pf_i(x)]\right\}^{\frac{1}{p}}. \tag{۱۲.۳}$$

از اینرو

$$F_p(x) = \ln\|V(x)\|_p \tag{۱۳.۳}$$

بنابراین

$$\lim_{p\to\infty}\|V(x)\|_p = \max_{1\le i\le m} v_i(x) = \exp[F(x)], \tag{۱۴.۳}$$

و

$$\lim_{p\to\infty} F_p(x) = F(x). \tag{۱۵.۳}$$

□

لم ۲.۲.۳. *$F_p(x)$ در X به صورت یکنواخت به $F(x)$ همگرا است، زمانی‌که $p \to +\infty$.*

برهان. با توجه به خواص نرم l_p برای $x \in X$، $F_p(x)$ یک تابع نزولی یکنواخت در p است، یعنی:

$$F_s(x) \le F_r(x)\ , \quad r \le s. \tag{۱۶.۳}$$

---

[۵] Smooth

[۶] Converges

از این رو از لم (۱.۲.۳)، برای هر $p > ۰$ داریم:

$$۰ \leq F_p(x) - F(x) = \ln \|V(x)\|_p - \ln\{\exp F(x)\} \tag{۱۷.۳}$$

$$= (\frac{۱}{p}) \ln \sum_{i=۱}^{m} \exp\{p[f_i(x) - F(x)]\} \tag{۱۸.۳}$$

$$\leq (\frac{۱}{p}) \ln(m). \tag{۱۹.۳}$$

به عبارت دیگر برای هر $p > ۰$

$$F(x) \leq F_p(x) \leq F(x) + \frac{\ln(m)}{p}, \quad \forall x \in X. \tag{۲۰.۳}$$

از این رو $F_p(x)$ همگرا به $F(x)$ به طوریکنواخت در X است، وقتی $p \to +\infty$. □

لم ۳.۲.۳. *$F_p(x)$ محدب است،اگر به ازای $i = ۱, ۲, \ldots, m$, $f_i(x)$ یک تابع محدب، تعریف شده روی یک مجموعه محدب X باشد.*

برهان. برای هر $x, y \in X$ و $\lambda \in (۰, ۱)$ داریم:

$$F_p(\lambda x + (۱ - \lambda)y) = \frac{۱}{p} \ln \sum_{i=۱}^{m} \exp[p f_i(\lambda x + (۱ - \lambda)y)]$$

$$\leq \frac{۱}{p} \ln \sum_{i=۱}^{m} \exp[\lambda p f_i(x) + (۱ - \lambda) p f_i(y)]$$

$$= \frac{۱}{p} \ln \sum_{i=۱}^{m} \{\exp[p f_i(x)]\}^{\lambda} \{\exp[p f_i(y)]\}^{۱-\lambda}. \tag{۲۱.۳}$$

نامساوی هلدر[۷] [۹] نشان می دهد که

$$\sum_{i=۱}^{m} \{\exp[p f_i(x)]\}^{\lambda} \{\exp[p f_i(y)]\}^{۱-\lambda} \leq \left\{\sum_{i=۱}^{m} \exp[p f_i(x)]\right\}^{\lambda} \left\{\sum_{i=۱}^{m} \exp[p f_i(y)]\right\}^{۱-\lambda}. \tag{۲۲.۳}$$

در نتیجه

$$F_p(\lambda x + (۱ - \lambda)y) \leq \frac{\lambda}{p} \ln \sum_{i=۱}^{m} \exp[p f_i(x)] + \frac{۱ - \lambda}{p} \ln \sum_{i=۱}^{m} \exp[p f_i(y)]$$

$$= \lambda F_p(x) + (۱ - \lambda) F_p(y). \tag{۲۳.۳}$$

این نشان می دهد که $F_p(x)$ یک تابع محدب روی X است. □

به طور خلاصه ما تمام نتایج را در قضیه اصلی زیر داریم:

قضیه ۴.۲.۳. *فرض کنید $X = \mathbb{R}^n$ و $f_i(x)$ محدب است، برای $i = ۱, ۲, \ldots, m$. یک جواب بهینه تقریبی ازمساله مین - ماکس (۱.۳) را با حل یک مساله برنامه ریزی محدب نا مقید $\{\min_x F_p(x)\}$ با p به اندازه کافی بزرگ بدست می آید.*

[۷] Holder

۳.۳ کاربردهای مرتبط

در این قسمت برخی کاربردها از آنالیز عددی و مسائل بهینه سازی معادل مساله را بررسی میکنیم.

۱.۳.۳ سیستم های نامساوی و مساوی

یک سیستم نامساوی در $\mathbb{R}^n$ را به صورت زیر درنظر بگیرید:

$$f_i(x) \leq ۰, \quad i = ۱, ۲, \ldots, m. \tag{۲۴.۳}$$

تعریف می کنیم:

$$F(x) \equiv \max_{۱\leq i\leq m} \{f_i(x)\}. \tag{۲۵.۳}$$

می دانیم که سیستم (۲۴.۳)جواب دارد اگر و تنها اگر $F(x)$ یک مینیمم کننده x^* داشته باشد آنچنانکه $F(x^*) \leq ۰$. با استفاده از $F_p(x)$ بجای $F(x)$، حل مساله (۲۴.۳) معادل با حل مساله مینیمم سازی نامقید با تابع هدف هموار $F_p(x)$ است. همچنین داریم:

$$\nabla F_p(x) = J(x)\lambda, \tag{۲۶.۳}$$

و

$$\nabla^۲ F_p(x) = \sum_{i=۱}^{m} \lambda_i \nabla^۲ f_i(x) + pJ(x)(\Gamma - \lambda\lambda^T)J(x)^T, \tag{۲۷.۳}$$

که در آن $J(x)$ ماتریس ژاکوبین $[\partial f_i/\partial x_j]$، λ یک بردار با مولفه های $\lambda_i = \exp[pf_i(x)]\{\sum_{i=۱}^{m} \exp[pf_i(x)]\}^{-۱}$ است و $\Gamma = diag(\lambda_i)$.

برای حل یک سیستم نا مساوی خطی به صورت

$$f_i(x) \equiv A_i(x) - b_i \leq ۰, \quad i = ۱, \ldots, m, \tag{۲۸.۳}$$

تابع آنتروپی متناظر بصورت زیر در نظر گرفته می شود:

$$F_p(x) = \frac{۱}{p} \ln\left\{\sum_{i=۱}^{m} \exp[p(A_i x - b_i)\right\}. \tag{۲۹.۳}$$

گرادیان[۸] و هسین[۹]بصورت زیر تعریف می شود:

$$\nabla F_p(x) = A^T\lambda, \tag{۳۰.۳}$$

و

$$\nabla^۲ F_p(x) = pA^T(\Gamma - \lambda\lambda^T)A, \tag{۳۱.۳}$$

که در آن

$$\lambda_i = \frac{\exp[p(A_i x - b_i)]}{\sum_{i=۱}^{m} \exp[p(A_i x - b_i)]}. \tag{۳۲.۳}$$

[۸] Gradient

[۹] Hessian

برای حل دستگاه معادلات تعریف شده توسط

$$f_i(x) = ۰, \quad i = ۱, ۲, \ldots, m, \tag{۳۳.۳}$$

ما می توانیم هر معادله را توسط دو نامعادله $f_i(x) \leq ۰$ و $-f_i(x) \leq ۰$ یا با یک نامعادله منحصر به فرد $f_i^۲(x) \leq ۰$ جایگزین کنیم. البته اولی افزایش ابعاد مساله و دومی پیچیده گی توابع غیر خطی ایجاد شده را در بر دارد.

