دانشگاه صنعتی شاهرود
دانشکده ریاضی
گروه ریاضی کاربردی

یک کاربرد از مدلهای شبکه‌های عصبی برای حل مسائل بهینه‌سازی سبد سرمایه

نگارش

نرگس طهماسبی

استاد راهنما

جناب آقای دکتر ناظمی

پایان‌نامه جهت اخذ درجه کارشناسی ارشد

شهریور ۱۳۹۱

چکیده

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

در چند دهه اخیر روش‌های بهینه‌سازی که بر پایه رویکرد هوش مصنوعی توسعه یافته اند، موفقیت‌های چشم‌گیری در حل مؤثر و کارای مسائل بهینه‌سازی به‌دست آورده‌اند. روش‌هایی چون الگوریتم ژنتیک، جستجوی ممنوع، شبیه‌سازی تبریدی، شبکه عصبی و ... قابلیت‌های خود را در حل مسائل بزرگ عملی به‌خوبی نشان داده‌اند. امتیازات ویژه‌ی موجود در شبکه‌های عصبی امکان کاربرد آنها را در حوزه وسیعی از تحقیقات فراهم ساخته است. از جمله آن امتیازات می‌توان به امکان یادگیری و بهبود عملکرد براساس داده‌های ورودی اشاره کرد. همچنین امکان انجام محاسبات به‌صورت موازی در شبکه‌های عصبی امتیاز دیگری است که با توجه به گسترش سخت افزارهای موازی، امکان حل مسائل بسیار بزرگ را توسط این رویکرد ممکن می‌سازد.

در این پایان نامه دو مدل مختلف شبکه عصبی بازگشتی برای حل رده‌ای از مسائل بهینه‌سازی ارائه می‌شود. تحلیل وجود یکتایی، پایداری و همگرایی سراسری جواب‌ها مورد بررسی قرار می‌گیرند و عملکرد روش‌های ارائه شده با به‌کارگیری چند مثال از مسائل برنامه‌ریزی تصادفی، برنامه‌ریزی کسری، بهینه سازی مقاوم و بهینه سازی سبد سرمایه نشان داده می‌شود.

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

واژه‌های کلیدی: شبکه‌های عصبی، پایداری، مدل میانگین-واریانس، برنامه‌ریزی مخروط مرتبه دوم، برنامه‌ریزی تصادفی، برنامه‌ریزی کسری، بهینه‌سازی مقاوم.

فهرست مطالب

لیست جداول

لیست تصاویر

فصل ۱

مقدمه و مفاهیم اولیه

۱.۱ مقدمه

در این فصل تعاریف مورد نیاز در فصل‌های بعدی به‌طور مختصر آورده شده است. ابتدا تعاریف مرتبط با تابع محدب [۱] را بیان نموده و در خصوص مسائل بهینه‌سازی مقید و نامقید به ذکر شرایط لازم و کافی بهینگی [۲] می‌پردازیم. سپس مفاهیم پایداری [۳] و تابع انرژی در سیستم‌های دینامیکی را ارائه می‌دهیم. در ادامه مسائل برنامه‌ریزی مخروط مرتبه دوم را معرفی کرده، و در پایان نیز برخی مفاهیم آماری و اقتصادی را بیان خواهیم کرد.

۲.۱ توابع محدب

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

تعریف ۱.۲.۱. مجموعه $E \subset \mathbb{R}^n$ را محدب می‌گوییم اگر:

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

تعریف ۲.۲.۱. تابع $f(x) = f(x_۱, x_۲, ..., x_n) \in \mathbb{R}$ بر روی یک مجموعه نقاط واقع در یک مجموعه محدب $E \subset \mathbb{R}^n$ یک تابع محدب نامیده می‌شود اگر:

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

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

تعریف ۳.۲.۱. در مسأله (۱.۱) ، اگر $\bar{x} \in X$ و $\varepsilon-$همسایگی از $\bar{x}$ مثل $N_\varepsilon(\bar{x})$ موجود باشد که به ازای هر $x \in X \cap N_\varepsilon(\bar{x})$ داشته باشیم $f(\bar{x}) \leq f(x)$ آنگاه $\bar{x}$ یک کمینه موضعی برای مسأله (۱.۱) است.

[۱] Convex
[۲] Optimality
[۳] Stability

تعریف ۴.۲.۱. تابع $\|.\| : \mathbb{R}^n \to \mathbb{R}$ را تابع نرم گوییم هرگاه

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

و برای بردار $x \in \mathbb{R}$ داریم:

$$\|x\|_۱ = |x_۱| + ... + |x_n|$$

$$\|x\|_۲ = \sqrt{x^T x} = \sqrt{{x_۱}^۲ + ... + {x_n}^۲}$$

$$\|x\|_r = \sqrt[r]{{x_۱}^r + ... + {x_n}^r}, \quad r \geq ۱$$

تعریف ۵.۲.۱. یک زیر مجموعه E از فضای متریک X را یک مجموعه فشرده می‌گوییم اگر هر پوشش باز E یک زیر پوشش باز متناهی داشته باشد.

شرایط لازم و کافی برای مسائل نامقید

قضیه ۶.۲.۱. (شرط لازم مرتبه اول [۱]): فرض کنید X یک مجموعه باز در $\mathbb{R}^n$ باشد. مسأله برنامه‌ریزی نامقید زیر را در نظر بگیرید:

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &\text{subject to} \\ &\quad x \in X, \end{aligned} \tag{۱.۱}$$

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

قضیه ۷.۲.۱. (شرط لازم مرتبه دوم[۱]): فرض کنید مشتق دوم $f : \mathbb{R}^n \to \mathbb{R}$ در $\hat{x}$ موجود باشد. اگر $\hat{x}$ یک کمینه موضعی باشد، آنگاه $\nabla f(\hat{x}) = ۰$ و ماتریس هسین [۴] f در $\hat{x}$ نیمه معین مثبت است.

[۴] Hessian

قضیه ۸.۲.۱. (شرط کافی [۱]): فرض کنید مشتق دوم $f : \mathbb{R}^n \to \mathbb{R}$ در $\hat{x}$ وجود باشد. اگر $\nabla f(\hat{x}) = ۰$ و ماتریس هسین f در $\hat{x}$ معین مثبت باشد، آنگاه $\hat{x}$ یک کمینه موضعی اکید است.

قضیه ۹.۲.۱. ([۱]): اگر $f(x)$ بر مجموعه محدب E تابعی محدب باشد آنگاه $f(x)$ دارای یک کمینه موضعی[۵] است. اگر چنین کمینه‌ای یافت شود، یک کمینه سراسری [۶] است و بر روی مجموعه محدب به‌دست آورده می‌شود.

قضیه ۱۰.۲.۱. ([۱]): فرض کنید E مجموعه نقاطی از $\mathbb{R}^n$ باشد که در قیود $g_k(x) \leq ۰,\ x \leq ۰$ برای $k = (۱, ..., m)$ صدق کند. اگر $g_k(x)$ ها توابعی محدب باشند آنگاه E یک مجموعه محدب است.

دو قضیه بالا بیان می‌کند که برای یک مسأله برنامه ریزی کلی کمینه سازی، اگر $f(x)$ و همه $g_k(x)$ ها توابعی محدب باشند آنگاه یک کمینه موضعی از $f(x)$ تحت این مجموعه از قیود، یک کمینه سراسری است.

تعریف ۱۱.۲.۱. یک ماتریس $M(X)$ در اندازه $n \times n$ که عناصر آن m_{ij} $(i, j = ۱, ..., n)$ توابعی روی $E \subset \mathbb{R}^n$ هستند،

۱. نیمه معین مثبت [۷] روی E نامیده می‌شود، اگر

$$\forall\, V \in \mathbb{R}^n, V \neq ۰, X \in E \ :\ V^T M(X) V \geq ۰.$$

۲. معین مثبت [۸] روی E نامیده می‌شود، اگر

$$\forall\, V \in \mathbb{R}^n, V \neq ۰, X \in E \ :\ V^T M(X) V > ۰.$$

[۵] Local minimum
[۶] Global minimum
[۷] Positive semi-definite
[۸] Positive definite

۳. معین مثبت قوی روی E نامیم ، اگر عددی مثبت مانند α وجود داشته باشد به‌طوری که

$$\forall\, V \in \mathbb{R}^n, X \in E \;:\; V^T M(X) V \geq \alpha \|V\|^2.$$

تعاریف فوق را می‌توان بر اساس مفهوم مقدار ویژه نیز ارائه داد.

اگر $\gamma(x)$ کوچکترین مقدار ویژه ماتریس $M(X)_{n\times n}$ باشد آنگاه:

۱. $M(X)_{n\times n}$ روی E نیمه معین مثبت اگر و فقط اگر به ازای هر $\gamma(x) \geq 0, X \in E$

۲. $M(X)_{n\times n}$ روی E معین مثبت اگر و فقط اگر به ازای هر $\gamma(x) > 0, X \in E$

۳. $M(X)_{n\times n}$ روی E معین مثبت قوی اگر و فقط اگر به ازای هر $\gamma(x) \geq \alpha \geq 0, X \in E$

شرایط لازم و کافی برای مسائل مقید

قضیه ۱.۲.۱. (شرایط لازم کروش-کان-تاکر [۹] $(K.K.T)$،[۱]): فرض کنید X یک مجموعه باز در $\mathbb{R}^n$ باشد. مسأله کمینه‌سازی مقید زیر را در نظر بگیرید:

$$\begin{aligned} &\text{Minimize} \quad f(x_1, x_2, ..., x_n) \\ &\text{subject to} \\ &\begin{cases} g_k(x_1, x_2, ..., x_n) \leq 0, & k = 1, ..., m, \\ h_j(x_1, x_2, ..., x_n) = 0, & j = 1, ..., l, \\ x = (x_1, x_2, ..., x_n)^T \in X. \end{cases} \end{aligned} \tag{۲.۱}$$

که در آن توابع $f : \mathbb{R}^n \to \mathbb{R}$ ، $g_k : \mathbb{R}^n \to \mathbb{R}$ و $h_j : \mathbb{R}^n \to \mathbb{R}$. فرض کنید $\hat{x}$ یک جواب شدنی این مسأله باشد و $K = \{k : g_k(\hat{x}) = 0\}$. همچنین فرض کنید f و g_k برای $k \in K$ در $\hat{x}$ مشتق پذیر ، g_k برای $k \notin K$ در $\hat{x}$ پیوسته و h_j برای $(j = 1, ..., l)$ دارای مشتق پیوسته باشند. به‌علاوه فرض کنید $\nabla g_k(\hat{x})$ برای $k \in K$ و $\nabla h_j(\hat{x})$ برای $(j = 1, ..., l)$ مجموعاً مستقل‌خطی باشند. اگر $\hat{x}$ کمینه

[۹] Krvsh-Kuhn-Tucker

موضعی برای (۲.۱) باشد، آنگاه اسکالرهای یکتای $u_k \geq ۰$ برای $k \in K$ و $v_j \geq ۰$ برای $(j = ۱, ..., l)$ موجود هستند به‌طوری که:

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

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

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

همچنین شرایط $K.K.T$ می‌تواند به‌فرم ماتریسی زیر نوشته شود:

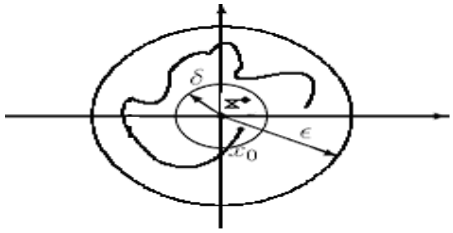
$$\begin{cases} \nabla f(\hat{x}) + \nabla g(\hat{x})^T u + \nabla h(\hat{x})^T v = ۰, \\ u^T g(\hat{x}) = ۰, \\ u \geq ۰. \end{cases}$$

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

قضیه ۱۳.۲.۱. (شرایط کافی کروش-کان-تاکر $(K.K.T)$،[۱]): فرض کنید X یک مجموعه باز در $\mathbb{R}^n$ باشد، مسأله (۲.۱) را در نظر بگیرید. فرض کنید $\hat{x}$ یک جواب شدنی باشد و $K = \{k : g_k(\hat{x}) = ۰\}$، اگر شرایط $K.K.T$ در $\hat{x}$ برقرار باشد، یعنی اسکالرهای $u_k \geq ۰$ برای $k \in K$ و $v_j \geq ۰$ برای $(j = ۱, ..., l)$ موجود باشند به‌طوری که ،

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

آنگاه $\hat{x}$ یک کمینه موضعی برای (۲.۱) خواهد بود.



شکل ۱.۱: پایداری لیاپانوف

۳.۱ پایداری

سیستم دینامیکی زیررا در نظر بگیرید:

$$\begin{cases} \dot{x} = f(x(t)), \\ x(t_\circ) = x_\circ \in \mathbb{R}^n \end{cases} \tag{۳.۱}$$

که در آن $f : \mathbb{R}^n \to \mathbb{R}^n$ است. x^* یک نقطهٔ تعادل [۱۰] برای (۳.۱) نامیده می‌شود اگر $f(x^*) = \circ$.

تعریف ۱.۳.۱. فرض کنید که $x(t)$ یک جواب (۳.۱) باشد، نقطهٔ تعادل تنها x^*، پایدار به مفهوم لیاپانوف [۱۱] است با توجه به شکل (۱.۱) ، اگر برای هر $x(t_\circ) = x_\circ$ و هر $\varepsilon > \circ$ ، یک $\delta > \circ$ وجود داشته باشد به‌طوری که اگر $\| x(t) - x_\circ \| \leq \delta$ آنگاه

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

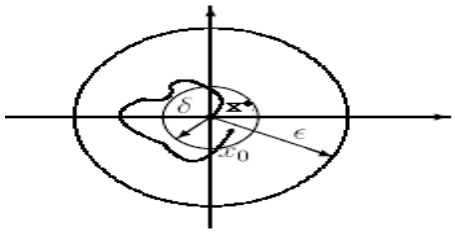
تعریف ۲.۳.۱. سیستم دینامیکی (۳.۱) همگرای سراسری [۱۲] به مجموعه

$$\Omega^* = \{x \mid \text{سیستم (۳.۱) جواب دارد}\}$$

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

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

[۱۰] Equilibrium point
[۱۱] Lyapunov
[۱۲] Global convergance



شکل ۲.۱: پایداری مجانبی

که در آن

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

تعریف ۳.۳.۱. سیستم دینامیکی در نقطهٔ تعادل یکتا x^* پایدار مجانبی سراسری [۱۳] نامیده می‌شود اگر x^* به مفهوم لیاپانوف پایدار باشد و در شرط زیر صدق کند:

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

شکل (۲.۱) را ببینید.

تعریف ۴.۳.۱. مجموعه $M \subseteq \mathbb{R}^n$ ، یک مجموعه پایدار [۱۴] نسبت به سیستم (۳.۱) گفته می‌شود اگر $x(t_0) \in M$ و $t_0 \geq 0$، آنگاه به ازای هر $t \geq t_0$، $x(t) \in M$ باشد.

تعریف ۵.۳.۱. نگاشت F در شرط لیپ شیتز [۱۵] صدق می‌کند اگر عدد ثابت L وجود داشته باشد، به‌طوری که برای هر دو نقطه $x, y \in \mathbb{R}^n$ داشته باشیم:

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

F، را لیپ شیتز محلی روی $\mathbb{R}^n$ نامیم اگر به ازای هر نقطه از $\mathbb{R}$، یک همسایگی مانند $D_0 \subset \mathbb{R}^n$ وجود داشته باشد به‌طوری که نامساوی بالا برای هر دو نقطه $x, y \in D_0$ برقرار باشد.

[۱۳] Globally asympotically stable
[۱۴] Invariant set
[۱۵] Lipschitz

تعریف ۶.۳.۱. نگاشت $F : \mathbb{R}^n \to \mathbb{R}^n$ یکنوا[۱۶] گفته می‌شود، اگر برای هر زوج از نقاط $x, y \in \mathbb{R}^n$ داشته باشیم:

$$(x-y)^T(F(x)-F(y)) \geq ۰.$$

همچنین روی $\mathbb{R}^n$ اکیداً یکنوا [۱۷] است، اگر نامساوی فوق به‌صورت اکید برای هر $x \neq y$ برقرار باشد. و F روی $\mathbb{R}^n$ یکنوای قوی است اگر برای هر زوج از نقاط $x, y \in \mathbb{R}^n$ ثابت $\beta > ۰$ وجود داشته باشد، به‌طوری که:

$$(x-y)^T(F(x)-F(y)) \geq \beta \parallel x-y \parallel^۲ .$$

قضیه ۷.۳.۱. ([۳]): فرض می‌کنیم که $F : K \subset \mathbb{R}^n \to \mathbb{R}^n$ ، یک تابع به‌طور پیوسته مشتق پذیر روی K باشد و ماتریس ژاکوبین F که لزوماً متقارن نیست، نیمه معین مثبت (معین مثبت) باشد، آنگاه $F(x)$ یکنوا (یکنوای قوی) است.

قضیه ۸.۳.۱. ([۴]): در سیستم دینامیکی (۳.۱) فرض می‌کنیم که $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ یک تابع پیوسته، آنگاه برای هر $t_۰ > ۰$ و $x_۰ \in \mathbb{R}^n$ ، یک جواب محلی [۱۸] $x(t)$ به ازای $t \in [t_۰, \tau)$ که $\tau > t_۰$ وجود دارد. علاوه بر این اگر f در $x_۰$ در شرط لیپ شیتز محلی صدق کند آنگاه جواب یکتا خواهد بود و اگر f در $\mathbb{R}^n$ در شرط لیپ شیتز صدق کند آنگاه τ می‌تواند تا $+\infty$ توسعه داده شود.

تذکر:

همیشه می‌توان نقطهٔ تعادل را به‌صورت $x^* = ۰$ در نظر گرفت. برای هر نقطهٔ تعادل دیگر می‌توان با استفاده از تغییرمتغیر، یک سیستم جدید با نقطهٔ تعادل y^* تعریف کرد. برای این منظور تعریف می‌کنیم:

$$y = x - x^* \Rightarrow \dot{y} = \dot{x} = f(x) \Longrightarrow f(x) = f(y + x^*) = g(y).$$

[۱۶] Monotone
[۱۷] Strictly monotonic
[۱۸] local solution

در نتیجه نقطه تعادل سیستم جدید $\dot{y} = g(y), y^* = ۰$ است زیرا

$$g(۰) = f(۰ + x^*) = f(x^*) = ۰.$$

در نتیجه x^* نقطهٔ تعادل برای سیستم $\dot{x} = f(x)$ است اگر و فقط اگر $y = ۰$ نقطهٔ تعادل سیستم $\dot{y} = g(y)$ باشد.

۴.۱ تابع انرژی

قضیه ۱.۴.۱. ([۵]): فرض کنید که $x = ۰$ یک نقطهٔ تعادل برای سیستم (۳.۱) و تابع $E : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ به‌طور پیوسته مشتق پذیر باشد اگر

۱. $E(۰) = ۰$،

۲. به ازای هر $x \in \Omega \subseteq \mathbb{R}^n \setminus \{۰\}$ ، $E(X) > ۰$ ،

۳. به ازای هر $x \in \Omega \subseteq \mathbb{R}^n \setminus \{۰\}$ ، $\dot{E}(X) \leq ۰$ ،

آنگاه $x = ۰$ نقطهٔ پایداری سیستم خواهد بود و $E(x)$ را« تابع لیاپانوف» یا «تابع انرژی» برای سیستم (۳.۱) نامیم.

قضیه ۲.۴.۱. ([۵]): فرض کنید که $x = ۰$ یک نقطهٔ تعادل برای سیستم (۳.۱) باشد و $E : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ یک تابع به‌طور پیوسته مشتق پذیر باشد اگر

۱. $E(۰) = ۰$،

۲. به ازای هر $x \in \Omega \subseteq \mathbb{R}^n \setminus \{۰\}$ ، $E(X) > ۰$ ،

۳. به ازای هر $x \in \Omega \subseteq \mathbb{R}^n \setminus \{۰\}$ ، $\dot{E}(X) < ۰$ ،

آنگاه $x = ۰$ پایدار مجانبی خواهد بود.

قضیه ۳.۴.۱. ([۵]): فرض کنید که $x = ۰$ یک نقطهٔ تعادل برای سیستم (۳.۱) و تابع $E : \Omega \subseteq \mathbb{R}^n \to$ $\mathbb{R}$به‌طور پیوسته مشتق‌پذیر باشد اگر

۱. $E(۰) = ۰$،

۲. به ازای هر $x \in \mathbb{R}^n \setminus \{۰\}$ ، $E(X) > ۰$ ،

۳. به ازای هر $x \in \mathbb{R}^n \setminus \{۰\}$ ، $\dot{E}(X) < ۰$ ، ،

آنگاه نقطهٔ $x = ۰$ پایدار مجانبی سراسری است.

قضیه ۴.۴.۱. ([۵۴]): اصل تغییرناپذیری (ناوردای) لازال [۱۹] : فرض می‌کنیم تابع $V : \mathbb{R}^n \to \mathbb{R}$ به‌طور پیوسته مشتق پذیر باشد، همچنین فرض می‌کنیم

۱. $M \subseteq \mathbb{R}^n$ یک مجموعه فشرده و پایدار نسبت به جواب سیستم (۳.۱) باشد،

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

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

۴. E بزرگترین مجموعه پایدار در N ،باشد

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

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

[۱۹] Lasalle's invariance principle

۵.۱ برنامه‌ریزی مخروط مرتبه دوم

تعریف ۱.۵.۱. مجموعه $C \subset \mathbb{R}^n$ یک مخروط [۲۰] نامیده می‌شود اگر:

$$\forall x \in C, \quad \lambda \geq ۰ \Rightarrow \lambda x \in C.$$

تعریف ۲.۵.۱. یک مخروط $C \subset \mathbb{R}^n$، مخروط مناسب نامیده می‌شود اگر:

۱. C محدب و بسته باشد،

۲. C بسته،

۳. C توپر باشد (درون آن ناتهی باشد)،

۴. C نوک تیز باشد $(x \in C, -x \in C \Rightarrow x = ۰)$.

فرض کنید که $\mathcal{K}$ یک مخروط مناسب است، در این صورت نامساوی مخروط به‌صورت زیر تعریف می‌شود:

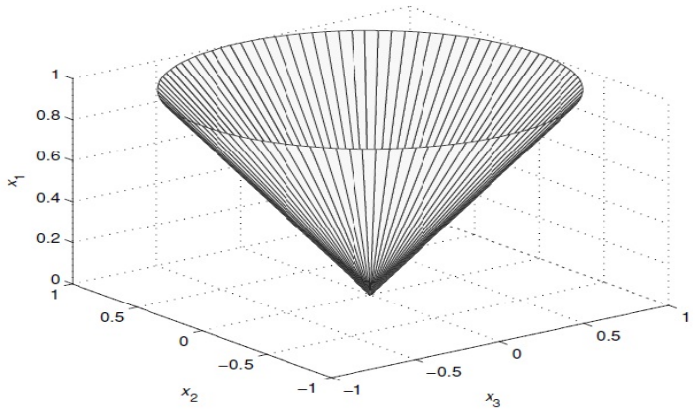
$$x \preceq_{\mathcal{K}} y \Leftrightarrow x - y \in \mathcal{K}.$$

توجه کنید که نامساوی مخروط $\preceq_{\mathcal{K}}$ به‌طور کلی نسبت به مخروط مناسب $\mathcal{K}$ معنی می‌دهد.

به منظور تعریف کردن مخروط مرتبه دوم [۲۱] در $\mathbb{R}^n$ $(n \geq ۱)$ ، افرازی از یک بردار $x \in \mathbb{R}^n$ را معرفی کرده که توسط $x = \begin{bmatrix} x_۰ \\ x_۱ \end{bmatrix}$ داده می‌شود که در آن $x_۰ \in \mathbb{R}$ و $x_۱ \in \mathbb{R}^{n-۱}$. بنابراین مخروط مرتبهٔ دوم به‌صورت مجموعه‌ای از $x \in \mathbb{R}^n$ به‌طوری که $|x_۰| \geq \|x_۱\|$ تعریف می‌شود، که این مخروط نیز شرایط مخروط مناسب را دارد. پس نامساوی ذکر شده همان نامساوی $\succeq_{\mathcal{K}}$ است. همچنین مخروط مرتبهٔ دوم C_q را به‌صورت زیر تعریف می‌کنیم:

$$C_q = \{x \in \mathbb{R}^n \mid {x_۱}^۲ \geq \| x_{۲n} \|^۲, x_۱ \geq ۰\},$$

[۲۰] Cone
[۲۱] Second Order Cone



شکل ۳.۱: مخروط مرتبه دوم در $\mathbb{R}^3$

که در آن $x_{2n} = (x_2, ..., x_n)$ است، به مخروط مرتبه دوم، مخروط لورنتس (مخروط بستنی) نیز گفته می‌شود. مخروط مرتبه دوم در فضای سه بعدی در شکل (۳.۱) نشان داده شده است. اکنون حالت خاصی از برنامه‌ریزی مخروط مرتبه دوم را به‌صورت زیر تعریف می‌کنیم که برای حل بعضی مسائل بهینه‌سازی استفاده می‌شود:

$$\begin{aligned} \text{Minimize} \quad & f^T x \\ \text{subject to} \quad & \\ & \|A_i x + b_i\| \le c_i^T x + d_i, \quad i = 1, ..., N \end{aligned} \tag{۴.۱}$$

که در آن $x \in \mathbb{R}^n$ متغیر بهینه‌سازی و $f \in \mathbb{R}^n$ ، $A_i \in \mathbb{R}^{(n_i-1)\times n}$، $b_i \in \mathbb{R}^{n_i-1}$، $c_i \in \mathbb{R}^n$ و $d_i \in \mathbb{R}$ پارامترهای مسأله هستند. نرمی که در قیود به کار رفته است، نرم استاندارد اقلیدسی است، یعنی $||u||_2 = (u^T u)^{\frac{1}{2}}$. یک مخروط مرتبه دوم استاندارد از بعد k به‌صورت زیر تعریف می‌شود:

$$c_k = \left\{ \left[\begin{array}{c} u \\ t \end{array} \right] \;\middle|\; u \in \mathbb{R}^{k-1}, t \in \mathbb{R}, \; \|u\| \le t \right\}.$$

برای $k = 1$ مخروط مرتبه دوم بالا به‌صورت زیر

$$c_1 = \left\{ \; t| \quad t \in \mathbb{R}, \; t \ge 0 \; \right\}.$$

مجموعه نقاطی که در یک قید مخروط مرتبه دوم صدق می‌کنند، نقش معکوس مخروط مرتبه دوم تحت یک نگاشت آفین هستند:

$$||A_i x + b_i|| \leq c_i^T x + d_i \Leftrightarrow \begin{bmatrix} A_i \\ c_i^T \end{bmatrix} x + \begin{bmatrix} b_i \\ d_i \end{bmatrix} \in C_{n_i}.$$

که منظور از C_{n_i} مخروط مرتبه دوم ازبعد n_i است که مجموعه‌ای محدب است. بنابراین برنامه‌ریزی مخروط مرتبه دوم (۴.۱) یک برنامه‌ریزی محدب است، برای مثال هنگامی که $n_i = ۱$ ، برنامه‌ریزی مخروط مرتبه دوم به برنامه‌ریزی خطی [۲۲] تبدیل می‌شود:

$$\begin{aligned} &\text{Minimize} \quad f^T x \\ &\text{subject to} \\ &c_i^T x + d_i \geq ۰, \quad i = ۱, ..., N. \end{aligned} \tag{۵.۱}$$

حالت خاص و جالب دیگری که می‌شود به‌دست آورد زمانی که $c_i = ۰$ است، که در این حالت i امین قید مخروط مرتبه دوم به $\|A_i x + b_i\| \leq d_i$ کاهش می‌یابد،که آن معادل (با فرض $d_i \geq ۰$) قید درجهٔ دو (محدب) $\|A_i x + b_i\|^۲ \leq d_i{}^۲$ است. بنابراین وقتی‌که c_i ها صفر هستند، برنامه‌ریزی مخروط مرتبه دوم به یک برنامه‌ریزی خطی با قیود درجه دوم [۲۳] منجر می‌شود. همچنین می‌بینیم که برنامه‌ریزی درجه دوم محدب و برنامه‌ریزی درجه دوم مقید شده و تعداد زیاد دیگری از مسائل بهینه سازی را می‌توان به‌صورت برنامه ریزی مخروط مرتبه دوم فرمول بندی کرد.

۶.۱ مفاهیم آماری

تعریف ۱.۶.۱. تابع $f(x) = \begin{cases} > ۰ & x \in X(s) \\ = ۰ & x \notin X(s) \end{cases}$ را تابع جرم احتمال متغیر تصادفی گسسته X گویند اگر $\sum_{x \in X(s)} f(x) = ۱$ و برای پیشامد A داریم، $\mathrm{P}(A) = \sum_{x \in A} f(x)$.

[۲۲] Linear programming(LP)
[۲۳] Quadratic Constrained Linear Programming

تعریف ۲.۶.۱. امید ریاضی یا میانگین متغیر تصادفی گسسته X، که با $E(X)$ یا μ_x نمایش داده شده، به‌صورت زیر تعریف می‌شود:

$$E(X) = \mu_x = \sum_x xP(X = x).$$

تعریف ۳.۶.۱. واریانس متغیر تصادفی X، که با $\mathrm{Var(X)}$ یا σ^2 نمایش داده می‌شود، به‌صورت زیر تعریف می‌شود:

$$\sigma^2 = \mathrm{Var(X)} = E(X - \mu)^2,$$

همچنین σ را انحراف معیار متغیر تصادفی X گوییم.

تعریف ۴.۶.۱. کوواریانس دو متغیر تصادفی X و Y عبارت است از:

$$\mathrm{Cov(X, Y)} = E[(X - \mu_x)(Y - \mu_y)] = E(XY) - E(X)E(Y).$$

تعریف ۵.۶.۱. ضریب همبستگی [۲۴] دو متغیر تصادفی X و Y عبارت است از:

$$\rho = \frac{\mathrm{Cov(X, Y)}}{\sqrt{\mathrm{Var(X)}\ \mathrm{Var(Y)}}}.$$

تعریف ۶.۶.۱. تابع $f(x) = \begin{cases} > 0 & x \in X(s) \\ = 0 & x \notin X(s) \end{cases}$ را تابع چگالی احتمال متغیر تصادفی پیوسته X گویند اگر $\int_{-\infty}^{\infty} f(x)dx = 1$ و برای پیشامد B داریم، $\mathrm{P}(B) = \int_B f(x)dx$.

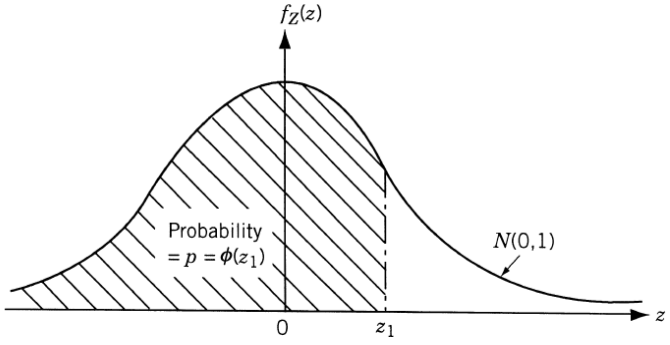
تعریف ۷.۶.۱. برای هر متغیر تصادفی X (اعم از گسسته یا پیوسته) تابع $F_X : \mathbb{R} \to [0, 1]$ ، تابع توزیع تجمعی X است که به‌صورت $F_X(x) = \mathrm{P}(X \leq x)$ تعریف می‌شود.

توزیع نرمال

فرض کنید تابع چگالی متغیر پیوسته X به‌صورت

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < +\infty,$$

[۲۴] Correlation Coefficient



شکل ۴.۱: تابع چگالی نرمال استاندارد

باشد، گوئیم X توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ دارد و می‌نویسم $X \sim N(\mu, \sigma^۲)$. همچنین اگر $\mu = ۰$ ، $\sigma^۲ = ۱$ باشد، گوئیم Z توزیع نرمال استاندارد دارد یا $Z \sim N(۰, ۱)$. تابع توزیع تجمعی نرمال استاندارد را با

$$\Phi(x) = P(Z \leq x) = \frac{۱}{\sqrt{۲\pi}} \int_{-\infty}^{x} e^{-\frac{z^۲}{۲}} dz,$$

نشان می‌دهند و مقادیر تقریبی آن را از جدول نرمال استاندارد حساب می‌کنند. با داشتن این جدول و با توجه به اینکه

$$X \sim N(\mu, \sigma^۲) \Longleftrightarrow z = \frac{X - \mu}{\sigma} \sim N(۰, ۱),$$

می‌توان مقادیر تقریبی احتمالات را برای هر توزیع نرمال یافت. طبق شکل (۴.۱) داریم:

$$\Phi(z_۱) = p, \quad z_۱ = \Phi^{-۱}(p),$$

و همچنین مقادیر $z_۱$ مربوط به $p < ۰/۵$ را می‌توان به‌صورت زیر به‌دست آورد:

$$z_۱ = \Phi^{-۱}(p) = -\Phi^{-۱}(۱ - p)$$

۷.۱ مفاهیم اقتصادی

با توجه به وجود فضای نامطمئن [۲۵] ، مهمترین مفاهیم در تصمیم‌گیری برای سرمایه‌گذاری، ریسک و بازده[۲۶]هستند.

۱.۷.۱ بازده

بازده در فرایند سرمایه‌گذاری، نیروی محرکی است که ایجاد انگیزه کرده و به‌عبارت بهتر پاداش سرمایه‌گذاران محسوب می‌گردد، چرا که تمام هدف سرمایه گذاری کسب بازده مطلوب است. از طرف دیگر تعیین تفاوت میان بازده تحقق یافته [۲۷] و بازده مورد انتظار [۲۸] از اهمیت بالایی برخوردار است. بازده تحقق یافته، بازده‌ای است که مربوط به گذشته بوده، حال آنکه بازده مورد انتظار، بازده تخمینی برای یک دارایی[۲۹] در دوره‌ای از آینده است که این بازده با عدم اطمینان[۳۰] همراه بوده و احتمال برآورده نشدن آن نیز وجود دارد. در مورد دارایی‌ها و مخصوصا سهام، بازده از دو جزء تشکیل شده است:

۱. بازده ناشی از تغییر قیمت [۳۱] ، این جزء از اختلاف بین قیمت خرید دارایی و فروش دارایی حاصل می‌شود.

۲. بازده ناشی از دریافت سود [۳۲] ، این بازده ناشی از صورت جریانات نقدی دریافتی به دلیل تملک دارایی در طول سرمایه‌گذاری است (همانند سود تقسیمی).

[۲۵] Uncertainity Space
[۲۶] Return
[۲۷] Realized Return
[۲۸] Expected Return
[۲۹] Asset
[۳۰] Uncertainty
[۳۱] Capital Gain
[۳۲] Yeild

بازده مورد انتظار هر سهم

با معلوم بودن توزیع احتمال بازده بالقوه هر سهم داریم

$$E(R_i) = \sum_{k=۱}^{m} (p_k) PR_k,$$

که در آن، p_k احتمال وقوع هر نرخ بازده بالقوه و PR_k بازده بالقوه یک سهم است.

سبد سرمایه

معنی ساده واژه سبد سرمایه (پورتفوی) [۳۳] ، سبد سرمایه‌گذاری به‌طور عام و سبد سهام به‌طور خاص، عبارت از ترکیب دارایی‌های سرمایه‌گذاری شده توسط یک سرمایه‌گذار اعم از فرد یا مؤسسه است. به لحاظ فنی، یک سبد سرمایه، مجموعه کامل دارایی‌های حقیقی و مالی سرمایه‌گذار را در بر می‌گیرد.

بازده مورد انتظار سبد سرمایه

در صورتی که سبد سرمایه با N دارایی داشته باشیم که وزن دارایی امi در آن x_i باشد، بازده مورد انتظار سبد سرمایه به‌صورت زیر محاسبه می‌شود:

$$E(R_p) = \sum_{i=۱}^{N} x_i E(R_i). \qquad (۶.۱)$$

فروش استقراضی

فروش استقراضی [۳۴] فروش اوراق بهادار به‌وسیله کسی که چنین اوراقی را ندارد. بنابراین باید آنها را قرض بگیرد و به خریدار تحویل دهد. فروش به این فرم باعث می‌شود که فروشنده از کاهش قیمت اوراق در آینده سود ببرد. فروشنده باید در زمان آینده اوراق را خریداری و بدهی خود را پرداخت نماید.

[۳۳] Portfolio
[۳۴] Short-Selling

۲.۷.۱ ریسک و مفهوم آن

به طور ساده می‌توان ریسک را عدم اطمینان به نتایج آتی تعریف کرد. به‌عبارت دیگر ریسک پدیده‌ای مربوط به آینده است که نمی‌توان آن را به‌طور دقیق پیش بینی نمود، زیرا با عدم اطمینان همراه است. هر چه این عدم اطمینان بیشتر باشد، ریسک هم بیشتر است. در بازار های مالی به‌همراه هر فرصتی ریسک هم وجود دارد و اصولا نمی‌توان کلیه ریسکها را از بین برد زیرا کلیه فرصتها هم از بین خواهند رفت. سرمایه گذار موفق کسی است که سطح قابل قبولی از ریسک را بپذیرد.

تعاریف مختلفی از ریسک بیان شده از جمله، ریسک را معادل نوسان بازدهی دارایی در نظر می‌گیرند. استفاده از واریانس بازده (نوسان حول میانگین) به‌عنوان معیار ریسک بر مبنای همین نظر است. همچنین ریسک عبارت است از میزان اختلاف بازده واقعی یک سرمایه‌گذاری از بازده مورد انتظار آن، معمولا در اقتصاد فرض بر اینکه سرمایه‌گذاران منطقی عمل می‌کنند و اطمینان را بر عدم‌اطمینان ترجیح می‌دهند و سرمایه‌گذاران ریسک گریزند[۳۵] .

ریسک سبد سرمایه

ریسک سبد سهام، تابعی از سهام منفرد و کواریانس بین بازده‌های سهام منفرد است که به‌صورت زیر بیان می‌شود:

$$\mathrm{Var}(R_p) = \sum_{i=۱}^{N} x_i \mathrm{Var}(R_i) + \sum_{i=۱}^{N} \sum_{i\neq j}^{N} \sum_{j=۱}^{N} x_i x_j \mathrm{Cov}(R_i, R_j). \qquad (۷.۱)$$

که در آن $\mathrm{Cov}(R_i, R_j) = \rho_{ij}\sigma_i\sigma_j$ کواریانس بین بازده‌های سهم‌های i و j است و ρ_{ij} ضریب همبستگی بازده‌های i و j است.

[۳۵] Risk-averse Investor

اندازه‌گیری ریسک

تاکنون معیارهای مختلفی برای تعیین ریسک معرفی شده، شاخص‌های اندازه‌گیری ریسک اولین بار از طریق مطالعات شاخص‌های پراکندگی آماری محاسبه گردیدند و از آن به بعد روش‌های جدیدتری از جمله ریسک نامطلوب [۳۶] ، استفاده از دیرش [۳۷] برای محاسبه حساسیت ارزش اوراق قرضه و در نهایت ارزش در معرض ریسک [۳۸] معرفی گردیدند که همگی از روش‌های آماری استفاده می‌کنند.

در سال ۱۹۵۲ ، هری مارکوویتز با ارائه مدلی کمی به اندازه‌گیری ریسک پرداخت و با معرفی مدلی مبتنی بر ریسک و بازده و ارائه مجموعهٔ کارا [۳۹] برای اولین بار ریسک را در کنار بازده به‌عنوان متغیری دیگر جهت انتخاب دارایی برای سرمایه‌گذاری قرار داد. وی انحراف معیار را به‌عنوان شاخص پراکندگی، معیار عددی ریسک خواند. شاگرد او ویلیام شارپ، شاخص بتا (ضریب حساسیت) را برای تغییرات نسبی ارزش یک سهم در قبال تغییر ارزش بازار با معرفی خط مشخصات ارائه کرد و با معرفی مدل قیمت‌گذاری دارایی‌های سرمایه‌ای، مدیریت علمی سبد سرمایه را پایه‌گذاری کرد. مک کالی [۴۰] معیار دیرش را به‌عنوان ملاک اندازه‌گیری ریسک اوراق بهادار با درآمد ثابت معرفی کرد، در ادامهٔ کارش معیار تحدب را به‌عنوان شاخصی دقیق‌تر معرفی کرد. در سال ۱۹۹۶ وترستون [۴۱]مدیر موسسهٔ $(J.P.Morgan)$ مدل ارزش در معرض ریسک را معرفی کرد. می‌توان طبقه بندی زیر را در مورد معیارهای اندازه‌گیری ریسک ارائه داد:

۱. حساسیت: تغییر یک متغیر وابسته بر اثر تغییر یک متغیر مستقل، مثل تغییرات قیمت در قبال تغییر یک واحد نرخ سود (دیرش، تحدب و بتا).

۲. نوسان: عبارت است از نوسان یک متغیر در اطراف میانگین و یا یک پارامتر تصادفی دیگر مثل

[۳۶] Downside Risk
[۳۷] Duration
[۳۸] Value at Risk
[۳۹] Efficient Frontier Line
[۴۰] McCauley
[۴۱] Weterstone

(واریانس و انحراف معیار).

۳. معیارهای ریسک نامطلوب: این معیارها که برعکس معیارهای نوسان، تنها بر بخش مخرب ریسک تمرکز دارند و در حقیقت نوسانات زیر سطح میانگین و یا متغیر هدف را مورد محاسبه قرار می‌دهند (نیم واریانس [۴۲] ،نیم انحراف معیار و ارزش در معرض ریسک).

تعریف ۱.۷.۱. نیم واریانس: شاخصی برای محاسبه پراکندگی صفت متغیر و برای انحرافات نامطلوب به‌کار برده می‌شود. به‌عبارت دیگر اگر ریسک را میزان ضرر تعریف کنیم، آنگاه تغییرات مطلوب (افزایش نرخ بازدهی دارایی مالی) به‌عنوان ریسک محسوب نمی‌شود و فقط آن دسته از مشاهداتی که کمتر از میانگین نرخ بازدهی هستند، به‌عنوان ریسک به‌حساب می‌آید:

$$SV = \frac{۱}{K}\sum_{i=۱}^{K}\{\max(۰, E - r_T)\}^۲,$$

که درآن، r_T بازدهی دارایی در طول دوره، K تعداد مشاهدات، E امید ریاضی نرخ بازدهی است.

نظریه مدرن سبد سرمایه MPT براساس رابطهٔ بازدهی و ریسک محاسبه شده از طریق واریانس و انحراف معیار بازدهی تبیین می‌شود. نظریه فرامدرن سبد سرمایه $PMPT$ [۴۳] براساس رابطهٔ ریسک نامطلوب به تبیین رفتار سرمایه‌گذار و معیار انتخاب سبد سرمایه بهینه می‌پردازد (برای توضیحات بیشتر، مراجع [۶، ۷] را ببینید.)

[۴۲] Semi Variance
[۴۳] Post-Modern Portfolio Theory

فصل ۲
مقدمه‌ای بر شبکه‌های عصبی

۱.۲ مقدمه

در این فصل به بیان مقدمه‌ای از مدل‌های شبکه‌های عصبی می‌پردازیم. در ابتدا ساختار شبکه عصبی و مدل ریاضی یک سلول عصبی را بیان می‌کنیم و در ادامه به بیان تاریخچه‌ای از شبکه‌های عصبی در حل مسائل بهینه سازی می‌پردازیم. در انتها به بیان مدل‌های ارائه شده پیشین برای حل مسائل بهینه سازی خواهیم پرداخت.

کار بر روی شبکه‌های عصبی مصنوعی، با الهام از عملکرد مغز انسان از آنجا شروع شد که دانشمندان دریافتند مغز بشر به‌گونه‌ای کاملاً متفاوت از کامپیوترهای دیجیتالی مرسوم محاسبات را انجام می‌دهد. مغز یک سیستم پردازش داده بسیار پیچیده است که از واحدهای ساختاری به نام سلول عصبی یا نرون[۱] تشکیل شده است. شبکه عصبی سیستمی پویا و غیرخطی است که از تعداد زیادی واحد پردازش (نرون ها) و اتصالات بین این واحدهای پردازش تشکیل می‌شود.

یک شبکه عصبی بر خلاف کامپیوتر که نیازمند دستورهای کاملاً صریح و مشخص است، به مدل‌های ریاضی محض نیازی ندارد، بلکه مانند انسان تجربه کسب کرده و سپس نتیجه این تجربیات را تعمیم می‌دهد. امروزه این شبکه‌ها تبدیل به ابزاری قدرتمند و عمومی شده‌اند که برای حل مسائل از قبیل تخمین (تقریب)، تشخیص الگو[۲] ، رباتیک، کنترل، و به‌طور کلی در هر جا که نیاز به یادگیری نگاشت خطی و یا غیرخطی باشد، به‌کار می‌روند.[۸]

برای حل هر مسأله‌ای، شبکه‌های عصبی سه مرحله را طی می‌کنند: آموزش [۳] ، تعمیم[۴] و اجرا[۵]. آموزش، فرایندی است که طی آن شبکه می‌آموزد تا الگوی موجود در ورودی‌ها را (که به‌صورت مجموعه داده‌های آموزشی است) بشناسد. برای این منظور هر شبکه عصبی از مجموعه‌ای از قوانین یادگیری

[۱] Neuron
[۲] Pattern recognition
[۳] Traning
[۴] generalization
[۵] Operation

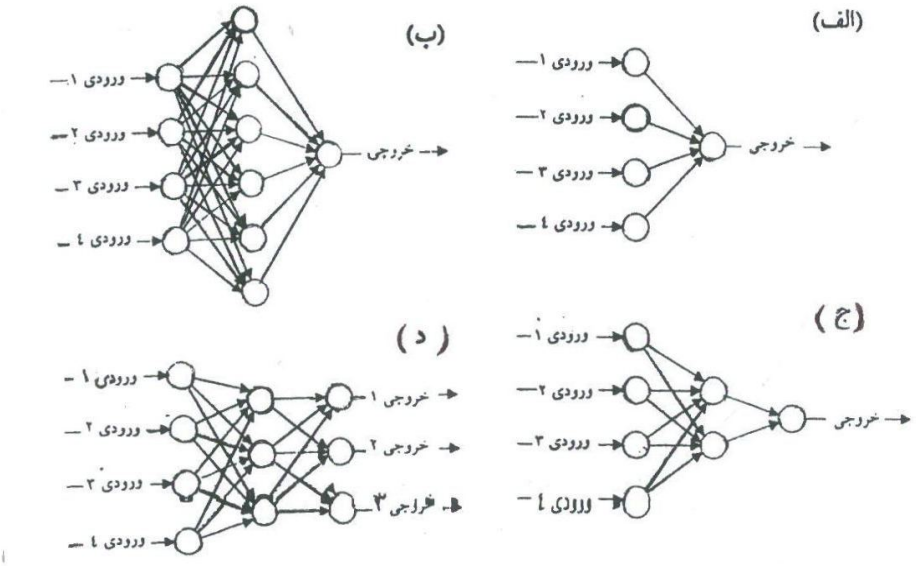
که نحوه یادگیری را تعریف می‌کنند استفاده می‌کند. تعمیم، توانایی شبکه برای ارائه جواب قابل قبول در قبال ورودی‌هایی است که در مجموعه آموزشی نبوده‌اند. یعنی پس از آنکه مثال‌های اولیه به شبکه آموزش داده شد، شبکه می‌تواند در قبال یک ورودی آموزش داده نشده قرار گیرد و یک خروجی مناسب را ارائه نماید. این خروجی براساس مکانیسم تعمیم که همانا چیزی جز فرایند درونیابی نیست به‌دست می‌آید. استفاده از شبکه برای انجام عملکردی که به آن منظور طراحی شده است را اجرا می‌گویند.

