

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده علوم ریاضی

پایان نامه کارشناسی ارشد بهینه سازی

یک رویکرد شبکه عصبی برای حل مسائل بهینه سازی چندهدفه محدب

نگارنده: زهرا کریمی

استادان راهنما

دکتر علیرضا ناظمی
دکتر مهرداد غزنوی

اسفند ۱۴۰۰

تقدیم بہ

مادر، پدر، خواہر، برادر و ہمسر عزیز و مہربانم کہ در ہمہ مراحل زندگی پشتیبانیم بودہ اند و ہر آن چہ را بہ دست آوردہ ام، مدیون حضور پر مہرشان ہستم.

سپاس‌گزاری...

سپاس مخصوص پروردگاری است که بدون توجهات و نگاه پرمهرش، انجام این مهم از عهده‌ام خارج بود.

در اینجا لازم می‌دانم از تمامی افرادی که به طریقی مرا در انجام ثایان‌نامه یاری دادند، تشکر نمایم. به خصوص از آقایان دکتر علیرضا ناظمی و دکتر مهرداد غزنوی که راهنمایی‌هایشان بسیار روشنگر و آموزنده بوده است، کمال تشکر و قدردانی را داشته باشم.

زهره کریمی

اسفند ۱۴۰۰

تعهد نامه

اینجانب زهرا کریمی دانشجوی کارشناسی ارشد رشته ریاضی کاربردی دانشکده علوم ریاضی دانشگاه صنعتی شاهرود، نویسنده پایان نامه با عنوان یک رویکرد شبکه عصبی برای حل مسائل بهینه سازی چندهدفه محدب، تحت راهنمایی دکتر علیرضا ناظمی متعهد می شوم:

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

زهرا کریمی

اسفند ۱۴۰۰

مالکیت نتایج و حق نشر

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

چکیده

در این پایان نامه یک رویکرد شبکه عصبی جدید گرادیان محور برای حل مسائل بهینه سازی غیرخطی، دوهدفه، چندهدفه و چندسطحی پیشنهاد شده است. ایده اصلی تبدیل مسائل بهینه سازی غیر خطی و دوهدفه با روش مجموع وزن دار، به مسائل بهینه سازی محدب، مسائل بهینه سازی چند هدفه با استفاده از نقطه مرجع به مسائل تک هدفه و مسائل بهینه سازی چندسطحی با شرایط کروش-کان-تاگر به مسئله تک سطحی می باشد. در واقع با روش مجموع وزن دار جواب های بهینه پارتو با وزن های مختلف بدست می آیند. همچنین فضای پارامتری برای مسائل بهینه سازی دوهدفه با استفاده از مجموعه پایداری نوع اول تحلیل می شود.

در نهایت یک شبکه عصبی برای حل مسائل برنامه ریزی محدب و مسائل بهینه سازی تک هدفه ارائه می شود که بر اساس قضیه لیاپانوف پایدار سراسری است و به یک جواب بهینه پارتو مسائل بهینه سازی چندهدفه و مسائل بهینه سازی چند سطحی همگرا می باشد. مهم ترین ویژگی رویکرد پیشنهادی این است که می تواند به سرعت به نقطه تعادل (جواب بهینه) همگرا شود. نتایج شبیه سازی نشان می دهند که شبکه عصبی پیشنهادی شدنی و کارا است.

کلمات کلیدی: مسائل بهینه سازی دوهدفه، مسائل برنامه ریزی غیر خطی، مسائل بهینه سازی چندهدفه، مسائل بهینه سازی چندسطحی، جواب های بهینه پارتو، شرایط کروش-کان-تاگر، شبکه عصبی.

فهرست مطالب

م	فهرست تصاویر
س	فهرست جداول
۱	مقدمه‌ای بر بهینه‌سازی چندهدفه و چندسطحی
۱	۱.۱ نمادها
۲	۲.۱ تعاریف
۴	۳.۱ مسائل بهینه‌سازی چندهدفه (MOOP)
۴	۱.۳.۱ مسائل بهینه‌سازی چند هدفه
۶	۲.۳.۱ فضای تصمیم‌گیری و فضای توابع هدف (معیار)
۱۰	۳.۳.۱ روش مجموع وزنی
۱۱	۴.۳.۱ روش بهینه‌سازی توافقی
۱۳	۴.۱ مسائل بهینه‌سازی چندسطحی
۱۳	۱.۴.۱ روش کروش-کان-تاکر برای حل مسائل برنامه‌ریزی دوسطحی خطی
۲۳	۲ شبکه‌های عصبی
۲۳	۱.۲ شبکه‌های عصبی طبیعی
۲۵	۲.۲ شبکه‌های عصبی مصنوعی
۲۶	۳.۲ توپولوژی شبکه‌های عصبی مصنوعی
۲۹	۱.۳.۲ تاریخچه‌ای از حل مسائل بهینه‌سازی با استفاده از شبکه‌های عصبی
۳۰	۲.۳.۲ مروری بر سیستم‌های دینامیکی
۳۳	۳ تجزیه فضای پارامتریک برای مسائل بهینه‌سازی دوهدفه
۳۴	۱.۳ مجموعه پایداری نوع اول
۳۵	۱.۱.۳ محاسبه مجموعه پایداری نوع اول
۳۶	۲.۳ رویکرد شبکه عصبی پیشنهادی برای مسائل بهینه‌سازی دوهدفه
۳۶	۱.۲.۳ شرح رویکرد شبکه عصبی پیشنهادی

۴۰	تجزیه و تحلیل پایداری	۲۰۲۰۳
۴۲	مثال‌های عددی	۳۰۳
۴۲	مسائل برنامه‌ریزی غیرخطی	۱۰۳۰۳
۴۲	تجزیه و تحلیل نتایج برای مسائل برنامه‌ریزی غیرخطی	۲۰۳۰۳
۴۸	مسائل بهینه‌سازی دوهدفه	۳۰۳۰۳
۴۸	تجزیه و تحلیل نتایج مسائل بهینه‌سازی دوهدفه	۴۰۳۰۳
۵۷	تجزیه فضاها برای پارامتری مسائل بهینه‌سازی دوهدفه	۵۰۳۰۳
۴	ترکیب شرایط کان-تاکر و رویکرد شبکه عصبی برای مسائل بهینه‌سازی چندسطحی و چندهدفه	۶۱
۶۱	بررسی رویکرد شبکه عصبی پیشنهادی	۱۰۴
۶۶	مثال‌های عددی	۲۰۴
۷۵	آزمون رتبه‌ی علامت‌دار ویلکاکسون	۳۰۴
۸۱	مراجع	

فهرست تصاویر

۱.۱	نمودار توابع هدف مثال ۳.۳.۱	۶
۲.۱	فضای هدف (معیار) در مثال ۱.۳.۱	۷
۳.۱	فضای معیار مثال ۳.۳.۱	۸
۴.۱	نمودار هندسی مثال بهینه‌سازی دوسطحی (۴.۱)	۲۲
۱.۲	ساختار نرون	۲۴
۲.۲	مدل ریاضی نرون عصبی مصنوعی	۲۷
۳.۲	انواع تابع فعال‌سازی	۲۸
۱.۳	چهار مسیر با نقاط اولیه متفاوت برای p_1	۴۵
۲.۳	چهار مسیر با نقاط اولیه متفاوت برای p_2	۴۶
۳.۳	چهار مسیر با نقاط اولیه متفاوت برای p_3	۴۶
۴.۳	چهار مسیر با نقاط اولیه متفاوت برای p_4	۴۷
۵.۳	چهار مسیر با نقاط اولیه متفاوت برای p_5	۴۷
۶.۳	هشت مسیر با دو نقطه اولیه برای مسئله p_6 در برابر تعداد تکرارها	۴۸
۱.۴	مسیرهایی با نقاط مختلف برای مثال ۱.۲.۴ (آ) مسیره‌ای جواب با $p = 1$	
۶۹	(ب) مسیره‌ای جواب با $p = 2$ (ج) مسیره‌ای جواب با $p = \infty$	
۲.۴	مسیرهایی با دو نقطه اولیه مختلف برای $p = 1, p = 2, p = \infty$ و (آ) دو مسیر با نقاط اولیه متفاوت برای $p = 1$ (ب) دو مسیر با نقاط اولیه متفاوت برای $p = 2$ (ج) دو مسیر با نقاط اولیه متفاوت برای $p = \infty$	
۷۰	رفتار همگرایی مسیره‌ها از طریق تکرارها (آ) مقایسه دو مسیر تکرار برای $p = 1$	
۳.۴	(ب) مقایسه دو مسیر تکرار برای $p = 2$ (ج) مقایسه دو مسیر تکرار برای $p = \infty$	
۷۴	رفتار مسیره‌ها برای مثال ۱.۳.۴ (آ) مسیره‌ای جواب برای نقاط اولیه مختلف	
۷۸	(ب) مسیره‌ها در تعداد دفعات	
۸۰	مسیره‌ای تکرار برای مثال ۲.۳.۴	

فهرست جداول

۱۰۱	معیارها و متغیرها در مثال ۱۰.۳.۱	۵
۱۰۳	مسائل برنامه‌ریزی غیرخطی	۴۳
۲۰۳	مقادیر بهینه مسائل بهینه‌سازی محدب	۴۵
۳۰۳	مسائل بهینه‌سازی دوهدفه	۴۹
۴۰۳	ادامه مسائل بهینه‌سازی دوهدفه	۵۰
۵۰۳	ادامه مسائل بهینه‌سازی دوهدفه	۵۱
۶۰۳	مسائل بهینه‌سازی دوهدفه	۵۳
۷۰۳	ادامه مسائل بهینه‌سازی دوهدفه	۵۴
۸۰۳	ادامه مسائل بهینه‌سازی دوهدفه	۵۵
۱۰۴	مقایسه حداقل فاصله با استفاده از روش پیشنهادی با روش های ارائه شده در منابع [۲۹، ۲۱]	۶۸
۲۰۴	مقایسه مینیم فاصله با روش پیشنهادی و آنچه در [۳۰] ارائه شده است.	۷۳
۳۰۴	نتایج آماری روش شبکه عصبی پیشنهادی برای فواصل مختلف	۷۵
۴۰۴	آزمون مقایسه‌ای رتبه‌ی علامت‌دار ویلکاکسون برای مقایسه نتایج فاصله‌های مختلف	۷۶

فصل ۱

مقدمه‌ای بر بهینه‌سازی چندهدفه و چندسطحی

در این پایان‌نامه به معرفی شبکه عصبی برای حل مسائل بهینه‌سازی چندهدفه و چندسطحی می‌پردازیم. در دهه‌های اخیر مسائل بهینه‌سازی چندهدفه (MOOPs)^۱ برای دانشمندان و محققان به یک زمینه تحقیقاتی مهم تبدیل شده است. در مسائل بهینه‌سازی چندهدفه، توابع باید به طور هم‌زمان بهینه شوند. در واقع، یافتن جوابی که با توجه به همه اهداف بهترین باشد، دشوار است. جواب ممکن است برای یک هدف بهترین اما برای دیگری بدترین باشد. بنابراین، در این زمینه، معمولاً می‌توان مجموعه‌ای از جواب‌ها را یافت. چنین مجموعه‌ای را مجموعه‌ی جواب‌های غیرمغلوب یا جواب‌های بهینه پارتو می‌نامند. در این فصل یک سری مفاهیم اولیه، نمادها، تعاریف اساسی که در طول این پایان‌نامه لازم است را بیان می‌کنیم.

۱.۱ نمادها

فرض کنید $\mathbb{R}$ نمایش اعداد حقیقی و $\mathbb{R}_{\geq} = [0, \infty)$ باشد. ناحیه نامنفی^۲ از $\mathbb{R}^p$ را با $\mathbb{R}_{\geq}^p$ نشان می‌دهیم به عبارت دیگر

$$\mathbb{R}_{\geq}^p := \{y \in \mathbb{R}^p : y \geq 0\}$$

¹Multi Objective Optimization Problems

²Nonnegative orthant

و ناحیه نامنفی به جز مبدا را به صورت زیر

$$\mathbb{R}_{\geq}^p := \{y \in \mathbb{R}^p : y \geq 0\} = \mathbb{R}_{\geq}^p \setminus \{0\}$$

و ناحیه مثبت^۱ را به صورت

$$\mathbb{R}_{>}^p := \{y \in \mathbb{R}^p : y > 0\}$$

نمایش خواهیم داد. فرض کنید S_1 و S_2 زیر مجموعه‌های ناتهی و معین از $\mathbb{R}^p$ باشند. مجموع جبری S_1 و S_2 را با $S_1 + S_2$ نمایش می‌دهیم و به صورت زیر تعریف می‌کنیم.

$$S_1 + S_2 = \{s = s_1 + s_2 \mid s_1 \in S_1, s_2 \in S_2\}$$

. برای مجموعه‌های دلخواه $S \subset \mathbb{R}$ ، نمادهای زیر را تعریف می‌کنیم.

۱. $int(S)$ ، درون S است.

۲. $bd(S)$ ، مرز S است.

۳. $conv(S)$ ، پوسته محدب^۲ S است.

۴. $cone(S)$ ، مخروط تولید شده به وسیله S است.

۲.۱ تعاریف

در این بخشی برخی از تعاریف را که در ادامه نیاز داریم، بیان می‌کنیم. یک رابطه دوتایی (باینری)^۳ روی S ، یک زیر مجموعه R از $S \times S$ است. برخی از خواص رابطه دوتایی را در زیر بیان می‌کنیم [۱۰].

تعریف ۱.۲.۱. یک رابطه دوتایی R روی S را می‌گوییم [۱۰]:

۱. بازتابی^۴ اگر $(s, s) \in R$ برای هر $s \in S$.

۲. غیربازتابی^۵ اگر $(s, s) \notin R$ برای هر $s \in S$.

۳. متقارن^۶ اگر $(s^1, s^2) \in R$ آنگاه $(s^2, s^1) \in R$ ، برای هر $s^1, s^2 \in S$.

۴. غیرمقارن^۷ اگر $(s^1, s^2) \in R$ آنگاه $(s^2, s^1) \notin R$ ، برای هر $s^1, s^2 \in S$.

¹Positive orthant

²Convex hull

³Binary relation

⁴Reflexive

⁵Irreflexive

⁶Symmetric

⁷Assymmetric

۵. پادمتقارن^۱ اگر $(s^1, s^2) \in R$ و $(s^2, s^1) \in R$ آنگاه $s_1 = s_2$ ، برای هر $s^1, s^2 \in S$.
۶. تعدی^۲ گویند هر گاه $(s^1, s^2) \in R$ و $(s^2, s^3) \in R$ آنگاه $(s^1, s^3) \in R$ ، برای هر $s^1, s^2, s^3 \in S$.
۷. تعدی منفی^۳ گویند هر گاه $(s^1, s^2) \notin R$ و $(s^2, s^3) \notin R$ آنگاه $(s^1, s^3) \in R$ برای هر $s^1, s^2, s^3 \in S$.
۸. همبند^۴ اگر $(s^1, s^2) \in R$ یا $(s^2, s^1) \in R$ و برای هر $s^1, s^2 \in S$ با $s^1 \neq s^2$.

تعریف ۲.۲.۱. یک رابطه دوتایی R روی یک مجموعه $[10]$

۱. یک رابطه هم ارزی است اگر بازتابی، متقارن و تعدی باشد.

۲. یک رابطه شبه مرتب است هرگاه بازتابی و تعدی باشد.

تعریف ۳.۲.۱. یک رابطه دوتایی $\preceq [10]$ را:

۱. مرتب جزئی^۵ گویند هر گاه بازتابی، تعدی و پاد متقارن باشد.

۲. مرتب جزئی اکید^۶ گویند هر گاه غیر متقارن و تعدی باشد.

تعریف ۴.۲.۱. برخی ترتیب‌ها روی $\mathbb{R}^P$ عبارتند از:

۱. $y^1 \leq y^2$ را ترتیب مولفه‌ای ضعیف گویند هر گاه $y_k^1 \leq y_k^2 \quad k = 1, 2, \dots, p$ باشد.

۲. $y^1 \leq y^2$ را ترتیب مولفه‌ای گویند هر گاه $y_k^1 \leq y_k^2 \quad k = 1, 2, \dots, p; y^1 \neq y^2$ باشد.

۳. $y^1 < y^2$ را ترتیب مولفه‌ای اکید گویند هر گاه $y_k^1 < y_k^2 \quad k = 1, 2, \dots, p$ باشد.

تعریف ۵.۲.۱. زیر مجموعه k از $\mathbb{R}^P$ را محدب گویند هر گاه به ازای هر $c_1, c_2 \in k$ و هر $\alpha \in [0, 1]$ داشته باشیم، $\alpha c_1 + (1 - \alpha)c_2 \in k$ [۱۴].

تعریف ۶.۲.۱. یک تابع $f: \mathbb{R}^P \rightarrow \mathbb{R}$ محدب نامیده می‌شود اگر برای هر $x_1, x_2 \in \mathbb{R}^P$ و هر $\alpha \in [0, 1]$ داشته باشیم، $f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$ [۱۴].

تعریف ۷.۲.۱. یک مخروط k در $\mathbb{R}^P$ را نابدهی نامیم هر گاه $k \neq \emptyset$ و $k \neq \mathbb{R}^P$ باشد [۱۴].

گزاره ۱. یک زیر مجموعه k از $\mathbb{R}^P$ مخروط محدب است اگر و فقط اگر

¹Antisymmetric

²Transitive

³Negative transitive

⁴Connected

⁵Quasi-order

⁶Partial order

Strict partial order

۱. به ازای هر $d \in k$ و هر $\alpha > 0$ داشته باشیم $\alpha d \in k$

۲. $x^1, x^2 \in k$ ایجاب کند $x^1 + x^2 \in k$ [۱۴].

گزاره ۲. یک زیر مجموعه k از $\mathbb{R}^P$ را محدب است اگر و فقط اگر $d_i \in k, (i = 1, 2, \dots, n)$ و هر $\sum_{i=1}^n \alpha_i d_i \in k$ نتیجه دهد $\sum_{i=1}^n \alpha_i = 1$ و $\alpha_i \geq 0, (i = 1, 2, \dots, n)$ یعنی شامل تمام ترکیبات محدب از عناصرش باشد [۱۴].

تعریف ۸.۲.۱. مجموعه تمام ترکیبات محدب نقاط در k را پوسته محدب مجموعه k گوئیم و با $conv(k)$ نمایش می‌دهیم.

$$conv(k) = \{\theta_1 x_1 + \dots + \theta_n x_n | x_i \in k, \theta_i \geq 0, i = 1, 2, \dots, n, \theta_1 x_1 + \dots + \theta_n x_n = 1\}$$

به علاوه پوسته محدب همیشه محدب است و $conv(k)$ کوچکترین مجموعه محدب شامل k است.

تعریف ۹.۲.۱. یک تابع حقیقی f پیوسته لیپ‌شیتز است اگر ثابت حقیقی مثبت K وجود داشته باشد، به‌طوری که برای همه مقادیر حقیقی x_1 و x_2 داشته باشیم.

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2|$$

۳.۱ مسائل بهینه‌سازی چندهدفه (MOOP)

مسائل بهینه‌سازی چندهدفه به مسائلی گفته می‌شود که بیشتر از یک تابع هدف دارند و نمیتوان توابع هدف را با هم ترکیب کرد چون متغیر و ضریب مشترکی ندارند. برای مثال هزینه‌ی جابجایی در دو مسیر متفاوت توسط پارامتری به نام هزینه‌ی جابجایی و متغیری به نام فاصله‌ی بین دو نقطه تعریف می‌شوند پس میتوان این دو تابع هدف را با هم جمع کرد و تحت عنوان تابع هدف کل هزینه‌ی جابجایی تعریف کرد اما اگر بخواهیم هزینه‌ی جابجایی و زمانبندی دو مسیر مختلف را بهینه کنیم چون یک تابع هدف از جنس هزینه و دیگری از جنس زمان است نمیتوان با هم ترکیب کرد، در نتیجه با یک مسئله‌ی چند هدفه مواجه هستیم. در این بخش ابتدا چند مثال برای تشخیص انواع مسائل تصمیم‌گیری معرفی می‌کنیم. در واقع مسائل بهینه‌سازی را می‌توان به عنوان مدل‌های ریاضی مسائل تصمیم‌گیری معرفی نمود. مفاهیم فضای تصمیم (یا متغیرها) و معیارها (یا توابع هدف) و اصول بهینگی را شرح می‌دهیم. مطالب این بخش عمدتاً از کتاب ارگات [۱۰] گرفته شده است.

۱.۳.۱ مسائل بهینه‌سازی چند هدفه

ابتدا سه مثال از مسائل تصمیم‌گیری را بررسی می‌کنیم.

مثال ۱.۳.۱. فرض کنید بخواهیم یک ماشین جدید بخریم و در بازار چهار نوع ماشین با نام‌های آپل، واگن گلف، فورد و تویوتا موجود باشد. از بین این چهار برند می‌خواهیم یکی را انتخاب کنیم.

معیارهای ما قیمت، مصرف سوخت و قدرت ماشین هستند. واضح است که به دنبال ماشینی هستیم که کمترین قیمت، کمترین مصرف سوخت و بیشترین قدرت را داشته باشد. در اینجا در واقع یک مسئله بهینه‌سازی داریم که دارای سه تابع هدف (معیار^۱ و چهار متغیر^۲ است. جدول ۱.۱ نشان دهنده داده‌های ماست.

جدول ۱.۱: معیارها و متغیرها در مثال ۱.۳.۱

جایگزین‌ها				
تویوتا	فورد	اوپل	واگن گلف	
۱۵/۲	۱۴/۰	۱۴/۹	۱۶/۲	قیمت (۱۰۰۰ یورو)
۸/۲	۷/۵	۷/۰	۷/۲	مصرف سوخت ($\frac{l}{100km}$)
۷۱/۰	۵۵/۰	۶۲/۰	۶۶/۰	قدرت (kW)

مثال ۲.۳.۱. فرض کنید می‌خواهیم سدی بسازیم که در عین حال که انرژی به دست آمده از این سد ماکزیمم شود میزان تبخیر آب در آن باید مینیمم شود. واضح است که یک مسئله بهینه‌سازی با دو هدف متناقض داریم.

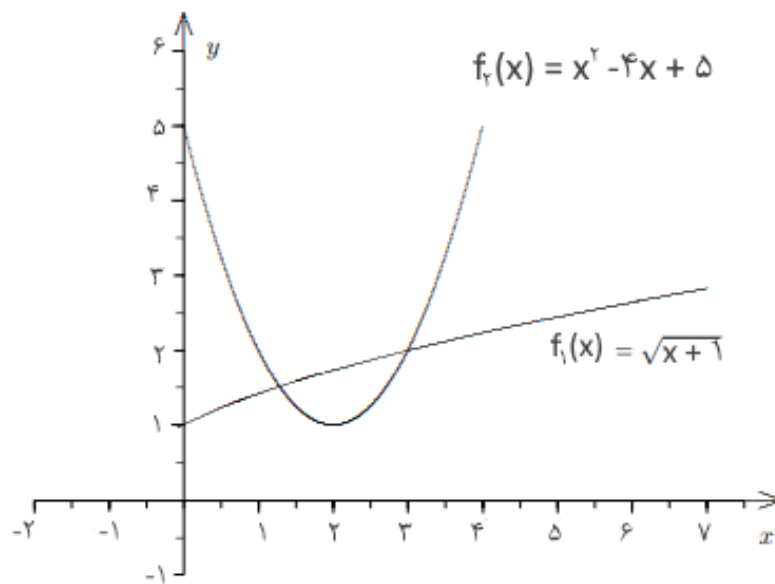
مثال ۳.۳.۱. فرض کنیم می‌خواهیم توابع $f_1(x) = \sqrt{x+1}$ و $f_2(x) = x^2 - 4x + 5 = (x-2)^2 + 1$ را روی خط حقیقی مثبت به طور همزمان مینیمم کنیم. سوال این است که جوابی که به طور همزمان این دو را مینیمم کند، کدام است؟ در واقع داریم:

$$\text{Min}_{x \geq 0}(f_1(x), f_2(x)) = \text{Min}(\sqrt{x+1}, (x-2)^2 + 1) \quad (1.1)$$

به آسانی می‌توان هر کدام از توابع هدف را به تنهایی مینیمم کرد. $x_1 = 0$ و $x_2 = 2$ به ترتیب نقاط بهینه توابع f_1 و f_2 روی $x \in \mathbb{R}: x \geq 0$ هستند، که نمودار توابع هدف در شکل ۱.۱ نشان داده شده است.

¹criteria

²alternative



شکل ۱.۱: نمودار توابع هدف مثال ۳.۳.۱

مسائل بهینه‌سازی چندهدفه (MOOP) را می‌توان به صورت زیر فرمول‌بندی کرد.

$$\begin{aligned} \text{Min} \quad & f(x) = (f_1(x), f_2(x), \dots, f_K(x))^T \\ \text{subject to} \quad & x \in \Omega \\ \Omega = \{ & x \in \mathbb{R}^n | g_p(x) \geq 0, h_j(x) = a_j^T(x) - b_j = 0, x_i^L \leq x_i \leq x_i^U \} \\ & (p = 1, 2, \dots, P), (j = 1, 2, \dots, J), (J < n), (i = 1, 2, \dots, n). \end{aligned} \quad (2.1)$$

که $f(x)$ ، $g_p(x)$ ، $h_j(x)$ به ترتیب توابع هدف، نامساوی‌ها و تساوی‌های توابع محدودیت با تعداد K ، P و J می‌باشند.

در مثال ۱.۳.۱ انتخاب ماشینی که دارای بالاترین قدرت و پایین‌ترین قیمت باشد، غیر ممکن است لذا به جای بهینه‌سازی، مجبوریم به یک جواب توافقی برسیم یعنی از یک تا چند معیار صرف نظر کنیم تا به آنچه در سایر معیارها می‌خواهیم، دست یابیم.

از نظر تاریخی پارتو^۱ اولین کسی است که به بررسی چنین مسائلی با اهداف و معیارهای متفاوت پرداخت. امروزه به جواب بهینه چنین مسائلی جواب‌های بهینه پارتو می‌گویند.

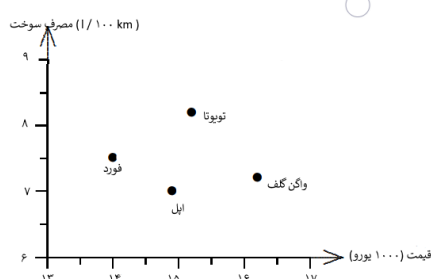
۲.۳.۱ فضای تصمیم‌گیری و فضای توابع هدف (معیار)

مسائل بهینه‌سازی چندهدفه بخاطر اینکه بیشتر از یک هدف را دنبال میکنند کاربردی‌تر از مسائل بهینه‌سازی تک‌هدفه هستند اما در پیدا کردن جواب بهینه با چالش بزرگی دست و پنجه نرم می‌کنند

¹Pareto

که نمی‌توان گفت که برای مسائل بهینه‌سازی چندهدفه روش حل قطعی وجود دارد! برای درک این چالش به مثال زیر توجه کنید: فرض کنید در یک مسئله ماکزیم‌سازی تک‌هدفه، دو نقطه با اندازه‌های «۱» و «۲» پیدا شده‌است، بدیهی است که چون به دنبال بیشترین مقدار هستیم، نقطه ی «۲» را انتخاب می‌کنیم. در این شرایط می‌گوییم نقطه‌ی ۲، نقطه‌ی ۱ را مغلوب کرده‌است. حال مسئله‌ی ماکزیم‌سازی چندهدفه‌ای را فرض کنید که سه نقطه‌ی $(1, 1)$ ، $(1, 2)$ و $(2, 1)$ را پیدا کرده‌است. اکنون سوال اینجاست که کدام نقطه بهتر است؟ نقاط $(1, 2)$ و $(2, 1)$ نقطه‌ی $(1, 1)$ را مغلوب می‌کنند اما در مورد خود دو نقطه ی $(1, 2)$ و $(2, 1)$ چیزی نمی‌توان گفت پس این نقاط نامغلوب هستند.

فرض کنید در مثال ۱.۳.۱ تنها دو هدف قیمت و مصرف سوخت داشته باشیم. فضای هدف (معیار) این مسئله به فرم زیر خواهد بود:



شکل ۲.۱: فضای هدف (معیار) در مثال ۱.۳.۱

از شکل ۲.۱ به راحتی می‌توان دریافت که اوپل و فورد کارآمدتر هستند. مجموعه $X = \{\text{فورد، تویوتا، واگن گلف}\}$ مجموعه شدنی^۱ مسئله ۱.۳.۱ و فضایی که مجموعه شدنی X یک زیر مجموعه از آن است، فضای تصمیم نامیده می‌شود. در واقع مسئله فوق به صورت دو هدف زیر نوشته می‌شود:

$$\text{Min}_{x \geq 0} (f_1(x), f_2(x)) \quad (3.1)$$

تعریف ۱.۳.۱. تصویر X تحت $F = (f_1, f_2)$ را با

$$Y = F(X) = \{y \in \mathbb{R}^2 | y = (f_1(x), f_2(x)) : x \in X\}$$

نشان می‌دهند و به آن تصویر فضای شدنی^۲ یا مجموعه شدنی در فضای تصویر (معیار)^۳ می‌گویند.

مثال ۴.۳.۱. فضای تصویر مثال ۳.۳.۱ به صورت زیر به دست می‌آید.