۲.۳.۳ حل مسائل برنامه ریزی خطی

مساله برنامه ریزی خطی زیر را در نظر بگیرید:

$$\min_x f(x) \equiv c^T x \tag{۳۴.۳}$$

$$\text{s.t.} \quad Ax = b, \tag{۳۵.۳}$$

$$e^T x = M, \tag{۳۶.۳}$$

$$x \geq ۰, \tag{۳۷.۳}$$

که در آن e یک بردار یکه و M یک ثابت است. دوگان خطی (۳۴.۳) - (۳۷.۳) به صورت زیر تعریف می شود:

$$\min_y d(y) \equiv -b^T y - M y_{m+۱} \tag{۳۸.۳}$$

$$\text{s.t.} \quad \sum_{i=۱}^{m} a_{ij} y_i - c_j \leq -y_{m+۱}\ , \quad j = ۱, ۲, \ldots, n. \tag{۳۹.۳}$$

به وضوح، این مساله معادل با مساله مینیم سازی بدون محدودیت زیر است:

$$\min_y \left\{ F(y) \equiv -b^T y + M \ \max_{۱\leq j\leq n} \left\{ \sum_{i=۱}^{m} a_{ij} y_i - c_j \right\} \right\}. \tag{۴۰.۳}$$

یادآوری می کنیم که فرم استاندارد برنامه ریزی خطی کارماکار[۱۰] [۴] با فرض $b = ۱$ و $M = ۱$ در نظر گرفته می شود. در این مورد، دوگان آن یک مساله مین - ماکس به صورت زیر است:

$$\min_y \left\{ F(y) \equiv \max_{۱\leq j\leq n} \left\{ \sum_{i=۱}^{m} a_{ij} y_i - c_j \right\} \right\}. \tag{۴۱.۳}$$

روش منظم سازی آنتروپی برای حل مساله (۴۱.۳) منجر به مساله زیر می شود:

$$\min_y \left\{ F_p(y) \equiv \frac{۱}{p} \sum_{j=۱}^{n} \exp\left[p \left(\sum_{i=۱}^{m} a_{ij} y_i - c_j \right) \right] \right\}. \tag{۴۲.۳}$$

[۱۰] Karmarkar's

۳.۳.۳ حل مسائل مین - ماکس مقید

مساله مین - ماکس مقید زیر را درنظر بگیرید:

$$\min_x \left\{ f(x) \equiv \max_{۱\le i\le m} \{f_i(x)\} \right\} \tag{۴۳.۳}$$

$$\text{s.t.} \quad g_j(x) \le ۰, \quad j = ۱, ۲, \ldots, l. \tag{۴۴.۳}$$

توجه داشته باشید که زمانی که $m = ۱$، با یک مساله برنامه ریزی غیر خطی با l محدودیت مواجه هستیم. مساله (۴۳.۳) و (۴۴.۳) معادل با مساله زیر است:

$$\min_{x,\alpha} \alpha \tag{۴۵.۳}$$

$$\text{s.t.} \quad f_i(x) - \alpha \le ۰, \quad ۱ \le i \le m, \tag{۴۶.۳}$$

$$g_j(x) \le ۰, \quad ۱ \le j \le l. \tag{۴۷.۳}$$

هنگامی که از روش ذکر شده در [۷] استفاده می شود، کار اصلی در هر تکرار عبارت است از پیدا کردن مرکز یک مجموعه تراز بصورت زیر

$$G_\alpha(x) \equiv \{x | f_i(x) - \alpha \le ۰,\ ۱ \le i \le m;\ g_j(x) \le ۰,\ ۱ \le j \le l\}, \tag{۴۸.۳}$$

با تعریف یک تابع فاصله

$$d(x) \equiv \min\{-f_i(x) + \alpha,\ ۱ \le i \le m;\ -g_j(x),\ ۱ \le j \le l\}, \tag{۴۹.۳}$$

مرکز $G_\alpha(x)$ را می توان با ماکزیمم کردن $d(x)$ یا به طور معادل با مینیمم کردن $-d(x)$ از طریق تقریب هموار آن بصورت زیر پیدا کرد:

$$d_p(x) = \frac{۱}{p} \ln\left\{ \sum_{i=۱}^{m} \exp[p(f_i(x) - \alpha)] + \sum_{j=۱}^{l} \exp[p g_j(x)] \right\}. \tag{۵۰.۳}$$

فصل ۴

یک مدل شبکه عصبی تصویری برای حل رده ای از مسائل بهینه سازی مشتق ناپذیر

۱.۴ مقدمه

در این فصل مابه تشریح مسائل بهینه سازی مشتق ناپذیر[۱] به صورت زیر میپردازیم

$$\min_{x\in\Omega} f(x) = \max_{۱\le i\le s}\{f_i(x)\} \tag{۱.۴}$$

$$\text{s.t.} \quad h_j(x) \le ۰, \quad j = ۱, ۲, \ldots, t \tag{۲.۴}$$

که در آن

$$\Omega = \{x \in \mathbb{R}^n | l_i \le x_i \le h_i,\ i = ۱, ۲, \ldots, n\},$$

$$f_i(x)(i = ۱, ۲, \ldots, s), h_j(x)(j = ۱, ۲, \ldots, t)$$

توابع محدب دوبار مشتق پذیر هستند. هرچند توابع $f_i(x)(i = ۱, ۲, \ldots, s)$ به طور پیوسته مشتق پذیر هستند ولی $f(x) = \max_{۱\le i\le s}\{f_i(x)\}$

یک تابع ناهموار[۲] است. مساله (۱.۴) و (۲.۴)، بنابراین به عنوان یک مساله بهینه سازی ناهموار طبقه بندی میشود. این نوع مسائل بهینه سازی دربسیاری از زمینه های مختلف در ریاضیات مهندسی و مسائل مالی مطرح میشود ([۱۵]-[۱۱]) در بسیاری از مقالات قبلی این مسائل را با الگوریتم های عددی حل میکنند ([۲۲]-[۱۳]). از آنجا که زمان محاسبات تا حد زیادی به ابعاد وساختار مسائل بستگی دارد الگوریتم های عددی معمولا در مقیاس بزرگ و برای مسائل بهینه سازی زمان واقعی، کمتر موثر است. شبکه های عصبی با محاسبات موازی و همگرایی سریع می توانند به عنوان یک روش