در اثر آموزش دیدن شبکه، وزن‌های داخلی[۶] (که بر روی ورودی‌های هر سلول اعمال می‌شود) تغییر می‌کند و به وضعیت مناسب می‌رسند. یکی از نقاط ضعف شبکه‌های عصبی این است که نتایج آموزش، یعنی وزن‌های داخلی، هیچگونه تصویر روشنی از اعتبار جواب‌های مسأله بدست نمی‌دهد. این وزن‌ها کاملاً قابل درک نیستند. با این وجود جواب‌های تولید شده توسط شبکه، اغلب صحیح است و با شرایط کمّی حاکم بر محیط سازگاری دارد. این صحت جوابها و صدق شرایط کمّی حاکم بر محیط، گاهی مهم‌تر از توضیح‌پذیر بودن آن است. تاکنون اسامی متفاوتی برای شبکه‌های عصبی به‌کار رفته است. از جمله آنها می‌توان به مدل‌های جعبه سیاه [۷]، مدل‌های تجربی ، تقریب زن‌های عام[۸] و مدل‌های موازی[۹] اشاره نمود.[۹]

۲.۲ ساختار شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی از یک سری واحدهای ساختمانی اولیه تشکیل می‌شوند. هر واحد دارای چندین ورودی است که با هم ترکیب شده و پس از انجام عملیات پردازش یک خروجی را به‌دست می‌دهند. این واحدهای اولیه مانند شکل (۱.۲) به‌هم متصل هستند به‌طوری‌که خروجی هر واحد به عنوان ورودی واحد های دیگر مورد استفاده قرار می‌گیرد. واحدهای ساختمانی را سلول عصبی، نرون یا گره[۱۰] نیز می‌نامند.

[۶] Internal Weight
[۷] Black-Box models
[۸] Universal Approximators
[۹] Parallel models
[۱۰] Node

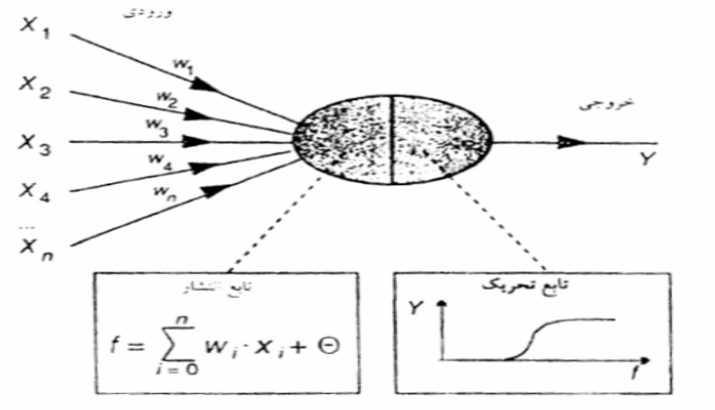


شکل ۱.۲: شبکه‌های عصبی مصنوعی

در شکل (۱.۲) ،(الف)-یک شبکه بسیار ساده شامل چهار ورودی و یک خروجی، نتیجه آموزش این شبکه تکنیک رگرسیون لجستیک[۱۱] در آمار را نشان می‌دهد. (ب)-شبکه دارای دو لایه میانی است که قدرت بیشتری در تشخیص الگو دارد. (ج)- افزایش تعداد لایه پنهان شبکه را قوی‌تر می‌کند ولی خطر بیش برازشی [۱۲] در آن وجود دارد. (د)- شبکه عصبی با چندین خروجی.

شبکه‌های عصبی اولیه تنها دو لایه (ورودی وخروجی) داشتند و به‌همین دلیل تنها در مسائل خطی مفید بودند. ساختارهای توسعه یافته امروزی شامل سه و یا تعداد بیشتری لایه هستند. رایج‌ترین ساختار شبکه، سه لایه‌ای است (لایه ورودی، یک لایه پنهان ولایه خروجی) و عموماً قادر به حل اکثر مسائل پیچیده است.

[۱۱] Logistic regression
[۱۲] Over fitting



شکل ۲.۲: مدل ساده یک سلول عصبی

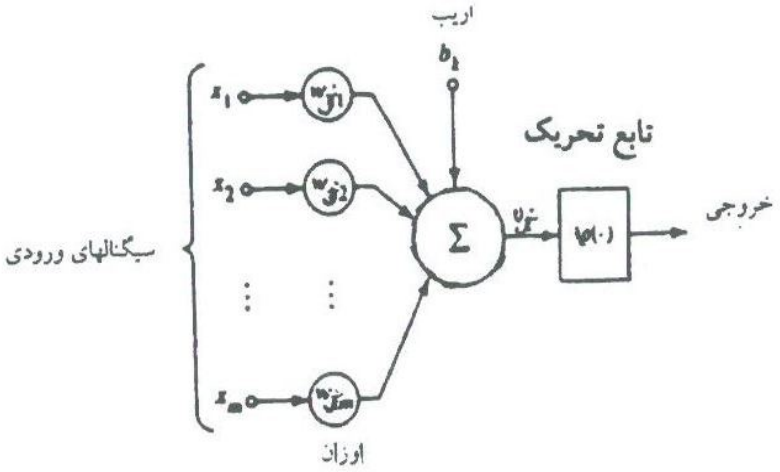
۳.۲ مدل ریاضی یک سلول عصبی

شکل (۲.۲) یک مدل سلول عصبی را نشان می‌دهد. بدنه این سلول از دو بخش تشکیل شده است اولین بخش را تابع ترکیب [۱۳] می‌گویند. وظیفه تابع ترکیب این است که تمام ورودی‌ها را ترکیب و یک عدد تولید کند. مطابق شکل (۲.۲) هر ورودی دارای وزن مختص به خود است . ورودی‌ها در وزن‌های مربوطه ضرب و سپس با هم جمع می‌شوند. مجموع حاصل را مجموع موزون [۱۴] گویند که رایج ترین تابع ترکیب است.

بخش دوم سلول عصبی، تابع انتقال [۱۵] نام داردکه تابع ترکیب را به خروجی سلول تبدیل و سپس انتقال می‌دهد. تابع انتقال را تابع تحریک [۱۶] نیز می‌نامند.

رایج‌ترین توابع تحریک، بر پایه مدل‌های بیولوژیک استوار گردیده است (برای توضیحات بیشتر مراجع [۹، ۱۱، ۱۰]را ببینید). اگر ورودی سلول i ام واقع در لایه ورودی را x_i در نظر بگیریم، این ورودی برای اتصال به سلول j ام لایه بعد در وزن w_{ji} ضرب می‌شود. حرف j در w_{ji} نشان دهنده شماره سلول در لایه

[۱۳] Combination function
[۱۴] Weighted sum
[۱۵] Transfer function
[۱۶] Activation function



شکل ۳.۲: مدل بلوکی سلول عصبی

بعد و حرف دوم اندیس معرف شماره سلول لایه قبلی است. شکل (۳.۲) مدل بلوکی یک سلول عصبی را نشان می‌دهد سلول مذکور دارای یک ورودی اضافی است که به آن اریبی [۱۷] می‌گویند و با b_j نشان داده می‌شود. نقش اریبی افزایش یا کاهش مجموع موزون است.

به بیان ریاضی، سلول j بوسیله دو رابطه زیر تعریف می‌شود:

$$u_j = \sum_{i=۱}^{m} w_{ji} x_i,$$

$$y_j = \phi(u_j + b_j),$$

که $x_۱, x_۲, ..., x_m$ داده‌های ورودی، $w_{j۱}, w_{j۲}, ...$ وزن‌های اتصال ورودی‌های ۲،۱ و ... به سلولj ، u_j ترکیب خطی ورودی‌ها، b_j اریبی، ϕ تابع تحریک و y_j خروجی سلول است. می‌توان مجموع u_j و b_j را به‌صورت زیر نمایش داد:

$$v_j = (u_j + b_j).$$

تابع تحریک یا تابع انتقال که با ϕ نشان داده می‌شود بر روی مجموع وزن‌ها (شامل اریبی) عمل می‌کند.

[۱۷] Bias

این تابع می‌تواند خطی یا غیرخطی باشد و بر اساس نیاز خاص حل یک مسأله انتخاب می‌شود. انواع مختلفی از توابع تحریک وجود دارند که مهم‌ترین آنها بشرح زیر می‌باشند:

۱. تابع حد آستانه [۱۸] : این تابع در شکل (۴.۲) قسمت (a) نشان داده شده و به‌صورت زیر بیان می‌شود

$$\phi(\nu) = \begin{cases} ۱ & \nu \geq ۰, \\ ۰ & \nu < ۰. \end{cases}$$

۲. تابع قسمتی خطی [۱۹] این تابع که در شکل : (۴.۲) قسمت (b) نشان داده شده است به‌صورت زیر تعریف می‌شود:

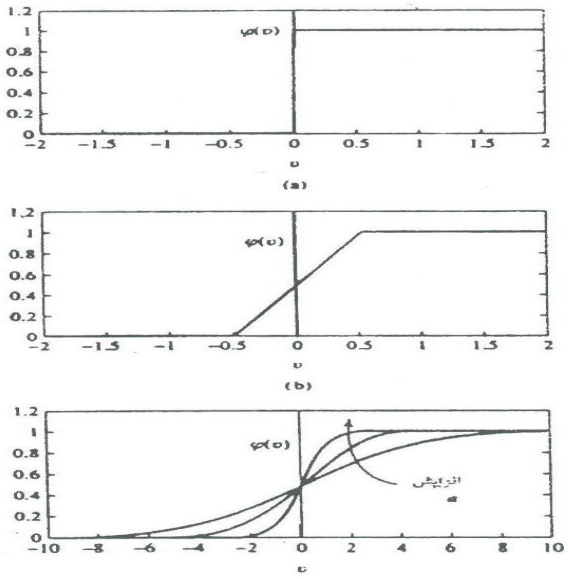
$$\phi(\nu) = \begin{cases} ۱ & \nu \geq ۰/۵, \\ \nu & -۰/۵ < \nu < ۰/۵, \\ ۰ & \nu < ۰/۵. \end{cases}$$

۳. تابع سیگموئید [۲۰] نام این تابع به علت شکل S مانند آن است و پرکاربردترین تابع تحریک محسوب می‌شود. این تابع بین رفتار خطی و غیرخطی، به آرامی تغییر می‌کند. مثالی از تابع سیگموئید تابع لجستیک [۲۱] است که به‌صورت زیر بیان می‌شود:

$$\phi(\nu) = \frac{۱}{۱ + exp(-a\nu)},$$

که در آن a پارامتر شیب تابع سیگموئید است، با تغییر پارامترa ، توابع مختلفی با شیب‌های متفاوت مانند آنچه که در شکل (۴.۲) قسمت (c) نشان داده شده بدست می‌آید.

[۱۸] Threshold function
[۱۹] Piecewise-Linear function
[۲۰] Sigmoid function
[۲۱] Logistic function



شکل ۴.۲: انواع تابع تحریک: الف) تابع حد آستانه ،ب) تابع قسمتی خطی ،ج) تابع سیگموئید با شیب‌های متغیر

۴.۲ تاریخچه کاربرد شبکه های عصبی مصنوعی در حل مسائل بهینه‌سازی

مسائل بهینه‌سازی درطیف وسیعی از کاربردهای علوم و مهندسی تحلیل پردازش تصویر[۲۲] ، شناسایی سیستم ها[۲۳] ، حل معادلات، طراحی فیلتر[۲۴] ، تقریب توابع [۲۵]، آنالیز رگرسیون[۲۶] و... پدیدار می‌شود. در بسیاری از کاربردهای علوم و مهندسی جواب زمان واقعی[۲۷] مسائل بهینه‌سازی بسیار مورد نیاز می‌شود. الگوریتم های قدیمی برای کامپیوترهای دیجیتال چندان موثر نیستند زیرا زمان مورد نیاز برای بدست آوردن جواب وابستگی زیادی به بُعد [۲۸] و ساختار[۲۹] مسأله دارد، این یک رهیافت ممکن و بسیار امیدوارکننده

[۲۲] Signal processing
[۲۳] System identification
[۲۴] Filter design
[۲۵] Function approximation
[۲۶] Regression analysis
[۲۷] Real time
[۲۸] Dimension
[۲۹] Structure

کاربرد شبکه‌های عصبی مصنوعی است.

به واسطه ذات شبکه‌های عصبی که می‌توانند حجم عظیمی از کارهای موازی را با هم انجام دهند رهیافت شبکه عصبی می‌تواند مسائل بهینه سازی در زمان واقعی را با سرعت بیشتری نسبت به سایر الگوریتم‌های بهینه‌سازی موجود حل کند. هاپفیلد [۳۰] و تنک [۳۱] در [۱۲] ابتدا شبکه عصبی را برای حل مسائل برنامه‌ریزی خطی پیشنهاد دادند،کار بنیادین آنها الهام بخش تحقیقات زیادی برای ارائه شبکه‌های عصبی دیگر برای حل مسائل برنامه‌ریزی خطی و غیرخطی شد. بدین ترتیب شبکه‌های عصبی بهینه‌سازی زیادی گسترش یافتند. به‌عنوان مثال کندی[۳۲] و چاو [۳۳] در [۱۳] یک شبکه عصبی برای حل مسائل برنامه‌ریزی غیرخطی پیشنهاد کردند که از روش گرادیان و روش تابع جریمه استفاده می‌نمودند و نقاط تعادلشان متناظر با تقریبی از جواب‌های بهینه بود. با استفاده از روش گرادیان و تصویر، بازِردوم[۳۴] و پاتیسن[۳۵] [۱۴] یک شبکه عصبی برای حل مسائل بهینه‌سازی درجه دوم تنها با متغیرهای کراندار معرفی کردند. این شبکه عصبی تحت شرایط خاصی به تنها جواب بهینه همگرا بود.

بر پایه روش‌های گرادیان و تصویر، ایکسیا[۳۶] و همکارانش[۲۰]-[۱۵] چندین شبکه عصبی برای حل مسائل برنامه‌ریزی درجه دوم و خطی با جواب‌های بهینه غیریکتا معرفی کردند و همگرایی سراسری آنها را به جواب‌های دقیق ثابت نمودند.

بر پایه بعضی از ملاحظات در اجرای سخت افزاری، یک شبکه عصبی با عملکرد محاسباتی خوب باید سه ویژگی اساسی داشته باشد:

۱. همگرایی شبکه عصبی به ازای هر وضعیت دلخواه داده شده تضمین شود.

[۳۰] Hopfield
[۳۱] Tank
[۳۲] Kennedy
[۳۳] Chua
[۳۴] Bouzerdoum
[۳۵] Pattison
[۳۶] Xia

۲. در طراحی شبکه ترجیحاً هیچ پارامتر مجهولی وجود نداشته باشد.

۳. نقاط تعادل شبکه متناظر با جواب‌های دقیق یا تقریبی باشد.

از دیدگاه ریاضی این خصوصیات مربوط به تکنیک‌های بهینه‌سازی هستند که برای طراحی مدل‌های شبکه‌های عصبی بهینه‌سازی به‌کار می‌روند.

در مجموع برای فرمول‌بندی یک مسأله بهینه‌سازی برحسب یک شبکه عصبی دو نوع روش وجود دارد. یک رهیافت که عموماً در ایجاد یک شبکه عصبی بهینه‌سازی استفاده می‌شود این است که در ابتدا مسأله بهینه‌سازی را به یک مسأله بهینه‌سازی نامقید تبدیل کرده و سپس یک شبکه عصبی طراحی می‌کنند که مسأله نامقید را با استفاده از روش گرادیان حل کند. روش گرادیان یک مزیت برای آن دسته از شبکه‌های عصبی است که مستقیماً از مشتق تابع انرژی استفاده می‌کنند اما عیب آنها این است که همگرایی در این موارد تضمین نمی‌شود بویژه درحالتی که مجموعه‌های جواب کراندار نیستند.

رهیافت دیگر این است که یک مجموعه از معادلات دیفرانسیل ایجاد می‌کنند به‌طوری‌که نقاط تعادل آنها متناظر با جواب مطلوب باشد و بعد یک تابع لیاپانوف مناسب پیدا کرده، به‌طوری‌که همه مسیرهای سیستم به نقاط تعادل همگرا باشند.

۵.۲ مدل‌های شبکه عصبی ارائه شده پیشین برای حل مسائل بهینه‌سازی

بحث وبررسی در مورد کاربرد شبکه‌های عصبی (شبکه‌های عصبی مصنوعی) در بهینه‌سازی از اوایل سال ۱۹۸۰ میلادی آغاز شده است. نتایج پژوهش‌های انجام شده از آن زمان تاکنون، مواردی چون برنامه‌ریزی خطی، برنامه‌ریزی درجه دوم، برنامه‌ریزی هندسی و برنامه‌ریزی غیرخطی را در برمی‌گیرد. ایده اصلی در استفاده از شبکه‌های عصبی مصنوعی برای مسائل بهینه سازی استفاده از یک تابع انرژی (نامنفی) و یک سیستم دینامیکی است که این دو بیان کننده مدل‌های شبکه‌های عصبی مصنوعی متناظر مسائل

بهینه‌سازی هستند. سیستم دینامیکی بیان شده معمولاً یک دستگاه معادلات دیفرانسیل غیرخطی مرتبه اول است. انتظار می‌رود که برای یک نقطهٔ آغازین نقطهٔ پایداری دستگاه معادلات دیفرانسیل غیرخطی به‌دست آمده، جواب بهینه مسأله بهینه‌سازی اصلی باشد. یک اصل اساسی در استفاده از تابع انرژی این است که برای همگرایی دستگاه معادلات دیفرانسیل به نقطهٔ پایداری آن، تابع انرژی متناظر با آن حتماً باید نامنفی باشد. البته می‌توان مدل شبکه‌های عصبی نظیر مسائل بهینه سازی را حتی بدون در نظر گرفتن تابع انرژی نیز بیان کرد. بنابراین برای یک مدل متناظر با مسائل بهینه‌سازی، اصل اساسی استفاده از شبکه‌های عصبی در اینگونه مسائل به‌صورت زیر بیان می‌شود:

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

مدل‌های مطرح شده متناظر مسائل مختلف بهینه‌سازی را می‌توان به دو قسمت مدل‌های دوگانی و مدل‌های جریمه‌ای تقسیم نمود. نظریه دوگانی و توابع جریمه‌ای دو نظریه بسیار مهم در مسائل بهینه‌سازی هستند که اکثر این مسائل بر مبنای این دو روش کلاسیک قابل حل هستند. در نظریه توابع جریمه‌ای معمولاً از مدل‌های گرادیانی برای معرفی مدل مورد نظر استفاده می‌شود، ولی در نظریه دوگانی از مدل‌های اولیه -دوگان برای حل این مسائل استفاده می‌شود. اگر بتوان متناظر با هر مسأله بهینه‌سازی، روش مشخصی را ارائه نمود که آن روش برای حل آن مسأله شرایط لازم وکافی را برآورده سازد، آنگاه می‌توان متناظر با آن روش، یک مدل شبکه عصبی برای مسأله مورد نظر بسازیم. در زیر به بیان چند مدل که تاکنون برای مسائل برنامه‌ریزی خطی و درجه دوم و غیرخطی ارائه شده‌اند می‌پردازیم.

مدل اول([۲۱])

مسأله برنامه‌ریزی خطی زیر را در نظر بگیرید:

$$\begin{cases} \text{Minimize} \quad f(x) = a^T x \\ \text{s.t.} \\ \quad g(x) = Dx - b = ۰, \\ \quad x \geq ۰, \end{cases} \tag{۱.۲}$$

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

$$\begin{cases} \text{Maximize} \quad \bar{f}(y) = b^T y \\ \text{s.t.} \\ \quad \bar{g}(y) = \bar{D}^T y - a \leq ۰, \\ \quad y \in \mathbb{R}^n. \end{cases} \tag{۲.۲}$$

در مرجع [۲۱] مدل متناظر با (۱.۲) و (۲.۲) به‌صورت زیر نمایش داده شده است:

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

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

مدل دوم([۲۱])

مسأله برنامه‌ریزی درجه دوم زیر را در نظر بگیرید:

$$\begin{cases} \text{Minimize} \quad f(x) = \frac{۱}{۲} x^T A x + a^T x \\ \text{s.t.} \\ \quad Dx - b = ۰, \\ \quad x \geq ۰, \end{cases} \tag{۳.۲}$$

که در آن A یک ماتریس متقارن معین مثبت است. دوگان (۳.۲) به‌صورت زیر است:

$$\begin{cases} \text{Maximize} \quad \bar{f}(y) = b^T y - \frac{۱}{۲} x^T A x \\ \text{s.t.} \\ \quad \bar{D} y - Ax - a \leq ۰. \end{cases} \tag{۴.۲}$$

در مرجع [۲۱] مدل متناظر با (۳.۲) و (۴.۲) به‌صورت زیر نمایش داده شده است:

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

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

مدل سوم([۱۲])

مسأله برنامه‌ریزی غیرخطی زیر را در نظر بگیرید:

$$\begin{cases} \text{Minimize} \quad f(x) \\ \text{s.t.} \\ \quad g(x) = [g_۱(x), g_۲(x), ..., g_m(x)]^T \leq ۰, \end{cases} \tag{۵.۲}$$

که در آن $f(x), g(x) \in C^۲$ و $f(x), g_i(x) : \mathbb{R}^n \to \mathbb{R}^۱$. برای حل (۵.۲) هاپفیلد و تنک در سال ۱۹۸۶، یک مدل شبکه عصبی به‌صورت زیر ارائه دادند:

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

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

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

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

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

مدل چهارم([۱۳])

برای حل (۵.۲) ، کندی و چوآ [۱۳] مدلی دیگر به‌صورت زیر ارائه داده‌اند:

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

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

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

مدل پنجم([۲۲])

مجدداً برای حل (۵.۲) توسط رودریگز-واسکوئز مدلی به‌صورت زیر ارائه شده است:

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

که در آن u_x متناظر اندیس‌های شدنی x است چنان‌که

$$u_x = \begin{cases} ۱ & g(x) \leq ۰, \\ ۰ & ۰.w. \end{cases}$$

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

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

همان طور که ملاحظه می‌کنید، مدل‌های بیان شده به‌طور متوالی اصلاح شده‌اند و این به‌دلیل مشکلاتی بوده است که هرکدام از این مدل‌ها داشته‌اند.([۲۱])

مدل ششم ([۴])

برنامه‌ریزی محدب درجه دوم (CQP)[۳۷] زیر را در نظر بگیرید:

$$\begin{cases} \text{Minimize} \quad f(x) = \frac{۱}{۲} x^T Q x + D^T x \\ \text{s.t.} \\ \quad g(x) = Ax - b \leq ۰, \\ \quad h(x) = Ex - f = ۰. \end{cases} \tag{۶.۲}$$

که در آن $Q \in \mathbb{R}^{n \times n}$ یک ماتریس متقارن معین مثبت، $A \in \mathbb{R}^{m \times n}$ ، $b \in \mathbb{R}^m$، $E \in \mathbb{R}^{l \times n}$، $f \in \mathbb{R}^l$ و $x \in \mathbb{R}^n$ است. دوگان برنامه‌ریزی محدب درجه دوم (۶.۲) به‌صورت زیر نوشته می‌شود:

$$\begin{cases} \text{Maximize} \quad \bar{f}(x,u,v) = f(x) + u^T(Ax - b) + v^T(Ex - f) \\ \text{s.t.} \\ \quad \bar{g}(x,u,v) = Qx + D + A^T u + E^T v = ۰, \\ \quad u \geq ۰. \end{cases} \tag{۷.۲}$$

[۳۷] Convex quadratic programming

بر طبق شرایط $K.K.T$ برای (۶.۲) و (۷.۲) ، $x^* \in \mathbb{R}^n$ جواب بهینه (۶.۲) است اگر وتنها اگر بردارهای $u^* \in \mathbb{R}^m$ و $v^* \in \mathbb{R}^l$ وجود داشته باشد به‌طوری‌که در رابطه زیر صدق کند:

$$\begin{cases} u^* \geq ۰, Ax^* - b \leq ۰, u^{*T}(Ax^* - b) = ۰, \\ Qx^* + D + A^T u^* + E^T v^* = ۰, \\ Ex^* - f = ۰. \end{cases} \tag{۸.۲}$$

برای حل مسأله (۶.۲) تابع زیر را که معروف به تابع فیشر-برمیستر[۳۸] است، معرفی می‌کنیم:

$$\begin{cases} \phi : R^۲ \to R^۱ \\ \phi(a,b) = \sqrt{a^۲ + b^۲} - a - b. \end{cases}$$

قضیه ۱.۵.۲. ([۲۳]):

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

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

۳. $\phi(a,b)$ همه جا به جز در مبدأ دو بار به‌طور پیوسته مشتق پذیر است، ولی در مبدأ قویاً [۳۹] هموار است.

تابع ϕ که در شرط (۱) قضیه (۱.۵.۲) صدق کند، تابع NCP [۴۰] نامیده می‌شود.

متناظر با دستگاه(۷.۲) و (۶.۲) دستگاه زیر را در نظر می‌گیریم:

$$\phi(x,u,v) = \begin{bmatrix} \phi(\alpha, Qx + D + A^T u + E^T v) \\ \phi(u, Ax - b) \\ \phi(\beta, Ex - f) \end{bmatrix}, \tag{۹.۲}$$

که در آن $\alpha = (\alpha_۱, ..., \alpha_m)^T > ۰$، $u = (u_۱, ..., u_m)$ و $\beta = (\beta_۱, ..., \beta_l)^T$ است. مدل شبکه عصبی متناظر با (۹.۲) به‌صورت زیر بیان شده است:

$$\begin{cases} \dfrac{dy}{dt} = -\tau \bigtriangledown E(y(t)), \quad \tau > ۰ \\ y(t_۰) = y_۰ \in \mathbb{R}^{n+m+۱}. \end{cases} \tag{۱۰.۲}$$

[۳۸] Fisher-Bermister
[۳۹] Stringly smooth
[۴۰] Nonlinear Complementarity Problem

که در آن $y = (x^T, u^T, v^T)^T$ و τ ضریبی برای افزایش سرعت همگرایی و کاهش تعداد تکرار ها در روش عددی انتخاب شده برای حل دستگاه (۹.۲) است. تابع انرژی متناظر این مدل به‌صورت زیر انتخاب شده است:

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

مدل هفتم ([۴])

برای طراحی مدل هفتم، مجدداً شرایط $K.K.T$ را برای حل مسائل (۷.۲) و (۶.۲) به‌صورت زیر در نظر می‌گیریم:

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

آنگاه متناظر (۶.۲) و(۷.۲) دستگاه زیر تعریف می‌شود:

$$U(x, u, v) = \begin{bmatrix} -(Qx + D + A^T u + E^T v) \\ (u + Ax - b)^+ - u \\ (Ex - f) \end{bmatrix}. \tag{۱۱.۲}$$

مدل شبکه عصبی متناظر برای حل (۶.۲) و(۷.۲) به‌صورت زیر بیان شده است:

$$\begin{cases} \frac{dy}{dt} = \kappa U(y), \\ y(t_۰) = y_۰. \end{cases} \tag{۱۲.۲}$$

مدل هشتم ([۲۴])

مسأله برنامه ریزی درجه دوم (۶.۲) را در نظر بگیرید. دوگان و شرایط بهینگی برای آن به ترتیب در فرم (۷.۲) و (۸.۲) بیان شده اند که در آن $Q \in \mathbb{R}^{n \times n}$ یک ماتریس نیمه معین مثبت است. برای مسائل (۶.۲) و (۷.۲) یک مدل شبکه عصبی به‌صورت زیر ارائه می‌شود:

$$\frac{dy}{dt} = (I + M^T) \begin{bmatrix} -(Qx + D + A^T u + E^T v) \\ (u + Ax - b)^+ - u \\ (Ex - f) \end{bmatrix}, \tag{۱۳.۲}$$

که در آن κ نرخ همگرایی است و

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

مدل نهم ([۲۵])

مسأله برنامه‌ریزی درجه دوم زیر را در نظر بگیرید:

$$\begin{cases} \text{Minimize} \quad f(x) = \frac{1}{2} x^T A x + a^T x \\ \text{s.t.} \\ \quad Dx \le b, \\ \quad x \ge \circ. \end{cases} \tag{۱۴.۲}$$

دوگان آن عبارت است از:

$$\begin{cases} \text{Maximize} \quad \bar{f}(w) = b^T w - \frac{1}{2} x^T A x \\ \text{s.t.} \\ \quad \bar{D}^T w - Ax \le a, \\ \quad w \ge \circ. \end{cases} \tag{۱۵.۲}$$

که در آن $A \in \mathbb{R}^{n\times n}$ متقارن و نیمه معین مثبت است، $a \in \mathbb{R}^n$ و $b \in \mathbb{R}^m$، $D \in \mathbb{R}^{m\times n}$ است. با توجه به شرایط KKT برای مسائل برنامه‌ریزی محدب، $(x^{*T}, w^{*T})^T$ به‌ترتیب یک جواب بهینه برای مسائل (۱۴.۲) و (۱۵.۲) است اگر و فقط اگر $(x^{*T}, w^{*T})^T$ در شرایط $K.K.T$ زیر صدق کند:

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

مدل متناظر با (۱۴.۲) و (۱۵.۲) در [۲۵] به‌صورت زیر بیان شده است:

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

که در آن

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

M یک ماتریس نیمه معین مثبت است زیرا $u^T M u = x^T A x \ge \circ$. همچنین $B = \lambda I$ است که $\lambda > \circ$ و با افزایش λ سرعت همگرایی افزایش می‌یابد.

فصل ۳

بهینه‌سازی سبدسرمایه

۱.۳ مقدمه

در این فصل ابتدا مقدمه‌ای بر سرمایه‌گذاری سبد سرمایه داریم سپس مسأله بهینه‌سازی سبد سرمایه را بیان کرده و مروری بر کارهای انجام شده در این زمینه داریم. در ادامه چندین مدل جایگزین برای مدل مبنا را معرفی می‌کنیم. برای مسائل درجه دوم محدب مدل شبکه عصبی را معرفی کرده و به اثبات پایداری و همگرایی آن می‌پردازیم. در پایان کارایی مدل شبکه عصبی را با چندین مثال عددی از مسائل بهینه‌سازی سبد سرمایه نشان می‌دهیم.

بحث سرمایه‌گذاری در اوراق بهادار و تحلیل آن را می‌توان به دو چارچوب کلی و متفاوت تقسیم کرد([۷]):

۱. تحلیل و انتخاب دارایی به‌طور جداگانه با استفاده از ابزارها و روشهای تحلیل تکنیکی [۱] و تحلیل بنیادی [۲].

۲. تشکیل سبد دارایی‌ها با استفاده از نظریه نوین سبد سرمایه [۳](MPT) و سایر تئوریهای بازار سرمایه.

مسأله انتخاب مجموعه‌ی بهینه‌ای از دارایی‌ها، یکی از مسائل بازار سرمایه است که اهمیت زیادی برخوردار است. از حوزه اقتصاد خرد، اهمیت تصمیمات سرمایه‌گذاری ناشی از این مسأله است که در واقع فرد سرمایه‌گذار، مصرف امروز را با امید مصرف بیشتر به زمانی در آینده موکول می‌کند. در واقع تصمیم بهینه سرمایه‌گذاری، میزان مطلوبیت هر فرد است که، با توجه به ترجیحات شخصی وی تعیین می‌گردد و لزوما با سایر افراد یکسان نخواهد بود، ولی می‌توان با معیارهایی مطلوبیت فرد را از انتخاب مجموعه‌ای از دارایی‌های سرمایه‌ای، مشخص کرد. از جمله مهمترین این معیارها، ریسک و بازده کل مجموعه انتخاب

[۱] Technical Analysis
[۲] Fundamental Analysis
[۳] Portfolio

شده است. هدف مدیریت مجموعه دارایی‌ها به‌طور عام و مجموعه سهام (سبد سهام) به‌طور خاص، تعیین دارایی‌های درون سبد بگونه‌ای که ریسک حداقل و بازده حداکثر شود. هر چه ریسک سبد سهام بیشتر باشد، احتمال دریافت بازده بالاتر هم بیشتر است.

در دنیای واقعی با توجه به‌اینکه درجه ریسک‌پذیری افراد با یکدیگر متفاوت است، بازده دارایی‌ها نیز به‌دلیل وجود عوامل متعدد مؤثر برآن، غیر قابل پیش‌بینی خواهد بود. به دلیل وجود این عدم اطمینان در بازار، مسأله تنوع بخشی در دارایی‌ها از اهمیت ویژه‌ای برخوردار می‌گردد. طبقات مختلف دارایی شامل انواع اوراق بهادار،پول نقد،زمین، طلاو...هستند که ریسک و بازده هر یک از آنها با هم متفاوت است.

تئوریهای سرمایه‌گذاری در چند دهه اخیر از پیشرفتهای زیادی برخوردار بودند و در سیر تاریخی تکاملی خود، به فرمولهای کاربردی زیادی دست یادفتند. حجم تجارت و سرمایه گذاری طی قرن بیستم و اوایل بیست ویکم گسترش فراوانی داشته است. این گسترش و تغییرات حاصل از آن، موجب شده است که معیارهای متفاوتی برای اخذ تصمیم سرمایه‌گذار در مقایسه با دوره‌های گذشته به‌کار گرفته شود. از عوامل مؤثر در انتخاب و انجام سرمایه‌گذاری، توجه سرمایه‌گذار به مفهوم ریسک و بازده و رابطه این دو با یکدیگر است. در قرن ۱۸ میلادی برنولی و کرامر[۴] به این نتیجه رسیدند که تصمیمات تحت شرایط عدم اطمینان نباید صرفا براساس بازده موردانتظار صورت پذیرد.

تا اوایل قرن بیستم میلادی، سرمایه‌گذاران جهت اخذ تصمیم در فرآیند سرمایه‌گذاری از نسبتهای بازده سرمایه‌گذاری استفاده می‌کردند و توجهی به مفاهیم ارزش زمانی پول و ریسک سرمایه‌گذاری نداشتند. از دهه ۱۹۲۰ مفهوم ارزش زمانی پول با روشهای تنزیلی [۵] وارد حوزه ادبیات مالی و سرمایه‌گذاری شد. این روشها، تحولی قابل توجه در انتخاب طرحهای سرمایه‌گذاری به‌وجود آوردند. لیکن همچنان رفتار متفاوت سرمایه‌گذاران در مواجه با ریسک نادیده گرفته می‌شد. در واقع با وجود اینکه نظریه مطلوبیت پول تا

[۴] Bernoulli and Cramer
[۵] Discounting Methods

حدودی به تکامل معیارهای انتخاب کمک نموده بود، لیکن هنوز از جامعیت کافی برخوردار نبودند. تا دهه ۱۹۵۰، ریسک یک عامل کیفی به‌شمار می‌رفت تا اینکه هری مارکویتز برای نخستین بار ریسک را کمیت‌پذیر نمود و انحراف معیار جریانات نقدی طرحهای سرمایه‌گذاری را به‌عنوان کمیت سنجش ریسک معرفی کرد. چندی بعد ویلیام شارپ در [۲۶] با تبیین ضریب حساسیت بتا (β) به‌عنوان معیار ریسک، مدل ساده‌تر و کاربردی‌تری را به دنیای تئوریهای سرمایه‌گذاری عرضه کرد. این روش امروز به‌عنوان مدل تک شاخصی معروف است. در ادامه این روند و در اواسط دهه ۶۰ شارپ و لینتر [۲۷] بر پایه تئوری بازار سرمایه، مدلی را توسعه دادند که امروزه تحت عنوان مدل قیمت‌گذاری دارایی‌های سرمایه‌ای[۶] شناخته می‌شود. این مدل ریسک سیستماتیک و غیرسیستماتیک را به‌عنوان اجزای اصلی ریسک از یکدیگر تفکیک می‌کند. در دهه ۷۰ ، پروفسور راس، مدل آربیتراژ [۷]را در [۲۸] پایه‌گذاری کرد. (برای توضیحات بیشتر به [۳۰] ،[۲۹] مراجعه شود).

۲.۳ مسأله بهینه‌سازی سبد سرمایه

در سال۱۹۵۲ میلادی، هری مارکوویتز مدل بنیادی سبد سرمایه را ارائه داد که مبنایی بر تئوری مدرن سبد سرمایه گردید. پیش از ارائه این تئوری سرمایه گذاران علیرغم آشنایی با مفاهیم ریسک و بازده ، قادر به اندازه‌گیری آن به روش کمی نبودند. سرمایه گذاران ضمن ایجاد تنوع، قرار ندادن همه تخم مرغها را در یک سبد را نیز قبول داشتند. مارکوویتز نخستین کسی بود که مفهوم تنوع بخشی در سبد سرمایه گذاری به‌طور عام و سبد سهام به‌طور خاص را بسط و گسترش داد. او این موضوع را که چگونه متنوع سازی سبد سرمایه می‌تواند منجر به کاهش ریسک سرمایه‌گذاری یک سرمایه‌گذار شود را به‌صورت کمی بیان نمود.

تئوری مدرن سبد سرمایه، یک نگرش کل‌گرا به بازار سهام است. این نظریه بر خلاف روشهای نموداری

[۶] Capital Asset Pricing Model(CAPM)
[۷] Arbitrage Pricing Theory(APT)

(تکنیکال) و بنیادی ، به مجموعه‌ی سهام در سبد توجه دارد. مارکوویتز در فرمول‌بندی مدل میانگین-واریانس[۸] خود، به هدف سرمایه‌گذاری توجه خاصی داشت. به نظر وی، سرمایه‌گذار عاقل به دنبال سرمایه‌گذاری در طرحهایی است که بازدهی بیشتر و ریسک کمتر داشته باشد. وی ریسک یک سرمایه‌گذاری را تنها در انحراف معیار آن طرح جستجو نمی‌کرد، بلکه به رابطه بین دارایی‌های مختلف در سبد سرمایه و تأثیر این رابطه بر ریسک کل هم توجه داشت.

یکی دیگر از مفاهیمی که مارکوویتز مطرح نمود، سبد سرمایه کارا (سبد سهام کارا)[۹] بود، که به‌معنای ترکیب مطلوب اوراق بهادار است به‌نحوی که ریسک آن سبد سرمایه در ازای بازده معینی حداقل شود یا بازده سبد سرمایه در ازای ریسک معین حداکثر شود. به این ترتیب، مجموعه این سبد سرمایه‌ها با عنوان مجموعه کارا شناخته شده و توسط حل مسأله پارامتریک غیرخطی قابل شناسایی است. در نمودار ریسک-بازدهی مجموعه کارا، مرز کارا نیز نامیده می‌شود.

مرز کارا مجموعه تعداد زیادی سبد سرمایه سهام ارائه می‌دهد که انتخاب هر کدام از این سبد سرمایه‌ها به منحنی مطلوبیت [۱۰] سرمایه‌گذار بستگی دارد. هر سرمایه‌گذار پورتفوئی که در نقطه مماس شدن منحنی مرز کارای مارکوویتز و منحنی مطلوبیت وی قرار دارد را به‌عنوان بهترین سبد سرمایه انتخاب خواهد کرد. شکل(۱.۳) مرز کارای مارکوویتز و منحنی مطلوبیت سرمایه‌گذار را نشان می‌دهد.

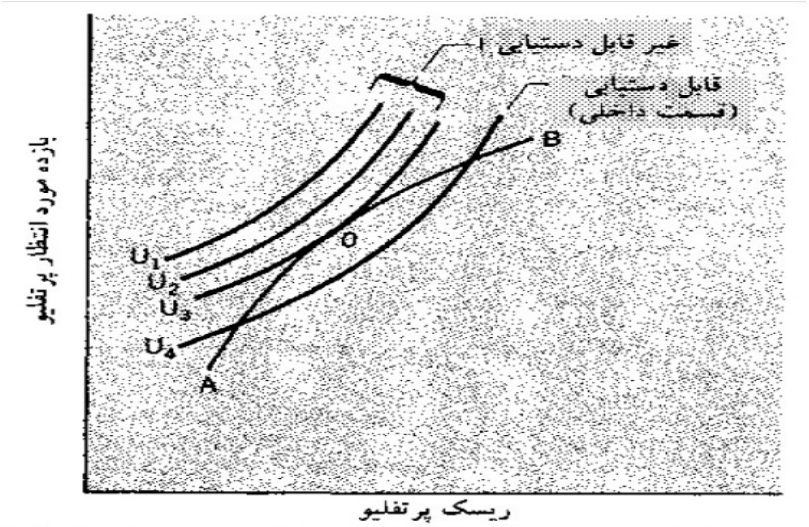
مدل مارکوویتز براساس مفروضاتی شکل گرفته است. این مفروضات عبارتند از:

۱. سرمایه‌گذار دارای رفتار عقلایی است.

۲. سرمایه‌گذاران، ریسک‌گریزند و دارای مطلوبیت مورد انتظار افزایشی هستند.

۳. سرمایه‌گذاران سبد سرمایه خود را بر مبنای میانگین و واریانس مورد انتظار بازدهی انتخاب می‌کنند.

[۸] Mean-Variance
[۹] Efficient Portfolio
[۱۰] Utility Curve



شکل ۱.۳: مرز کارایی و منحنی مطلوبیت

بنابراین منحنیهای بی‌تفاوتی آنها تابعی از نرخ بازده و واریانس مورد انتظار است.

۴. هر گزینه سرمایه‌گذاری تا بی‌نهایت قابل تقسیم است.

۵. سرمایه‌گذاران افق زمانی یک دوره‌ای داشته و این برای همه مشابه است.

۶. سرمایه‌گذاران در یک سطح مشخص از ریسک، بازده بالاتری را ترجیح می‌دهند و بلعکس برای سطح معینی از بازدهی، خواهان کمترین ریسک هستند.

۷. بازارها کامل هستند (هزینه مالیات و معاملات وجود ندارد).

در رویکرد مارکوویتز بازده دارایی‌ها به‌صورت متغیر تصادفی در نظر گرفته می‌شود که احتمال آنها از تابع توزیع نرمال به‌دست می‌آید. میانگین به‌عنوان معیار کارایی سبد سرمایه و انحراف معیار [۱۱] هم به‌عنوان شاخص ریسک سبد سرمایه در نظر گرفته می‌شود. ورودیهای اصلی مدل مارکوویتز عبارتند از بازده مورد انتظار ، ماتریس واریانس-کواریانس بازدهی دارایی‌ها و ضریب همبستگی بین بازدهی‌ها. مدل کلاسیک

[۱۱] Standard Deviation

MV به‌صورت زیر بیان می‌شود([۳۱]):

$$\text{Minimize} \quad Z_{Qp} = \frac{۱}{۲} X^T \Sigma X \tag{۱.۳}$$

subject to

$$\begin{cases} \mu^T X = R, \\ \sum_{i=۱}^{n} x_i = ۱, \\ x_i \geq ۰, \quad i = ۱, ۲, ..., n, \end{cases} \tag{۲.۳}$$

که در آن

$$E[X] = x_۱\mu_۱ + ... + x_n\mu_n = \mu^T X,$$

$$V[X] = \sum_{i,j} \rho_{ij}\sigma_i\sigma_j x_i x_j = X^T \Sigma X,$$

وقتی که ρ_{ij} ضریب همبستگی بین سهم i ام و j ام، μ_i بازده مورد انتظار سهم i ام است و قید اول نشان دهنده سطح مشخصی از بازدهی مورد انتظار یعنیR، قید دوم این اطمینان را می‌دهد که تمام بودجه سرمایه‌گذاری شود و قید آخر هم نشان‌دهنده این است که سرمایه‌گذار اجازه فروش استقراضی [۱۲] را ندارد. متأسفانه مدل مارکوویتز یک سری معایب و فرضیات غیرواقعی دارد که از جمله می‌توان به‌موارد زیر اشاره کرد:

۱. با افزایش تعداد دارایی، حجم محاسبات ماتریس کواریانس بیش از اندازه بزرگ می‌شود.

۲. محدویتهایی که در دنیای واقعی وجود دارد، در مدل مارکوویتز در نظر گرفته نشده است، همانند محدودیت سهام ثابت در سبد سرمایه[۱۳]، هزینه‌های معاملاتی[۱۴]، خرید آستانه‌ای [۱۵] ، معاملات بلوکی [۱۶] ، معادلات سهام در دسته‌های ثابت [۱۷] و ... ، اکثر این محدودیتها از توابع غیرخطی پیروی می‌کنند که مدل‌ها را برای حل دچار پیچیدگی می‌نماید.

[۱۲] Short Selling
[۱۳] Cardinality Constraint
[۱۴] Transaction Cost
[۱۵] Buy-in Thresholdy
[۱۶] Roundlots
[۱۷] Transaction Lot

۳. فرض نرمال بودن تابع توزیع بازده دارایی‌ها فرض قابل قبولی نیست.

۴. معیار عمومی ریسک، واریانس و یا انحراف معیار است که این معیار برای یک دارایی با توزیع نرمال و در بازار کارا قابل قبول است. اگر این دو خصوصیت برای دارایی وجود نداشته باشد، واریانس معیار مناسبی نیست.

۳.۳ مدل‌های جایگزین

در سال ۱۹۵۹ مارکویتز [۱۸] به مزایای ریسک نامطلوب اشاره کرد، وی در [۳۲] به این نتیجه رسید که سرمایه‌گذاران به دو علت مایل به حداقل ساختن ریسک نامطلوب هستند، نخست اینکه سرمایه‌گذاران ابتدا به امنیت اصل سرمایه می‌اندیشند و دوم اینکه اگر توزیع متغیرهای تصادفی (نرخ بازدهی) از نوع نرمال نباشد، آنگاه معیار ریسک نامطلوب مفید خواهد بود. بعدها در سال ۱۹۹۱ وی هنگام دریافت جایزه نوبل اقتصاد برای ارائه نظریه مدرن سبد سرمایه و مدل (MV) نیز، به این مسأله اشاره داشت.