¹Feasible set

²image of the feasible set

³the feasible set in criterion space

مسئله ۳.۳.۱ به شکل زیر می‌باشند.

$$\text{Min}(f_1(x), f_2(x)) = \text{Min}(\sqrt{x+1}, (x-2)^2 + 1) \quad (4.1)$$

که در شکل ۱.۱ توابع f_1 و f_2 نشان داده شده‌است. فضای شدنی عبارتست از:

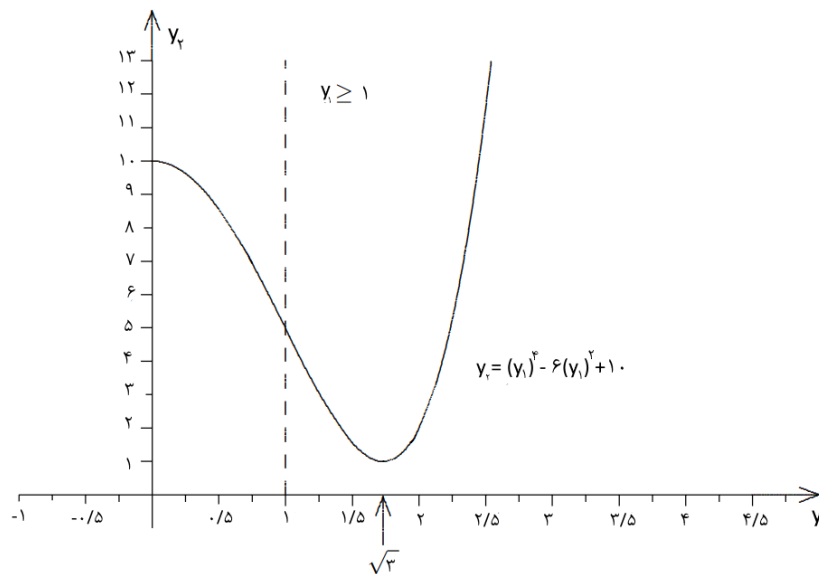
$$X = \{x \in \mathbb{R} : x \geq 0\} \quad (5.1)$$

و فضای تصمیم‌گیری $\mathbb{R}$ است زیرا $X \subset \mathbb{R}$ و فضای معیار $\mathbb{R}^2$ است، زیرا $f(X) \subset \mathbb{R}^2$. برای به دست آوردن تصویر مجموعه شدنی در فضای معیار کافیت قرار دهید $y_1 = f_1(x)$ و $y_2 = f_2(x)$ و سپس x را بر حسب y_1 به دست آورید و در $f_2(x)$ جایگذاری کنید.

$$y_1 = \sqrt{x+1} \rightarrow x = y_1^2 - 1$$

$$y_2 = (x-2)^2 + 1 \rightarrow y_2(x) = ((y_1)^2 - 1)^2 - 4(y_1)^2 + 4 + 1 = (y_1)^4 - 6(y_1)^2 + 10$$

در شکل ۳.۱ نمودار این تابع نشان داده شده است. توجه کنید که $x \geq 0$ به $y_1 \geq 1$ تبدیل شده است، بنابراین $Y = f(X)$ بخشی از نمودار است که در سمت راست خط عمودی $y_1 = 1$ قرار دارد.



شکل ۳.۱: فضای معیار مثال ۳.۳.۱

با محاسبه مینیمم تابع y_2 بر حسب تابع y_1 مشاهده می‌شود که $y_1 = f_1(x) \in [1, \sqrt{3}]$ و $y_2 = f_2(x) \in [1, 5]$ نقاط بهینه نمودار هستند، که متناظر با $x \in [0, 2]$ می‌شود. این نقاط y_1 و (y_2) نقاط غیرمغلوب (غیرتسلطی)^۱ نامیده می‌شوند.

¹nondominated

نمادگذاری ۱. اگر $(s^1, s^2) \in R$ آن گاه به جای $s^1 R s^2$ یا $(s^1, s^2) \in R$ از نماد $s^1 < s^2$ استفاده می‌کنیم و می‌گوییم s^1 بر s^2 ترجیح داده می‌شود.

$$s^1 < s^2 \iff s^1 \preceq s^2 \quad \text{و} \quad s^2 \not\prec s^1$$

$$s^1 \sim s^2 \iff s^1 \preceq s^2 \quad \text{و} \quad s^2 \prec s^1.$$

تعریف ۲.۳.۱ (بهینگی پارتو). مسئله بهینه‌سازی چندهدفه

$$\text{Min}_{x \in X} (f_1(x), f_2(x), \dots, f_K(x))$$

را در نظر بگیرید. نقطه $\hat{x} \in X$ را بهینه پارتو گویند هرگاه هیچ $x \in X$ موجود نباشد، به طوری که $f(x) \leq f(\hat{x})$. یا به عبارتی دیگر $\hat{x}$ بهینه پارتو (کارا) است هرگاه هیچ $x \in X$ دیگری وجود نداشته باشد که:

$$f_i(x) \leq f_i(\hat{x}), \quad \forall i = 1, 2, \dots, K$$

و

$$f_j(x) \leq f_j(\hat{x}), \quad j \in \{1, 2, \dots, K\} \text{ برای حداقل یک}$$

اگر $\hat{x}$ یک نقطه بهینه پارتو (کارا) باشد، به $f(\hat{x})$ یک نقطه غیرمغلوب (غیر تسلطی) گویند. مجموعه تمام نقاط بهینه پارتو را با X_E نشان می‌دهند و تصویر آن‌ها در فضای تصویر را با Y_N نمایش می‌دهند و به آن مجموعه نقاط غیرمغلوب می‌گویند.

تعریف ۳.۳.۱. اگر $x^1, x^2 \in X$ و $f(x^1) \leq f(x^2)$ باشد، گویند x^1 غالب بر x^2 است.

تعریف ۴.۳.۱ (مجموعه نقاط غیرمغلوب). تصویر مجموعه نقاط کارا تحت نگاشت توابع هدف را مجموعه نقاط غیرمغلوب (غیرتسلطی) می‌نامند.

تعریف ۵.۳.۱. (نقاط کارا ضعیف)^۱ مسئله بهینه‌سازی چندهدفه

$$\text{Min}_{x \in X} (f_1(x), f_2(x), \dots, f_K(x))$$

را در نظر بگیرید. نقطه $\hat{x} \in X$ را یک نقطه کارای ضعیف گویند هرگاه هیچ $x \in X$ دیگری وجود نداشته باشد که $f(x) < f(\hat{x})$. مجموعه نقاط کارا ضعیف را با X_{WE} نشان می‌دهند. اگر $\hat{x}$ یک نقطه کارا ضعیف باشد آنگاه $\hat{y} = f(\hat{x})$ را یک نقطه غیر مغلوب ضعیف گویند. مجموعه تمام نقاط غیرمغلوب ضعیف را با Y_{WE} نشان می‌دهند.

¹Weakly efficient

تعریف ۶.۳.۱. (نقاط کارای اکید)^۱ مسئله بهینه‌سازی چندهدفه

$$\text{Min}_{x \in X} (f_1(x), f_2(x), \dots, f_K(x))$$

را در نظر بگیرید. نقطه $\hat{x} \in X$ را یک نقطه کارای اکید گویند هر گاه هیچ $x \in X$ و $x \neq \hat{x}$ وجود نداشته باشد به طوری که $f(x) \leq f(\hat{x})$. مجموعه نقاط کارا اکید را با X_{SE} نشان می‌دهند.

۳.۳.۱ روش مجموع وزنی

مسئله چندهدفه $\text{Min}_{x \in X} (f_1(x), f_2(x), \dots, f_K(x))$ را در نظر بگیرید. یکی از مهمترین و متداول‌ترین روش‌ها برای حل این مسئله استفاده از تکنیک‌های اسکالرسازی^۲ می‌باشد. در واقع یک مسئله اسکالرسازی، یک مسئله تک هدفه متناظر با مسئله چندهدفه است. بعد از این که مسئله تک هدفه را به دست آوردیم رابطه بین جواب‌های بهینه مسئله تک هدفه و جواب‌های کارا مسئله چند هدفه را بررسی می‌کنیم. روش‌های اسکالرسازی زیادی برای حل مسائل بهینه‌سازی چندهدفه وجود دارد. در این بخش روش مجموع وزنی را برای حل این مسائل توضیح می‌دهیم. یکی از معروف‌ترین و پرکاربردترین روش‌های اسکالرسازی روش مجموع وزن دار است. در این روش به هر یک از توابع هدف $f_i(x)$ ، $i = 1, 2, \dots, K$ یک وزن w_i اختصاص داده می‌شود، که در آن $\sum_{i=1}^K w_i = 1$ ، $w_i \geq 0$ در واقع مسئله اسکالرسازی مجموع وزن دار به فرم زیر می‌باشد:

$$\mathbf{P}(w): \min F(\mathbf{x}) = \sum_{i=1}^K w_i f_i(\mathbf{x}) = w_1 f_1(\mathbf{x}) + w_2 f_2(\mathbf{x}) + \dots + w_K f_K(\mathbf{x}) \quad (6.1)$$

subject to $\mathbf{x} \in \Omega$,

که در آن

$$w \in W = \left\{ w \in \mathbb{R}^P \mid \sum_{i=1}^K w_i = 1, w_i \geq 0 \right\} \quad (7.1)$$

به منظور رسیدن به چندین نقطه در مرز بهینه پارتو، مسائل بهینه‌سازی چندهدفه با بردارهای مختلف وزن (w) اجرا می‌شوند. این روش ساده‌ترین و مستقیم‌ترین راه برای به دست آوردن چندین نقطه در مرز بهینه پارتو است.

گزاره ۳. اگر $\hat{x}$ یک جواب کارای مسئله مجموع وزن دار با $w \in \mathbb{R}_{\geq}^P$ باشد آنگاه گزاره‌های زیر برقرارند:

۱. اگر $w \in \mathbb{R}_{\geq}^P$ آنگاه $\hat{x} \in X_{WE}$

¹Strictly efficient

²Scalarisation methods

۲. اگر $w \in \mathbb{R}_{\geq}^P$ آنگاه $\hat{x} \in X_E$.

۳. اگر $w \in \mathbb{R}_{\geq}^P$ و $\hat{x}$ جواب بهینه منحصر به فرد مسئله مجموع وزن دار باشد آنگاه $\hat{x} \in X_{SE}$.

تعریف ۷.۳.۱. نقطه $y^I = (y_1^I, \dots, y_p^I)$ که به صورت

$$y_k^I = \min_{x \in X} f_k(x) \quad (۸.۱)$$

تعریف می‌شود را یک نقطه ایده‌آل برای مسئله بهینه‌سازی چندهدفه (۲.۱) می‌نامیم.

تعریف ۸.۳.۱. نقطه ندیر^۱ نقطه $y^N = (y_1^N, \dots, y_p^N)$ که به صورت

$$y_k^N = \max_{x \in X_E} f_k(x)$$

تعریف می‌شود را یک نقطه ندیر برای مسئله بهینه‌سازی چندهدفه (۲.۱) می‌نامیم.

تعریف ۹.۳.۱. مجموعه جواب‌های بهینه پارتو برای مسائل بهینه‌سازی چندهدفه با روش مجموع وزن دار به صورت زیر تعریف می‌کنیم.

$$E(w) = \left\{ x^* \in R^n \mid \sum_{k=1}^K w_k f_k(x^*) = \min_{x \in \Omega} F(x) \right\}. \quad (۹.۱)$$

گزاره ۴. یک نقطه مانند x^* جواب بهینه پارتو (POS)^۲ مسئله بهینه‌سازی چندهدفه است اگر وجود داشته باشد $\sum_{k=1}^K w_k = 1$, $0 \leq w_k \leq 1$ به طوری که x^* یک جواب منحصر به فرد مسئله $P(w)$ باشد یعنی $E(w) = \{x^*\}$.

گزاره ۵. یک نقطه مانند $\bar{x}$ یک جواب بهینه پارتو سره برای مسئله بهینه‌سازی چندهدفه است اگر وجود داشته باشد $\sum_{k=1}^K \bar{w}_k = 1$, $0 < \bar{w}_k < 1$ به طوری که $\bar{x} \in E(\bar{w})$.

۴.۳.۱ روش بهینه‌سازی توافقی

برخی از رایج‌ترین روش‌های اسکالرسازی عمومی برای مسائل تصمیم‌گیری چندهدفه (MCDM)^۳، روش معیارسراسری، برنامه‌ریزی آرمانی، برنامه‌ریزی فازی و روش‌های تعاملی است. همه آنها برای رسیدن به جواب(های) توافقی، به نقاط مرجع اولیه نیاز دارند. جواب توافقی به عنوان جوابی تعریف می‌شود، که تحت یک تابع فاصله نزدیکترین به جواب ایده‌آل است. ما اکنون به دنبال رسیدن به جواب توافقی مسائل بهینه‌سازی در معادله (۲.۱) هستیم. بردار توابع هدف $f(x) = (f_1(x), f_2(x), \dots, f_K(x))$ را در نظر بگیرید. فرض کنید که

^۱Nadir point

^۲Pareto Optimal Solution

^۳Multiple Criteria Decision Making

$F^+ = (f_1^+, f_2^+, \dots, f_K^+)$ جواب ایده‌آل یا جواب ایده‌آل مثبت (PIS)^۱ بردار توابع هدف باشد به طوری که $f_k^+ = \text{Min}_{x \in \Omega} f_k(x)$, $k = 1, 2, \dots, K$

هم‌چنین فرض کنید $F^- = (f_1^-, f_2^-, \dots, f_K^-)$ جواب ضدایده‌آل یا جواب ایده‌آل منفی (NIS)^۲ بردار توابع هدف باشد به طوری که $f_k^- = \text{Max}_{x \in \Omega} f_k(x)$, $k = 1, 2, \dots, K$

اکنون از معیار L_P برای رسیدن به معیار «نزدیکی» استفاده می‌کنیم. معیار L_P فاصله بین $F(x)$ و F^+ را به صورت زیر تعریف می‌کند:

$$d_p(x) = \left(\sum_{k=1}^K w_k^p (f_k(x) - f_k^+)^p \right)^{\frac{1}{p}} \quad p = 1, 2, \dots, \infty \quad (10.1)$$

که در آن w_k نشان‌دهنده اهمیت نسبی تابع هدف k ام است. $\forall k = 1, 2, \dots, K$ و اگر توابع هدف $f_k(x)$, $k = 1, 2, \dots, K$ در واحدهای ثابت و معادل بیان نشوند، آنگاه معمولاً برای هر یک از توابع هدف از یک تابع مقیاس‌گذاری، در بازه‌ی $[0, 1]$ استفاده می‌شود. در این حالت، از معیار زیر می‌توان استفاده کرد:

$$d_p(x) = \left(\sum_{k=1}^K w_k^p \left(\frac{f_k(x) - f_k^+}{f_k^- - f_k^+} \right)^p \right)^{\frac{1}{p}} \quad p = 1, 2, \dots, \infty \quad (11.1)$$

روش‌های مختلف توابع چندهدفه مانند روش معیار سراسری، روش برنامه‌ریزی هدف، روش برنامه‌ریزی فازی و روش تعاملی، از خانواده فاصله (۱۰.۱) و (۱۱.۱) برای رسیدن به جواب توافقی یک مسئله تصمیم‌گیری چندمعیاری استفاده می‌کنند، که جواب ایده‌آل $F^+ = (f_1^+, f_2^+, \dots, f_K^+)$ همان نقطه مرجع است. مطابق با [۳۸] مسئله بهینه‌سازی چندهدفه به مسئله کمکی زیر کاهش می‌یابد.

$$P_1 : \text{Min} \quad d_p(x) = \left(\sum_{k=1}^K w_k^p \left(\frac{f_k(x) - f_k^+}{f_k^- - f_k^+} \right)^p \right)^{\frac{1}{p}} \quad p = 1, 2, \dots, \infty \quad (12.1)$$

که در آن $x \in \Omega$

مقدار انتخاب‌شده برای p ، روش رسیدن به جواب (های) توافقی را، با به حداقل رساندن مجموع وزنی توابع هدف از نقطه مرجع مربوطه‌ی آنها (نقطه ایده‌آل) تعیین می‌کند. پارامتر p نقش «عامل متعادل‌کننده» بین کارایی گروه و حداکثر بهینگی هر یک از توابع را ایفا می‌کند. با افزایش p ، کارایی گروه (فاصله d_p) کاهش می‌یابد. یعنی $d_1 \geq d_2 \geq \dots \geq d_P$ که تأکید اصلی به بیشترین انحراف در شکل کلی داده می‌شود. به طور خاص، $p = 1$ نشان می‌دهد همه این انحرافات از اهمیتی معادل (وزن) برخوردار هستند، در حالی که $p = 2$ نشان می‌دهد وزن با انحرافات متناسب است یعنی بیشترین انحراف دارای بیشترین وزن است. سرانجام برای $p = \infty$ بزرگترین انحراف کاملاً بر

¹Positive Ideal Solution

²Anti-ideal Solution or Negative Ideal Solution

ضریب فاصله غلبه دارد، معیار l_∞ به شکل است:

$$d_\infty = \text{Max}_k \{w_k [f_k(\mathbf{x}) - f_k^+]\} \quad \text{یا} \quad d_\infty = \text{Max}_k \left\{ w_k \left[\frac{f_k(\mathbf{x}) - f_k^+}{f_k^- - f_k^+} \right] \right\} \quad (13.1)$$

۴.۱ مسائل بهینه‌سازی چندسطحی

مسئله دوسطحی خطی زیر را در نظر بگیرید [۶، ۱، ۲۵]: برای $\mathbf{x} \in X \subset \mathbb{R}^n$ ، $\mathbf{y} \in Y \subset \mathbb{R}^m$ ، $F: X \times Y \rightarrow \mathbb{R}$ و $f: X \times Y \rightarrow \mathbb{R}$ یک مسئله دوسطحی خطی به صورت زیر است:

$$\begin{aligned} P_2 : \quad & \text{Min}_{\mathbf{x}} \quad F(\mathbf{x}, \mathbf{y}) = c_1 \mathbf{x} + d_1 \mathbf{y} \\ & A_1 \mathbf{x} + B_1 \mathbf{y} \leq b_1 \end{aligned}$$

به طوری که $\mathbf{y}$ جوابی از

$$\begin{aligned} P_1 : \quad & \text{Min}_{\mathbf{y}} \quad f(\mathbf{x}, \mathbf{y}) = c_2 \mathbf{x} + d_2 \mathbf{y} \\ & A_2 \mathbf{x} + B_2 \mathbf{y} \leq b_2 \\ & \mathbf{y} \geq \mathbf{0} \end{aligned} \quad (14.1)$$

است. به طوری که در آن $c_1, c_2 \in \mathbb{R}^n$ ، $d_1, d_2 \in \mathbb{R}^m$ ، $b_1 \in \mathbb{R}^p$ ، $b_2 \in \mathbb{R}^q$ ، $A_1 \in \mathbb{R}^{p \times n}$ ، $A_2 \in \mathbb{R}^{q \times n}$ ، $B_1 \in \mathbb{R}^{p \times m}$ و $B_2 \in \mathbb{R}^{q \times m}$.

روش‌های مختلفی برای حل مسائل بهینه‌سازی چندسطحی (MLPPs)^۱ پیشنهاد شده است [۶]. یکی از روش‌ها شرایط کروش-کان-تاکر^۲ است که مسئله را به مسئله یک‌سطحی کاهش می‌دهد.

۱.۴.۱ روش کروش-کان-تاکر برای حل مسائل برنامه‌ریزی دوسطحی خطی

تعریف ۱.۴.۱. ایده اصلی روش کروش-کان-تاکر برای مسائل برنامه‌ریزی دو سطحی خطی بدین صورت است که مسئله سطح پایین را به مسئله سطح بالا اضافه می‌کنیم. فرض کنید $u \in \mathbb{R}^q$ و $v \in \mathbb{R}^m$ متغیرهای دوگان متناظر با قیدهای سطح پایین و $y \geq 0$ باشند. [۲۸]

تعریف ۲.۴.۱:

۱. ناحیه قیود از مسئله دو سطحی:

$$N = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \in X, \mathbf{y} \in Y, A_1 \mathbf{x} + B_1 \mathbf{y} \leq b_1, A_2 \mathbf{x} + B_2 \mathbf{y} \leq b_2\}. \quad (15.1)$$

¹Multi-level Linear Programming Problems

²Karush-Kuhn-Tucker

۲. مجموعه شدنی پیرو برای هر ثابت $x \in X$:

$$N(x) = \{y \in Y | B_2 y \leq b_2 - A_2 x\}. \quad (16.1)$$

۳. تصویر N به توی فضای تصمیم پیشرو:

$$N(X) = \{x \in X | \exists y \in Y, A_1 x + B_1 y \leq b_1, A_2 x + B_2 y \leq b_2\}. \quad (17.1)$$

۴. مجموعه واکنش منطقی پیرو^۱ برای $x \in N(X)$:

$$\psi(x) = \{y \in Y | y \in \operatorname{argmin}[f(x, \hat{y}) | \hat{y} \in N(x)]\}. \quad (18.1)$$

به قسمی که

$$\operatorname{argmin}[f(x, \hat{y}) | \hat{y} \in N(x)] = \{y \in N(x) | f(x, y) \leq f(x, \hat{y}), \text{ for all } \hat{y} \in N(x)\} \quad (19.1)$$

۵. ناحیه القایی^۲ (ناحیه شدنی):

$$IR = \{(x, y) \in S, y \in \psi(x)\}. \quad (20.1)$$

تعریف بالا و فرضیه‌های زیر تضمین می‌کند که مسئله (۱۴.۱) دارای جواب است.

فرضیات ۱ :

۱. N ناتهی و فشرده باشد.

۲. برای تصمیماتی که پیشرو می‌گیرد پیرو واکنش نشان دهد. یعنی $\psi(x) \neq \emptyset$

۳. $\psi(x)$ یک نگاشت نقطه به نقطه^۳ باشد.

مجموعه واکنش منطقی $\psi(x)$ ، تعریف پاسخ‌های پیرو است. در حالیکه ناحیه القایی IR مجموعه‌ای را بیان می‌کند که تابع پیشرو هدف روی آن بهینه می‌شود. بنابراین با توجه به نمادهای بالا، مسئله دوسطحی خطی را می‌توان به صورت زیر نوشت:

$$\operatorname{Min}\{F(x, y) | (x, y) \in IR\} \quad (21.1)$$

اگر فرض (۳) برقرار نباشد مسئله (۲۱.۱) ممکن است جواب نداشته باشد. به ویژه اگر $\psi(x)$ برای هر $x \in X$ تک مقداری نباشد، تابع پیشرو ممکن است روی ناحیه القایی IR مینیم خود را بدست نیاورید. در ادامه فرض می‌کنیم فرضیات بالا برقرار باشند.

¹Followers rational reaction

²Inducible region

³Point-to-point

تعریف ۳.۴.۱. (x^*, y^*) یک جواب بهینه کامل مسئله دو سطحی است اگر $(x^*, y^*) \in N$ و $F(x^*, y^*) \leq F(x, y)$ و $f(x^*, y^*) \leq f(x, y)$ برای هر $(x, y) \in N$.

هر چند معمولاً چنین جواب بهینه کاملی که به طور همزمان هر دو تابع هدف پیشرو و پیرو را بهینه کند، وجود ندارد. اما با استفاده از این مفهوم، بهینگی پارتو را برای مسئله دوسطحی خطی بیان کنیم.

تعریف ۴.۴.۱. (x^*, y^*) یک جواب بهینه پارتو مسئله دو سطحی است اگر و فقط اگر وجود نداشته باشد $(x, y) \in N$ به طوری که $F(x, y) \leq F(x^*, y^*)$ ، $f(x, y) \leq f(x^*, y^*)$ و $F(x, y) \neq F(x^*, y^*)$ یا $f(x, y) \neq f(x^*, y^*)$. [۱۱]

مسئله برنامه‌ریزی خطی زیر را در نظر بگیرید.

$$\begin{aligned} \text{Min} \quad & f(x) = cx \\ \text{subject to} \quad & Ax \geq b, \\ & x \geq 0 \end{aligned} \quad (22.1)$$

که c یک بردار سطری n بعدی، b یک بردار سطری m بعدی، A یک ماتریس $m \times n$ با $m \leq n$ و $x \in \mathbb{R}^n$ است.

فرض کنید $\lambda \in \mathbb{R}^m$ و $\mu \in \mathbb{R}^n$ به ترتیب متغیر دوگان وابسته به محدودیت‌های $Ax \geq b$ و $x \geq 0$ باشند. گزاره زیر برقرار است.

گزاره ۶. شرط لازم و کافی برای اینکه (x^*) جواب مسئله خطی ۲۲.۱ باشد این است که بردارهای سطری λ^* و μ^* وجود داشته باشند به طوری که (x^*, λ^*, μ^*) جواب سیستم زیر باشد [۱۱].

$$\begin{aligned} \lambda A - \mu &= -c, \\ Ax - b &\geq 0, \\ \lambda(Ax - b) &= 0, \\ \mu x &= 0, \\ x \geq 0, \lambda &\geq 0, \mu \geq 0. \end{aligned} \quad (23.1)$$

لم ۱.۴.۱. یک شرط لازم برای اینکه (x^*, y^*) جواب مسئله برنامه‌ریزی دوسطحی (۱۴.۱) باشد این است که بردارهای سطری u^* و v^* وجود داشته باشند به طوری که (x^*, y^*, u^*, v^*) جواب مسئله زیر

باشد [۱۱].

$$\begin{aligned}
 & \text{Min } c_1x + d_2y \\
 & \text{subject to } A_1x + B_1y \leq b_1 \\
 & \text{subject to } uB_2 - v = -d_2 \\
 & \text{subject to } u(b_2 - A_2x - B_2y) + vy = 0 \\
 & \text{subject to } A_2x + B_2y \leq b_2 \\
 & x \leq 0, y \leq 0, u \leq 0, v \leq 0.
 \end{aligned} \tag{۲۴.۱}$$

قضیه ۱۰۴.۱. شرط لازم و کافی برای اینکه (x^*, y^*) جواب مسئله برنامه‌ریزی دوسطحی (۱۴.۱) باشد این است که بردارهای سطری u^* ، v^* و w^* وجود داشته باشند به طوری که $(x^*, y^*, u^*, v^*, w^*)$ جواب مسئله زیر باشد [۲۸].

$$\begin{aligned}
 & \text{Min } F(x, y) = c_1x + d_1y \\
 & \text{subject to } A_1x + B_1y \leq b_1, \\
 & A_2x + B_2y \leq b_2, \\
 & uB_1 - vB_2 - w = -d_2, \\
 & u(b_1 - A_1x - B_1y) + v(b_2 - A_2x - B_2y) + wy = 0, \\
 & A_2x + B_2y \leq b_2 \\
 & x \leq 0, y \leq 0, u \leq 0, v \leq 0, w \leq 0.
 \end{aligned} \tag{۲۵.۱}$$

برهان. مسئله دوسطحی (۲۱.۱) در نظر بگیرید. این مسئله را به صورت زیر بازنویسی می‌کنیم:

$$\begin{aligned}
 & \text{Min } F(x, y) \\
 & \text{subject to } (x, y) \in IR.
 \end{aligned}$$

با توجه به تعریف ناحیه القایی ۲۰.۱ داریم

$$\begin{aligned}
 & \text{Min } F(x, y) \\
 & \text{subject to } (x, y) \in N, \\
 & y \in \psi(x).
 \end{aligned}$$

با توجه به تعریف مجموعه واکنش منطقی پیرو ۱۹.۱ داریم

$$\begin{aligned}
 & \text{Min } F(x, y) \\
 & \text{subject to } (x, y) \in N, \\
 & y \in \operatorname{argmin}[f(x, \hat{y}) : \hat{y} \in N(x)].
 \end{aligned}$$

یعنی داریم

$$\begin{aligned} \text{Min} \quad & F(x, y) \\ \text{subject to} \quad & (x, y) \in N, \\ & \text{Min} \quad f(x, y), \\ & \text{subject to } y \in N(x). \end{aligned}$$

با توجه به تعریف مجموعه شدنی پیرو ۱۶.۱ داریم

$$\begin{aligned} \text{Min} \quad & F(x, y) \\ \text{subject to} \quad & (x, y) \in N, \\ & \text{Min}_{y \in Y} \quad f(x, y), \\ & \text{subject to } (x, y) \in N(x). \end{aligned}$$

بنابراین، با توجه به تعریف ناحیه قیود مسئله دوسطحی ۱۵.۱ داریم

$$\begin{aligned} \text{Min}_{x \in X} \quad & F(x, y) = c_1 x + d_1 y \\ \text{subject to} \quad & A_1 x + B_1 y \leq b_1, \\ & A_2 x + B_2 y \leq b_2, \\ & \text{Min}_{y \in Y} \quad f(x, y) = c_2 x + d_2 y, \\ & \text{subject to } A_1 x + B_1 y \leq b_1, \\ & A_2 x + B_2 y \leq b_2, \end{aligned} \quad (26.1)$$

این تبدیل ساده نشان می‌دهد

۱. که جواب مسئله (۱۴.۱) معادل با جواب (۲۶.۱) است.

۲. شرط لازم قضیه با توجه به شکل مسئله (۲۶.۱) واضح است.