[۱] Non-differentiable

[۲] non-smooth

کارامد برای حل مسائل با مقیاس بزرگ و یا مسائل بهینه سازی زمان واقعی در نظر گرفته شوند. به تازگی، شبکه های عصبی برای حل مسائل بهینه سازی به طور گسترده مورد مطاله قرار گرفته و برخی نتایج مهم نیز بدست آمده است. ([۳۲]-[۲۳]) به عنوان مثال، چن[۳] و همکاران، [۲۶]،شیا[۴] و همکاران[۲۹، ۲۸]و گائو[۵] و همکاران[۳۰]، شبکه های عصبی را برای حل مسائل برنامه ریزی محدب مشتق پذیر بر اساس روش اولیه - دوگان و عملگر تصویر بکار برده اند. در این فصل ابتدا مساله (۱.۴) و (۲.۴) به یک برنامه ریزی غیر خطی محدب با محدودیت های نامساوی تبدیل می شود و بر اساس شرایط بهینگی کروش-کان-تاکر (KKT) برای بهینگی برنامه ریزی محدب، یک مدل شبکه عصبی برای حل مساله بدست آمده پیشنهاد خواهد شد. با ساختن یک تابع لیاپانف مناسب، پایداری مجانبی وهمگرایی سراسری شبکه اثبات می شود.

۲.۴ مدل شبکه عصبی

یک روش مناسب برای تبدیل مساله(۱.۴) و (۲.۴) بصورت زیر است:

$$\min_{x\in\Omega} z$$

$$\text{s.t.} \begin{cases} f_i(x) \le z, & i = ۱, ۲, \dots, s, \\ h_j(x) \le ۰ & j = ۱, ۲, \dots, m. \end{cases} \tag{۳.۴}$$

واضح است که این تبدیل منجر به برنامه ریزی محدب مشتق پذیر با تعداد محدودیت ها وپیچیدگی محاسبات بالا میشود.

به تازگی، یک روش تقریبی برای حل (۱.۴) و (۲.۴) ارائه شده است. با استفاده از تابع آنتروپی معرفی شده در فصل ۳ مساله زیر را در نظر میگیریم:

$$\min_{x\in\Omega} F_p(x) = \frac{۱}{p} \ln \sum_{i=۱}^{s} \exp(p f_i(x))$$

$$\text{s.t.} \quad h_j(x) \le ۰ \quad j = ۱, ۲, \dots, m, \ p \to +\infty. \tag{۴.۴}$$

از تجزیه وتحلیل بالا، واضح است که مساله مشتق ناپذیر (۱.۴) و (۲.۴) معادل یک مساله برنامه ریزی محدب است. بنابراین مساله برنامه ریزی غیرخطی محدب زیر را با محدودیت های نامساوی غیر خطی بصورت زیر در نظر می گیریم:

$$\begin{aligned} \min_{x\in\Omega} \quad & F(x) \\ \text{s.t.} \quad & g_j(x) \le ۰, \quad j = ۱, ۲, \dots, m, \end{aligned} \tag{۵.۴}$$

---

[۳] Chen

[۴] Xia

[۵] Gao

که $(j = ۱, ۲, \ldots, m), g_j(x), F(x)$ محدب وبه طورپیوسته مشتق پذیر هستند.
فرض کنید

$$S = \{x \in \mathbb{R}^n : \ g_j(x) \leq ۰; \ j = ۱, ۲, \ldots, m\}$$

به خوبی می دانیم $x^* \in S$ جواب بهینه (۵.۴) است اگر و تنها اگر
$\lambda^* = (\lambda_۱^*, \lambda_۲^*, \ldots, \lambda_m^*)^T \in \mathbb{R}^m$وجود داشته باشد به طوریکه $(x^{*T}, \lambda^{*T})^T$در شرایط کروش-کان-تاکر زیر صدق کند:

$$\begin{cases} (x - x^*)^T(\nabla F(x^*) + \nabla g(x^*)^T\lambda^*) \geq ۰, & \forall x \in \Omega, \\ \lambda_j^* g_j(x^*) = ۰, \quad \lambda_j^* \geq ۰, \quad g_j(x^*) \leq ۰, & j = ۱, ..., m. \end{cases} \tag{۶.۴}$$

لم ۱.۲.۴. *[۲۸] x^*جواب بهینه (۵.۴) است اگر و تنها اگر وجود داشته باشد $\lambda^* = (\lambda_۱^*, \lambda_۲^*, \ldots, \lambda_m^*)^T \in \mathbb{R}^m$ به طوریکه:*

$$\begin{cases} x^* - P_\Omega[x^* - \nabla F(x^*) - \nabla g(x^*)^T\lambda^*] = ۰, \\ \lambda^* - [\lambda^* + g(x^*)]^+ = ۰. \end{cases} \tag{۷.۴}$$

که

$$P_\Omega(x_i) = \begin{cases} l_i, & x_i < l_i, \\ x_i, & l_i \leq x_i \leq h_i, \\ h_i, & x_i > h_i. \end{cases} \qquad (x_i)^+ = \begin{cases} x_i, & x_i \geq ۰, \\ ۰, & x_i < ۰. \end{cases}$$

لم ۱.۲.۴ نشان می دهد که جواب بهینه (۵.۴) باحل (۷.۴) به دست می‌آید.

لم ۲.۲.۴. *[۳۳] فرض کنید مجموعه $\Omega \subset \mathbb{R}^n$ یک مجموعه محدب بسته است. آنگاه*

$$(v - P_\Omega(v))^T(P_\Omega(v) - u) \geq ۰, \quad v \in \mathbb{R}^n, u \in \Omega, \tag{۸.۴}$$

$$\|P_\Omega(u) - P_\Omega(v)\| \leq \|u - v\|, \quad u, v \in \mathbb{R}^n. \tag{۹.۴}$$

شبکه عصبی تصویر فریز[۶] برای حل (۵.۴) به صورت زیر بیان می شود:

$$\begin{cases} \frac{dx}{dt} = -(x - P_\Omega[x - \nabla F(x) - \nabla g(x)^T\lambda]), \\ \frac{d\lambda}{dt} = -(\lambda - [\lambda + g(x)]^+). \end{cases} \tag{۱۰.۴}$$

متأسفانه شبکه عصبی (۱۰.۴) برای برخی مسائل برنامه ریزی محدب ناپایدار است [۲۸]. برای سادگی

[۶] Friesz

فرض می کنیم:

$$
\begin{aligned}
&u(t) = (x^T, \lambda^T)^T \in \mathbb{R}^{n+m},\\
&D = \Omega \times \mathbb{R}^m_+,\\
&\bar{\lambda} = [\lambda + g(x)]^+,\\
&\bar{x} = P_\Omega[x - \nabla F(x) - \nabla g(x)^T \lambda],
\end{aligned}
$$