سال ۱۹۷۳ لی و لرو [۱۹] در [۳۳]، مدلی به‌منظور بهینه‌سازی سبدسرمایه با به کارگیری برنامه‌ریزی آرمانی ارائه نمودند. چندی بعد در سال ۱۹۸۰ لی و چیسیر [۲۰] در [۳۴]، نگرش برنامه‌ریزی آرمانی چند معیاره [۲۱] را مطرح کردند. کونو و یامازاکی در سال ۱۹۹۱، مدلی از طریق برنامه‌ریزی خطی برای بهینه‌سازی سبد سرمایه‌گذاری ارائه کردند. پس از آن، در سال ۱۹۹۷ ، اسپرانزا و مانسینی [۲۲] در [۳۵]، مدلی از برنامه‌ریزی صحیح آمیخته [۲۳] را با خصوصیات واقعی مثل هزینه‌های معاملات و حداقل واحدهای معاملات ارائه داد و برای حل آن از الگوریتم‌های ابتکاری کمک گرفتند و مدل را برای بازار سهام میلان آزمایش کردند. سال ۱۹۹۸ یونگ [۲۴] در [۳۶]، مدل بهینه‌سازی سبد سرمایه قابل حل توسط برنامه‌ریزی خطی را که در آن

[۱۸] Markowitz
[۱۹] Lee and Lerro
[۲۰] Lee and Chesser
[۲۱] Multi-criteria goal programming
[۲۲] Mansini and Speranza
[۲۳] Mixed integer programming
[۲۴] Young

ریسک بر مبنای بدترین حالت تعریف شده و آن را رویکرد "حداقل نمودن حداکثر" [۲۵] نامید. سال ۱۹۹۹ اکسیا و همکاران در [۳۷] مدلی خاص را معرفی کردند. در فوریه همین سال چانک [۲۶] و همکاران، در [۳۸] به ارائه مدلی پارامتریک و غیرخطی جهت بهینه‌سازی سهام پرداختند. در ادامه در نوامبر همین سال، جوناس پالمکویست [۲۷] و همکارانش در [۳۹] به‌معرفی بهینه‌سازی سبد سرمایه با محدودیت‌ها و تابع هدف ارزش درمعرض ریسک شرطی [۲۸] ، مدلی از نوع برنامه‌ریزی خطی ارائه دادند. در سال ۲۰۰۲ لوبو [۲۹] و همکاران در [۴۰] به بهینه‌سازی سبد با اعمال هزینه‌های معاملات خطی و ثابت پرداختند. در همین سال، کریستوس پاپاریستودولو [۳۰] در [۴۱] به ارائه مدل‌های موجود در این زمینه پرداخت و سپس با طرح مسأله ای متشکل از ۵ سهام در دوره زمانی ۱۲ ماهه به مقایسه سبدهای به‌دست آمده در هر مدل پرداخت. در ژانویه ۲۰۰۳ آلکسی کخلاو [۳۱] و همکاران در مقاله‌ای تحت عنوان "بهینه‌سازی سبد سرمایه با محدودیت‌های کاهش دهنده" [۳۲] به ارائه مدلی در این زمینه پرداخت. در ژولای ۲۰۰۳ مانسینی [۳۳] و همکارانش در [۴۲] به‌معرفی مدل‌های LP پرداخته‌اند. در سال ۲۰۰۴ گاندریو [۳۴] و همکاران برای حل مسائل بهینه‌سازی غیرخطی سبد سرمایه ، روش نقطه درونی [۳۵] اولیه-دوگان را معرفی کردند. در ادامه به بیان چندین مدل از مدل‌های نامبرده می‌پردازیم.

[۲۵] Minimax
[۲۶] Chang
[۲۷] Palmquist
[۲۸] Conditional Value at Risk
[۲۹] Lobo
[۳۰] Papahristodoulou
[۳۱] Alexi Chekhlov
[۳۲] Draw Down
[۳۳] Mansini
[۳۴] Gandryv
[۳۵] Interior-point method

۱.۳.۳ مدل‌های درجه دوم

فرناندز و گومز [۳۶] در [۴۳] ، مدل اصلاح یافته مارکویتز را تحت عنوان"مدل میانگین-واریانس با مؤلفه‌های محدود [۳۷] " معرفی کردند که شکل عمومی آن به‌صورت زیر است:

$$\text{Minimize} \quad Z = \frac{۱}{۲} X^T \Sigma X \tag{۳.۳}$$

subject to

$$\begin{cases} \mu^T X = R, \\ \sum_{i=۱}^{n} x_i = ۱, \\ \varepsilon_i \leq x_i \leq \delta_i, \quad i = ۱, ۲, ..., n. \end{cases} \tag{۴.۳}$$

در صورتی‌که محدودیت مربوط به دارایی‌های منتخب به مسأله فوق اضافه شود، مدل مربوطه به‌صورت زیر در می‌آید:

$$\text{Minimize} \quad \lambda[\sum_{i=۱}^{n}\sum_{j=۱}^{n} z_i x_i z_j x_j \rho_{ij}] - (۱ - \lambda)[\sum_{i=۱}^{n} z_i x_i \mu_i] \tag{۵.۳}$$

subject to

$$\begin{cases} \sum_{i=۱}^{n} x_i = ۱, \\ \sum_{i=۱}^{n} z_i = k, \\ \varepsilon_i z_i \leq x_i \leq \delta_i z_i, \\ z_i \in [۰, ۱], \\ x_i \geq ۰, \quad i = ۱, ۲, ..., n. \end{cases} \tag{۶.۳}$$

که در آن ρ_{ij} ضریب همبستگی بین سهم i ام و j ام، μ_i بازده مورد انتظار سهم i ام است و $\lambda \in [۰, ۱]$ و z_i متغیر تصمیم در مورد سرمایه‌گذاری در سهم i ام است. مجموع تعداد سهامی که در سبد خواهد بود بنا به محدودیت دوم، k تا خواهد بود. برای حل این مسأله از الگوریتم های ابتکاری مثل تکنیک بهینه سازی جمعی ذرات [۳۸] کمک گرفته شده است.

[۳۶] Fernandez and Gomez
[۳۷] Cardinality Constrained Mean-Variance(CCMV)
[۳۸] Particle Swarm Optimization(PSO)

مدلهایی با فرم درجه دوم قطری

تجزیه چولسکی [۴۴] ماتریس کواریانس V به‌صورت زیر بیان می‌شود:

$$V = L^T L$$

که درآن $L_{n\times n}$ یک ماتریس پایین مثلثی است، با تعریف بردار جدید $Y_{n\times ۱}$ به‌صورت $Y = LX$ با عناصر،

$$y_i = \sum_{j=۱}^{i} l_{ij} x_j \quad i = ۱, ..., n,$$

فرمول‌بندی یکسان از مسأله مارکوویتز به‌دست می‌آید. بنابراین مدل قطری به‌صورت زیر بیان می‌شود:

$$\text{Minimize} \quad Z_{diag۱} = \sum_{i=۱}^{n} y_i^۲ \tag{۷.۳}$$

subject to

$$\begin{cases} y_i = \sum_{j=۱}^{i} l_{ij} x_j \quad i = ۱, ۲, ..., n, \\ \sum_{i=۱}^{n} x_i \mu_i = R, \\ \sum_{i=۱}^{n} x_i = ۱, \\ x_i \geq ۰, \qquad i = ۱, ۲, ..., n, \\ g_i^y \leq y_i \leq h_i^y \qquad i = ۱, ۲, ..., n. \end{cases} \tag{۸.۳}$$

برای فرم درجه دوم کلی، y_i متغیرهای آزاد $(-\infty < y_i < +\infty)$ هستند.

رویکرد مشابه در تجزیه ماتریس کواریانس V ، بازدهی‌های R مشاهده شده در کل دوره T را نیز در این تجزیه اعمال کنیم. اگر $\bar{R}$ را ماتریس میانگین بازدهی‌ها در نظر بگیریم آنگاه ماتریس V به‌صورت زیر محاسبه می‌شود:

$$V = \frac{۱}{n-۱}(R - \bar{R})^T (R - \bar{R}).$$

با تعریف $S_{T\times n} = \frac{۱}{\sqrt{n-۱}}(R - \bar{R})$ ، ماتریس V به‌صورت $V = S^T S$ تعریف می‌شود، بنابراین مدل قطری دوم به‌صورت زیر بیان می‌شود:

$$\text{Minimize} \quad Z_{diag۲} = \sum_{t=۱}^{T} y_t^۲ \tag{۹.۳}$$

subject to

$$\begin{cases} y_t = \sum_{i=۱}^{n} s_{ti}x_i \quad t = ۱, ۲, ..., T, \\ \sum_{i=۱}^{n} x_i\mu_i = R, \\ \sum_{i=۱}^{n} x_i = ۱, \\ x_i \geq ۰, \qquad i = ۱, ۲, ..., n, \\ g_t^y \leq y_t \leq h_t^y \qquad t = ۱, ۲, ..., T. \end{cases} \tag{۱۰.۳}$$

۲.۳.۳ مدل‌های خطی

مدل میانگین قدرمطلق انحرافات [۳۹]

بعد از مارکویتز، شارپ در سال ۱۹۷۱ تلاش فراوانی کرد تا مدل خطی از بهینه‌سازی سبد سرمایه ارائه دهد، بعد از آن کونو [۴۰] در [۴۵]، مدلی را با استفاده از توابع قطعه‌ای خطی پیشنهاد کرد. در مدل MAD ریسک به جای واریانس توسط میانگین قدرمطلق انحرافات اندازه‌گیری می‌شود. هم چنین آنها در [۴۶] نشان دادند که مدل مارکویتز تحت این فرض که بازدهی‌ها دارای توزیع نرمال چند متغیره باشند، معادل مدل آنهاست. فرض کنید m_t انحرافات مطلق بازدهی سبد سرمایه از میانگین آن در زمان t باشد، بنابراین مدل را به‌صورت زیر بیان می‌کنیم:

$$\text{Minimize} \qquad Z_{mad} = \frac{۱}{T}\sum_{t=۱}^{T} m_t \tag{۱۱.۳}$$

subject to

$$\begin{cases} \sum_{i=۱}^{n}(r_{it} - \mu_i)x_i \leq m_t \quad t = ۱, ۲, ..., T, \\ \sum_{i=۱}^{n}(r_{it} - \mu_i)x_i \geq -m_t \quad t = ۱, ۲, ..., T, \\ \sum_{i=۱}^{n} x_i\mu_i = R, \\ \sum_{i=۱}^{n} x_i = ۱, \\ m_t \geq ۰, \qquad t = ۱, ۲, ..., T, \\ x_i \geq ۰, \qquad i = ۱, ۲, ..., n. \end{cases} \tag{۱۲.۳}$$

سیمن [۴۱] در [۴۷] ، درباره معایب و مزایا دو مدل MV و MAD بحث کرده است.

[۳۹] Mean-absolute Divation(MAD)
[۴۰] Konno
[۴۱] Simaan

مدل Maxmin

یانگ در سال ۱۹۹۸ در [۳۶] برای مسأله انتخاب سبد سرمایه، از حداقل بازدهی که در دوره‌های گذشته حاصل شده است به‌عنوان تخمین ریسک استفاده می‌کنند. به‌عبارتی دیگر به دنبال حداقل کردن، حداکثر زیانی که در طول دوره ممکن است رخ دهد، هستیم. اکنون متغیر M_p را که نشان‌دهنده حداقل بازدهی بدست آمده از سبد سرمایه را به‌فرم $M_p = min_t \sum_{i=۱}^{n} r_{it}x_i$ معرفی می‌کنیم، بنابراین مدل به‌صورت زیر قابل بیان است:

$$\text{Maximize} \quad Z_{mm} = M_p \tag{۱۳.۳}$$

subject to

$$\begin{cases} \sum_{i=۱}^{n} r_{it}x_i \geq M_p \quad t = ۱, ۲, ..., T, \\ \sum_{i=۱}^{n} x_i\mu_i = R, \\ \sum_{i=۱}^{n} x_i = ۱, \\ x_i \geq ۰, \quad i = ۱, ۲, ..., n, \\ l^{M_p} \leq M_p \leq u^{M_p}. \end{cases} \tag{۱۴.۳}$$

مدل Minimax

کای [۴۲] و همکاران در [۴۸]، مدلی جدید ریسک را بر پایه MAD ارائه کردند، قبل از معرفی مدل یک تعریف ارائه می‌دهیم:

تعریف ۱.۳.۳. مدل L_∞، ماکسیمم انحراف مطلق ریسک به‌صورت زیر است:

$$L_\infty(x) = \max_{۱\leq j\leq n} E(\mid r_j x_j - \mu_j x_j \mid).$$

بنابراین مدل $Minimax$ به‌صورت زیر بیان می‌شود:

$$\text{Minimize} \quad Z_{mm} = l_\infty(x) = max_{۱\leq j\leq n} E(|r_j x_j - \mu_j x_j|) \tag{۱۵.۳}$$

[۴۲] Cai

subject to

$$\begin{cases} \sum_{j=۱}^{n} x_j \mu_j = R, \\ \sum_{j=۱}^{n} x_j = ۱, \\ x_j \geq ۰, \quad j = ۱, ۲, ..., n. \end{cases} \tag{۱۶.۳}$$

این مدل را می‌توان به‌فرم خطی زیر تبدیل کرد:([۴۹])

$$\text{Minimize} \quad y \tag{۱۷.۳}$$

subject to

$$\begin{cases} q_j x_j \leq y, \\ \sum_{j=۱}^{n} x_j \mu_j = R, \\ \sum_{j=۱}^{n} x_j = ۱, \\ x_j \geq ۰, \quad j = ۱, ۲, ..., n. \end{cases} \tag{۱۸.۳}$$

که‌در آن $q_j = E(|r_j - \mu_j|)$ و $y = \max q_j$ برای $(j = ۱, ۲, ..., n)$ است.

همچنین، جایگزین دیگر تابع ریسک L_∞ را که توسط تئو [۴۳] و یانگ در [۵۰] پیشنهاد شد، به‌صورت زیر است: که درآن r_{jt} متغیرهای تصادفی و μ_{jt} مقادیر مورد انتظار از r_{jt} برای $(i = ۱, ۲, ..., n)$ و $(t = ۱, ۲, ..., T)$ هستند. این تابع توسعه یافته L_∞ است. بنابراین مدل دیگر از **Minimax** به صورت زیر ارائه شده است:

$$\text{Minimize} \quad Z_{mm} = H_\infty^T(x) = \frac{۱}{T} \sum_{t=۱}^{T} \max_{۱ \leq j \leq n} E(|r_{jt} x_j - \mu_{jt} x_j|) \tag{۱۹.۳}$$

subject to

$$\begin{cases} \sum_{j=۱}^{n} x_j \mu_j = R, \\ \sum_{j=۱}^{n} x_j = ۱, \\ x_j \geq ۰, \quad j = ۱, ۲, ..., n. \end{cases} \tag{۲۰.۳}$$

این مدل را نیز، می‌توان به‌فرم خطی زیر تبدیل کرد:([۴۹])

$$\text{Minimize} \quad \frac{۱}{T} \sum_{t=۱}^{T} y_t \tag{۲۱.۳}$$

subject to

[۴۳] Teo

$$\begin{cases} a_{jt}x_j \leq y_t, \quad i = ۱, ۲, ..., n \ , t = ۱, ۲, ..., T, \\ \sum_{j=۱}^{n} x_j \mu_j = R, \\ \sum_{j=۱}^{n} x_j = ۱, \\ x_j \geq ۰, \quad j = ۱, ۲, ..., n. \end{cases} \tag{۲۲.۳}$$

که‌در آن $a_{jt} = E|r_{jt} - \mu_{jt}|$، برای $(i = ۱, ۲, ..., n)(t = ۱, ۲, ..., T)$ هستند. برای یافتن مدلهای بیشتر مراجع [۳۲، ۵۱، ۵۲] را ببینید.

۴.۳ مدل شبکه عصبی

با توجه به تحلیل‌های بالا، واضح است که بیشتر مسائل بهینه سازی سبد سرمایه به‌صورت برنامه‌ریزی خطی و درجه دوم هستند، بنابراین ما فرم کلی مسائل برنامه‌ریزی درجه دوم را به‌صورت زیر در نظر می‌گیریم:

$$\text{Minimize} \quad \frac{۱}{۲} x^T Q x + D^T x \tag{۲۳.۳}$$

$$\text{subject to} \quad \begin{cases} Ax - b \leq ۰, \\ Ex - f = ۰. \end{cases} \tag{۲۴.۳}$$

که درآن $Q \in \mathbb{R}^{n \times n}$ ماتریس متقارن و نیمه معین مثبت، $A \in \mathbb{R}^{m \times n}$ ، $b \in \mathbb{R}^n$، $E \in \mathbb{R}^{l \times n}$ ، $x \in \mathbb{R}^n$ و $rank(A, E) = m + l$. با استفاده از تکنیک‌های استاندارد بهینه‌سازی، مسأله (۲۳.۳)-(۲۴.۳) را به سیستم دینامیکی غیرخطی تبدیل می‌کنیم.

با توجه به قضایای ۱۲.۲.۱ و ۱۳.۲.۱ ، x^* جواب بهینه (۲۳.۳)-(۲۴.۳) است اگر و تنها اگر بردارهای $u^* \in \mathbb{R}^m$ و $v^* \in \mathbb{R}^l$ موجود باشند به‌طوری‌که $(x^{*T}, u^{*T}, v^{*T})^T$ در سیستم $K.K.T$ زیر صدق کند:

$$\begin{cases} Qx^* + D^T + A^T u^* + E^T v^* = ۰ \\ u^* \geq ۰, \ Ax^* - b \leq ۰, \ u^{*T}(Ax^* - b) = ۰, \\ Ex^* - f = ۰ \end{cases} \tag{۲۵.۳}$$

x^* را یک نقطه $K.K.T$ برای (۲۳.۳) و (۲۴.۳) می‌گویند. بنابراین x^* جواب بهینه (۲۳.۳)-(۲۴.۳) است اگر و تنها اگر x^* یک نقطه $K.K.T$، (۲۳.۳)-(۲۴.۳) باشد.

اکنون فرض کنید که $x(.)$ ، $u(.)$ و $v(.)$ متغیرهای وابسته به زمان هستند. هدف بیان یک ساختار دینامیکی پیوسته‌ای است که به نقطه $K.K.T$ مسأله (۲۳.۳)-(۲۴.۳) هدایت شود. ما مدل شبکه عصبی بازگشتی را با معادله دینامیکی به‌صورت زیر بیان می‌کنیم:

$$\begin{cases} \frac{dx}{dt} = -(Qx + D + A^T(u + Ax - b)^+ + E^T v), \\ \frac{du}{dt} = (u + Ax - b)^+ - u, \\ \frac{dv}{dt} = Ex - f. \end{cases} \tag{۲۶.۳}$$

با نقطه آغازین $(x_۰^T, u_۰^T, v_۰^T)^T$ ،که‌در آن

$$(u + Ax - b)^+ = ([(u + Ax - b)_۱]^+, [(u + Ax - b)_۲]^+, ..., [(u + Ax - b)_m]^+),$$

$$[(u + Ax - b)_k]^+ = \max\{(u + Ax - b)_k, ۰\},\ k = ۱, ۲, ..., m.$$

برای سادگی، تعریف می‌کنیم $y = (x^T, u^T, v^T)^T \in \mathbb{R}^{n+m+l}$ ، D^* مجموعه نقاط بهینه از (۲۳.۳)-(۲۴.۳) و دوگانش و

$$\zeta(y) = \begin{bmatrix} -(Qx + D + A^T(u + Ax - b)^+ + E^T v \\ (u + Ax - b)^+ - u \\ Ex - f \end{bmatrix} \tag{۲۷.۳}$$

بنابراین شبکه عصبی (۲۶.۳) را می‌توان به‌صورت زیر نوشت:

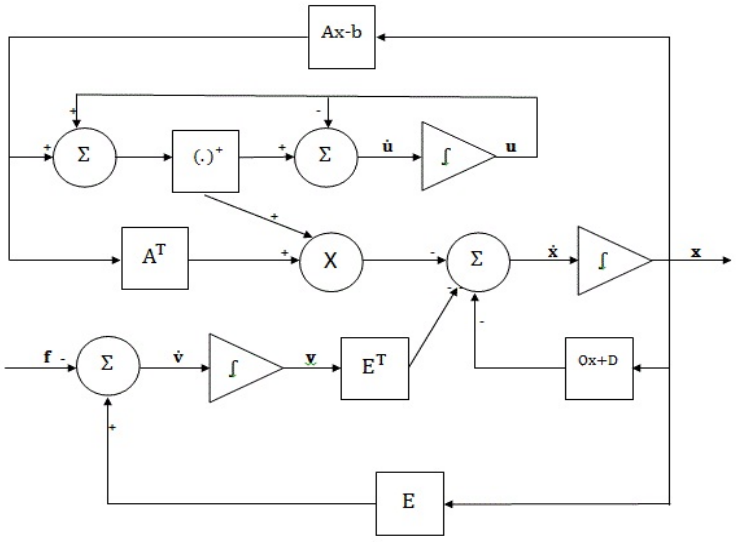
$$\frac{dy}{dt} = \kappa\zeta(y), \tag{۲۸.۳}$$

$$y(t_۰) = y_۰. \tag{۲۹.۳}$$

وقتی κ پارامتر اسکالر و بیان کننده نرخ همگرایی شبکه عصبی (۲۸.۳)-(۲۹.۳) است که برای سادگی تحلیلمان، κ را برابر ۱ در نظر می‌گیریم. شکل (۲.۳) بیانگر این است که مدل فوق روی سخت افزار چگونه اجرا می‌شود.

۵.۳ تحلیل پایداری و همگرایی

در این بخش پایداری و همگرایی شبکه عصبی پیشنهاد شده (۲۹.۳)-(۲۸.۳) بررسی می‌شود.



شکل ۲.۳: بلوک دیاگرام شبکه عصبی (۲۹.۳)-(۲۸.۳)

قضیه ۱.۵.۳. فرض کنید $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ ، نقطه تعادل از شبکه عصبی (۲۹.۳)-(۲۸.۳) باشد. آنگاه x^* نقطه $K.K.T$ از مسأله (۲۳.۳)-(۲۴.۳) است و برعکس، اگر $x^* \in \mathbb{R}^n$ جواب بهینه مسأله (۲۳.۳)-(۲۴.۳) باشد، آنگاه بردارهای $u^* \in \mathbb{R}^m$ و $v^* \in \mathbb{R}^l$ وجود دارد به‌طوری که $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ یک نقطه تعادل از شبکه عصبی پیشنهاد شده (۲۹.۳)-(۲۸.۳) است.

اثبات. فرض کنید $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ نقطه تعادل شبکه عصبی فوق باشد، آنگاه $\frac{dx^*}{dt} = ۰$، $\frac{du^*}{dt} = ۰$ و $\frac{dv^*}{dt} = ۰$. بنابراین

$$\begin{cases} Qx^* + D + A^T(u^* + Ax^* - b)^+ + E^T v^* = ۰, \\ (u^* + Ax^* - b)^+ - u^* = ۰, \\ Ex^* - f = ۰. \end{cases} \tag{۳۰.۳}$$

از طرفی، داریم $(u^* + Ax^* - b)^+ = u^*$ اگر و تنها اگر

$$u^* \geq ۰, \quad Ax^* - b \leq ۰, \quad u^{*T}(Ax^* - b) = ۰. \tag{۳۱.۳}$$

با جایگذاری (۳۱.۳) در (۳۰.۳) داریم:

$$Qx^* + D + A^T u^* + E^T v^* = ۰ \qquad (۳۲.۳)$$

از (۳۴.۳) و (۳۰.۳)، به‌نظر می‌رسد که $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ در شرایط $K.K.T$ صدق می‌کند و این اثبات را کامل می‌کند. □

لم ۲.۵.۳. برای هر نقطه آغازین $y(t_۰) = (x(t_۰)^T, u(t_۰)^T, v(t_۰)^T)^T$ ، سیستم (۲۸.۳)-(۲۹.۳) دارای جواب یکتای پیوسته $y(t) = (x(t)^T, u(t)^T, v(t)^T)^T$ است.

اثبات. واضح است که $Qx + D + A^T(u + Ax - b)^+ + E^T v$ ، $(u + Ax - b)^+ - u$ و $Ex - f$ لیپ شیتز محلی پیوسته روی مجموعه باز $D \subseteq \mathbb{R}^{n+m+l}$ هستند، با توجه به قضیه (۸.۳.۱) ، شبکه عصبی (۲۸.۳)-(۲۹.۳) دارای جواب یکتای پیوسته $y(t)$ برای $t \in [t_۰, \tau)$ وقتی $\tau \to \infty$ است. □

لم ۳.۵.۳. ماتریس ژاکوبی $\nabla\zeta(y)$ از نگاشت ζ تعریف شده در (۲۷.۳) ، ماتریس نیمه معین منفی است.

اثبات. بدون کاستن ازکلیت مسأله، فرض می‌کنیم $۰ < p < m$ موجود است به‌طوری که:

$$(u + Ax - b)^+ = \left((u + Ax - b)_۱, (u + Ax - b)_۲, ..., (u + Ax - b)_p, \underbrace{۰, ۰, ..., ۰}_{m-p} \right).$$

با محاسبات ساده می‌توان نشان داد:

$$\nabla\zeta(y) = \begin{bmatrix} -Q_{n\times n} - (A^p)^T A^p & -(A^p)^T & -E^T \\ A^p & \mathcal{T}_{m\times m} & O_{m\times l} \\ E & O_{l\times m} & O_{l\times l} \end{bmatrix},$$

که درآن

$$A^p = \begin{bmatrix} U_{p\times n} \\ O_{(m-p)\times n} \end{bmatrix} = \begin{bmatrix} A_{۱.} \\ A_{۲.} \\ ... \\ ... \\ A_{p.} \\ O_{۱\times n} \\ O_{۱\times n} \\ ... \\ ... \\ O_{۱\times n} \end{bmatrix},$$

و

$$\mathcal{T}_{m\times m}=\begin{bmatrix} O_{p\times p} & O_{p\times(m-p)} \\ O_{(m-p)\times p} & -I_{(m-p)\times(m-p)} \end{bmatrix}.$$

که O ماتریس صفر است و $(A^p)^TA^p$ ماتریس نیمه معین مثبت است. و از طرفی ماتریس $\mathcal{T}_{m\times m}$ ، نیمه معین منفی است. بنابراین ماتریس ژاکوبی $\nabla\zeta(y)$ ، نیمه معین منفی است.

اگر $p=m$ یعنی $(u+Ax-b)^+=((u+Ax-b)_۱,(u+Ax-b)_۲,...,(u+Ax-b)_m)$ ، آنگاه

$$\nabla\zeta(y)=\begin{bmatrix} -Q_{n\times n}-A^TA & -A^T & -E^T \\ A & O_{m\times m} & O_{m\times l} \\ E & O_{l\times m} & O_{l\times l} \end{bmatrix}.$$

مشابه حالت قبلی، می‌توان نشان داد که $\nabla\zeta(y)$ ماتریس نیمه معین منفی است. سرانجام، اگر $p=۰$ یعنی $(u+Ax-b)^+=(\underbrace{۰,۰,...,۰}_{m})$ ، آنگاه داریم:

$$\nabla\zeta(y)=\begin{bmatrix} O_{n\times n} & O_{n\times m} & -E^T \\ O_{m\times n} & -I_{m\times m} & O_{m\times l} \\ E & O_{l\times m} & O_{l\times l} \end{bmatrix}.$$

در این مورد هم، به‌سادگی می‌توان نشان داد که $\nabla\zeta(y)$ ماتریس نیمه معین منفی است، این اثبات را کامل می‌کند. □

قضیه ۴.۵.۳. اگر فرض لم (۳.۵.۳) برقرار باشد، آنگاه شبکه عصبی پیشنهاد شده در (۲۹.۳)-(۲۸.۳) ، پایدار مجانبی به مفهوم لیاپانوف و همگرایی سراسری به $y^*=(x^{*T},u^{*T},v^{*T})^T$ است که در آن x^* جواب بهینه (۲۳.۳)-(۲۴.۳) است.

اثبات. تابع لیاپانوف $(E:\mathbb{R}^{m+n+l}\to\mathbb{R})$ را به‌صورت زیر را در نظر بگیرید:

$$E(y)=\|\zeta(y)\|^۲+\frac{۱}{۲}\|y-y^*\|^۲. \qquad (۳۳.۳)$$

از (۲۷.۳) داریم:

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

با محاسبه مشتق $E(y(t))$ از شبکه عصبی (۲۹.۳)-(۲۸.۳) ، داریم:

$$\frac{dE(y(t))}{dt} = (\frac{d\zeta}{dt})^T\zeta + \zeta^T(\frac{d\zeta}{dt}) + (y-y^*)^T\frac{dy(t)}{dt} =$$

$$\zeta^T(\nabla\zeta(y)^T + \nabla\zeta(y))\zeta + (y-y^*)^T\zeta(y).$$

با به‌کار بردن لم (۳.۵.۳) ، داریم

$$\zeta^T(y)(\nabla\zeta(y)^T + \nabla\zeta(y))\zeta(y) \leq ۰, \quad \forall y \neq y^*. \tag{۳۴.۳}$$

همچنین با استفاده از تعریف (۶.۳.۱) و قضیه (۷.۳.۱) داریم:

$$(y-y^*)^T(\zeta(y)-\zeta(y^*)) = (y-y^*)^T\zeta(y) \leq ۰, \quad \forall y \neq y^*.$$

بنابراین

$$\frac{dE(y(t))}{dt} \leq ۰. \tag{۳۵.۳}$$

و این بدان معنی است که شبکه عصبی (۲۹.۳)-(۲۸.۳) ، پایدار مجانبی به مفهوم لیاپانوف است. در ادامه چون

$$E(y) \geq \frac{۱}{۲}\|y-y^*\|^۲ \tag{۳۶.۳}$$

زیر دنباله‌های همگرای

$$\{(x(t_k)^T, u(t_k)^T, v(t_k)^T)^T | t_۰ < t_۱ < ... < t_k < t_{k+۱}\}$$

وقتی $t_k \to \infty$ از $k \to \infty$ وجود دارد، به‌طوری‌که

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

وقتی $(\bar{x}^T, \bar{u}^T, \bar{v}^T)^T$ در رابطه زیر صدق می‌کند:

$$\frac{dE(y(t))}{dt} = ۰.$$

این نشان می‌دهد که $(\bar{x}^T, \bar{u}^T, \bar{v}^T)^T$ ، یک نقطه ω–حدی از $\{(x(t)^T, u(t)^T, v(t)^T)^T | t \geq t_0\}$ است (مرجع [۵۴] ،صفحه ۲۲۵ و پیوست [۵۵] را ببینید). با استفاده از قضیه مجموعه ناوردای لازال در (۴.۴.۱) ، در می‌یابیم $\{(x(t)^T, u(t)^T, v(t)^T)^T \to M\}$، وقتی‌که $t \to \infty$ و M، بزرگترین مجموعه ناوردا در $K = \{(x(t)^T, u(t)^T, v(t)^T)^T | \frac{dE(y(t))}{dt} = 0\}$ است. از (۲۹.۳)-(۲۸.۳) و (۳۵.۳) داریم، $M \subseteq K \subseteq D^*$ با $(\bar{x}^T, \bar{u}^T, \bar{v}^T)^T \in D^*$ بنابراین $\frac{dx}{dt} = 0$ $\frac{du}{dt} = 0$، $\frac{dv}{dt} = 0 \Leftrightarrow \frac{dE(y(t))}{dt} = 0$.

با جایگذاری $x^* = \bar{x}$ ، $u^* = \bar{u}$ و $v^* = \bar{v}$ در(۳۳.۳) ، ما تابع لیاپانوف دیگری، تعریف می‌کنیم:

$$\bar{E}(y) = \|\zeta(y)\|^2 + \frac{1}{2}\|y - \bar{y}\|^2. \tag{۳۷.۳}$$

آنگاه $\bar{E}(y)$ به‌طور پیوسته مشتق پذیر و $\bar{E}(\bar{y}) = 0$. توجه داریم که

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

از اینرو داریم $\lim_{k\to\infty} \bar{E}(x(t_k)^T, u(t_k)^T, v(t_k)^T)^T = \bar{E}(\bar{x}, \bar{u}, \bar{v})$.پس، $\forall \epsilon > 0$ ، $q > 0$ موجود است به‌طوری‌که برای هر $t \geq t_q$ داریم $\bar{E}(y(t)) < \epsilon$. مشابها ما می‌توانیم بدست آوریم $\frac{d\bar{E}(y(t))}{dt} \leq 0$. برای $t \geq t_q$ داریم،

$$\frac{1}{2}\|y(t) - \bar{y}\|^2 \leq \bar{E}(y(t)) \leq \epsilon.$$

در نتیجه، $\lim_{t\to\infty} \|y(t) - \bar{y}\| = 0$ و $\lim_{t\to\infty} y(t) = \bar{y}$. بنابراین، شبکه عصبی پیشنهاد شده (۲۹.۳)-(۲۸.۳) همگرای سراسری به نقطه تعادل $\bar{y} = (\bar{x}^T, \bar{u}^T, \bar{v}^T)^T$ است که در آن، $\bar{x}$ جواب بهینه (۲۴.۳)-(۲۳.۳) است. □

از قضیه (۴.۵.۳) می‌توانیم گزاره زیر را نتیجه بگیریم.

گزاره ۵.۵.۳. اگر فرضیات لم (۳.۵.۳) برقرار باشد و $D^* = \{(x^{*T}, u^{*T}, v^{*T})^T\}$ ، آنگاه شبکه عصبی (۲۹.۳)-(۲۸.۳) برای حل کردن (۲۴.۳)-(۲۳.۳) ، پایدار مجانبی سراسری به نقطه تعادل یکتای $y^* = (x^{*T}, u^{*T}, v^{*T})^T$ است.

۶.۳ مثال‌های عددی

به منظور نشان دادن کارایی و تأثیر شبکه عصبی پیشنهاد شده، در این بخش ما به ذکر۷ مثال از ۲ مسأله مختلف می‌پردازیم. برای برخی مثال‌ها، عملکرد شبکه عصبی را برای مقادیر مختلف $y(۰)$ بررسی می‌کنیم، همچنین ما مقایسه عملکرد عددی از شبکه عصبی را با مقادیر گوناگون κ و مقادیر مختلف $\|x(t)-x^*\|^۲$ با نقطه آغازین مشخص $y_۰$ داریم. نتایج شبیه‌سازی در پکیج $ode۴۵s$ نرم افزار $Matlab$ انجام شده و جواب واقعی مسائل نیز توسط نرم افزار $Lingo۱۱$ محاسبه شده است.

مسأله I

مدیر سبد سرمایه مؤسسه سیگما، قصد دارد یک سبد سرمایه بهینه برای یک مشتری انتخاب کند. به‌جای در نظر گرفتن تعداد زیاد دارایی‌ها در سبد، برای سادگی فرض می‌کنیم ۵ سهم (A,B,C,D,E) از بازار بورس استکهلم در اختیار باشد. داده‌های تاریخی ۱۲ ماهه، ماتریس واریانس و انحرافات مطلق هر ماه در[۴۱] داده شده است. مشتری خواهان سرمایه گذاری $۱۰۰۰۰۰SEK$، حداقل بازدهی دریافتی $(۳,۰۰۰SEK)$ ۳ % است و در نظر دارد، در هر سهم حداکثر ۰/۷۵% سرمایه گذاری داشته باشد. بنابراین مسأله فوق، حداقل کردن ریسک سبد، با در نظر گرفتن خواسته‌های مشتری و اجازه نداشتن فروش استقراضی در هرسهام است.

مثال ۱.۶.۳. مسأله برنامه‌ریزی درجه دوم به فرم (۱.۳)-(۲.۳) را در نظر بگیرید:

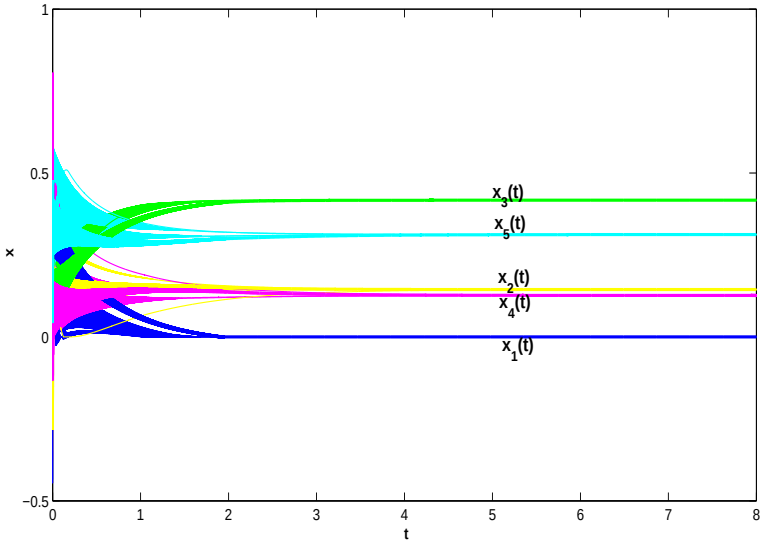
$$
\begin{cases}
\text{Minimize} \\
۰/۰۰۱۴۶۷۳۹x_A^۲ + ۲[۰/۰۰۰۲۶۱۷۹x_Ax_B + ۰/۰۰۰۵۲۱۰۳x_Ax_c + ۰/۰۰۰۱۷۳۸۲x_Ax_D \\
+ ۰/۰۰۰۴۳۲۷۲x_Ax_E - ۰/۰۰۰۰۱۱۳۳x_Bx_C + ۰/۰۰۰۵۹۰۱۰x_Bx_D + ۰/۰۰۰۴۸۶۹۹x_Bx_E \\
- ۰/۰۰۰۰۲۹۹۰x_Cx_D - ۰/۰۰۰۱۷۰۶۴x_Cx_E + ۰/۰۰۰۵۵۲۶۴x_Dx_E] + ۰/۰۰۱۱۱۶۴۰x_B^۲ \\
+ ۰/۰۰۱۰۵۷۹۰x_c^۲ + ۰/۰۰۱۰۸۵۴۳x_D^۲ + ۰/۰۰۱۰۰۳۷۱x_E^۲ \\
\text{s.t.} \\
\quad ۰/۰۲۰۷x_A + ۰/۰۳۱۶x_B + ۰/۰۳۲۳x_C + ۰/۰۳۳۷x_D + ۰/۰۳۷۶x_E \geq ۰/۰۳, \\
\quad x_A + x_B + x_C + x_D + x_E = ۱, \\
\quad ۰ \leq x_A, x_B, x_C, x_D, x_E \leq ۰/۷۵.
\end{cases}
$$

جواب بهینه این مسأله $x^* = (۰, ۰/۱۴۴, ۰/۴۱۷, ۰/۱۲۷, ۰/۳۱۲)^T$. ما برای حل این مسأله شبکه عصبی پیشنهاد شده در (۲۸.۳)-(۲۹.۳) را به‌کار می‌بریم. شکل (۳.۳) نشان می‌دهد که مسیرهای شبکه عصبی پیشنهادی با ۱۰ نقطه آغازین متفاوت، همگرا به جواب بهینه مسأله است.

مثال ۲.۶.۳. مسأله برنامه‌ریزی خطی به فرم (۱۱.۳)-(۱۲.۳) را در نظر بگیرید:

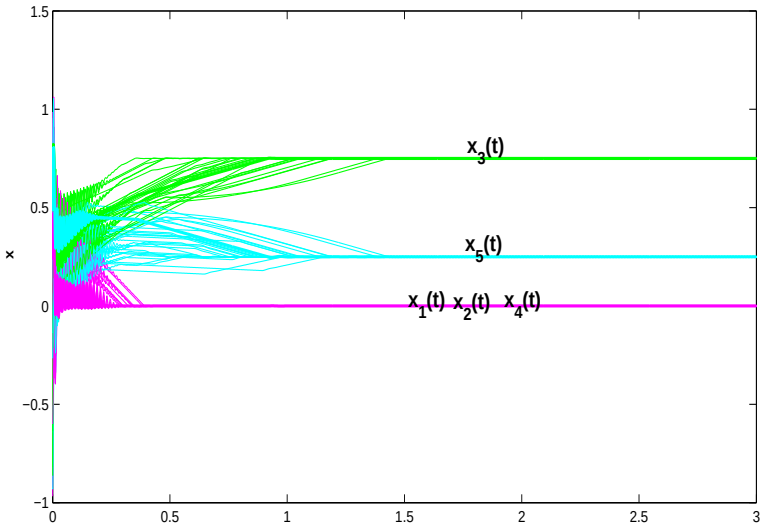
$$
\begin{cases}
\text{Minimize} & \frac{۱}{۲}\sum_{t=۱}^{۱۲} y_t \\
\text{s.t.} & \\
& y_۱ + ۰/۰۳۳۳x_A + ۰/۰۰۰۴x_B + ۰/۰۰۸۳x_C + ۰/۰۰۴۳x_D + ۰/۰۱۱۴x_E \geq ۰, \\
& y_۱ - ۰/۰۳۳۳x_A - ۰/۰۰۰۴x_B - ۰/۰۰۸۳x_C - ۰/۰۰۴۳x_D - ۰/۰۱۱۴x_E \geq ۰, \\
& y_۲ + ۰/۰۲۴۳x_A + ۰/۰۲۳۴x_B + ۰/۰۲۰۳x_C + ۰/۰۲۸۳x_D + ۰/۰۲۹۴x_E \geq ۰, \\
& y_۲ - ۰/۰۲۴۳x_A - ۰/۰۲۳۴x_B - ۰/۰۲۰۳x_C - ۰/۰۲۸۳x_D - ۰/۰۲۹۴x_E \geq ۰, \\
& y_۳ + ۰/۰۵۰۷x_A + ۰/۰۶۷۶x_B + ۰/۰۱۲۳x_C + ۰/۰۷۰۷x_D + ۰/۰۷۶۶x_E \geq ۰, \\
& y_۳ - ۰/۰۵۰۷x_A - ۰/۰۶۷۶x_B - ۰/۰۱۲۳x_C - ۰/۰۷۰۷x_D - ۰/۰۷۶۶x_E \geq ۰, \\
& y_۴ + ۰/۰۳۸۷x_A + ۰/۰۲۰۴x_B + ۰/۰۰۸۷x_C + ۰/۰۱۶۳x_D + ۰/۰۱۳۴x_E \geq ۰, \\
& y_۴ - ۰/۰۳۸۷x_A - ۰/۰۲۰۴x_B - ۰/۰۰۸۷x_C - ۰/۰۱۶۳x_D - ۰/۰۱۳۴x_E \geq ۰, \\
& y_۵ + ۰/۰۲۲۳x_A + ۰/۰۱۵۴x_B + ۰/۰۱۷۳x_C + ۰/۰۳۱۳x_D + ۰/۰۱۱۴x_E \geq ۰, \\
& y_۵ - ۰/۰۲۲۳x_A - ۰/۰۱۵۴x_B - ۰/۰۱۷۳x_C - ۰/۰۳۱۳x_D - ۰/۰۱۱۴x_E \geq ۰, \\
& y_۶ + ۰/۰۲۶۳x_A + ۰/۰۰۲۴x_B + ۰/۰۰۰۳x_C + ۰/۰۷۶۷x_D + ۰/۰۰۰۶x_E \geq ۰, \\
& y_۶ - ۰/۰۲۶۳x_A - ۰/۰۰۲۴x_B - ۰/۰۰۰۳x_C - ۰/۰۷۶۷x_D - ۰/۰۰۰۶x_E \geq ۰, \\
& y_۷ + ۰/۰۳۳۴x_A + ۰/۰۳۱۴x_B + ۰/۰۰۱۳x_C + ۰/۰۲۸۳x_D + ۰/۰۱۷۴x_E \geq ۰, \\
& y_۷ - ۰/۰۳۳۴x_A - ۰/۰۳۱۴x_B - ۰/۰۰۱۳x_C - ۰/۰۲۸۳x_D - ۰/۰۱۷۴x_E \geq ۰, \\
& y_۸ + ۰/۰۱۵۳x_A + ۰/۰۱۶۴x_B + ۰/۰۱۱۳x_C + ۰/۰۰۰۳x_D + ۰/۰۱۲۶x_E \geq ۰, \\
& y_۸ - ۰/۰۱۵۳x_A - ۰/۰۱۶۴x_B - ۰/۰۱۱۳x_C - ۰/۰۰۰۳x_D - ۰/۰۱۲۶x_E \geq ۰, \\
& y_۹ + ۰/۰۵۹۷x_A + ۰/۰۰۶۶x_B + ۰/۰۷۴۷x_C + ۰/۰۰۱۳x_D + ۰/۰۱۴۴x_E \geq ۰, \\
& \text{‘}y_۹ - ۰/۰۵۹۷x_A - ۰/۰۰۶۶x_B - ۰/۰۷۴۷x_C - ۰/۰۰۱۳x_D - ۰/۰۱۴۴x_Eg \geq ۰, \\
& y_{۱۰} + ۰/۰۶۳۷x_A + ۰/۰۰۸۴x_B + ۰/۰۱۱۳x_C + ۰/۰۲۲۳x_D + ۰/۰۱۷۶x_E \geq ۰, \\
& y_{۱۰} - ۰/۰۶۳۷x_A - ۰/۰۰۸۴x_B - ۰/۰۱۱۳x_C - ۰/۰۲۲۳x_D - ۰/۰۱۷۶x_E \geq ۰, \\
& y_{۱۱} + ۰/۰۲۵۳x_A + ۰/۰۰۴۴x_B + ۰/۰۰۴۳x_C + ۰/۰۲۳۳x_D + ۰/۰۰۷۴x_E \geq ۰, \\
& y_{۱۱} - ۰/۰۲۵۳x_A - ۰/۰۰۴۴x_B - ۰/۰۰۴۳x_C - ۰/۰۲۳۳x_D - ۰/۰۰۷۴x_E \geq ۰, \\
& y_{۱۲} + ۰/۰۳۱۳x_A + ۰/۰۴۸۶x_B + ۰/۰۰۳۷x_C + ۰/۰۰۸۷x_D + ۰/۰۰۲۴x_E \geq ۰, \\
& y_{۱۲} - ۰/۰۳۱۳x_A - ۰/۰۴۸۶x_B - ۰/۰۰۳۷x_C - ۰/۰۰۸۷x_D - ۰/۰۰۲۴x_E \geq ۰, \\
& ۰/۰۲۰۷x_A + ۰/۰۳۱۶x_B + ۰/۰۳۲۳x_C + ۰/۰۳۳۷x_D + ۰/۰۳۷۶x_E \geq ۰/۰۳, \\
& x_A + x_B + x_C + x_D + x_E = ۱, \\
& ۰ \leq x_A, x_B, x_C, x_D, x_E \leq ۰/۷۵, \\
& y_t \geq ۰, \ t = ۱, ..., ۱۲.
\end{cases}
$$

جواب بهینه این مسأله $x^* = (۰, ۰, ۰/۷۵, ۰, ۰/۲۵)^T$. ما برای حل این مسأله شبکه عصبی پیشنهاد شده در (۲۸.۳)-(۲۹.۳) را به‌کار می‌بریم. شکل (۴.۳) نشان می‌دهد که مسیرهای شبکه عصبی پیشنهادی

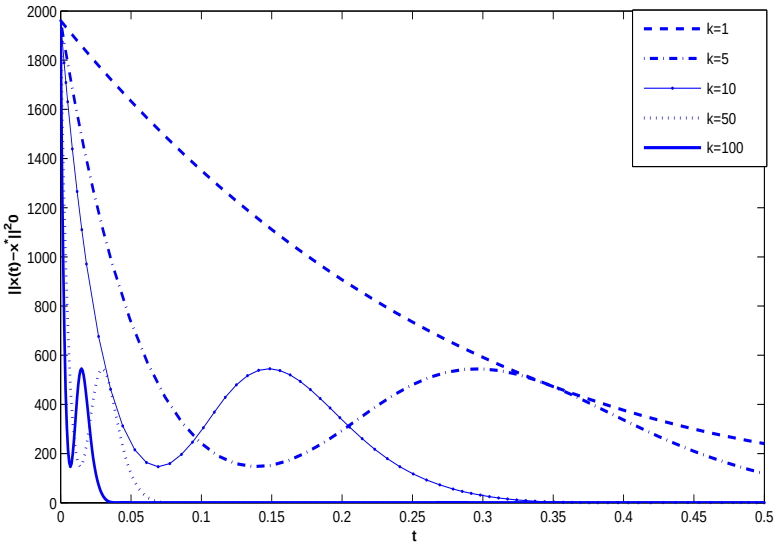


شکل ۳.۳: رفتار گذرا شبکه عصبی (۲۸.۳)-(۲۹.۳) با ۱۰ نقطه آغازین دلخواه در مثال ۱.۶.۳.

