



## یک شبکه عصبی بازگشتی تک لایه کارا برای حل دسته‌ای از مسائل بهینه‌سازی محدب ناهموار

محمدجواد عبادی<sup>۱\*</sup>، علی‌رضا حسینی<sup>۲</sup>، حسین جعفری<sup>۳</sup>

<sup>(۱)</sup> استادیار، گروه ریاضی، دانشگاه دریاوردی و علوم دریایی چابهار، چابهار، ایران

<sup>(۲)</sup> دانشکده ریاضی، آمار و علوم کامپیوتر، پردیس علوم، دانشگاه تهران، تهران، ایران

<sup>(۳)</sup> گروه ریاضی، دانشگاه دریاوردی و علوم دریایی چابهار، چابهار، ایران

تاریخ ارسال مقاله: ۹۷/۰۹/۱۳ تاریخ پذیرش مقاله: ۹۸/۱۱/۲۷

### چکیده

مسائل بهینه‌سازی مقید دارای کاربردهای وسیعی در علوم، مهندسی و اقتصاد می‌باشند. در این مقاله یک مدل شبکه عصبی برای حل دسته‌ای از مسائل بهینه‌سازی مقید ناهموار با تابع هدف محدب ناهموار و قیود نامساوی‌های غیرخطی و خطی آفین پیشنهاد شده است. آن یک شبکه عصبی بازگشتی تک لایه غیر جریمه‌ای مبتنی بر شمول دیفرانسیلی است. برخلاف اکثر مدل‌های شبکه عصبی موجود برای حل مسائل بهینه‌سازی، در ساختار مدل پیشنهادی هیچ پارامتر جریمه‌ای یا تابع جریمه وجود نداشته و مدل از پیچیدگی کمتری برخوردار است که منجر به پیاده‌سازی آسان‌تر مدل پیشنهادی می‌شود. معادل بودن مجموعه جواب‌های بهینه مسأله بهینه‌سازی اصلی و مجموعه نقاط تعادلی مدل شبکه عصبی پیشنهادی اثبات گردیده است. به‌علاوه همگرایی سراسری و پایداری شبکه عصبی پیشنهادی نشان داده شده‌اند. به منظور روشن ساختن کارایی و اثربخشی مدل ارائه شده تعدادی مثال شامل مسأله مینیمم‌سازی نرم  $L_1$  ارائه و حل شده‌است.

**واژه‌های کلیدی:** مسأله بهینه‌سازی محدب ناهموار، معادله خطی، نامساوی غیرخطی، شبکه عصبی بازگشتی.

## ۱- مقدمه

مسائل بهینه‌سازی مقید در بسیاری از کاربردهای علوم، مهندسی و اقتصاد رخ می‌دهند. برای حل چنین مسائلی روش‌های مختلفی ارائه و پیشنهاد شده است. یکی از این روش‌های ارائه شده روش شبکه عصبی است که از دهه ۱۹۶۰ شروع به پیشنهاد شد. اولین مدل شبکه عصبی برای حل مسائل برنامه‌ریزی خطی توسط تانک و هاپفیلد معرفی شد [1]. مدل مذکور شالوده‌ی انواع دیگر شبکه‌های عصبی پیشنهادی برای حل مسائل بهینه‌سازی خطی و غیرخطی است. با الهام از شبکه تانک و هاپفیلد، کندی و چوا یک شبکه عصبی را برای حل مسائل برنامه‌ریزی غیرخطی معرفی نمودند [2]. این مدل شبکه عصبی شامل پارامترهای جریمه‌ی متناهی و براساس روش کاهش نیوتن بود. وقتی که پارامتر جریمه خیلی بزرگ است این شبکه با مشکل مواجه می‌شود. برای حل برخی مسائل برنامه‌ریزی اکیداً محدب، یک شبکه لاگرانژی دو لایه معرفی شد [3]. یک شبکه عصبی اولیه-دوگان با ساختاری تک و دو لایه برای حل مسائل برنامه‌ریزی خطی و درجه دو پیشنهاد شد [4]. مسائل برنامه‌ریزی محدب کلی و غیرخطی با برخی شبکه‌های عصبی پیشنهادی از نوع تصویری حل شده‌اند که به طور سراسری به یک جواب بهینه دقیق همگرایند [5-6]. برخی شبکه‌های عصبی بازگشتی تک‌لایه برای حل مسائل برنامه‌ریزی خطی و درجه دو در [7] و [8] پیشنهاد گردیدند. در میان مسائل بهینه‌سازی، ناهموارها به دلیل نقش مهم‌شان در کاربردهای علوم و مهندسی برای محققین از جذابیت خاصی برخوردارند، به‌عنوان نمونه می‌توان از کاربردهای آن‌ها در پردازش سیگنال، کنترل ربات، مد لغزشی و مانند آن نام برد [9-14] و منابع داخل آن‌ها را ملاحظه نمایید. به طور کلی حل این‌گونه مسائل حتی در حالت‌های نامقید نیز مشکل است. برای اولین بار فورتی و همکاران [15] یک شبکه عصبی برای حل مسائل بهینه‌سازی ناهموار که توسیعی از شبکه کندی و چوا از حالت هموار به ناهموار بود را پیشنهاد دادند. مدل براساس روش جریمه است و خط سیرهای جواب فقط وقتی پارامتر جریمه به بی‌نهایت میل می‌کند همگرایند. بعداً بیان و ژو [16] این شبکه را برای حل مسائل بهینه‌سازی با قیود نامساوی غیرخطی و

خطی آفین توسیع دادند. برخی شبکه‌های عصبی بازگشتی برای حل مسائل بهینه‌سازی محدب ناهموار که به ترتیب دارای قیود مساوی خطی، کران‌دار و هردوی مساوی خطی و کران‌دار هستند، پیشنهاد شده است [17-19]. اخیراً مدل‌هایی برای حل مسائل بهینه‌سازی محدب ناهموار به ترتیب با قیود نامساوی، مساوی خطی و نامساوی غیرخطی پیشنهاد شده است [13]، [20] و [21]. مسأله بهینه‌سازی غیرخطی مقید ناهموار زیر را در نظر بگیرید:

$$\begin{aligned} \min & f(x) \\ \text{s. t} & g_i(x) \leq 0, \\ & Ax = b \end{aligned} \quad (1)$$

که در آن

$$\begin{aligned} x &= (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n \\ A &\in \mathbb{R}^{m \times n} \quad S_1 = \{x \in \mathbb{R}^n : g_i(x) \leq 0\}, \\ S_2 &= \{x \in \mathbb{R}^n : Ax = b\} \end{aligned}$$

9

$$S = S_1 \cap S_2$$

ناحیه شدنی مسأله بهینه‌سازی (۱) است.  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  غیرخطی است و لزوماً بر ناحیه شدنی  $S$  هموار نیست و  $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$  ( $i = 1, 2, \dots, m$ ) توابعی ناهموارند. فرض می‌کنیم که یا حداقل یک جواب بهینه از مسأله (۱) در درون ناحیه  $S_1$  وجود دارد یا  $g_i$ ها در ناحیه  $S_1$  اکیداً محدبند. در این مقاله یک شبکه عصبی تک لایه بازگشتی ساده شده برای حل مسائل بهینه‌سازی محدب ناهموار با قیود خطی آفین و نامساوی غیرخطی مانند مسأله بهینه‌سازی (۱) پیشنهاد شده است. همگرایی سراسری شبکه عصبی پیشنهادی را اثبات نموده و پایداری شمول دیفرانسیلی متناظر را نشان می‌دهیم. در مقایسه با مدل‌های پیشنهادی دیگر برای حل مسأله بهینه‌سازی (۱)، به عنوان نمونه [22] و [23] ساختار مدل پیشنهادی دارای پارامتر یا تابع جریمه‌ای نیست. در [13] و [14] نشان داده‌ایم که وقتی پارامتر جریمه به درستی انتخاب نشده باشد، ممکن است مشکلات متعددی از قبیل واگرایی، عدم رسیدن به جواب بهینه و دور شدن از ناحیه

می‌شود هرگاه در مجاورت هر نقطه  $x \in \mathfrak{R}^n$  لیپ شیتس باشد.