۳. و برای شرط کافی داریم:

اگر (x^*, y^*) جواب بهینه مسئله (۱۴.۱) باشد، باید نشان دهیم که بردارهایی مانند u^*, v^* و w^* وجود دارند به طوری که $(x^*, y^*, u^*, v^*, w^*)$ جواب بهینه مسئله (۲۵.۱) باشند. برای اولین مرحله باید ثابت کنیم که بردارهایی مانند u^*, v^* و w^* وجود دارند به طوری که $(x^*, y^*, u^*, v^*, w^*)$ در شرایط زیر صدق کند.

$$\begin{cases} uB_1 + vB_2 - w = -d_2 \\ u(b_1 - A_1x - B_1y) = 0 \\ v(b_2 - A_2x - B_2y) = 0 \\ wy = 0 \end{cases} \quad (27.1)$$

که $w \in \mathbb{R}^m, v \in \mathbb{R}^q, u \in \mathbb{R}^P$ و همگی نامنفی هستند. چون (x^*, y^*) جواب بهینه مسئله (۱۴.۱) و با توجه به ۲۱.۱ داریم:

$$(x^*, y^*) \in IR$$

. بنابراین با توجه به ۲۰.۱ داریم

$$y^* \in \psi(x^*)$$

. با توجه به تعریف مجموعه واکنش منطقی پیرو ۱۸.۱ y^* جواب بهینه مسئله زیر است

$$\min(f(x^*, y) : y \in N(x^*)) \quad (28.1)$$

. مسئله (۲۸.۱) را به صورت زیر می‌نویسیم

$$\begin{aligned} \min \quad & f(x, y) \\ \text{subject to} \quad & y \in N(x), \\ & x = x^*. \end{aligned}$$

. با توجه به تعریف تصویر N به توی فضای تصمیم پیشرو ۱۷.۱ داریم

$$\begin{aligned} \text{Min} \quad & F(x, y) = c_1x + d_2y \\ \text{subject to} \quad & A_1x + B_1y \leq b_1, \\ & A_2x + B_2y \leq b_2, \\ & x = x^*, \\ & y \geq 0. \end{aligned} \quad (29.1)$$

مسئله (۲۹.۱) را می‌توان به شکل زیر مختصر کرد.

$$\begin{aligned} \text{Min} \quad & F(y) = d_2y \\ \text{subject to} \quad & - \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} y \geq - \begin{pmatrix} b_1 - A_1x^* \\ b_2 - A_2x^* \end{pmatrix}, \\ & y \geq 0. \end{aligned} \quad (30.1)$$

حال می‌بینیم که y^* جواب بهینه (۳۰.۱) که یک مسئله بهینه‌سازی خطی است، می‌باشد. با توجه به گزاره ۶ بردارهای λ^* و μ^* وجود دارند به طوری که (y^*, λ^*, μ^*) در سیستم زیر صدق می‌کنند.

$$\begin{aligned} \lambda \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} - \mu &= d_2 y \\ \lambda \left(- \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} y + \begin{pmatrix} b_1 - A_1 x^* \\ b_2 - A_2 x^* \end{pmatrix} \right) &= 0, \\ \mu y &\geq 0. \end{aligned} \quad (31.1)$$

که $\lambda \in \mathbb{R}^{p+q}$ و $\mu \in \mathbb{R}^m$. فرض کنید $w \in \mathbb{R}^m$, $v \in \mathbb{R}^q$, $\mu \in \mathbb{R}^p$ و λ و w را به صورت زیر

$$\lambda = (\mu, v)$$

$$w = \mu$$

تعریف کنیم. بنابراین بردار $(x^*, y^*, \mu^*, v^*, w^*)$ در شرایط ۲۷.۱ صدق می‌کند و اثبات کامل می‌شود.

□

مثال ۱.۴.۱. مسئله دوسطحی زیر را با فرض $x \in \mathbb{R}^1$, $y \in \mathbb{R}^1$, $X = \{x \geq 0\}$ و $Y = \{y \geq 0\}$ در نظر بگیرید.

$$\begin{aligned} \min_{x \in X} \quad & F(x, y) = x - 4y \\ \text{subject to} \quad & -x - y \leq -3, \\ & -3x - y \geq -4, \\ \min_{y \in Y} \quad & f(x, y) = x + y \\ \text{subject to} \quad & -2x + y \leq 0, \\ & +2x + y \leq 12. \end{aligned}$$

با توجه به رویکرد کان-تاکر نامساوی‌های محدودیت‌ها را به صورت زیر می‌نویسیم.

$$\begin{aligned} g_1(x, y) &= x + y - 3 \geq 0, \\ g_2(x, y) &= -3x + 2y + 4 \geq 0, \\ g_3(x, y) &= 2x - y \geq 0, \\ g_4(x, y) &= -2x - y + 12 \geq 0, \\ g_5(x, y) &= y \geq 0. \end{aligned} \quad (32.1)$$

با توجه به ۲۵.۱ داریم.

$$\min F(x, y) = x - 4y \quad (33.1)$$

$$\text{subject to } -x - y \leq -3,$$

$$-3x - y \geq -4,$$

$$-2x + y \leq 0,$$

$$+2x + y \leq 12,$$

$$-u_1 - 2u_2 + u_3 + u_4 - u_5 = -1, \quad (34.1)$$

$$u_1 g_1(x, y) + u_2 g_2(x, y) + u_3 g_3(x, y) + u_4 g_4(x, y) + u_5 g_5(x, y) = 0, \quad (35.1)$$

$$x \geq 0, y \geq 0, u_1 \geq 0, u_2 \geq 0, u_3 \geq 0, u_4 \geq 0, u_5 \geq 0. \quad (36.1)$$

با توجه به (۳۴.۱) و (۳۶.۱) سه حالت به شکل زیر داریم.

۱.

$$(u_1^*, u_2^*, u_3^*, u_4^*, u_5^*) = (1, 0, 0, 0, 0).$$

با توجه به (۳۵.۱)، داریم

$$g_1(x, y) = 0$$

. بنابراین دستگاه (۳۳.۱) به صورت زیر تبدیل می‌شود:

$$\min F(x, y) = x - 4y$$

$$\text{subject to } -x - y = -3,$$

$$-3x - y \geq -4,$$

$$-2x + y \leq 0,$$

$$+2x + y \leq 12,$$

$$x \geq 0, y \geq 0.$$

با استفاده از الگوریتم سیمپلکس، درمی‌یابیم که جواب بهینه این سیستم در نقطه $(x^*, y^*) = (1, 2)$ با $F^* = -7$ و $f^* = 3$ رخ می‌دهد.

۲.

$$(u_1^*, u_2^*, u_3^*, u_4^*, u_5^*) = (0, 0.5, 0, 0, 0).$$

با توجه به (۳۵.۱)، داریم

$$g_2(x, y) = 0$$

. بنابراین دستگاه (۳۳.۱) به صورت زیر تبدیل می‌شود:

$$\begin{aligned} \min \quad & F(x, y) = x - 4y \\ \text{subject to} \quad & -x - y \leq -3, \\ & -3x - y = -4, \\ & -2x + y \leq 0, \\ & +2x + y \leq 12, \\ & x \geq 0, y \geq 0. \end{aligned}$$

با استفاده از الگوریتم سیمپلکس، درمی‌یابیم که جواب بهینه این سیستم در نقطه $(x^*, y^*) = (4, 4)$ با $F^* = -12$ و $f^* = 8$ رخ می‌دهد.

۳.

$$(u_1^*, u_2^*, u_3^*, u_4^*, u_5^*) = (0, 0, 0, 0, 1).$$

با توجه به (۳۵.۱)، داریم

$$g_5(x, y) = 0$$

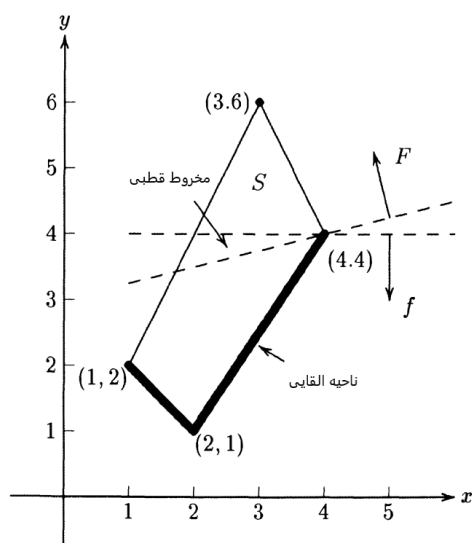
. بنابراین دستگاه (۳۳.۱) به صورت زیر تبدیل می‌شود:

$$\begin{aligned} \min \quad & F(x, y) = x - 4y \\ \text{subject to} \quad & -x - y \leq -3, \\ & -3x - y \geq -4, \\ & -2x + y \leq 0, \\ & +2x + y \leq 12, \\ & x \geq 0, y = 0. \end{aligned}$$

با استفاده از الگوریتم سیمپلکس، درمی‌یابیم که این سیستم حل نشدنی است.

در نقطه $(x^*, y^*) = (4, 4)$ با $F^* = -12$ و $f^* = 8$ رخ می‌دهد.

چندوجهی شکل ۴.۱ ناحیه محدودیت‌های N را برای این مثال نشان می‌دهد.



شکل ۴.۱: نمودار هندسی مثال بهینه‌سازی دوسطحی (۴.۱)

مسئله؟؟ با محدودیت $wy = 0$ غیر خطی است. فورتالی و مک‌کارل^۱ [۱۲] با معرفی بردار صفر و یک، شرایط مکمل زائد را دوباره فرمول بندی و یک مسئله برنامه ریزی عدد صحیح صفر و یک با متغیر دوتایی $\xi \in \{0, 1\}$ ایجاد کردند. بنابراین، محدودیت ضربی $wy = 0$ با دونا مساوی ساده $y \leq M\xi$ و $w \leq (1 - \xi)M$ که M به اندازه کافی بزرگ است، جایگزین می‌شود. اکنون فرمول بندی جدید به صورت زیر بدست می‌آید.

$$\begin{aligned}
 \text{Min} \quad & F(x, y) = c_1x + d_1y \\
 \text{s.t} \quad & A_1x + B_1y \leq b_1, \\
 \text{s.t} \quad & A_2x + B_2y \leq b_2, \\
 & uB_1 - vB_2 - w = -d_2, \\
 & w \leq (1 - \xi)M \\
 & y \leq M\xi, \\
 & x \leq 0, y \leq 0, u \leq 0, v \leq 0, w \leq 0, \xi \in \{0, 1\}.
 \end{aligned} \tag{۳۷.۱}$$

¹Fortuny and McCarl

فصل ۲

شبکه‌های عصبی

در این فصل ابتدا ساختار شبکه‌های عصبی مغز انسان را معرفی و عملکرد سیستمی مغز را به صورت کلی بررسی می‌کنیم. سپس با الگو برداری از رفتار سیستمی مغز، درک می‌کنیم که می‌توانیم برخی از محاسبات را برحسب عملکرد مغز انجام دهیم که به آن هوش مصنوعی می‌گویند. یکی از روش‌های کارا برای حل مسائل پیچیده، تبدیل مسئله به زیر مسئله‌های ساده‌تر است. در حقیقت یک شبکه مجموعه‌ای از این مسئله‌های ساده است که در کنار همدیگر، سیستم نهایی را توصیف می‌کنند. در این فصل ابتدا نرون‌های مغز و طرز کار آنها نشان داده شده‌است. سپس مدل مصنوعی این نرون‌ها و ساختار آنها، مدل ریاضی آنها، شبکه‌های عصبی مصنوعی و نحوه آموزش و بکارگیری این شبکه‌ها نشان داده شده‌است. در این پایان‌نامه تمرکز بیشتر بر نوعی از این شبکه‌ها به نام شبکه‌های عصبی مصنوعی بازگشتی است. اغلب مطالب این فصل از [۱۷] گرفته شده است.

۱.۲ شبکه‌های عصبی طبیعی

مغز انسان در حالیکه درصد کمی از وزن بدن را دارد اما بیست درصد انرژی بدن را مصرف می‌کند. ساده‌ترین و اصلی‌ترین اجزای تشکیل دهنده شبکه عصبی مغز نرون‌ها^۱ یا عصب‌ها هستند. این سلول‌ها با یکدیگر در ارتباط بوده و اطلاعات را با سرعتی باورنکردنی در بدن انتقال می‌دهند. شکل و ساختار، سرعت انتقال پیام، عملکرد و... سبب شده تا این سلول‌های کوچک، پیشرفته‌ترین ماشین

¹Neuron

بیولوژیک جهان یعنی مغز را بسازند و خودشان در رده‌ی متمایزترین سلول‌ها قرار بگیرند. هر نرون مغز از صد تریلیون (۱۰^{۱۱}) نرون به هم مرتبط با تعداد ۱۰^{۱۶} پیوند تشکیل شده‌است و این ساختار، امکان پردازش موازی اطلاعات را موجب می‌شود. هر نرون از سه قسمت اصلی تشکیل شده‌است.

۱. جسم سلولی^۱ که هسته و اندامک‌ها را در خود جای داده است و به نوعی مرکز فرماندهی یاخته‌ی عصبی می‌باشد.

۲. دندریت‌ها^۲ که زائده‌هایی هستند که از جسم سلولی منشا گرفته و معمولاً به صورت منشعب و چندتایی هستند.

۳. آکسون^۳ که یک رشته بلند که از سمت دیگر یاخته‌ی عصبی بیرون زده است و به سلول بعدی (نرون - ماهیچه - غده) می‌رسد.

در شکل ۱.۲ ساختار یک نرون را مشاهده می‌کنید.



شکل ۱.۲: ساختار نرون

پیام الکتریکی پس از آنکه به جسم سلولی رسید، در آنجا پردازش شده و سپس به آکسون‌ها فرستاده می‌شود. پیام الکتریکی در درازای آکسون هدایت شده و به پایان آکسون که پایانه آکسونی نام دارد، می‌رسد. پایانه آکسونی یا با دندریت نرون دیگر ارتباط برقرار کرده یا به یک سلول ماهیچه‌ای یا یاخته غده‌ای می‌رسد. در پایانه‌های آکسونی ریزساختار ویژه‌ای به نام سیناپس^۴ وجود دارد که به آن فضای سیناپسی می‌گویند. از راه سیناپس پیام سلول عصبی به سلول بعدی راه می‌یابد. همه سیگنال‌ها در تنه نرون ترکیب شده و هنگامی که وسعت سیگنال‌های به آستانه نرون برسد، مرحله تحریک شدن^۵ اتفاق می‌افتد و یک سیگنال خروجی تولید می‌شود. این سیگنال به صورت یک پالس منفرد به پایانه‌های سیناپسی انتقال و باعث ترشح موادی به نام عصب-رسانه^۶ می‌شود. این مواد

¹Soma

²Dendrite

³Axon

⁴synaps

⁵Firing

⁶Neuro-transmitter

در داخل شکاف سیناپسی پخش شده و موجب تحریک نرون بعدی می‌شوند و به این ترتیب یک سیگنال از یک نرون به نرون دیگر انتقال می‌یابد. محققان ابتدا فکر می‌کردند که یک پیام عصبی ماهیت الکتریکی دارد اما بعدها مشخص شد که این پیام‌ها ماهیت الکتروشیمیایی دارند. در شبکه عصبی مغز انسان، سلول‌ها قادر به یادگیری اند. هنگامی که سلول‌های عصبی لامسه دچار سوزش می‌شوند، یاد می‌گیرند که به طرف جسم داغ نروند. در واقع با این الگوریتم سیستم می‌آموزد که خطای خود را اصلاح کند. یادگیری در این سیستم‌ها با مطابقت دادن صورت می‌گیرد، یعنی با استفاده از نمونه‌های قبلی، وزن سیناپس‌ها طوری تغییر می‌کند که در صورت وقوع رخ داده‌های مشابه و جدید سیستم عصبی پاسخ درستی تولید کند. همانطور که سلول‌های سیستم عصبی ما می‌توانند از داده‌های گذشته، تجربه کسب کنند، به همین صورت شبکه عصبی‌های مصنوعی نیز قادر به یادگیری از داده‌ها و ارائه پاسخ در قالب پیش‌بینی‌ها و طبقه‌بندی‌ها هستند.

۲.۲ شبکه های عصبی مصنوعی

شبکه عصبی مصنوعی^۱ ایده‌ای است که مانند مغز به پرداخت اطلاعات می‌پردازد و از سیستم عصبی طبیعی الهام گرفته شده است. این موضوع شاخه‌ای از هوش مصنوعی است که سعی دارد سیستمی مشابه نرون‌های مغز انسان را به صورت مصنوعی بسازد. در واقع سیستم شبکه عصبی مصنوعی متشکل از تعداد زیادی عناصر پردازشی به هم پیوسته است که برای حل یک مسئله با یکدیگر هماهنگ کار می‌کنند، در این شبکه‌ها به کمک علم برنامه‌نویسی، ساختاری طراحی می‌شود که داده‌های آن می‌توانند مانند نرون‌ها عمل کنند. قابلیت یادگیری مهم‌ترین ویژگی شبکه‌های عصبی مصنوعی است. در این معنا که این شبکه بر پایه‌ی یک قاعده کلی و مشخص طراحی نشده و با گذر زمان و یا در هر تکرار شبکه، تغییر وضعیت آن مشاهده می‌شود یا به اصطلاح شبکه آموزش می‌بیند. شبکه‌های عصبی مصنوعی مرتبط با الگوبرداری از ویژگی یادگیری در انسان و اجرای آن به صورت الگوریتم‌های کامپیوتری است. در نتیجه شبکه‌های عصبی این توانایی را دارند که آن چه برای سیستم‌های بر مبنای قاعده ثابت سخت است را، بیاموزند. آغاز تحقیقات در این زمینه به دهه ۱۹۴۰ میلادی بر می‌گردد اما گسترش آن در دهه ۱۹۸۰ رخ داد. تولد دوباره شبکه‌های عصبی مصنوعی در سال‌های ۱۹۸۲ تا ۱۹۸۶ اتفاق افتاد. دو اتفاق مهم که باعث گشایش افق جدیدی در زمینه شبکه‌های عصبی مصنوعی شد عبارتند از ارائه شبکه عصبی بازگشتی و مفهوم تابع انرژی توسط جان هاپفیلد^۲ و حل مسئله فروشنده دوره گرد^۳ توسط شبکه هاپفیلد. او نشان داد که از این شبکه بازگشتی می‌توان برای حل مسائل بهینه‌سازی استفاده کرد.

^۱Artificial neural network

^۲John Hopfield

^۳Travelling salesman problem

۳.۲ توپولوژی شبکه‌های عصبی مصنوعی

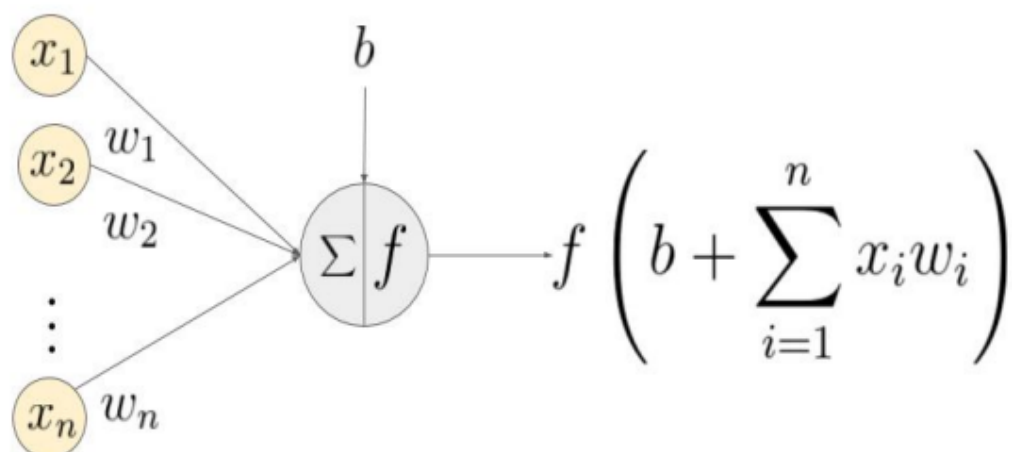
وضعیت نسبی نرون‌ها در شبکه (تعداد و گروه‌بندی و نوع اتصالات آن‌ها) را توپولوژی شبکه گویند. توپولوژی در واقع سیستم اتصال سخت‌افزار نرون‌ها به همدیگر است که همراه با نرم‌افزار مربوطه نوع عملکرد شبکه‌عصبی را تعیین می‌کند. این توپولوژی با قرار دادن یک لایه‌ی ورودی که اطلاعات را دریافت می‌کند، تعدادی لایه‌ی مخفی که اطلاعات را از لایه‌های قبلی می‌گیرند و در نهایت یک لایه‌ی خروجی که نتیجه محاسبات به آنجا می‌رود، کامل می‌شود. در یک نگاه ساده مدل یک نرون باید شامل n ورودی x_i ($i = 1, 2, \dots, n$) باشد. این ورودی‌ها در وزن‌های سیناپس w_i ($i = 1, 2, \dots, n$) که روی آکسون‌ها واقع هستند، ضرب می‌شوند. حاصل ضرب ورودی‌ها در وزن‌های متناظرشان با عددی مانند b که بایاس^۱ نامیده می‌شود، جمع می‌شوند. این حاصل جمع را خروجی خطی نرون می‌نامیم و با z نشان می‌دهیم. نهایتاً تابعی به نام تابع فعال‌ساز^۲ به z اثر می‌کند و به این ترتیب خروجی نرون حاصل می‌شود. معادلات ریاضی چنین نرونی به صورت زیر است.

$$f(z) = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

در نتیجه شبکه عصبی مصنوعی با استفاده از مدل ساده‌شده نرون به پردازش و بررسی اطلاعات می‌پردازد. بنابراین برای توصیف یک نرون (یک گره در شبکه عصبی مصنوعی) می‌توان مدل ساده‌ای پیشنهاد کرد. این مدل در شکل ۲.۲ نشان داده شده است.

^۱Bias

^۲Activation function



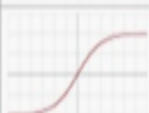
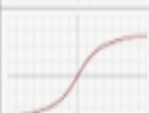
شکل ۲.۲: مدل ریاضی نرون عصبی مصنوعی

شاخصه‌های اصلی یک شبکه عصبی که در طراحی آن باید مد نظر قرار گیرند عبارتند از:

۱. معماری شبکه عصبی: طریقه اتصالات بین نرون‌ها، تعداد آنها و تعداد لایه‌های متشکل از بخش نرون‌های ارتباطی را معماری شبکه عصبی گویند.

۲. تابع فعال‌ساز: نوع تابعی که انتخاب می‌شود تا روی نرون قرار گرفته و خروجی نرون را تولید نماید، تابع فعال‌ساز آن نرون گویند. ماهیت تابع f بسته به نوع خروجی‌هایی که انتظار داریم، شبکه آنها را حاصل کند، تعیین می‌شود. در شکل ۳.۲ انواع توابع فعال‌ساز مشاهده می‌شود.

۳. یادگیری شبکه عصبی: یادگیری عبارت است از یافتن وزن‌ها، بایاس‌ها و توابع فعال‌ساز به‌طوری‌که در یک دامنه‌کاری از هر ورودی، خروجی مطلوب حاصل گردد.

نام	نمودار	معادله	مشتق
تابع همانی		$f(x) = x$	$f'(x) = 1$
تابع پله‌ای دودویی		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
تابع لوجستیک (پله‌ای گام نرم)		$f(x) = \frac{1}{1 + e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
تابع تانژانت هذلولوی		$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
تابع آرگ تانژانت		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
تابع واحد خطی یکسوسازی شده (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
تابع واحد خطی یکسوسازی شده پارامتریک (PReLU)		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
تابع واحد خطی نمایی (ELU)		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$

شکل ۳.۲: انواع تابع فعال‌سازی

۱.۳.۲ تاریخچه‌ای از حل مسائل بهینه‌سازی با استفاده از شبکه‌های عصبی

در سال ۱۹۸۶ تانک و هاپفیلد^۱ [۱۵] به معرفی اولین شبکه عصبی برای حل مسائل برنامه‌ریزی خطی پرداختند. آن‌ها نشان دادند که وضعیت شبکه در هر تکرار به نحوی تغییر می‌کند که تابع انرژی متناظر با آن، به طور یکنواخت کاهش می‌یابد تا جایی که به نقطه کمینه خود می‌رسد که این نقطه کمینه متناظر با نقطه تعادل شبکه عصبی است. آنها این شبکه عصبی را توسط یک مدار الکتریکی اجرا نمودند. همچنین از این شبکه برای حل مساله فروشنده دوره گرد با ۳۰ شهر استفاده کردند. یکی از نقص‌های این شبکه این بود که نقطه تعادل شبکه در شرایط بهینگی صدق نمی‌کرد. با این وجود کارهای هاپفیلد مشوق بسیاری از محققان بود تا در این زمینه فعالیت کنند. کندی^۲ و چا^۳ با افزودن یک پارامتر جریمه متناهی به شبکه هاپفیلد آن را گسترش دادند [۱۸] و از آن برای حل مسئله برنامه‌ریزی غیرخطی محدب استفاده نمودند. با پارامترهای جریمه توانایی این شبکه در پیدا کردن نقطه بهینه دقیق کم شد مخصوصاً اگر پارامتر جریمه بزرگ باشد، این شبکه به سختی کار می‌کند. برای جلوگیری از به کار بردن پارامتر جریمه رودریگز-وازکز^۴ و همکاران یک شبکه عصبی سوئیچ-خازن^۵ معرفی کردند [۲۷].

ما^۶ و شانبلات^۷ یک شبکه عصبی دو فازي ارائه کردند که فاز اول شبکه مشابه شبکه کندی و چا بود و در فاز دوم مسیر شبکه به جواب دقیق مساله همگرا می‌شد. در نتیجه این روش جواب‌های دقیق‌تری نسبت به شبکه کندی و چا ارائه داد. مشکل این شبکه در این بود که پایداری فاز دوم شبکه وابسته به انتخاب یک مقدار بزرگ برای پارامتر جریمه بود و لذا اگرچه تاثیر پارامتر جریمه در این روش کاهش یافته و جواب‌های حاصل دقیق‌تر بودند اما هنوز مستقل از پارامتر نبود. نویسندگان دیگر بر مبنای روش لاگرانژ یک شبکه عصبی ارائه نمودند که کاملاً مستقل از پارامتر جریمه بود و توانایی حل مسائل غیرخطی را داشت. نقطه تعادل این شبکه عصبی در شرایط بهینگی مرتبه اول و دوم صدق می‌کرد و هم چنین شبکه حاصل همگرا بود. در سال ۱۹۹۳ بوزردوم^۸ و پتیسن^۹ [۷] شبکه‌ای را بر مبنای گرادیان و روش تصویر و مستقل از پارامتر جریمه ابداع کردند که توانایی حل مسائل درجه دوم با متغیرهای کران‌دار را داشت. این روش در عمل روش کارایی بود اما قادر به حل حالت کلی مسائل برنامه‌ریزی خطی و درجه دوم نبود.

هم‌چنین، شبکه عصبی اولیه-دوگان برای بررسی مسائل درجه دوم توسط شیا^{۱۰} [۳۷] ارائه

¹Hopfield and Tank

²Kennedy

³Chua

⁴Rodriguez-Vazquez

⁵Switched-Capacitor

⁶Maa

⁷Shanblatt

⁸Bouzerdoun

⁹Pattison

¹⁰Xia

شد. تائو و فانگ^۱ [۱۳] و تائو و همکاران^۲ [۳۱] یک روش شبکه عصبی بر مبنای روش‌های تصویرسازی و گرادیان برای حل مسائل برنامه‌ریزی غیرخطی محدب ارائه دادند. هم‌چنین، تعدادی شبکه‌های عصبی را بر اساس یک تابع انرژی تعریف‌شده و برخی دیگر بر اساس شرایط وجود بهینگی، مانند شرایط لازم کان-تاکر [۸، ۳۲، ۳۳، ۱۹] طراحی کرده‌اند. به‌علاوه برخی از کارهای اخیر در [۱۶، ۳۴، ۵، ۲۵] ارائه شده‌است.

۲.۳.۲ مروری بر سیستم‌های دینامیکی

در این بخش سیستم‌های دینامیکی^۳ را معرفی کرده و قضیه‌ی اساسی وجود و یگانگی وجود جواب مسئله مقدار اولیه و قضایای مربوط به پایداری را بیان می‌کنیم. دستگاه معادلات دیفرانسیل به شکل زیر را در نظر بگیرید:

$$\frac{dx}{dt} = f(x(t), t) \quad (۱.۲)$$

که در آن $x(t) = [x_i(t)]^T$ ، $t \geq 0$ بردار وضعیت یا حالت^۴ می‌نامند. چنین دستگاهی که در آن $t \leq 0$ ، یک دستگاه وابسته به زمان یا غیرخودگردان^۵ است. به این معنی که متغیر t به طور واضح در ضابطه f ظاهر می‌شود. اگر f آشکارا به t وابسته نباشد به عبارتی

$$\frac{dx}{dt} = f(x(t)) \quad (۲.۲)$$

در این صورت سیستم خودگردان^۶ یا مستقل از زمان نامیده می‌شود. به این دستگاه معادلات دیفرانسیلی که بردار وضعیت آن با تغییر زمان تغییر می‌کند، یک سیستم دینامیکی گفته می‌شود. در ادامه از نماد $\dot{x}(t)$ برای نمایش مشتق x نسبت به زمان استفاده می‌نماییم.

قضیه ۱.۳.۲. (قضیه‌ی اساسی وجود و یگانگی جواب مسئله مقدار اولیه) فرض کنید D زیرمجموعه‌ی باز از R^n شامل نقطه x_0 باشد و f_i تابعی پیوسته و مشتق‌پذیر روی R^n باشد یا به عبارتی $f_i \in C^1(R)$ ، در این صورت به ازای هر $a > 0$ که $a \in R$ ، مسئله مقدار اولیه زیر روی بازه $[-a, a]$ جواب یگانه $x(t)$ دارد:

$$\begin{cases} \dot{x}(t) = f(x(t)), \\ x(t_0) = x_0. \end{cases}$$

¹Tao and Fang

²Tao et al.