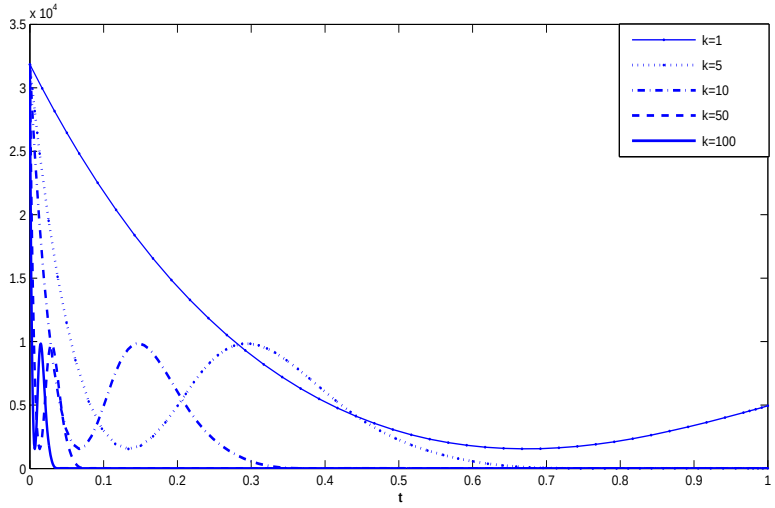
با ۳۰ نقطه آغازین متفاوت، همگرا به جواب بهینه مسأله است. ما همچنین تأثیر پارامتر κ از شبکه عصبی (۲۸.۳)-(۲۹.۳) روی مقدار $\|x(t) - x^*\|^2$ را در شکل (۵.۳) بررسی کردیم. واضح است که κ بزرگتر، نرخ همگرایی بهتری از $\|x(t) - x^*\|^2$ نتیجه می‌دهد.



شکل ۴.۳: رفتار گذرا شبکه عصبی (۲۸.۳)-(۲۹.۳) با ۳۰ نقطه آغازین دلخواه در مثال ۲.۶.۳.



شکل ۵.۳: رفتار همگرایی از $\|x(t)-x^*\|^2$ با $y_0 = (40,-40,40,-40,40,-40,40)^T$ در مثال ۲.۶.۳.



شکل ۶.۳: رفتار همگرایی از $\|x(t)-x^*\|^2$ با $y_0=(40,-40,40,-40,40,-40,40)^T$ در مثال ۳.۶.۳.

مثال ۳.۶.۳. مسأله برنامه‌ریزی خطی به فرم (۱۳.۳)-(۱۴.۳) را در نظر بگیرید:

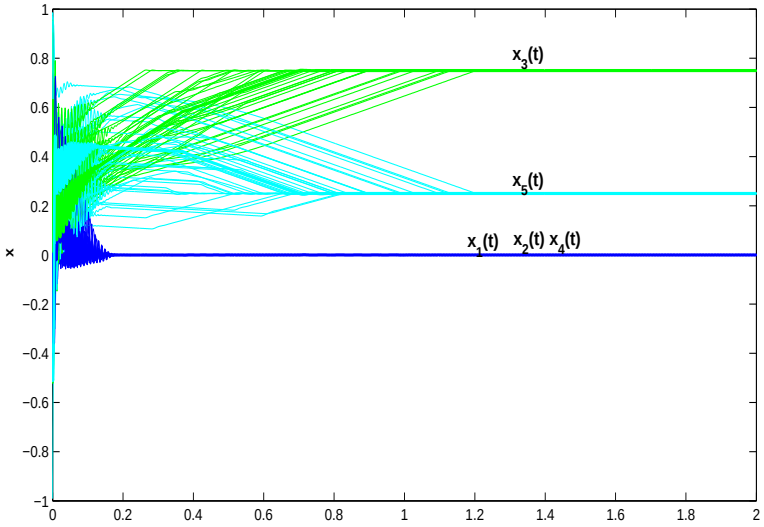
$$
\begin{cases}
\text{Maximize} & z \\
\text{s.t.} & \\
& ۰/۰۵۴x_A + ۰/۰۳۲x_B + ۰/۰۶۴x_C + ۰/۰۳۸x_D + ۰/۰۴۹x_E \geq z, \\
& ۰/۰۴۵x_A + ۰/۰۵۵x_B + ۰/۰۵۶x_C + ۰/۰۶۲x_D + ۰/۰۶۷x_E \geq z, \\
& -۰/۰۳۰x_A - ۰/۰۳۶x_B - ۰/۰۴۸x_C - ۰/۰۳۷x_D - ۰/۰۳۹x_E \geq z, \\
& -۰/۰۱۸x_A + ۰/۰۵۲x_B + ۰/۰۰۷x_C + ۰/۰۵۰x_D + ۰/۰۵۱x_E \geq z, \\
& ۰/۰۴۳x_A + ۰/۰۴۷x_B + ۰/۰۵۳x_C + ۰/۰۶۵x_D + ۰/۰۴۹x_E \geq z, \\
& ۰/۰۴۷x_A + ۰/۰۳۴x_B + ۰/۰۳۶x_C - ۰/۰۴۳x_D + ۰/۰۳۷x_E \geq z, \\
& ۰/۰۵۵x_A + ۰/۰۶۳x_B + ۰/۰۱۷x_C + ۰/۰۶۲x_D + ۰/۰۵۵x_E \geq z, \\
& ۰/۰۳۶x_A + ۰/۰۴۸x_B + ۰/۰۴۷x_C + ۰/۰۳۴x_D + ۰/۰۲۵x_E \geq z, \\
& -۰/۰۳۹x_A + ۰/۰۲۵x_B - ۰/۰۵۹x_C + ۰/۰۳۵x_D + ۰/۰۵۲x_E \geq z, \\
& -۰/۰۴۳x_A + ۰/۰۴۰x_B + ۰/۰۴۷x_C + ۰/۰۵۶x_D + ۰/۰۲۰x_E \geq z, \\
& ۰/۰۴۶x_A + ۰/۰۳۶x_B + ۰/۰۴۰x_C + ۰/۰۵۷x_D + ۰/۰۴۵x_E \geq z, \\
& ۰/۰۵۲x_A - ۰/۰۱۷x_B + ۰/۰۳۲x_C + ۰/۰۲۵x_D + ۰/۰۴۰x_E \geq z, \\
& ۰/۰۲۰۷x_A + ۰/۰۳۱۶x_B + ۰/۰۳۲۳x_C + ۰/۰۳۳۷x_D + ۰/۰۳۷۶x_E \geq ۰/۰۳, \\
& x_A + x_B + x_C + x_D + x_E = ۱, \\
& ۰ \leq x_A, x_B, x_C, x_D, x_E \leq ۰/۷۵, \\
& z \geq ۰.
\end{cases}
$$

جواب بهینه این مسأله $x^*=(۰,۰,۰/۴۵۹۶,۰,۰/۵۴۰۴)^T$ است. تأثیر پارامتر κ شبکه فوق روی $\|x(t)-x^*\|^2$ در شکل (۶.۳) نشان داده شده است.

مثال ۴.۶.۳. مسأله برنامه‌ریزی خطی به‌فرم (۲۱.۳)-(۲۲.۳) را در نظر بگیرید:

$$
\begin{cases}
\text{Minimize} \quad \frac{۱}{۱۲}\sum_{t=۱}^{۱۲} y_t \\
\text{s.t.} \\
\text{۰/۰۳۳۳}x_A + \text{۰/۰۰۰۴}x_B + \text{۰/۰۰۸۳}x_C + \text{۰/۰۰۴۳}x_D + \text{۰/۰۱۱۴}x_E \le y_۱, \\
\text{۰/۰۲۴۳}x_A + \text{۰/۰۲۳۴}x_B + \text{۰/۰۲۰۳}x_C + \text{۰/۰۲۸۳}x_D + \text{۰/۰۲۹۴}x_E \le y_۲, \\
\text{۰/۰۵۰۷}x_A + \text{۰/۰۶۷۶}x_B + \text{۰/۰۱۲۳}x_C + \text{۰/۰۷۰۷}x_D + \text{۰/۰۷۶۶}x_E \le y_۳, \\
\text{۰/۰۳۸۷}x_A + \text{۰/۰۲۰۴}x_B + \text{۰/۰۰۸۷}x_C + \text{۰/۰۱۶۳}x_D + \text{۰/۰۱۳۴}x_E \le y_۴, \\
\text{۰/۰۲۲۳}x_A + \text{۰/۰۱۵۴}x_B + \text{۰/۰۱۷۳}x_C + \text{۰/۰۳۱۳}x_D + \text{۰/۰۱۱۴}x_E \le y_۵, \\
\text{۰/۰۲۶۳}x_A + \text{۰/۰۰۲۴}x_B + \text{۰/۰۰۰۳}x_C + \text{۰/۰۷۶۷}x_D + \text{۰/۰۰۰۶}x_E \le y_۶, \\
\text{۰/۰۳۳۴}x_A + \text{۰/۰۳۱۴}x_B + \text{۰/۰۰۱۳}x_C + \text{۰/۰۲۸۳}x_D + \text{۰/۰۱۷۴}x_E \le y_۷, \\
\text{۰/۰۱۵۳}x_A + \text{۰/۰۱۶۴}x_B + \text{۰/۰۱۱۳}x_C + \text{۰/۰۰۰۳}x_D + \text{۰/۰۱۲۶}x_E \le y_۸, \\
\text{۰/۰۵۹۷}x_A + \text{۰/۰۰۶۶}x_B + \text{۰/۰۷۴۷}x_C + \text{۰/۰۰۱۳}x_D + \text{۰/۰۱۴۴}x_E \le y_۹, \\
\text{۰/۰۶۳۷}x_A + \text{۰/۰۰۸۴}x_B + \text{۰/۰۱۱۳}x_C + \text{۰/۰۲۲۳}x_D + \text{۰/۰۱۷۶}x_E \le y_{۱۰}, \\
\text{۰/۰۲۵۳}x_A + \text{۰/۰۰۴۴}x_B + \text{۰/۰۰۴۳}x_C + \text{۰/۰۲۳۳}x_D + \text{۰/۰۰۷۴}x_E \le y_{۱۱}, \\
\text{۰/۰۳۱۳}x_A + \text{۰/۰۴۸۶}x_B + \text{۰/۰۰۳۷}x_C + \text{۰/۰۰۸۷}x_D + \text{۰/۰۰۲۴}x_E \le y_{۱۲}, \\
\text{۰/۰۲۰۷}x_A + \text{۰/۰۳۱۶}x_B + \text{۰/۰۳۲۳}x_C + \text{۰/۰۳۳۷}x_D + \text{۰/۰۳۷۶}x_E \le \text{۰/۰۳}, \\
x_A + x_B + x_C + x_D + x_E = ۱, \\
۰ \le x_A, x_B, x_C, x_D, x_E \le \text{۰/۷۵}, \\
y_t \ge ۰, \ t = ۱, ..., ۱۲.
\end{cases}
$$

جواب بهینه این مسأله $x^* = (۰, ۰, \text{۰/۷۵}, ۰, \text{۰/۲۵})^T$. ما برای حل این مسأله از مدل شبکه عصبی پیشنهاد شده در (۲۸.۳)-(۲۹.۳) استفاده می‌کنیم. شکل (۷.۳) نشان می‌دهد که مسیرهای شبکه عصبی پیشنهادی با ۴۰ نقطه آغازین متفاوت، همگرا به جواب بهینه مسأله است.

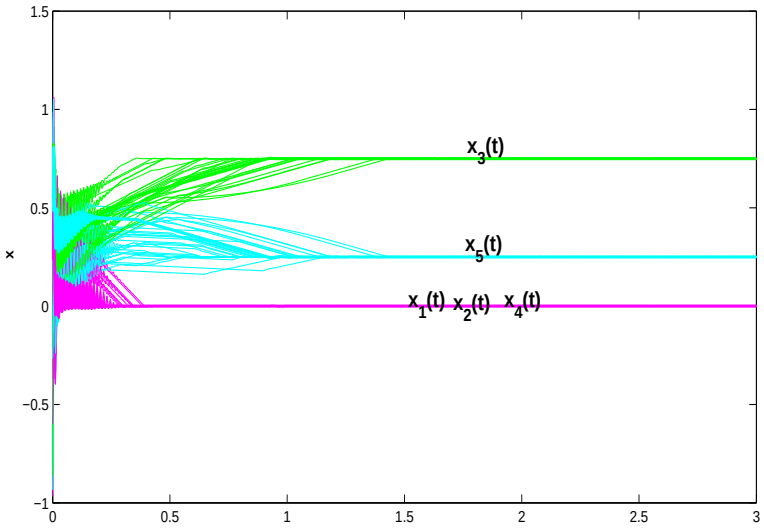


شکل ۷.۳: رفتار گذرا شبکه عصبی (۲۸.۳)-(۲۹.۳) با ۴۰ نقطه آغازین دلخواه در مثال ۴.۶.۳.

مثال ۵.۶.۳. مسأله برنامه‌ریزی درجه دوم به فرم (۹.۳)-(۱۰.۳) را در نظر بگیرید:

$$
\begin{cases}
\text{Minimize} \quad \sum_{t=۱}^{۱۲} y_t^۲ \\
\text{s.t.} \\
y_۱ = \frac{۱}{۲}(۰/۰۳۳۳x_A + ۰/۰۰۰۴x_B + ۰/۰۰۸۳x_C + ۰/۰۰۴۳x_D + ۰/۰۱۱۴x_E), \\
y_۲ = \frac{۱}{۲}(+۰/۰۲۴۳x_A + ۰/۰۲۳۴x_B + ۰/۰۲۰۳x_C + ۰/۰۲۸۳x_D + ۰/۰۲۹۴x_E), \\
y_۳ = \frac{۱}{۲}(-۰/۰۵۰۷x_A - ۰/۰۶۷۶x_B + ۰/۰۱۲۳x_C - ۰/۰۷۰۷x_D - ۰/۰۷۶۶x_E), \\
y_۴ = \frac{۱}{۲}(-۰/۰۳۸۷x_A + ۰/۰۲۰۴x_B - ۰/۰۰۸۷x_C + ۰/۰۱۶۳x_D + ۰/۰۱۳۴x_E), \\
y_۵ = \frac{۱}{۲}(+۰/۰۲۲۳x_A + ۰/۰۱۵۴x_B + ۰/۰۱۷۳x_C + ۰/۰۳۱۳x_D + ۰/۰۱۱۴x_E), \\
y_۶ = \frac{۱}{۲}(+۰/۰۲۶۳x_A + ۰/۰۰۲۴x_B + ۰/۰۰۰۳x_C - ۰/۰۷۶۷x_D - ۰/۰۰۰۶x_E), \\
y_۷ = \frac{۱}{۲}(+۰/۰۳۳۴x_A + ۰/۰۳۱۴x_B + ۰/۰۰۱۳x_C + ۰/۰۲۸۳x_D + ۰/۰۱۷۴x_E), \\
y_۸ = \frac{۱}{۲}(+۰/۰۱۵۳x_A + ۰/۰۱۶۴x_B + ۰/۰۱۱۳x_C + ۰/۰۰۰۳x_D - ۰/۰۱۲۶x_E), \\
y_۹ = \frac{۱}{۲}(-۰/۰۵۹۷x_A - ۰/۰۰۶۶x_B - ۰/۰۷۴۷x_C + ۰/۰۰۱۳x_D + ۰/۰۱۴۴x_E), \\
y_{۱۰} = \frac{۱}{۲}(-۰/۰۶۳۷x_A + ۰/۰۰۸۴x_B + ۰/۰۱۱۳x_C + ۰/۰۲۲۳x_D - ۰/۰۱۷۶x_E), \\
y_{۱۱} = \frac{۱}{۲}(+۰/۰۲۵۳x_A + ۰/۰۰۴۴x_B + ۰/۰۰۴۳x_C + ۰/۰۲۳۳x_D + ۰/۰۰۷۴x_E), \\
y_{۱۲} = \frac{۱}{۲}(+۰/۰۳۱۳x_A - ۰/۰۴۸۶x_B - ۰/۰۰۳۷x_C - ۰/۰۰۸۷x_D + ۰/۰۰۲۴x_E), \\
۰/۰۲۰۷x_A + ۰/۰۳۱۶x_B + ۰/۰۳۲۳x_C + ۰/۰۳۳۷x_D + ۰/۰۳۷۶x_E \geq ۰/۰۳, \\
x_A + x_B + x_C + x_D + x_E = ۱, \\
۰ \leq x_A, x_B, x_C, x_D, x_E \leq ۰/۷۵, \\
y_t \geq ۰, \ t = ۱, ..., ۱۲.
\end{cases}
$$

جواب بهینه این مسأله $x^* = (۰, ۰, ۰/۵۱۵۵, ۰, ۰/۴۸۴۵)^T$. ما برای حل این مسأله شبکه عصبی پیشنهاد شده در (۲۸.۳)-(۲۹.۳) را به‌کار می‌بریم. شکل (۸.۳) نشان می‌دهد که مسیرهای شبکه عصبی



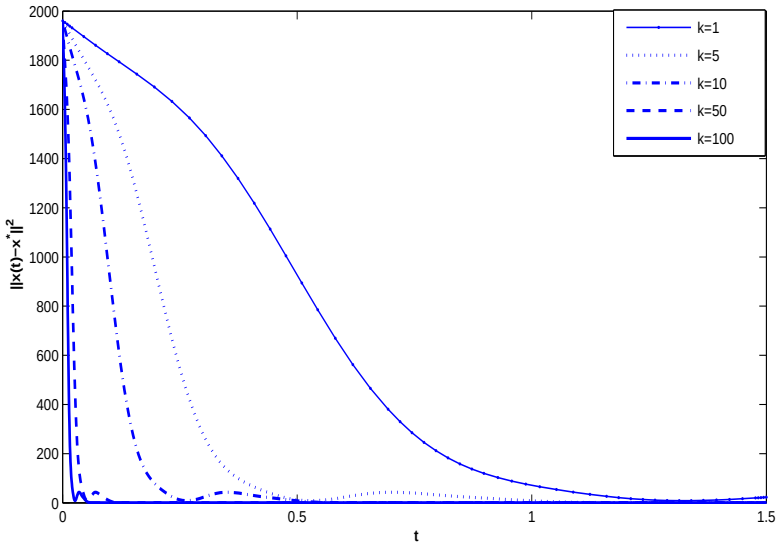
شکل ۸.۳: رفتار گذرا شبکه عصبی (۲۸.۳)-(۲۹.۳) با ۳۰ نقطه آغازین دلخواه در مثال ۵.۶.۳.

پیشنهادی با ۳۰ نقطه آغازین متفاوت، همگرا به جواب بهینه مسأله است. تأثیر پارامتر κ شبکه فوق روی $\|x(t) - x^*\|^2$ هم در شکل (۹.۳) نشان داده شده است.

مسأله II

مسأله فرضی دیگر به این صورت بیان می‌شود، فرض کنید سبد شامل سهام‌های US، اوراق قرضه و وجوه نقد باشد. ما می‌خواهیم با استفاده از داده‌های بازدهی تاریخی برای این ۳ کلاس از دارایی‌ها، بازده مورد انتظار آتی آنها را تخمین بزنیم. از شاخص $S\&P۵۰۰$ برای بازدهی سهام‌ها، شاخص اوراق خزانه ۱۰ ساله [۴۴] برای بازدهی اوراق استفاده کردیم و فرض می‌کنیم وجه ما در بازار پول سرمایه‌گذاری شده که بازدهی آن با شاخص $(۱ - dayfederalfundrate)$ داده شده است. سری زمانی سالانه برای بازدهی کل ، میانگین و ماتریس کواریانس هر کلاس دارایی بین سال‌های ۱۹۶۰ و ۲۰۰۳ در [۵۷] داده شده

[۴۴] 10-year Treasury Bond Index



شکل ۹.۳: رفتار همگرایی از $\|x(t)-x^*\|^2$ در مثال ۵.۶.۳.

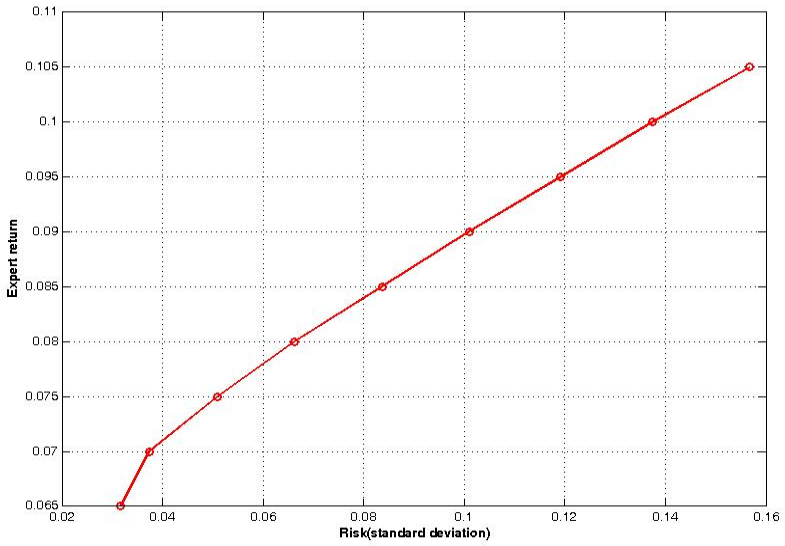
است.

مثال ۶.۶.۳. مسأله برنامه‌ریزی درجه دوم به فرم (۱.۳)-(۲.۳) را در نظر بگیرید:

$$\begin{cases} \text{Minimize} \\ ۰/۰۲۷۷۸x_S^2 + ۰/۰۱۱۱۲x_B^2 + ۰/۰۰۱۱۵x_M^2 + ۲(۰/۰۰۳۸۷x_Sx_B + ۰/۰۰۰۲۱x_Sx_M - ۰/۰۰۰۲۰x_Bx_M) \\ \text{s.t.} \\ ۰/۱۰۷۳x_S + ۰/۰۷۳۷x_B + ۰/۰۶۲۷x_M \geq R, \\ x_S + x_B + x_M = ۱, \\ x_S, x_B, x_M \geq ۰. \end{cases}$$

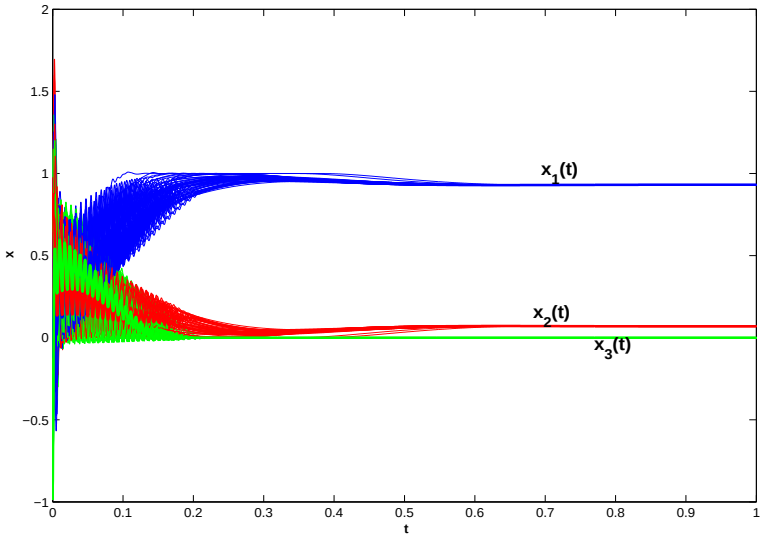
جدول ۱.۳: سبد سرمایه‌های کارا

MM	$Bonds$	$Stocks$	واریانس	نرخ بازدهی(R)
۰.۸۷	۰.۱۰	۰.۰۳	۰.۰۰۱۰	۰.۰۶۵
۰.۷۵	۰.۱۲	۰.۱۳	۰.۰۰۱۴	۰.۰۷۰
۰.۶۲	۰.۱۴	۰.۲۴	۰.۰۰۲۶	۰.۰۷۵
۰.۴۹	۰.۱۶	۰.۳۵	۰.۰۰۴۴	۰.۰۸۰
۰.۳۷	۰.۱۸	۰.۴۵	۰.۰۰۷۰	۰.۰۸۵
۰.۲۴	۰.۲۰	۰.۵۶	۰.۰۱۰۲	۰.۰۹۰
۰.۱۱	۰.۲۲	۰.۶۷	۰.۰۱۴۲	۰.۰۹۵
۰	۰.۲۲	۰.۷۸	۰.۰۱۸۹	۰.۱۰۰
۰	۰.۰۷	۰.۹۳	۰.۰۲۴۶	۰.۱۰۵

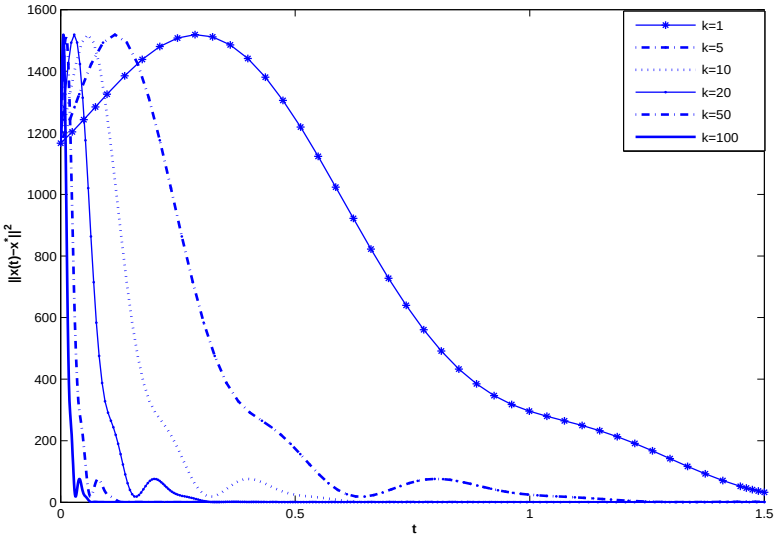


شکل ۱۰.۳: مرز کارایی سبد سرمایه‌های بدست آمده در مثال ۶.۶.۳.

به ازای $R = ۶/۵\%$ الی $R = ۱۱/۵\%$ با افزایش $R = ۰/۵\%$درصد و با حل این مدل به ازای مقادیر مختلف R ، سبد سرمایه‌های بهینه مطابق جدول (۱.۳) بدست می‌آید. شکل (۱۰.۳) مرز کارایی این سبد سرمایه‌ها را نشان می‌دهد. جواب بهینه این مسأله برای $R = ۱۰/۵\%$ ، $x^* = (۰/۹۳۱۵, ۰/۰۶۸۵, ۰)^T$ است. ما برای حل این مسأله شبکه عصبی پیشنهاد شده در (۲۸.۳)-(۲۹.۳) را به‌کار می‌بریم. شکل (۱۱.۳) نشان می‌دهد که مسیرهای شبکه عصبی پیشنهادی با ۵۰ نقطه آغازین متفاوت، همگرا به جواب بهینه مسأله است. تأثیر پارامتر κ شبکه فوق روی $\|x(t) - x^*\|^۲$ با نقطه آغازین $y_۰ = (۲۰, -۲۰, ۲۰, -۲۰, ۲۰, -۲۰, ۲۰, -۲۰)^T$، در شکل (۱۲.۳) نشان داده شده است.



شکل ۱۱.۳: رفتار گذرا شبکه عصبی (۲۸.۳)-(۲۹.۳) با ۳۰ نقطه آغازین دلخواه در مثال ۶.۶.۳.



شکل ۱۲.۳: رفتار همگرایی از $\|x(t) - x^*\|^2$ در مثال ۶.۶.۳.

مثال ۷.۶.۳. مسأله برنامه‌ریزی درجه دوم به فرم (۷.۳)-(۸.۳) را در نظر بگیرید:

$$\begin{cases} \text{Minimize} \quad (y_۱)^۲ + (y_۲)^۲ + (y_۳)^۲ \\ \text{s.t.} \\ y_۱ = ۰/۱۶۶۶۷x_S, \\ y_۲ = ۰/۰۲۳۲۲x_S + ۰/۱۰۲۸۶x_B, \\ y_۳ = ۰/۰۰۱۲۶x_S + ۰/۰۲۴۶۸x_M - ۰/۰۰۲۲۴x_B, \\ ۰/۱۰۷۳x_S + ۰/۰۷۳۷x_B + ۰/۰۶۲۷x_M \geq ۰/۱۰۵, \\ x_S + x_B + x_M = ۱, \\ x_S, x_B, x_M \geq ۰, \\ y_۱, y_۲, y_۳ \geq ۰. \end{cases}$$

جواب بهینه این مسأله $x^* = (۰/۹۳۱۵, ۰/۰۶۸۵, ۰)^T$. ما برای حل این مسأله شبکه عصبی پیشنهاد شده در (۲۸.۳)-(۲۹.۳) را به‌کار می‌بریم.

مقایسه عملکرد مدل‌ها با یکدیگر

در مورد مقایسه عملکرد این ۵ مدل، با ۲ فرضیه در نظر می‌گیریم:

۱. سرمایه گذار، خواهان کمترین ریسک بدون توجه به میزان بازدهی دریافتی خود است. با توجه به این شرط طبق جدول (۲.۳) ، مدل MV بهترین انتخاب برای وی است.

۲. سرمایه گذار خواهان بیشترین بازدهی است، بنابراین مدل $Maximin$ را انتخاب می‌کند.

جدول ۲.۳: مقایسه مدل‌ها

مدل	وزن‌ها	$E(r)(\%)$	$\sigma(\%)$
MV	$x^* = (۰, ۰/۱۴۴, ۰/۴۱۷, ۰/۱۲۷, ۰/۳۱۲)$	۳.۴۰	۰.۰۳۸۳
MAD	$x^* = (۰, ۰, ۰/۷۵, ۰, ۰/۲۵)$	۳.۳۶	۰.۰۵۹۴
$Maximin$	$x^* = (۰, ۰, ۰/۴۵۹۶, ۰, ۰/۵۴۰۴)$	۳.۵۲	۰.۰۴۳۲
$Minimax$	$x^* = (۰, ۰, ۰/۷۵, ۰, ۰/۲۵)$	۳.۳۶	۰.۰۵۹۴
$Diago۲$	$x^* = (۰, ۰, ۰/۵۱۵۵, ۰, ۰/۴۸۴۵)$	۳.۴۹	۰.۰۴۳۱

فصل ۴
مسائل برنامه‌ریزی تصادفی

۱.۴ مقدمه

در این قطعه از پایان نامه ابتدا برنامه‌ریزی خطی تصادفی را معرفی کرده و مروری بر کارهای انجام شده داریم، سپس مسائل بهینه‌سازی با قیود تصادفی را مطرح می‌کنیم. یک مدل شبکه عصبی برای مسائل برنامه‌ریزی مخروط مرتبه دوم ارائه می‌دهیم و تحلیلی بر پایداری و همگرایی مدل پیشنهاد شده می‌آوریم. برای نشان دادن کارایی مدل شبکه عصبی پیشنهاد شده، ۴ مثال عددی از مسائل برنامه‌ریزی تصادفی که حالت خاصی از مسائل برنامه‌ریزی مخروط مرتبه دوم محدب هستند، را حل می‌کنیم.

۱.۱.۴ برنامه‌ریزی خطی تصادفی

مسأله برنامه‌ریزی خطی تصادفی را می‌توان به‌صورت زیر فرمولبندی کرد:

$$\text{Minimize} \quad f(x) = \sum_{j=۱}^{n} c_j x_j, \tag{۱.۴}$$

subject to

$$\sum_{j=۱}^{n} a_{ij} x_j \geq b_i, \quad i = ۱, ۲, ..., m, \tag{۲.۴}$$

$$x_j \geq ۰, \quad j = ۱, ۲, ..., n. \tag{۳.۴}$$

وقتیکه c_j و a_{ij} و b_i متغیرهای تصادفی با توزیعهای احتمال معلوم هستند. (برای سادگی، متغیرهای تصمیم x_j ،قطعی فرض می‌شوند)، مدل‌های برنامه‌ریزی تصادفی [۵۸، ۵۹، ۶۰] سبب ایجاد طرح‌هایی می‌شوند که در برابر زیان‌ها و عدم موفقیت محصول (نقص‌های فاجعه آمیز) بهتر می‌توانند مقاومت کنند. چنین مدلهایی برای کاربردهای مختلف شامل تولید انرژی الکتریکی [۱] [۶۱] ، نقشه‌های مالی [۲] [۶۲، ۶۳، ۶۴] ، ارتباط تلفنی [۳] [۶۵، ۶۶]، مدیریت زنجیره تأمین [۴] [۶۷] ، صنایع نفت[۶۸]، تولید

[۱] Electric power generation
[۲] Financial planning
[۳] Telecommunication network planning
[۴] Supply chain management

کنندگان وسایل نقلیه[۶۹]، تأمین کنندگان الکتریسیته[۷۰، ۷۱]، محیط[۷۲]، حمل ونقل [۷۳، ۷۴]، ساختمان، انرژی، فرآیندهای شیمیایی[۷۵]، هوا-فضا و سیستم نظامی[۷۶][۵] به‌کار می‌روند.

مدل‌های برنامه‌ریزی تصادفی شامل متغیرهای تصمیم قابل پیش‌بینی [۶] و تطبیقی [۷] هستند. متغیرهای قابل پیش‌بینی مربوط به تصمیماتی می‌شوند که باید در زمان حال گرفته شوند و نمی توانند به پارامترهای تصادفی تشخیص جزئی و مشاهدات آینده وابسته باشند. متغیرهای تطابقی به تصمیمات آتی پس از مشاهده تمام یا برخی از پارامترهای تصادفی بستگی دارند. مدل‌های قابل پیش‌بینی می‌توانند در برنامه‌ریزی تصادفی با منابع [۸] و یا مدل‌های برنامه‌ریزی با قیود تصادفی [۹] فرمول بندی شوند.

در فرمول‌بندی برنامه‌ریزی با قیود تصادفی که توسط چارنز و کوپر [۱۰] در [۵۹] ارائه شد، ما به یک متغیر خاص(تابع) درون دامنهٔ هدف به همراه احتمال مشخص نیاز داریم. این مدل‌ها معمولا به قیود غیرخطی و نامحدب منتهی می‌شوند. به کارگیری اولیه محاسباتی مسائل برنامه‌ریزی با قیود تصادفی محدود به متغیرهای تصادفی نرمال می‌شود. پریکوپا [۱۱] [۷۷] یک الگوریتم دوگانه برای حل مسأله کلی قیود تصادفی پیشنهاد داد، او در[۷۸] به بحث درباره روشهای مختلف برای حل مدل‌های قیود تصادفی شامل روش‌های گرادیانی و استفاده از توابع جریمه می‌پردازد. هیلر [۱۲] در [۷۹] فرآیندی برای تقریب قیود تصادفی با قیود خطی را پیشنهاد داد. وی قیود تصادفی را با قیود تفکیک‌پذیر محدب جایگزین کرد و سپس مسأله را با روش برنامه‌ریزی تفکیک‌پذیر محدب با استفاده از الگوریتم سیمپلکس حل کرد. سپالا در [۸۰] مجموعه‌ای از قیود خطی یکنواخت سخت [۱۳] که جایگزین یک قید تصادفی منفرد می‌شود را معرفی کرد. براساس نظر سپالا این روش بسیار دقیق تراست اما نسبت به روش هیلر کارایی کمتری دارد.

[۵] Military system
[۶] Anticipative
[۷] Adaptive
[۸] Stochastic Programming with Recourse (SPR)
[۹] Chance Constrained Programming (CCP)
[۱۰] Charnes and Cooper
[۱۱] Prekopa
[۱۲] Hillier
[۱۳] uniformly tighter linear constraints

این مدل محدود به مسائل دارای متغیر کران‌دار نمی‌شود، با این حال به محدودیت‌های بیشتری نسبت به روش هیلر نیازمند است. سپالا و اورپانا [۱۴] در [۸۱] کارایی و دقت الگوریتم سپالا را امتحان کردند. اولسون و سونس [۱۵] در [۸۲] یک فرمول تقریبی برای حل مسائل برنامه‌ریزی با قیود تصادفی پیشنهاد دادند که سبب ایجاد کران برای قیود تصادفی می‌گردد، که این(هم) ممکن است فرم غیر خطی داشته باشد، با این وجود این تکنیک خیلی دقیق نیست ([۸۳] را ببینید). وینتراب و ورا [۱۶] در [۸۴] روش صفحه برشی برای حل معادل قطعی مسائل برنامه‌ریزی با قیود تصادفی برای ضرایب محدود شده تصادفی با توزیع نرمال را ارائه کردند. دنچوا [۱۷] و همکاران در [۸۵] ترتیب جزئی روی مجموعه‌ای از سناریوها تعریف کردند و نشان دادند که تعداد محدودی از سناریوها (که نقاط مؤثر $(1-\alpha)$ نامیده می‌شوند) ، نقش کلیدی دارند. آنها الگوریتمی ایجاد کردند که به‌طور مدوام مجموعه‌ای از نقاط مؤثر را به‌منظور تولید کران‌های بالا و پائین به‌روز رسانی می‌کند. این الگوریتم تا حد برنامه‌ریزی محدب کلی با قیود تصادفی در[۸۶] توسعه یافته است. برالدی و روژنسکی [۱۸] در [۸۷] به بحث درباره روش‌هایی پرداختند که براساس شمارش جزئی یا کامل نقاط مؤثر برای برنامه‌ریزی صحیح تصادفی است. روژنسکی [۱۹] در [۸۸] مسائل برنامه‌ریزی با قیود تصادفی چون مسائل برنامه‌ریزی صحیح آمیخته بزرگ مقیاس [۲۰] را با در نظر گرفتن محدودیت‌های کوله پشتی [۲۱] دوباره فرمول‌سازی کرد و نامعادله‌های صحیح برای مسأله برنامه‌ریزی صحیح آمیخته ایجاد کرد. آرینگری [۲۲] در [۸۹] جستجوی ممنوع [۲۳] و فرآیند شبیه سازی را برای دو مسأله بهینه‌سازی شبکه ارتباط تلفنی با فرمول‌بندی از برنامه‌ریزی با قیود تصادفی توسعه داد.

[۱۴] Seppala and Orpana
[۱۵] Olson and Swenseth
[۱۶] Weintraub and Vera
[۱۷] Dentcheva
[۱۸] Beraldi and Ruszczynski
[۱۹] Ruszczynski
[۲۰] Large scale Mixed integer programming
[۲۱] Knapsack Constraints
[۲۲] Aringheri
[۲۳] Tabu Search

کالافیور و کامپی [۲۴] در [۹۰] ، دی‌فاریاس و ون‌روی [۲۵] در [۹۱] روش‌های تقریبی بر مبنای نمونه‌گیری محدودیت و روش‌های یادگیری آماری را پیشنهاد دادند. کالافیور و کامپی از روش‌های یادگیری آماری برای ایجاد کردن کران‌هایی روی اندازه (احتمال یا حجم) از مجموعهٔ قیود اصلی که احتمالا از طریق حل تصادفی شده‌اند استفاده کردند. نتایج بدست آمده توسط آنها با استفاده از این اصل اساسی که جواب بهینه برای مسأله محدب از طریق تعداد متناهی از قیود پشتیبانی می‌شود. دی‌فاریاس و ون‌روی نشان دادند که احتمال وقوع می‌تواند برای یک مسأله برنامه‌ریزی خطی از طریق نمونه‌گیری تعداد قابل ردیابی [۲۶] از قیود فراهم گردد. اندازه نمونه با زیاد شدن تعداد متغیرها به صورت چندجمله ای افزایش می‌یابد و مستقل از مجموعهٔ قید اصلی است. اردوگان و لینگار [۲۷] در [۹۲] نمونه نتایج پیچیدگی برای مسائل برنامه‌ریزی با قیود تصادفی با مسائل برنامه‌ریزی با قیود تصادفی مبهم گسترش دادند که در آن مجموعه غیر قطعی با گوی‌های تعریف شده در جملاتی از تابع متریک پروهورف [۲۸] داده شده است. با استفاده از قضیه نمایش استراسن-دوادلی، یک مسأله نمونه مقاوم [۲۹] را ساختند که تقریب خوبی برای مسائل برنامه‌ریزی با قیود تصادفی مبهم با احتمال بالا است. ایوامورا و لیو [۳۰] در [۹۳] براساس الگوریتم ژنتیک یک شبیه سازی تصادفی ایجاد کردند که در آن شبیه‌سازی، بررسی قیود تصادفی و محاسبه توابع هدف مورد استفاده قرار می‌گرفت. پوجاری و وارگس در [۹۴] [۳۱] با الهام گرفتن از کار ایوامورا و لیو، چارچوبی شامل الگوریتم ژنتیک و شبیه‌سازی مونت کارلو [۳۲] برای حل مسأله برنامه‌ریزی با قیود تصادفی پیشنهاد کردند.

در مهندسی و بسیاری از کاربردهای علمی، جوابهای زمان واقعی مسائل قیود تصادفی معمولا مطلوب

[۲۴] Calafiore and Campi
[۲۵] De Farias and V. Roy
[۲۶] tractable
[۲۷] Erdogan and Iyengar
[۲۸] Prohorov metric
[۲۹] Robust
[۳۰] Iwamura and Liu
[۳۱] Poojari and Varghese
[۳۲] Monte Carlo simulation

است. در نتیجه به روشهای مؤثر برای حل برنامه‌ریزی با قیود تصادفی زمان واقعی نیاز داریم. نتایج مطالعات قبلی [۹۵]-[۱۲۴] حاکی از آن است که از شبکه‌های عصبی می‌توان برای حل مسائل مختلف بهینه‌سازی استفاده کرد. ایده اصلی شبکه‌های عصبی برای بهینه‌سازی ایجاد یک تابع انرژی نامنفی و پایه‌گذاری یک سیستم دینامیکی است که یک شبکه عصبی مصنوعی را ایجاد می‌کند. سیستم دینامیکی معمولا به فرم معادلات دیفرانسیل مرتبه اول است. به‌علاوه انتظار می‌رود که سیستم دینامیکی به یک وضعیت پایدار (یا نقطهٔ تعادل) که متناظر با جواب مسأله بهینه‌سازی اصلی است که با یک نقطه اولیه آغاز می‌شود نزدیک گردد.

با توجه به مباحث فوق، در این فصل یک مدل از شبکه‌های عصبی برای حل مسائل برنامه‌ریزی تصادفی خطی ارائه می‌دهیم. ابتدا مسأله برنامه‌ریزی تصادفی به‌صورت یک مسأله برنامه ریزی مخروط مرتبه دوم محدب در می‌آید سپس این مسأله را تبدیل به برنامه‌ریزی غیرخطی محدب با قیود نامساوی می‌کنیم، زمانیکه با استفاده از تکنیک هموار سازی، قیود مخروط مرتبه دوم ناهموار به‌فرم هموار آن تبدیل شوند. براساس شرایط بهینگی $(K.K.T)$ برنامه‌ریزی محدب ، مدل شبکه عصبی برای حل مسائل برنامه‌ریزی غیرخطی محدب [۳۳] پیشنهاد خواهد شد. همچنین ثابت می‌شود نقطهٔ تعادل شبکه عصبی پیشنهاد شده برابر نقطه $K.K.T$ از مسأله برنامه ریزی مخروط مرتبه دوم محدب است. وجود و یکتایی نقطه تعادل شبکه عصبی نیز مورد آنالیز قرار خواهد گرفت. با ساختن یک تابع لیاپانوف مناسب، شرط کافی برای تضمین وجود و پایدار مجانبی سراسری برای نقطه تعادل یکتای شبکه پیشنهاد شده بدست خواهد آمد. مدل شبکه عصبی پیشنهاد شده برای مسائل $Minimax$ در [۱۰۸]، $maximum\ flow$ در [۱۰۹] ، $shortest\ Path\ problem$ در [۱۱۰] و برای برنامه‌ریزی هندسی در [۱۱۱] نیز با موفقیت استفاده شده است.

[۳۳] Convex Nonlinear Programming

۲.۴ مسائل بهینه‌سازی با قیود تصادفی

در روش برنامه‌ریزی با قیود تصادفی، مسأله برنامه‌ریزی تصادفی خطی (۱.۴)-(۳.۴) به‌صورت زیر بیان می‌شود:

$$\text{Minimize} \quad f(x) = \sum_{j=۱}^{n} c_j x_j \tag{۴.۴}$$

subject to

$$\mathrm{P}[\sum_{j=۱}^{n} a_{ij} x_j \leq b_i] \geq p_i, \quad i = ۱, ۲, ..., m, \tag{۵.۴}$$

$$x_j \geq ۰, \quad j = ۱, ۲, ..., n. \tag{۶.۴}$$

وقتیکه c_j ، b_i ، a_{ij} ، متغیرهای تصادفی و p_i ها احتمالات مشخص شده هستند. توجه کنید که (۲.۴) نشان می‌دهد که iامین قید

$$\sum_{j=۱}^{n} a_{ij} x_j \leq b_i,$$

باید دست کم با احتمال p_i وقتیکه $۰ \leq p_i \leq ۱$ است، برآورده شود. برای سادگی، فرض می‌کنیم که متغیرهای تصمیم x_j قطعی هستند. ابتدا حالات خاصی را که تنها c_j یا a_{ij} یا b_i ها متغیرهای تصادفی هستند را بررسی کرده، سپس به حالت کلی که a_{ij} ، c_jو b_i همگی متغیرهای تصادفی هستند، می‌پردازیم. علاوه بر این فرض می‌کنیم که همهٔ متغیرهای تصادفی دارای توزیع نرمال با میانگین و انحراف معیار معلوم هستند.