**تعریف ۲-۴.** [24] فرض کنید  $f$  در نزدیکی  $x \in \mathfrak{R}^n$  لیپ شیتس باشد. در این صورت  $f^0(x; v)$  مشتق جهتی تعمیم یافته  $f$  در  $x$  در جهت هر بردار دلخواه  $v \in \mathfrak{R}^n$  با

$$f^0(x; v) = \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t}$$

تعریف می‌شود و گرادیان تعمیم یافته کلارک  $f$  در  $x$  را به صورت

$$\partial f(x) = \{y \in \mathfrak{R}^n : f^0(x; v) \geq y^T v, \forall v \in \mathfrak{R}^n\}$$

تعریف می‌کنیم.

**تعریف ۲-۵.** [25] تابع  $f$  در  $x$  منظم نامیده می‌شود هرگاه برای هر  $v \in \mathfrak{R}^n$ :

(۱) مشتق جهتی یک طرفه معمولی  $f'(x; v)$  که با

$$f'(x; v) = \lim_{\xi \rightarrow 0^+} \frac{f(y + \xi v) - f(y)}{\xi}$$

داده شده، وجود داشته باشد

$$(۲) f'(x; v) = f^0(x; v) \text{ باشد.}$$

به‌وضوح هر تابع محدب منظم است.

**لم ۲-۱.** [25] (قاعده‌ی زنجیره‌ای کلارک) اگر  $W : \mathfrak{R}^n \rightarrow \mathfrak{R}$  در  $x(t)$  منظم و

$x : \mathfrak{R} \rightarrow \mathfrak{R}^n$  در  $t$  مشتق‌پذیر و در حوالی آن لیپ شیتس باشد، آن‌گاه به‌ازای تقریباً هر  $t \in [0, \infty)$

$$\frac{dW(x(t))}{dt} = \xi^T \dot{x}, \quad \forall \xi \in \partial W(x(t)), \text{ است.}$$

**لم ۲-۲.** [25] اگر  $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$  تابعی محدب بر مجموعه محدب ناتهی  $S \subset \mathfrak{R}^n$  باشد، آن‌گاه  $x^*$  یک

جواب بهینه از  $f(x)$  در  $S$  است اگر و فقط اگر  $f$  یک زیرگرادیان  $\xi^*$  در  $x^*$  به گونه‌ای داشته باشد که

$$\xi^{*T}(x - x^*) \geq 0, \quad \forall x \in S$$

شدنی رخ دهد. ساختار بدنه مقاله به صورت زیر است. برخی از تعاریف و قضایای مقدماتی در بخش ۲ ارائه شده است. در بخش ۳، مدل شبکه عصبی برای حل (۱) ساخته شده است. پایداری و همگرایی سراسری مدل شبکه عصبی پیشنهادی را در بخش ۴ اثبات می‌کنیم. برخی مثال‌های گویا و مسأله مینیمم‌سازی نرم  $L_1$  در بخش ۵ بیان و حل شده است. در نهایت در بخش ۶ به نتیجه‌گیری می‌پردازیم.

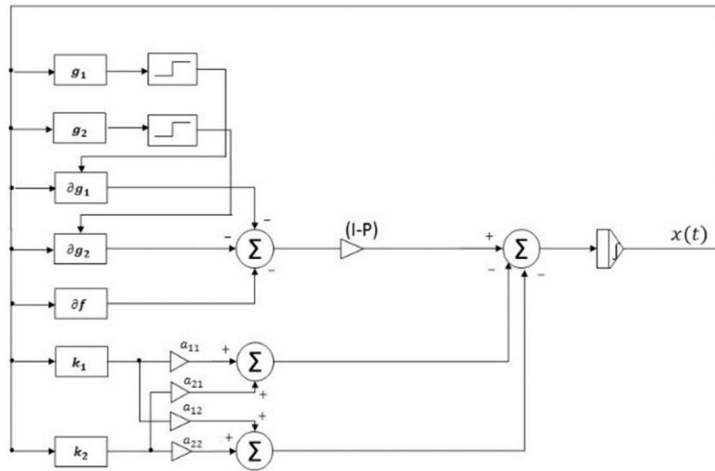
## ۲- تعاریف مقدماتی

در این بخش تعدادی از تعاریف و لم‌های مورد نیاز در ادامه فصل آمده است. توجه داشته باشید که  $l_1$  و  $l_2$  به ترتیب نشان دهنده نرم‌های  $\|\cdot\|_1$  و  $\|\cdot\|_2$  از یک بردار در  $\mathfrak{R}^n$  هستند.

**تعریف ۲-۱.** [24] مجموعه‌های  $X$  و  $Y$  را در نظر بگیرید. یک نگاشت مجموعه مقدار  $F$  از  $X$  به  $Y$  نگاشتی است که به هر نقطه  $x \in X$  یک زیرمجموعه  $F(x)$  از  $Y$  را نسبت می‌دهد.

**تعریف ۲-۲.** [24] نگاشت مجموعه مقدار  $F$  با مقادیر ناتهی نیم‌پیوسته بالایی (U.S.C.) در  $x^0 \in X$  نامیده می‌شود هرگاه برای هر مجموعه باز  $N$  شامل  $F(x^0)$  یک همسایگی  $M$  از  $x^0$  به گونه‌ای وجود داشته باشد که  $F(M) \subseteq N$  باشد. همچنین  $F$  نیم‌پیوسته بالایی است اگر و فقط اگر در هر نقطه  $x^0 \in X$  نیم‌پیوسته بالایی باشد.

**تعریف ۲-۳.** [24] به‌ازای هر تابع  $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$  اگر برای هر  $\epsilon > 0$  داده شده  $\delta > 0$  چنان وجود داشته باشد که برای هر  $x_1, x_2 \in \mathfrak{R}^n$  که  $\|x_1 - x_2\|_2 < \delta$  و  $\|x_2 - x\|_2 < \delta$  را برآورده نمایند،  $|f(x_1) - f(x_2)| \leq \epsilon \|x_1 - x_2\|_2$  را داشته باشیم، آن‌گاه گوئیم تابع  $f$  در نزدیکی  $x \in \mathfrak{R}^n$  لیپ شیتس است. تابع  $f$  در  $\mathfrak{R}^n$  لیپ شیتس موضعی نامیده



شکل ۱: نمودار شماتیک شبکه عصبی (۲) برای مسأله بهینه‌سازی (۱)

$$\begin{aligned}
 s. t. \quad & g_1(x_1, x_2) \leq 0, \\
 & g_2(x_1, x_2) \leq 0, \\
 & a_{11}x_1 + a_{12}x_2 = b_1, \\
 & a_{21}x_1 + a_{22}x_2 = b_2.
 \end{aligned} \tag{۴}$$

معماری شبکه عصبی (۲) برای مسأله بهینه‌سازی (۴) در شکل ۱. نشان داده شده است، که

$$(k_1, k_2)^T = Ax - b$$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

می‌باشند.