و D^* رامجموعه نقاط بهینه (۵.۴) در نظر می گیریم.
براساس نتایج فوق مدل شبکه عصبی جدید زیر را معرفی می کنیم:

$$
\begin{cases}
\frac{dx}{dt} = -(x - P_\Omega[x - \nabla F(x) - \nabla g(x)^T \bar{\lambda}]),\\
\frac{d\lambda}{dt} = -(\lambda - [\lambda + g(x)]^+).
\end{cases}
\tag{۱۱.۴}
$$

راساس نتایج فوق مدل شبکه عصبی جدید زیر را معرفی می کنیم:

$$
\begin{cases}
\frac{dx}{dt} = -(x - P_\Omega[x - \nabla F(x) - \nabla g(x)^T \bar{\lambda}]),\\
\frac{d\lambda}{dt} = -(\lambda - [\lambda + g(x)]^+).
\end{cases}
\tag{۱۲.۴}
$$

فرض کنید D^eمجموعه نقاط تعادل شبکه عصبی (۱۲.۴) باشد. طبق لم ۱.۲.۴،$x \in D^e$یک نقطه تعادل شبکه عصبی (۱۲.۴) است اگر و تنها اگر$x \in D^*$ یک جواب بهینه (۵.۴) باشد. در این صورت $D^* = D^e$.

۳.۴ تجزیه وتحلیل پایداری

دراین بخش پایداری نقطه تعادل و همگرایی جواب بهینه شبکه عصبی (۱۲.۴) را مطالعه می کنیم.

قضیه ۱.۳.۴. *برای هر نقطه اولیه* $u(t_\circ) = (x(t_\circ)^T, \lambda(t_\circ)^T)^T \in \mathbb{R}^{n+m}$ *یک جواب پیوسته منحصر به فرد* $u(t) = (x(t)^T, \lambda(t)^T)^T$ *برای سیستم* (۱۲.۴) *وجود دارد. علاوه بر این* $x(t) \in \Omega$ *و* $\lambda(t_\circ) \geq \circ$.

برهان. $\nabla F(x)$ و $\nabla g_j(x)$ روی مجموعه محدب باز$D \subseteq R^{n+m}$ شامل $\Omega \times R^m_+$ به طور پیوسته مشتق پذیر هستند. پس $x - P_\Omega[x - \nabla F(x) - \nabla g(x)^T \lambda]$ و $\lambda - [\lambda + g(x)]^+$ به طور موضعی پیوسته لیپ شیتز هستند. طبق قضیه وجود و یکتایی جواب برای معادلات دیفرانسیل معمولی مسئله مقدار اولیه سیستم (۱۲.۴) یک جواب منحصربه فرد دارد.
فرض کنید نقطه اولیه $x_\circ = x(t_\circ) \in \Omega$ و $\lambda_\circ = \lambda(t_\circ) \geq \circ$داده شده است. چون:

$$
\begin{cases}
\frac{dx}{dt} + x = P_\Omega[x - \nabla F(x) - \nabla g(x)^T \bar{\lambda}],\\
\frac{d\lambda}{dt} + \lambda = [\lambda + g(x)]^+.
\end{cases}
\tag{۱۳.۴}
$$

درنتیجه

$$\begin{cases} \int_{t_\circ}^{t}(\frac{dx}{dt}+x)e^t dt = \int_{t_\circ}^{t} P_\Omega[x-\nabla F(x)-\nabla g(x)^T\bar{\lambda}]e^t dt, \\ \int_{t_\circ}^{t}(\frac{d\lambda}{dt}+\lambda)e^t dt = \int_{t_\circ}^{t}[\lambda+g(x)]^+ e^t dt. \end{cases} \tag{۱۴.۴}$$

آنگاه

$$\begin{cases} x(t) = e^{-(t-t_\circ)}x_\circ + e^{-t}\int_{t_\circ}^{t} e^t P_\Omega[x-\nabla F(x)-\nabla g(x)^T\bar{\lambda}]dt, \\ \lambda(t) = e^{-(t-t_\circ)}\lambda_\circ + e^{-t}\int_{t_\circ}^{t} e^t[\lambda+g(x)]^+ dt. \end{cases} \tag{۱۵.۴}$$

با استفاده از قضیه مقدار میانگین برای انتگرال ها داریم:

$$\begin{cases} x(t) = e^{-(t-t_\circ)}x_\circ + (۱-e^{-(t-t_\circ)})P_\Omega[\hat{x}-\nabla F(\hat{x})-\nabla g(\hat{x})^T\hat{\lambda}], \\ \lambda(t) = e^{-(t-t_\circ)}\lambda_\circ + (۱-e^{-(t-t_\circ)})[\hat{\lambda}+g(\hat{x})]^+. \end{cases} \tag{۱۶.۴}$$

چون $x(t_\circ)\in\Omega$ و $\lambda(t_\circ)\geq \circ$ درنتیجه $x(t)\in\Omega$ و $\lambda(t)\geq\circ$ و این اثبات را کامل می کند. □

ملاحظه ۲.۳.۴. *نتایج قضیه ۱.۳.۴ نشان می دهدکه شبکه عصبی (۱۲.۴) خوش تعریف است. دراثبات قضیه ۱.۳.۴ فرض کردیم $(x_\circ^T,\lambda_\circ^T)^T\in\Omega\times\mathbb{R}_+^m$ یعنی نقطه اولیه باید شدنی باشد.*
اما اگر $(x_\circ^T,\lambda_\circ^T)^T\notin\Omega\times\mathbb{R}_+^m$ ، وقتی $t\to+\infty$ $(x(t,x_\circ)^T,\lambda(t,x_\circ)^T)^T\in\Omega\times\mathbb{R}_+^m$ می شود. بنابراین یک لحظه $t_\circ$ وجود دارد به طوریکه به ازای $t>t_\circ$ ، $(x(t,x_\circ)^T,\lambda(t,x_\circ)^T)^T$ در ناحیه شدنی قرار می‌گیرد و $(x(t,x_\circ)^T,\lambda(t,x_\circ)^T)^T$ در $\Omega\times\mathbb{R}_+^m$ می ماند.