³Dynamical systems

⁴State vector

⁵Non-autonomous

⁶Autonomous

تعریف ۱.۳.۲. (نقطه‌ی تعادل^۱) جواب سیستم دینامیکی (۲.۲) را نقطه‌ی تعادل می‌نامیم و با x^e نمایش می‌دهیم، هرگاه $f(x^e) = 0$. لذا نقطه تعادل جوابی از سیستم $\dot{x}(t) = f(x(t))$ است که مستقل از زمان می‌باشد.

تعریف ۲.۳.۲. (پایداری^۲) نقطه تعادل x^e ، نقطه پایدار نامیده می‌شود هرگاه

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad s.t. \quad \|x(t_0) - x^e\| < \delta \implies \|x(t) - x^e\| < \varepsilon \quad \forall t \geq t_0.$$

تعریف ۳.۳.۲. (پایداری مجانبی^۳) نقطه تعادل x^e به طور مجانبی پایدار است هرگاه

$$\lim_{t \rightarrow +\infty} x(t) = x^e \quad \text{و} \quad \frac{dE(x(t))}{dt} < 0 \quad \forall x \in \Omega - \{x^e\}$$

روش لیاپانف یکی از بهترین روش‌ها برای تحلیل پایداری سیستم‌های دینامیکی است. این روش برای سیستم‌های خطی و غیرخطی از هر مرتبه‌ای قابل استفاده است و شرط کافی برای پایداری مجانبی نقطه تعادل یک سیستم دینامیکی غیرخطی را ارائه می‌دهد. یعنی بدون حل معادلات وضعیت سیستم، می‌توان پایداری نقطه تعادل سیستم غیرخطی را بررسی نمود. در واقع لیاپانف پیشنهاد کرد که ممکن است توابعی وابسته به مسیر وجود داشته باشند که ویژگی آنها تعیین می‌کند که آیا مسیر به یک نقطه‌ی ساکن همگرا می‌شود یا خیر. چنین تابعی به عنوان تابع انرژی برای سیستم دینامیکی در نظر گرفته می‌شود که به تابع لیاپانف نیز معروف است. قضیه‌ی لیاپانف به صورت خلاصه بیان می‌کند که نقطه‌ی تعادل یک سیستم دینامیکی به طور مجانبی پایدار است هرگاه یک تابع لیاپانف در همسایگی این نقطه وجود داشته باشد.

تعریف ۴.۳.۲. تابع لیاپانف تابعی مانند $E(x(t))$ است که در شرایط زیر صدق می‌کند:

۱. $E(x(t))$ و مشتقات جزئی مرتبه اول آن همگی پیوسته باشند، به عبارت دیگر

$$E(x(t)) \in C^1$$

۲. $E(x(t)) \geq 0$ ، مخصوصاً در همسایگی باز نقطه تعادل x^e ، $E(x(t)) > 0$ و

$$E(x) = 0$$

۳. مشتق تابع $E(x(t))$ نسبت به زمان منفی است، به عبارت دیگر:

$$\frac{dE(x(t))}{dt} = [\nabla_{x(t)} E(x(t))]^T \cdot \dot{x}(t) = [\nabla_{x(t)} E(x(t))]^T \cdot f(x(t)) \leq 0, \quad x \in \Omega - \{x^e\}$$

$$\frac{dE(x^e)}{dt} = 0 \quad \text{و در نقطه‌ی تعادل } x^e \text{ داریم}$$

قضیه ۲.۳.۲. نقطه‌ی تعادل سیستم $\dot{x}(t) = f(x(t))$ پایدار است اگر یک تابع لیاپانف متناظر با این سیستم وجود داشته باشد.

¹Equilibrium

²Stability

³Asymptotic

قضیه ۳.۳.۲. نقطه‌ی تعادل سیستم $\dot{x}(t) = f(x(t))$ به طور مجانبی پایدار است اگر یک تابع لیپانف متناظر با این سیستم وجود داشته باشد و در یک همسایگی نقطه‌ی تعادل x^e ، $\frac{dE(x(t))}{dt} < 0$ ، برای $x \neq x^e$.

ملاحظه می‌شود که اگر $\frac{dE(x(t))}{dt} \leq 0$ باشد، آنگاه x^e پایدار است ولی لزوماً به طور مجانبی پایدار نیست.

فصل ۳

تجزیه فضای پارامتریک برای مسائل بهینه‌سازی دوهدفه

در این فصل، یک روش شبکه عصبی (NNA)^۱ مبتنی بر گرادیان برای حل مسائل برنامه‌ریزی غیرخطی (NLPPs)^۲ و مسائل بهینه‌سازی دوهدفه (BOOPs)^۳ ارائه می‌شود. روش پیشنهادی دارای دو ویژگی بارز است. در مرحله اول، مسئله بهینه‌سازی دوهدفه با استفاده از روش مجموع وزنی به مسئله برنامه‌ریزی غیرخطی تبدیل می‌شود. سپس، بر اساس نظریه‌های لیاپانوف و لاسال^۴، برای یک نقطه اولیه دلخواه، ثابت شده‌است که روش پیشنهادی پایدار سراسری است و می‌تواند جواب‌های بهینه را برای مسائل برنامه‌ریزی غیرخطی و مسائل برنامه‌ریزی دوهدفه به دست آورد. به علاوه، جواب‌های بهینه پارتو مسائل برنامه‌ریزی دوهدفه با استفاده از وزن‌های مختلف تولید می‌شوند. علاوه بر این، فضای پارامتری مسائل برنامه‌ریزی دوهدفه بر اساس مجموعه پایداری نوع اول، مورد تجزیه و تحلیل قرار می‌گیرد. اغلب مطالب این فصل از [۴] گرفته شده است. هدف اصلی این فصل به شرح زیر است :

۱. یک روش شبکه عصبی مبتنی بر گرادیان برای حل مسائل برنامه‌ریزی غیرخطی و مسائل برنامه‌ریزی دوهدفه ارائه شده است. در روش شبکه عصبی مبتنی بر گرادیان می‌توان مسائل

¹neural network approach

²nonlinear programming problems

³bi-objective optimization problems

⁴ Lyapunov and LaSalle

اولیه و دوگان را به‌طور هم‌زمان حل کرد.

۲. یک تابع جدید به تابع انرژی اضافه شده‌است به گونه‌ای که این تابع محدب و تفکیک‌پذیر است. بر اساس تابع انرژی، می‌توان سیستمی از معادلات دیفرانسیل را ایجاد کرد که مربوط به یک شبکه عصبی است.

۳. با استفاده از روش مجموع وزنی، می‌توان با استفاده از وزن‌های مختلف، جواب‌های بهینه پارتو را به‌دست آورد.

۴. از مجموعه پایداری به‌عنوان اولین مورد برای تجزیه و تحلیل فضای پارامتری مسائل برنامه‌ریزی دوهدفه استفاده می‌شود.

۵. کارایی روش شبکه عصبی مبتنی بر گرادینان با حل شش مسئله برنامه‌ریزی غیرخطی و شش مسئله برنامه‌ریزی دوهدفه نشان داده می‌شود.

۱.۳ مجموعه پایداری نوع اول

مسئله اسکالرسازی مجموع وزن‌دار برای مسائل بهینه‌سازی دوهدفه که به فرم زیر می‌باشد، را در نظر بگیرید:

$$\begin{aligned} P(w): \min F(x) &= \sum_{i=1}^{i=K} w_i f_i(x) = w_1 f_1(x) + w_2 f_2(x) \\ \text{subject to } x &\in \Omega, \end{aligned} \quad (1.3)$$

که در آن

$$w \in W = \left\{ w \in \mathbb{R}^P \mid \sum_{i=1}^K w_i = 1, w_i \geq 0 \right\}. \quad (2.3)$$

تعریف ۱.۱.۳. فرض کنید مسئله $P(w)$ به ازای $w^* \in W$ دارای جواب بهینه پارتو x^* باشد. آن‌گاه مجموعه پایداری نوع اول مسئله (۲.۱) متناظر با جواب x^* ، با $S(x^*)$ نشان داده می‌شود و بنا به تعریف عثمان^۱ [۲۴] و ابوسینا و حسین^۲ [۲، ۳] به شکل زیر تعریف می‌شود.

$$S(x^*) = \{w \in W \mid x^* \text{ جواب بهینه پارتو مسئله (۲.۱) است.}\} \quad (3.3)$$

به‌سادگی می‌توان نشان داد که:

۱. مجموعه $S(x^*)$ بسته و محدب است.

۲. اگر $\text{int}[S(\bar{x}) \cap S(x^*)] \neq \emptyset$ آنگاه $S(\bar{x}) = S(x^*)$.

¹Osman

²Abo-sinna and Hussein

۱.۱.۳ محاسبه مجموعه پایداری نوع اول

اگر نقطه $x^* \in \Omega$ جواب بهینه پارتو مسئله (۲.۱) باشد، آن‌گاه وجود دارد $w^* \in W$ به طوری که x^* جواب بهینه پارتو $P(w^*)$ است. بنابراین از پایداری مسئله $(P(w^*))$ نتیجه می‌شود که وجود دارد $\mu \geq 0$ ، $\mu \in \mathbb{R}^{p+J}$ (یعنی قیدهای تساوی می‌توانند مانند $h_j(x) - \varepsilon \leq 0$ به نامساوی تبدیل شوند) به طوری که (منگاسریان [۲۳] را ببینید)

$$w^T \frac{\partial F}{\partial x}(x^*) + \mu^T \frac{\partial g}{\partial x}(x^*) = 0, \quad g(x^*) \leq 0, \quad \mu^T g(x^*) = 0 \quad (4.3)$$

که در آن η^T ترانهاده برادر η است. مجموعه قیدهای فعال در x^* را با $A(x^*)$ به صورت زیر نشان می‌دهیم:

$$A(x^*) = \{p | g_p(x^*) = 0\}. \quad (5.3)$$

بنابراین دستگاه معادلات مستقل خطی به صورت زیر فرمول‌بندی می‌شود

$$w^T \frac{\partial F}{\partial x}(x^*) + \sum_{p \in A(x^*)} \mu_p \frac{\partial g_p(x^*)}{\partial x} = 0, \quad (6.3)$$

که به شکل ماتریسی زیر نوشته می‌شود:

$$\begin{bmatrix} C' & D' \end{bmatrix} \begin{bmatrix} w \\ \mu \end{bmatrix} = 0 \quad (7.3)$$

که $C' = [c'_{ai}]$ یک ماتریس $s \times K$ ، $D' = [d'_{ai}]$ یک ماتریس $s \times k$ ، $w \in \mathbb{R}^K$ ، $\mu \in \mathbb{R}^k$ ، $w \neq 0$ و $\mu \neq 0$ است، که در آن $s \leq \sum_{i=1}^n k_i$ و k عدد کاردینال^۱ مجموعه $A(x^*)$ است. فرض کنید $a \in I \subset \{1, 2, \dots, k\}$ ، $i = 1, 2, 3, \dots, k$ ، $d'_{ai} = 0$ که عدد کاردینال I معادل با $s - l$ فرض شده است. بنابراین، گشتاور دستگاه باقیمانده به شکل زیر است

$$\begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} w \\ \mu \end{bmatrix} = 0 \quad (8.3)$$

که C و D به ترتیب ماتریس‌هایی از مرتبه $l \times K$ و $l \times k$ هستند. بنابراین دستگاه (۸.۳) همراه با شرط $a \in I$ ، $\sum_{i=1}^K C'_{ai} w_i = 0$ ، سیستم (۷.۳) را می‌دهد که معادل سیستم (۴.۳) است. بنابراین، ما دو قضیه زیر را از عثمان^۲ [۲۴] و ابوسینا و حسین^۳ [۳، ۲] داریم.

قضیه ۱.۱.۳. اگر $k \geq l$ آن‌گاه

$$S(x^*) = \left\{ w \in \mathbb{R}^K \mid \left(w^T C^T (D^T)^{-1} \right)_i \leq 0, i = 1, 2, \dots, l, \sum_{i=1}^K C'_{ai} w_i = 0, a \in I \right\}. \quad (9.3)$$

^۱cardinal number

^۲Osman

^۳Obo-sinna and Hussein

قضیه ۲.۱.۳. اگر $k < l$ آن‌گاه

$$S(x^*) = \left\{ w \in \mathbb{R}^K \mid \begin{cases} (w^T C_1^T - C_1^T (D_1^T)^{-1} D_2^T)_i = 0, i = 1, 2, \dots, k-l, \\ (w^T C_1^T (D_1^T)^{-1}) \leq 0, i = 1, 2, \dots, k, \sum_{i=1}^K C'_{ai} w_i = 0, a \in I \end{cases} \right\}. \quad (10.3)$$

که $D = \begin{bmatrix} D_1 & D_2 \end{bmatrix}$ ، D_1 و D_2 به ترتیب ماتریس‌هایی $l \times l$ و $l \times k - 1$ هستند و η_i درایه ستون i ام بردار ردیف η است.

اگر w با شرط $\sum_{i=1}^n w_i = 1$ نرمال شود آن‌گاه می‌توانیم الگوریتم زیر را برای محاسبه مجموعه $S(x^*)$ به کار ببریم.

الگوریتم ۱

- ۱: با $w_1 = w^0 = 0$ و $w_2 = 1 - w^0$ شروع کنید.
- ۲: مسئله (Pw^0) را حل کرده و جواب بهینه پارتو x^* مسئله (۲.۱) بدست آورید.
- ۳: جایگذاری شرایط کان-تاکر برای به دست آوردن دستگاه (۶.۳)
- ۴: اگر $s = p + k - 1$ ، آن‌گاه $S(x^*) = \{tw^0 \mid t \geq 0\}$.
- ۵: در پایان مرحله ۴، سیستم (۸.۳) به سادگی به دست می‌آید. مجموعه I را مشخص کنید.
- ۶: اگر $k \geq l$ باشد آن‌گاه $S(x^*)$ با معادله (۹.۳) به دست می‌آید و اگر $k < l$ باشد آن‌گاه $S(x^*)$ با معادله (۱۰.۳) به دست می‌آید.

۲.۳ رویکرد شبکه عصبی پیشنهادی برای مسائل بهینه‌سازی دوهدفه

در این بخش، ابتدا درباره روش کلی برای بهینه‌سازی روش شبکه عصبی بحث می‌کنیم، و سپس اهمیت نظری آن را نشان می‌دهیم.

۱.۲.۳ شرح رویکرد شبکه عصبی پیشنهادی

برای حل مسئله بهینه‌سازی از طریق روش پیشنهادی، گام اصلی ایجاد یک تابع انرژی $E(z)$ مناسب است، تا کمترین سطح انرژی متناظر با جواب بهینه z^* به دست آید. براساس تابع انرژی، ما یک سیستم‌گرادیان از معادلات دیفرانسیل ایجاد می‌کنیم که روش شبکه عصبی را شبیه‌سازی می‌کند. فرض کنید مدل تبدیل‌شده‌ی مسائل بهینه‌سازی دوهدفه، بر اساس روش مجموع وزنی، به صورت

مسئله برنامه‌ریزی غیرخطی محدب زیر است.

$$\begin{aligned} \text{CNLPP: } \min F &= \sum_{k=1}^K w_k f_k(x), \\ \text{subject to } x &\in \Omega, w \in W \end{aligned} \quad (11.3)$$

که در آن

$$W = \left\{ w \in \mathbb{R}^P \mid w_k \geq 0, \sum_{k=1}^K w_k = 1 \right\}$$

فرض می‌کنیم مسئله برنامه‌ریزی غیرخطی محدب یک جواب بهینه دارد و این جواب پایدار قوی است، یعنی وجود دارد x^* به‌طوری‌که $(p=1, 2, \dots, P), g_p(x^*) > 0$ ، $(j=1, 2, \dots, J), h_j(x^*) = 0$ (شرط اسلاتر). با توجه به نتایج داده‌شده در منگاساریان^[۲۳]، مسئله دوگان (DP) می‌تواند به شرح زیر محاسبه می‌شود

$$\begin{aligned} \text{DP : } \max_{\mathbf{x}, \lambda, \mu} L(\mathbf{x}, \lambda, \mu) \\ \text{subject to} \end{aligned} \quad (12.3)$$

$$\nabla_x L(\mathbf{x}, \lambda, \mu) = 0, \lambda \geq 0$$

که در آن

$$\begin{aligned} \lambda &= (\lambda_1, \lambda_2, \dots, \lambda_P)^T, \mu = (\mu_1, \mu_2, \dots, \mu_J)^T, \\ L(\mathbf{x}, \lambda, \mu) &= F(\mathbf{x}) - \sum_{p=1}^P \lambda_p g_p(\mathbf{x}) - \sum_{j=1}^J \mu_j h_j(\mathbf{x}), \\ \nabla_{\mathbf{x}} L(\mathbf{x}, \lambda, \mu) &= \nabla F(\mathbf{x}) - \sum_{p=1}^P \lambda_p \nabla g_p(\mathbf{x}) - \sum_{j=1}^J \mu_j \nabla h_j(\mathbf{x}), \end{aligned} \quad (13.3)$$

یا به شکل ماتریسی زیر است

$$\begin{aligned} L(\mathbf{x}, \lambda, \mu) &= F(\mathbf{x}) - \lambda^T g(\mathbf{x}) - \mu^T h(\mathbf{x}), \\ \nabla_{\mathbf{x}} L(\mathbf{x}, \lambda, \mu) &= \nabla F(\mathbf{x}) - \lambda^T \nabla g(\mathbf{x}) - A^T \mu, \end{aligned} \quad (14.3)$$

که در آن

$$\begin{aligned} g(\mathbf{x}) &= (g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_p(\mathbf{x}))^T, \quad h(\mathbf{x}) = (h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_J(\mathbf{x}))^T, \\ \nabla g(\mathbf{x}) &= [\nabla g_1(\mathbf{x}), \nabla g_2(\mathbf{x}), \dots, \nabla g_p(\mathbf{x})], \quad z = (\mathbf{x}^T, \lambda^T, \mu^T)^T, \end{aligned} \quad (15.3)$$

$$\nabla g(\mathbf{x}) = \begin{bmatrix} \frac{\partial g_1(\mathbf{x})}{\partial x_1} & \frac{\partial g_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial g_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial g_2(\mathbf{x})}{\partial x_1} & \frac{\partial g_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial g_2(\mathbf{x})}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial g_p(\mathbf{x})}{\partial x_1} & \frac{\partial g_p(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial g_p(\mathbf{x})}{\partial x_n} \end{bmatrix}_{(P \times n)}, \quad ; A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_J^T \end{bmatrix}_{(J \times n)},$$

قضیه ۱.۲.۳. [۲۳] فرض کنید x^* یک جواب بهینه مسئله بهینه‌سازی غیرخطی محدب باشد، آنگاه $(\lambda^{*T}, \mu^{*T})^T \in \mathbb{R}^{P+J}$ وجود دارد، به طوری که $(x^{*T}, \lambda^{*T}, \mu^{*T})^T$ یک جواب بهینه برای مسئله دوگان، برای یک مقدار مشخص $w \in W$ است، که در آن W نشان‌دهنده مجموعه‌ای از بردارهای ضریب وزن $w = (w_1, w_2, \dots, w_K)^T$ باشد.

قضیه ۲.۲.۳. [۲۲] فرض کنید $g(x) \in \mathbb{R}^P$ و $g(x) \in C^1$ (که C^1 مجموعه توابع پیوسته مشتق‌پذیر است) باشد. آنگاه

$$g(x) \geq 0 \Leftrightarrow \frac{1}{P} g(x)^T [g(x) - |g(x)|] = 0 \quad (16.3)$$

و $\frac{1}{P} g(x)^T [g(x) - |g(x)|] \in C^1$ که $|x| = (|x_1|, |x_2|, \dots, |x_n|)^T$ است.

با کمک قضیه ۲.۲.۳، محدودیت‌های نامساوی می‌توانند به محدودیت‌های تساوی تبدیل شوند و پس از تبدیل، مشتق‌پذیری تابع جدید همانند تابع $g(x)$ است. به راحتی می‌توان ثابت کرد که $\frac{1}{P} g(x)^T [g(x) - |g(x)|] \in C^1$ زمانی تابع محدب است که $g(x)$ یک تابع مقعر باشد.

قضیه ۳.۲.۳. [۲۳] $x^* \in \mathbb{R}^n$ و $(x^{*T}, \lambda^{*T}, \mu^{*T})^T$ برای یک مقدار خاص W به ترتیب، جواب‌های بهینه مسائل بهینه‌سازی غیرخطی محدب و مسئله دوگان هستند، اگر و فقط اگر

$$\begin{cases} F(x^*) = L(x^*, \lambda^*, \mu^*), & \nabla_x L(x^*, \lambda^*, \mu^*) = 0, \\ g(x^*) \geq 0, & h(x^*) = 0, \quad \lambda^* \geq 0, \end{cases}$$

یا

$$(17.3)$$

$$\begin{cases} \lambda^{*T} g(x^*) = 0, & \nabla_x L(x^*, \lambda^*, \mu^*) = 0, \\ g(x^*) \geq 0, & Ax^* - b = 0, \quad \lambda^* \geq 0, \end{cases}$$

که در آن $b = (b_1, b_2, \dots, b_J)^T$.

برهان. بنا به قضیه‌های ۲.۲.۳ و ۳.۲.۳، تابع انرژی مسئله بهینه‌سازی غیرخطی محدب می‌تواند به شرح زیر ساخته شود

$$E(z) = E(x, \lambda, \mu) = \frac{1}{4} [\lambda^T g(x)]^2 + \frac{1}{4} \|\nabla_x L(x, \lambda, \mu)\|^2 + \frac{1}{4} g(x)^T [g(x) - |g(x)|] \quad (18.3)$$

$$+ \frac{1}{4} \|Ax - b\|^2 + \frac{1}{4} \lambda^T (\lambda - |\lambda|)$$

که $z = (x^T, \lambda^T, \mu^T)^T \in \mathbb{R}^{n+P+J}$. صفر بودن هر جمله سمت راست معادله (۱۸.۳) متناظر با هر تساوی یا نامساوی معادله (۱۷.۳) و $E(z) \geq 0$ است. بعلاوه $E(z)$ تابعی مشتق‌پذیر است. □

قضیه ۴.۲.۳. [۲۳] $z^* = (x^{*T}, \lambda^{*T}, \mu^{*T})^T$ یک جواب صفر برای $E(z)$ است اگر و فقط اگر z^* یک جواب بهینه برای مسائل بهینه‌سازی غیرخطی محدب و مسئله دوگان باشد. (یعنی x^* و $(x^{*T}, \lambda^{*T}, \mu^{*T})^T$ به‌ترتیب جواب های بهینه مسائل بهینه‌سازی غیرخطی محدب و مسئله دوگان برای مقدار مشخص W باشند).

برهان. بنا به قضیه ۳.۲.۳، $z^* = (x^{*T}, \lambda^{*T}, \mu^{*T})^T$ یک جواب صفر برای $E(z)$ است اگر و فقط اگر z^* در معادله (۱۱.۳) صدق کند. اگر و فقط اگر x^* و $(x^{*T}, \lambda^{*T}, \mu^{*T})^T$ به‌ترتیب جواب های بهینه مسائل بهینه‌سازی غیرخطی محدب و مسئله دوگان برای مقدار مشخص W باشند. بنابراین قضیه ۴.۲.۳ اثبات می‌شود.

□

با استفاده از ایده اصلاح‌شده در قضیه لیانگ و همکاران^۱ [۲۲]، می‌توانیم از سیستم گرادیان برای ایجاد رویکرد شبکه‌عصبی جدید برای حل مسائل بهینه‌سازی غیرخطی محدب و مسائل بهینه‌سازی دوهدفه به شرح زیر استفاده کنیم

$$\frac{dz}{dt} = -\nabla E(z), \quad (19.3)$$

با فرمول‌های محاسبه گرادیان در لیانگ و همکاران [۲۲]، روش شبکه‌عصبی جدید در معادله (۱۹.۳) را می‌توان با شرح جزئیات زیر نوشت:

$$\begin{aligned} \frac{dx}{dt} &= -\nabla_x E(z) = -\lambda^T g(x) \cdot \nabla g(x)^T \lambda - \nabla g(x)^T [g(x) - |g(x)|] \\ &\quad - \nabla_{xx}^2 L(z) \nabla_x L(z) - A^T (Ax - b) \\ \frac{d\lambda}{dt} &= -\nabla_\lambda E(z) = -\lambda^T g(x) \cdot \nabla g(x) + \nabla g(x) \nabla_x L(z) - (\lambda - |\lambda|) \\ \frac{d\mu}{dt} &= -\nabla_\mu E(z) = A \nabla_x L(z) \end{aligned} \quad (20.3)$$

که $\nabla_{xx}^2 L(z)$ ماتریس هسین^۲ تابع $L(z)$ نسبت به x است. فرض کنید $\nabla E(z)$ پیوسته لیپ‌شیتز^۳ باشد، آنگاه مسئله معادلات دیفرانسیل با مقدار اولیه، در معادله (۱۹.۳) یک جواب بهینه منحصر به فرد

¹Leung et al.

²Hessian

³Lipschitz

دارد. از آنجایی که تابع سمت راست معادلات (۱۹.۳) یا (۲۰.۳) پیوسته هستند، این معادلات را می‌توان با اجرای نرم‌افزار شبکه MATLAB به راحتی به‌دست آورد. بنابراین، این یک شبکه عصبی شدنی است.

۲.۲.۳ تجزیه و تحلیل پایداری

برای یک مسئله بهینه‌سازی محدب، اگر گرادیان تابع هدف پیوسته لیپ‌شیتز و از پایین کران‌دار باشد، ثابت خواهیم کرد که:

۱. هر جواب (۲۰.۳) همگرا به یک نقطه تعادل است.

۲. در رویکرد شبکه عصبی پیشنهادی پایداری لیاپانف^۱ معادل پایداری مجانبی است. برای مسائل بهینه‌سازی غیرخطی محدب یا مسائل بهینه‌سازی دوهدفه، ما نشان می‌دهیم هر جواب (۲۰.۳)، به نقطه تعادل پایدار مجانبی همگرا است.

قضیه ۵.۲.۳. [۲۲] مجموعه‌ای از نقاط تعادل معادله (۱۹.۳) را به‌صورت

$$M = \{z = (x^T, \lambda^T, \mu^T)^T \in \mathbb{R}^{n+P+J} | \nabla E(z) = 0\}$$

$\Theta = \{z = (x^T, \lambda^T, \mu^T)^T \in \mathbb{R}^{n+P+J}\}$ که در آن x و $(x^T, \lambda^T, \mu^T)^T$ به ترتیب جواب‌های بهینه مسئله بهینه‌سازی غیرخطی محدب و مسئله بهینه‌سازی دوگان هستند، آنگاه $\Theta \subset M$ است. (یعنی هر جواب بهینه z از مسائل بهینه‌سازی غیرخطی محدب و دوگان، نقطه تعادل شبکه عصبی جدید در معادله (۱۹.۳) است.)

برهان. فرض کنید $z = (x^T, \lambda^T, \mu^T)^T \in E$ ، یعنی $\nabla E(z) = 0$ و $z = (x^{*T}, \lambda^{*T}, \mu^{*T})^T$ و x^* و z^* به‌ترتیب جواب‌های بهینه مسائل بهینه‌سازی غیرخطی محدب و مسائل بهینه‌سازی دوگان باشند. آنگاه بنا به قضیه ۳.۲.۳، $E(z^*) = 0$ است، یعنی z^* نقطه مینیم تابع مشتق‌پذیر $E(z)$ است. از آنجاکه $E(z)$ یک تابع مشتق‌پذیر است، بنابراین شرایط لازم و کافی توابع مشتق‌پذیر نتیجه می‌گیریم، $\nabla E(z^*) = 0$ یعنی z^* یک نقطه تعادل برای شبکه معادله (۱۹.۳) است. و از این‌رو $z \in M$ است. بنابراین $\Theta \subset M$. $\square$

قضیه ۶.۲.۳. [۳۶] فرض کنید شبکه عصبی (۱۹.۳) دارای نقطه تعادل منحصربه‌فرد z^* است، آنگاه z^* به‌طور مجانبی و به‌طور یکنواخت پایدار است.

برهان. فرض کنید $V(z) = E(z) - E(z^*) \geq 0$. یعنی $V(z)$ یک تابع مثبت‌معین است. لذا $V(z) > 0, \forall z \neq z^*$. علاوه‌براین، به ازای یک نقطه اولیه دلخواه مانند z^0 ، یک مسیر منحصربه‌فرد $z = z(t, z^0)$ از سیستم معادله (۱۹.۳) وجود دارد، به‌طوری‌که

$$\begin{aligned} \frac{d}{dt} V(z) &= \frac{d}{dt} V[z(t, z^0)] \\ &= \nabla V(z)^T \cdot \frac{dz}{dt} = -\|\nabla E(z)\|^2 < 0, \text{ if } z \neq z^* \end{aligned}$$

¹Lyapunov

با استفاده از قضیه لیاپانف [۳۶]، z^* به‌طور مجانبی و به‌طور یکنواخت پایدار است. □

قضیه زیر نشان می‌دهد که تحت برخی شرایط، مسیر سیستم (۱۹.۳) به نقطه تعادل همگرا می‌شود.