**تعریف ۳-۶.** نقطه  $x_{eq}$  یک نقطه تعادلی شمول دیفرانسیلی (۲) نامیده می‌شود، هرگاه توابع اندازه‌پذیر  $\eta \in \partial f(x_{eq})$  و  $\gamma_i \in \partial g_i(x_{eq})$  چنان موجود باشند که

$$(I - P) \left( \eta + \sum_{i=1}^m \theta(g_i(x_{eq})) \gamma_i \right) + A^T(Ax_{eq} - b) = 0. \tag{۵}$$

**تعریف ۳-۷.** تابع به‌طور مطلق پیوسته  $x(\cdot): [0, T] \rightarrow \mathbb{R}^n$  یک جواب از شبکه عصبی (۲) بر بازه  $[0, T]$  نامیده می‌شود هرگاه شرط اولیه  $x(0) = x_0$  را برآورده نماید و برای تقریباً همه  $t \in [0, T]$  داشته باشیم:

### ۳- مدل شبکه عصبی

در این بخش یک مدل شبکه عصبی ساده را برای حل مسأله بهینه‌سازی (۱) ارائه می‌نماییم. مدل پیشنهادی را با الهام از مدل ارائه شده در [21] و مبتنی بر آن مدل و با استفاده از اثبات قضایا در [13]، [20]، [21] و [24] به‌صورت زیر برای حل مسأله بهینه‌سازی (۱) ارائه می‌نماییم.

$$\begin{aligned}
 \dot{x}(t) \in & -(I - P) \left( \partial f(x(t)) + \right. \\
 & \left. \sum_{i=1}^m \theta(g_i(x)) \partial g_i(x) \right) - A^T(Ax - b) \tag{۲}
 \end{aligned}$$

که در آن  $I, P = A^T(AA^T)^{-1}A$  ماتریس همانی و  $\theta(\cdot)$  یک تابع فعال‌سازی است که به‌صورت زیر تعریف می‌گردد:

$$\theta(x) = \begin{cases} 1 & x > 0, \\ [0,1] & x = 0, \\ 0 & x < 0. \end{cases} \tag{۳}$$

شبکه عصبی (۲) می‌تواند توسط یک مدار تعمیم یافته پیاده‌سازی شود. برای توضیحات و جزئیات بیشتر در مورد مدارهای تعمیم یافته، خوانندگان می‌توانند به [8]، [15] و [28] مراجعه نمایند. حال یک پیاده‌سازی مدار تعمیم یافته از شبکه عصبی پیشنهادی (۲) برای حل مسأله بهینه‌سازی ساده زیر را ارائه می‌نماییم.

$$\min f(x_1, x_2),$$

$$\left\{ \begin{array}{l} 0 \in \partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^* + A^T v^* \\ g_i(x^*) \leq 0, \\ Ax^* - b = 0, \\ \mu_i^* g_i(x^*) = 0, \\ v^* \quad \text{آزاد} \\ \mu_i^* \geq 0 \end{array} \right. \quad (8)$$

$$\frac{dx(t)}{dt} \in -(I - P) \left( \partial f(x(t)) + \sum_{i=1}^m \theta \left( g_i(x(t)) \right) \partial g_i(x(t)) \right) - A^T (Ax(t) - b).$$

به بیانی دیگر، توابع اندازه‌پذیر  $\eta(t) \in \partial f(x(t))$  و  $\gamma_i(t) \in \partial g_i(x(t))$  به گونه‌ای وجود داشته باشند که به ازای هر  $t \in [0, T]$ :

$$\dot{x}(t) = -(I - P) \left( \eta(t) + \sum_{i=1}^m \theta \left( g_i(x(t)) \right) \gamma_i(t) \right) - A^T (Ax(t) - b), \quad (6)$$

می‌توانیم مسأله (۱) را به‌طور معادل زیر بنویسیم:

$$\begin{aligned} \min & f(x) \\ \text{s.t. } & g(x) = (g_1(x), g_2(x), \dots, g_m(x))^T \leq 0, \\ & Ax = b, \end{aligned} \quad (7)$$

که

$$g(x) = (g_1(x), g_2(x), \dots, g_m(x))^T : \mathcal{R}^n \rightarrow \mathcal{R}^m$$

یک تابع برداری مقدار  $m$ -بعدی و  $g_i$ ها  $(i = 1, 2, \dots, m)$  همان‌طور که قبلاً تعریف شده‌اند، محدب می‌باشند ولی ممکن است ناهموار باشند.

در این‌جا شرایط کروش-کان-تاکر را برای مسأله بهینه‌سازی ناهموار غیرخطی (۱) به‌صورت زیر بیان می‌کنیم:

**لم ۳-۳.** اگر توابع  $f$  و  $g$  محدب باشند، آن‌گاه  $x^*$  یک نقطه بهینه از مسأله (۱) است، اگر و فقط اگر  $\mu_i^* \in \mathcal{R}^m$  و  $v^* \in \mathcal{R}^n$  به گونه‌ای وجود داشته باشند که

#### ۴- تحلیل همگرایی

در این بخش همگرایی سراسری شبکه عصبی (۲) مورد تجزیه و تحلیل و اثبات قرار گرفته است. ابتدا ثابت می‌کنیم هنگامی که نقطه آغازین به‌طور دلخواه انتخاب شود خط سیر جواب شبکه عصبی (۲) به  $S_2 = \{x \in \mathcal{R}^n \mid Ax = b\}$  همگرا خواهد بود. سپس پایداری شبکه عصبی (۲) و همگرایی سراسری به یک جواب بهینه از مسأله بهینه‌سازی (۱) اثبات خواهند شد.