قضیه ۳.۳.۴. *فرض کنید $f(x)$ و $g_j(x)$ ، $j=۱,۲,\dots,m$ ، $x\in\mathbb{R}^n$ روی مجموعه محدب باز $N\subseteq\mathbb{R}^{n+m}$ شامل $\Omega\times\mathbb{R}_+^m$ محدب و مشتق پذیر باشند. آنگاه شبکه عصبی (۱۲.۴) به مفهوم لیاپانوف پایدار است و برای هر نقطه اولیه $(x(t_\circ)^T,\lambda(t_\circ)^T)^T\in\mathbb{R}^{n+m}$ مسیر جواب (۱۲.۴) به یک نقطه در D^* همگرا خواهد بود. به خصوص شبکه عصبی (۱۲.۴) هنگامی‌که D^* فقط یک نقطه دارد پایدار مجانبی سراسری است.*

برهان. با استفاده از قضیه ۱.۳.۴ ، $\forall(x_\circ^T,\lambda_\circ^T)^T\in\Omega\times\mathbb{R}_+^m$ یک جواب پیوسته منحصر به فرد $(x(t)^T,\lambda(t)^T)^T\subseteq\Omega\times\mathbb{R}_+^m$ برای سیستم (۱۲.۴) وجود دارد.
یک تابع لیاپانوف به صورت زیر تعریف می کنیم:

$$\begin{aligned} V(x,\lambda) =& F(x)-F(x^*)+\frac{۱}{۲}[(\bar{\lambda})^۲-(\lambda^*)^۲]-(x-x^*)^T(\nabla F(x^*)+\nabla g(x^*)^T\lambda^*) \\ & -(\lambda-\lambda^*)^T\lambda^*+\frac{۱}{۲}\|x-x^*\|^۲+\frac{۱}{۲}\|\lambda-\lambda^*\|^۲. \end{aligned} \tag{۱۷.۴}$$

باتوجه به اینکه $\|\bar{\lambda}\|^۲=\Sigma_{j=۱}^m[(\lambda_j+g_j(x))^+]^۲$ و

$$[(\lambda_j+g_j(x))^+]^۲ = \begin{cases} [(\lambda_j+g_j(x))]^۲, & \lambda_j+g_j(x)\geq\circ, \\ \circ, & o.w. \end{cases} \tag{۱۸.۴}$$

داریم:

$$\nabla\|\bar{\lambda}\|^{۲} = \nabla(\sum_{j=۱}^{m}[(\lambda_j + g_j(x))^{+}]^{۲}) = \begin{pmatrix} ۲\nabla g(x)^T\bar{\lambda} \\ ۲\bar{\lambda} \end{pmatrix}. \qquad (۱۹.۴)$$

با محاسبه مشتق $V(t)$ درطول مسیر سیستم (۱۲.۴) با $-\lambda + \bar{\lambda} = g(x) - (\lambda + g(x))^{-}$ داریم:

$$\begin{aligned}
\frac{dV(x,\lambda)}{dt} &= [\nabla F(x) + \nabla g(x)^T\bar{\lambda} - \nabla F(x^*) - \nabla g(x^*)^T\lambda^* + x - x^*]^T(-x + \bar{x}) \\
&+ \frac{(\bar{\lambda} + \lambda - ۲\lambda^*)^T(-\lambda + \bar{\lambda})}{۲} \\
&= -\|x - \bar{x}\|^{۲} \\
&+ [\nabla F(x) + \nabla g(x)^T\bar{\lambda} - \nabla F(x^*) - \nabla g(x^*)^T\lambda^* + \bar{x} - x^*]^T(-x + \bar{x}) \\
&- \frac{۱}{۲}\|\lambda - \bar{\lambda}\|^{۲} + (\bar{\lambda} - \lambda^*)^T(-\lambda + \bar{\lambda}) \\
&= -\|x - \bar{x}\|^{۲} - [\nabla F(x) - \nabla F(x^*)]^T(x - x^*) - (\nabla F(x^*) \\
&+ \nabla g(x^*)^T\lambda^*)^T(\bar{x} - x^*) \\
&- [x - \nabla F(x) - \nabla F(x) - \nabla g(x)^T\bar{\lambda} - \bar{x}]^T(\bar{x} - x^*) \\
&- (\nabla g(x)^T\bar{\lambda} - \nabla g(x^*)^T\lambda^*)^T(\bar{x} - x^*) \\
&- \frac{۱}{۲}\|\lambda - \bar{\lambda}\|^{۲} + (\bar{\lambda} - \lambda^*)^T[g(x) - (\lambda + g(x))^{-}] \\
&= -\|x - \bar{x}\|^{۲} - \frac{۱}{۲}\|\lambda - \bar{\lambda}\|^{۲} - [\nabla F(x) - \nabla F(x^*)]^T(x - x^*) \\
&- [\nabla F(x^*) + \nabla g(x^*)^T\lambda^*]^T(\bar{x} - x^*) - [x - \nabla F(x) - \nabla g(x)^T\bar{\lambda} - \bar{x}]^T(\bar{x} - x^*) \\
&- \bar{\lambda^T}[-\nabla g(x)(x^* - x) - g(x) + g(x^*)] + \bar{\lambda}^T g(x^*) \\
&- \bar{\lambda}(\lambda + g(x))^{-} + (\lambda^*)^T[\nabla g(x^*)(x - x^*) - g(x) + g(x^*)] \\
&- (\lambda^*)^T g(x^*) + (\lambda^*)^T(\lambda + g(x))^{-} \\
&= -\|x - \bar{x}\|^{۲} - \frac{۱}{۲}\|\lambda - \bar{\lambda}\|^{۲} - [\nabla F(x) - \nabla F(x^*)]^T(x - x^*) \\
&- [\nabla F(x^*) + \nabla g(x^*)^T\lambda^*]^T(\bar{x} - x^*) - [x - \nabla F(x) - \nabla g(x)^T\bar{\lambda} - \bar{x}]^T(\bar{x} - x^*) \\
&- \bar{\lambda^T}[g(x^*) - g(x) - \nabla g(x)(x^* - x)] + \bar{\lambda}^T g(x^*) \\
&- (\lambda^*)^T[g(x) - g(x^*) - \nabla g(x^*) - (x - x^*)] + (\lambda^*)^T(\lambda + g(x))^{-} \\
&\leq -\|x - \bar{x}\|^{۲} - \frac{۱}{۲}\|\lambda - \bar{\lambda}\|^{۲} - [\nabla F(x) - \nabla F(x^*)]^T(x - x^*) \\
&- [\nabla F(x^*) + \nabla g(x^*)^T\lambda^*]^T(\bar{x} - x^*) - [x - \nabla F(x) - \nabla g(x)^T\bar{\lambda} - \bar{x}]^T(\bar{x} - x^*) \\
&- \bar{\lambda^T}[g(x^*) - g(x) - \nabla g(x)(x^* - x)] \\
&- (\lambda^*)^T[g(x) - g(x^*) - \nabla g(x^*)(x - x^*)].
\end{aligned} \qquad (۲۰.۴)$$

در نامساوی لم ۲.۲.۴ فرض کنید $v = x - \nabla f(x) - \nabla g(x)^T\bar{\lambda}$ و $u = x^*$ لذا داریم:

$$(x - \nabla F(x) - \nabla g(x)^T\bar{\lambda} - \bar{x})^T(\bar{x} - x^*) \geq ۰. \qquad (۲۱.۴)$$

به واسطه مشتق پذیری و تحدب $F(x)$ و $g(x)$، $\forall x \in \Omega$ داریم:

$$\begin{cases} [\nabla F(x) - \nabla F(x^*)]^T(x - x^*) \geq ۰, \\ g(x^*) - g(x) - \nabla g(x)^T(x^* - x) \geq ۰, \\ g(x) - g(x^*) - \nabla g(x^*)^T(x - x^*) \geq ۰. \end{cases} \qquad (۲۲.۴)$$

با جایگذاری (۶.۴) ، (۲۱.۴) و (۲۲.۴) در (۲۰.۴) داریم:

$$\frac{dV(x,\lambda)}{dt} \leq -\|x - \bar{x}\|^۲ - \frac{۱}{۲}\|\lambda - \bar{\lambda}\|^۲ \leq ۰. \qquad (۲۳.۴)$$

این به این معناست که شبکه عصبی (۱۲.۴) به مفهوم لیاپانوف پایدار است.