هنگامی‌که تنها a_{ij} ها متغیرهای تصادفی هستند

فرض کنید $\bar{a_{ij}}$، میانگین و $\mathrm{Var}(a_{ij}) = \sigma^۲$ واریانس متغیر تصادفی a_{ij} با توزیع نرمال باشد. همچنین فرض کنید توزیع چند متغیره a_{ij} ها $(i = ۱, ..., m;\ j = ۱, ..., n)$ و در نتیجه کواریانس بین متغیرهای

تصادفی a_{ij} و a_{kl} یعنی $\mathrm{cov}(a_{ij}, a_{kl})$ معلوم باشد . کمیت d_i را به‌صورت زیر تعریف می‌کنیم:

$$d_i = \sum_{j=۱}^{n} a_{ij} x_j, \quad i = ۱, ۲, ..., m. \tag{۷.۴}$$

چون $a_{i۱}, ..., a_{in}$ دارای توزیع نرمال و $x_۱, ..., x_n$ مقادیر ثابت و نامعلوم هستند ، d_i هم دارای توزیع نرمال با میانگین و واریانس

$$\bar{d_i} = \sum_{j=۱}^{n} \bar{a}_{ij} x_j, \quad i = ۱, ۲, ..., m, \tag{۸.۴}$$

$$\mathrm{Var}(d_i) = \sigma^{۲}_{d_i} = X^T V_i X, \tag{۹.۴}$$

است، که در آن V_i ، i امین ماتریس کواریانس، به‌صورت زیر تعریف می‌شود:

$$V_i = \begin{bmatrix} \mathrm{Var}(a_{i۱}) & \mathrm{Cov}(a_{i۱}, a_{i۲}) & ... & \mathrm{Cov}(a_{i۱}, a_{in}) \\ \mathrm{Cov}(a_{i۲}, a_{i۱}) & \mathrm{Var}(a_{i۲}) & ... & \mathrm{Cov}(a_{i۲}, a_{in}) \\ \vdots & \vdots & & \\ \mathrm{Cov}(a_{in}, a_{i۱}) & \mathrm{Cov}(a_{in}, a_{i۲}) & ... & \mathrm{Var}(a_{in}) \end{bmatrix}. \tag{۱۰.۴}$$

قیود (۵.۴) را می‌توان به‌صورت زیر بیان کرد:

$$\mathrm{P}[d_i \leq b_i] \geq p_i,$$

یعنی

$$\mathrm{P}[\frac{d_i - \bar{d_i}}{\sqrt{\mathrm{Var}(d_i)}} \leq \frac{b_i - \bar{d_i}}{\sqrt{\mathrm{Var}(d_i)}}] \geq p_i \quad i = ۱, ۲, .., m, \tag{۱۱.۴}$$

وقتی $\frac{d_i - \bar{d_i}}{\sqrt{\mathrm{Var}(d_i)}}$ یک متغیر تصادفی[۳۴] نرمال استاندارد با میانگین صفر و واریانس یک است. بنابراین احتمال اینکه d_iکوچکتر یا مساوی با b_i باشد را می‌توان به‌صورت زیر نوشت:

$$\mathrm{P}[d_i \leq b_i] = \phi(\frac{b_i - \bar{d_i}}{\sqrt{\mathrm{Var}(d_i)}}), \quad i = ۱, ..., m, \tag{۱۲.۴}$$

[۳۴] Variate

که‌در آن $\phi(x)$ بیانگر تابع توزیع تجمعی، توزیع نرمال استاندارد در نقطه x است. اگر e_i نشان دهندهٔ مقدار توزیع نرمال استانداردی باشد که در آن

$$\phi(e_i) = p_i \geq ۰/۵ \quad (e_i \geq ۰), \tag{۱۳.۴}$$

آنگاه قیود رابطهٔ (۱۱.۴) را می توان به‌صورت زیر بیان کرد:

$$\phi(\frac{b_i - \bar{d}_i}{\sqrt{\mathrm{Var}(d_i)}}) \geq \phi(e_i), \quad i = ۱, ۲, ..., m, \tag{۱۴.۴}$$

این نامساوی تنها وقتی برقرار است که

$$(\frac{b_i - \bar{d}_i}{\sqrt{\mathrm{Var}(d_i)}}) \geq e_i,$$

یا

$$\bar{d}_i + e_i\sqrt{\mathrm{Var}(d_i)} - b_i \leq ۰, \quad i = ۱, ۲, ..., m. \tag{۱۵.۴}$$

با قرار دادن روابط (۸.۴) و (۹.۴) در (۱۵.۴)، داریم

$$\sum_{j=۱}^{n} \bar{a}_{ij}x_j + e_i\sqrt{X^T V_i X} - b_i \leq ۰, \quad i = ۱, ۲, ..., m. \tag{۱۶.۴}$$

اینها قیود غیرخطی قطعی معادل با قیود خطی تصادفی هستند. بنابراین جواب مسأله برنامه‌ریزی احتمالی بیان شده در (۴.۴)-(۶.۴) را می‌توان با حل مسأله قطعی زیر به دست آورد.

$$\text{Minimize} \qquad f(x) = \sum_{j=۱}^{n} c_j x_j \tag{۱۷.۴}$$

subject to

$$\begin{cases} \sum_{j=۱}^{n} \bar{a}_{ij}x_j + e_i\sqrt{X^T V_i X} - b_i \leq ۰, & i = ۱, ۲, ..., m, \\ x_j \geq ۰, & j = ۱, ۲, ..., n. \end{cases} \tag{۱۸.۴}$$

اگر متغیرهای تصادفی با توزیع نرمال a_{ij} مستقل باشند، رابطه (۱۰.۴) به یک ماتریس قطری به‌صورت زیر کاهش می‌یابد:

$$V_i = \begin{bmatrix} \mathrm{Var}(a_{i1}) & 0 & \dots & 0 \\ 0 & \mathrm{Var}(a_{i2}) & \dots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \dots & \mathrm{Var}(a_{in}) \end{bmatrix}. \qquad (۱۹.۴)$$

در این حالت قیود رابطهٔ (۱۶.۴) به شکل زیر ساده می‌شوند:

$$\sum_{j=1}^{n} \bar{a}_{ij}x_j + e_i\sqrt{\sum_{j=1}^{n}[\mathrm{Var}(a_{ij})x_j^2]} - b_i \le 0, \quad i = 1, 2, ..., m. \qquad (۲۰.۴)$$

هنگامی‌که تنها b_i ها متغیرهای تصادفی هستند

فرض کنید $\bar{b_i}$ و $Var(b_i)$ به‌ترتیب نشان دهنده میانگین و واریانس متغیر تصادفی b_i با توزیع نرمال باشند. ضرایب قیود (۵.۴) را می‌توان به‌صورت زیر بیان کرد:

$$\begin{aligned} \mathrm{P}[\sum_{j=1}^{n} a_{ij}x_j \le b_i] &= P[\frac{\sum_{j=1}^{n} a_{ij}x_j - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \le \frac{b_i - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}] \\ &= \mathrm{P}[\frac{b_i - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \ge \frac{\sum_{j=1}^{n} a_{ij}x_j - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}] \ge p_i \quad i = 1, 2, ..., m, \end{aligned} \qquad (۲۱.۴)$$

که در آن $\frac{b_i - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}$ یک متغیر نرمال استاندارد با میانگین صفر و واریانس یک است. نامساوی (۲۱.۴) را به‌صورت زیر هم می‌توان نوشت:

$$\mathrm{P}[\frac{\sum_{j=1}^{n} a_{ij}x_j - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \le \frac{b_i - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}] \le 1 - p_i, \quad i = 1, 2, ..., m. \qquad (۲۲.۴)$$

اگر E_i بیانگر مقدار تغییر استانداردی باشد که در آن

$$\phi(E_i) = 1 - p_i < 0/5, \quad (E_i < 0)$$

قیود رابطهٔ (۲۲.۴) را می‌توان به‌صورت زیر بیان کرد:

$$\phi\left(\frac{\sum_{j=1}^{n} a_{ij}x_j - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}}\right) \le \phi(E_i), \quad i = 1, 2, ..., m. \qquad (۲۳.۴)$$

این نامساوی تنها وقتی برقرار است که:

$$\frac{\sum_{j=۱}^{n} a_{ij}x_j - \bar{b}_i}{\sqrt{\mathrm{Var}(b_i)}} \leq E_i < ۰, \quad i = ۱, ۲, ..., m,$$

یا

$$\sum_{j=۱}^{n} a_{ij}x_j - \bar{b}_i - E_i\sqrt{\mathrm{Var}(b_i)} \leq ۰, \quad i = ۱, ۲, ..., m. \tag{۲۴.۴}$$

بنابراین مسالهٔ برنامه‌ریزی خطی بیان شده در (۴.۴)-(۶.۴) با مسأله قطعی زیر معادل است:

$$\text{Minimize} \qquad f(x) = \sum_{j=۱}^{n} c_j x_j \tag{۲۵.۴}$$

subject to

$$\begin{cases} \sum_{j=۱}^{n} a_{ij}x_j - \bar{b}_i - E_i\sqrt{\mathrm{Var}(b_i)} \leq ۰, & i = ۱, ۲, ..., m \\ x_j \geq ۰, & j = ۱, ۲, ..., n \end{cases} \tag{۲۶.۴}$$

هنگامی‌که تنها c_j ها متغیرهای تصادفی هستند

چون c_j ها متغیرهایی تصادفی با توزیع نرمال هستند، تابع هدف $f(x)$ هم یک متغیر تصادفی با توزیع نرمال است. میانگین و واریانس f عبارتند از:

$$\bar{f} = \sum_{j=۱}^{n} \bar{c}_j x_j, \tag{۲۷.۴}$$

$$\mathrm{Var}(f) = X^T V X, \tag{۲۸.۴}$$

وقتیکه $\bar{c}_j$ ، میانگین c_j و ماتریس کواریانس V به‌صورت زیر تعریف می‌شود:

$$V = \begin{bmatrix} \mathrm{Var}(c_۱) & \mathrm{Cov}(c_۱, c_۲) & ... & \mathrm{Cov}(c_۱, c_n) \\ \mathrm{Cov}(c_۲, c_۱) & \mathrm{Var}(c_۲) & ... & \mathrm{Cov}(c_۲, c_n) \\ \vdots & \vdots & & \\ \mathrm{Cov}(c_n, c_۱) & \mathrm{Cov}(c_n, c_۲) & ... & \mathrm{Var}(c_n) \end{bmatrix}. \tag{۲۹.۴}$$

حال یک تابع هدف کمینه‌سازی قطعی جدید را می‌توان به‌صورت زیر فرمولبندی کرد:

$$F(x) = k_۱\bar{f} + k_۲\sqrt{\mathrm{Var}(f)}, \tag{۳۰.۴}$$

وقتیکه k_1 و k_2 مقادیر ثابت نامنفی هستند که مقادیر آنها نشان دهندهٔ اهمیت نسبی $\bar{f}$، و انحراف معیار f برای کمینه سازی است. بنابراین $k_2 = 0$ نشان می‌دهد که مقدار مورد انتظار fباید بدون توجه به انحراف معیار f کمینه سازی شود. از طرف دیگر اگر $k_1 = 0$ باشد، نشان می دهد که باید کمینه سازی قابلیت تغییر f حول میانگینش را بدون توجه به مقدار میانگین f انجام دهیم. به همین ترتیب، اگر $k_1 = k_2 = 1$ باشد کمینه سازی میانگین و انحراف معیار f از اهمیت یکسانی برخوردار هستند. توجه کنید که تابع هدف جدید که در رابطهٔ (۳۰.۴) بیان شده است، از دیدگاه واریانس f ، یک تابع غیر خطی از x است.

بنابراین جواب، مسالهٔ برنامه‌ریزی خطی تصادفی (۴.۴)-(۶.۴) را می‌توان با حل مسأله برنامه‌ریزی غیرخطی معادل زیر به‌دست آورد:

$$\text{Minimize} \quad F(x) = k_1 \sum_{j=1}^{n} \bar{c}_j x_j + k_2 \sqrt{X^T V X} \tag{۳۱.۴}$$

subject to

$$\begin{cases} \sum_{j=1}^{n} a_{ij} x_j - b_i \leq 0, & i = 1, 2, ..., m, \\ x_j \geq 0, & j = 1, 2, ..., n. \end{cases} \tag{۳۲.۴}$$

اگر همهٔ متغیرهای تصادفی c_j مستقل خطی باشند، تابع هدف به‌صورت زیر در می‌آید:

$$F(x) = k_1 \sum_{j=1}^{n} \bar{c}_j x_j + k_2 \sqrt{\sum_{j=1}^{n} \mathrm{Var}(c_j) x_j^2}. \tag{۳۳.۴}$$

هنگامی‌که c_j ، a_{ij} و b_i ها متغیرهای تصادفی هستند

چون متغیرهای تصادفی c_j تنها در تابع هدف ظاهر می‌شوند، می‌توانیم تابع هدف جدید $F(X)$ را به‌همان صورت داده شده در رابطهٔ (۳۰.۴) در نظر بگیریم. قیود (۵.۴) را به‌صورت زیر می‌توان بیان کرد:

$$\mathrm{P}[h_i \leq 0] \geq p_i \qquad i = 1, 2, ..., m, \tag{۳۴.۴}$$

که درآن h_i یک متغیر تصادفی جدید است که به‌صورت زیر تعریف می‌شود

$$h_i = \sum_{j=۱}^{n} a_{ij}x_j - b_i = \sum_{k=۱}^{n+۱} q_{ik}y_k, \tag{۳۵.۴}$$

که درآن

$$\begin{cases} q_{ik} = a_{ik}, & k = ۱, ۲, ..., n, \\ q_{i,n+۱} = b_i, & \\ y_k = x_k, & k = ۱, ۲, ..., n, \\ y_{n+۱} = -۱. & \end{cases}$$

توجه کنید که ثابت $y_{n+۱}$ برای راحتی کار تعریف می‌شود. چون h_i به‌صورت ترکیب خطی متغیرهای تصادفی q_{ik} با توزیع نرمال داده می‌شود، پس دارای توزیع نرمال است. میانگین و واریانس h_i عبارتند از:

$$\bar{h}_i = \sum_{k=۱}^{n+۱} \bar{q}_{ik}y_k = \sum_{j=۱}^{n} \bar{a}_{ij}x_j - \bar{b}_i, \tag{۳۶.۴}$$

$$\mathrm{Var}(h_i) = Y^T V_i Y, \tag{۳۷.۴}$$

که درآن

$$V_i = \begin{bmatrix} \mathrm{Var}(q_{i۱}) & \mathrm{Cov}(q_{i۱}, q_{i۲}) & ... & \mathrm{Cov}(q_{i۱}, q_{i,n+۱}) \\ \mathrm{Cov}(q_{i۲}, q_{i۱}) & \mathrm{Var}(q_{i۲}) & ... & \mathrm{Cov}(q_{i۲}, q_{i,n+۱}) \\ \vdots & & & \\ \mathrm{Cov}(q_{i,n+۱}, q_{i۱}) & \mathrm{Cov}(q_{i,n+۱}, q_{i۲}) & ... & \mathrm{Var}(q_{i,n+۱}) \end{bmatrix}.$$

این را می‌توان به‌صورت صریح‌تر زیر نوشت:

$$\begin{aligned} \mathrm{Var}(h_i) &= \sum_{k=۱}^{n+۱}[y_k^۲\mathrm{Var}(q_{ik}) + ۲\sum_{l=k+۱}^{n+۱} y_k y_l \mathrm{Cov}(q_{ik}, q_{il})] \\ &= \sum_{k=۱}^{n}[y_k^۲\mathrm{Var}(q_{ik}) + ۲\sum_{l=k+۱}^{n} y_k y_l \mathrm{Cov}(q_{ik}, q_{il})] \\ &+ y_{n+۱}^۲\mathrm{Var}(q_{i,n+۱}) + \sum_{l=۱}^{n}[۲y_l y_{n+۱}\mathrm{Cov}(q_{il}, q_{i,n+۱})] \\ &= \sum_{k=۱}^{n}[x_k^۲\mathrm{Var}(a_{ik}) + ۲\sum_{l=k+۱}^{n} x_k x_l \mathrm{Cov}(a_{ik}, a_{il})] \\ &+ \mathrm{Var}(b_i) - ۲\sum_{k=۱}^{n} x_k \mathrm{Cov}(a_{ik}, b_i). \end{aligned} \tag{۳۸.۴}$$

بنابراین قیود روابط (۳۴.۴) را می‌توان به‌صورت زیر بیان کرد:

$$\mathrm{P}[\frac{h_i-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}} \leq \frac{-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}}] \geq p_i, \quad i=۱,۲,...,m, \tag{۳۹.۴}$$

که درآن $[\frac{h_i-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}}]$ بیان کنندهٔ یک متغیر نرمال استاندارد با میانگین صفر و واریانس یک است. بنابراین اگر e_i نشان دهندهٔ مقدار متغیر استانداردی باشد که در آن

$$\phi(e_i)=p_i \geq ۰/۵ \ \ (e_i \geq ۰), \tag{۴۰.۴}$$

قیود رابطهٔ (۳۹.۴) قابل بیان به‌صورت زیر هستند:

$$\phi(\frac{-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}}) \geq \phi(e_i) \quad i=۱,۲,...,m. \tag{۴۱.۴}$$

این نامساویها تنها در صورتی برقرارند که، نامساویهای غیرخطی زیر برقرار باشند:

$$\frac{-\bar{h}_i}{\sqrt{\mathrm{Var}(h_i)}} \geq e_i > ۰, \quad i=۱,۲,...,m,$$

یا

$$\bar{h}_i + e_i\sqrt{\mathrm{Var}(h_i)} \leq ۰, \quad i=۱,۲,...,m. \tag{۴۲.۴}$$

بنابراین مسأله برنامه‌ریزی احتمالی (۴.۴)-(۶.۴) را می‌توان به‌عنوان یک مسأله برنامه‌ریزی غیرخطی قطعی معادل زیر نوشت:

$$\text{Minimize} \quad F(x)=k_۱\sum_{j=۱}^{n}\bar{c}_j x_j + k_۲\sqrt{X^TVX}, \ \ k_۱,k_۲ \geq ۰ \tag{۴۳.۴}$$

subject to

$$\begin{cases} \bar{h}_i + e_i\sqrt{\mathrm{Var}(h_i)} \leq ۰, & i=۱,۲,...,m, \\ x_j \geq ۰, & j=۱,۲,...,n. \end{cases} \tag{۴۴.۴}$$

از تحلیل بالا و [۱۲۵] ، واضح است که مسائل بهینه سازی با قیود تصادفی قابل تبدیل به مسائل برنامه ریزی مخروط مرتبه دو محدب [۳۵] هستند. بنابراین ، ما فرم کلی از مسائل مسائل برنامه ریزی مخروط مرتبه دو محدب را به‌صورت زیر در نظر می‌گیریم:

$$\text{Minimize} \quad f(x) \tag{۴۵.۴}$$

$$\text{subject to}$$

$$Ex + d \succeq_{\mathcal{K}} ۰, \tag{۴۶.۴}$$

$$x \geq ۰. \tag{۴۷.۴}$$

وقتیکه $f : \mathbb{R}^n \to \mathbb{R} \bigcup \{+\infty\}$، تابع سره بستهٔ محدب ، E یک ماتریس $m \times n$ با $m \geq n$ ، d برداری در $\mathbb{R}^m$ و $x \succeq_{\mathcal{K}} ۰$ یعنی $x \in \mathcal{K}$ و $\mathcal{K}$ حاصلضرب دکارتی از مخروط‌های مرتبه دوم که مخروط‌های لورنتس یا مخروط بستنی هم نامیده می‌شوند . به عبارت دیگر

$$\mathcal{K} = \mathcal{K}^{m_۱} \times \mathcal{K}^{m_۲} \times ... \times \mathcal{K}^{m_r},$$

که درآن $r, m_۱, ..., m_r \geq ۱$ با $m_۱ + ... + m_r = n$ و

$$\mathcal{K}^{m_i} := \{(x_۱, x_۲)^T \in \mathbb{R} \times \mathbb{R}^{m_i - ۱} \mid x_۱ \geq \|x_۲\|\}, \tag{۴۸.۴}$$

که درآن $\|.\|$ نرم اقلیدسی و $\mathcal{K}^۱$ مجموعه از $\mathbb{R}_+$ حقیقی نامنفی است. در قطعه بعد ما سعی داریم یک مدل شبکه عصبی با عملکرد بالا برای حل مسأله (۴۵.۴)-(۴۷.۴) پیشنهاد دهیم .

۳.۴ مدل شبکه عصبی

ما قیود غیرهموار مخروط مرتبه دوم $x \succeq_{\mathcal{K}^{m_i}} ۰$ را برای $(i = ۱, ..., r)$ با یک هموار سازی از [۱۲۶] به‌صورت زیر جایگزین می‌کنیم:

$$\sqrt{\epsilon^۲ + \|x_۲\|^۲} \leq x_۱, \tag{۴۹.۴}$$

[۳۵] CSOCP

که درآن ϵ ثابت کوچکی است که اغلب حدود $۱۰^{-۶}$ انتخاب می‌شود. اغتشاش به‌دست آمده در قید مخروط مرتبه دوم ، باعث می‌شود قیود مسأله (۴۵.۴)-(۴۷.۴) محدب و هموار شود. بنابراین مسأله محدب و غیر هموار (۴۵.۴)-(۴۶.۴) می‌تواند به یک مسأله بهینه‌سازی محدب هموار معادل به‌صورت زیر تبدیل شود:

$$\text{Minimize} \qquad f(x) \tag{۵۰.۴}$$

$$\text{subject to}$$

$$g(x) \leq ۰. \tag{۵۱.۴}$$

که درآن $g(x) = \begin{bmatrix} h(x) \\ x \end{bmatrix}$ و $h(x) \in \mathbb{R}^m$, یک هموارسازی از $Ex + d \succeq_{\mathcal{K}} ۰$ با روش ϵ–اغتشاش است. واضح است که توابع برداری $h(x)$ محدب و دوبار مشتق‌پذیر هستند.

از [۱]، می‌بینیم که $x^* \in \mathbb{R}^n$ یک جواب بهینه از (۵۰.۴)-(۵۱.۴) اگر و تنها اگر بردار $u^* \in \mathbb{R}^m$ وجود داشته باشد به‌طوری‌که $(x^{*T}, u^{*T})^T$ در سیستم $K.K.T$ زیر صدق کند:

$$\begin{cases} u^* \geq ۰, \;\; g(x^*) \leq ۰, \;\; u^{*T} g(x^*) = ۰, \\ \nabla f(x^*) + \nabla g(x^*)^T u^* = ۰. \end{cases} \tag{۵۲.۴}$$

x^* را نقطه $K.K.T$ از (۵۰.۴)-(۵۱.۴) و u^* را بردار ضریب لاگرانژ متناظر با x^* نامیم. واضح است که x^* جواب بهینه از (۵۰.۴)-(۵۱.۴) است، اگر و تنها اگر x^* نقطه $K.K.T$ از (۵۰.۴)-(۵۱.۴) باشد.

اکنون فرض کنید که $x(.)$ و $u(.)$ متغیرهای وابسته به زمان باشند. هدف ما بیان یک ساختار دینامیکی زمان پیوسته است که به نقطه $K.K.T$ مسأله (۵۰.۴)-(۵۱.۴) و دوگان آن هدایت شود. ما مدل شبکه عصبی بازگشتی برای حل (۵۰.۴)-(۵۱.۴) و دوگان آن را با معادله دینامیکی به‌صورت زیر بیان می‌کنیم:

$$\frac{dx}{dt} = -\left(\nabla f(x) + \nabla g(x)^T (u + g(x))^+\right), \tag{۵۳.۴}$$

$$\frac{du}{dt} = (u + g(x))^+ - u, \tag{۵۴.۴}$$

با نقطه آغازین $y_۰ = (x_۰^T, u_۰^T)^T$، که درآن

$$(u + g(x))^+ = ([u_۱ + g_۱(x)]^+, [u_۲ + g_۲(x)]^+, ..., [u_m + g_m(x)]^+),$$

$$[u_k + g_k(x)]^+ = \max\{u_k + g_k(x), ۰\},\ k = ۱, ۲, ..., m.$$

برای سادگی تعریف می‌کنیم، $y = (x^T, u^T)^T \in \mathbb{R}^{n+m}$ ، D^* مجموعه نقاط بهینه از (۵۰.۴)-(۵۱.۴) و دوگان آن ،

$$\Phi(y) = \begin{bmatrix} -\left(\nabla f(x) + \nabla g^T (u + g(x))^+\right) \\ (u + g(x))^+ - u \end{bmatrix}. \tag{۵۵.۴}$$

بنابراین شبکه عصبی (۵۳.۴)-(۵۴.۴) را می‌توان به‌صورت زیر نوشت:

$$\frac{dy}{dt} = \kappa \Phi(y), \tag{۵۶.۴}$$

$$y(t_۰) = y_۰. \tag{۵۷.۴}$$

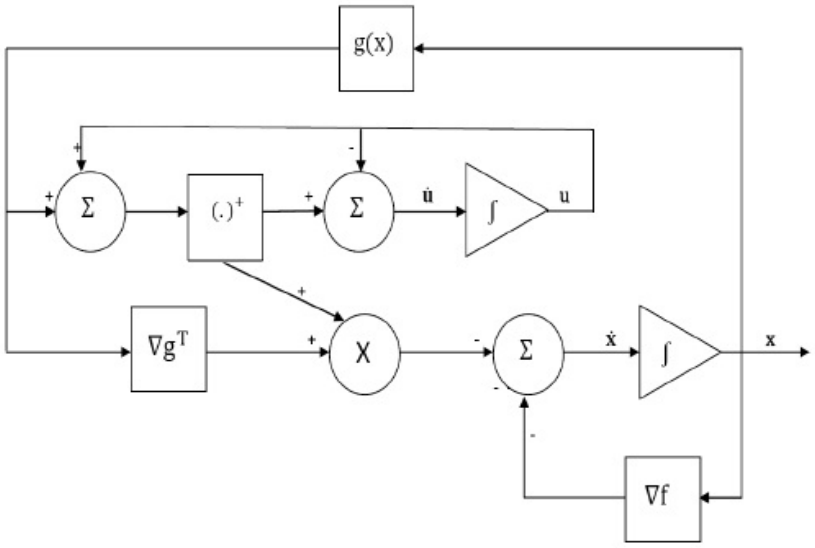
که در آن κ پارامتر اسکالر [۳۶] است و نرخ همگرایی شبکه عصبی (۵۶.۴)-(۵۷.۴) را نشان می‌دهد. برای سادگی، $\kappa = ۱$ در نظر می‌گیریم. شکل (۱.۴) بیانگر این است که شبکه فوق روی سخت افزار چگونه اجرا می‌شود.

۴.۴ تحلیل پایداری و همگرایی

در این قطعه پایداری و همگرایی شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) بررسی می‌شود.

قضیه ۱.۴.۴. فرض کنید $y^* = (x^{*T}, u^{*T})^T$ ، نقطه تعادل از شبکه عصبی (۵۷.۴)-(۵۶.۴) باشد. آنگاه x^* نقطه $K.K.T$ از مسأله (۵۱.۴)-(۵۰.۴) ، برعکس اگر $x^* \in \mathbb{R}^n$ جواب بهینه مسأله (۵۱.۴)-(۵۰.۴)

[۳۶] Scale



شکل ۱.۴: بلوک دیاگرام شبکه عصبی (۵۶.۴)-(۵۷.۴)

باشد، آنگاه بردار $u^* \in \mathbb{R}^m$ وجود دارد به‌طوری که $y^* = (x^{*T}, u^{*T})^T$ نقطه تعادل از شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) است.

اثبات. فرض کنید $y^* = (x^{*T}, u^{*T})^T$ نقطه تعادل شبکه عصبی (۵۶.۴)-(۵۷.۴) باشد، آنگاه $\frac{dx^*}{dt} = ۰$ و $\frac{du^*}{dt} = ۰$. در نتیجه

$$\nabla f\left(x^*\right) + \nabla g(x^*)^T (u^* + g(x^*))^+ = ۰, \tag{۵۸.۴}$$

$$(u^* + g(x^*))^+ = u^*. \tag{۵۹.۴}$$

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

$$u^* \geq ۰, \quad g(x^*) \leq ۰, \quad u^{*T} g(x^*) = ۰. \tag{۶۰.۴}$$

بنابراین، با جایگذاری (۵۹.۴) در (۵۸.۴) داریم

$$\nabla f(x^*) + \nabla g(x^*)^T u^* = ۰ \tag{۶۱.۴}$$

از (۶۰.۴) و (۶۱.۴) ،نتیجه می‌شود که $y^* = (x^{*T}, u^{*T})^T$ در شرایط $K.K.T$ صدق می‌کند و این اثبات را کامل می‌کند. □

لم ۲.۴.۴. نقطه تعادل مدل شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) یکتاست.

اثبات. چون مسأله (۵۰.۴)-(۵۱.۴) دارای جواب بهینه یکتای x^* است، شرایط لازم وکافی $K.K.T$ دارای جواب یکتای $y^* = (x^{*T}, u^{*T})^T$ هستند، بنابراین از قضیه (۱.۴.۴) می‌بینیم که نقطه تعادل مدل پیشنهاد شده (۵۷.۴)-(۵۶.۴) در سیستم $K.K.T$، (۵۲.۴) صدق می‌کند. بنابراین نقطه تعادل شبکه عصبی (۵۷.۴)-(۵۶.۴) یکتاست. □

لم ۳.۴.۴. برای هر نقطه آغازین $y(t_۰) = (x(t_۰)^T, u(t_۰)^T)^T$ ، سیستم (۵۷.۴)-(۵۶.۴) دارای جواب یکتای پیوسته $y(t) = (x(t)^T, u(t)^T)^T$ است.

اثبات. چون $\nabla f(x)$ و $\nabla g_k(x), (k = ۱, \ldots, m)$ ، توابع به‌طور پیوسته مشتق پذیر روی مجموعه باز $D \subseteq \mathbb{R}^{n+m}$ هستند، بنابراین $\nabla f(x) + \nabla g(x)^T (u + g(x))^+$ و $(u + g(x))^+ - u$ لیپ شیتز محلی پیوسته روی مجموعه باز $D \subseteq \mathbb{R}^{n+m+l}$ هستند، با توجه به قضیه (۸.۳.۱) ، شبکه عصبی (۵۷.۴)-(۵۶.۴) دارای جواب یکتای پیوسته $y(t)$ برای $t \in [t_۰, \tau)$ وقتی $\tau \to \infty$ است. □

لم ۴.۴.۴. ماتریس ژاکوبی $\nabla \Phi(y)$ از نگاشت Φ تعریف شده در (۵۵.۴) نیمه معین منفی است.

اثبات. بدون کاستن ازکلیت مسأله، فرض می‌کنیم $۰ < p < m$ وجود دارد به‌طوری که

$$(u + g)^+ = (u_۱ + g_۱(x), u_۲ + g_۲(x), ..., u_p + g_p(x), \underbrace{۰, ۰, ..., ۰}_{m-p}) \geq ۰.$$

با محاسبات ساده می‌توان نشان داد که

$$\nabla \Phi(y) = \begin{bmatrix} -(\nabla^۲ f(x) + \sum_{k=۱}^{p} ((u_k + g_k) \nabla^۲ g_k^p(x)) + \nabla g^p(x)^T \nabla g^p(x)) & -\nabla g^p(x)^T \\ \nabla g^p(x) & S_{m \times m} \end{bmatrix},$$

که درآن

$$\nabla g^p(x) = \begin{bmatrix} U_{p\times n} \\ O_{(m-p)\times n} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_{p-1}} & \frac{\partial g_1}{\partial x_p} & \frac{\partial g_1}{\partial x_{p+1}} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \dots & \frac{\partial g_2}{\partial x_{p-1}} & \frac{\partial g_2}{\partial x_p} & \frac{\partial g_2}{\partial x_{p+1}} & \dots & \frac{\partial g_2}{\partial x_n} \\ & \dots & & & & \dots & \\ & \dots & & & & \dots & \\ \frac{\partial g_p}{\partial x_1} & \dots & \frac{\partial g_p}{\partial x_{p-1}} & \frac{\partial g_p}{\partial x_p} & \frac{\partial g_p}{\partial x_{p+1}} & \dots & \frac{\partial g_p}{\partial x_n} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & & & & & \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

و $\nabla^2 g_k^p(x)$ برای $(k = 1, 2, ..., p)$

$$\nabla^2 g_k^p(x) = \begin{bmatrix} \frac{\partial^2 g_k}{\partial x_1^2} & \dots & \frac{\partial^2 g_k}{\partial x_1 \partial x_{p-1}} & \frac{\partial^2 g_k}{\partial x_1 \partial x_p} & \frac{\partial^2 g_k}{\partial x_1 \partial x_{p+1}} & \dots & \frac{\partial^2 g_k}{\partial x_1 \partial x_n} \\ \frac{\partial^2 g_k}{\partial x_2 \partial x_1} & \dots & \frac{\partial^2 g_k}{\partial x_2 \partial x_{p-1}} & \frac{\partial^2 g_k}{\partial x_2 \partial x_p} & \frac{\partial^2 g_k}{\partial x_2 \partial x_{p+1}} & \dots & \frac{\partial^2 g_k}{\partial x_2 \partial x_n} \\ & \dots & & & & \dots & \\ & \dots & & & & \dots & \\ \frac{\partial^2 g_k}{\partial x_p \partial x_1} & \dots & \frac{\partial^2 g_k}{\partial x_p \partial x_{p-1}} & \frac{\partial^2 g_k}{\partial^2 x_p} & \frac{\partial^2 g_k}{\partial x_p \partial x_{p+1}} & \dots & \frac{\partial^2 g_k}{\partial x_p \partial x_n} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & & & & & \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

و

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

که درآن O ماتریس صفر است. ماتریس $\nabla g^p(x)^T \nabla g^p(x)$ ، ماتریس نیمه معین مثبت است . چون توابع $f, g_1, \dots, g_m$ محدب و دوبار مشتق پذیر فرض شده، بنابراین ماتریس‌های هسین $\nabla^2 f(x)$ و $\nabla^2 g_k(x)$ برای $(k = 1, 2, ..., p)$، ماتریس‌های نیمه‌معین مثبت هستند. به‌علاوه از معین مثبت بودن $\nabla^2 g_k(x)$ می‌توان گفت که ماتریس‌های $\nabla^2 g_k^p(x)$ نیمه معین مثبت هستند. همچنین واضح است که ماتریس $S_{m\times m}$ ، نیمه معین منفی است.

طبق توضیحات گفته شده، ما می‌توانیم نتیجه بگیریم ماتریس ژاکوبی $\nabla\Phi(y)$ ، نیمه معین منفی است.

اگر $p = m$ یعنی $(u+g)^+ = (u_1 + g_1(x), u_2 + g_2(x), ..., u_m + g_m(x))$ ، آنگاه داریم:

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

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

سرانجام، اگر $p = 0$ یعنی $(u+g)^+ = (\underbrace{0, 0, ..., 0}_{m})$ ، آنگاه داریم:

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

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

قضیه ۵.۴.۴. مدل شبکه عصبی پیشنهاد شده در (۵۶.۴)-(۵۷.۴) ، پایدار مجانبی به مفهوم لیاپانوف و همگرایی سراسری به $y^* = (x^{*T}, u^{*T})^T$ ، وقتی x^* جواب بهینه از (۵۰.۴)-(۵۱.۴) است.

اثبات. تابع لیاپانوف زیر را در نظر بگیرید:

$$E(y) = \|\Phi(y)\|^2 + \frac{1}{2}\|y - y^*\|^2. \qquad (۶۲.۴)$$

از (۵۵.۴) ، داریم:

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

بنابراین

$$\frac{dE(y(t))}{dt} = (\frac{d\Phi}{dt})^T\Phi + \Phi^T(\frac{d\Phi}{dt}) + (y-y^*)^T\frac{dy(t)}{dt} =$$
$$\Phi^T(\nabla\Phi(y)^T + \nabla\Phi(y))\Phi + (y-y^*)^T\Phi(y).$$

با بکار بردن لم (۴.۴.۴) داریم:

$$\Phi^T(y)(\nabla\Phi(y)^T + \nabla\Phi(y))\Phi(y) \leq ۰, \quad \forall y \neq y^*. \tag{۶۳.۴}$$

با استفاده از تعریف (۶.۳.۱) و قضیه (۷.۳.۱) داریم:

$$(y-y^*)^T(\Phi(y) - \Phi(y^*)) = (y-y^*)^T\Phi(y) \leq ۰, \quad \forall y \neq y^*.$$

بنابراین

$$\frac{dE(y(t))}{dt} \leq ۰. \tag{۶۴.۴}$$

این بدان معنی است که شبکه عصبی (۵۷.۴)-(۵۶.۴) ، پایدار مجانبی به مفهوم لیاپانوف است. در ادامه چون

$$E(y) \geq \frac{۱}{۲}\|y-y^*\|^۲, \tag{۶۵.۴}$$

زیردنباله‌های همگرای

$$\{(x(t_k)^T, u(t_k)^T)^T | t_۰ < t_۱ < ... < t_k < t_{k+۱}\},$$

وقتی $t_k \to \infty$ از $k \to \infty$ وجود دارد، به‌طوری‌که

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

وقتی $(\bar{x}^T, \bar{u}^T)^T$ در رابطه زیر صدق می‌کند:

$$\frac{dE(y(t))}{dt} = ۰,$$

این نشان می‌دهد که $(\bar{x}^T, \bar{u}^T)^T$ ، یک نقطه $-\omega$ حدی از $\{(x(t)^T, u(t)^T)^T | t \geq t_0\}$ است. با استفاده از قضیه (۴.۴.۱) داریم $\{(x(t)^T, u(t)^T)^T \to M$ برای $t \to \infty\}$ ، که درآن M بزرگترین مجموعه ناوردا در $\{\frac{dE(y(t))}{dt} = 0\}$ است. از (۵۶.۴)-(۵۷.۴) و (۶۴.۴) داریم $K = \{(x(t)^T, u(t)^T)^T |$

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

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

با جایگذاری $x^* = \bar{x}$ و $u^* = \bar{u}$ در(۶۲.۴) ، ما تابع لیاپانوف دیگری به‌صورت زیر تعریف می‌کنیم،

$$\bar{E}(y) = \|\Phi(y)\|^2 + \frac{1}{2}\|y - \bar{y}\|^2. \qquad (۶۶.۴)$$

آنگاه $\bar{E}(y)$ به‌طور پیوسته مشتق پذیر و $\bar{E}(\bar{y}) = 0$. از طرفی

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

از اینرو داریم

$$\lim_{k\to\infty} \bar{E}(x(t_k)^T, u(t_k)^T)^T = \bar{E}(\bar{x}, \bar{u}).$$

بنابراین، به ازای هر $\epsilon > 0$ ، $q > 0$ وجود دارد به‌طوری‌که برای هر $t \geq t_q$ ، داریم $\bar{E}(y(t)) < \epsilon$. مشابها می‌توانیم به دست بیاوریم $\frac{d\bar{E}(y(t))}{dt} \leq 0$. برای $t \geq t_q$ داریم،

$$\frac{1}{2}\|y(t) - \bar{y}\|^2 \leq \bar{E}(y(t)) \leq \epsilon.$$

در نتیجه $\lim_{t\to\infty} \|y(t) - \bar{y}\| = 0$ یا $\lim_{t\to\infty} y(t) = \bar{y}$. بنابراین، شبکه عصبی پیشنهاد شده در (۵۶.۴)-(۵۷.۴) همگرای سراسری به نقطه تعادل $\bar{y} = (\bar{x}^T, \bar{u}^T)^T$ است که در آن $\bar{x}$ جواب بهینه از (۵۰.۴)-(۵۱.۴) است. □

ما از قضیه (۵.۴.۴) ،گزاره زیر را نتیجه می گیریم.

گزاره ۶.۴.۴. اگر $D^* = \{(x^{*T}, u^{*T})^T\}$ ، آنگاه شبکه عصبی (۵۶.۴)-(۵۷.۴) برای حل کردن (۵۰.۴)-(۵۱.۴) ، پایدار مجانبی سراسری به نقطه تعادل یکتای $y^* = (x^{*T}, u^{*T})^T$ است.

۵.۴ مثال‌های عددی

به منظور نشان دادن کارایی و تأثیر شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) ، در این قطعه، ما به ذکر چندین مثال عددی می‌پردازیم. برای برخی مثال‌ها، ما مقایسه عملکرد عددی از شبکه عصبی پیشنهادی را با مقادیر گوناگون از حالت اولیه $y(۰)$ داریم. برای برخی دیگر از مسائل ما مقایسه عملکرد عددی از شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) با مقادیر گوناگون از نرخ همگرایی κ و مقادیر مختلف $\|x(t) - x^*\|^۲$ با نقطه آغازین مشخص $y_۰$ انجام دادیم. نتایج شبیه‌سازی در پکیج $ode۴۵s$، نرم افزار $Matlab$ انجام شده و جواب واقعی مسائل نیز توسط نرم افزار $Lingo۱۱$ محاسبه شده است.

مثال ۱.۵.۴. یک بنگاه تولیدی دو قطعه را با استفاده از دستگاههای تراش، فرز و مته تولید می‌کند. زمان موجود دستگاه‌های مختلف در هفته، سود هر قطعه در جدول زیر داده شده است. زمان ماشینکاری مورد نیاز هر قطعه بر روی دستگاهای مختلف به طور دقیق معلوم نیست (به کارگر مربوط بستگی دارد)، اما می‌دانیم که دارای توزیع نرمال با میانگین و انحراف معیار داده شده در جدول (۱.۴) است. تعداد قطعات I و II را که باید در هفته تولید شود به‌گونه‌ای تعیین کنید که سود بیشینه شود و در زمان صرف شده برای ماشینکاری بیش از ۱% از مقدار موجود تجاوز نکند.

حل: اگر تعداد قطعات I و IIتولید شده در هفته را به‌ترتیب با $x_۱$ و $x_۲$ نشان دهیم، مسأله را می‌توان به‌صورت زیر بیان کرد:

$$\text{Maximize} \qquad f = ۵۰x_۱ + ۱۰۰x_۲ \tag{۶۷.۴}$$

جدول ۱.۴: مقادیر مثال ۱.۵.۴

	زمان ماشینکاری مورد نیاز هر قطعه(دقیقه)				
نوع دستگاه	قطعه I		قطعه II		حداکثر زمان
	میانگین	انحراف معیار	میانگین	انحراف معیار	موجود در هفته (دقیقه)
تراش	$\bar{a}_{۱۱} = ۱۰$	$\sigma_{a_{۱۱}} = ۶$	$\bar{a}_{۱۲} = ۵$	$\sigma_{a_{۱۲}} = ۴$	$b_۱ = ۲۵۰۰$
فرز	$\bar{a}_{۲۱} = ۴$	$\sigma_{a_{۲۱}} = ۴$	$\bar{a}_{۲۲} = ۱۰$	$\sigma_{a_{۲۲}} = ۷$	$b_۲ = ۲۰۰۰$
مته	$\bar{a}_{۳۱} = ۱$	$\sigma_{a_{۳۱}} = ۲$	$\bar{a}_{۳۲} = ۱/۵$	$\sigma_{a_{۳۲}} = ۳$	$b_۳ = ۴۵۰$
سود هر واحد(R_s)	$c_۱ = ۵۰$		$c_۲ = ۱۰۰$		

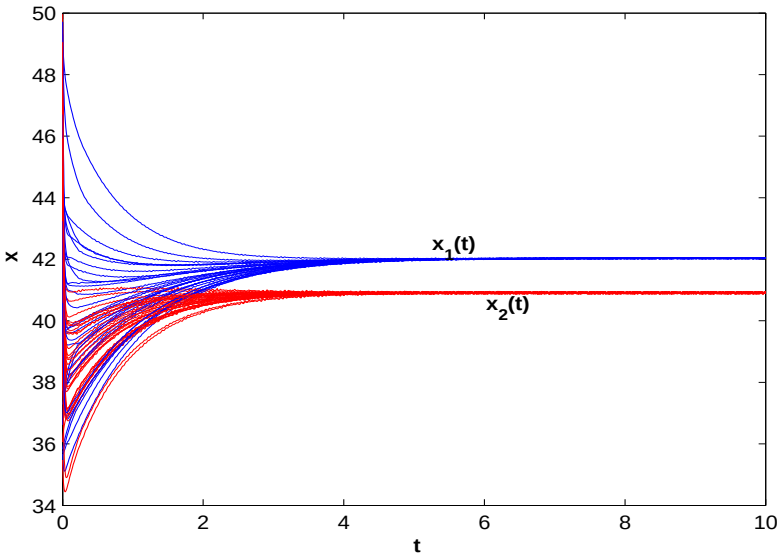
$$\text{subject to} \quad \begin{cases} \mathrm{P}[a_{۱۱}x_۱ + a_{۱۲}x_۲ \le ۲۵۰۰] \ge ۰/۹۹, \\ \mathrm{P}[a_{۲۱}x_۱ + a_{۲۲}x_۲ \le ۲۰۰۰] \ge ۰/۹۹, \\ \mathrm{P}[a_{۳۱}x_۱ + a_{۳۲}x_۲ \le ۴۵۰] \ge ۰/۹۹, \\ x_۱, x_۲ \ge ۰. \end{cases} \tag{۶۸.۴}$$

ار آنجا که اطلاعاتی درباره کواریانس داده نشده است، ضرایب a_{ij} را می‌توان به عنوان متغیرهای مستقل در نظر گرفت. مقدار متغیر تصادفی نرمال استاندارد e_i برای $(i = ۱, .., ۳)$ مربوط به احتمال ۹۹% را می‌توان از جدول توزیع نرمال استاندارد در [۱۲۷] به‌دست آورد. بنابراین معادل قطعی مسأله (۶۷.۴)-(۶۸.۴) با کمک (۱۷.۴)-(۱۸.۴) به‌صورت بیان می‌شود:

$$\text{Maximize} \quad f = ۵۰x_۱ + ۱۰۰x_۲ \tag{۶۹.۴}$$

$$\text{subject to} \quad \begin{cases} ۱۰x_۱ + ۵x_۲ + ۲/۳۳\sqrt{۳۶x_۱^۲ + ۱۶x_۲^۲} - ۲۵۰۰ \le ۰, \\ ۴x_۱ + ۱۰x_۲ + ۲/۳۳\sqrt{۱۶x_۱^۲ + ۴۹x_۲^۲} - ۲۰۰۰ \le ۰, \\ x_۱ + ۱/۵x_۲ + ۲/۳۳\sqrt{۴x_۱^۲ + ۹x_۲^۲} - ۴۵۰ \le ۰, \\ x_۱, x_۲ \ge ۰. \end{cases} \tag{۷۰.۴}$$

جواب بهینه این مسأله $x^* = (۴۲/۰۳۲۲۷, ۴۰/۹۰۸۹۳)^T$ است. ما برای حل این مسأله شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) را به‌کار می‌بریم. شکل (۲.۴) نشان می‌دهد که مسیرهای شبکه عصبی پیشنهادی با ۳۰ نقطه آغازین متفاوت و $\epsilon = ۱۰^{-۶}$ همگرا به جواب بهینه مسأله (۶۹.۴)-(۷۰.۴) است. منحنی فاز ، شبکه پیشنهادی با ۱۴حالت اولیه گوناگون در شکل (۳.۴) نشان داده شده است. ما همچنین



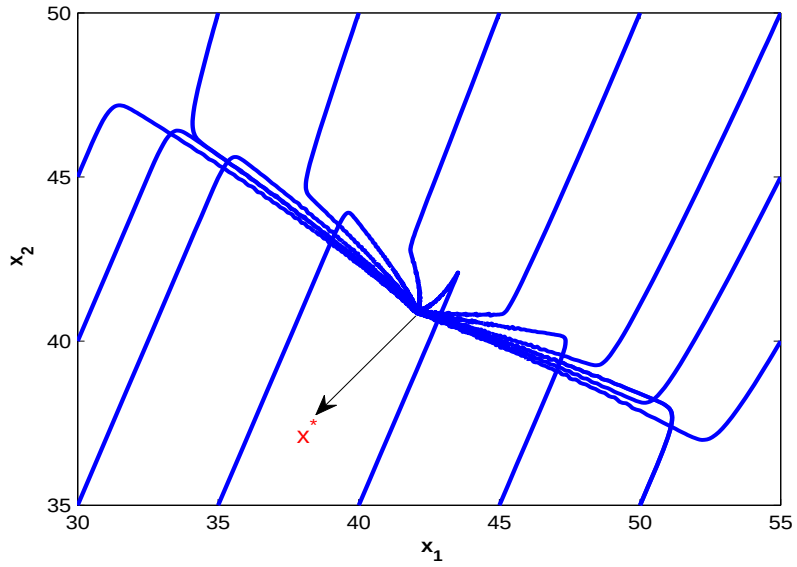
شکل ۲.۴: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۳۰ نقطه آغازین دلخواه در مثال ۱.۵.۴.