**قضیه ۴-۱.** حداقل یک جواب موضعی  $x(t)$  برای هر نقطه آغازین  $x(0) = x_0 \in \mathcal{R}^n$  برای شبکه عصبی (۲) وجود دارد. همچنین  $x(t)$  به‌طور نمایی به  $S_2 = \{x \in \mathcal{R}^n \mid Ax = b\}$  همگراست و برای هر  $t \geq 0$  وقتی  $x_0 \in S_2$  باشد، داریم:

$$x(t) \in S_2$$

**اثبات:** به‌وضوح سمت راست شمول دیفرانسیلی (۲) نیم پیوسته بالایی است و مقادیرش ناتهی، محدب و فشرده‌اند. بنابراین برای هر نقطه آغازین  $x_0 \in \mathcal{R}^n$  با توجه به قضیه ۳ در صفحه ۹۸ از [24] یک جواب موضعی  $x(t)$  از شبکه عصبی (۲) با  $x(0) = x_0$  وجود دارد. مشابه اثبات قضیه ۱.۴ در مرجع [13] فرض می‌کنیم که  $B(x) = \frac{1}{2} \|Ax - b\|_2^2$  باشد. به آسانی می‌توان نشان داد که  $B(x)$  مشتق‌پذیر و محدب است و  $\nabla B(x) = A^T (Ax - b)$ . با استفاده از قاعده زنجیره‌ای داریم:

$$v^* = (AA^T)^{-1}A(\partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^*) + (Ax^* - b)$$

داریم:

$$0 \in (I-P) \left( \partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^* \right) + A^T(Ax^* - b) \quad (10)$$

$$= \partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^* + A^T v^*$$

بنابراین برطبق لم ۳-۳،  $x^*$  یک جواب بهینه از مسأله (۱) است. برای اثبات عکس قضیه فرض کنید که  $x^*$  یک جواب بهینه از مسأله بهینه‌سازی (۱) باشد. از این رو بر طبق لم ۳-۳،  $\mu_i^*$  و  $v^*$  به گونه‌ای وجود دارند که مساوی‌ها در (۸) برقرارند. بنابراین

$$0 \in (I-P) \left( \partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^* \right) + A^T v^* \quad (11)$$

$$= (I-P) [\partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^*].$$

به‌علاوه چون  $Ax^* - b = 0$  داریم:

$$0 \in (I-P) \left( \partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^* \right) + A^T(Ax^* - b) \quad (12)$$

با ترکیب این با معادله واقع در (۸)، به‌وضوح  $x^*$  یک نقطه تعادلی شمول دیفرانسیلی (۲) است.

**لم ۴-۴.** اگر  $S^* \cap S \neq \emptyset$  باشد، آن‌گاه به ازای هر  $x \in \mathfrak{R}^n$ ،  $\frac{dV(x(t), x^*)}{dt} \leq 0$  است.

**اثبات.** فرض کنید تابع انرژی  $V(x, x^*) : \mathfrak{R}^n \times S^* \rightarrow [0, \infty)$  به‌صورت  $V(x, x^*) = \frac{1}{2} \|x - x^*\|_2^2$  تعریف شده باشد. به‌وضوح این تابع انرژی نسبت به متغیر  $x$  به‌طور پیوسته مشتق‌پذیر و نسبت به متغیر  $x^*$  پیوسته است. توابع اندازه‌پذیر  $\eta(t) \in \partial f(x(t))$  و  $\gamma_i(t) \in \partial g_i(x(t))$  به گونه‌ای وجود دارند که مشتق  $V$  نسبت به  $t$  به‌صورت

$$\frac{d}{dt} B(x(t)) = \nabla B(x(t))^T x'(t)$$

$$= (Ax - b)^T A x'(t)$$

$$\leq \sup_{\eta(t) \in \partial f(x(t)), \gamma_i(t) \in \partial g_i(x(t))} (Ax - b)^T$$

$$A \left[ -A^T(Ax - b) - (I - P)(\eta(t) + \sum_{i=1}^m \theta(g_i(x(t)))\gamma_i(t)) \right]$$

$$= -(Ax - b)^T A A^T (Ax - b)$$

$$\leq -\lambda_{\max}(A A^T) \|Ax - b\|_2^2$$

$$= -2\lambda_{\max}(A A^T) B(x(t))$$

که در آن  $\lambda_{\max}(AA^T)$  مقدار ویژه ماکزیمم  $AA^T$  و از این حقیقت که

$$A(I-P) = A(I - A^T(AA^T)^{-1}A) = A - A = 0$$

است، استفاده می‌کنیم. به‌دلیل این که  $A$  یک ماتریس رتبه کامل سطری است و به‌وضوح  $\lambda_{\max}(AA^T) > 0$  است، داریم:

$$B(x(t)) \leq e^{-2\lambda T_{\max 0}}$$

در حالتی که  $x(0) = x_0 \notin S_2$ ،  $x(t)$  به‌طور نمایی به  $S_2 = \{x \in \mathfrak{R}^n \mid Ax = b\}$  همگرا خواهد شد. اگر  $x(0) = x_0 \in S_2$  باشد، آن‌گاه  $B(x_0) = 0$  است و در این حالت به‌ازای  $t \geq 0$ ،  $B(x(t)) = 0$  است. از این رو به ازای  $t \geq 0$  وقتی  $x_0 \in S_2$  است،  $x(t) \in S_2$  است.

**قضیه ۴-۹.** اگر  $x^*$  یک نقطه تعادلی شمول دیفرانسیلی (۲) باشد، در این صورت  $x^*$  یک نقطه بهینه مسأله بهینه‌سازی (۱) است و بالعکس.

**اثبات.** فرض کنید  $x^*$  یک نقطه تعادلی شمول دیفرانسیلی (۲) باشد، در این صورت داریم:

$$0 \in (I-P)(\partial f(x^*) + \sum_{i=1}^m \partial g_i(x^*)^T \mu_i^*) + A^T(Ax^* - b)$$

برطبق قضیه ۴-۸ داریم:  $Ax^* = b$  و  $\mu_i^* \geq 0$  با انتخاب

$$\begin{aligned} & \sum_{i=1}^m \theta(g_i(x))(g_i(x^*) - g_i(x)) \\ &= \sum_{i \in I^+(x)} \theta(g_i(x))(g_i(x^*) - g_i(x)) \\ &+ \sum_{i \in I^0(x)} \theta(g_i(x))(g_i(x^*) - g_i(x)) \\ &+ \sum_{i \in I^-(x)} \theta(g_i(x))(g_i(x^*) - g_i(x)). \end{aligned} \quad (۱۶)$$

چون  $x \in S$ ، به‌وضوح  $I^+(x) = \emptyset$  است و برای  $i \in I^-(x)$ ، داریم:

$$\sum_{i \in I^-(x)} \theta(g_i(x))(g_i(x^*) - g_i(x)) = 0.$$

هم‌چنین اگر  $i \in I^0(x)$  باشد، در این صورت

$$\begin{aligned} & g_i(x^*) - g_i(x) \leq 0 \\ & \sum_{i \in I^0(x)} \theta(g_i(x))(g_i(x^*) - g_i(x)) \leq 0. \end{aligned}$$

بر طبق روابط فوق و از (۱۶-۱۴) در می‌یابیم که

$$\sum_{i=1}^m \theta(g_i(x))(g_i(x^*) - g_i(x)) \leq 0, \quad (۱۷)$$

بر طبق لم ۲-۲، داریم:

$$-\eta^T(x - x^*) \leq 0 \quad (۱۸)$$

چون  $x, x^* \in S$  و  $Px = Px^* = q$  است، پس

$$P \left[ \eta(t) + \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t) (x - x^*) \right] = 0 \quad (۱۹)$$

و جواب بدیهی عبارت است از

$$-\|Ax - b\|_2^2 \leq 0 \quad (۲۰)$$

با ترکیب (۲۱) - (۱۷) با (۱۳) داریم:

$$\begin{aligned} \frac{dV(x, x^*)}{dt} &= \frac{dV(x, x^*)}{dx} \frac{dx}{dt} \\ &= -[(I - P)(\eta(t) + \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)) - A^T(Ax - b)](x - x^*) \\ &= -(I - P)(\eta(t) + \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t))(x - x^*) - A^T(Ax - b)(x - x^*) \\ &= -(I - P)(\eta(t) + \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t))(x - x^*) \\ &\quad - \|Ax - b\|_2^2 \\ &= -\eta(t)(x - x^*) - \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)(x - x^*) \quad (۱۳) \end{aligned}$$

$$+ P \left[ \eta(t) + \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t) (x - x^*) \right] - \|Ax - b\|_2^2$$

از تحدب  $g_i$  ها به‌ازای هر  $i \in \{1, 2, \dots, m\}$  داریم:  $\gamma_i(t)^T(x - x^*) \leq g_i(x^*) - g_i(x)$  به‌دلیل این‌که  $\theta(g_i(x))$  برای هر  $i$ ، نامنفی است پس داریم:  $-\theta(g_i(x)) \gamma_i(t)^T(x - x^*) \leq \theta(g_i(x))(g_i(x^*) - g_i(x))$ .