چون

$$V(x,\lambda) \geq \frac{۱}{۲}(\|x - x^*\|^۲ + \|\lambda - \lambda^*\|^۲).$$

بنابراین $V(x,\lambda)$ معین مثبت و شعاعی بی کران است. بنابراین یک زیر دنباله همگرا به صورت زیر وجود دارد.

$$\{(x(t_k)^T, \lambda(t_k)^T)^T \mid t_۰ < t_۱ < \cdots < t_k < t_{k+۱} < \dots\}, \quad t_k \to \infty, \quad k \to \infty,$$

به طوریکه

$$\lim_{k\to\infty}(x(t_k)^T, \lambda(t_k)^T)^T = (\hat{x}^T, \hat{\lambda}^T)^T.$$

ازقضیه لسال داریم:

$\{(x(t)^T, \lambda(t)^T)^T \to M\}$ وقتی $t \to \infty$ که M بزرگترین مجموعه تغییر ناپذیر در

$$K = \{(x(t)^T, \lambda(t)^T)^T \mid \frac{dV(x,\lambda)}{dt} = ۰\},$$

است.

از طرفی (۱۲.۴) و (۲۳.۴) نشان می دهند که

$$\frac{d\lambda}{dt} = ۰, \frac{dx}{dt} = ۰ \iff \frac{dV(x,\lambda)}{dt} = ۰.$$

بنابراین $(\hat{x}^T, \hat{\lambda}^T)^T \in D^*$ با $M \subseteq k \subseteq D^*$. با جایگذاری $x^* = \hat{x}$ و $\lambda^* = \hat{\lambda}$ در (۱۷.۴) تابع لیاپانوف دیگری را به صورت زیر تعریف می کنیم:

$$\begin{aligned} \hat{V}(x,\lambda) =& F(x) - F(\hat{x}) + \frac{۱}{۲}[(\bar{\lambda})^۲ - (\hat{\lambda})^۲] - (x - \hat{x})^T(\nabla F(\hat{x}) + \nabla g(\hat{x})\hat{\lambda}) \\ & - (\lambda - \hat{\lambda})^T\hat{\lambda} + \frac{۱}{۲}\|x - \hat{x}\|^۲ + \frac{۱}{۲}\|\lambda - \hat{\lambda}\|^۲. \end{aligned} \qquad (۲۴.۴)$$

پس $\hat{V}(x,\lambda)$ به طور پیوسته مشتق پذیر است و $\hat{V}(\hat{x},\hat{\lambda}) = ۰$. با توجه به اینکه

$$\lim_{k\to\infty}(x(t_k)^T, \lambda(t_k)^T)^T = (\hat{x}^T, \hat{\lambda}^T)^T,$$

بنابراین

$$\lim_{k\to\infty}\hat{V}(x(t_k)^T, \lambda(t_k)^T)^T = \hat{V}(\hat{x},\hat{\lambda}) = ۰.$$

بنابراین $\forall \epsilon > ۰$ وجود دارد $q > ۰$ به طوریکه برای هر $t > t_q$ داریم $\hat{V}(x,\lambda) < \epsilon$. به طور مشابه با تجزیه و تحلیل بالا می توانیم ثابت کنیم که تابع $\frac{d\hat{V}(x,\lambda)}{dt} \leq ۰$ در نتیجه برای $t \geq t_q$ داریم:

$$\|x(t) - \hat{x}\|^۲/۲ + \|\lambda(t) - \hat{\lambda}\|^۲/۲ \leq \hat{V}(x,\lambda) \leq \epsilon.$$

این یعنی $\lim_{t\to\infty} x(t) = \hat{x}$ و $\lim_{t\to\infty}\lambda(t) = \hat{\lambda}$ است. بنابراین مسیر جواب شبکه عصبی (۱۲.۴) به نقطه تعادل $(\hat{x}^T, \hat{\lambda}^T)^T$ همگرای سراسری است. همچنین $(\hat{x}^T, \hat{\lambda}^T)^T$ یک جواب بهینه (۵.۴) است. به خصوص اگر $D^* = \{((x^*)^T, (\lambda^*)^T)^T\}$ آنگاه برای هر $x_۰ \in \Omega$ و $\lambda_۰ \geq ۰$ جواب $(x^T, \lambda^T)^T$ با نقطه اولیه $(x_۰^T, \lambda_۰^T)^T$ با تجزیه و تحلیل بالا به $((x^*)^T, (\lambda^*)^T)^T$ همگرا می شود. این یعنی شبکه عصبی (۱۲.۴) پایدار مجانبی سراسری است و این اثبات را کامل می کند. □

۴.۴ مثال شبیه سازی شده

مثال ۱.۴.۴. برنامه ریزی مین ماکس زیر را در نظر بگیرید:

$$\min \quad f(x) = \max\{x_۱^۲ + x_۲^۴; (۲ - x_۱)^۲ + (۲ - x_۲)^۲; ۲e^{-x_۱+x_۲}\}$$
$$\text{s.t.} \begin{cases} x_۱^۲ + x_۲^۲ - ۲x_۱ + x_۲ \leq ۴, \\ |x_۱| \leq ۲, \quad |x_۲| \leq ۲. \end{cases} \tag{۲۵.۴}$$

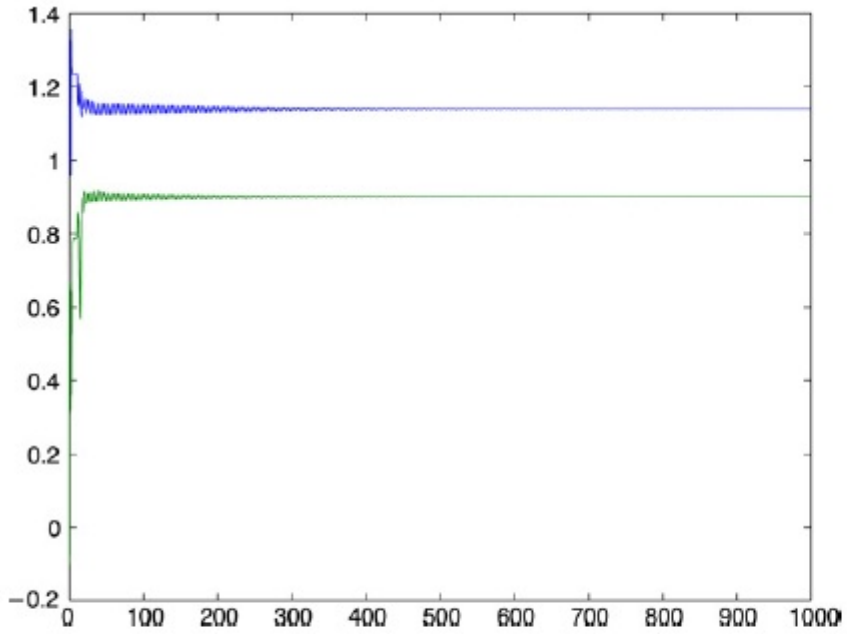
مساله (۲۵.۴) متعلق به یک رده از مسائل بهینه سازی مشتق ناپذیر است.