قضیه ۷.۲.۳. فرض کنید یک مجموعه تراز $L(z^\circ) = \{z \in \mathbb{R}^{n+P+J} : E(z) \leq E(z^\circ)\}$ کران‌دار باشد، آنگاه یک نقطه تعادل $\bar{z} \in M$ و یک دنباله اکیدا صعودی $\{t_n\}$ ، $(t_n \geq 0)$ وجود دارد به‌طوری‌که

$$\lim_{n \rightarrow +\infty} z(t_n, z^\circ) = \bar{z}, \quad \nabla E(\bar{z}) = 0.$$

برهان. ۱. ثابت می‌کنیم که $E[z(t, z^\circ)]$ در طول مسیر $z = z(t, z^\circ)$ یکنوای نزولی است. در حقیقت، در امتداد مسیر $z = z(t, z^\circ)$ داریم

$$\dot{E}(z) = \frac{d}{dt} E[z(t, z^\circ)] = \nabla E(z)^T \frac{dz}{dt} = -\|\nabla E(z)\|^2 \leq 0.$$

۲. ثابت می‌کنیم که $\gamma^+(z^\circ) = \{z(t, z^\circ) : t \geq 0\}$ یک مسیر نیمه‌مثبت کران‌دار است. درواقع، $E(z) \geq 0$ از پایین کران‌دار و یک تابع پیوسته است، بنابراین باتوجه به قسمت (۱)، مجموعه تراز $L(z^\circ)$ مجموعه‌ای بسته و کران‌دار است و $\gamma^+(z^\circ) = \{z(t, z^\circ) : t \geq 0\} \subset L(z^\circ)$ از این رو $z = z(t, z^\circ)$ ($t \geq 0$) یک نیمه‌مسیر مثبت معین در $\mathbb{R}^{n+P+J}$ است. اکنون، ثابت می‌کنیم که وجود دارد $\bar{z} \in L(z^\circ)$ و دنباله‌ی $\{t_n\}$ ، $t_n \rightarrow +\infty$ ، $t_n \geq 0$ ، اکیداًصعودی، به‌طوری‌که $\lim_{n \rightarrow +\infty} z(t_n, z^\circ) = \bar{z}$. در واقع $\gamma^+(z^\circ)$ مجموعه‌ای کران‌دار از نقاط یک دنباله اکیداً صعودی $\{t_n\}$ ، $0 \leq \bar{t}_1 < \bar{t}_2 < \dots < \bar{t}_n < t_{n+1}^- < \dots$ ، $t_n \rightarrow +\infty$ ، اکیداًصعودی است، آنگاه فرض کنید $z(\bar{t}_n, z^\circ)$ کران‌دار و دنباله‌ای مرکب از بی‌نهایت نقطه است. لذا، یک نقطه حدی $\bar{z}$ وجود دارد، یعنی یک دنباله اکیداًصعودی $\{t_n\} \subset \{\bar{t}_n\}$ ، $t_n \rightarrow +\infty$ وجود دارد به‌طوری‌که

$$\lim_{n \rightarrow +\infty} z(t_n, z^\circ) = \bar{z}$$

درانتها، ما $\nabla E(\bar{z}) = 0$ را ثابت می‌کنیم.

با توجه به قسمت (۱)، $E(z)$ یک تابع لیاپانف در $\mathbb{R}^{n+P+J}$ است.

$$\dot{E}(\bar{z})|_{(19.3)} = 0 \Leftrightarrow \nabla E(\bar{z}) = 0. \quad [20]$$

$z(t_n, z^\circ) \rightarrow S'(t \rightarrow \infty)$ بزرگترین مجموعه‌ی ناورد در مجموعه نقاط تعادل است)، یعنی

$$\nabla E(\bar{z}) = 0, \quad \lim_{n \rightarrow +\infty} z(t_n, z^\circ) = \bar{z} \text{ به‌طوری‌که } t_n \geq 0, \{t_n\} \text{ یک دنباله}$$

بنابراین اثبات کامل می‌شود. □

قضیه زیر نشان می‌دهد که شبکه عصبی معادله (۱۹.۳) پایدارمجانبی و سراسری است.

قضیه ۸.۲.۳. فرض کنید شبکه عصبی (۱۹.۳) دارای یک نقطه تعادل منحصربه‌فرد باشد و منحنی تراز $L(z^\circ) = \{z \in \mathbb{R}^{n+P+J} : E(z) \leq E(z^\circ)\}$ کران‌دار است، آنگاه هر مسیر شبکه معادله (۱۹.۳)

به جواب بهینه پارتو z^* همگرا می‌شود، یعنی جواب بهینه z^* ، پایدارمجانبی و سراسری برای فضای پارامتری تجزیه‌شده وزن‌ها است.

برهان. بنا به قضیه‌های ۶.۲.۳ و ۷.۲.۳، این قضیه بدیهی است. $\square$

از روش اویلر برای حل معادلات دیفرانسیل غیرخطی توصیف‌شده با معادله (۲۰.۳) [۲۲] استفاده شده است. شبکه‌کد اجرای الگوریتم پیشنهادی در الگوریتم ۲، در صفحه بعد آمده‌است.

۳.۳ مثال‌های عددی

به منظور بررسی کارایی و شدنی بودن روش شبکه عصبی پیشنهادی، دوازده مسئله در جدول‌های ۱.۳ و ۶.۳ ارائه شده است. روش اویلر را برای حل مسئله متناظر با (۲۰.۳) به کار گرفتیم. جدول ۱.۳ مسائل بهینه‌سازی غیرخطی و جدول ۶.۳ مسائل بهینه‌سازی دوهدفه را نشان می‌دهد. در تمام جداول، ما نقاط اولیه مختلفی، $(x_1(t), x_2(t), \dots, x_n(t), \lambda_1, \lambda_2, \dots, \lambda_p)$ (شدنی یا نشدنی) و اندازه گام، Δt را در نظر گرفتیم.

۱.۳.۳ مسائل برنامه‌ریزی غیرخطی

اکنون، برای نشان دادن رویکرد شبکه عصبی پیشنهادی برای مسائل برنامه‌ریزی غیرخطی، مثال‌هایی مانند جدول ۱.۳ ارائه می‌دهیم.

۲.۳.۳ تجزیه و تحلیل نتایج برای مسائل برنامه‌ریزی غیرخطی

برای یافتن جواب‌های مسائل برنامه‌ریزی غیرخطی از رویکرد شبکه عصبی جدید استفاده شده است. جدول ۲.۳ بدون در نظر گرفتن نقاط اولیه در داخل یا خارج از ناحیه شدنی، مقادیر بهینه برای همه مسائل را نشان می‌دهد. شکل ۱.۳ مسیرهای با نقاط اولیه $(0, 0, 0)$ ، $(0, 3, 0)$ ، $(0, 5, 1, 0)$ ، $(3, 0, 0)$ و $(3, 3, 0)$ برای مسئله اول را نشان می‌دهد. همه مسیرها بعد از حدود ۴۰۰ تکرار به جواب بهینه $(x_1^*, x_2^*) = (1, 2)$ و $\lambda_1^* = 1$ همگرا می‌شوند. رفتار متناظر با مسئله دوم در شکل ۲.۳ نشان داده شده است که در آن مسیرهایی با نقاط اولیه $(5, 5, 0, 0)$ ، $(-5, -5, 0, 0)$ ، $(5, 5, 0, 0)$ و $(5, -5, 0, 0)$ پس از حدود ۴۰۰ تکرار به جواب بهینه نظری $(x_1^*, x_2^*) = (0.8944, 0.4472)$ و $(x_1^*, x_2^*) = (-1/118, 0)$ نزدیک می‌شوند. با در نظر گرفتن نقاط اولیه $(4, 4, 0)$ ، $(-4, -4, 0)$ ، $(4, -4, 0)$ و $(-4, 4, 0)$ برای مسئله سوم، مشاهده می‌کنیم که چهار مسیر پس از حدود ۵۰۰ تکرار به جواب بهینه نظری $(x_1^*, x_2^*) = (0, 1)$ و $\lambda_1^* = -0.5$ همگرا می‌شود. رفتار متناظر با آن در شکل ۳.۳ نشان داده شده است. شکل ۴.۳ مسیرهای مسئله چهارم با نقاط اولیه دلخواه به ترتیب $(6, 6, 0, 0)$ ، $(-6, -6, 0, 0)$ ، $(-6, 6, 0, 0)$ و $(6, -6, 0, 0)$ نشان می‌دهد. که به $(x_1^*, x_2^*) = (1, 2)$ و $(\lambda_1^*, \lambda_2^*) = (0, -1)$ همگرا هستند، که جواب بهینه مسئله ۱.۳ پس از ۶۵۰ تکرار است.

اندازه گام Δt	نقاط اولیه مسیره‌ها $(x_1(t), x_2(t), \dots, x_n(t), \lambda_1, \lambda_2, \dots, \lambda_p)$	مسائل
-----------------------	---	-------

۰/۰۸ $(\circ/\circ\Delta, ۱, \circ)$ و $(\circ, ۳, \circ)$ و $(۳, ۳, \circ)$

P۱

$$\text{Min}f(x) = -x_1x_2$$

subject to $x \in \Omega$,

$$\Omega = \{x \in \mathbb{R}^n | g_1(x) = x_1^2 + x_2 - ۳ \leq \circ\}$$

P۲

$$(\circ\Delta, \circ, \circ) \text{ و } (-\circ\Delta, -\circ\Delta, \circ, \circ) \text{ و } (\circ\Delta, -\circ\Delta, \circ, \circ)$$

$$\text{Min}f(x) = -۲x_1 - x_2$$

subject to $x \in \Omega$,

$$\Omega = \left\{x \in \mathbb{R}^n | g_1(x) = x_1^2 + x_2^2 - ۱ \leq \circ, \right\}$$

P۳

$$(۴, ۴, \circ) \text{ و } (-۴, -۴, \circ) \text{ و } (-۴, ۴, \circ) \text{ و } (۴, -۴, \circ)$$

$$\text{Min}f(x) = x_1 - x_2$$

subject to $x \in \Omega$,

$$\Omega = \left\{x \in \mathbb{R}^n | g_1(x) = ۳x_1^2 - ۲x_1x_2 + x_2^2 - ۱ \leq \circ\right\}$$

P۴

$$(۶, ۶, \circ, \circ) \text{ و } (-۶, -۶, \circ, \circ) \text{ و } (-۶, ۶, \circ, \circ) \text{ و } (۶, -۶, \circ, \circ)$$

$$\text{Min}f(x) = x_1^2 + ۲x_2^2 + ۲x_1x_2 - ۱ \circ x_1 - ۱۲x_2$$

subject to $x \in \Omega$,

$$\Omega = \{x \in \mathbb{R}^n | g_1(x) = x_1 + ۳x_2 - ۸ \leq \circ, \}$$

P۵

$$(۴, ۴, \circ, \circ) \text{ و } (-۴, -۴, \circ, \circ) \text{ و } (-۴, ۴, \circ, \circ) \text{ و } (۴, -۴, \circ, \circ)$$

$$\text{Min}f(x) = \circ/۴x_1 + x_2^2 + x_1^2 - x_1x_2 - ۱ \circ x_1 - ۱۲x_2 + x_1^2/۳ \circ$$

subject to $x \in \Omega$,

$$\Omega = \{x \in \mathbb{R}^n | g_1(x) = \circ/۴ - x_1 - \circ/\Delta x_2 \leq \circ, \}$$

الگوریتم ۲ الگوریتم پیشنهادی

۱: آغاز

یک بردار اولیه دلخواه $x \in \mathbb{R}^n$ ، $\lambda \in \mathbb{R}^P$ ، $\mu \in \mathbb{R}^J$ ، انتخاب کنید و قرار دهید $\Delta t > 0$ ،
 $t = 0$ و $\varepsilon = 10^{-19}$.

۲: مسئله بهینه‌سازی دوهدفه را به صورت یک مسئله بهینه‌سازی غیرخطی محدب با استفاده از معادله (۶.۱) فرمول‌بندی کنید.

۳: تا هنگامی که $t \leq T$ (حداکثر تکرار) مراحل زیر را انجام دهید.

۴: گرادیان تابع $E(z)$ را برای به‌دست‌آوردن سیستم معادلات دیفرانسیل زیر محاسبه کنید:

$$e(t) = \nabla_x E(z) = \lambda^T g(x) \cdot \nabla g(x)^T \lambda + \nabla g(x)^T [g(x) - |g(x)|] + \nabla_{xx}^2 L(z) \nabla_x L(z) \\ + A^T(Ax - b)$$

$$\nu(t) = \nabla_\lambda E(z) = \lambda^T g(x) \cdot g(x) - \nabla g(x) \nabla_x L(z) + (\lambda - |\lambda|)$$

$$u(t) = \nabla_\mu E(z) = -A \nabla_x L(z)$$

۵: بردار اولیه را به روز کنید

$$x(t + \Delta t) = x(t) - \Delta t \cdot e(t)$$

$$\lambda(t + \Delta t) = \lambda(t) - \Delta t \cdot \nu(t)$$

$$\mu(t + \Delta t) = \mu(t) - \Delta t \cdot u(t)$$

۶: مقادیر زیر را محاسبه کنید

$$r(t) = \sum_{i=1}^n e_i^2(t); \quad l(t) = \sum_{p=1}^P \nu_p^2(t); \quad m(t) = \sum_{j=1}^J u_j^2(t)$$

۷: معیارهای توقف: اگر r ، l و m کوچکتر از ε هستند، توقف کنید؛ در غیر این صورت قرار دهید $t = t + 1$.

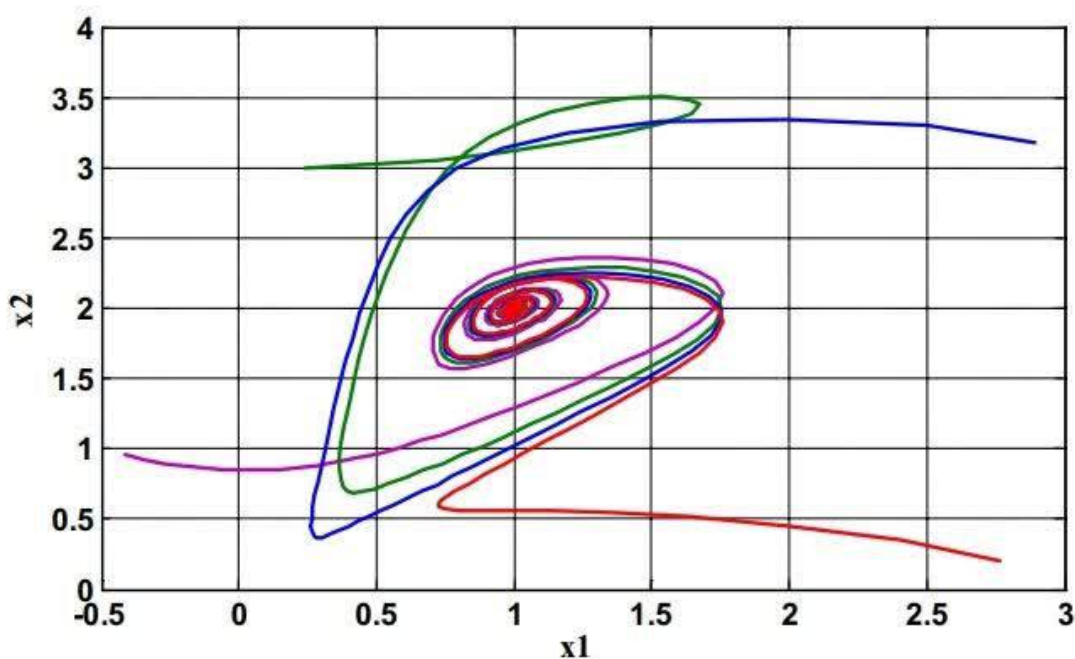
خروجی: $x(t + \Delta t)$ ، $\lambda(t + \Delta t)$ ، $\mu(t + \Delta t)$.

شکل ۵.۳ نشان می‌دهد که مسیرهای مسئله (P۵) با نقاط اولیه $(-4, -4, 0, 0)$ ، $(4, 4, 0, 0)$ ، $(\lambda_1^*, \lambda_2^*) = (-0.1372, -0.5321)$ و $(x_1^*, x_2^*) = (0.2, 0.4)$ هم‌چنین به $(4, -4, 0, 0)$ و $(-4, 4, 0, 0)$ همگرا هستند، که جواب بهینه مسئله ۱.۳ پس از ۵۰۰ تکرار است. شکل ۶.۳ نشان می‌دهد که

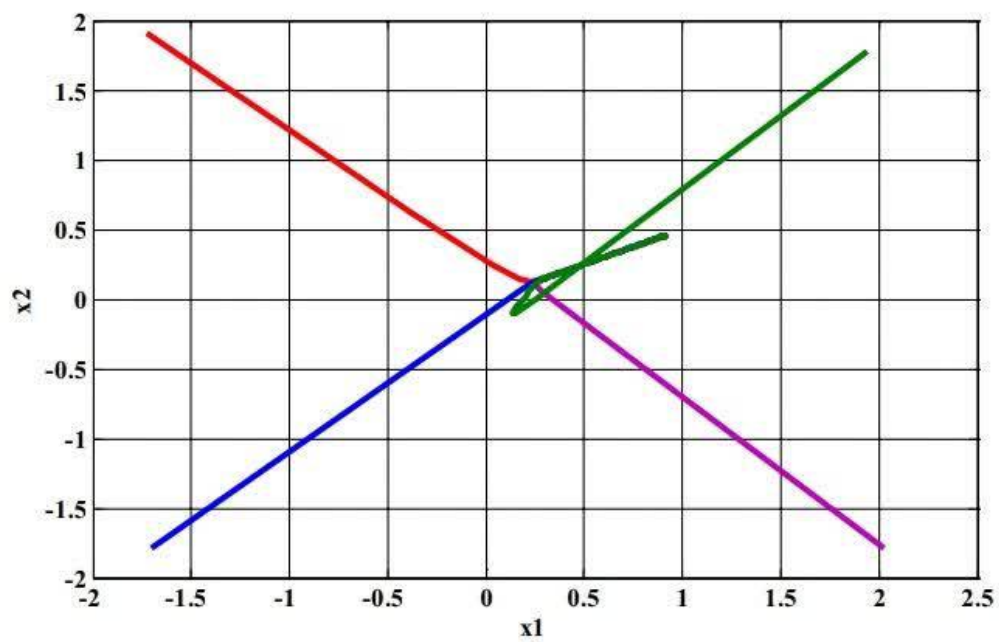
مسیرها پس از ۵۰۰ تکرار به یک جواب بهینه (۰/۱۱۰۹, ۲/۲۴۰۶, ۰/۱۸۵۱, ۰/۵۹۲۵) همگرا می‌شوند که در آن نقاط اولیه دلخواه (۶, ۶, ۶, ۶, ۰, ۰, ۰) و (۰, ۰, ۰, ۰, ۰, ۰, ۰) انتخاب شدند. بنابراین، همه اشکال نشان داده‌اند که تمام مسیرهای جواب روش پیشنهادی به جواب بهینه همگرا هستند.

جدول ۲.۳: مقادیر بهینه مسائل بهینه‌سازی محدب

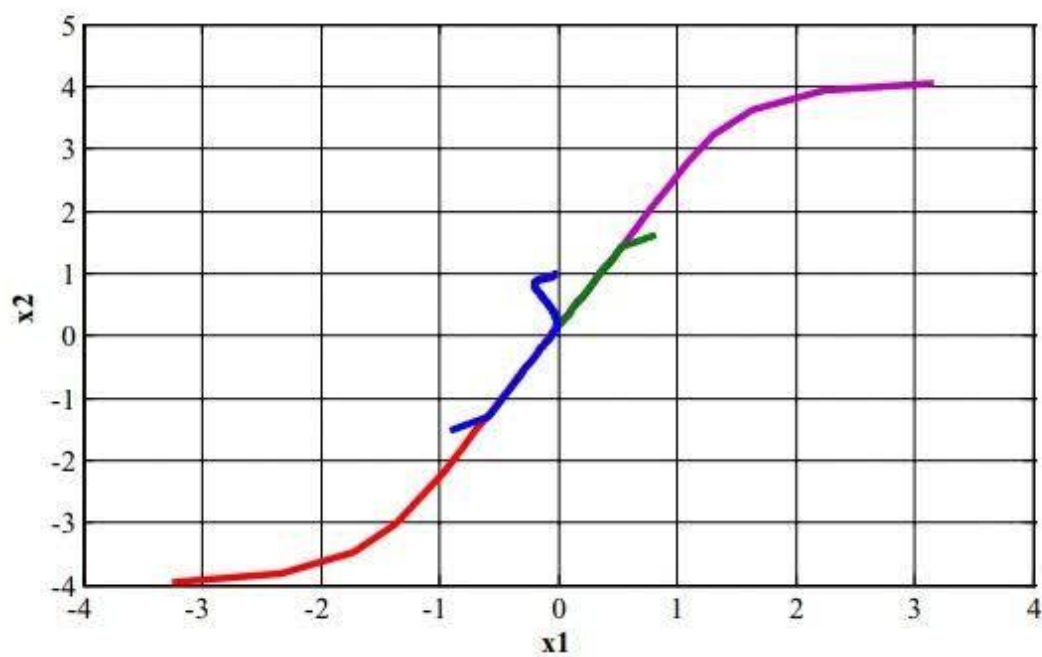
مسائل	نقاط اولیه برای بردارها	مقدار بهینه تابع هدف
(P)	$(x_1(t), x_2(t), \dots, x_n(t), \lambda_1, \lambda_2, \dots, \lambda_p)$	$(f(x^*))$
P_1	(۱, ۲, -۱)	-۲
P_2	(۰/۸۹۴۴, ۰/۴۴۷۲, -۱/۱۱۱۸, ۰)	-۲/۲۳۶۰
P_3	(۰, ۱, -۰/۵)	-۱
P_4	(۱, ۲, ۰, -۱)	-۲۱
P_5	(۰/۲, ۰/۴, -۰/۱۳۷۲, -۰/۵۳۲۱)	۰/۲۰۰۳
P_6	(۰/۱۱۰۹, ۲/۲۴۰۶, -۰/۱۸۵۱, ۰/۵۹۵۲, ۰, -۴/۲۶۲)	-۴/۹۶۷۲



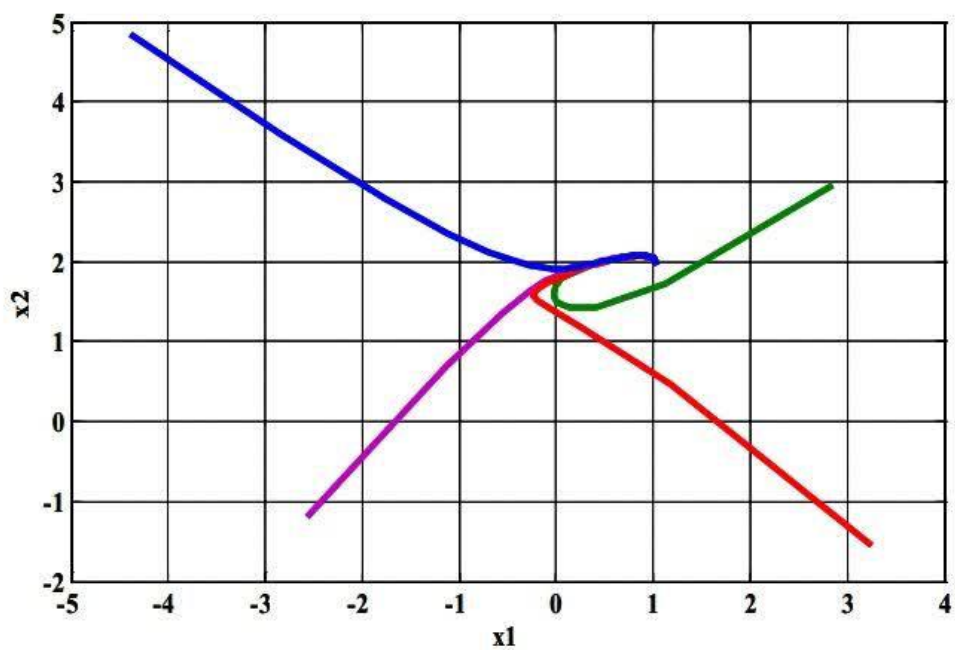
شکل ۱.۳: چهار مسیر با نقاط اولیه متفاوت برای P_1



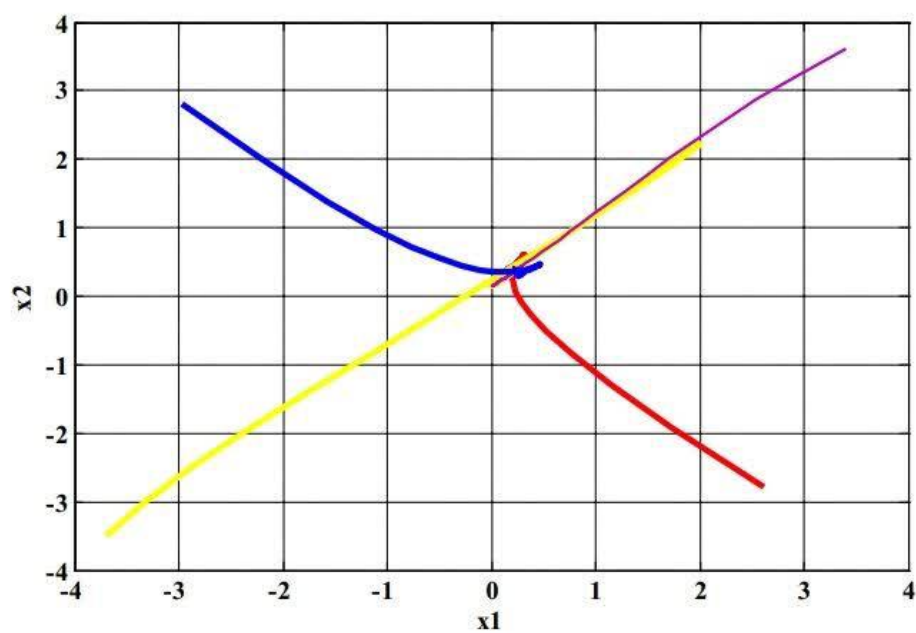
شکل ۲.۳: چهار مسیر با نقاط اولیه متفاوت برای p_2



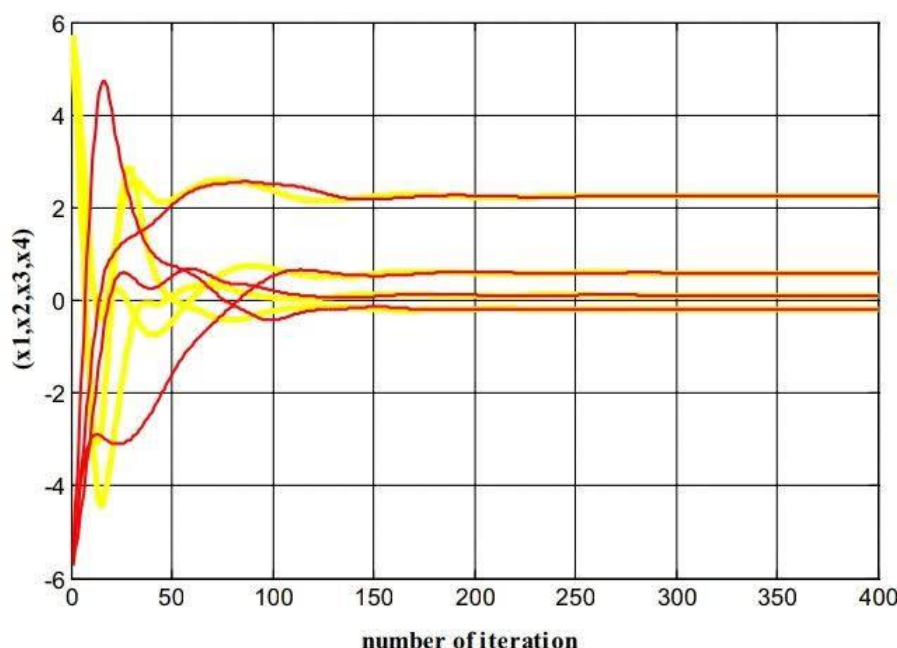
شکل ۳.۳: چهار مسیر با نقاط اولیه متفاوت برای p_3



شکل ۴.۳: چهار مسیر با نقاط اولیه متفاوت برای p_4



شکل ۵.۳: چهار مسیر با نقاط اولیه متفاوت برای p_5



شکل ۶.۳: هشت مسیر با دو نقطه اولیه برای مسئله p_6 در برابر تعداد تکرارها

۳.۳.۳ مسائل بهینه‌سازی دوهدفه

در این بخش، ما مثال‌هایی را در جدول ۶.۳ ارائه می‌دهیم تا روش شبکه عصبی پیشنهادی برای مسائل بهینه‌سازی دوهدفه را بررسی کنیم.