می‌توان نتیجه گرفت که

$$\begin{aligned} & - \left[ \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t) \right]^T (x - x^*) \\ & \leq \sum_{i=1}^m \theta(g_i(x))(g_i(x^*) - g_i(x)) \end{aligned} \quad (۱۴)$$

فرض می‌کنیم که

$$\begin{aligned} I^+(x) &= \{i \in \{1, \dots, m\} \mid g_i(x) > 0\}, \\ I^0(x) &= \{i \in \{1, \dots, m\} \mid g_i(x) = 0\}, \\ I^-(x) &= \{i \in \{1, \dots, m\} \mid g_i(x) < 0\}. \end{aligned} \quad (۱۵)$$

بنابراین

$$S^c \cup \partial S = \bigcup_{i=1}^m \{x \in \mathfrak{R}^n : g_i(x) \geq 0\}$$

اگر  $L_i = \{x \in \mathfrak{R}^n : g_i(x) \geq 0\}$  را انتخاب نماییم، به آسانی می‌توان نشان داد که  $L_i$ ها مجموعه‌هایی بسته-اند و داریم:

$$C \cap (S^c \cup \partial S) = \bigcup_{i=1}^m C \cap L_i$$

به‌علاوه برای هر  $x \in C \cap L_i$  و هر  $i \in \{1, 2, \dots, m\}$  داریم:

$$g_i(x^*) - g_i(x) < 0.$$

به‌دلیل این‌که  $C \cap L_i$  فشرده است،  $\delta_{2,i} > 0$  به گونه‌ای وجود دارد که به‌ازای هر  $x \in C \cap L_i$

$$g_i(x^*) - g_i(x) < -\delta_{2,i} \quad (26)$$

با انتخاب

$$\delta_2 = \min\{\delta_{2,1}, \delta_{2,2}, \dots, \delta_{2,m}\} \quad (27)$$

چون به‌ازای هر  $i \in \{1, 2, \dots, m\}$   $g_i$ ها محدبند و  $x \in \mathfrak{R}^n$  داریم

$$-\gamma_i^T(t)(x - x^*) \leq g_i(x^*) - g_i(x). \quad (28)$$

با ترکیب (26)، (27) و (28)، به‌ازای هر  $x \in C \cap L_i$  داریم:

$$-\gamma_i^T(t)(x - x^*) < -\delta_2 \quad (29)$$

برای حالتی که  $x \in C \cap \partial S$  است، بدون کم شدن از کلیت می‌توان فرض نمود که به‌ازای  $i \in \{1, 2, \dots, k : k \leq m\}$   $x \in L_i$  است و به‌ازای  $i \in \{k+1, \dots, m : k < m\}$   $x \notin L_i$  است. از (25) با (29) داریم:

$$-\sum_{i=1}^m \theta(g_i(x)) \gamma_i^T(t)(x - x^*) \quad (30)$$

$$< -\sum_{i=1}^k \theta(g_i(x)) \delta_2$$

$$\frac{dV(x, x^*)}{dt} \leq 0 \quad (21)$$

لم ۴-۵. برای هر مجموعه فشرده  $C \subset \mathfrak{R}^n$  به‌طوری که  $C \cap S^* = \emptyset$  باشد،  $\zeta > 0$  چنان موجود است که

$$\frac{dV(x(t), x^*)}{dt} < -\zeta.$$

اثبات. بنا بر اثبات لم ۴-۴، اگر  $x(t)$  یک خط سیر جواب از شمول دیفرانسیلی (۲) باشد، به‌وضوح توابع اندازه‌پذیر  $\eta(t) \in \partial f(x(t))$  و  $\gamma_i(t) \in \partial g_i(x(t))$  به گونه‌ای وجود دارند که داریم:

$$\begin{aligned} \frac{dV(x(t), x^*)}{dt} &= -\eta(t)(x - x^*) - \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)(x - x^*) \quad (22) \\ &+ P \left[ \eta(t) + \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)(x - x^*) \right] \\ &- \|Ax - b\|_2^2. \end{aligned}$$

از تحدب  $f$  و بهینگی  $x^*$ ، به‌ازای هر  $x \in C \cap S_1$  داریم:

$$-\eta^T(x - x^*) \leq f(x^*) - f(x) < 0 \quad (23)$$

به دلیل این‌که  $f(x^*) - f(x)$  پیوسته و  $C \cap S_1$  فشرده است، پس  $\delta_1 > 0$  به گونه‌ای وجود دارد که به‌ازای هر  $x \in C \cap S_1$  داریم:

$$-\eta^T(x - x^*) \leq f(x^*) - f(x) < -\delta_1 \quad (24)$$

حال می‌خواهیم وجود  $\delta_2 > 0$  به گونه‌ای که برای هر  $x \in C \cap (S^c \cup \partial S)$  و  $\gamma_i \in \partial g_i(x)$  داشته باشیم

$$-\sum_{i=1}^m \theta(g_i(x)) \gamma_i^T(t)(x - x^*) < -\delta_2 \quad (25)$$

را نشان دهیم. به وضوح  $C \cap (S^c \cup \partial S)$  یک مجموعه فشرده است و داریم:



که در آن  $N = \|A^T(AA^T)^{-1}\|$  است. بنابراین  $\delta_4 > 0$  چنان موجود است که

$$|\gamma_i(t)^T P(x(t) - x^*)| < \frac{\delta_4}{m} \quad (37)$$

به عبارت دیگر داریم:

$$0 \leq \sum_{i=1}^m \theta(g_i(x)) < m \quad (38)$$

از روابط (37) و (38) داریم:

$$|\sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)^T P(x(t) - x^*)| < \delta_4 \quad (39)$$

به شیوه‌ای مشابه

$$|\eta^T(t) P(x(t) - x^*)| < \delta_4 \quad (40)$$

را می‌توان ثابت کرد. در نهایت بر طبق روابط (32) و (33) و با انتخاب  $\delta = \min\{\frac{\delta_1}{4}, \frac{\delta_2}{4}\}$  داریم:

$$-\eta^T(x - x^*) - \sum_{i=1}^m \theta(g_i(x)) \gamma_i^T(t) (x - x^*) < -\min\{\frac{\delta_1}{4}, \frac{\delta_2}{4}\} = -4\delta. \quad (41)$$

از (39) و (40) خواهیم داشت:

$$|\sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)^T P(x(t) - x^*) + \eta^T(t) P(x(t) - x^*)| < \delta_4 < \delta \quad (42)$$

از این رو

$$\begin{aligned} \frac{dV(x(t), x^*)}{dt} &= -\eta(t)(x - x^*) - \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)(x - x^*) + \sum_{i=1}^m \theta(g_i(x)) \gamma_i(t)^T P(x(t) - x^*) + \eta^T(t) P(x(t) - x^*) - \|Ax - b\|_2^2 < \delta - \min\{\delta_1, \delta_2\} = -3\delta = -\zeta. \end{aligned}$$

**قضیه ۴-۱۰.** اگر  $S^* \neq \emptyset$  باشد، آن‌گاه هر خط سیر جواب از شبکه عصبی (۲)، به یک جواب بهینه از مسأله (۱) همگراست.