حالت ۱: فرض کنید $z = \max\{x_۱^۲ + x_۲^۴; (۲ - x_۱)^۲ + (۲ - x_۲)^۲; ۲e^{-x_۱+x_۲}\}$, (۲۵.۴) را می توان به برنامه ریزی محدب معادل زیر تبدیل کرد:

$$\min \ z$$
$$\text{s.t.} \begin{cases} x_۱^۲ + x_۲^۴ - z \leq ۰ \\ (۲ - x_۱)^۲ + (۲ - x_۲)^۲ - z \leq ۰ \\ ۲e^{-x_۱+x_۲} - z \leq ۰ \\ x_۱^۲ + x_۲^۲ - ۲x_۱ + x_۲ \leq ۴ \\ |x_۱| \leq ۲; \quad |x_۲| \leq ۲ \end{cases} \tag{۲۶.۴}$$

نتیجه شبیه سازی نشان می دهد که مسیر جواب همگرا به نقطه تعادل $x^e = (۱/۱۳۹۲, ۰/۸۹۹۶)^T$ است.

شکل۱.۴ نشان می دهد که مسیر شبکه به طور سراسری همگرا است.



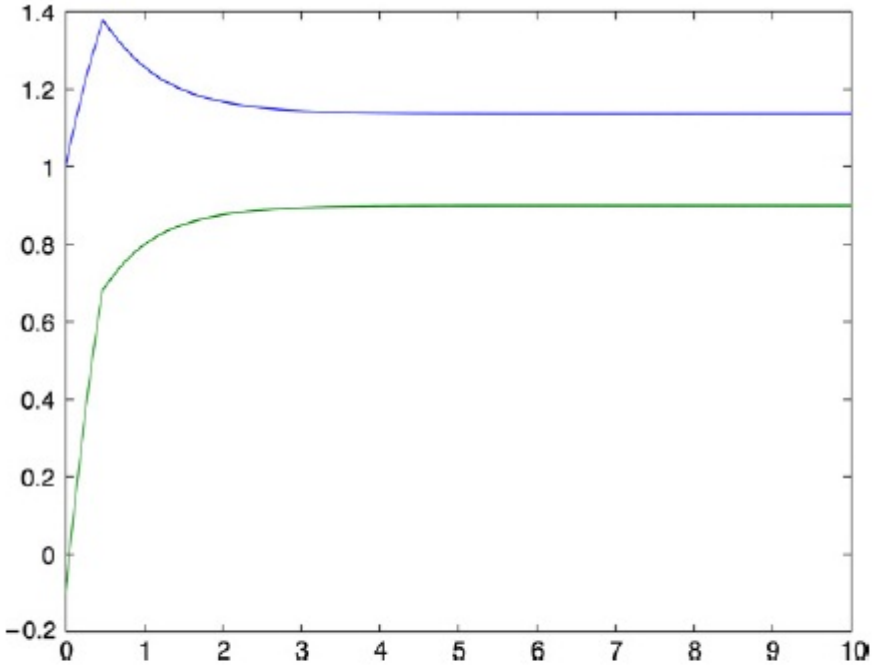
شکل ۱.۴: مسیرهای شبکه های عصبی در حالت (۱)

بنابراین، جواب بهینه بدست آمده عبارتست از $x^* = (۱/۱۳۹۲, ۰/۸۹۹۶)^T$. همچنین
$f_۱(x^*) = ۱/۹۵۲۷,\ f_۲(x^*) = ۱/۹۵۱۹,\ f_۳(x^*) = ۱/۵۷۴۰$.

حالت ۲ :فرض کنید $F(x) = \frac{۱}{p}\ln\sum_{i=۱}^{۳}\exp(pf_i(x))$. واضح است که، $F(x)$ یک تابع محدب هموار است برای هر $p > ۰$. مساله برنامه ریزی محدب زیر را در نظر بگیرید:

$$\min\ F(x) = \frac{۱}{p}\ln\sum_{i=۱}^{۳}\exp(pf_i(x))$$
$$\text{s.t.}\ \begin{cases} x_۱^۲ + x_۲^۲ - ۲x_۱ + x_۲ \le ۴, \\ |x_۱| \le ۲; \quad |x_۲| \le ۲. \end{cases} \qquad (۲۷.۴)$$

با $p = ۱۰^۵$ نتیجه شبیه سازی نشان می دهد مسیر جواب همگرا به نقطه تعادل $x^e = (۱/۱۳۷۷, ۰/۹۰۰۶)^T$ است. شکل ۲.۴ نشان می دهد که مسیر شبکه عصبی به طور سراسری همگراست. بنابراین، جواب بهینه بدست آمده عبارتست از $x^* = (۱/۱۳۹۲, ۰/۰/۸۹۹۶)^T$. همچنین داریم:
$f_۱(x^*) = ۱/۹۵۲۲, f_۲(x^*) = ۱/۹۵۲۳, f_۳(x^*) = ۱/۵۷۷۹$



شکل ۲.۴: مسیر های شبکه عصبی در حالت (۲)

۵.۴ نتیجه گیری و پیشنهادات برای کارهای آتی

در این فصل، ما روش های بهینه سازی برای حل رده ای از مسائل بهینه سازی مشتق ناپذیر را بررسی کردیم. ابتدا مساله برنامه ریزی مشتق ناپذیر به یک مساله برنامه ریزی محدب مشتق پذیر تبدیل می شود. با ساختن یک مدل شبکه عصبی برای حل برنامه ریزی محدب، پایداری نقطه تعادل و همگرایی مسیرها به جواب بهینه بررسی می شود. کارایی مدل پیشنهادی با یک مثال داده می شود.
از جمله مزیت‌های مدل ارائه شده می‌توان به موارد زیر اشاره کرد:

- حل مسأله ربات چند انگشتی
- کاربرد مدل ارائه شده در مسائل پردازش تصویر
- حل مسأله رگرسیون و طبقه‌بندی داده‌ها
- حل مسائل بهینه‌سازی سبد سهام مشتق ناپذیر
- 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

۱.۵.۴ پیشنهاداتی برای کارهای آتی

- در زمینه آنالیز عددی و محاسبات عددی ، مینیمم کردن ماکزیمم خطای چند جمله ای درونیاب با استفاده از تابع آنتروپی و چند جمله ای های چبیشف بر مبنای مسائل مین ماکس
- در زمینه مهندسی پزشکی، مینیمم کردن خطای دستگاه رادیو تراپی با ماکزیمم تابش پرتو بر اساس شبکه عصبی و مسائل مین ماکس
- در زمینه نظامی مینیمم کردن خطای موشک با ماکزیمم برد پرتاب، بر اساس مسائل مین ماکس

مراجع

[1] A. R. Nazemi, A dynamic system model for solving convex nonlinear optimization problems, comman Nonlinear Sci Number Simulat 17, 2012, 1696-1705

[2] A. R. Nazemi, Solving general convex nonlinear optimization problems by an efficient neurodynamic model, Engineering Application of Artificial Intelligence 26, 2013, 685-696

[3] Bazaraa. M. S. Sherali, H. D. and Shetty, C. M., 1979. Nonlinear Programming, Theory and Algorithms, John Wiley Sons, New York.