۴.۳.۳ تجزیه و تحلیل نتایج مسائل بهینه‌سازی دوهدفه

در مسائل بهینه‌سازی دوهدفه، به دست آوردن مرز پارتو برای فرایند تصمیم‌گیری بسیار مهم است. مرزهای پارتو تولید شده با استفاده از رویکرد شبکه عصبی پیشنهادی در شکل‌های ۷.۳ (آ)، ۷.۳ (ب)، ۷.۳ (ج)، ۷.۳ (د)، ۷.۳ (ه) و ۷.۳ (و) نشان داده شده است. در بخش قبل، وجود یک نقطه تعادل را که متناظر با جواب بهینه مسائل بهینه‌سازی غیرخطی است و پایداری مجانبی نقطه تعادل رویکرد شبکه عصبی پیشنهادی را نشان دادیم. بنابراین، برای هر نقطه اولیه دلخواه داخل یا خارج از ناحیه شدنی، روش شبکه پیشنهادی همیشه به جواب بهینه نظری همگرا می‌شود. بنابراین، همانطور که در شکل‌های ۷.۳ (آ)، ۷.۳ (ب)، ۷.۳ (ج)، ۷.۳ (د)، ۷.۳ (ه) و ۷.۳ (و) نشان داده شده است، رویکرد پیشنهادی قادر به تولید جواب‌های بهینه پارتو برای مسائل بهینه‌سازی دوهدفه می‌باشد. شکل‌های ۷.۳ (آ)، ۷.۳ (ب)، ۷.۳ (ج)، ۷.۳ (د)، ۷.۳ (ه) و ۷.۳ (و) نشان می‌دهد که ویژگی رویکرد پیشنهادی این است که با در نظر گرفتن وزن‌های مختلف جواب‌های بهینه پارتو حاصل می‌شوند به گونه‌ای که مرز پارتو به طور یکنواخت تشکیل شده است. هم‌چنین آزمایشات تجربی گسترده‌ای برای مطالعه تأثیر نقطه اولیه بر رفتار رویکرد شبکه عصبی پیشنهادی انجام شده است. نتایج نشان می‌دهد که رویکرد

جدول ۳.۳: مسائل بهینه‌سازی دوحده

مسئله (P)	نقاط اولیه برای مسیرها	اندازه گام Δt
$\begin{aligned} & \text{Min}_{\Omega} \begin{cases} f_1(x) = (x_1 - 1)^2 + (x_2 - 1)^2 \\ f_2(x) = (x_1 - 2)^2 + (x_2 - 3)^2 \end{cases} \\ & \text{subject to } x \in \Omega, \\ & \Omega = \left\{ x \in \mathbb{R}^n \mid g_1(x) = x_1 + 2x_2 - 1 \leq 0 \right. \\ & \quad \left. x_1 \geq 0, 0 \leq x_2 \leq 4 \right\} \end{aligned}$	$(x_1(t), x_2(t), \dots, x_n(t)),$ $\lambda_1, \lambda_2, \dots, \lambda_p)$ $(3, 0, 0)$ و $(3, 3, 0)$ $(0.5, 1, 0)$ و $(0, 3, 0)$	۰/۰۸
$\begin{aligned} & \text{Min}_{\Omega} \begin{cases} f_1(x) = 4(x_1)^2 + 4(x_2)^2 \\ f_2(x) = (x_1 - 5)^2 + (x_2 - 5)^2 \end{cases} \\ & \text{subject to } x \in \Omega, \\ & \Omega = \left\{ \begin{aligned} & x \in \mathbb{R}^n \mid g_1(x) = (x_1 - 5) + x_2^2 - 25 \leq 0 \\ & g_2(x) = 2x_1 - (x_1 - 8)^2 - (x_2 + 3)^2 \leq 0 \\ & 0 \leq x_1 \leq 5, 0 \leq x_2 \leq 3 \end{aligned} \right. \end{aligned}$	$(-5, 5, 0, 0)$ و $(5, -5, 0, 0)$ $(5, 5, 0, 0)$ و $(-5, -5, 0, 0)$	۰/۰۸

جدول ۴.۳: ادامه مسائل بهینه‌سازی دوحده

مسئله (P)	نقاط اولیه برای مسیرها	اندازه گام Δt
	$(x_1(t), x_2(t), \dots, x_n(t),$ $\lambda_1, \lambda_2, \dots, \lambda_p)$	
P_3 :	$(-4, 4, 0)$ و $(4, -4, 0)$	0.4
$\text{Min}_{\Omega} \begin{cases} f_1(x) = 2 + (x_1 - 2)^2 + (x_2 - 1)^2 \\ f_2(x) = 9x_1 + (x_2 - 1)^2 \end{cases}$ subject to $x \in \Omega$,	$(4, 4, 0)$ و $(-4, -4, 0)$	
$\Omega = \begin{cases} x \in \mathbb{R}^n g_1(x) = x_1^2 + x_2^2 - 225 \leq 0 \\ g_2(x) = x_1 - 2x_2 + 10 \leq 0 \\ -20 \leq x_1 \leq 20, -20 \leq x_2 \leq 20 \end{cases}$		
P_4 :		
$\text{Min}_{\Omega} \begin{cases} f_1(x) = x_1^2 \\ f_2(x) = (x_1 - 2)^2 \end{cases}$ subject to $x \in \Omega, \quad \Omega = \{x \in \mathbb{R}^n 10^{-6} \leq x_2 \leq 10^6\}$	$(-6, 6, 0, 0)$ و $(6, -6, 0, 0)$ $(6, 6, 0, 0)$ و $(-6, -6, 0, 0)$	0.4

جدول ۵.۳: ادامه مسائل بهینه‌سازی دوحده

مسئله (P)	نقاط اولیه برای مسیرها	اندازه گام Δt
$\begin{aligned} &P_{\Delta}: \\ &\text{Min}_{\Omega} \begin{cases} f_1(x) = 1 - x_1 x_2 \\ f_2(x) = \frac{1}{2}(2x_1 + x_2) \end{cases} \\ &\text{subject to } x \in \Omega, \\ &\Omega = \{x \in \mathbb{R}^n \mid 0 < x_1, x_2 < 1\} \end{aligned}$	$(x_1(t), x_2(t), \dots, x_n(t), \lambda_1, \lambda_2, \dots, \lambda_p)$ $(-4, -4, 0, 0)$ و $(-4, 4, 0, 0)$ و $(4, -4, 0, 0)$ $(4, 4, 0, 0)$	۰/۰۱
$\begin{aligned} &P_{\varphi}: \\ &\text{Min}_{\Omega} \begin{cases} f_1(x) = x_1^2 - x_2 \\ f_2(x) = -0.5x_1 - x_2 - 1 \end{cases} \\ &\text{subject to } x \in \Omega, \\ &\Omega = \left\{ \begin{aligned} &x \in \mathbb{R}^n \mid g_1(x) = \frac{x_1}{2} + x_2 - 9.5 \leq 0 \\ &g_2(x) = \frac{x_1}{4} + x_2 - 7.5 \leq 0 \\ &g_3(x) = 5x_1 + x_2 - 30 \leq 0 \\ &-7 \leq x_1 \leq 4, -7 \leq x_2 \leq 4 \end{aligned} \right\} \end{aligned}$	$(-6, -6, -6, -6, 0, 0, 0, 0, 0)$ و $(4, -4, 0, 0)$ $(4, 4, 0, 0)$ و $(6, 6, 6, 6, 0, 0, 0, 0)$	۰/۰۴

پیشنهادی پایدار است و جواب‌های بهینه پارتو همگرا تولید می‌کند. در این بخش، یک مطالعه مقایسه‌ای برای ارزیابی رویکرد پیشنهادی در تولید جواب‌های بهینه پارتو و زمان محاسبه انجام شده است. این اولین بار است که از رویکرد شبکه عصبی پیشنهادی برای بررسی مسائل بهینه‌سازی دوهدفه و هم‌چنین به کارگیری مجموعه پایداری نوع اول استفاده می‌شود. با این حال، رویکرد پیشنهادی می‌تواند به یک تصمیم‌گیرنده کمک کند تا تصمیم صحیحی را با تجسم مرز پارتو با کمک نقاط ترسیم‌شده به دست آورد. در نهایت، امکان استفاده از رویکرد پیشنهادی برای بررسی مسائل بهینه‌سازی دوهدفه به صورت تجربی تأیید شده و مباحث فوق تأیید می‌کند که رویکرد پیشنهادی برای مسائل بهینه‌سازی دوهدفه بهتر است.

جدول ۶.۳: مسائل بهینه‌سازی دوهدفه

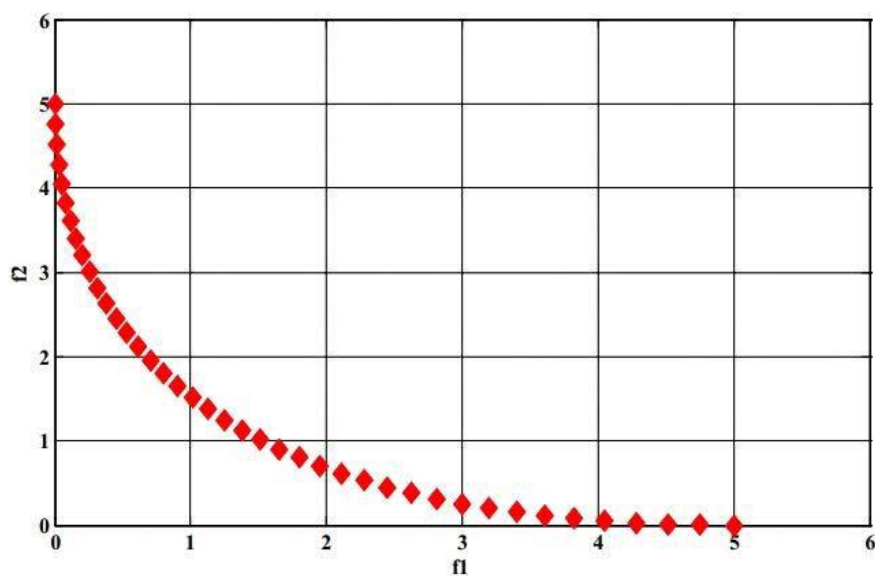
مسئله (P)	نقاط اولیه برای مسیرها	اندازه گام اندازه Δt
$P_1: \begin{cases} \text{Min}_{\Omega} \begin{cases} f_1(x) = (x_1 - 1)^2 + (x_2 - 1)^2 \\ f_2(x) = (x_1 - 2)^2 + (x_2 - 3)^2 \end{cases} \\ \text{subject to } x \in \Omega, \\ \Omega = \left\{ x \in \mathbb{R}^n \mid g_1(x) = x_1 + 2x_2 - 1 \leq 0 \right. \\ \left. x_1 \geq 0, 0 \leq x_2 \leq 4 \right\} \end{cases}$	$(x_1(t), x_2(t), \dots, x_n(t), \lambda_1, \lambda_2, \dots, \lambda_p)$	۰/۰۸
$P_2: \begin{cases} \text{Min}_{\Omega} \begin{cases} f_1(x) = 4(x_1)^2 + 4(x_2)^2 \\ f_2(x) = (x_1 - 5)^2 + (x_2 - 5)^2 \end{cases} \\ \text{subject to } x \in \Omega, \\ \Omega = \left\{ x \in \mathbb{R}^n \mid g_1(x) = (x_1 - 5) + x_2^2 - 25 \leq 0 \right. \\ \left. g_2(x) = 44 - (x_1 - 8)^2 - (x_2 + 3)^2 \leq 0 \right. \\ \left. 0 \leq x_1 \leq 5, 0 \leq x_2 \leq 3 \right\} \end{cases}$	$(5, 5, 0, 0)$ و $(-5, -5, 0, 0)$ و $(0, 3, 0)$ و $(0, 5, 1, 0)$ و $(3, 3, 0)$	۰/۰۸

جدول ۷.۳: ادامه مسائل بهینه‌سازی دوحده

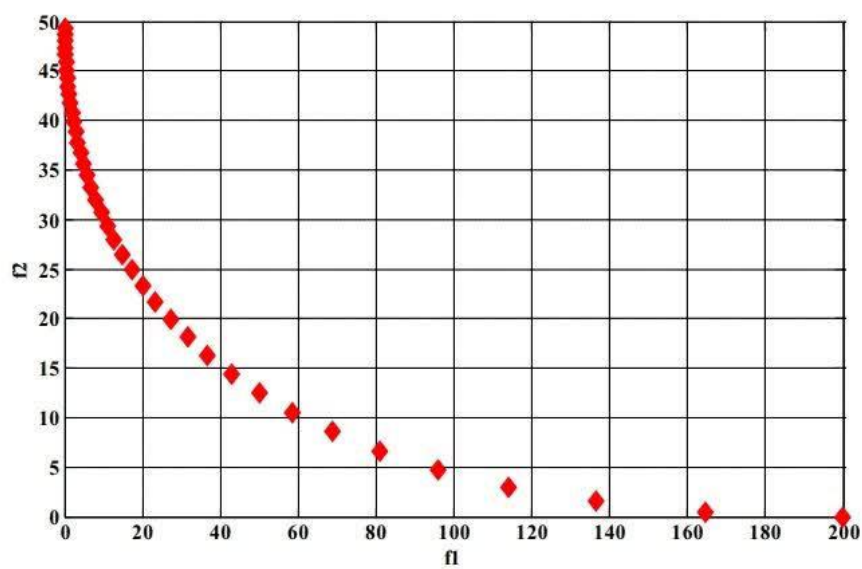
مسئله (P)	نقاط اولیه برای مسیرها	اندازه گام Δt
$P_3:$	$(x_1(t), x_2(t), \dots, x_n(t), \lambda_1, \lambda_2, \dots, \lambda_p)$	Δt
$\text{Min}_{\Omega} \begin{cases} f_1(x) = 1 + (x_1 - 2)^2 + (x_2 - 1)^2 \\ f_2(x) = 9x_1 + (x_2 - 1)^2 \end{cases}$	$(-4, 4, 0)$ و $(4, -4, 0)$	0.4
subject to $x \in \Omega,$	$(4, 4, 0)$ و $(-4, -4, 0)$	
$\Omega = \begin{cases} x \in \mathbb{R}^n g_1(x) = x_1^2 + x_2^2 - 22 \leq 0 \\ g_2(x) = x_1 - 3x_2 + 1 \leq 0 \\ -2 \leq x_1 \leq 2, -2 \leq x_2 \leq 2 \end{cases}$		
$P_4:$	$(-6, 6, 0, 0)$ و $(6, -6, 0, 0)$	0.4
$\text{Min}_{\Omega} \begin{cases} f_1(x) = x_1^2 \\ f_2(x) = (x_1 - 2)^2 \end{cases}$	$(6, 6, 0, 0)$ و $(-6, -6, 0, 0)$	
subject to $x \in \Omega, \quad \Omega = \{x \in \mathbb{R}^n 1 \leq x_2 \leq 1.5\}$		
$P_5:$	$(-4, 4, 0, 0)$ و $(4, -4, 0, 0)$	0.1
$\text{Min}_{\Omega} \begin{cases} f_1(x) = 1 - x_1 x_2 \\ f_2(x) = \frac{1}{2}(2x_1 + x_2) \end{cases}$	$(4, 4, 0, 0)$ و $(-4, -4, 0, 0)$	
subject to $x \in \Omega,$		
$\Omega = \{x \in \mathbb{R}^n 0 < x_1, x_2 < 1\}$		

جدول ۸.۳: ادامه مسائل بهینه‌سازی دوهدفه

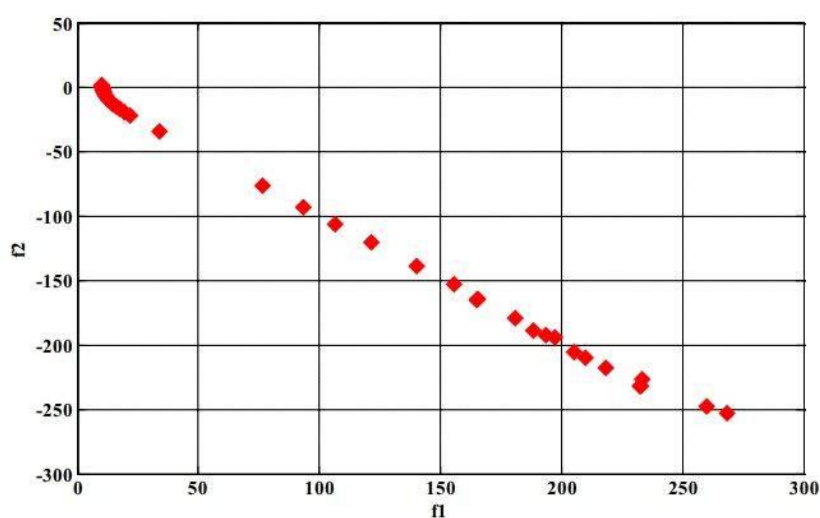
مسئله (P)	نقاط اولیه برای مسیرها	اندازه گام
	$(x_1(t), x_2(t), \dots, x_n(t), \lambda_1, \lambda_2, \dots, \lambda_p)$	اندازه Δt
P_2 :	$(-6, -6, -6, -6, 0, 0, 0, 0, 0, 0)$ و $(4, -4, 0)$	0.04
$\text{Min}_{\Omega} \begin{cases} f_1(x) = x_1^2 - x_2 \\ f_2(x) = -0.5x_1 - x_2 - 1 \end{cases}$	$(4, 4, 0)$ و $(6, 6, 6, 6, 0, 0, 0, 0)$	
subject to $x \in \Omega$,		
$\Omega = \begin{cases} x \in \mathbb{R}^n g_1(x) = \frac{x_1}{2} + x_2 - 6/5 \leq 0 \\ g_2(x) = \frac{x_1}{3} + x_2 - 7/5 \leq 0 \\ g_3(x) = 5x_1 + x_2 - 30 \leq 0 \\ -7 \leq x_1 \leq 4, -7 \leq x_2 \leq 4 \end{cases}$		



(آ) جواب‌های بهینه پارتو POS با استفاده از روش پیشنهادی NNA برای مسئله p_1

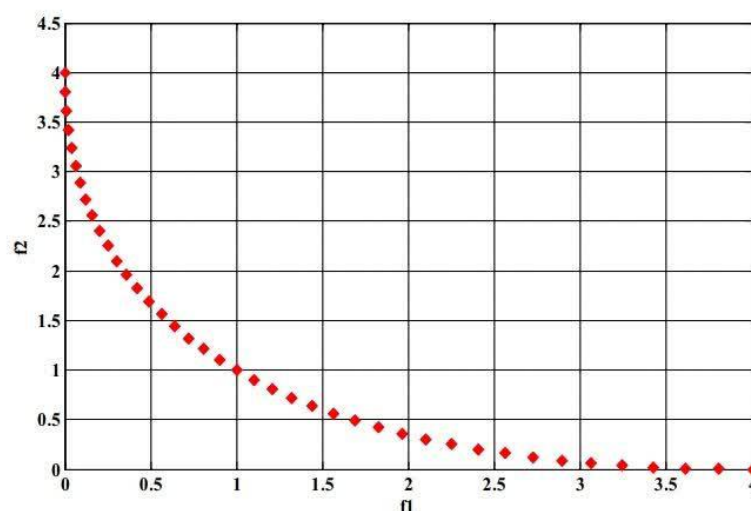


(ب) جواب‌های بهینه پارتو POS با استفاده از روش پیشنهادی NNA برای مسئله p_2



(ج) جواب‌های بهینه پارتو POS با استفاده از روش پیشنهادی NNA برای مسئله

P۳

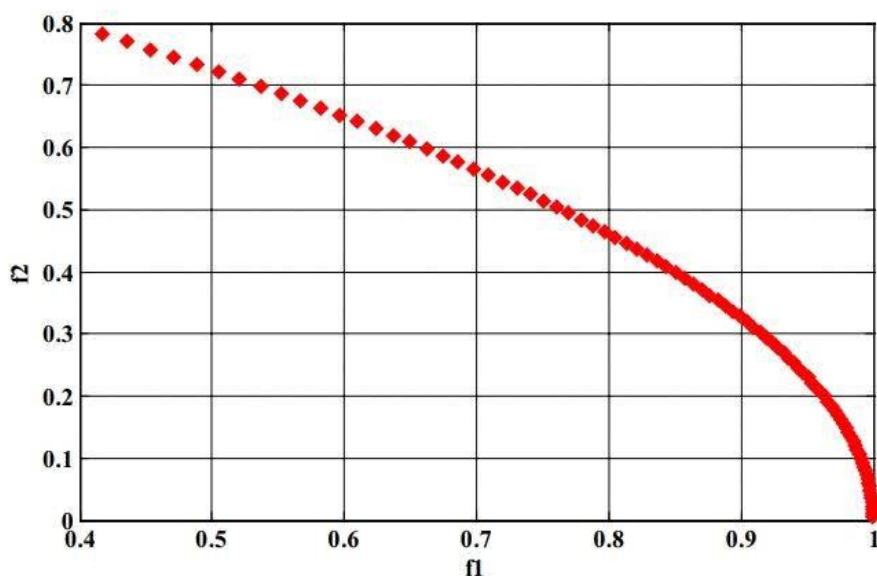


(د) جواب‌های بهینه پارتو POS با استفاده از روش پیشنهادی NNA برای مسئله

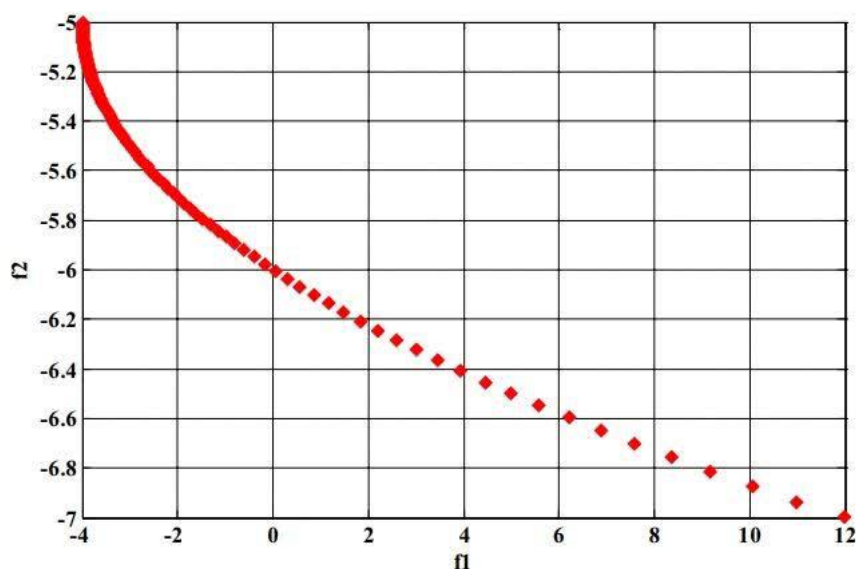
P۴

۵.۳.۳ تجزیه فضاهای پارامتری برای مسائل بهینه‌سازی دوهدفه

در این بخش یک مرحله جدید برای پایداری مسائل بهینه‌سازی دوهدفه از دیدگاه پارامتری را مورد بررسی قرار می‌دهیم. این مرحله جدید پایداری برای به‌دست‌آوردن هر زیرمجموعه‌ای از فضای پارامتریک، که دارای جواب‌های بهینه پارتوی یکسانی برای مسائل بهینه‌سازی دوهدفه است، استفاده می‌شود.



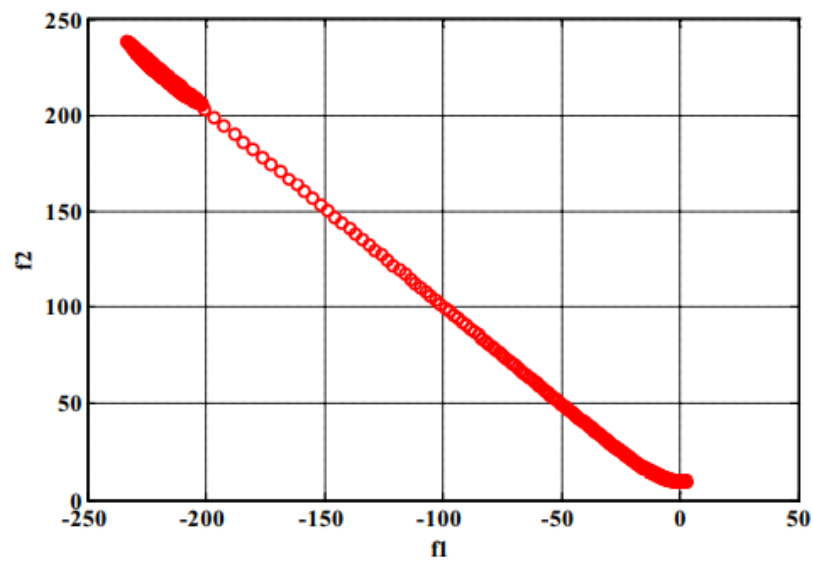
(ه) جواب‌های بهینه پارتو POS با استفاده از روش پیشنهادی NNA برای مسئله p_5



(و) جواب‌های بهینه پارتو POS با استفاده از روش پیشنهادی NNA برای مسئله p_6

این رویکرد با یک نقطه اولیه مناسب و وزن‌های اولیه برای مسئله بهینه‌سازی دوهدفه شروع می‌شود. تکرارها با اجرای $P(w)$ شروع و سپس با استفاده از رویکرد شبکه عصبی پیشنهادی $P(w)$ حل می‌شود.

با تکرار مکرر این رویکرد و کاهش گام، می‌توانیم تمام مرز پارتو را که به خوبی توزیع شده و به خوبی متنوع است، مانند شکل ۷.۳ (ز) به دست آوریم. نتایج به دست آمده ثابت می‌کند که همگرایی رویکرد پیشنهادی به جواب بهینه سریع و قوی است و مشاهده می‌شود که به جواب‌های بهینه پارتو همگرا است.



(ز) نقاط به‌دست‌آمده با استفاده از روش شبکه عصبی پیشنهادی

فصل ۴

ترکیب شرایط کان-تاکر و رویکرد شبکه عصبی برای مسائل بهینه‌سازی چندسطحی و چندهدفه

در فصل حاضر، نخست روش شبکه عصبی پیشنهادی برای حل مسائل بهینه‌سازی چندسطحی با شرایط کان-تاکر ترکیب می‌شود و برای مسائل بهینه‌سازی چندهدفه از مدل تبدیل‌شده مسائل چندهدفه بر اساس نقطه مرجع استفاده و در نهایت اهمیت نظری آن ارائه می‌گردد. اغلب مطالب این فصل از [۲۶] گرفته شده است.

۱.۴ بررسی رویکرد شبکه عصبی پیشنهادی

در فصل قبل گفتیم که برای حل مسائل بهینه‌سازی با روش پیشنهادی، گام اصلی ساختن یک تابع انرژی $E(z)$ مناسب است تا کمترین سطح انرژی متناظر با جواب بهینه z^* باشد. بر اساس تابع انرژی، یک سیستم گرادیان از معادلات دیفرانسیل ایجاد می‌کنیم تا روش شبکه عصبی را شبیه‌سازی کند.