واضح است که در این حالت  $\theta(g_i(x)) \in [0, 1]$ ، و با کمک استقرا می‌توان نشان داد

$$0 \leq \sum_{i=1}^k \theta(g_i(x)) < k \quad (31)$$

است. از (30) و (31) داریم:

$$-\sum_{i=1}^m \theta(g_i(x)) \gamma_i^T(t) (x - x^*) < -\delta_2 \quad (32)$$

برطبق قضیه ۴-۸،  $x(t)$  به  $S_2$  همگراست. بنابراین

$$0 < \delta_3 < \min\{\frac{\delta_1}{4}, \frac{\delta_2}{4}\} \quad (33)$$

$$\|Ax(t) - b\|_2^2 < \delta_3$$

برای ادامه اثبات، روابط زیر را ثابت می‌کنیم.

$$\begin{aligned} \eta^T P(x - x^*) &< \delta_4, \\ P \sum_{i=1}^m \theta(g_i(x)) \gamma_i^T(t) (x - x^*) &< \delta_4, \end{aligned} \quad (34)$$

توابع  $f$  و  $g_i$  موضعاً لیپ‌شیتس‌اند. همچنین می‌دانیم که  $\|x - x^*\|$  نزولی است. از لم ۴-۴ کران‌داری خط سیر جواب  $x(t)$  را داریم. بنابراین می‌توان فرض نمود که یک مجموعه فشرده  $K \subset \mathfrak{R}^n$  به گونه‌ای وجود دارد که  $x(t) \in K, \forall t > 0$ . از این رو یک ثابت مثبت  $M$  به طوری وجود دارد که برای هر  $x \in K$  و  $\gamma_i \in \partial g_i(x)$  و  $\eta \in \partial f(x)$   $\|\gamma_i(t)\|_2 < M$  و  $\|\eta(t)\|_2 < M$  را داریم. علاوه بر این داریم:

$$\lim_{t \rightarrow \infty} Ax(t) - b = 0. \quad (35)$$

بنابراین بر طبق توضیحات فوق برای هر  $i \in \{1, 2, \dots, m\}$

$$\begin{aligned} &\lim_{t \rightarrow \infty} |\gamma_i^T(t) P(x - x^*)| \\ &= \lim_{t \rightarrow \infty} |\gamma_i(t)^T A^T (AA^T)^{-1} A(x(t) - x^*)| \\ &= \lim_{t \rightarrow \infty} |\gamma_i(t)^T A^T (AA^T)^{-1} (Ax(t) - b)| \\ &\leq \lim_{t \rightarrow \infty} \|\gamma_i(t)\| \|A^T (AA^T)^{-1}\| \|Ax(t) - b\| \\ &< MN \|Ax(t) - b\| \rightarrow 0 \end{aligned} \quad (36)$$

$$2x_4^2 - 3x_5 \leq 0, \\ \sum_{i=1}^5 |x_i| \leq 10, x_1 + x_2 + x_3 + x_4 + \\ x_5 = 1$$

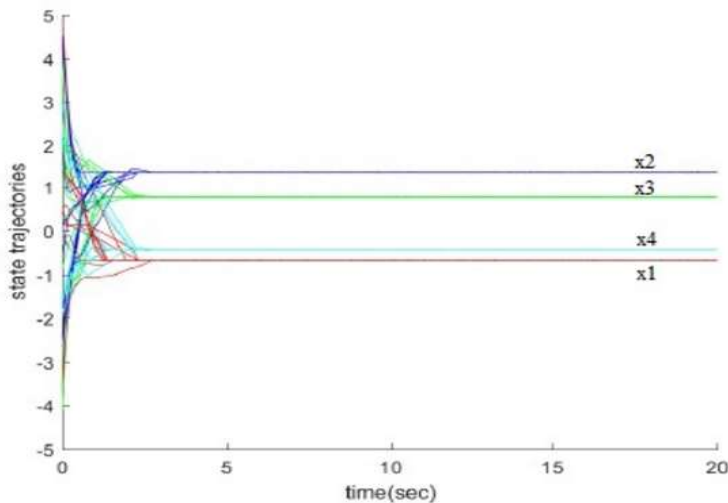
به‌وضوح تابع هدف  $f(x)$  محدب و ناهموار است. تنها جواب بهینه مسأله عبارت است از

$$x^* = (-0.3868, -0.2322, 0.5950, 0.7034, 0.3313)^T$$

شکل ۲. نشان می‌دهد که همه خط سیرهای جواب شبکه عصبی (۲) با ۱۰ نقطه آغازین تصادفی به جواب بهینه  $x^*$  هم‌گرایی دارند. همچنین  $x^*$  درون ناحیه شدنی واقع است و دقیقاً بر قید تساوی واقع است. برخلاف [23] نیازی به انتخاب تابع جریمه  $\epsilon(t)$  در مدل مان نداریم.

**مثال ۲.۵.** مسأله بهینه‌سازی نامحدب زیر که دارای قیود نامساوی و آفین مساوی است را در نظر بگیرید:

$$\min f(x) = 2|x_1 - x_2 + 2| + |x_2 - x_3 + 2x_4| + |x_3 + x_4 - 4| - \cos(0.2(x_1 - x_4)), \\ s. t. \quad g_1(x) = |2x_1 - x_2| + x_3^2 - 4 \leq 0, \\ g_2(x) = -x_2 - x_3 - |x_4 + 2| - 2 \leq 0, \\ g_3(x) = x_1 - x_4 - 5 \leq 0, \\ g_4(x) = x_4 - x_1 - 5 \leq 0, \\ x_1 + 2x_2 = 2, \\ -2x_3 + x_4 = -2.$$



شکل ۲. رفتار خطوط سیر جواب شبکه عصبی (۲) با ۱۰ نقطه آغازین تصادفی در مثال ۱.۵

**اثبات.** اثبات به آسانی مشابه قضیه ۹.۳ از [20] به دست می‌آید و آن را به خواننده واگذار می‌کنیم.

**قضیه ۴-۱۱.** هر نقطه تعادلی شمول دیفرانسیلی (۲) به‌طور نمایی پایدار سراسری است.

**اثبات.** اثبات به آسانی از ترکیب قضایای ۹-۴ و ۱۰-۴ حاصل می‌شود.

## ۵- مثال‌های عددی

در این بخش مثال‌هایی برای نشان دادن کارایی شبکه عصبی پیشنهادی (۲) برای حل مسأله بهینه‌سازی (۱) ارائه نموده‌ایم. در تمام مثال‌ها، شمول دیفرانسیلی تعریف شده با (۲) با نرم‌افزار متلب ورژن ۲۰۱۲ بر روی یک سیستم با مشخصات سی پی یو اینتل ۴.۲ گیگا هرتز پنج هسته‌ای

با سیستم عامل ویندوز ۷ و رم ۴ گیگا بایت حل شده‌اند.