تأثیر پارامتر κ از شبکه عصبی (۵۶.۴)-(۵۷.۴) روی مقدار $\|x(t)-x^*\|^2$ بررسی می‌کنیم. در شکل (۴.۴)، مشاهده می‌کنیم κ بزرگتر، نرخ همگرایی بهتری برای $\|x(t)-x^*\|^2$ را نتیجه می‌دهد.

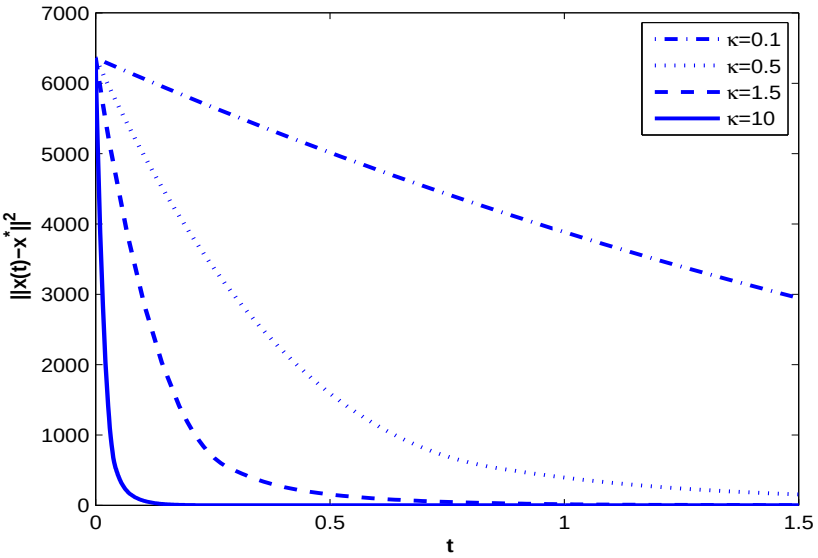
مثال ۲.۵.۴. اگر زمانهای ماشینکاری موجود بر روی دستگاه های مختلف، احتمالی(با توزیع نرمال) و پارامترهای مشخص شده در جدول (۲.۴) باشند، تعداد قطعات I و II را که باید در هفته ساخته شوند تا سود بیشینه شود، پیدا کنید.

جدول ۲.۴: مقادیر مثال ۲.۵.۴

نوع دستگاه	زمان ماشینکاری مورد نیاز هر قطعه(دقیقه)		حداکثر زمان موجود در هفته(دقیقه)	
	قطعه I	قطعه II	میانگین	انحراف معیار
تراش	$a_{۱۱}=۱۰$	$a_{۱۲}=۵$	$\bar{b}_۱=۲۵۰۰$	$\sigma_{b_۱}=۵۰۰$
فرز	$a_{۲۱}=۴$	$a_{۲۲}=۱۰$	$\bar{b}_۲=۲۰۰۰$	$\sigma_{b_۲}=۴۰۰$
مته	$a_{۳۱}=۱$	$a_{۳۲}=۱/۵$	$\bar{b}_۳=۴۵۰$	$\sigma_{b_۳}=۵۰$
سود هر واحد(Rs)	$c_۱=۵۰$	$c_۲=۱۰۰$		



شکل ۳.۴: منحنی فاز شبکه عصبی(۵۶.۴)-(۵۷.۴) با ۱۴ نقطه آغازین در مثال ۱.۵.۴.



شکل ۴.۴: رفتار همگرایی از $\|x(t) - x^*\|^2$ با $y_0 = (40, -40, 40, -40, 40, -40, 40)^T$ در مثال ۱.۵.۴.

مقدار واریتهٔ نرمال استاندارد (E_i) برای $(i = ۱, ۲, ۳)$ که در آن $\phi(E_i) = ۱ - p_i = ۰/۰۱$ است را نمی‌توان مستقیما از جدول توزیع نرمال استاندارد به‌دست آورد. لاکن توجه داریم که چون $۱ - p_i < ۰/۵$ در نتیجه $E_i < ۰$ و بنابراین

$$\phi(-E_i) = ۱ - \phi(E_i) = p_i = ۰/۹۹ \Longrightarrow E_i = -۲/۳۳.$$

بنابراین از (۲۶.۴)-(۲۵.۴) داریم:

$$\text{Maximize} \quad f = ۵۰x_۱ + ۱۰۰x_۲ \tag{۷۱.۴}$$

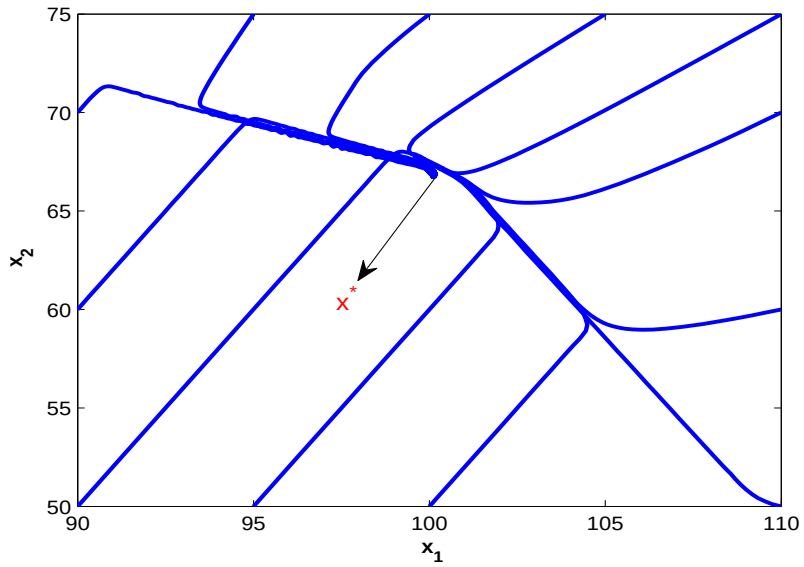
$$\text{subject to} \quad \begin{cases} ۱۰x_۱ + ۵x_۲ - (-۲/۳۳)(۵۰۰) - ۲۵۰۰ \leq ۰, \\ ۴x_۱ + ۱۰x_۲ - (-۲/۳۳)(۴۰۰) - ۲۰۰۰ \leq ۰, \\ x_۱ + ۱/۵x_۲ - (-۲/۳۳)(۵۰) - ۴۵۰ \leq ۰, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۷۲.۴}$$

جواب بهینه این مسأله $x^* = (۱۰۰/۱۲۵۰, ۶۶/۷۵۰۰)^T$ است. ما شبکه عصبی پیشنهاد شده در (۵۶.۴)-(۵۷.۴) را برای حل این مسأله بکار می‌بریم، شکل(۵.۴) منحنی فاز از (۵۶.۴)-(۵۷.۴) با ۱۲ نقطه اولیه مختلف با $\epsilon = ۱۰^{-۶}$ نشان می‌دهد. واضح است حتی وقتی نقطه اولیه یک نقطه نشدنی باشد، مسیر از شبکه پیشنهادی هنوز به x^* همگراست. نرم خطای $L_۲$ بین x و x^* با نقاط اولیه گوناگون هم در شکل (۶.۴) رسم شده است.

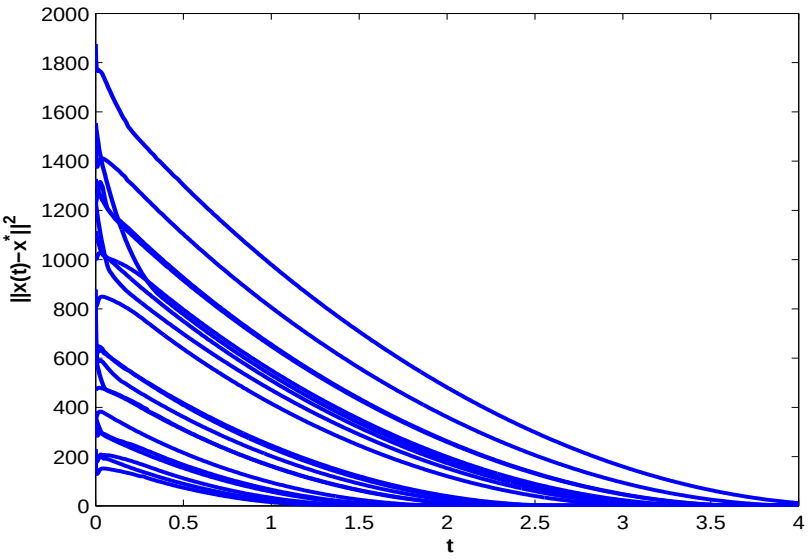
مثال ۳.۵.۴. فرض کنید که سود هر یک از قطعات I و II در مثال (۱.۵.۴) متغیری تصادفی با توزیع نرمال باشد. تعداد قطعاتی که باید در هر هفته تولید کرد تا سود بیشینه شود را بدست آورید. میانگین و انحراف سود به ترتیب برای هر قطعه I ، $۵۰, ۲۰$ و برای هر قطعه II ، $۱۰۰, ۵۰$ روپیه است.

حل: با این فرض که سودها متغیرهای تصادفی مستقل هستند، مسأله قطعی معادل را با استفاده از (۳۱.۴)-(۳۲.۴) می‌توان به‌صورت زیر بیان کرد:

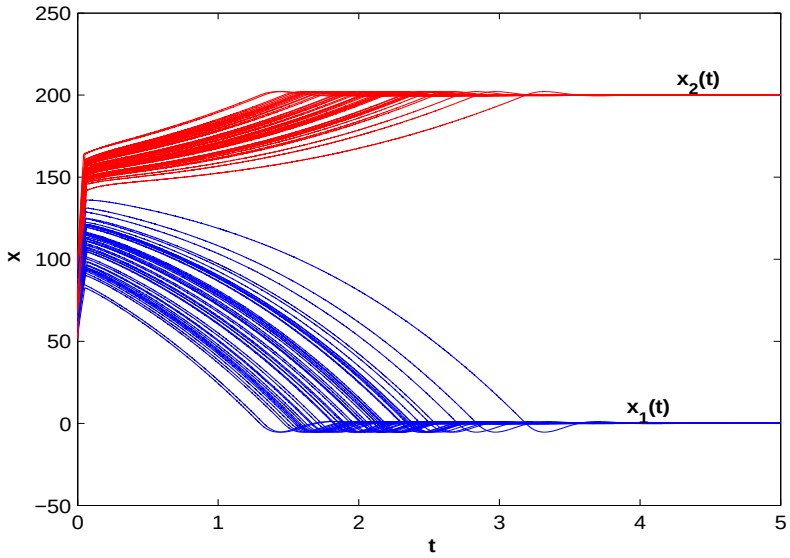
$$\text{Maximize} \quad F = k_۱(۵۰x_۱ + ۱۰۰x_۲) + k_۲\sqrt{۴۰۰x_۱^۲ + ۲۵۰۰x_۲^۲} \tag{۷۳.۴}$$



شکل ۵.۴: منحنی فاز شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۱۲ نقطه آغازین در مثال۲.۵.۴.



شکل ۶.۴: رفتار همگرایی از $\|x(t) - x^*\|^2$ با ۲۰ نقطه اولیه دلخواه در مثال ۲.۵.۴.



شکل ۷.۴: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۵۰ نقطه آغازین دلخواه در مثال ۳.۵.۴.

$$\text{subject to} \quad \begin{cases} ۱۰x_۱ + ۵x_۲ - ۲۵۰۰ \leq ۰, \\ ۴x_۱ + ۱۰x_۲ - ۲۰۰۰ \leq ۰, \\ x_۱ + ۱/۵x_۲ - ۴۵۰ \leq ۰, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۷۴.۴}$$

وقتیکه $k_۱$ ، $k_۲$ مقادیر ثابتی هستند که اهمیت نسبی میانگین و انحراف معیار سود را برای کمینه سازی نشان می‌دهند. جواب بهینه مسأله با $k_۱ = k_۲ = ۱$ و $\epsilon = ۱۰^{-۶}$ است. قضیه (۵.۴.۴) $x^* = (۰, ۲۰۰)^T$ و گزاره (۶.۴.۴) تضمین می‌کند که مدل بیان شده در (۵۶.۴)-(۵۷.۴) ، همگرای سراسری به جواب بهینه یکتاست. شکل (۷.۴) نشان می‌دهد که مسیرهای شبکه پیشنهاد شده برای حل مسأله بالا با ۵۰ نقطه اولیه تصادفی همگرا به جواب بهینه مسأله است.

مثال ۴.۵.۴. اگر زمانهای ماشینکاری مورد نیاز، حداکثر زمان موجود و سودهای هر واحد همه متغیرهای تصادفی با توزیع نرمال در داده های جدول (۳.۴) فرض کنیم، تعداد قطعاتی را که باید ساخته شوند تا سود بیشینه شود را پیدا کنید. قیود باید با احتمال دست کم ۹۹٪ برقرار باشند.

جدول ۳.۴: مقادیر مثال ۴.۵.۴

	زمان ماشینکاری مورد نیاز هر قطعه(دقیقه)				
نوع دستگاه	قطعه I		قطعه II		حداکثر زمان موجود در هفته (دقیقه)
	میانگین	انحراف معیار	میانگین	انحراف معیار	
تراش	$\bar{a}_{۱۱} = ۱۰$	$\sigma_{a_{۱۱}} = ۶$	$\bar{a}_{۱۲} = ۵$	$\sigma_{a_{۱۲}} = ۴$	$b_۱ = ۲۵۰۰$ $\sigma_{b_۱} = ۵۰۰$
فرز	$\bar{a}_{۲۱} = ۴$	$\sigma_{a_{۲۱}} = ۴$	$\bar{a}_{۲۲} = ۱۰$	$\sigma_{a_{۲۲}} = ۷$	$b_۲ = ۲۰۰۰$ $\sigma_{b_۲} = ۴۰۰$
مته	$\bar{a}_{۳۱} = ۱$	$\sigma_{a_{۳۱}} = ۲$	$\bar{a}_{۳۲} = ۱/۵$	$\sigma_{a_{۳۲}} = ۳$	$b_۳ = ۴۵۰$ $\sigma_{b_۳} = ۵۰$
سود هر واحد(R_s)	$c_۱ = ۵۰$ $\sigma_{c_۱} = ۲۰$		$c_۲ = ۱۰۰$ $\sigma_{c_۲} = ۵۰$		

حل: با تعریف متغیرهای تصادفی جدیدh_i به‌صورت زیر

$$h_i = \sum_{j=۱}^{n} a_{ij}x_j - b_i,$$

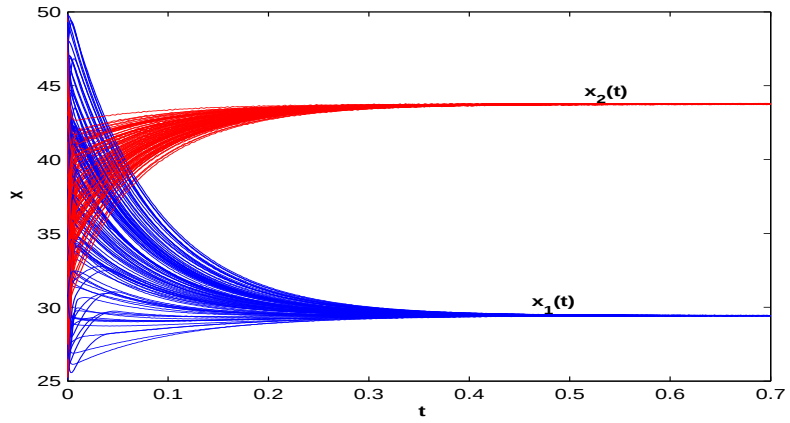
در می‌یابیم که h_i هم دارای توزیع نرمال است. با فرض اینکه بین a_{ij} و b_i ها همبستگی وجود ندارد، میانگین ها و واریانس های h_i را می‌توان از روابط (۳۶.۴) و (۳۷.۴) بدست آورد. تابع هدف را همان F در مثال قبل می‌گیریم و قیود هم به‌صورت زیر بیان می‌شوند:

$$P[h_i \leq ۰] \geq ۰/۹۹,\ i = ۱, ۲, ۳,$$

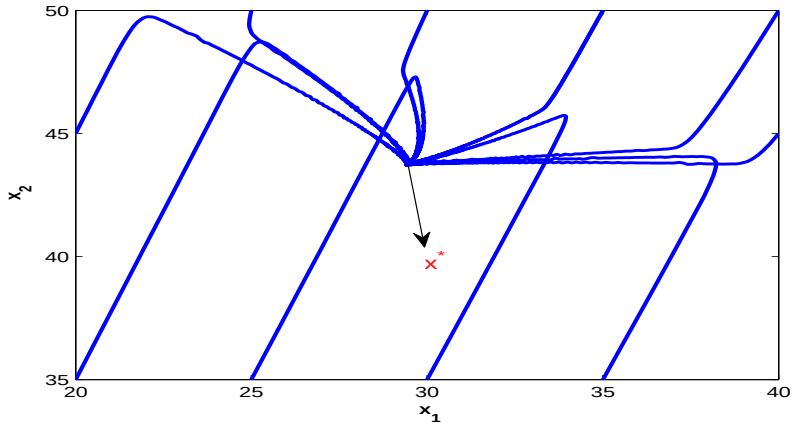
چون مقدار واریتهٔ نرمال استاندارد e_i مربوط به احتمال ۰/۹۹% برابر۲/۳۳ است، می‌توانیم مسأله بهینه‌سازی غیرخطی معادل را به‌صورت زیر بیان کنیم:

$$\begin{aligned}
&\text{Maximize } F = k_۱(۵۰x_۱ + ۱۰۰x_۲) + k_۲\sqrt{۴۰۰x_۱^۲ + ۲۵۰۰x_۲^۲} \\
&\qquad \text{subject to} \\
&\begin{cases}
۱۰x_۱ + ۵x_۲ + ۲/۳۳\sqrt{۳۶x_۱^۲ + ۱۶x_۲^۲ + ۲۵۰۰۰۰} - ۲۵۰۰ \leq ۰, \\
۴x_۱ + ۱۰x_۲ + ۲/۳۳\sqrt{۱۶x_۱^۲ + ۴۹x_۲^۲ + ۱۶۰۰۰۰} - ۲۰۰۰ \leq ۰, \\
x_۱ + ۱/۵x_۲ + ۲/۳۳\sqrt{۴x_۱^۲ + ۹x_۲^۲ + ۲۵۰۰} - ۴۵۰ \leq ۰, \\
x_۱, x_۲ \geq ۰.
\end{cases}
\end{aligned} \tag{۷۵.۴}$$

جواب بهینه این مسأله برای $k_۱ = k_۲ = ۱$ ، $x^* = (۲۹/۴۳۵۰۲, ۴۳/۷۶۳۸۶)^T$ است. شکل (۸.۴) نشان می‌دهد که مسیرهایی از شبکه پیشنهاد شده در (۵۶.۴)-(۵۷.۴) برای حل مسأله بالا با ۱۰۰ نقطه



شکل ۸.۴: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۱۰۰ نقطه آغازین دلخواه در مثال ۴.۵.۴.



شکل ۹.۴: منحنی فاز شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۱۰ نقطه آغازین در مثال ۴.۵.۴.

آغازین دلخواه همگرا به جواب بهینه مسأله است. منحنی فاز این شبکه عصبی نیز، در شکل (۹.۴) رسم شده است.

فصل ۵

مسائل برنامه‌ریزی کسری

۱.۵ مقدمه

در این فصل ابتدا برنامه‌ریزی کسری [۱] را معرفی می‌کنیم مروری بر کاربردهای آنها داریم سپس به معرفی رده خاصی از برنامه‌ریزی کسری که می‌توان آنها را به برنامه‌ریزی مخروط مرتبه دوم تبدیل کرد، می‌پردازیم. در انتها چندین مثال عددی از مسائل کسری را با شبکه عصبی (۵۶.۴)-(۵۷.۴) که برای حل مسائل برنامه‌ریزی مخروط مرتبه دوم پیشنهاد کردیم، حل می‌کنیم.

۲.۵ برنامه‌ریزی کسری

برنامه‌ریزی کسری یک موضوع جالب در بسیاری از مسائل بهینه سازی از جمله انتخاب سبدسرمایه، تولید، تئوری اطلاعات [۲] و تصمیم‌گیری‌های بسیار پیچیده در علم مدیریت است، به‌طور خاص در مهندسی و اقتصاد برای حداقل سازی توابع اقتصادی یا فیزیکی یا هردو از جمله هزینه/زمان، هزینه/حجم، هزینه/سود و غیره به‌منظور اندازه کارایی و بهره‌وری سیستم کاربرد دارد. بسیاری کاربردهای دیگر، مسائل برنامه ریزی کسری توسط کراون [۳] در [۱۲۹] ، ایباراکی [۴] و همکارش در [۱۳۰] ارائه شده است. برای بررسی گسترده‌تر برنامه‌ریزی کسری می‌تواند مراجع [۱۳۱]-[۱۳۵] را ببینید.

در مهندسی و بسیاری از کاربردهای علمی، جوابهای زمان واقعی مسائل برنامه‌ریزی کسری مطلوب است. در نتیجه به روشهای مؤثر برای حل برنامه‌ریزی کسری زمان واقعی نیاز داریم. نتایج مطالعات قبلی حاکی از آن است که از شبکه‌های عصبی می‌توان برای حل مسائل مختلف بهینه‌سازی استفاده کرد. ایده اصلی شبکه‌های عصبی برای بهینه‌سازی ایجاد یک تابع انرژی نامنفی و پایه‌گذاری یک سیستم دینامیکی است که یک شبکه عصبی مصنوعی را ایجاد می‌کند. سیستم دینامیکی معمولا به فرم معادلات دیفرانسیل مرتبه اول است. به‌علاوه انتظار می‌رود که سیستم دینامیکی به یک وضعیت پایدار (یا نقطهٔ تعادل) که

[۱] Fractional Programming
[۲] Information theory
[۳] Craven
[۴] Ibaraki

متناظر با جواب مسأله بهینه‌سازی اصلی است که با یک نقطه اولیه آغاز می‌شود نزدیک گردد.
با توجه به مباحث فوق، در این فصل یک مدل از شبکه‌های عصبی برای حل مسائل برنامه‌ریزی کسری ارائه می‌دهیم.

ابتدا رده‌ای از مسائل برنامه‌ریزی کسری را تبدیل به مسائل برنامه ریزی مخروط مرتبه دوم محدب کرده، سپس این مسائل را تبدیل به برنامه‌ریزی غیرخطی محدب با قیود نامساوی می‌کنیم، زمانیکه با استفاده از تکنیک هموار سازی، قیود مخروط مرتبه دوم ناهموار به‌فرم هموار آن تبدیل شوند. براساس شرایط بهینگی $(K.K.T)$ برنامه‌ریزی محدب ، مدل شبکه عصبی برای حل مسائل برنامه‌ریزی غیرخطی محدب پیشنهاد خواهد شد. همچنین ثابت می‌شود نقطهٔ تعادل شبکه عصبی پیشنهاد شده برابر نقطه $K.K.T$ مسأله برنامه ریزی مخروط مرتبه دوم محدب است. وجود و یکتایی نقطه تعادل شبکه عصبی نیز مورد آنالیز قرار خواهد گرفت. با ساختن یک تابع لیاپانوف مناسب، شرط کافی برای تضمین وجود و پایدار مجانبی سراسری برای نقطه تعادل یکتای شبکه پیشنهاد شده بدست خواهد آمد.
در این بخش رده‌های خاصی از مسائل برنامه‌ریزی کسری را مورد بررسی قرار می‌دهیم که بتوان آن‌ها را به مسائلی از نوع برنامه‌ریزی مخروط مرتبه دوم تبدیل کرد. برای تبدیل برنامه‌ریزی کسری به برنامه‌ریزی مخروط مرتبه دوم به نامساوی زیر نیاز داریم:([۱۲۵])

اگر $z, y \geq ۰$ و w یک عدد حقیقی دلخواه باشد، آن‌گاه داریم:

$$w^۲ \leq zy \iff \left\| \begin{bmatrix} ۲w \\ z-y \end{bmatrix} \right\| \leq z+y, \tag{۱.۵}$$

و در صورت کلی، وقتی w یک بردار حقیقی است:

$$w^T w \leq zy \iff \left\| \begin{bmatrix} ۲w \\ z-y \end{bmatrix} \right\| \leq z+y. \tag{۲.۵}$$

۱.۲.۵ برنامه‌ریزی کسری با صورت و مخرج خطی

اولین حالتی را که در برنامه‌ریزی کسری مورد بحث قرار می‌دهیم به‌صورت زیر بیان می‌شود([۱۲۸])

$$\text{Minimize} \quad \sum_{i=۱}^{p} \frac{c}{a_i^T x + b_i} \tag{۳.۵}$$

subject to

$$\begin{cases} a_i^T x + b_i > ۰, & i = ۱, \ldots, p, \\ c_j^T x + d_j \geq ۰, & j = ۱, \ldots, q. \end{cases} \tag{۴.۵}$$

برای تبدیل مسأله بالا به برنامه‌ریزی مخروط مرتبه دوم ابتدا، متغیرهای جدید u_i را معرفی می‌کنیم

بنابراین مسأله زیر را داریم:

$$\text{Minimize} \quad \sum_{i=۱}^{p} u_i \tag{۵.۵}$$

subject to

$$\begin{cases} \frac{c}{a_i^T x + b_i} \leq u_i, & i = ۱, \ldots, p, \\ u_i > ۰, & i = ۱, \ldots, p, \\ a_i^T x + b_i > ۰, & i = ۱, \ldots, p, \\ c_j^T x + d_j \geq ۰, & j = ۱, \ldots, q, \end{cases} \tag{۶.۵}$$

با فرض اینکه

$$z_i = a_i^T x + b_i, \ y_i = u_i, \ w_i = \sqrt{c}, \quad i = ۱, \ldots, p, \tag{۷.۵}$$

و با به‌کار بردن نامساوی (۱.۵) داریم:

$$c \leq (a_i^T x + b_i) u_i \iff \left\| \begin{bmatrix} ۲\sqrt{c} \\ a_i^T x + b_i - u_i \end{bmatrix} \right\| \leq a_i^T x + b_i + u_i, i = ۱, \ldots, p.$$

بنابراین جواب مسأله برنامه‌ریزی کسری (۳.۵)-(۴.۵) از حل مسأله معادل برنامه‌ریزی مخروط مرتبه دوم زیر به‌دست می‌آید:

$$\text{Minimize} \quad \sum_{i=۱}^{p} u_i \tag{۸.۵}$$

subject to

$$\begin{cases} \left\| \begin{bmatrix} ۲\sqrt{c} \\ a_i^T x + b_i - u_i \end{bmatrix} \right\| \leq a_i^T x + b_i + u_i, & i = ۱, \ldots, p, \\ c_j^T x + d_j \geq ۰, & j = ۱, \ldots, q. \end{cases} \tag{۹.۵}$$

۲.۲.۵ برنامه‌ریزی کسری با تابع هدف درجه دوم

مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \sum_{i=۱}^{p} \frac{c}{a_i^T x + b_i} + x^T p_۰ x + ۲q_۰{}^T x + r \tag{۱۰.۵}$$

$$\text{subject to}$$

$$\begin{cases} a_i^T x + b_i > ۰, & i = ۱, \dots, p, \\ c_j^T x + d_j \geq ۰, & j = ۱, \dots, q, \end{cases} \tag{۱۱.۵}$$

که درآن $q_۰ \in \mathbb{R}^n$ و $p_۰ \in \mathbb{R}^{n \times n}$ یک ماتریس معین مثبت است. چون $p_۰$ معین مثبت است، بنابراین $x^T p_۰ x + ۲q_۰{}^T x + r$ قابل تبدیل به فرم زیر است([۱۲۵])

$$\| p_۰{}^{\frac{۱}{۲}} x + p_۰{}^{\frac{-۱}{۲}} q_۰ \|^۲ + (r - q_۰{}^T p_۰{}^{-۱} q_۰). \tag{۱۲.۵}$$

اکنون با معرفی متغیرهای $t_۰$ و u_i برای $(i = ۱, ..., p)$ به‌طوری‌که

$$\frac{c}{a_i^T x + b_i} \leq u_i, \qquad i = ۱, ..., p, \tag{۱۳.۵}$$

$$x^T p_۰ x + ۲q_۰{}^T x + r \leq t_۰, \tag{۱۴.۵}$$

و با بکارگیری نامساوی (۲.۵) و (۱۲.۵)-(۱۴.۵) ، مسأله برنامه‌ریزی کسری (۱۰.۵)-(۱۱.۵) به فرم معادل برنامه‌ریزی مخروط مرتبه دوم زیر تبدیل می‌شود:

$$\text{Minimize} \quad \sum_{i=۱}^{p} u_i + t_۰ \tag{۱۵.۵}$$

$$\text{subject to}$$

$$\begin{cases} \left\| \begin{bmatrix} ۲\sqrt{c} \\ a_i^T x + b_i - u_i \end{bmatrix} \right\| \leq a_i^T x + b_i + u_i, & i = ۱, \dots, p, \\ \left\| \begin{bmatrix} ۲(p_۰{}^{\frac{۱}{۲}} x + p_۰{}^{\frac{-۱}{۲}} q_۰) \\ ۱ - t_۰ \end{bmatrix} \right\| \leq t_۰ + ۱, & \\ c_j^T x + d_j \geq ۰, & j = ۱, \dots, q. \end{cases} \tag{۱۶.۵}$$

۳.۲.۵ برنامه‌ریزی کسری محدب با صورت درجه دوم

مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \sum_{i=۱}^{p} \frac{\|F_i x + g_i\|^۲}{a_i^T x + b_i} \tag{۱۷.۵}$$

$$\text{subject to}$$

$$\begin{cases} a_i^T x + b_i > ۰, & i = ۱, \dots, p, \\ c_j^T x + d_j \geq ۰, & j = ۱, \dots, q, \end{cases} \tag{۱۸.۵}$$

که درآن $F_i \in \mathbb{R}^{q_i \times n} g_i \in \mathbb{R}^{q_i} (i = ۱, ...p)$ هستند. اکنون با معرفی متغیرهای جدید u_i به‌طوری‌که

$$\frac{\|F_i x + g_i\|^۲}{a_i^T x + b_i} \leq u_i, \quad i = ۱, \dots, p. \tag{۱۹.۵}$$

و با بکار بردن نامساوی (۲.۵) ، مسأله (۱۷.۵)-(۱۸.۵) به فرم برنامه‌ریزی مخروط مرتبه دوم زیر تبدیل می‌شود:

$$\text{Minimize} \quad \sum_{i=۱}^{p} u_i \tag{۲۰.۵}$$

$$\text{subject to}$$

$$\begin{cases} \left\| \begin{bmatrix} ۲\,(F_i x + g_i) \\ a_i^T x + b_i - u_i \end{bmatrix} \right\| \leq a_i^T x + b_i + u_i, & i = ۱, \dots, p, \\ c_j^T x + d_j \geq ۰, & j = ۱, \dots, q. \end{cases} \tag{۲۱.۵}$$

اکنون حالت خاصی از مسأله (۱۷.۵)-(۱۸.۵) با قیود درجه دوم به‌صورت زیر در نظر می‌گیریم:

$$\text{Minimize} \quad \sum_{i=۱}^{p} \frac{x^T P_i x + ۲ q_i^T x + r_i}{a_i^T x + b_i} \tag{۲۲.۵}$$

$$\text{subject to}$$

$$\begin{cases} a_i^T x + b_i > ۰, & i = ۱, \dots, p, \\ c_j^T x + d_j \geq ۰, & i = j, \dots, q, \\ x^T Q_k x + ۲ m_k^T x + l_k \leq ۰, & k = ۱, \dots, s, \end{cases} \tag{۲۳.۵}$$

که در آن $m_k \in \mathbb{R}^n$ و ماتریس‌های $P_i \in \mathbb{R}^{n\times n}$ $(i = ۱,\dots,p)$ ، $Q_k \in \mathbb{R}^{n\times n}$ $(k = ۱,\dots,s)$ ماتریس‌های معین مثبت هستند. فرم مخروط مرتبه دوم متناظر با توابع درجه دوم مسأله بنابراین (۲۲.۵)-(۲۳.۵) به‌صورت زیر بیان می‌شوند:([۱۲۵])

$$\left\|P_i^{\frac{۱}{۲}}x + P_i^{\frac{-۱}{۲}}q_i\right\|^۲ + r_i - q_i^T P_i q_i = x^T P_i x + ۲q_i{}^T x + r_i,$$
$$\left\|Q_k^{\frac{۱}{۲}}x + Q_k^{\frac{-۱}{۲}}m_k\right\| \le \left(m_k^T Q_k^{-۱} m_k - l_k\right)^{\frac{۱}{۲}} \equiv x^T Q_k x + ۲m_k^T x + l_k \le ۰.$$

بنابراین تابع هدف (۲۲.۵) را می‌توان به‌صورت زیر بیان کرد:

$$\frac{x^T P_i x + ۲q_i^T x + r_i}{a_i^T x + b_i} = \frac{\left\|P_i^{\frac{۱}{۲}}x + P_i^{\frac{-۱}{۲}}q_i\right\|^۲}{a_i^T x + b_i} + \frac{r_i - q_i^T P_i q_i}{a_i^T x + b_i}, \quad i = ۱,\dots,p.$$

با معرفی متغیرهای جدید u_i و v_i ، مشابه حالت قبلی، مسأله برنامه ریزی کسری (۲۲.۵)-(۲۳.۵) به فرم معادل زیر تبدیل می‌شود:

$$\text{Minimize} \quad \sum_{i=۱}^{p}(u_i + v_i) \tag{۲۴.۵}$$
$$\text{subject to}$$
$$\begin{cases} \left\|\begin{bmatrix} ۲\left(P_i^{\frac{۱}{۲}}x + P_i^{\frac{-۱}{۲}}q_i\right) \\ a_i^T x + b_i - u_i \end{bmatrix}\right\| \le a_i^T x + b_i + u_i, & i = ۱,\dots,p, \\ \left\|\begin{bmatrix} ۲\left(r_i - q_i^T P_i^{-۱} q_i\right)^{\frac{۱}{۲}} \\ a_i^T x + b_i - v_i \end{bmatrix}\right\| \le a_i^T x + b_i + v_i, & i = ۱,\dots,p, \\ \left\|Q_k^{\frac{۱}{۲}}x + Q_k^{\frac{-۱}{۲}}m_k\right\| \le \left(m_k^T Q_k^{-۱} m_k - l_k\right)^{\frac{۱}{۲}}, & k = ۱,\dots,s, \\ c_j^T x + d_j \ge ۰, & j = ۱,\dots,q. \end{cases} \tag{۲۵.۵}$$

۴.۲.۵ برنامه‌ریزی با تابع هدف مختلط از کسری و درجه دوم وهمراه با قیود درجه دوم

مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \sum_{i=۱}^{p}\frac{\|F_i x + g_i\|^۲}{a_i^T x + b_i} + \sum_{j=۱}^{l}\frac{c}{c_j^T x + d_j} + x^T P_۰ x + ۲q_۰^T x + r \tag{۲۶.۵}$$

subject to

$$\begin{cases} a_i^T x + b_i > ۰, & i = ۱, \dots, p, \\ c_j^T x + d_j > ۰, & j = ۱, \dots, q, \\ x^T P_k x + ۲ q_k^T x + r_k \leq ۰, & k = ۱, \dots, s, \end{cases} \tag{۲۷.۵}$$

که درآن ماتریس‌های $P_k (k = ۱, \dots, s)$ ، $p_۰$ معین مثبت هستند. با معرفی متغیرهای جدید $u_i (i =$ $۱, ..., p)$، $v_j (j = ۱, \dots, q)$ ، $t_۰$ و با استفاده ازروابط به‌کاررفته در حالت‌های قبل، مسأله (۲۶.۵)-(۲۷.۵) قابل بیان به‌صورت معادل زیر است:

$$\text{Minimize} \qquad \sum_{i=۱}^{p} u_i + \sum_{j=۱}^{q} v_j + t_۰ \tag{۲۸.۵}$$

subject to

$$\begin{cases} \left\| \begin{bmatrix} ۲ (F_i x + g_i) \\ a_i^T x + b_i - u_i \end{bmatrix} \right\| \leq a_i^T x + b_i + u_i, \quad i = ۱, \dots, p, \\ \left\| \begin{bmatrix} ۲(p_۰^{\frac{۱}{۲}} x + p_۰^{\frac{-۱}{۲}} q_۰) \\ ۱ - t_۰ \end{bmatrix} \right\| \leq t_۰ + ۱, \\ \left\| P_k^{\frac{۱}{۲}} x + P_k^{\frac{-۱}{۲}} q_k \right\| \leq \left(q_k^T P_k^{-۱} q_k - r_k \right)^{\frac{۱}{۲}}, \quad k = ۱, \dots, s, \\ \left\| \begin{bmatrix} ۲\sqrt{c} \\ c_j^T x + d_j - v_j \end{bmatrix} \right\| \leq c_j^T x + d_j + v_j, \quad j = ۱, \dots, l. \end{cases} \tag{۲۹.۵}$$

از تحلیل بالا و [۱۲۵، ۱۲۸]، واضح است که مسائل برنامه ریزی کسری قابل تبدیل به مسائل برنامه ریزی مخروط مرتبه دو محدب هستند. بنابراین ، ما فرم کلی از مسائل مسائل برنامه ریزی مخروط مرتبه دو محدب را به‌صورت زیر در نظر می‌گیریم:

$$\text{Minimize} \qquad f(x) \tag{۳۰.۵}$$

subject to

$$Ex + d \succeq_{\mathcal{K}} ۰, \tag{۳۱.۵}$$

$$x \geq ۰. \tag{۳۲.۵}$$

وقتیکه $f : \mathbb{R}^n \to \mathbb{R} \bigcup \{+\infty\}$، تابع سره بستهٔ محدب ، E یک ماتریس $m \times n$ با $m \geq n$ ، d برداری در $\mathbb{R}^m$ و $x \succeq_{\mathcal{K}} ۰$ یعنی $x \in \mathcal{K}$ و $\mathcal{K}$ حاصلضرب دکارتی از مخروط‌های مرتبه دوم که مخروط‌های لورنتس یا مخروط بستنی هم نامیده می‌شوند . به عبارت دیگر،

$$\mathcal{K} = \mathcal{K}^{m_۱} \times \mathcal{K}^{m_۲} \times ... \times \mathcal{K}^{m_r},$$

که درآن $r, m_۱, ..., m_r \geq ۱$ با $m_۱ + ... + m_r = n$ و

$$\mathcal{K}^{m_i} := \{(x_۱, x_۲)^T \in \mathbb{R} \times \mathbb{R}^{m_i - ۱} \mid x_۱ \geq \|x_۲\|\}, \qquad (۳۳.۵)$$

که درآن $\|.\|$ نرم اقلیدسی و $\mathcal{K}^۱$ مجموعه از $\mathbb{R}_+$ حقیقی نامنفی است. در فصل قبل مدل شبکه عصبی (۵۶.۴)-(۵۷.۴) با عملکرد بالا برای حل مسأله (۳۰.۵)-(۳۲.۵) پیشنهاد کردیم.

۳.۵ مثال‌های عددی

در این بخش برای روشن شدن مطالب شرح داده شده و به منظور نشان دادن کارایی و تأثیر شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) ، ما به ذکر چندین مثال از مسائل برنامه‌ریزی کسری می‌پردازیم. برای برخی مثال‌ها، ما همچنین مقایسه عملکرد عددی از شبکه عصبی پیشنهادی را با مقادیر گوناگون از حالت اولیه $y(۰)$ داریم . برای برخی دیگر از مسائل ما عملکرد عددی از شبکه عصبی پیشنهاد شده را با مقادیر گوناگون از نرخ همگرایی κ مقایسه می‌کنیم. در شبیه سازی همه مثال‌ها از دستور $ode۴۵s$ در نرم افزار $Matlab$ استفاده شده و جواب واقعی مسائل نیز توسط نرم افزار $Lingo۱۱$ محاسبه شده است..

مثال ۱.۳.۵. مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \frac{۱}{x_۱ + x_۲ + ۲} + \frac{۱}{۲x_۱ - ۳x_۲ - ۲} \qquad (۳۴.۵)$$

$$\text{subject to}$$

$$\begin{cases} x_۱ + x_۲ \leq ۴, \\ x_۱ + x_۲ + ۲ > ۰, \\ ۲x_۱ - ۳x_۲ - ۲ > ۰, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۳۵.۵}$$

در این مثال با معرفی متغیرهای جدید $u_۱, u_۲$ و بکار بردن نامساوی (۱.۵) مسأله برنامه ریزی مخروط مرتبه دوم محدب زیر به‌دست می‌آید:

$$\text{Minimize} \qquad u_۱ + u_۲ \tag{۳۶.۵}$$

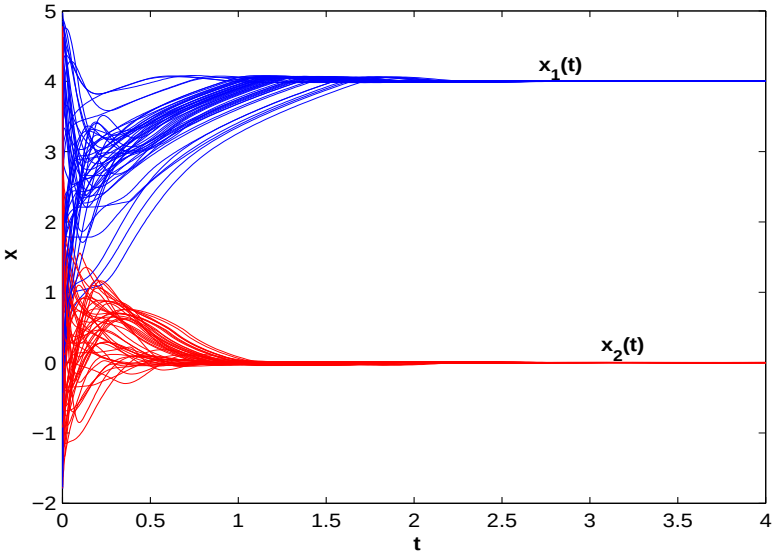
$$\text{subject to} \quad \begin{cases} \left\| \begin{bmatrix} ۲ \\ x_۱ + x_۲ + ۲ - u_۱ \end{bmatrix} \right\| \leq x_۱ + x_۲ + ۲ + u_۱, \\ \left\| \begin{bmatrix} ۲ \\ ۲x_۱ - ۳x_۲ - ۲ - u_۲ \end{bmatrix} \right\| \leq ۲x_۱ - ۳x_۲ - ۲ + u_۲, \\ x_۱ + x_۲ \leq ۴, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۳۷.۵}$$

جواب بهینه مسأله(۳۵.۵)-(۳۴.۵) $x^* = (۴, ۰)^T$ است . ما برای حل این مسأله شبکه عصبی پیشنهادی (۵۶.۴)-(۵۷.۴) را به‌کار می‌بریم، شکل (۱.۵) نشان می‌دهد که مسیرهای شبکه عصبی پیشنهادی با ۵۰ نقطه آغازین متفاوت و $\epsilon = ۱۰^{-۶}$، همگرا به جواب بهینه مسأله است. همچنین در شکل(۲.۵) تأثیر پارامتر κ از شبکه عصبی روی مقدار $\|x(t) - x^*\|^۲$ با حالت اولیه $y_۰ = (۱, -۱, ۱, -۱, ۱, -۱, ۱, -۱, ۱)^T$ بررسی شده است، در شکل فوق مشاهده می‌کنیم κ بزرگتر، نرخ همگرایی بهتری از خطای $\|x(t) - x^*\|^۲$ به‌دست می‌آورد.

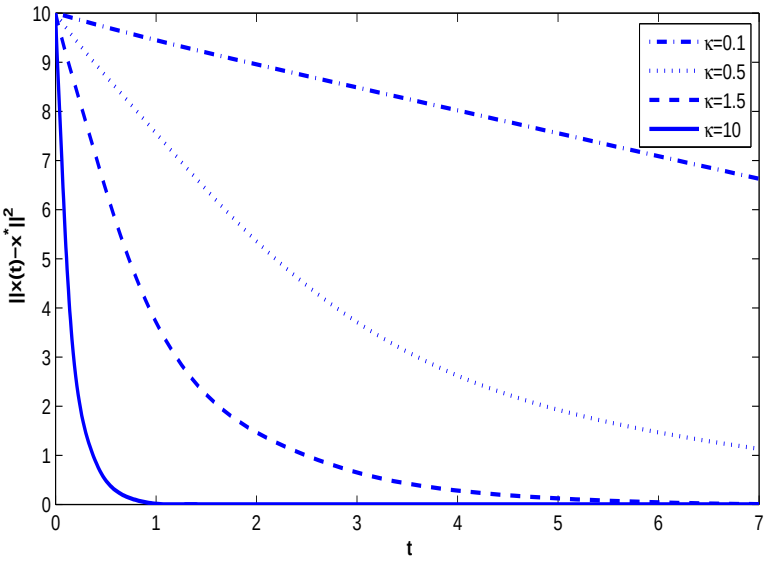
مثال ۲.۳.۵. مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \qquad \frac{۱}{۳x_۱ - ۴x_۲ + ۵} + \frac{۱}{۲x_۱ + ۵x_۲ + ۷} \tag{۳۸.۵}$$

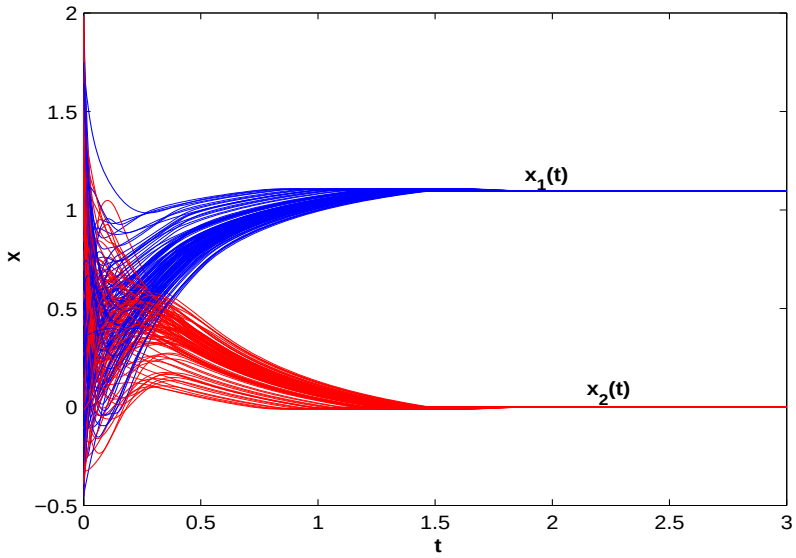
$$\text{subject to}$$



شکل ۱.۵: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۵۰ نقطه آغازین دلخواه در مثال ۱.۳.۵.



شکل ۲.۵: رفتار همگرایی از $\|x(t)-x^*\|^2$ با $y_0 = (40,-40,40,-40,40,-40,40)^T$ در مثال ۱.۳.۵.