مسئله بهینه‌سازی چندهدفه

$$\text{Min}_{x \in X} (f_1(x), f_2(x), \dots, f_K(x))$$

۶۲ ترکیب شرایط کان-تاکر و رویکرد شبکه عصبی برای مسائل بهینه‌سازی چندسطحی و چندهدفه

را در نظر بگیرید. همان‌طور که در فصل اول اشاره کردیم، از معیار L_P برای رسیدن به معیار «نزدیکی» استفاده می‌کنیم. در واقع معیار L_P فاصله بین $F(x)$ و F^+ را به صورت زیر بیان می‌کند:

$$d_p = \left(\sum_{k=1}^K w_k^p (f_k(x) - f_k^+)^p \right)^{\frac{1}{p}} \quad p = 1, 2, \dots, \infty \quad (1.4)$$

که در آن w_k نشان‌دهنده اهمیت نسبی تابع هدف k ام است. $\forall k = 1, 2, \dots, K$ و اگر توابع هدف $f_k(x)$ ، $k = 1, 2, \dots, K$ در واحدهای ثابت و معادل بیان نشوند، آنگاه معمولاً برای هر یک از توابع هدف از یک تابع مقیاس‌گذاری، در بازه‌ی $[0, 1]$ استفاده می‌شود. در این حالت، از معیار زیر می‌توان استفاده کرد:

$$d_p(x) = \left(\sum_{k=1}^K w_k^p \left(\frac{f_k(x) - f_k^+}{f_k^- - f_k^+} \right)^p \right)^{\frac{1}{p}} \quad p = 1, 2, \dots, \infty \quad (2.4)$$

که جواب‌ایده‌آل $F^+ = (f_1^+, f_2^+, \dots, f_K^+)$ همان نقطه مرجع است. مدل تبدیل‌شده‌ی مسائل بهینه‌سازی چندهدفه براساس نقطه مرجع یعنی

$$P_1 : \text{Min} \quad d_p(x) = \left(\sum_{k=1}^K w_k^p \left(\frac{f_k(x) - f_k^+}{f_k^- - f_k^+} \right)^p \right)^{\frac{1}{p}} \quad p = 1, 2, \dots, \infty \quad (3.4)$$

که در آن $x \in \Omega$ را در نظر بگیرید، مسئله بهینه‌سازی دوسطحی زیر

$$P_2 : \text{Min}_x \quad F(x, y) = c_1 x + d_1 y \\ A_1 x + B_1 y \leq b_1$$

به طوری که y جوابی از

$$\text{Min}_y \quad f(x, y) = c_2 x + d_2 y \\ A_2 x + B_2 y \leq b_2 \quad (4.4)$$

است و به طوری که در آن $A_1 \in \mathbb{R}^{p \times n}$ ، $b_1 \in \mathbb{R}^p$ ، $d_1, d_2 \in \mathbb{R}^m$ ، $c_1, c_2 \in \mathbb{R}^n$ ، $A_2 \in \mathbb{R}^{p \times n}$ ، $b_2 \in \mathbb{R}^p$ ، $B_1 \in \mathbb{R}^{p \times m}$

را با شرایط فروش-کان-تاکر به یک مسئله یک سطحی مانند

$$\begin{aligned}
 \text{P}_3 \text{Min} \quad & F(x, y) = c_1 x + d_1 y \\
 \text{s.t.} \quad & A_1 x + B_1 y \leq b_1, \\
 & A_2 x + B_2 y \leq b_2, \\
 & u B_1 - v B_2 - w = -d_2, \\
 & u(b_1 - A_1 x - B_1 y) + v(b_2 - A_2 x - B_2 y) + w y = 0, \\
 & A_2 x + B_2 y \leq b_2 \\
 & x \leq 0, y \leq 0, u \leq 0, v \leq 0, w \leq 0.
 \end{aligned} \tag{5.4}$$

در صورتی که بردارهای سطری u^* ، v^* و w^* وجود داشته باشند به طوری که $(x^*, y^*, u^*, v^*, w^*)$ جواب سیستم بالا باشند، کاهش دهید. حال با توجه به نتیجه موجود در مانگاریسن^۱ [۲۳] و فصل قبل، مسئله دوگان (DP) و مسئله اولیه (PP) (به عنوان مثال، P_1 برای مسائل بهینه سازی چندهدفه و P_3 برای مسائل بهینه سازی چندسطحی) را می توان به شرح زیر محاسبه کرد.

$$\begin{aligned}
 \text{DP:} \quad & \max_{\langle x, \lambda, \mu \rangle} L(x, \lambda, \mu) \\
 \text{s.t.} \quad & \nabla_x L(x, \lambda, \mu) = 0, \lambda \geq 0.
 \end{aligned} \tag{6.4}$$

که $\mu = (\mu_1, \mu_2, \dots, \mu_L)^T$ ، $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_L)^T$ باشد و

$$L(x, \lambda, \mu) = F(x) - \sum_{p=1}^P \lambda_p g_p(x) - \sum_{j=1}^J \mu_j h_j(x),$$

و

$$\nabla_x L(x, \lambda, \mu) = \nabla F(x) - \sum_{p=1}^P \lambda_p \nabla g_p(x) - \sum_{j=1}^J \mu_j \nabla h_j(x).$$

علاوه بر این، فرم ماتریسی $L(x, \lambda, \mu)$ و $\nabla_x L(x, \lambda, \mu)$ را می توان به صورت زیر نوشت:

$$L(x, \lambda, \mu) = F(x) - \lambda^T g(x) - \mu^T h(x), \quad \nabla_x L(x, \lambda, \mu) = \nabla F(x) - \nabla g(x)^T \lambda - A^T \mu.$$

و $g(x)$ و $\nabla g(x)^T$ را هم می توان به صورت زیر تعریف کرد:

$$g(x) = (g_1(x), g_2(x), \dots, g_n(x))^T, \quad h(x) = (h_1(x), h_2(x), \dots, h_n(x))^T,$$

$$\nabla g(x)^T = [\nabla g_1(x), \nabla g_2(x), \dots, \nabla g_n(x)], \quad z = (x^T, \lambda^T, \mu^T)^T.$$

$$\nabla g(\mathbf{x}) = \begin{bmatrix} \frac{\partial g_1(\mathbf{x})}{\partial \mathbf{x}_1} & \frac{\partial g_1(\mathbf{x})}{\partial \mathbf{x}_2} & \dots & \frac{\partial g_1(\mathbf{x})}{\partial \mathbf{x}_n} \\ \frac{\partial g_2(\mathbf{x})}{\partial \mathbf{x}_1} & \frac{\partial g_2(\mathbf{x})}{\partial \mathbf{x}_2} & \dots & \frac{\partial g_2(\mathbf{x})}{\partial \mathbf{x}_n} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial g_L(\mathbf{x})}{\partial \mathbf{x}_1} & \frac{\partial g_L(\mathbf{x})}{\partial \mathbf{x}_2} & \dots & \frac{\partial g_L(\mathbf{x})}{\partial \mathbf{x}_n} \end{bmatrix}_{(L \times n)}, \quad ; A = \begin{Bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_J^T \end{Bmatrix}_{(J \times n)}.$$

نتایج زیر را از فصل قبل بازنویسی می‌کنیم.

قضیه ۱.۱.۴ [۲۳] فرض کنید که $\mathbf{x}^* \in \mathbb{R}^n$ یک جواب بهینه برای مسئله اولیه باشد و آنگاه $(\lambda^{*T}, \mu^{*T})^T \in \mathbb{R}^{L+J}$ وجود دارد به‌طوری‌که یک جواب بهینه‌ی مسئله دوگان است.

قضیه ۲.۱.۴ [۲۳] جواب‌های بهینه $\mathbf{x}^* \in \mathbb{R}^n$ و $(\mathbf{x}^{*T}, \lambda^{*T}, \mu^{*T})^T$ جواب‌های مسئله اولیه و دوگان هستند، اگر فقط اگر

$$\begin{cases} F(\mathbf{x}^*) = L(\mathbf{x}^*, \lambda^*, \mu^*), & \nabla_{\mathbf{x}} L(\mathbf{x}^*, \lambda^*, \mu^*) = \circ, \\ g(\mathbf{x}^*) \geq \circ, & h(\mathbf{x}^*) = \circ, \quad \lambda^* \geq \circ \end{cases}$$

یا

$$\begin{cases} \lambda^{*T} g(\mathbf{x}^*) = \circ, & \nabla_{\mathbf{x}} L(\mathbf{x}^*, \lambda^*, \mu^*) = \circ, \\ g(\mathbf{x}^*) \geq \circ, & A\mathbf{x}^* - b = \circ, \quad \lambda^* \geq \circ \end{cases} \quad (۷.۴)$$

$$.b = (b_1, b_2, \dots, b_J)^T \text{ که}$$

با کمک قضیه‌های ۲.۲.۳ و ۲.۱.۴، می‌توان تابع انرژی مسئله اولیه را به صورت زیر ساخت:

$$\begin{aligned} E(z) = E(\mathbf{x}, \lambda, \mu) = & \frac{1}{2} [\lambda^T g(\mathbf{x})]^2 + \frac{1}{2} \|\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda, \mu)\|^2 + \frac{1}{2} g(\mathbf{x})^T [g(\mathbf{x}) - |g(\mathbf{x})|] \\ & + \frac{1}{2} \|A\mathbf{x} - b\|^2 + \frac{1}{2} \lambda^T (\lambda - |\lambda|) \end{aligned} \quad (۸.۴)$$

که

$z = (\mathbf{x}^T, \lambda^T, \mu^T) \in \mathbb{R}^{n+L+J}$ است. صفر بودن هر جمله سمت راست معادله (۸.۴) متناظر با هر تساوی یا نامساوی معادله (۲.۲.۳) است و در $E(z) \geq \circ$ صدق می‌کند. بنابراین $E(z)$ طبق نظریه لیانگ و همکاران^۱ [۲۲]، تابعی مشتق‌پذیر است.

قضیه ۳.۱.۴ [۲۳] $z^* = (\mathbf{x}^{*T}, \lambda^{*T}, \mu^{*T})$ یک جواب صفر برای $E(z)$ است اگر و فقط اگر z^* جواب بهینه مسائل اولیه و دوگان باشد. (به‌عنوان مثال $\mathbf{x}^*$ و $(\mathbf{x}^{*T}, \lambda^{*T}, \mu^{*T})^T$ جواب‌های بهینه مسائل اولیه و دوگان هستند).

با استفاده از ایده اصلاح شده در قضیه لیانگ و همکاران [۲۳]، می‌توانیم از سیستم گرادیان برای ایجاد روش شبکه عصبی جدید برای حل مسائل بهینه‌سازی چندهدفه و مسائل بهینه‌سازی چندسطحی استفاده کنیم:

$$\frac{dz}{dt} = -\nabla E(z) \quad (9.4)$$

با فرمول‌های محاسباتی گرادیان در نظریه لیانگ و همکاران [۲۲]، رویکرد جدید شبکه عصبی در معادله (۹.۴) را می‌توان با شرح جزئیات زیر نوشت:

$$\begin{aligned} \frac{dx}{dt} &= -\nabla_x E(z) = -\lambda^T g(x) \cdot \nabla g(x)^T \lambda - \nabla g(x)^T - \nabla g(x)^T [g(x) - |g(x)|] \\ &\quad - \nabla_{xx} L(z) \nabla_x L(z) - A^T (Ax - b) \\ \frac{d\lambda}{dt} &= -\nabla_\lambda E(z) = -\lambda^T g(x) \cdot g(x) + \nabla g(x) \nabla_x L(z) - (\lambda - |\lambda|) \\ \frac{d\mu}{dt} &= -\nabla_\mu E(z) = A \nabla_x L(z) \end{aligned} \quad (10.4)$$

که در آن $\nabla_{xx}^2 L(z)$ ماتریس هسین^۱ تابع $L(z)$ نسبت به x است. فرض کنید $\nabla E(z)$ پیوسته لپ‌شیتز^۲ باشد، آنگاه مسئله مقدار اولیه معادلات دیفرانسیل در معادله (۹.۴) یک جواب منحصر به فرد دارد از آنجایی که تابع سمت راست معادلات دیفرانسیل (۹.۴) یا (۱۰.۴) پیوسته هستند، این معادلات را می‌توان با اجرای سخت‌افزار شبکه عصبی به راحتی به دست آورد. بنابراین، این یک شبکه عصبی شدنی است.

قضیه ۴.۱.۴. [۲۲] مجموعه‌ای از نقاط تعادل معادله (۸.۴) را به صورت

$$M = \{z = (x^T, \lambda^T, \mu^T)^T \in \mathbb{R}^{n+L+J} \mid \nabla E(z) = 0\}$$

مساائل اولیه و دوگان هستند، آنگاه $M \subset \Theta$ است. (یعنی هر جواب بهینه مسائل اولیه و دوگان، نقطه تعادل روش جدید شبکه عصبی در معادله (۸.۴) است.)

□

برهان. مشابه با اثبات قضیه ۵.۲.۳، این قضیه هم ثابت می‌شود.

قضیه ۵.۱.۴. [۳۶] فرض کنید شبکه عصبی (۹.۴) دارای نقطه تعادل منحصر به فرد z^* است، آنگاه z^* یکنواخت و پایدار مجانبی است.

قضیه زیر نشان می‌دهد که تحت برخی شرایط، مسیر به نقطه تعادل (۹.۴) همگرا می‌شود.

قضیه ۶.۱.۴. [۳۶] فرض کنید یک منحنی تراز $L(z^0) = \{z \in \mathbb{R}^{n+L+J} : E(z) \leq E(z^0)\}$ کران‌دار باشد، آنگاه یک نقطه تعادل $\bar{z} \in M$ و یک دنباله اکید صعودی $\{t_n\}$ ($t_n \geq 0$) وجود دارد به طوری که $\lim_{n \rightarrow \infty} z(t_n, z^0) = \bar{z}$ و $\nabla E(\bar{z}) = 0$.

¹Hessian

²Lipschitz

قضیه زیر نشان می‌دهد که شبکه عصبی (۱۹.۳) پایدارمجبانی و سراسری است.

قضیه ۷.۱.۴. فرض کنید شبکه عصبی (۱۹.۳) دارای یک نقطه تعادل منحصربه‌فرد و منحنی تراز $L(z^\circ) = \{z \in \mathbb{R}^{n+L+J} : E(z) \leq E(z^\circ)\}$ کران‌دار است، آنگاه هر مسیر معادله شبکه معادله (۱۸.۳) به جواب بهینه z^* همگرا می‌شود. یعنی جواب بهینه z^* پایدارمجبانی و سراسری و برای مسائل بهینه‌سازی چندهدفه و چندسطحی است.

برهان. طبق قضیه‌های ۶.۲.۳ و ۷.۲.۳ این قضیه بدیهی است. $\square$

از روش اوایلر^۱ برای حل معادلات دیفرانسیل غیرخطی توصیف‌شده، با معادله (۱۹.۳) استفاده شده‌است. الگوریتم شبکه عصبی پیشنهادی این فصل همانند الگوریتم ۲ می‌باشد.

۲.۴ مثال‌های عددی

در این بخش چند مثال برای تأیید شدنی بودن و کارایی رویکرد شبکه عصبی پیشنهادی ارائه می‌دهیم. دو نمونه مختلف از مسائل بهینه‌سازی، یعنی مسائل بهینه‌سازی چندهدفه و مسائل بهینه‌سازی چندسطحی را بررسی می‌کنیم. روش پیشنهادی در نرم افزار MATLAB اجرا شده‌است.

مثال ۱.۲.۴. مسئله برنامه ریزی دوهدفه زیر را در نظر بگیرید.

$$\text{Max } f_1(x) = 2x_1 - x_2$$

$$\text{Max } f_2(x) = x_1 + 2x_2$$

$$\text{subject to } 3x_1 - 5x_2 \leq 15$$

$$3x_1 - x_2 \leq 21$$

$$3x_1 + x_2 \leq 27$$

$$3x_1 + 4x_2 \leq 45$$

$$x_1 + 3x_2 \leq 30$$

$$x_1 \geq 0, x_2 \geq 0.$$

ابتدا نقطه ایده‌آل $F^+ = (13/5, 21)$ و نقطه ضدایده‌آل $F^- = (-10, 0)$ را برای مثال ۱.۲.۴ به دست می‌آوریم. سپس مثال ۱.۲.۴ به مسئله‌ای مشابه معادله (۱۲.۱) با $p = 1, 2$ کاهش می‌یابد.

$$\text{Min } d_p \text{ subject to } x \in \Omega = \left\{ \left(\frac{13/5 - f_1(x)}{13/5 + 10} \right)^P + \left(\frac{21 - f_2(x)}{21 - 0} \right)^P \right\}^{\frac{1}{p}}$$

¹Euler

که در آن، Ω برای نشان دادن مجموعه محدودیت استفاده می‌شود. برای $p = \infty$ فرمول را می‌توان به صورت زیر بیان کرد

$$\begin{aligned} & \text{Min } d_p \\ & \text{subject to } x \in \Omega \\ & \frac{13/5 - f_1(x)}{13/5 + 1^0} \leq d_p, \\ & \frac{21 - f_2(x)}{21 - 0} \leq d_p. \end{aligned}$$

رویکرد شبکه عصبی پیشنهادی برای یافتن جواب توافقی برای مسائل بهینه‌سازی چندهدفه با استفاده از سه تعریف مختلف فاصله با پارامترهای $p = 1$ ، $p = 2$ و $p = \infty$ استفاده می‌شود. جواب بهینه مثال ۱.۲.۴ با الگوریتم پیشنهادی و مقایسه آن با روشهای مختلف [۲۹، ۲۱] در جدول ۱.۴ نشان داده شده‌است. جدول ۱.۴ تعداد تکرارها، مرحله، بردار تصمیم، بردار هدف و حداقل فاصله را با استفاده از سه مورد برای فاصله $p = 1$ ، $p = 2$ و $p = \infty$ نشان می‌دهد. جواب‌های به‌دست‌آمده برای این سه مورد بهتر از جواب‌های [۲۹، ۲۱] می‌باشد. اگرچه سه جواب مختلف با استفاده از سه فرمول مختلف به‌دست آمده‌است، اما همه آن‌ها در مرز غیرمغلوب واقع شده‌اند.

علاوه بر این، شکل ۱.۴ نشان می‌دهد که برای مجموعه‌ای از نقاط اولیه دلخواه، مسیرهای رویکرد شبکه عصبی پیشنهادی به جواب بهینه مثال ۱.۲.۴ که مربوط به بهترین جواب توافقی برای مسائل بهینه‌سازی چندهدفه با وزن‌های معادل است، همگرا می‌باشد. شکل ۱.۴ (آ) رفتار متغیرها را برای فاصله منتهن^۱ نشان می‌دهد (یعنی $p = 1$)، شکل ۱.۴ (ب) رفتار متغیرها را برای فاصله اقلیدسی^۲ (یعنی $p = 2$) نشان می‌دهد و فاصله چبی‌شف^۳ (یعنی $p = \infty$) در شکل ۱.۴ (ج) نشان داده شده‌است.

علاوه بر این، همگرایی مسیرهای جواب با متغیرهای تصمیم در مقابل تعداد تکرارها در شکل ۲.۴ برای $p = 1$ ، $p = 2$ و $p = \infty$ نشان داده شده‌است.

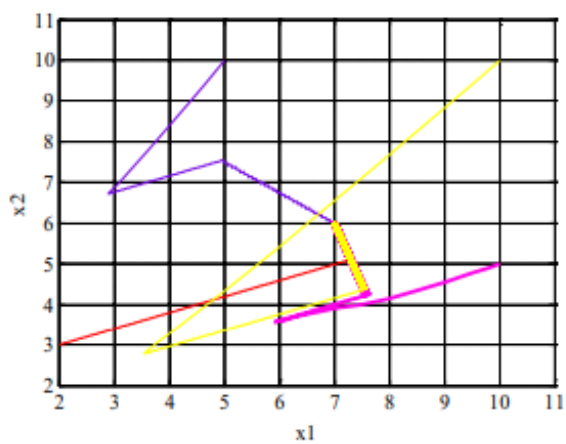
¹Manhattan

²Euclidian

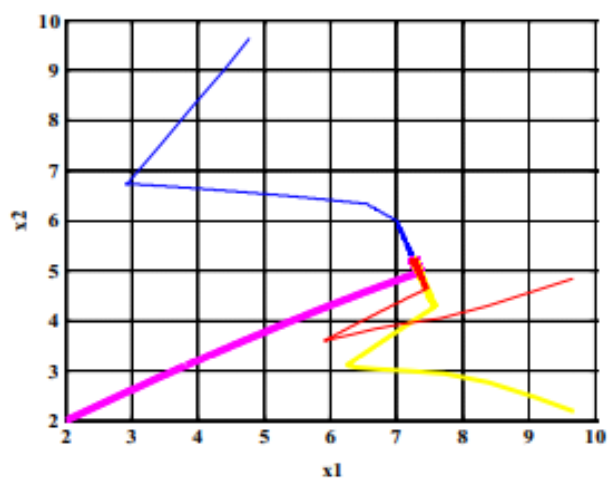
³Tchebycheff

جدول ۱۰۴: مقایسه حداقل فاصله با استفاده از روش پیشنهادی با روش‌های ارائه شده در منابع [۲۹، ۲۱]

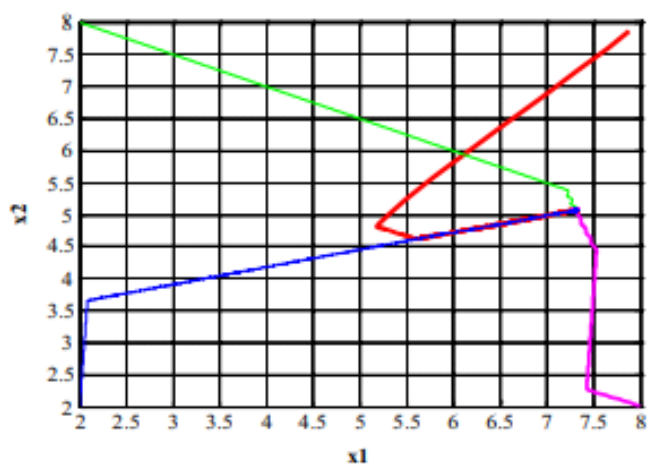
جواب بهینه	تکرار	Δt	x_1^*	x_2^*	f_1	f_2	d_p
$p = 1$ جواب بهینه با راه حل پیشنهادی برای $p = 1$	۵۰۰۰	۰/۰۰۱	۷	۶	۸	۱۹	۰/۳۲۹۳
$p = 2$ جواب بهینه با راه حل پیشنهادی برای $p = 2$	۵۰۰۰	۰/۰۰۱	۷/۳۰۳۸	۵/۱۸۸۸	۹/۴۱۸۸	۱۷/۶۸۱۴	۰/۲۳۴۸
$p = \infty$ جواب بهینه با راه حل پیشنهادی برای $p = \infty$	۲۰۰۰۰	۰/۰۰۱	۷/۳۲۴۶	۵/۰۸۰۳	۹/۵۶۸۹	۱۷/۴۸۵۲	۰/۱۶۷۴
جواب [۲۱] برای $p = 1$	-	-	۷	۶	۸	۱۹	۰/۳۲۹۳
جواب [۲۱] برای $p = 2$	-	-	۷/۲۷	۵/۲	۹/۳۳	۱۷/۶۷	۰/۳۳۵۶
جواب [۲۱] برای $p = \infty$	-	-	۷/۳۱	۵/۰۸	۹/۵۴	۱۷/۴۷	۰/۳۳۶۶
جواب [۲۹] برای $p = \infty$	-	-	۷/۳۰۷۳	۵/۰۷۸۲	۹/۵۳۶۴	۱۷/۴۵۲۹	۰/۱۶۸۵۲۶



(I)

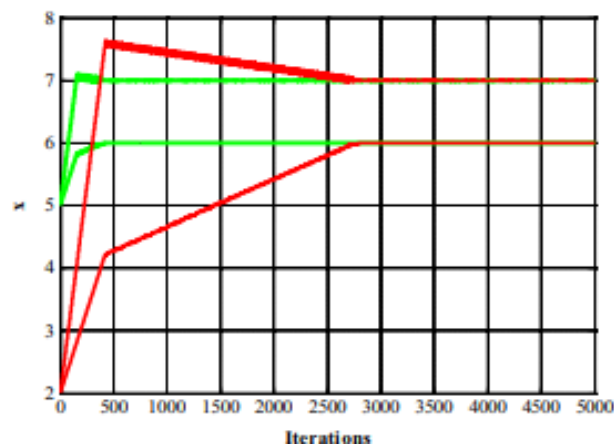


(ب)

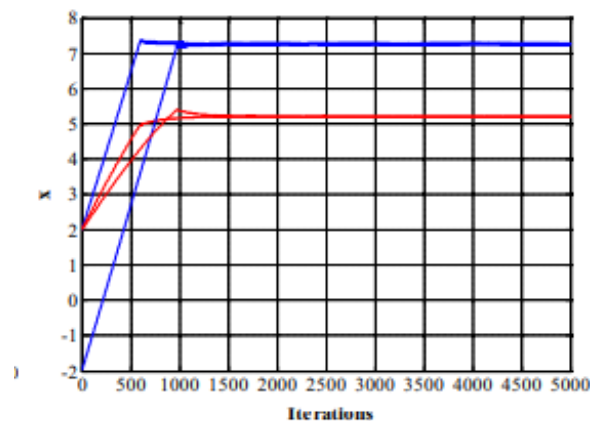


(ج)

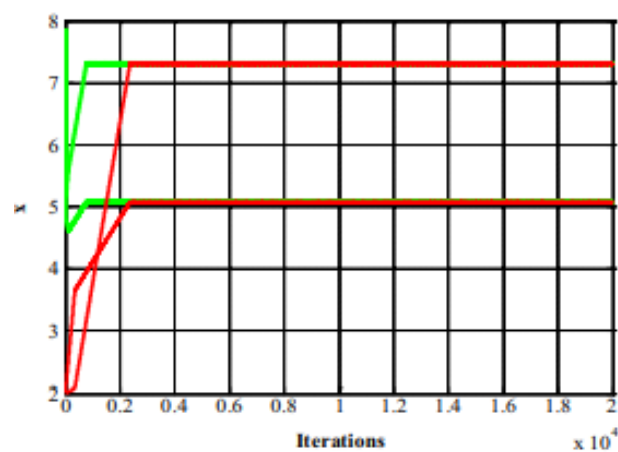
شکل ۱.۴: مسیرهایی با نقاط مختلف برای مثال ۱.۲.۴ (I) مسیرهای جواب با $p = 1$ (ب) مسیرهای جواب با $p = 2$ (ج) مسیرهای جواب با $p = \infty$



(ا)



(ب)



(ج)

شکل ۲.۴: مسیرهایی با دو نقطه اولیه مختلف برای $p = 1$, $p = 2$, و $p = \infty$ دو مسیر با نقاط اولیه متفاوت برای $p = 1$ (ب) دو مسیر با نقاط اولیه متفاوت برای $p = 2$ (ج) دو مسیر با نقاط اولیه متفاوت برای $p = \infty$

به این ترتیب، شبکه عصبی پیشنهادی طبق نظریه لیاپانوف پایدار است و می‌تواند به یک جواب دقیق همگرا شود.

مثال ۲.۲.۴. یک مسئله برنامه‌ریزی سه‌هدفه به صورت زیر در نظر بگیرید.

$$\text{Max } f_1(x) = 2x_1 + 5x_2 + 5x_3 - 2x_4 + 5x_5$$

$$\text{Max } f_2(x) = -x_1 - 2x_2 + 4x_3 - x_4$$

$$\text{Max } f_3(x) = 5x_1 + 3x_2 + 2x_3 + x_4 - x_5$$

$$\text{subject to } 7x_1 + 2x_2 + 6x_3 \leq 28$$

$$3x_1 + 4x_2 \leq 23$$

$$4x_1 + 4x_2 + x_3 \leq 23$$

$$3x_1 + 4x_2 \leq 45$$

$$x_1 + 6x_2 + 7x_3 + 4x_4 - 6 \leq 23$$

$$2x_1 + 5x_2 + 5x_3 + 5x_4 - 4 + 8x_5 \leq 29$$

$$x_1, x_2, x_3, x_4, x_5, x_6 \geq 0.$$

مثال ۲.۲.۴ نشان‌دهنده مسئله‌ای سه‌هدفه با شش متغیر تصمیم‌گیری است که برای به‌دست‌آوردن بهترین جواب نقطه تعادل به مسئله تک‌هدفه با وزن‌های برابر تبدیل می‌شود. برای حل این مثال کافی است نقطه ایده‌آل و نقطه ضد ایده‌آل را به دست آوریم، که به ترتیب برابر با $F^+ = (33/1, 14/5, 39/25)$ و $F^- = (-7/25, -16/412, -9/207)$ هستند. برای $p = 1, 2$ فرمول را می‌توان به صورت زیر بیان کرد:

$$\text{Min } d_p \text{ subject to } x \in \Omega = \left\{ \left(\frac{33/1 - f_1(x)}{33/1 + 7/25} \right)^P + \left(\frac{14/5 - f_2(x)}{14/5 + 16/412} \right)^P + \left(\frac{39/25 - f_3(x)}{39/25 + 9/207} \right)^P \right\}^{\frac{1}{P}}$$

برای $p = \infty$ ، فرمول را می‌توان به صورت زیر بیان کرد

$$\text{Min } d_p$$

$$\text{subject to } x \in \Omega$$

$$\frac{31/5 - f_1(x)}{33/1 + 7/25} \leq d_p,$$

$$\frac{14/5 - f_2(x)}{14/5 + 16/412} \leq d_p,$$

$$\frac{39/25 - f_3(x)}{39/25 + 9/207} \leq d_p.$$

جدول ۲.۴ بهترین جواب توافقی به‌دست‌آمده از رویکرد شبکه عصبی پیشنهادی را نشان می‌دهد که با جواب به‌دست‌آمده توسط ام.سان و همکاران^۱ [۳۰]، قابل مقایسه است. جدول ۲.۴ نشان می‌دهد که جواب به‌دست‌آمده با رویکرد شبکه عصبی پیشنهادی با استفاده از فرمول‌بندی‌های مختلف فاصله، بهتر از جواب به‌دست‌آمده با [۳۰]، برای تمام موارد فاصله است.

علاوه بر این، مسیرهای جواب با متغیرهای تصمیم و بر حسب تعداد تکرارها در شکل ۳.۴ نشان داده شده است. شکل ۳.۴ (ج) رفتار رویکرد شبکه عصبی پیشنهادی را با در نظر گرفتن دو نقطه اولیه مختلف برای $p = 1$ نشان می‌دهد. همچنین رفتار رویکرد شبکه عصبی پیشنهادی برای $p = 2$ در شکل ۳.۴ (ب) نشان داده شده است. به علاوه فاصله چپی شف برای $p = \infty$ در شکل ۳.۴ (ج) نشان داده شده است.

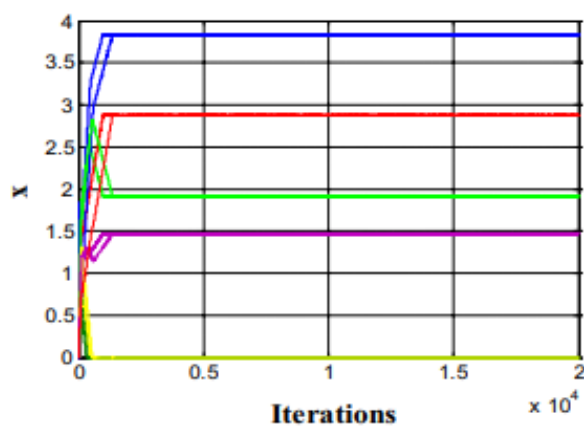
در این بخش، یک مطالعه مقایسه‌ای برای ارزیابی رویکرد پیشنهادی در مورد بهترین جواب نقطه تعادل و همگرایی انجام شده است. از طرف دیگر، تاکنون از رویکرد شبکه عصبی سنتی برای یافتن جواب توافقی استفاده نشده است. بنابراین، این اولین بار است که از رویکرد شبکه عصبی پیشنهادی برای بررسی مسائل بهینه‌سازی چندهدفه استفاده می‌شود. علاوه بر این، شذنی بودن رویکرد پیشنهادی برای مسائل مقیاس بزرگ، به ویژه برای مثال ۲.۲.۴ با پارامتر $p = 2$ (فاصله اقلیدسی) بررسی شده است. همچنین، الگوریتم پیشنهادی برای یک نقطه اولیه دلخواه دارای قابلیت همگرایی سریع است. مسیر رویکرد شبکه عصبی پیشنهادی چه در مسیرهای جواب و چه بر حسب تکرارها به بهترین نقطه تعادل همگرا می‌شود.