### ۵-۱. مسأله بهینه‌سازی ناهموار

**مثال ۵.۱.** مسأله بهینه‌سازی ناهموار محدب زیر را در نظر بگیرید. این مثال قبلاً در [23] حل شده‌است ولی در اینجا برای حل آن از شبکه عصبی (۲) استفاده می‌کنیم.

$$\min f(x) = e^{|x_1+1|} + 3e^{x_2} + x_1^2 + 2x_2^2 + x_3^2 + |x_4| + e^{x_5} \\ s. t. \quad -x_1 + x_2 \leq 2, \\ x_1 - 2x_3 \leq 3,$$

مهندسی ایفا می‌کند از جمله در پردازش سیگنال و تصویر، ماشین بردار پشتیبان، کنترل دستی و مانند آن [۲۲]. این مسأله می‌تواند به صورت زیر نشان داده شود:

$$\begin{aligned} \min \quad & \|Cx - d\|_1 \\ \text{s.t.} \quad & g(x) \leq 0, \\ & Ax = b, x \in \Omega \end{aligned} \quad (43)$$

که اجزای این مسأله مشابه مسأله (۱) و  $\Omega \subset \mathbb{R}^n$  یک مجموعه محدب است. به‌وضوح وقتی  $g(x)$  محدب است، این مسأله یک مسأله بهینه‌سازی ناهموار محدب است.

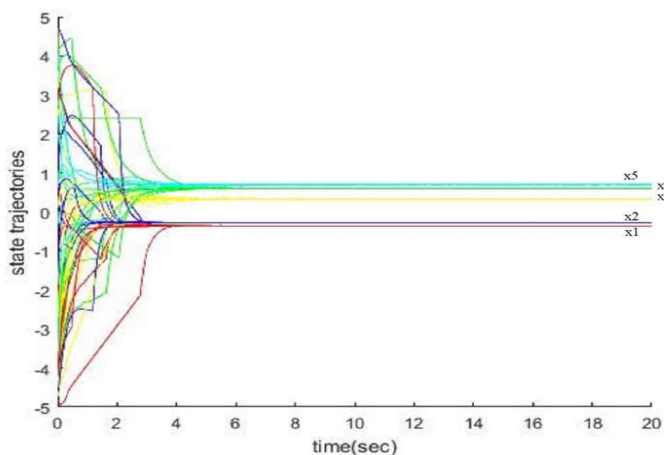
**مثال ۵.۳.** مسأله مینیمم‌سازی نرم  $L_1$  زیر با تعدادی قید که در [۲۲] مطرح و حل شده است را در نظر بگیرید:

$$\begin{aligned} \min \quad & \|Cx - d\|_1, \\ \text{s.t.} \quad & x_1^2 - x_2 + x_3 \leq 0, \\ & x_1 + x_2 + x_3 = 0, \\ & 2x_2 + x_3 = 2, \end{aligned} \quad (44)$$

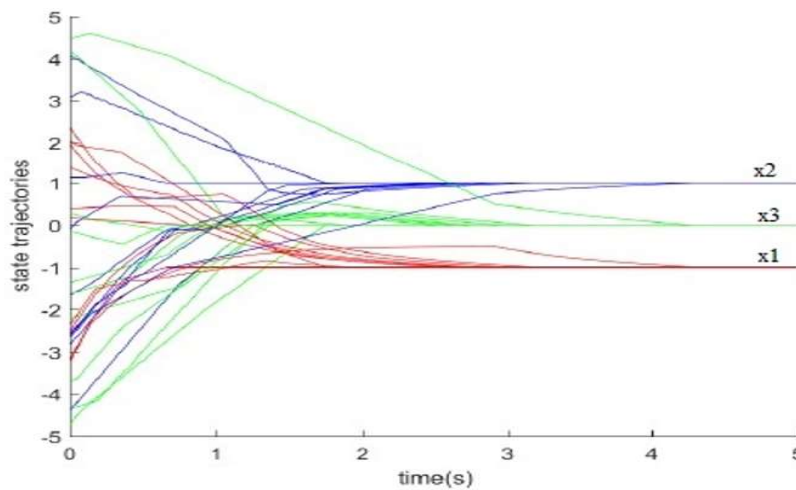
بر طبق [26]، وقتی  $-5 \leq x_1 - x_4 \leq 5$  است،  $\cos(0.2(x_1 - x_4))$  محدب است. از این‌رو تابع هدف بر  $S_1$  محدب است و می‌توانیم از شبکه عصبی (۲) برای حل آن استفاده کنیم. جواب بهینه مسأله عبارت است از  $x^* = (-0.6667, 1.3333, 0.8889, -0.2222)^T$ . که مقدار بهینه متناظر  $f(x^*) = 2.3373$  می‌باشد. مسأله را با ۱۰ نقطه آغازین نشدنی حل می‌کنیم. مشاهده می‌شود که تمام خط‌سیرهای جواب شبکه عصبی پیشنهادی به جواب بهینه همگرايند. به‌علاوه در ساختار شبکه پیشنهادی از هیچ تابع یا پارامتر جریمه‌ای استفاده نکرده‌ایم. همچنین نیازی به یافتن یک کران بالا برای ثابت‌های لیپ‌شیتس توابع  $f(x)$  و  $g_i$  نیست. رفتار گذرای مدل شبکه عصبی پیشنهادی برای حل این مثال در شکل ۳. نشان داده شده است.

### ۲-۵. مسأله مینیمم‌سازی نرم $L_1$

مسأله مینیمم‌سازی نرم  $L_1$  به مسأله حداقل انحراف مطلق مقید مشهور است و نقشی مهم در کاربردهای



شکل ۳. رفتار خطوط سیر جواب شبکه عصبی (۲) با ۱۰ نقطه آغازین تصادفی در مثال ۲.۵



شکل ۴. رفتار خطوط سیر جواب شبکه عصبی (۲) با ۱۰ نقطه آغازین تصادفی در مثال ۳.۵

متناظر با آن  $\|Cx^* - d\|_1 = 10.5$  می‌باشد که کاملاً سازگار و منطبق بر نتایج به‌دست آمده در [۲۲] است. در مقایسه با شبکه‌های عصبی یاد شده مدل شبکه عصبی پیشنهادی تک لایه است که دارای پیچیدگی کمتری است.

#### نتیجه‌گیری

در این مقاله یک شبکه عصبی بازگشتی تک لایه غیر جریمه‌ای را برای حل مسائل بهینه‌سازی محدب ناهموار با قیود نامساوی غیرخطی و مساوی خطی پیشنهاد نموده‌ایم. با استفاده از تجزیه و تحلیل شمول دیفرانسیلی، همگرایی شبکه پیشنهادی به جواب بهینه و همچنین پایداری شمول دیفرانسیلی پیشنهادی را اثبات نموده‌ایم. همچنین کارایی و اثربخشی مدل پیشنهادی را با دو مثال گویا نشان داده و به‌عنوان کاربرد از آن در حل مسئله مینیمم‌سازی نرم  $L_1$  استفاده کرده‌ایم. در آینده‌ای نزدیک به توسعه مدل پیشنهادی برای حل مسائل بهینه‌سازی نامحدب ناهموار با قیود کلی خواهیم پرداخت.