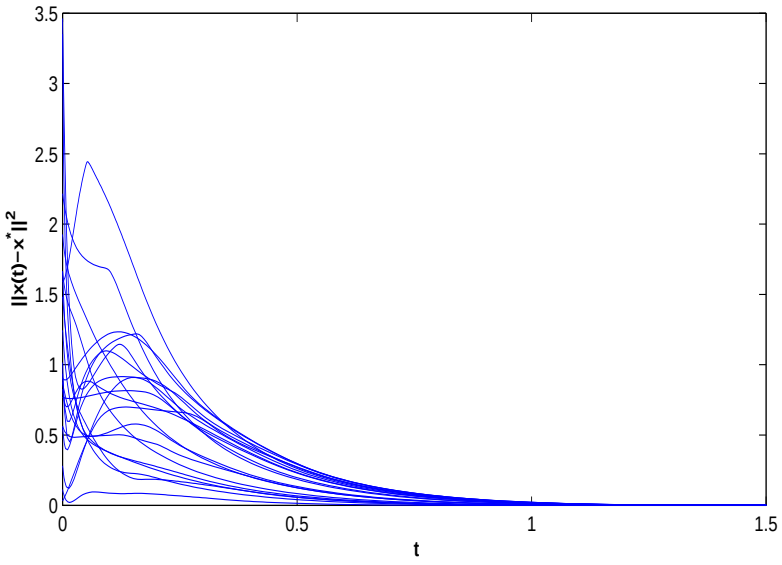
شکل ۳.۵: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۱۰۰ نقطه آغازین دلخواه در مثال ۲.۳.۵.

$$\begin{cases} ۳x_۱ - ۴x_۲ + ۵ > ۰, \\ ۲x_۱ + ۵x_۲ + ۷ > ۰, \\ ۴x_۱^۲ + ۲x_۲^۲ + ۴x_۱x_۲ + ۲x_۱ + ۴x_۲ - ۷ \leq ۰. \end{cases} \tag{۳۹.۵}$$

با تبدیل مسأله به برنامه‌ریزی مخروط مرتبه دوم محدب و حل آن، جواب بهینه مسأله، $x^* = (۱/۰۹۶۲, ۰, ۰/۱۲۰۶, ۰/۱۰۸۸)^T$ به‌دست می‌آید، بنابراین جواب مسأله (۳۸.۵)-(۳۹.۵) $x^* = (۱/۰۹۶۲, ۰)^T$ است. نتایج شبیه سازی نشان می‌دهد که مسیرهایی از (۵۶.۴)-(۵۷.۴) با هر نقطه اولیه همگرا به $y^* = (x^{*T}, u^{*T})^T$ است. برای مثال شکل (۳.۵) نشان دهنده رفتار گذرا شبکه عصبی با ۱۰۰ نقطه آغازین متفاوت و $\epsilon = ۱۰^{-۶}$ است، نرم خطای $L_۲$ بین x و x^* با ۲۰ نقطه اولیه متفاوت در شکل (۴.۵) رسم شده است.

مثال ۳.۳.۵. مسأله برنامه ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \frac{۱}{۵x_۱ + x_۲ + ۱} + \frac{۱}{۷x_۱ - x_۲ + ۴} + x_۱^۲ + ۳x_۲^۲ + ۳x_۱x_۲ + x_۱ - ۴x_۲ - ۸$$



شکل ۴.۵: رفتار همگرایی از $\|x(t)-x^*\|^2$ با ۲۰ نقطه اولیه متفاوت در مثال ۲.۳.۵.

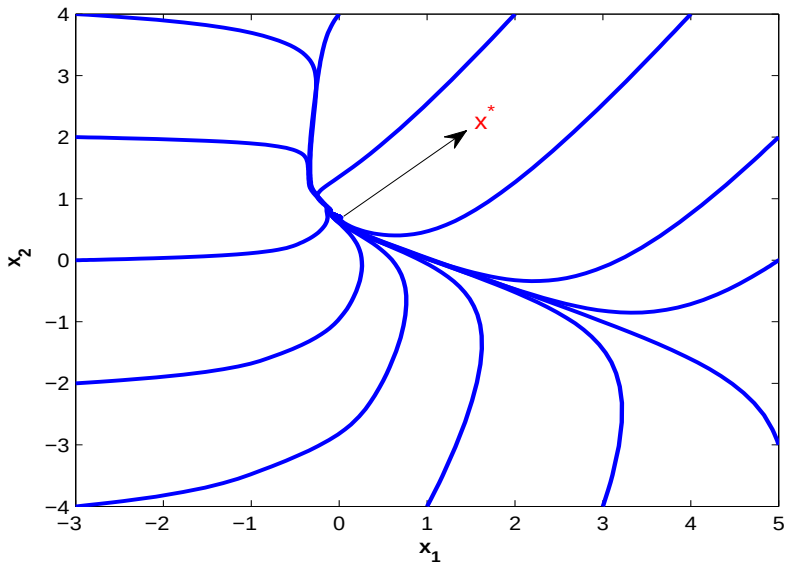
subject to

$$\begin{cases} 5x_1 + x_2 + 1 > 0, \\ 7x_1 - x_2 + 4 > 0, \\ x_1, x_2 \geq 0. \end{cases} \tag{۴۰.۵}$$

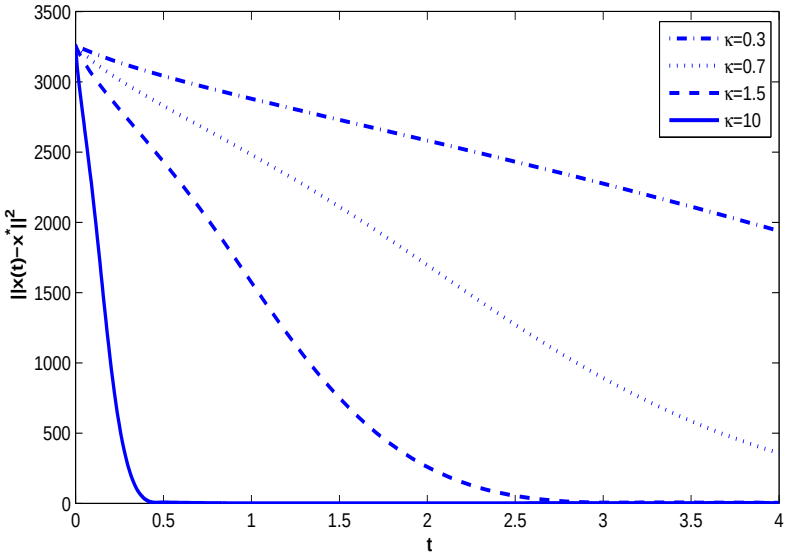
جواب بهینه این مسأله، $x^* = (0, 0/7084)^T$ است. قضیه (۵.۴.۴) و گزاره (۶.۴.۴) تضمین می‌کند که مدل بیان شده در (۵۶.۴)-(۵۷.۴) ، همگرای سراسری به جواب بهینه یکتاست. شکل (۵.۵) منحنی فاز متغیرهای حالت $(x_1(t), x_2(t))^T$ ، با ۱۳ نقطه اولیه متفاوت و $\epsilon = 10^{-6}$ نشان می‌دهد. این بردارها همگرای سراسری به جواب دقیق x^* هستند. شکل (۶.۵) رفتار همگرایی، خطای $\|x(t)-x^*\|^2$ با مقادیر مختلف κ و حالت اولیه $y_0 = (40, -40, 40, -40, 40, -40, 40, -40, 40, -40)^T$ نشان می‌دهد. واضح است که κ بزرگتر منجر به نرخ همگرایی بهتری شده است.

مثال ۴.۳.۵. مسأله برنامه ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \frac{\|3x_1 + x_2 + 1\|^2}{x_1 + x_2 + 2} + \frac{\|x_1 + x_2 + 1\|^2}{2x_1 - 3x_2 + 1} \tag{۴۱.۵}$$



شکل ۵.۵: منحنی فاز شبکه عصبی(۵۶.۴)-(۵۷.۴) با ۱۳ نقطه آغازین در مثال ۳.۳.۵.



شکل ۶.۵: رفتار همگرایی از $\|x(t) - x^*\|^2$ با y_0 اولیه در مثال ۳.۳.۵.

$$\text{subject to} \quad \begin{cases} x_۱ + x_۲ + ۲ > ۰, \\ ۲x_۱ - ۳x_۲ + ۱ > ۰, \\ x_۱ + x_۲ \leq ۴, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۴۲.۵}$$

با معرفی متغیرهای جدید $u_۱$ و $u_۲$ با فرض اینکه

$$\frac{\|۳x_۱ + x_۲ + ۱\|^۲}{x_۱ + x_۲ + ۲} \leq u_۱, \quad \frac{\|x_۱ + x_۲ + ۱\|^۲}{۲x_۱ - ۳x_۲ + ۱} \leq u_۲.$$

بنابراین مسأله بالا تبدیل به معادل برنامه‌ریزی مخروط مرتبه دوم محدب زیر می‌شود:

$$\text{Minimize} \qquad u_۱ + u_۲ \tag{۴۳.۵}$$

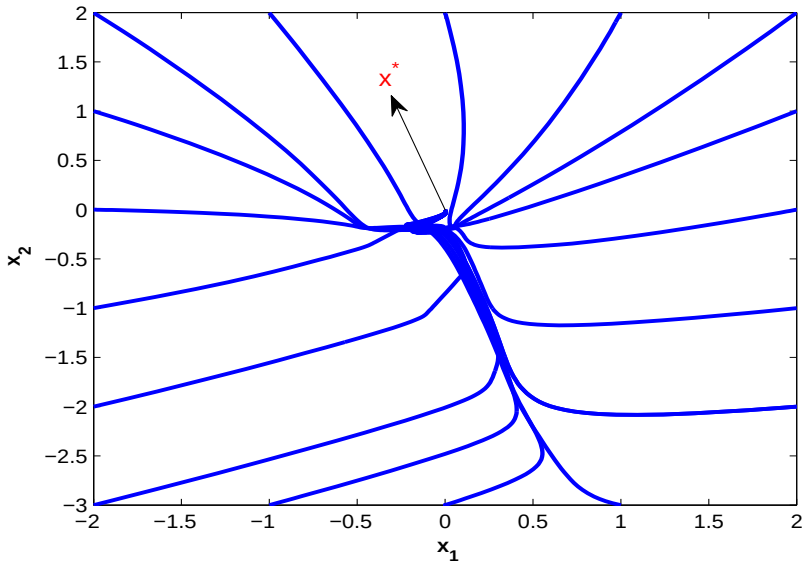
$$\text{subject to} \quad \begin{cases} \left\| \begin{array}{c} ۲\,(۳x_۱ + x_۱ + ۱) \\ x_۱ + x_۲ + ۲ - u_۱ \end{array} \right\| \leq x_۱ + x_۲ + ۲ + u_۱, \\ \left\| \begin{array}{c} ۲\,(x_۱ + x_۲ + ۱) \\ ۲x_۱ - ۳x_۲ + ۱ - u_۱ \end{array} \right\| \leq ۲x_۱ - ۳x_۲ + ۱ + u_۲, \\ x_۱ + x_۲ \leq ۴, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۴۴.۵}$$

جواب بهینه این مسأله $x^* = (۰, ۰)^T$ است. شکل (۷.۵) بیان کننده منحنی فاز متغیرهای حالت $(x_۱(t), x_۲(t))^T$ ، با ۱۷ حالت اولیه متفاوت و $\epsilon = ۱۰^{-۶}$ است. شکل (۸.۵) رفتار همگرا از خطای $\|x(t) - x^*\|^۲$ با κ های متفاوت و نقطه آغازین $y_۰ = (۰, -۱, -۲, -۳, -۴, -۵, -۶, -۷, -۸)^T$ را نشان می‌دهد.

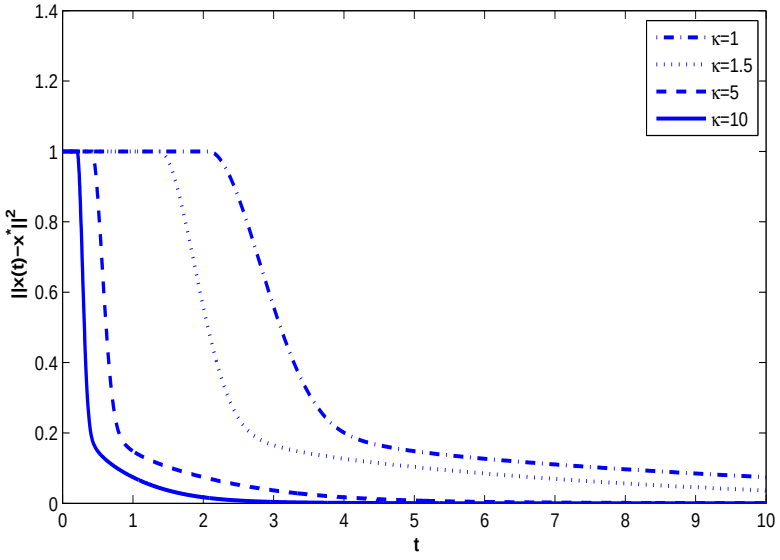
مثال ۵.۳.۵. مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \frac{x_۱^۲ + ۳x_۲^۲ + ۳x_۱x_۲ + x_۱ + x_۲ + ۲}{x_۱ + x_۲ + ۲} + \frac{\|x_۱ + x_۲ + ۱\|^۲}{۲x_۱ - ۳x_۲ + ۱} \tag{۴۵.۵}$$

$$\text{subject to}$$



شکل ۷.۵: منحنی فاز شبکه عصبی(۵۶.۴)-(۵۷.۴) با ۱۷ نقطه آغازین در مثال ۴.۳.۵.



شکل ۸.۵: رفتار همگرایی از $\|x(t) - x^*\|^2$ با $y_0 = (0, -1, -2, -3, -4, -5, -6, -7, -8)^T$ در مثال ۴.۳.۵.

$$\begin{cases} x_۱ + x_۲ + ۲ > ۰, \\ ۲x_۱ - ۳x_۲ + ۱ > ۰, \\ x_۱ + x_۲ \leq ۴, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۴۶.۵}$$

برنامه ریزی کسری در این مثال قابل بیان به‌صورت زیر است:

$$\text{Minimize} \quad \frac{\left\| \begin{bmatrix} ۰/۰۷۷۹۴x_۱ + ۰/۶۲۶۵x_۲ \\ ۰/۶۲۶۵x_۱ + ۱/۶۱۴۸x_۲ \end{bmatrix} + \begin{bmatrix} ۰/۵۷۰۶ \\ ۰/۰۸۸۳ \end{bmatrix} \right\|^۲}{x_۱ - x_۲ + ۲} + \frac{۱/۶۶۶۷۵}{x_۱ - x_۲ + ۲} + \frac{\|x_۱ + x_۲ + ۱\|^۲}{۲x_۱ - ۳x_۲ + ۲}$$

subject to

$$\begin{cases} x_۱ - x_۲ + ۲ > ۰, \\ ۲x_۱ - ۳x_۲ + ۱ > ۰, \\ x_۱ + x_۲ \leq ۴, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۴۷.۵}$$

جواب بهینه مسأله (۴۵.۵)-(۴۶.۵) ، $x^* = (۰, ۰)^T$ است. شکل (۹.۵) بیان کننده مسیرهای $x(t)$ با ۵۰ نقطه اولیه متفاوت و $\epsilon = ۱۰^{-۶}$ است. واضح است، وقتی نقاط اولیه از نقاط نشدنی انتخاب شوند، مسیرهای شبکه همچنان همگرا به x^* هستند. منحنی فاز متغیرهای حالت $(x_۱(t), x_۲(t))^T$ ، با ۱۷ حالت اولیه متفاوت در شکل (۱۰.۵) نشان داده شده. نرم خطای $L_۲$ بین x و x^* با ۳۰ نقطه اولیه متفاوت در شکل (۱۱.۵) رسم شده است.

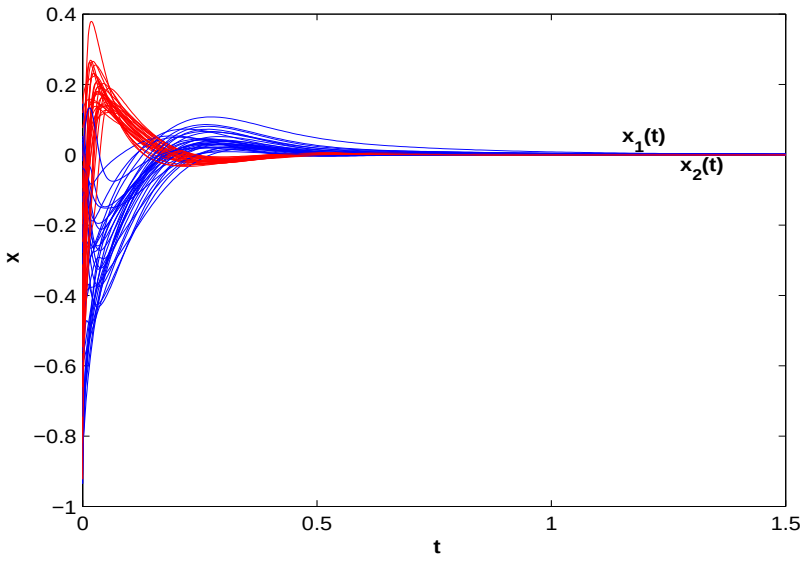
مثال ۶.۳.۵. مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \frac{\|-۳x_۱ + ۴x_۲ + ۵\|^۲}{-x_۱ + ۲x_۲ + ۴} + \frac{\|۴x_۱ - ۳x_۲ + ۴\|^۲}{x_۱ + x_۲ + ۵} \tag{۴۸.۵}$$

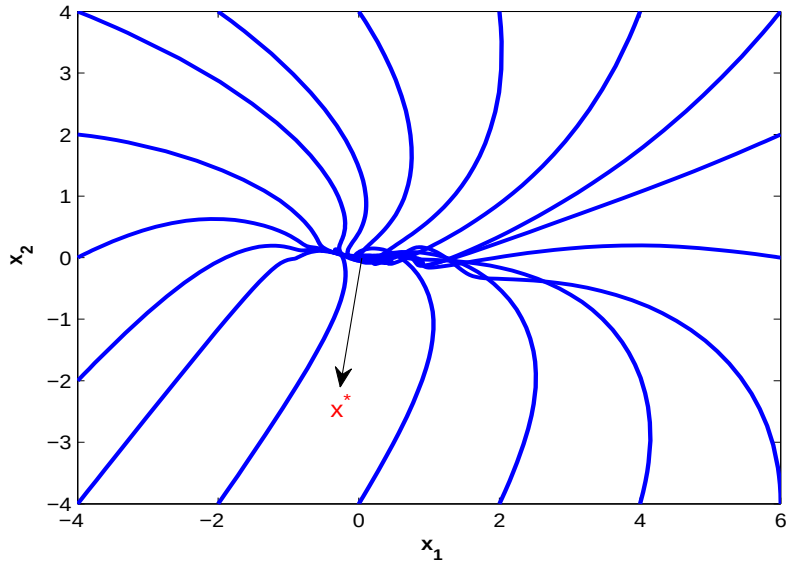
subject to

$$\begin{cases} -x_۱ + ۲x_۲ + ۴ > ۰, \\ x_۱ + x_۲ + ۵ > ۰, \\ ۲x_۱^۲ + x_۲^۲ - ۲x_۱x_۲ + ۲x_۱ - x_۲ - ۶ \leq ۰, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۴۹.۵}$$

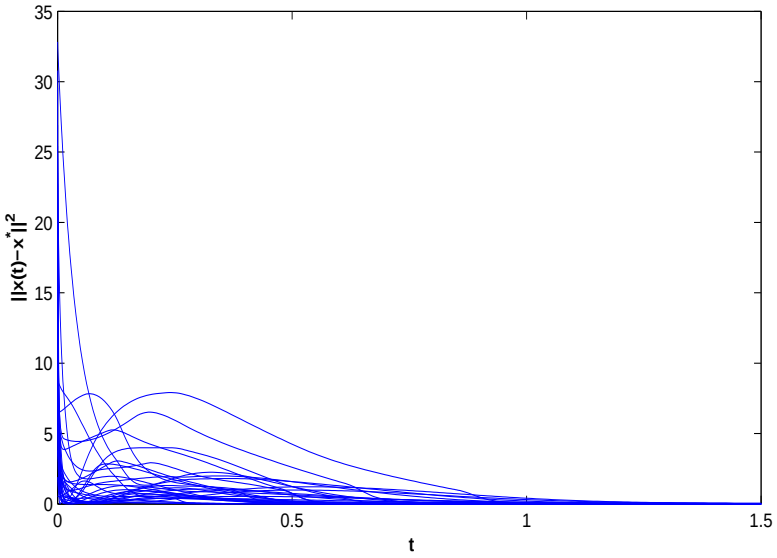
با تبدیل مسأله بالا به برنامه‌ریزی مخروط مرتبه دوم محدب جواب بهینه مسأله



شکل ۹.۵: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۵۰ نقطه آغازین دلخواه در مثال ۵.۳.۵.



شکل ۱۰.۵: منحنی فاز شبکه عصبی(۵۶.۴)-(۵۷.۴) با ۱۷ نقطه آغازین در مثال ۵.۳.۵.



شکل ۱۱.۵: رفتار همگرایی از $\|x(t) - x^*\|^2$ با ۳۰ نقطه اولیه متفاوت در مثال ۵.۳.۵.

$x^* = (۰/۳۱۸, ۰, ۶/۰۶۲۸, ۳/۳۸۴۴)^T$ به‌دست می‌آید. رفتار گذرا از شبکه عصبی (۵۶.۴)-(۵۷.۴) با $x_۰ = (-۱, ۱)^T$ و $\epsilon = ۱۰^{-۶}$ در شکل (۱۲.۵) و منحنی فاز این مدل با ۱۸ نقطه اولیه متفاوت در شکل (۱۳.۵) رسم شده است. شکل (۱۴.۵) ، تأثیر پارامتر κ از شبکه عصبی روی مقدار $\|x(t) - x^*\|^2$ نشان می‌دهد.

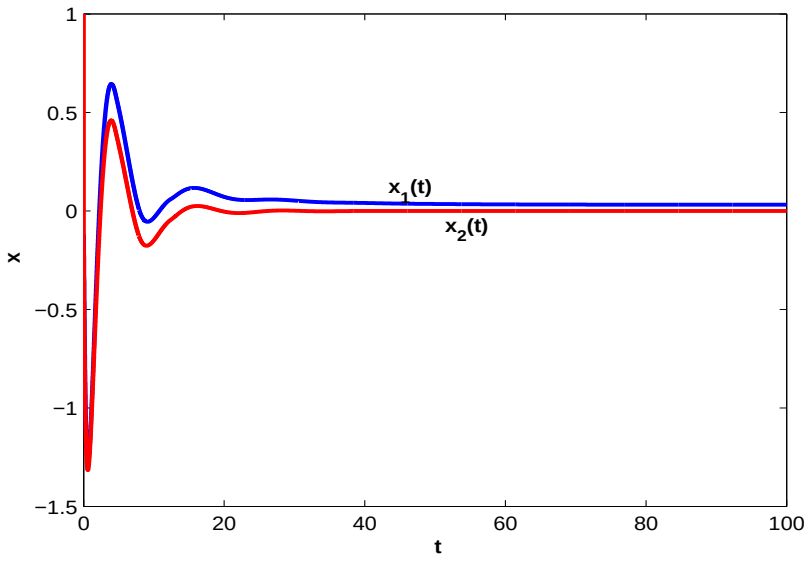
مثال ۷.۳.۵. مسأله برنامه‌ریزی کسری زیر را در نظر بگیرید:

$$\text{Minimize} \quad \frac{۴x_۱^۲ + ۴x_۲^۲ - ۴x_۱x_۲ - x_۱ - ۴x_۲ + ۶}{۳x_۱ - ۷x_۲ + ۴} + ۳x_۱^۲ + ۲x_۲^۲ + ۲x_۱x_۲ - ۴x_۱ - ۲x_۲ + ۴$$

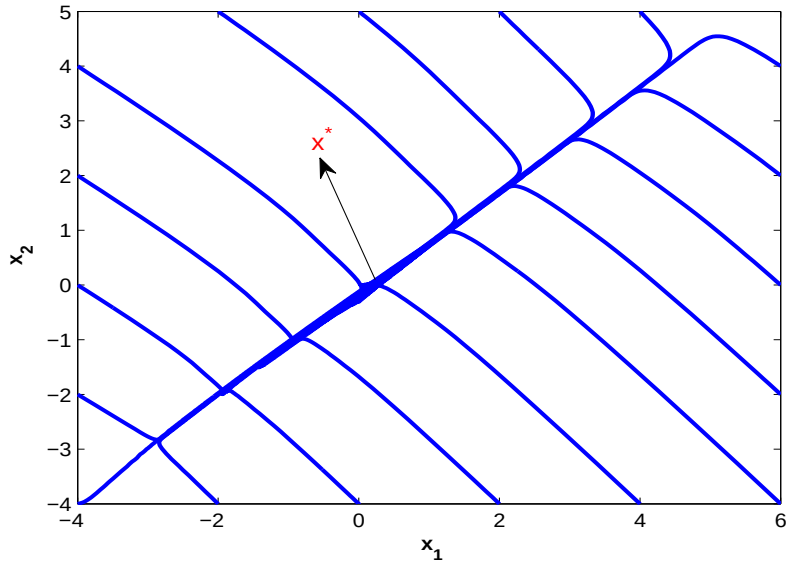
$$\text{subject to}$$

$$\begin{cases} ۳x_۱ - ۷x_۲ + ۴ > ۰, \\ x_۱, x_۲ \geq ۰. \end{cases} \tag{۵۰.۵}$$

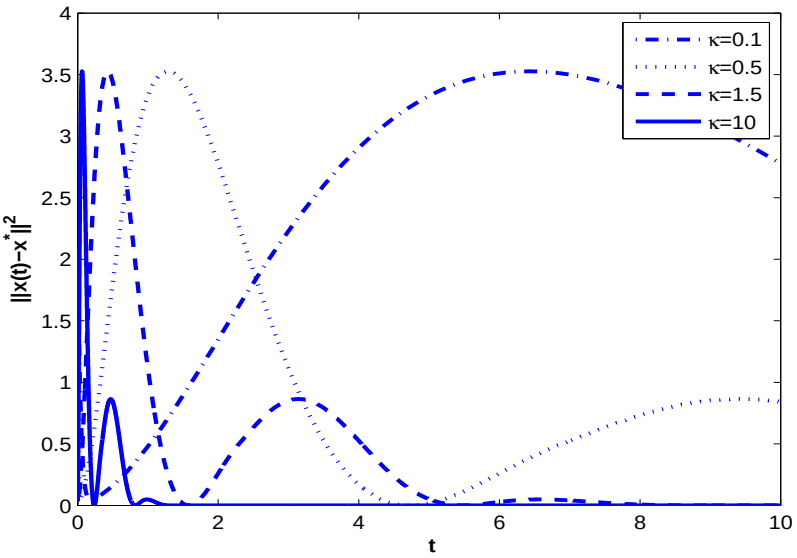
ابتدا مسأله را به برنامه‌ریزی مخروط مرتبه دوم محدب تبدیل کرده و آن را با شبکه عصبی (۵۶.۴)-(۵۷.۴) حل کردیم. شکل (۱۵.۵) نشان می دهد که مسیرهای شبکه پیشنهادی در (۵۶.۴)-(۵۷.۴) با



شکل ۱۲.۵: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با $x_0 = (-1, 1)^T$ در مثال ۶.۳.۵.

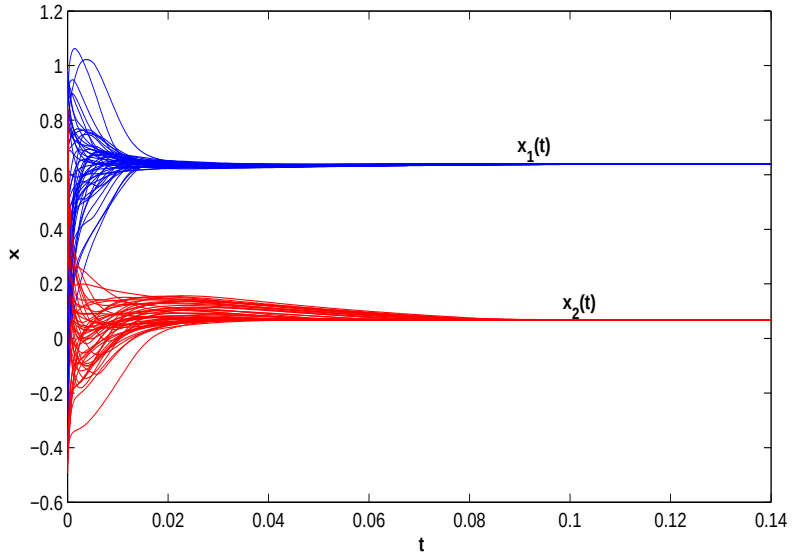


شکل ۱۳.۵: منحنی فاز شبکه عصبی(۵۶.۴)-(۵۷.۴) با ۱۸ نقطه آغازین در مثال ۶.۳.۵.

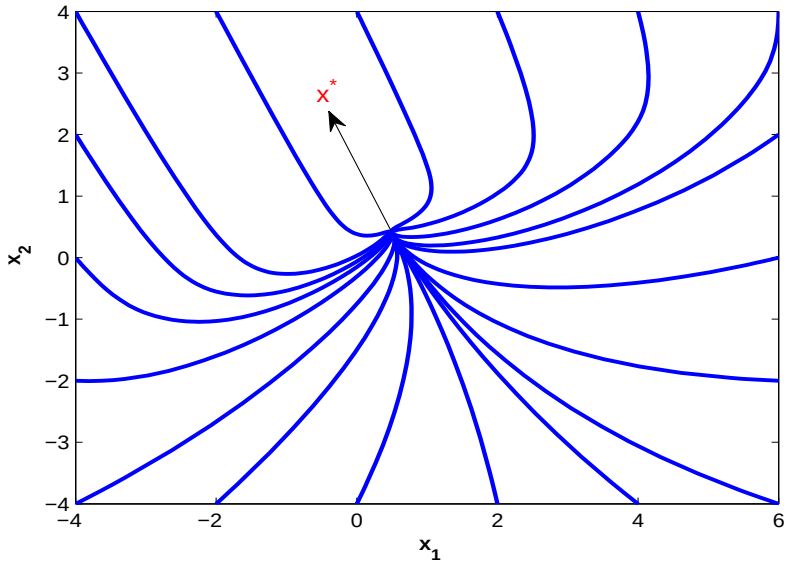


شکل ۱۴.۵: رفتار همگرایی از $\|x(t) - x^*\|^2$ با $y_0 = (-1, 1, -1, 1, -1, 1, -1, 1, -1)^T$ در مثال ۶.۳.۵.

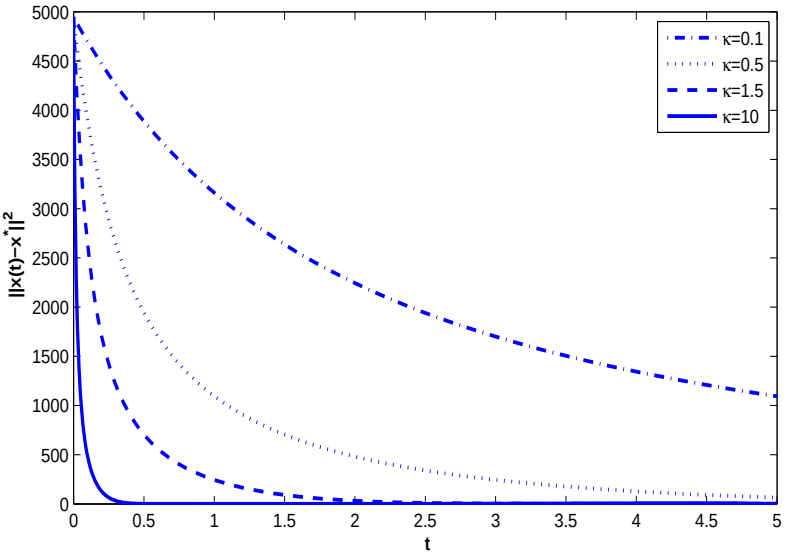
$\epsilon = 10^{-6}$ همگرا به $x^* = (0.6376, 0.0687)^T$ است. منحنی فاز متغیرهای حالت $(x_1(t), x_2(t))^T$ ، با ۱۸ حالت اولیه متفاوت و $\epsilon = 10^{-6}$ نیز در شکل (۱۶.۵) رسم شده است. واضح است که نقاط اولیه داخل و خارج ناحیه شدنی، مدل شبکه عصبی پیشنهادی همواره همگرا به جواب بهینه x^* است. رفتار همگرایی از $\|x(t) - x^*\|^2$ با نقطه اولیه $y_0 = (50, -50, 50, -50, 50, -50, 50, -50, 50, -50)^T$ هم در شکل (۱۷.۵) نشان داده شده است.



شکل ۱۵.۵: رفتار گذرا شبکه عصبی با ۵۰ نقطه آغازین دلخواه در مثال ۷.۳.۵.



شکل ۱۶.۵: منحنی فاز شبکه عصبی(۵۶.۴)-(۵۷.۴) با ۱۸ نقطه آغازین در مثال ۷.۳.۵.



شکل ۱۷.۵: رفتار همگرایی از $\|x(t) - x^*\|^2$ با y_0 در مثال ۷.۳.۵.

فصل ۶
بهینه‌سازی مقاوم

۱.۶ مقدمه

در این فصل از پایان نامه، ابتدا بهینه‌سازی داده‌های غیرقطعی و رویکردهای مقابله با عدم قطعیت در مدل‌ها را بیان می‌کنیم سپس چند نوع عمده از مجموعه‌های غیرقطعی را عنوان کرده و در ادامه با معرفی اشکال مختلف مقاوم بودن پرداخته و مروری بر کارهای انجام شده داریم. در ادامه با مسأله برنامه‌ریزی خطی مقاوم آشنا می‌شویم و در انتها به معرفی مسأله بهینه سازی مقاوم سبد سرمایه و ذکر یک مثال از آن می‌پردازیم.

۲.۶ عدم‌قطعیت در پارامترهای مدل

در هنگام ارائه یک مدل غالبا پارامترهای ورودی در آینده مشخص می‌شوند یا مقدار دقیق آن در زمان مدلسازی مسأله مشخص نیستند. عدم‌قطعیت[۱] در داده‌ها و پارامترهای مسأله، مشکلاتی در بسیاری از موارد عملی بخصوص مقادیر مالی مانند بازده سرمایه‌گذاری، ریسک و غیره به‌وجود می‌آورد. دو رویکرد کلی برای برخورد با این عدم‌قطعیت وجود دارد:

۱. برخورد واکنشی یا غیر فعال:

در این رویکرد که روش شناخته شده آن تحلیل حساسیت[۲] است، ابتدا مسأله را براساس تخمین نقطه‌ای پارامترهای غیرقطعی تحت عنوان مقادیر اسمی[۳] حل می‌کنند و بعد از حل به اندازه‌گیری حساسیت جواب مسأله نسبت به پارامترهای غیرقطعی می‌پردازند و محدوده تغییرات پارامترهای مذکور را طوری به‌دست می‌آورند که جواب مسأله همچنان شدنی باقی بماند. برای برآورد پارامترهای غیرقطعی کارهای زیادی صورت گرفته است. در این زمینه چاپرا [۴] و همکاران استفاده

[۱] Uncertainty
[۲] Sensitivity Analysis
[۳] Nominal Values
[۴] Chapra

از برآوردگر جیمز-استاین را برای میانگین‌ها پیشنهاد کردند، در حالیکه کلین و باوا [۵] ، فراست و ساوارینو[۶] ، بلک و لیترمن [۷] استفاده از برآوردگر بیزی [۸] را برای میانگینها و کواریانسها ارائه کردند اگر چه این روشها حساسیت سبد انتخابی را به مقدار برآورد شده پارامترها کاهش می‌دهد، ولی تضمینی برای کارایی میانگین-واریانس سبد سرمایه فراهم نمی‌کند. میچاد [۹] روشی را که مبتنی بر چند بار نمونه‌گیری از میانگین بازده‌ها (بوت استرپ [۱۰])و ماتریس کواریانس دارایی‌ها از فاصله اطمینان حول مقادیر اسمی پارامترها و سپس یکی کردن سبدهای به‌دست آمده از حل مسأله مارکویتز برای هر نمونه است پیشنهاد کرده است.

روش دیگر در برخورد واکنشی، حل مسأله با پارامتری در نظر گرفتن پارامترهای غیرقطعی است که به آن برنامه‌ریزی پارامتری می‌گویند.

در این رویکرد در فرایند حل مدل اقدامات لازم برای غیرقطعی بودن پارامترها در نظر گرفته نمی‌شود و کیفیت جواب از لحاظ برخورد با اختلالات و آشفتگی‌های موجود در پارامترهای قطعی بالا نمی‌رود. بنابراین به‌طور کلی جواب بهینه به‌دست آمده از رویکرد قطعی هم از نظر شدنی بودن و هم از نظر بهینگی تابع هدف مورد سوال است.

۲. برخورد سازنده یا فعال:

در این رویکرد به‌دنبال روشهایی هستیم که برخورد با عدم قطعیت در داده‌ها را قبل از حل و در متن مدل پیش‌بینی کند و چاره‌ای برای پایداری آنها ارائه کند. این روشها سعی در حل مسائل برنامه‌ریزی در شرایط عدم قطعیت دارد که هر یک مشکلات خاص خود را دارا هستند. لذا اعتقاد بر این است که به رویکردی نیاز است که در مقابل داده‌های غیرقطعی علاوه بر برخورد فعال و

[۵] Klein and Bava
[۶] Frost and Savarynv
[۷] Black and Lytrmn
[۸] Bayesian estimator
[۹] Mychad
[۱۰] Bootstrap

سازنده دارای خطا و هزینه کمی باشد. تکنیک مدل سازی اساسا متفاوت، برنامه‌ریزی تصادفی و بهینه سازی مقاوم [۱۱] در این رویکرد وجود دارد که در پایین شرح داده شده‌اند.

پس از رشد و گسترش مدل‌های برنامه‌ریزی خطی و روشهای حل آن، اولین برخورد فعال با داده‌های غیرقطعی رویکردی بود که اطلاعات احتمالی را درباره داده‌های مسأله به‌صورت برنامه‌ریزی تصادفی مطرح کرد که سابقه آن دنتزیگ [۱۲] در سال ۱۹۵۵ باز می‌گردد[۶۰]. در این روش مدل‌سازی، عدم‌قطعیت داده‌های مدل از همان ابتدا در مدل‌سازی در نظر گرفته می‌شوند. این روش برای حالتی استفاده می‌شود که عدم قطعیت آن طبیعت احتمالی داشته باشد و از توانایی تعیین توزیع احتمالی آن نیز برخوردار باشیم. لازم به ذکر است که در این نوع مدل‌سازی به دلیل عدم‌قطعیت موجود، امکان نقض محدودیت‌های مدل توسط جواب حاصله وجود دارد. حالت مهم آن برنامه‌ریزی خطی احتمالی است که در فصل قبل به بحث در مورد آن پرداختیم. همچنین برنامه‌ریزی پویای احتمالی نیز روش دیگری است که به‌نوعی به برنامه‌ریزی دارای روابط بازگشتی اطلاق می‌شود که برای نخستین بار توسط ریچارد بلمن [۱۳] در [۱۳۷] ارائه و توسعه داده شده است. این روش از روی مسائل برنامه‌ریزی سرچشمه گرفته است که تغییرات در طول زمان در آن‌ها دارای اهمیت است و برنامه‌ریزی پویا [۱۴] نامیده شده است. ایده این تکنیک تقسیم مسأله به چند مرحله بهینه‌سازی بازگشتی است که با ترکیب با عناصر احتمالی به برنامه‌ریزی پویای احتمالی تبدیل می‌شود.

در روش‌های بیان شده یکی از اصلی‌ترین مسائلی که در عمل به آن برخورد می‌کنیم دشواری جمع آوری اطلاعات جزئی درباره توزیع احتمالی عدم‌قطعیت‌های موجود در مدل است. در روشهای برنامه‌ریزی پویا و احتمالی با توجه به امکان افزایش تعداد سناریوهای (طرح‌های) [۱۵] درگیر در

[۱۱] Robust Optimization
[۱۲] Dantzing
[۱۳] Bellman
[۱۴] Dynamic programming
[۱۵] Scenarios

مسأله، معمولا جمع آوری این اطلاعات به‌شدت پرهزینه می‌شود به‌نحوی که استفاده از این روش مدل‌سازی را دچار مشکل می‌کند. باتوجه به این مطلب این روشها در عمل به‌طور گسترده مورد استفاده قرار نمی‌گیرند. بهینه‌سازی مقاوم یک تکنیک جدید توسعه داده شده برای مسائل مشابه مسائل برنامه‌ریزی احتمالی است. در این تکنیک یک فرض عمومی برای توزیع احتمالی پارامترهای دارای عدم‌قطعیت در نظر گرفته می‌شود تا بدین وسیله کار با مدل مسأله از نظر محاسباتی قابل کنترل شود.

در مرور ادبیات، معمولا بهینه‌سازی مقاوم به حوزه‌ای از بهینه‌سازی گفته می‌شود که ریشه در ادبیات کنترل مقاوم مهندسی [۱۶] دارد. در این حوزه باتوجه خاص به عدم‌نقض محدودیت‌های مدل (شرط موجه بودن مدل) به حل همتای مقاوم [۱۷] مدل اصلی با درنظر گرفتن مجموعه‌های عدم‌قطعیت پارامترها می‌پردازد. در حقیقت همتای مقاوم این مسائل، بدترین مدلسازی مسأله اصلی را در حوزه انحرافات پارامترها از مقدار اسمی آن‌ها بیان می‌کند. به هرحال به‌نوعی بدترین سناریوها به روشی هوشمندانه ایجاد می‌شود به‌نحوی‌که منجر به مدلی به‌شدت محافظه‌کارانه نشود. دربهینه‌سازی مقاوم، مدل‌سازی مسائل بادر نظر گرفتن مجموعه عدم‌قطعیت، تضمین کننده مناسب بودن جواب حاصل از بهینه‌سازی برای کلیه مقادیر ممکن پارامترهای مدل است. در این حالت، روش مذکور دارای فرضیات غیرتصادفی موجود در بهینه‌سازی احتمالی برای پارامترهای مدل نیست و اهمیت یکسانی برای تمام مقادیر ممکن از پارامترهای مدل در نظر می‌گیرد و به‌این ترتیب جواب به‌دست آمده در بهینه‌سازی به ازای تمام مقادیر پارامترهای مدل در فضای موجه [۱۸] قرار می‌گیرد. برنامه‌ریزی مقاوم برای هنگامی به‌کار می رود که تحلیلگر به‌دنبال رفتار مطلوب جواب به‌ازای تمامی رخدادهای ممکن برای داده‌هایی که دچار عدم‌اطمینان (قطعیت) هستند باشد

[۱۶] Robust Contorol Engineering
[۱۷] Robust Counterpart
[۱۸] Feasible Space

و براین اساس اهمیت یکسانی را برای تمام نمودهای مسأله قائل می‌شود. به‌عبارتی این تکنیک را برای زمانی که برخی از پارامترها با ریسک برآورد مواجه هستند، جواب بهینه به تغییرات کوچک حساس باشد و یا تصمیم‌گیرنده ریسک‌پذیری پایینی داشته باشد بسیار مناسب است.

۳.۶ انواع مجموعه‌های عدم‌قطعیت در بهینه‌سازی مقاوم

عدم اطمینان در پارامترها توسط مجموعه‌های عدم‌قطعیت منظور می‌شود، که شامل همه یا بیشتر مقادیر ممکنی است که می‌توان توسط پارامترهای نامعین رخ دهد. این مجموعه‌ها می‌توان براساس نظرات و تجربیات کارشناسان و یا برآوردهای گوناگون تکنیکهای آماری و یا بیزی بر مبنای داده‌های تاریخی تعیین کرد. دانستن چگونگی توزیع احتمالی داده‌ها برای تعیین مجموعه عدم‌قطعیت می‌تواند مفید باشد.
یک هدف مهم در رویکرد بهینه‌سازی مقاوم این است که معادل قطعی این مسائل به‌فرم صریحی در آیند که قابل حل با روش‌های حل موجود باشد. امکان وجود چنین فرمی برای هر مسأله مورد بررسی، علاوه بر ساختار مسأله به نوع مجموعه عدم‌قطعیت $\mathcal{U}$ نیز بستگی دارد. شکل و هندسه مجموعه عدم‌قطعیت، علاوه بر آن که بر خصوصیات معادل مقاوم مسأله تأثیر فراوانی می‌گذارد، می‌تواند در کنترل ماهیت محافظه کارانه رویکرد بهینه‌سازی مقاوم نیز نقش اساسی ایفا کند. تلاش‌های زیادی در جهت کنترل درجه محافظه‌کارانه بودن صورت مقاوم مسأله برنامه‌ریزی خطی به کمک تعریف مجموعه‌های عدم قطعیت جدید انجام شده است که در ادامه به چند مورد از آن‌ها اشاره خواهد شد.

۱. فضای نمونه گسسته: در این شکل بیشتر روشهای بهینه‌سازی با طرح‌های مجزا برای مقادیر ممکن پارامتر قابل طرح است که به‌صورت زیر بیان می‌شوند:

$$\mathcal{U} = \{p_1, p_2, ..., p_k\},$$

که در آن p_k بردارهای پارامترهای غیرقطعی برای سناریو k ام است.

۲. فضای نمونه پیوسته: عدم‌قطعیت بازه‌ای به‌واسطه سادگی و کاربرد فراوانش در شرایط عملی مورد توجه بسیاری از محققین قرار گرفته است.در این حالت، برای هر داده نامعلوم یک بازه تغییرات در نظر می‌گیریم که این بازه‌ها می‌توانند از روی فواصل اطمینان حاصل از نتایج آماری به‌دست آمده باشند:

$$\mathcal{U} = \{p : l \leq p \leq m\},$$

که در آن p بردار پارامترهای غیرقطعی است.

۳. مجموعه‌های عدم‌قطعیت براساس فضای نرمی: می‌توان فضای نرمی را به‌صورت زیر بیان کرد:

$$B_r(\theta) = \{x \in \mathbb{R} \mid \|x\|_r \leq \theta\}.$$

این فضا خصوصیات متقارن، محدب، بسته و کراندار بودن را داراست. در صورتی‌که θ را برابر یک در نظر بگیریم به آن دیسک واحد می‌گویند و اگر r را برابر دو فرض کنیم آنگاه این فضا یک بیضی‌گون را نشان می‌دهد. بنابراین مجموعه‌های عدم‌قطعیت را با استفاده از فضای نرمی به‌صورت زیر تعریف می‌کنیم:

$$\mathcal{U} = \{p : p = p_۰ + Mu \mid u \in B_r(\theta)\}.$$

در حقیقت این فضا میزان انحراف پارامترهای غیرقطعی از مقادیر اسمی خود یعنی بردار $p_۰$ را نشان می‌دهد. این انحراف تحت تأثیر شیب تغییرات، ماتریس M و به اندازه اختلالاتی که در فضای نرمی $B_r(\theta)$ می‌تواند رخ دهد، تعریف می‌شود.