علاوه بر مقایسه با روش‌های مختلف مانند جدول ۱.۴ و جدول ۲.۴، از معیارهای آماری (جدول ۳.۴ را ببینید) مانند بهترین، میانگین، متوسط و بدترین مقادیر، و همچنین انحراف معیار و میانگین زمان به‌دست‌آمده از ۳۰ بار تکرار برای یک آزمایش مشابه اجرا می‌شود.

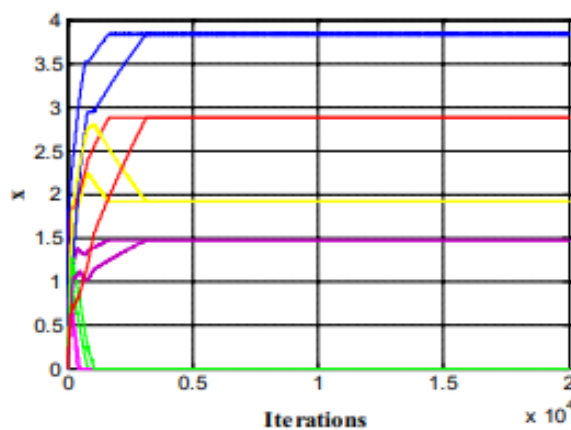
¹M. Sun et al.

جدول ۲.۴: مقایسه مینیم فاصله با روش پیشنهادی و آنچه در [۳۰] ارائه شده است.

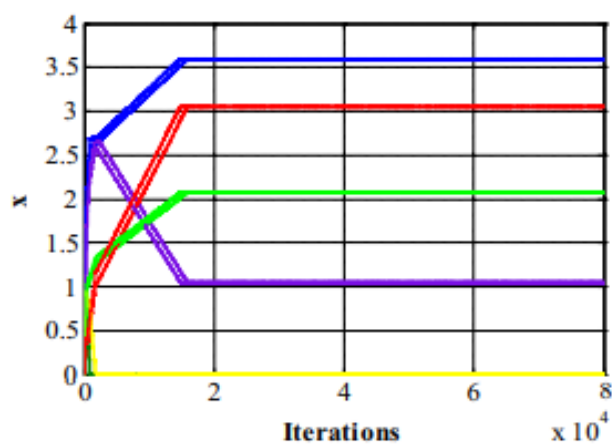
d_p	f_3	f_2	f_1	x_6	x_5	x_4	x_3	x_2	x_1	جواب‌های بهینه
۱/۲۶۵۴	۱۸/۹۵۶	-۰/۸۷۸	۱۶/۵۱۳	۲/۳۳۱	۱/۵۰۳	۱/۵۷۲	۰	۰	۴/۵۵۸	جواب بهینه [۳۰]
۱/۲۲۵۷	۲۲/۳۹۵	-۳/۱۷۸	۱۸/۷۱۹	۱/۹۲۳	۰/۹۶۶	۲/۰۷۶	۰/۱۲۸	۰/۰۰۹	۵/۱۰۳	جواب بهینه [۳۰] برای $p = 1$
۱/۱۵۹۲	۲۱/۵۸۲۸	۰/۸۲۵۸	۲۱/۰۴۲	۲/۸۷۹۴	۱/۴۷۱۲	۰	۱/۹۱۷۵	۰	۳/۸۳۱۲	راه حل پیشنهادی برای $p = 1$
۰/۶۸۴۲	۲۱/۵۹۵۲	-۰/۸۲۸۳	۲۱/۰۲۶۹	۲/۸۷۶۲	۱/۴۷۰۳	۰	۱/۹۱۷۳	۰	۳/۸۳۳۳	راه حل پیشنهادی برای $p = 2$
۰/۴۱۷۲	۱۹/۱۰۰۲	۱/۶۰۲۶	۱۶/۵۰۵۳	۳/۰۵۹۵	۲/۰۶۵۷	۰/۰۵۶۱	۱/۰۵۶۱	۰/۰۰۳۱	۳/۵۹۴۵	راه حل پیشنهادی برای $p = \infty$



(ا)



(ب)



(ج)

شکل ۳.۴: رفتار همگرایی مسیرها از طریق تکرارها (ا) مقایسه دو مسیر تکرار برای $p = 1$ (ب) مقایسه دو مسیر تکرار برای $p = 2$ (ج) مقایسه دو مسیر تکرار برای $p = \infty$

۳.۴ آزمون رتبه‌ی علامت‌دار ویلکاکسون

آزمون رتبه‌ی علامت‌دار ویلکاکسون^۱ یک روش غیرپارامتری است که در وضعیت آزمون فرض شامل یک طرح با دو نمونه است [۹]. این یک آزمون زوجی است که هدف آن یافتن تفاوت‌های معنادار بین رفتارهای دو روش است. احتمال درست بودن فرض صفر به مقدار ρ مربوط است. نتیجه آزمون در $\rho < 0.05$ برگشت داده می‌شود، که رد فرض صفر را نشان می‌دهد، درحالی‌که $\rho > 0.05$ نشان‌دهنده عدم موفقیت در رد فرض صفر است. $\mathbb{R}^+$ مجموع رتبه‌های مثبت است، درحالی‌که $\mathbb{R}^-$ مجموع رتبه‌های منفی است. در جدول ۴.۴، ما نتایج آزمون رتبه علامت‌دار ویلکاکسون را برای معیارهای مختلف فاصله ارائه می‌دهیم که در آن d_1 با d_2 مقایسه می‌شود، d_∞ با d_1 و d_2 مقایسه می‌شود. ما می‌توانیم از جدول ۳.۴ و ۴.۴ نتیجه بگیریم که d_∞ پیشنهادی یک متر بامعنا است و از متر فاصله بهتر می‌باشد.

جدول ۳.۴: نتایج آماری روش شبکه عصبی پیشنهادی برای فواصل مختلف

مثال ۲			مثال ۱			
d_∞	d_2	d_1	d_∞	d_2	d_1	
۰/۴۱۷۲	۰/۶۸۴۲	۱/۱۵۹۲	۰/۱۶۷۴	۰/۲۳۴۸	۰/۳۲۹۲	بهترین
۰/۴۱۷۲	۰/۶۸۴۳	۱/۱۵۹۵	۰/۱۶۷۴	۰/۲۳۵۳	۰/۳۲۹۴	بدترین
۰/۴۱۷۲	۰/۶۸۴۲۲	۱/۱۵۹۲۷۶	۰/۱۶۷۴	۰/۲۳۴۸۷	۰/۳۲۹۳	میانگین
۰/۴۱۷۲	۰/۶۸۴۲	۱/۱۵۹۲۵	۰/۱۶۷۴	۰/۲۳۴۸	۰/۳۲۹۳	میانه
۰/۰۰۰۰	۴/۰۶۸۳E - ۵	۸/۹۷۶۳E - ۵	۰/۰۰۰۰	۱/۳۶۸۳E - ۴	۰/۰۰۰۰	St.dev.
۱/۱۰۱۵	۰/۶۹۳۴	۰/۸۶۵۷	۱/۰۹۱۲	۰/۹۸۶۷	۰/۶۵۴۸	تعداد

پس از اعتبارسنجی الگوریتم پیشنهادی برای مسائل چندهدفه، ما می‌خواهیم الگوریتم پیشنهادی را برای حل مسئله برنامه‌ریزی چندسطحی نشان‌دهیم. برای نشان‌دادن روش، دو مثال یکی مسئله برنامه‌ریزی دوسطحی و دیگری مسئله برنامه‌ریزی غیرمتمرکز که با دوسطح حل می‌شود، بررسی شده‌است.

¹Wilcoxon

جدول ۴.۴: آزمون مقایسه‌ای رتبه‌ی علامت‌دار ویلکاکسون برای مقایسه نتایج فاصله‌های مختلف

مثال	روش‌های مقایسه‌ای		ارزیابی جواب‌ها		بهترین روش
	روش اول	روش دوم	R^-	R^+	مقدار ρ
مثال ۱	d_2	d_1	۴۶۵	۰	۰/۰۰۰۰۰۲
	d_∞	d_1	۴۶۵	۰	۰/۰۰۰۰۰۲
	d_∞	d_2	۴۶۵	۰	۰/۰۰۰۰۰۲
	d_2	d_1	۴۶۵	۰	۰/۰۰۰۰۰۲
مثال ۱	d_∞	d_1	۴۶۵	۰	۰/۰۰۰۰۰۲
	d_∞	d_2	۴۶۵	۰	۰/۰۰۰۰۰۲
	d_∞	d_2	۴۶۵	۰	۰/۰۰۰۰۰۲

مثال ۱.۳.۴. مسئله برنامه‌ریزی دوسطحی زیر را در نظر بگیرید

$$\text{Max } f_1(x) = -2x_1 + 11x_2$$

$$\text{Max } f_2(x) = -x_1 - 3x_2$$

$$\text{subject to } x_1 - 2x_2 \leq 4$$

$$x_1 - x_2 \leq 24$$

$$3x_1 + 4x_2 \leq 96$$

$$x_1 + 7x_2 \leq 126$$

$$-4x_1 + 5x_2 \leq 65$$

$$x_1 + 4x_2 \geq 8$$

$$x_1 \geq 0, x_2 \geq 0.$$

با استفاده از شرایط KKT، مسئله اصلی بالا به شکل زیر تبدیل می‌شود.

$$\text{Max } f_1(x) = -2x_1 + 11x_2$$

$$\text{subject to } x \in \Omega$$

$$x_1 - 2x_2 - 4 \leq M\xi_1, u_1 \leq (1 - \xi_1)M,$$

$$x_1 - x_2 - 24 \leq M\xi_2, u_2 \leq (1 - \xi_2)M,$$

$$3x_1 + 4x_2 - 96 \leq M\xi_3, u_3 \leq (1 - \xi_3)M,$$

$$x_1 + 7x_2 - 126 \leq M\xi_4, u_4 \leq (1 - \xi_4)M,$$

$$-4x_1 + 5x_2 - 65 \leq M\xi_5, u_5 \leq (1 - \xi_5)M,$$

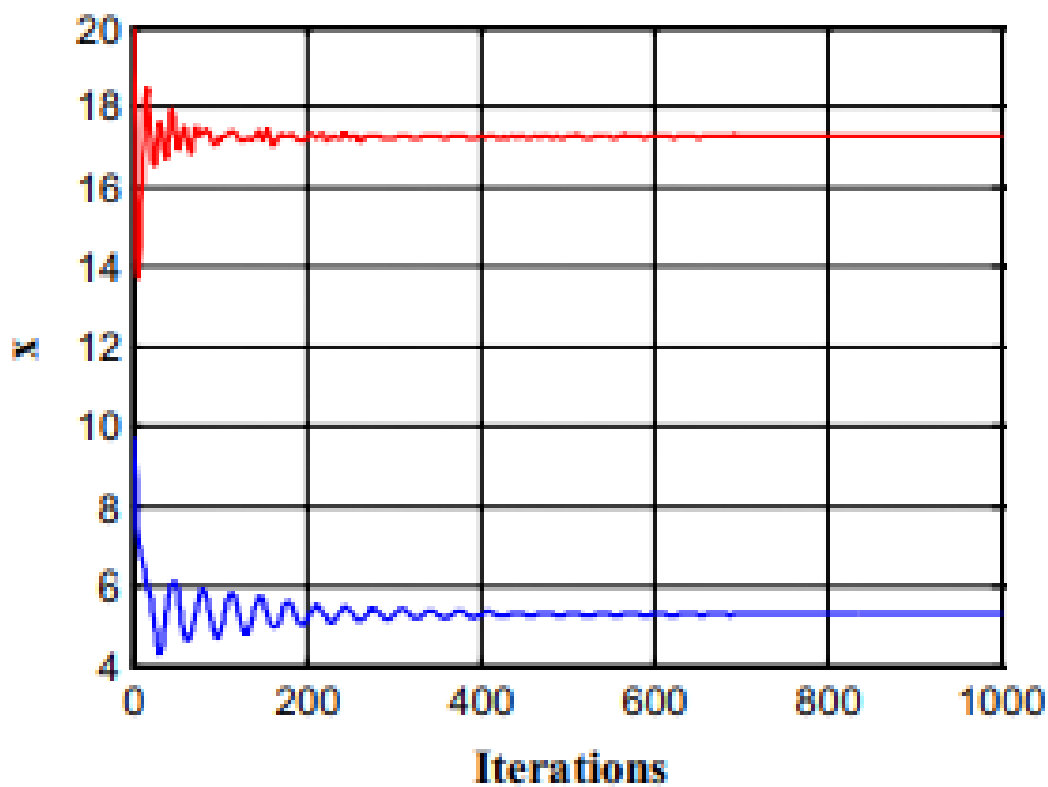
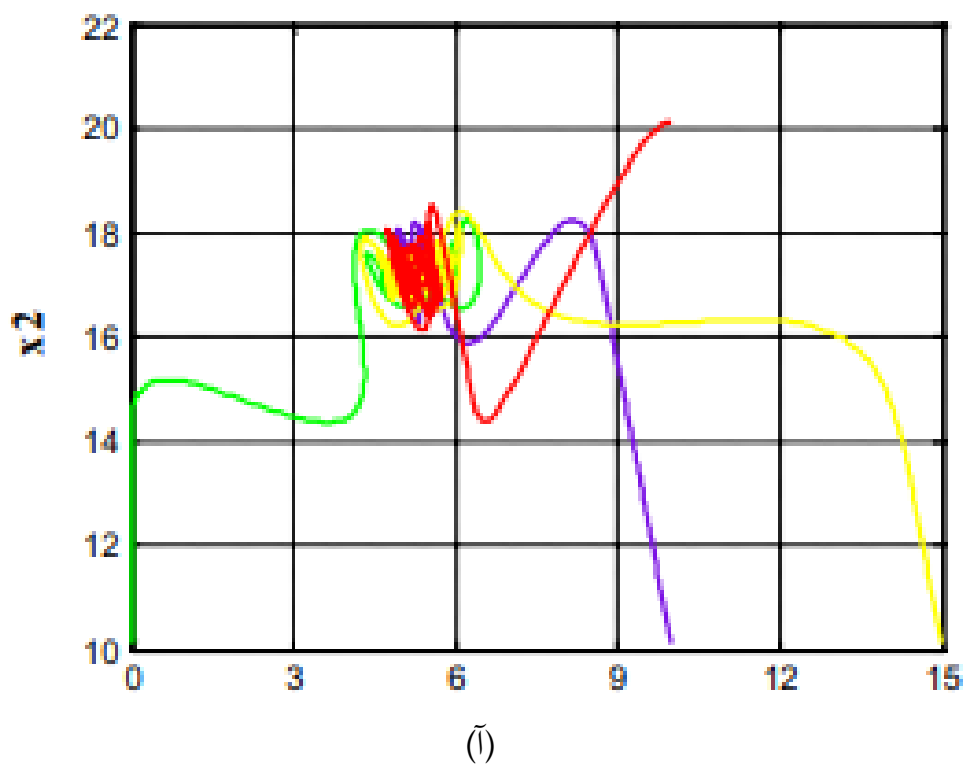
$$-x_1 - 4x_2 + 8 \leq M\xi_6, u_6 \leq (1 - \xi_6)M,$$

$$-2u_1 - u_2 + 4u_3 + 7u_4 + 5u_5 - 4u_6 = -3,$$

$$\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6 \in \{0, 1\},$$

$$u_1, u_2, u_3, u_4, u_5, u_6 \geq 0.$$

با استفاده از شرایط KKT، مسئله چندسطحی به یک مسئله تک‌هدفه تبدیل می‌شود. برای مسائل چندسطحی، با استفاده از روش مشابه مسائل چندهدفه از روش پیشنهادی شبکه عصبی استفاده می‌کنیم. هم‌چنین جواب بهینه را برای $M = 1000$ حساب می‌کنیم. نتیجه مثال ۱.۳.۴ در شکل ۴.۴ (آ) ارائه شده‌است. به طور تصادفی چهار نقطه اولیه مختلف را که در شکل ۴.۴ (آ) نشان داده شده‌است، در نظر می‌گیریم.



شکل ۴۰۴: رفتار مسیرها برای مثال ۱.۳.۴ (آ) مسیرهای جواب برای نقاط اولیه مختلف (ب) مسیرها در تعداد دفعات

مثال ۲.۳.۴. سیستم غیرمتمرکز دوسطحی با یک مرکز در سطح بالا و سه بخش مستقل در سطح پایین را در نظر بگیرید.

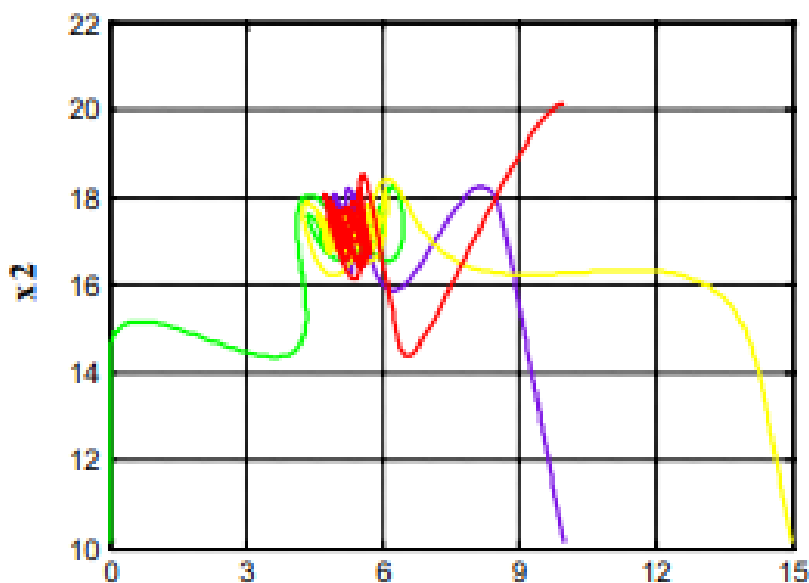
$$\begin{aligned} \text{Max } f_1(x) &= +x_1 + x_2 + 2x_3 + x_4 \\ \text{Max } f_{21}(x) &= x_1 + 3x_2 + x_3 + x_4 \\ \text{Max } f_{22}(x) &= x_1 + x_2 + 3x_3 + x_4 \\ \text{Max } f_{23}(x) &= x_1 + x_2 + x_3 + 3x_4 \\ \text{subject to } & 3x_1 + 3x_2 \leq 30 \\ & 2x_1 + x_2 \leq 20 \\ & x_1 + x_2 \leq 15 \\ & x_1 + 2x_2 + 2x_3 + x_4 \leq 40 \\ & x_1 \leq 10, x_4 \leq 10 \\ & x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

پس از اعمال شرایط KKT برای مسئله اصلی، یک مسئله کمکی جدید فقط با یک سطح به شرح زیر به دست می‌آید

$$\begin{aligned} \text{Max } f_1(x) &= x_1 + x_2 + 2x_3 + x_4 \\ \text{subject to } & x \in \Omega \\ & 3x_1 + 3x_2 - 30 \leq M\xi_1, u_1 \leq (1 - \xi_1)M, \\ & 2x_1 + x_2 - 20 \leq M\xi_2, u_2 \leq (1 - \xi_2)M, \\ & x_3 - 10 \leq M\xi_3, u_3 \leq (1 - \xi_3)M, \\ & x_3 + x_4 - 15 \leq M\xi_4, u_4 \leq (1 - \xi_4)M, \\ & x_4 - 10 \leq M\xi_5, u_5 \leq (1 - \xi_5)M, \\ & x_1 + 2x_2 + 2x_3 + x_4 - 40 \leq M\xi_6, u_6 \leq (1 - \xi_6)M, \\ & 3u_1 + u_2 + 2u_6 = 3, u_3 + u_4 + 2u_6 = 3, \\ & u_4 + u_5 + u_6 = 3 \\ & \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6 \in \{0, 1\}, \\ & u_1, u_2, u_3, u_4, u_5, u_6 \geq 0. \end{aligned}$$

با پیروی از همین روش برای مثال ۲.۳.۴، جواب به دست آمده با روش پیشنهادی برابر با $f_1^* = 35$ و مقدار تابع هدف برابر است با $(x_1^*, x_2^*, x_3^*, x_4^*) = (5, 3859, 46141, 10, 5)$

است. جواب به‌دست‌آمده از جواب کیوئن-مینگ $(f_{21}^*, f_{22}^*, f_{23}^*) = (34/2282, 45, 35)$ لان‌وهمکاران^۱ بهتر است [۱۹] و نزدیک به جواب بهینه $(x_1^*, x_2^*, x_3^*, x_4^*) = (5, 5, 10, 5)$ با مقدار تابع هدف $(f_{21}^*, f_{22}^*, f_{23}^*, f_{24}^*) = (35, 35, 45, 35)$ است. همچنین مسیر جواب x^* با تعداد تکرارهای مثال ۲.۳.۴ در شکل ۵.۴ نشان داده شده است. شکل ۵.۴ نشان می‌دهد که مسیر خروجی x^* روش پیشنهادی به جواب بهینه $(x_1^*, x_2^*, x_3^*, x_4^*) = (5/3859, 4/6141, 10, 5)$ همگرا می‌شود.



شکل ۵.۴: مسیرهای تکرار برای مثال ۲.۳.۴

مراجع

- [1] Abo-Sinna, M. and Baky, I. Interactive balance space approach for solving multi-level multi-objective programming problems. *Information Sciences*, 177(16):3397–3410, 2007.
- [2] Abo-Sinna, M. and Hussein, M. An algorithm for decomposing the parametric space in multiobjective dynamic programming problems. *European Journal of Operational Research*, 73(3):532–538, 1994.
- [3] Abo-Sinna, M. and Hussein, M. An algorithm for generating efficient solutions of multiobjective dynamic programming problems. *European Journal of Operational Research*, 80(1):156–165, 1995.
- [4] Abo-Sinna, Mahmoud A and Rizk-Allah, Rizk M. Decomposition of parametric space for bi-objective optimization problem using neural network approach. *Opsearch*, 55(2):502–531, 2018.
- [5] Arjmandzadeh, Z., Safi, M., and Nazemi, A. A new neural network model for solving random interval linear programming problems. *Neural Networks*, 89:11–18, 2017.
- [6] Bard, J. Practical bilevel optimization,. the netherlands, 1998.
- [7] Bouzerdoum, A. and Pattison, T. Neural network for quadratic optimization with bound constraints. *IEEE Transactions on Neural Networks*, 4(2):293–304, 1993.
- [8] Cheng, L., Hou, Z., Tan, M., and Wang, X. A simplified recurrent neural network for solving nonlinear variational inequalities. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pages 104–109. IEEE, 2008.
- [9] Derrac, J., García, S., Molina, D., and Herrera, F. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and

- swarm intelligence algorithms. *Swarm and Evolutionary Computation*, 1(1):3–18, 2011.
- [10] Ehrgott, M. *Multicriteria Optimization*, volume 491. Springer Science & Business Media, 2005.
- [11] Fliege, J. and Vicente, L. Multicriteria approach to bilevel optimization. *Journal of Optimization Theory and Applications*, 131(2):209–225, 2006.
- [12] Fortuny-Amat, J. and McCarl, B. A representation and economic interpretation of a two-level programming problem. *Journal of The Operational Research Society*, 32(9):783–792, 1981.
- [13] Gao, X., Liao, L., and Xue, W. A neural network for a class of convex quadratic min-max problems with constraints. *IEEE Transactions on Neural Networks*, 15(3):622–628, 2004.
- [14] Goudarzi, H. Multicriteria approach to bilevel optimization. *M.Sc Amirkabir University of Technology*, 2008.
- [15] Hopfield, J. and Tank, D. “neural” computation of decisions in optimization problems. *Biological Cybernetics*, 52(3):141–152, 1985.
- [16] Huang, W., Yan, C., Wang, J., and Wang, W. A time-delay neural network for solving time-dependent shortest path problem. *Neural Networks*, 90:21–28, 2017.
- [17] Karbasi, D., Nazemi, A., and Rabiei, M. A parametric recurrent neural network scheme for solving a class of fuzzy regression models with some real-world applications. *Soft Computing*, 24:11159–11187, 2020.
- [18] Kennedy, M. and Chua, L. Neural networks for nonlinear programming. *IEEE Transactions on Circuits and Systems*, 35(5):554–562, 1988.
- [19] Lan, K., Wen, U., Shih, H., and Lee, E. A hybrid neural network approach to bilevel programming problems. *Applied Mathematics Letters*, 20(8):880–884, 2007.
- [20] LaSalle, J. The stability of dynamical systems, regional conference series in applied mathematics, siam, philadelphia, 1976. *Khalid Hattaf Department of Mathematics and Computer Science, Faculty of Sciences Ben M’sik, Hassan II University, PO Box, 7955*, 2012.

- [21] Lee, E. and Shih, H. *Fuzzy and multi-level decision making: an interactive computational approach*. Springer Science & Business Media, 2012.
- [22] Leung, Y., Chen, K., Jiao, Y., Gao, X., and Leung, K. A new gradient-based neural network for solving linear and quadratic programming problems. *IEEE Transactions on Neural Networks*, 12(5):1074–1083, 2001.
- [23] Mangasarian, O. „nonlinear programming “mcgraw-hill book company. *New York*, 1969.
- [24] Osman, M. Qualitative analysis of basic notions in parametric convex programming. ii. parameters in the objective function. *Aplikace Matematiky*, 22(5):333–348, 1977.
- [25] Rizk-Allah, R. and Abo-Sinna, M. Integrating reference point, kuhn–tucker conditions and neural network approach for multi-objective and multi-level programming problems. *Opsearch*, 54(4):663–683, 2017.
- [26] Rizk-Allah, RM and Abo-Sinna, Mahmoud A. Integrating reference point, kuhn–tucker conditions and neural network approach for multi-objective and multi-level programming problems. *Opsearch*, 54(4):663–683, 2017.
- [27] Rodriguez-Vazquez, A., Dominguez-Castro, R., Rueda, A., Huertas, J., and Sanchez-Sinencio, E. Nonlinear switched capacitor’neural’networks for optimization problems. *IEEE Transactions on Circuits and Systems*, 37(3):384–398, 1990.
- [28] Shi, C., Lu, J., and Zhang, G. An extended kuhn–tucker approach for linear bilevel programming. *Applied Mathematics and Computation*, 162(1):51–63, 2005.
- [29] Shih, H., Wen, U., Lee, S., Lan, K., and Hsiao, H. A neural network approach to multiobjective and multilevel programming problems. *Computers & Mathematics with Applications*, 48(1-2):95–108, 2004.
- [30] Sun, M., Stam, A., and Steuer, R. Interactive multiple objective programming using tchebycheff programs and artificial neural networks. *Computers & Operations Research*, 27(7-8):601–620, 2000.
- [31] Tao, Q., Cao, J., Xue, M., and Qiao, H. A high performance neural network for solving nonlinear programming problems with hybrid constraints. *Physics Letters A*, 288(2):88–94, 2001.

- [32] Wang, J. Recurrent neural network for solving quadratic programming problems with equality constraints. *Electronics Letters*, 28(14):1345–1347, 1992.
- [33] Wang, J. Primal and dual neural networks for shortest-path routing. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 28(6):864–869, 1998.
- [34] Wang, J. and Wang, J. Forecasting stochastic neural network based on financial empirical mode decomposition. *Neural Networks*, 90:8–20, 2017.
- [35] Wen, U. and Hsu, S. Linear bi-level programming problems—a review. *Journal of the Operational Research Society*, 42(2):125–133, 1991.
- [36] Xia, J. The foundational theorems of liapunov stability. *Acta Math. Appl. Sin.*, 11(2):249–252, 1988.
- [37] Xia, Y. A new neural network for solving linear and quadratic programming problems. *IEEE Transactions on Neural Networks*, 7(6):1544–1548, 1996.
- [38] Yadav, V., Karmakar, S., Kalbar, P., and Dikshit, A. Pytops: A python based tool for topos. *SoftwareX*, 9:217–222, 2019.

Aabstract

In this paper, a new gradient-based neural network approach is proposed for solving nonlinear programming problems (NLPPs), bi-objective optimization problems (BOOPs), multi-objective programming problem (MOPP) and multi-level programming problem.

The main idea is to convert the NLPPs and the BOOPs into an equivalent optimization problem by the mean of the weighted sum method, the MOPP reference point, the MLPP into a to single level problem by the Kuhn–Tucker conditions. the Pareto optimal solutions are obtained by using different weights. Also the decomposition of parametric space for BOOP is analyzed in details based on the stability set of the first kind.

A neural network approach is constructed for solving the obtained convex programming problem. Based on employing Lyapunov theory, the proposed neural network approach is stable in the sense of Lyapunov and it is globally convergent to an exact optimal solution of the MOPP and the MLPP. The most prominent feature of the proposed approach is that it can converge rapidly to the equilibrium point (optimal solution), for an arbitrary initial point. The simulation results also demonstrate that the proposed neural network is feasible and efficient.

Keywords: Bi-objective optimization problem, Non-linear programming problem, Multi-objective programming, Multi-level programming, Karush–Kuhn–Tucker conditions, Pareto-optimal solutions, Neural networks.



Faculty Of Mathematical Sciences

MSc Thesis in Optimization

A neural network approach for solving convex multiobjective optimization problems

By: Zahra Karimi

Supervisors:

Dr. Alireza Nazemi
Dr. Mehrdad Ghaznavi

February 2022