که در آن

$$C = \begin{pmatrix} 1 & 0 & 3.5 \\ 0 & 2 & 1.6 \\ 1 & 1 & 1 \end{pmatrix}, \quad d = \begin{pmatrix} 1.5 \\ 3.8 \\ -6.2 \end{pmatrix}.$$

است.  $x \in \mathbb{R}^3$  و این مسئله ناهموار و محدب است. بنابراین می‌توانیم از شبکه عصبی (۲) برای حل این مسئله استفاده کنیم. با توجه به مفروضات مسئله تابع هدف به صورت زیر می‌باشد:

$$f(x) = |x_1 + x_2 + x_3 + 6.2| + |2x_2 + 1.6x_3 - 3.8| + |x_1 + 3.5x_3 - 1.5|$$

شبکه عصبی (۲) توسط نرم‌افزار متلب شبیه‌سازی شده است. شکل ۴. رفتار خط سیرهای جواب شبکه عصبی پیشنهادی (۲) با ۱۰ نقطه آغازین تصادفی را نشان می‌دهد. مشاهده می‌شود که حالت‌های شبکه عصبی (۲) با نقاط آغازین تصادفی به  $(-1, 1, 0)^T$  همگرایی، که یک نقطه تعادلی شبکه عصبی (۲) است. بنابراین جواب بهینه مسئله  $x^* = (-1, 1, 0)^T$  است و مقدار بهینه تابع هدف

Transactions on Neural Networks, 19(4), pp.558-570.

[9] Beliakov, G. and Bagirov, A., 2006. Non-smooth optimization methods for computation of the conditional value-at-risk and portfolio optimization Optimization, 55 (5), pp. 459-479.

[10] Soleimani-Damaneh, M., 2010. Nonsmooth optimization using Mordukhovich's subdifferential. SIAM Journal on Control and Optimization, 48(5), pp.3403-3432.

[11] Mordukhovich, B.S., 2006. Variational analysis and generalized differentiation I: Basic theory (Vol. 330). Springer Science & Business Media.

[12] Mordukhovich, B.S., 2006. Grundlehren Series (Fundamental Principles of Mathematical Sciences) 330. In Variations Analysis and Generalized Differentiation, II: Applications. Springer.

[13] Ebadi, M.J., Hosseini, A. and Hosseini, M.M., 2017. A projection type steepest descent neural network for solving a class of nonsmooth optimization problems. Neurocomputing, 235, pp.164-181

[14] Ebadi, M.J., Hosseini, M.M., Karbassi, S.M., 2018. An efficient one-layer recurrent neural network for solving a class of nonsmooth pseudoconvex optimization problems. Journal of Theoretical and Applied Information Technology, 96(7), pp.1999-2014.

[15] Forti, M., Nistri, P. and Quincampoix, M., 2004. Generalized neural network for nonsmooth nonlinear programming problems. IEEE Transactions on Circuits and Systems I: Regular Papers, 51(9), pp.1741-1754.

[1] Tank, D. and Hopfield, J.J., 1986. Simple neural optimization networks: An a/d converter, signal decision circuit, and a linear programming circuit. IEEE transactions on Circuits and Systems, 33(5), pp.533-541.

[2] Kennedy, M.P. and Chua, L.O., 1988. Neural networks for nonlinear programming. IEEE Transactions on Circuits and Systems, 35(5), pp.554-562.

[3] Zhang, S. and Constantinides, A.G., 1992. Lagrange programming neural networks. IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 39(7), pp.441-452.

[4] Wang, J. and Xia, Y., 1998. Analysis and design of primal-dual assignment networks. IEEE Transactions on Neural Networks, 9(1), pp.183-194.

[5] Xia, Y., Leung, H. and Wang, J., 2002. A projection neural network and its application to constrained optimization problems. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 49(4), pp.447-458.

[6] Xia, Y.S. and Wang, J., 2000. On the stability of globally projected dynamical systems. Journal of Optimization Theory and Applications, 106(1), pp.129-150.

[7] Liu, Q. and Wang, J., 2008a. A one-layer recurrent neural network with a discontinuous activation function for linear programming. Neural Computation, 20(5), pp.1366-1383.

[8] Liu, Q. and Wang, J., 2008b. A one-layer recurrent neural network with a discontinuous hard-limiting activation function for quadratic programming. IEEE

- network for pseudoconvex optimization problems with equality and inequality constraints. *IEEE transactions on cybernetics*.
- [24] Aubin, J.P. and Cellina, A., 1984. *Differential inclusions: set-valued maps and viability theory*, vol. 264. Grundlehren der mathematischen Wissenschaften.
- [25] Clarke, F.H., 1990. *Optimization and nonsmooth analysis*. Society for Industrial and Applied Mathematics.
- [26] Liu, Q. and Wang, J., 2011. A one-layer recurrent neural network for constrained nonsmooth optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41(5), pp.1323-1333.
- [27] Hosseini, A., Wang, J. and Hosseini, S.M., 2013. A recurrent neural network for solving a class of generalized convex optimization problems. *Neural Networks*, 44, pp.78-86.
- [28] Ebadi, M.J., Hosseini, M.M. and Ali-Akbari, M. 2017. Implementation of recurrent neural networks for solving constrained optimization problems by analog circuits, 4 th national Conference on Application of Novel Technologies in Engineering, Torbat Heydaryeh, Iran.
- [16] Bian, W. and Xue, X., 2009. Subgradient-based neural networks for nonsmooth nonconvex optimization problems. *IEEE Transactions on Neural Networks*, 20(6), pp.1024-1038.
- [17] Li, G., Song, S., Wu, C. and Du, Z., 2006. A neural network model for nonsmooth optimization over a compact convex subset. *Advances in Neural Networks*, pp.344-349.
- [18] Liu, Q. and Wang, J., 2006. A recurrent neural network for non-smooth convex programming subject to linear equality and bound constraints. In *Neural Information Processing* (pp. 1004-1013). Springer Berlin/Heidelberg.
- [19] Liu, Q. and Wang, J., 2008c, November. A one-layer recurrent neural network for non-smooth convex optimization subject to linear equality constraints. In *International Conference on Neural Information Processing* (pp. 1003-1010). Springer, Berlin, Heidelberg.
- [20] Hosseini, A., Hosseini, S.M. and Soleimani-damaneh, M., 2013. A differential inclusion-based approach for solving nonsmooth convex optimization problems. *Optimization*, 62(9), pp.1203-1226.
- [21] Hosseini, A., 2016. A non-penalty recurrent neural network for solving a class of constrained optimization problems. *Neural Networks*, 73, pp.10-25.
- [22] Qin, S. and Xue, X., 2015. A two-layer recurrent neural network for nonsmooth convex optimization problems. *IEEE transactions on neural networks and learning systems*, 26(6), pp.1149-1160.
- [23] Qin, S., Yang, X., Xue, X. and Song, J., 2016. A one-layer recurrent neural