[4] Fang, S. C, Puthenpura. S, (1993) Linear optimization and extensions: Theory and algorithms. Prentice Hall, Englewood Cliffs, New Jersey

[5] Gigola. C, Gomez. S, (1990) A regularization method for solving the finite convex min-max problem. SIAM Journal on Numerical Analysis 27:1621-1634

[6] Hiriart-Urruty J-B, Lemarechal. C, (1993) Convex analysis and minimization algorithm. Springer-Verlag, Berlin

[7] Huard. P, (1967) Resolution of mathematical programming with nonlinear constraints by the method of centers. In: Abadie J (ed) Nonlinear programming, North-Holland, Amsterdam: 207-219

[8] Jaynes ET (1957) Information theory and statistical mechanics. Physics Review 106:620-630

[9] Kazarinoff ND (1961) Analytic inequalities. Holt, Rinehart and Winston, New York

[10] Polak E, Mayne DQ, Higgins JE (1991) Superlinearly convergent algorithm for min-max problems. Journal of Optimization Theory and Applications 69:407-439

[11] E. Cherkaev, A. Cherkaev, Minimax optimization problem of structural design, Computers and Structures 86 (2008) 1426-1435.

[12] K.L. Teo, X.Q. Yang, Portfolio selection problem with minimax type risk function, Journal of Annals of Operations Research 101 (2001) 333-349.

[13] Y. H. Yu, L. Gao, Nonmonotone linear search algorithm for constrained minimax problem, Journal of Optimization Theory and Application 115 (2002) 419-446.

[14] X. Li, S. Fang, On the entropy regularization method for solving min-max problems with applications, Mathematics Methods of Operations Research 46 (1997) 119-130.

[15] X. Li, An efficient approach to a class of non-smooth optimization problems, Science in China (Series A) 37 (1994) 323-330.

[16] G. Di Pillo, L. Grippo, S. Lucidi, Smooth transformation of the generalized minimax problem, Journal of Optimization Theory and Application 95 (1997) 1-24.

[17] C. Charalambous, A.R. Conn, An efficient method to solve the minimax problem directly, SIAM Journal on Numerical Analysis 15 (1978) 162-187.

[18] C. Gigola, S. Gomez, A regularization method for solving the finite convex min-max problem, SIAM Journal on Numerical Analysis 27 (1990) 1621-1634.

[19] A. Vardi, New minimax algorithm, Journal of Optimization Theory and Application 75 (1992) 613-634.

[20] F. Ye, H. Liu, S. Zhou, S. Liu, A smoothing trust-region Newton-CG method for minimax problem, Applied Mathematics and Computation 199 (2008) 581-589.

[21] M. Chao, Z. Wang, Y. Liang, Q. Hu, Quadratically constraint quadratical algorithm model for nonlinear minimax problems, Applied Mathematics and Computation 205 (2008) 247-262.

[22] Z. Zhu, X. Cai, J. Jian, An improved SQP algorithm for solving minimax problems, Applied Mathematics Letters (in press).

[23] D.W. Tank, J.J. Hopfield, Simple 'neural' optimization network: An A/D converter signal decision circuit and a linear programming circuit, IEEE Transactions on Circuits and System 33 (1986) 533-541.

[24] M.P. Kennedy, L.O. Chua, Neural networks for nonlinear programming, IEEE Transactions on Circuits and System 35 (1988) 554562.

[25] S. Effati, A.R. Nazemi, Neural network models and its application for solving linear and quadratic programming problems, Applied Mathematics and Computation 172 (2006) 305-331.

[26] K.Z. Chen, Y. Leung, K.S. Leung, X.B. Gao, A neural network for nonlinear programming problems, Neural Computing Application 11 (2002) 103-111.

[27] Q. Tao, J. Cao, M. Xue, H. Qiao, A high performance neural network for solving nonlinear programming problems with hybrid constraints, Physics Letters A 288 (2001) 88-94.

[28] Y. Xia, H. Leung, J. Wang, A projection neural network and its application to constrained optimization problems, IEEE Transactions on Circuits and System-I 49 (2002) 447-458.

[29] Y. Xia, J. Wang, A recurrent neural networks for nonlinear convex optimization subject to nonlinear inequality constraints, IEEE Transactions on Circuits and System-I 51 (2004) 1385-1394.

[30] X.B. Gao, A novel neural network for nonlinear convex programming, IEEE Transactions on Neural Networks 15 (2004) 613-621.

[31] M. Forti, P. Nistri, M. Quincampoix, Generalized neural network for nonsmooth nonlinear programming problems, IEEE Transactions on Circuits and System-I 51 (2004) 1741-1754.

[32] X. Xue, W. Bian, Subgradient-based neural network for nonsmooth convex optimization problems, IEEE Transactions on Circuits and System-I 55 (2008) 2378-2391.

[33] D. Kinderlehrer, G. Stampcchia, An Introduction to Variational Inequalities and their Applications, Academic, New York, 1980.

واژه‌نامه فارسی به انگلیسی

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

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

واژه‌نامه انگلیسی به فارسی

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

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

Aabstract

In this thesis, we describe a neural network model based on a dynamic optimization technique for solving a class of non-smooth optimization problems with min-max objective function. The basic idea is to replace the min-max function by a smooth one using an entropy function. With this smoothing technique, the non-smooth problem is converted into an equivalent differentiable convex programming problem. A neural network model is then constructed based on Karush-Kuhn-Tucker optimality conditions. It is investigated that the proposed neural network is stable in the sense of Lyapunov and can converge to an exact optimal solution of the original problem. The effectiveness of the method are demonstrated by several numerical simulations.
keywords: Stability, Convergence, Non-differentiable Programming, Optimization, Neural network, Min-Max Problem, Linear and Nonlinear Programming, Entropy

Shahrood University

Faculty of Mathematical Sciences

An application of dynamic systems for solving a class of non-smooth optimization problems

Abdolsakhi Nazari

Supervisor

Dr. Alireza Nazemi

2015