۴. مجموعه‌های عدم‌قطعیت چندوجهی‌گونه (چند سقفی) [۱۹]: این مجموعه‌ها پوسته محدبی [۲۰] از تعداد محدودی سناریوی تولید شده برای مقادیر ممکن پارامترها هستند.

$$\mathcal{U} = \text{conv}\{p_۱, p_۲, ..., p_k\},$$

[۱۹] Polytopic
[۲۰] Convex Hull

۵. مجموعه‌های عدم‌قطعیت بیضی‌گون: مهم ترین کارها در زمینه عدم قطعیت بیضی گون توسط بن تال [۲۱] و نمیروفسکی[۲۲] و به‌طور مستقل از آن‌ها ال‌گوی [۲۳]و همکارانش انجام شده است. آن‌ها معادل مقاوم رده‌های مهمی از مسائل حوزه بهینه‌سازی محدب را تحت مجموعه عدم‌قطعیت توسعه دادند. در واقع می‌توان کارهای بن-تال و نمیروفسکی را هسته اولیه و اساسی تئوری بهینه‌سازی مقاوم قلمداد کرد. بن-تال و نمیروفسکی در[۱۳۸] اثبات کردند که تحت چندین فرض معقول، معادل مقاوم یک مسأله برنامه‌ریزی محدب خود یک برنامه‌ریزی محدب است. اما عمده کار آن‌ها در این مرجع و در [۱۳۹] به یافتن فرمی صریح برای معادل مقاوم چند رده عمده مسائل بهینه‌سازی محدب بر می‌گردد، آن ها نشان دادندکه:

۱. معادل مقاوم یک مسأله برنامه‌ریزی خطی تحت یک مجموعه عدم قطعیت بیضی‌گون یا اشتراک تعداد محدودی بیضی‌گون دارای فرم صریح برنامه ریزی مخروط مرتبه دوم است.

۲. معادل مقاوم یک مسأله برنامه‌ریزی محدب درجه دوم با قیود درجه دوم تحت یک مجموعه عدم قطعیت بیضی‌گون دارای فرم صریح برنامه‌ریزی نیمه معین [۲۴] است.

۳. معادل مقاوم یک مسأله برنامه‌ریزی مخروط مرتبه دوم تحت یک مجموعه عدم قطعیت بیضی‌گون و چند فرض نه چندان محدود کننده، یک برنامه‌ریزی نیمه معین است.

جزئیات قضایا و روش‌های حل ارائه شده توسط بن-تال و نمیروفسکی را می‌توانید در مراجع [۱۳۸]-[۱۴۶] ببینید.

[۲۱] Ben-Tal
[۲۲] Nemirovski
[۲۳] El Ghaoui
[۲۴] Semi-Definite programming(SDP)

۴.۶ برنامه‌ریزی خطی مقاوم

مسأله زیر را در نظر بگیرید:

$$\begin{aligned} &\text{Minimize} \qquad c^T x \\ &\text{subject to} \\ &\qquad a_i{}^T x \le b_i, \quad i = ۱, ..., m, \end{aligned} \tag{۱.۶}$$

که در آن برخی از پارامترهای c، a_i و b_i غیرقطعی هستند. برای سادگی فرض می‌کنیم c و b_i ها ثابت و a_i ها هم با مجموعه غیرقطعی بیضی‌گونهایی به‌فرم زیر تعریف می‌شوند:

$$a_i \in \varepsilon_i = \{\bar{a}_i + P_i u \mid \|u\| \le ۱\},$$

وقتی $P_i \succeq ۰$ ، یعنی ماتریس‌های متقارن و نیمه معین مثبت هستند. در بدترین حالت، ما نیاز داریم که قیود برای همه مقادیر ممکن از پارامترهای a_i صدق کند که منجر به برنامه‌ریزی خطی مقاوم زیر می‌شود:

$$\begin{aligned} &\text{Minimize} \qquad c^T x \\ &\text{subject to} \\ &\qquad a_i{}^T x \le b_i \quad \forall a_i \in \varepsilon_i, \;\; i = ۱, ..., m. \end{aligned} \tag{۲.۶}$$

قیود مقاوم خطی بالا را می‌توان به‌صورت زیر بیان کرد:

$$\max\{a_i{}^T x \mid a_i \in \varepsilon_i\} = \bar{a}_i{}^T x + \|P_i x\| \le b_i, \tag{۳.۶}$$

واضح است که قیود مخروط مرتبه دوم می‌باشد. بنابراین (۸.۶) را می‌توان به‌صورت برنامه‌ریزی مخروط مرتبه دوم زیر بیان کرد:

$$\text{Minimize} \qquad c^T x$$

subject to

$$\bar{a_i}^T x + \|P_i x\| \leq b_i \quad i = ۱, ..., m. \tag{۴.۶}$$

همچنین برنامه ریزی‌خطی مقاوم را می‌توان در چارچوب آماری در نظر گرفت. در این حالت ما فرض می‌کنیم پارامترهای a_i متغیرهای تصادفی نرمال مستقل، با میانگین $\bar{a_i}$ و واریانس Σ_i باشند. ما نیاز داریم که هر قید با احتمال حداقل η وقتی $\eta \geq ۰$ باشد، برقرار باشد یعنی:

$$\mathrm{P}(a_i^T x \leq b_i) \geq \eta. \tag{۵.۶}$$

می‌خواهیم نشان دهیم که قید احتمالی فوق را می‌توان به‌صورت یک قید مخروط مرتبه دوم بیان کرد. فرض کنید $u = a_i^T x$، با $\sigma^۲$ که نشان دهنده واریانس آنها است، بنابراین داریم:

$$\mathrm{P}(\frac{u - \bar{u}}{\sqrt{\sigma^۲}} \leq \frac{b_i - \bar{u}}{\sqrt{\sigma^۲}}) \geq \eta.$$

چون $z = \frac{u - \bar{u}}{\sqrt{\sigma^۲}} \sim N(۰, ۱)$ و همچنین،

$$\Phi(z) = \frac{۱}{\sqrt{۲\pi}} \int_{-\infty}^{z} e^{-\frac{t^۲}{۲}} dt,$$

بنابراین قید (۵.۶) را می‌توان به‌صورت زیر بیان کرد:

$$\frac{b_i - \bar{u}}{\sqrt{\sigma^۲}} \geq \Phi^{-۱}(\eta),$$

یا

$$\bar{u} + \Phi^{-۱}(\eta)\sqrt{\sigma^۲} \leq b_i.$$

از $\bar{u} = \bar{a_i}^T x$ و $\sigma^۲ = x^T \Sigma_i x$ داریم:

$$\bar{a_i}^T x + \Phi^{-۱}(\eta)\|\Sigma_i^{\frac{۱}{۲}} x\| \leq b_i.$$

اکنون در صورتیکه $\eta \geq ۰/۵$ یعنی $(\Phi^{-۱}(\eta) \geq ۰)$ باشد، این قید یک قید مخروط درجه دوم است. بنابراین، مسأله

$$\text{Minimize} \qquad c^T x$$

$$\begin{aligned} &\text{subject to} \\ &\mathrm{P}(\bar{a_i}^T x \leq b_i) \geq \eta \quad i = ۱, ...m. \end{aligned} \tag{۶.۶}$$

می‌تواند با برنامه‌ریزی مخروط مرتبه دوم زیر بیان شود:

$$\begin{aligned} &\text{Minimize} \qquad c^T x \\ &\text{subject to} \\ &\bar{a_i}^T x + \Phi^{-۱}(\eta)\|\Sigma_i^{\frac{۱}{۲}} x\| \leq b_i \quad i = ۱, ...m. \end{aligned} \tag{۷.۶}$$

که این مسائل با روشهای نقطه درونی قابل حل بوده، ما نیز برای حل این مسائل از شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) کمک می‌گیریم.

۵.۶ اشکال گوناگون مقاوم بودن

۱.۵.۶ مقاوم بودن محدودیت‌ها

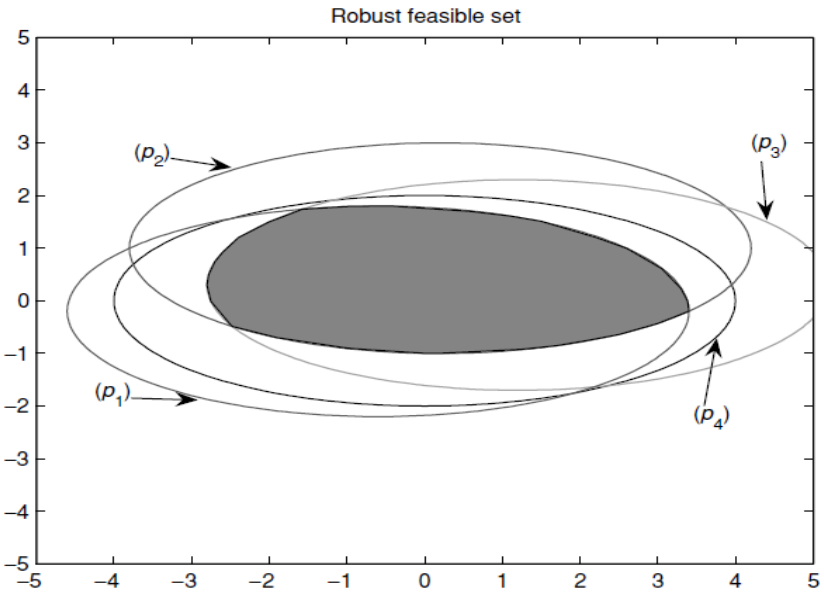
یک مفهوم مهم در این تکنیک مقاوم بودن محدودیت‌ها[۲۵] است که اغلب به آن مقاوم بودن مدل[۲۶] می‌گویند. این مقاوم بودن مربوط به به‌دست آوردن جوابهایی است که برای تمامی مقادیر ممکن ورودی‌های غیرقطعی شدنی باقی بماند. مدل زیر را در نظر بگیرید:

$$\begin{aligned} &\text{Minimize}_{\mathrm{x}} \qquad f(x) \\ &\text{subject to} \\ &G(x, p) \in K. \end{aligned} \tag{۸.۶}$$

در صورتیکه p را پارامترهای نامطمئن مسأله و $\mathcal{U}$ را مجموعه عدم قطعیت در نظر بگیریم، مدل مقاوم مسأله به‌صورت زیر خواهد بود:

$$\text{Minimize}_{\mathrm{x}} \qquad f(x)$$

[۲۵] Constraint Robustness
[۲۶] Model Robustness



شکل ۱.۶: مقاوم بودن محدودیتها

$$\text{subject to}$$

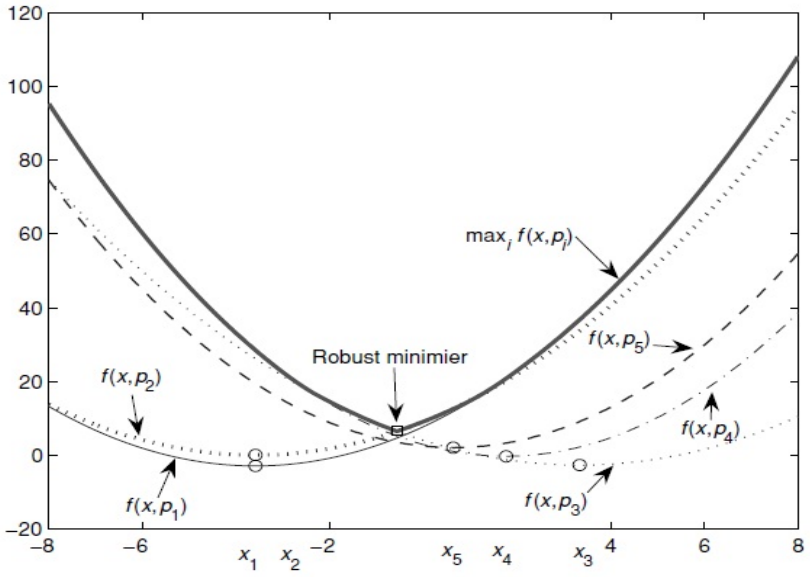
$$G(x,p) \in K, \quad \forall p \in \mathcal{U}. \tag{۹.۶}$$

مسأله (۸.۶) نشان می‌دهد، مجموعه شدنی مقاوم اشتراک مجموعه‌های شدنی $S(p) = \{x : G(x,p) \in K\}$ ، با مجموعه عدم قطعیت $\mathcal{U}$ است. ما آن را در شکل (۱.۶) برای مجموعه بیضی‌گون $\mathcal{U} = \{p_۱, p_۲, p_۳, p_۴\}$ ، وقتیکه p_i متناظر با مرکز عدم قطعیت بیضی است نشان داده‌ایم.

۲.۵.۶ مقاوم بودن تابع هدف

مفهوم دیگری که در مقاوم بودن مطرح است، مقاوم بودن تابع هدف است. در این حالت به‌دنبال جواب‌هایی هستیم که برای تمام حالت‌های ممکن پارامترهای غیرقطعی، نزدیک به بهینگی باقی بماند. مدل زیر را برای بردار p در تابع هدف در نظر بگیرید:

$$\text{Minimize}_{\text{x}} \quad f(x,p)$$



شکل ۲.۶: مقاوم بودن تابع هدف

$$\text{subject to}$$

$$x \in S, \tag{۱۰.۶}$$

که درآن S مجموعه شدنی و تابع هدف که وابسته به پارامترهای غیرقطعی p است. بنابراین با حل مسأله زیر جواب مقاوم تابع هدف به‌دست می‌آید:

$$\text{Minimize}_{\text{x}\in\text{S}} \ \ \text{Maximize}_{\text{p}\in\mathcal{U}} \ \ f(x,p) \tag{۱۱.۶}$$

مسأله مقاوم بودن تابع هدف (۱۱.۶) در شکل (۲.۶) نشان داده شده است. در این مثال، S مجموعه شدنی خط حقیقی، $\mathcal{U} = \{p_۱, p_۲, p_۳, p_۴, p_۵\}$ و $f(x,p_i)$ تابع درجه دوم محدب است. برای بررسی اشکال دیگر مرجع [۱۴۷] را ببینید.

۶.۶ مدل های بهینه سازی مقاوم در مسائل مالی

بسیاری از مسائل مالی شامل مقادیر آتی از قیمت دارایی‌ها، نرخ بهره، نرخ ارز و غیره است که در زمان حال معلوم نیستند اما می‌توان آنها را پیش‌بینی یا برآورد کرد. از جمله فرمول‌های بهینه‌سازی مقاوم برای مسائل مالی می‌توان به مسائل انتخاب سبد سرمایه[۲۷]، مدیریت ریسک[۲۸]، پوشش و قیمت‌گذاری مشتقات مالی[۲۹] اشاره کرد.

۱.۶.۶ انتخاب سبد سرمایه مقاوم

بن-تال و نمیروفسکی در [۱۴۸] یک مسأله انتخاب سبد سرمایه چند دوره‌ای مقاوم را با استفاده از رویکرد برنامه‌ریزی خطی فرمول بندی کردند که با روش سیمپلکس یا روشهای نقطه درونی قابل حل است. گلدفارب و آینگار [۳۰] [۱۴۹] و نیز لوبو و بوید [۳۱] در [۱۲۵] مسأله سبد مالی مقاوم را در چارچوب میانگین-واریانس مورد مطالعه قرار دادند و برای بازده دارایی‌های ریسکی، برخی از انواع عدم‌قطعیت مانند بیضی‌گون و بازه‌ای را معرفی کردند. آن‌ها مسأله را به برنامه‌ریزی مخروط مرتبه دوم و برنامه‌ریزی نیمه معین تبدیل کردند، که توسط الگوریتم‌های نقطه درونی به‌صورت کارایی قابل حل هستند.

در این بخش با یادآوری مدل میانگین-واریانس مارکویتز می‌توان ترکیبی از بازدهی و ریسک را در تابع هدف به‌صورت زیر داشته باشیم:

$$\mathrm{Max}_{\mathrm{x}\in\mathcal{X}} \;\; \mu(x) - \lambda x^T \Sigma x \qquad (۱۲.۶)$$

وقتیکه μ_i برآورد از بازدهی موردانتظار از دارایی iام، σ_{ij} کواریانس بین بازدهی دارایی i و j ام و λ هم نرخ ریسک گریزی [۳۲] که تعادلی بین ریسک و بازده برقرار می‌سازد، مجموعه $\mathcal{X}$، مجموعه همه سبد

[۲۷] Portfolio Selection
[۲۸] Risk Managment
[۲۹] Dreivatives Pricing/Hedging
[۳۰] Goldfarb and Iyengar
[۳۱] Lobo and Boyd
[۳۲] Risk-Aversion

سرمایه‌های شدنی که شامل اطلاعاتی چون، محدود کردن موقعیت کوتاه، استفاده از کل بودجه و غیره است. چون قیود ما مشخص شده هستند، ما می‌توانیم فرض کنیم مجموعه $\mathcal{X}$ بدون هیچ عدم قطعیتی در زمانی که مسأله حل می‌شود. یکی از محدودیت‌های این مدل، برآورد دقیق و صحیح از بازدهی‌های مورد انتظار و واریانس‌هاست. در [۱۵۰] باوا [۳۳] و همکارانش درباره استفاده برآوردهایی از بازدهی‌های مورد انتظار نامعلوم و کواریانس بحث کردند که منجر به برآورد ریسک در انتخاب سبد سرمایه شد. بنابراین، جواب بهینه به آشفتگی در پارامترهای ورودی-تغییرات کوچک در برآورد بازدهی و واریانس حساس است، که شاید منجر به تغییرات بزرگ در جواب متناظر باشد. برای مثال [۱۵۱، ۱۵۲] را ببینید. برای بهینه‌سازی مقاوم سبد سرمایه، ما مدلی که اطلاعات بازدهی و ماتریس کواریانس دارای فرم بازه‌ای هستند را به‌صورت زیر نظر می‌گیریم:

$$\mathcal{U} = \{(\mu, \Sigma) : \mu^L \leq \mu \leq \mu^U,\ \Sigma^L \leq \Sigma \leq \Sigma^U,\ \Sigma \succeq 0\}. \qquad (۱۳.۶)$$

با استفاده از مدل مقاوم بودن تابع هدف در (۱۱.۶)، ما می‌توانیم سبد سرمایه را پیدا کنیم که تابع هدف در (۱۲.۶) را در بدترین حالت تشخیص پارامترهای ورودی از مجموعه $\mathcal{U}$ در (۱۳.۶) بیشینه کند. با در نظر گرفتن این شرایط مسأله بهینه‌سازی مقاوم به‌صورت زیر بیان می‌شود:

$$\mathrm{Max}_{\mathrm{x}\in\mathcal{X}}\ \ \{\mathrm{Min}_{(\mu,\Sigma)\in\mathcal{U}} - \mu(x) + \lambda x^T \Sigma x\}. \qquad (۱۴.۶)$$

چون $\mathcal{U}$ کراندار است، با استفاده از نتایج کلاسیک از آنالیز محدب می‌توان نشان داد که (۱۴.۶) معادل دوگان آن است وقتی که جای Max و Min با هم عوض می‌شود:

$$\mathrm{Min}_{(\mu,\Sigma)\in\mathcal{U}}\ \ \{\mathrm{Max}_{\mathrm{x}\in\mathcal{X}}\mu(x) - \lambda x^T \Sigma x\}.$$

بنابراین، جواب مسأله (۱۴.۶) نقطه زینی تابع $f(x, \mu, \Sigma) = \mu(x) - \lambda x^T \Sigma x$ است و با استفاده از تکنیک‌های نقطه درونی در [۱۵۳] تعیین می‌شود.

[۳۳] Bawa

۲.۶.۶ مثال عددی

در این بخش به ذکر یک مثال عددی که از [۱۵۴] گرفته شده می‌پردازیم، مسأله زیر را در نظر بگیرید:

$$\text{Maximize} \quad \mu_۱x_۱ + \mu_۲x_۲ + \mu_۳x_۳ \tag{۱۵.۶}$$

$$\text{subject to}$$

$$\begin{cases} TE(x_۱, x_۲, x_۳) \leq ۰/۱۰, \\ x_۱ + x_۲ + x_۳ = ۱, \\ x_۱, x_۲, x_۳ \geq ۰, \end{cases} \tag{۱۶.۶}$$

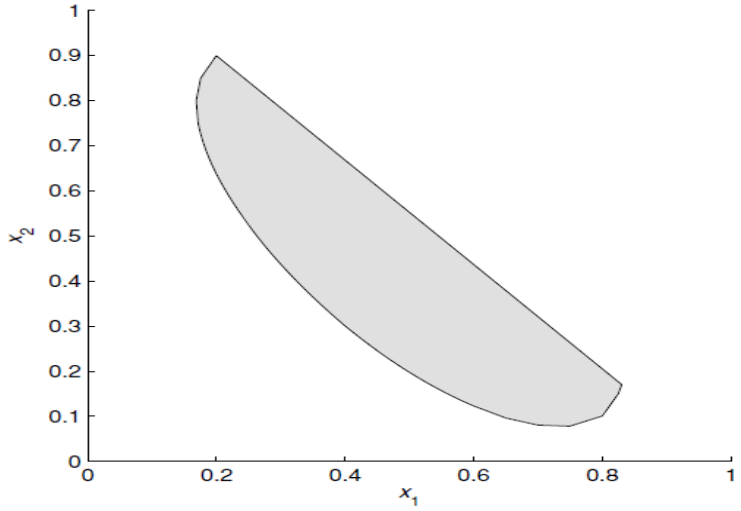
که درآن

$$TE(x_۱, x_۲, x_۳) = \sqrt{\begin{bmatrix} x_۱ - ۰/۵ \\ x_۲ - ۰/۵ \\ x_۳ \end{bmatrix}^T \begin{bmatrix} ۰/۱۷۶۴ & ۰/۰۹۷۰۲ & ۰ \\ ۰/۰۹۷۰۲ & ۰/۱۰۸۹ & ۰ \\ ۰ & ۰ & ۰ \end{bmatrix} \begin{bmatrix} x_۱ - ۰/۵ \\ x_۲ - ۰/۵ \\ x_۳ \end{bmatrix}}.$$

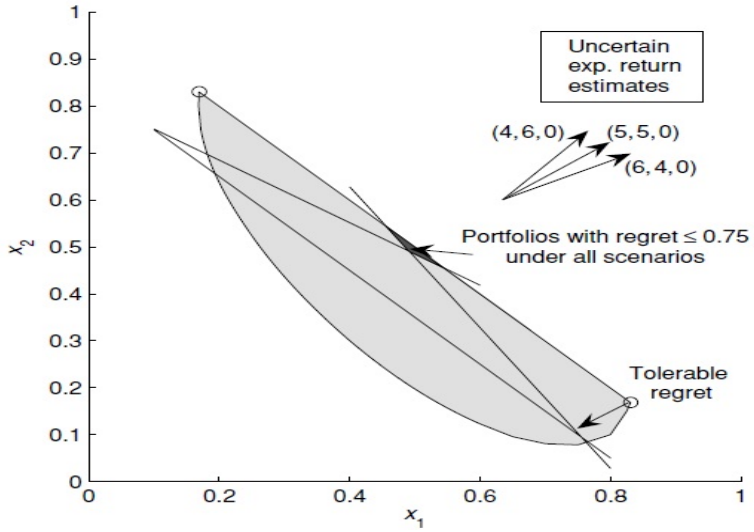
در این مسأله سبد سرمایه شامل ۳ دارایی است که دارایی سوم متناظر با $x_۳$، نسبتی از بودجه است که سرمایه‌گذاری نشده است. دو دارایی اول دارای انحراف معیارهای ۴۲٪ ، ۳۳٪ و ضریب همبستگی ۰/۷ هستند. محک [۳۴] در سبد سرمایه، سرمایه‌گذاری نصف-نصف میان دو دارایی اول است. تابع $TE(x)$ ، خطای ردیابی (پیگردی)[۳۵] است که باید کمتر از ۱۰٪ باشد، این تابع که در آن x_{BM} مقدار محک است به‌صورت زیر تعریف می‌شود:

$$TE(x) = \sqrt{(x - x_{BM})^T \Sigma (x - x_{BM})}.$$

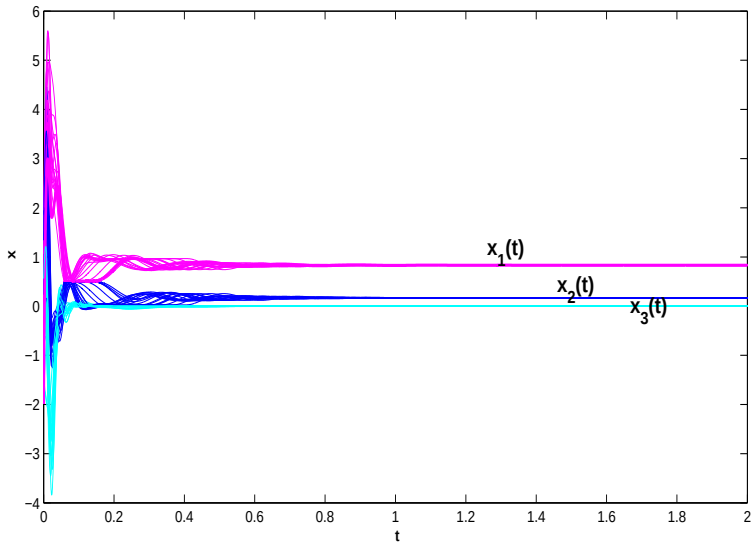
ما در شکل (۳.۶) ناحیه شدنی این مسأله را که توسط فضای متغیرهای $x_۱$ و $x_۲$ پوشانده شده را می‌بینیم. اکنون یک مدل مقاوم بودن نسبی برای این مسأله سبد سرمایه می‌سازیم، فرض کنید برآورد (تخمین) ماتریس کواریانس، قطعی است و مجموعه عدم‌قطعیت ساده‌ای شامل ۳ سناریو که در شکل (۴.۶) با پیکان‌هایی نشان داده شده است برای تخمین بازدهی موردانتظار در نظر می‌گیریم. این ۳ سناریو متناظر



شکل ۳.۶: ناحیه شدنی مسأله (۱۵.۶)-(۱۶.۶)



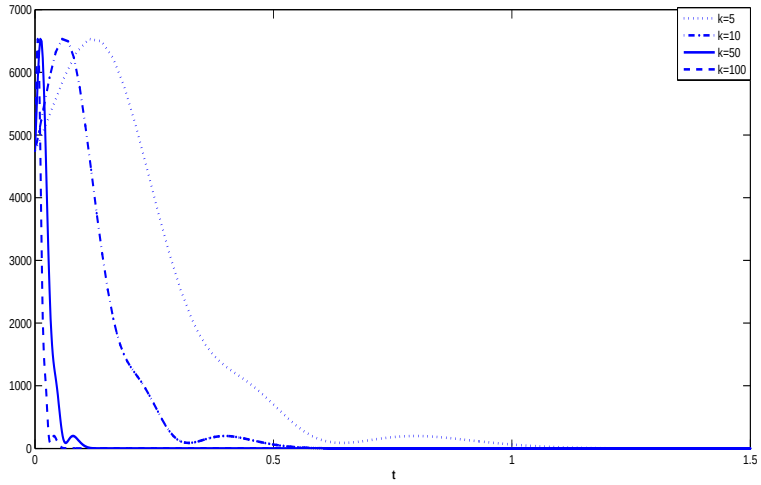
شکل ۴.۶: مجموعه جوابهایی با پشیمانی کمتر از ۰/۷۵ مسأله (۱۵.۶)-(۱۶.۶)



شکل ۵.۶: رفتار گذرا شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۲۰ نقطه آغازین دلخواه

با مقادیر $(\mu_۱, \mu_۲, \mu_۳) = (۶, ۴, ۰), (۵, ۵, ۰), (۴, ۶, ۰)$ هستند. به نظر می‌رسد مسأله فوق یک حالت خاص از برنامه‌ریزی مخروط مرتبه دوم است، بنابراین ما برای حل این مسأله از مدل شبکه عصبی پیشنهاد شده (۵۶.۴)-(۵۷.۴) استفاده می‌کنیم. جواب بهینه این مسأله وقتی $(\mu_۱, \mu_۲, \mu_۳) = (۶, ۴, ۰)$ باشد، $X^* = (۰/۸۳۱, ۰/۱۶۹, ۰)$ است، با مقدار تابع هدف ۵/۶۶۲ . شکل (۵.۶) نشان می‌دهد که مسیرهایی از شبکه عصبی (۵۶.۴)-(۵۷.۴) با ۲۰ نقطه اولیه متفاوت و $\epsilon = ۱۰^{-۶}$ همگرا به جواب بهینه است. همچنین تأثیر پارامتر κ از شبکه روی $\|x(t) - x^*\|^۲$ در شکل (۶.۶) نشان داده شده است. مشابها، وقتی $(\mu_۱, \mu_۲, \mu_۳) = (۴, ۶, ۰)$ باشد، $X^* = (/ ۰/۱۶۹, ۰/۸۳۱, ۰)$ است، با مقدار تابع هدف ۵/۶۶۲، همچنین وقتی $(\mu_۱, \mu_۲, \mu_۳) = (۵, ۵, ۰)$ است تمام نقاط بین دو جواب بهینه قبلی، بهینه بوده با مقدار

[۳۴] Benchmark
[۳۵] Tracking Error



شکل ۶.۶: رفتار همگرایی از $\|x(t) - x^*\|^2$ با $y_0 = (40, -40, 40, -40, 40, -40, 40)^T$

هدف مشترک ۵. بنابراین فرم مقاوم بودن نسبی برای این مسأله می‌تواند به‌فرم زیر نوشته شود:([۱۴۷])

$$\text{Minimize}_{x,t} \quad t \tag{۱۷.۶}$$

$$\text{subject to}$$

$$\begin{cases} 5/662 - (6x_1 + 4x_2) \leq t, \\ 5/662 - (4x_1 + 6x_2) \leq t, \\ 5 - (5x_1 + 5x_2) \leq t, \\ TE(x_1, x_2, x_3) \leq 0/10, \\ x_1 + x_2 + x_3 = 1, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{cases} \tag{۱۸.۶}$$

به‌جای حل کردن مسأله وقتی که سطح پشیمانی [۳۶] بهینه ما متغیر است (t)، یک استراتژی ساده انتخاب، سطحی از تأسف که می‌تواند پورتفوهایی را بیابد که از سطح پشیمانی در هر سناریو تجاوز نکند. برای مثال حداکثر سطح پشیمانی قابل‌تحمل [۳۷] را ۰/۷۵ می‌گیریم بنابراین مسأله شدنی زیر را داریم:

$$\text{Find} \quad x \tag{۱۹.۶}$$

$$\text{subject to}$$

[۳۶] Regret Level
[۳۷] Tolerable

$$\begin{cases} ۵/۶۶۲ - (۶x_۱ + ۴x_۲) \leq ۰/۷۵, \\ ۵/۶۶۲ - (۴x_۱ + ۶x_۲) \leq ۰/۷۵, \\ ۵ - (۵x_۱ + ۵x_۲) \leq ۰/۷۵, \\ TE(x_۱, x_۲, x_۳) \leq ۰/۱۰, \\ x_۱ + x_۲ + x_۳ = ۱, \\ x_۱ \geq ۰, x_۲ \geq ۰, x_{۳.} \geq ۰ \end{cases} \quad (۲۰.۶)$$

این مسأله و مجموعه شدنی از جواب‌های آن در شکل (۴.۶) نشان داده شده است. گوشه مثلث که تیره‌تر نشان داده شده است بیانگر پورتفوهایی است که دارای سطح پشیمانی کمتر و مساوی ۰/۷۵ تحت هر سه طرح هستند.

نتایج و پیشنهاداتی برای کارهای آتی

نتایج

در این پایان نامه سعی کرده‌ایم مروری بر شبکه‌های عصبی داشته و با استفاده از این نظریه به حل رده‌ای از مسائل بهینه‌سازی بپردازیم. یکی از انواع مسائل بهینه سازی، مسائل مالی هستند که به چهار دسته کلی مدیریت ریسک، انتخاب سبد سرمایه، قیمت‌گذاری اختیارمعامله‌ها، مدیریت بدهی و دارایی تقسیم می‌شوند. این مسائل شامل برنامه‌ریزی خطی، برنامه‌ریزی درجه دوم، برنامه‌ریزی مخروط، بهینه‌سازی مقاوم و برنامه‌ریزی تصادفی هستند. برای مسائل انتخاب سبد سرمایه، مدل‌هایی با فرم برنامه‌ریزی خطی و برنامه‌ریزی درجه دوم معرفی کردیم و برای حل آنها به ارائه یک مدل از شبکه‌های عصبی پرداختیم. همچنین برای مسائل برنامه‌ریزی مخروط مرتبه دوم نیز مدل دیگری از شبکه‌های عصبی ارائه کردیم. این مدل پیشنهادی را برای حل مسائل قابل تبدیل به برنامه‌ریزی مخروط مرتبه دوم محدب از جمله برنامه‌ریزی تصادفی، رده‌ای از مسائل برنامه‌ریزی کسری و بهینه سازی مقاوم بکار بردیم. همچنین پایداری و همگرایی سراسری مدل‌ها مورد بررسی قرار گرفت و کارایی مدل‌های پیشنهاد شده را با حل چندین مثال عددی نشان دادیم.

از جمله مزیت‌های مدل‌های ارائه شده می‌توان به موارد زیر اشاره کرد:

* همگرایی سراسری مدل و وابسته نبودن به نقطه اولیه لزوما شدنی برخلاف بیشتر روشهای بهینه سازی از جمله روشهای نقطه درونی و سیمپلکس که ممکن است مسأله جواب یکتا نداشته باشد

و در دور بیافتد.

* نداشتن پارامتر جریمه بر خلاف مدل‌های گرادیانی و سادگی مدل پیشنهاد شده.

* محدود نبودن مدل به تعداد متغیرها و قیود بر خلاف روش‌های دیگر و نرم افزارها.

پیشنهادات برای کارهای آتی

از جمله تحقیقات دیگری که در ادامه می‌توان انجام داد عبارتند از:

۱. کلاس‌های مختلفی از مسائل بهینه‌سازی سبد و دیگر مسائل مالی از جمله مدیریت ریسک و غیره که قابل بیان به فرم برنامه‌ریزی محدب درجه دوم باشند را با این مدل حل نماییم.

۲. مسائل مدیریت ریسک شامل تابع ارزش در معرض خطر و مسائل مدیریت بدهی و دارایی‌ها که حالت خاصی از مسائل برنامه‌ریزی تصادفی هستند را با کمک مدل‌های شبکه عصبی حل نمائیم.

۳. مسأله برنامه‌ریزی چند هدفه برای انتخاب سبد سرمایه که با روش بهینه‌سازی با قیود تصادفی قابل تبدیل به‌فرم برنامه‌ریزی مخروط مرتبه دوم است را با این مدل پیشنهادی حل نماییم.

۴. امید می‌رود که بتوان مسائل برنامه‌ریزی کسری با قیود کسری، مسائل برنامه‌ریزی کسری با ضرایب تصادفی، مسائل چندهدفه و مسائل مینیماکس کسری را با روش مذکور و یا با روشهای دیگر حل کرد.

۵. مسائلی دیگر از بهینه‌سازی مقاوم و مسائل مالی مثل مدیریت ریسک، قیمت‌گذاری اختیار معامله‌ها، آربیتراژ و غیره را که قابل بیان شدن به فرم برنامه‌ریزی مخروط مرتبه دوم محدب را حل نمود. امید می‌رود بتوان مسأله بهینه سازی مقاوم سبد مالی یک (چند دوره‌ای) با استفاده از ارزش در

معرض خطر که همتای مقاوم آنها منجر به برنامه‌ریزی مخروط درجه دوم شوند را با مدل شبکه عصبی پیشنهاد شده حل کرد.

۶. همچنین با کمک مدل‌های ارائه شده می‌توان مسائل حرکت روبات‌ها و تابع رگرسیون را نیز حل نمود.

مراجع و مآخذ

[1] M. S. Bazaraa, H. D. Sherali, C. M. Shetty, (1993), **"Nonlinear Programming-Theory and Algorithms"**, 2nd ed. New York: Wiley.

[2] B. N. Datta,(2003), **"Numerical Methods for Linear Control Systems"**, Elsevier Academic Press, San Diego, CA.

[3] G. Xing-Bao, (2003), "Exponential Stability of Globally Projected Dynamic Systems", **IEEE Trans.On Neural Networks**,14,2, 426-431.

[4] S. Effati, A. R. Nazemi, (2006), "Neural network models and its application for solving linear and quadratic programming problems", **Appl. Math. Comput**, 172, 305-331.

[5] G. Simmons, (2006), **"Differential Equations and Applications"**, McGraw-Hill Education (India).

[6] J. C. Hull, (2006), **"Options,Futures,And Other Derivatives"**, 6th Edition, Perntice Hall,New Jersey.

[7] C. P. Jones, (2001), **"Investments: Analysis and Management"**, New York: Wiley.

[8] M. J. A. Berry, G. Linoff, (1997), **"Data Mining Techniques"**, John wiley and sons, 450-460.

[9] B. Shelly, S. Stephenson, (2000), "The use of artificial neural network in completion simulation and design", **Computer and Geosciences**, 26, 941-951.

[10] S. Haykin, (1999), **"Neural Network: A Comprehensive Foundation"**, prentice hall, pp 842.

[11] D. L. Hudoson, M. E. Cohen, (1999), "Neural Network and Artificial Intellegence for Biomedicall Engineering", **IEEE Press Series on Biomedical Engineering**.

[12] J. J. Hopfield, D. W. Tank, (1986), "Simple neural optimization network:An A/D converter, signal decision circuit, and a linear programming pircuit", **IEEE Trans on Circuits and Systems**, 33, 533-541.

[13] M. P. Kennedy, L. O. Chua, (1988), "Neural networks for nonlinear programming", **IEEE Trans. on Neural Networks**, 35, 554-562.

[14] A. Bouzerdoum, T. R. pattison, (1993),"Neural network for quadratic optimization with bound constraints",**IEEE Trans. on Neural Networks**, 4, 293-304.

[15] Y. Xia, (1996), "A neural network for solving linear and quadratic programming problems", **IEEE Trans. on Neural Networks**,7, 1544-1547.

[16] Y. Xia, G. Feng, (2005), "An improved network for convex quadratic optimization with application to real-time beam forming", **Neurocomputing**, 64, 359-374.

[17] Y. Xia, G. Feng, J. Wang, (2004), "A recurrent neural network with exponential convergence for solving convex quadratic program and related linear piecewise equation", **Neural Networks**, 17, 1003-1015.

[18] Y. Xia, J. Wang, (2004), "A general projection neural network for solving monotone variational inequality and related optimization problems" , **IEEE Trans. On Neural Networks**, 15, 318-328.

[19] Y. Xia, H. Leung, J. Wang, (2002), "A projection neural network and its application to constrained optimization problems", **IEEE Transactions On Circuits And Systems**, 44, 4, 447-458.

[20] Y. Xia, J. Wang, (2000), "A recurrent neural network for solving linear projection equation", **Neural Networks**, 13, 337-350.

[21] X. Wu, Y. Xia, J. Li and W. Chen, (1996), "A High Performance Neural Network for solving Linear and Quadratic Programming Problems", **IEEE Trans. On Neural Networks**, 7, 3, 643-651.

[22] A. Rodriguez-Vazquez, R. Dommnguer-Castro, A. Rueda, J. L. Huertas and E. Sahnchez-Sinencio, (1990), "Nonlinear switched-capacitor neural network for optimization problems", **IEEE Trans on Circuits and Systems**, 384-390.

[23] L. Qi, J. Sun, (1993), "A nonsmooth version of Newton method, **Math Program**, 58, 353-368.

[24] S. Azarbek, A. R. Nazemi, D. SHahsavani, (2010), "An Application of Neural networks for solving linear and quadratic programming problems", **3th International Conference of Iranian Operations Research**.

[25] S. Effati, A. Ghomashi, A. R. Nazemi, (2007), "Application of projection neural network in solving convex programming problems", **Appl. Math. Comput**, 188, 1103-1114.

[26] W. F. Sharpe, (1964), "Capital Asset Price", **Journal of Finance**, 19, 3, 16-18.

[27] J. lintner, (1965), "The Valuation of Risk Asset and the selection of risky Investments in stock: Portfolio and Capital Budgets", **Reviwe of Economics and Statistics**, 47, 22-26.

[28] S.A. Ross, (1976), "The arbitrage theory of capital asset pricing", **Journal of Economic Theory**, 13, 341-360.

[29] J. Mossin, (1966), "Equilibrium in a capital asset market", **Econometrica**, 34(4), 14-16.

[30] F. Black, R. Litterman, (1990), "Asset allocation: combing investor views with market equilibrium", **Goldman Sachs Fixed Income Research**.

[31] H. Markowitz, (1952), "Portfolio Selection", **Journal of Finance**, 7, 77-91.

[32] H. Markowitz, (1959), "Portfolio selection: Efficient Diversification of investment", John Wiley and Sons, New York.

[33] S. M. Lee, A. J. Lerro, (1973), "Optimization the portfolio selection for Mutual Funds", **The Journal of finance**, 28, 1087-1099.

[34] S. M. Lee, D. L. Chesser, (1980), "Goal programming for portfolio", **The Journal of portfolio manegmant**, 22-26.

[35] R. Mansini, M.G. Speranza, (1997), " Heuristic Algorithms for the portfolio selection problem with Minimum Transaction Lots", **European Journal of Opeartional Resarch**, 114, 219-233.

[36] M.R. Young, (1998), "A minimax portfolio selection rule with linear programming solution", **Management Science**, 44, 673-683.

[37] Y. Xia, B. Liu, W. Shoayang, K. Lai, (1999), "A model for portfoilo selection with order of Expected Returns", **Computers and Opeartionals Resarch**, 27, 409-424.

[38] T. Chang, T. Meade, J. E. Beasley, Y. M. Sharaihan, (2000), " Heuristics for Cardinay Constrained portfoilo optimization", **Computers and Opeartionals Resarch**, 27, 1271-1302.

[39] J. Palmquist, S. Uryaser, P. Krokhmal, (1999), **"Portfolio optimization with conditional Value-at-Risk objective and constraints"**.

[40] S. M. Lobo, M. Fazel, S. Boyd, (2002), **"Portfolio optimization with Linear and Fixed Transaction Coasts"**, California Institue of Technology.

[41] C. Papahristodoulou, E. Dotzauer, (2004), "Optimal portfolios using linear programming models", **Journal of Operations Research Society**, 55, 1169-1177.

[42] R. Mansini, W. Oyryczak, M. G. Speranza, (2003), "LP solvable models for portfolio optimization: Aclassification and Computational comparison", **IMA Journal of management mathematics**, 14, 187-220.

[43] A. Fernandez, S. Gomez, (2007), "Portfolio selection using neural networks", **Computers and Operations Research**, 34, 1177–1191.

[44] D. R. Kincard, (1942), **"Numerical Analysis Mathimatics of Scientific Computing"**.

[45] H. Konno, (1988), **"Portfolio Optimization using L_1 Risk Function"**, Institute of Human and Social Sciences, IHSS Report Tokyo Institute of Technology, 88-89.

[46] H. Konno, H. Yamazaki, (1991), "Mean-absolute deviation portfolio optimization model and its application to Tokyo Stock Market", **Management Science**, 37, 519-531.

[47] Y. Simaan, (1997), "Estimation Risk in Portfolio Selection:The Mean Variance Model Versus the Mean Absolute Deviation Model", **Management Science**, 43, 1437-1446.

[48] X. Q. Cai, K. L. Teo, X. Q. Yang, X. Y. Zhou, (2000), "Portfolio optimization under a minimax rule", **Management Science**, 46, 957-972.

[49] M. Yu, H. Inouez, J. Shi, **"Portfolio optimization problems with linear programming models"**, 14-16.

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

[51] O. Ogryczak, A. Ruszczynski, (1999), "From stochastic dominance mean-risk model: semideviation as risk measure", **European Journal of Operational Research**, 116, 33-50.

[52] P. C. Fishburn, (1977), "Mean-risk analysis with risk associated with below-target returns", **American Economics Review**, 67, 116-126.

[53] R. K. Miller, A. N. Michel, (1982), **"Ordinary Differential Equations"**, Academic Press, New York.

[54] H. Amann, (1990), **"Ordinary Differential Equations: An Introduction to Nonlinear Analysis**, Walter de Gruyter, Berlin.

[55] J. M. Vegas, P. J. Zufiria, (2004), "Generalized neural networks for spectral analysis: dynamics and Liapunov functions", **Neural Networks**, 17, 233–245.

[56] C. Papahristodoulou, E. Dotzauer, (2004), "Optimal portfolios using linear programming models", **Journal of Operations Research Society**, 55, 1169-1177.

[57] G. Cornuejols, R. Tutuncu, (2007), **Optimization Methods in finance**, New York, 140-144.

[58] E. Beale, (1955), "On minimizing a convex function subject to linear inequalities", **Journal of Royal Statistical Society**, Series B, 173–184.

[59] A. Charnes, W. Cooper, (1959), "Chance constrained programming", **Management Science**, 5, 73–79.

[60] G. Dantzig, (1955), "Linear programming under uncertainty", **Management Science**, 1, 197– 206.

[61] F. Murphy, S. Sen, A. L. Soyster, (1982), "Electric utility capacity expansion planning with uncertain load forecasts", **AIIE Transaction**, 14, 52–59.

[62] M. Dempster, G. Consigli, (1998), "The CALM stochastic programming model for dynamic assetliability management", **World Wide Asset and Liability Modelling**, Cambridge University Press, 464–500.

[63] C. Dert, (1995), **Asset liability management for pension funds: A multistage chance constrained programming approach**, Ph.D. thesis, Erasmus University, Rotterdam, The Netherlands.

[64] M. vander Vlerk, (2003), "Integrated chance constraints in an ALM model for pension funds", **Technical Report 03A21, University of Groningen, Research Institute SOM (Systems,Organisations and Management)**.

[65] S. Sen, R. Doverspike, S. Cosares, (1994), "Network planning with random demand", **Telecommunication Systems**, 3, 11–30.

[66] A. Tomasgard, S. Dye, S. Wallace, J. Audestad, L. Stougie, (1998), "Modelling aspects of distributed processing in telecommunication networks", **Annals of Operations Research**, 82, 161–184.

[67] S. MirHassani, C. Lucas, G. Mitra, E. Messina, C. Poojari, Computational solution of capacity planning models under uncertainty,Parallel Computing 26 (2000) 511–538.

[68] M. Dempster, N. H. Pedron, E. Medova, J. Scott, A. Sembos, (2000), "Planning logistic operations in the oil industry", **Journal of Operational Research Society**, 51, (11), 1271–1288.

[69] G. Eppen, R. Martin, L. Schrage, (1989), "A scenario approach to capacity planning", **Operational Research**, 37, (4), 517–525.

[70] J. R. Robinson, (1988), "Loaded questions: New approaches to utlity forecasting", **Energy Policy** , 16, (1), 58–68.

[71] S. Takriti, J. Birge, E. Long, (1996), "A stochastic model for the unit commitment problem", **IEEE Transactions on Power Systems**,11, 1497–1508.

[72] A. Kampas, B. White, (2003), "Probabilistic programming for nitrate pollution control: Comparing different probabilistic constraint approximations", **European Journal of Operational Research**, 147, (1), 217–228.

[73] G. Laporte, F. Louveaux, L. V. Hamme, (1994), "Exact solution to a location problem with stochastic demands", **Transportation Science**, 28, 95–103.

[74] S. E. Elmaghraby, H. Soewandi, M. J. Yao, (2001), "Chance-constrained programming in activity networks: A critical evaluation", **European Journal of Operational Research**, 131,(1), 440–458.

[75] M. Ierapetritou, J. Acevedo, E. Pistikopoulos, (1996), "An optimization approach for process engineering problems under uncertainty", **Computers and Chemical Engineering**, 20, 703–709.

[76] J. Smith, (1999), "Optimizing platform survivability using a chance constrained linear program", in: US Army Ground Vehicle Survivability Symposium. Survivability/lethality Analysis Directorate, Munitions and Platform Branch, White Sands Missile Range, NM.

[77] A. Prekopa, (1990), "Dual method for a one-stage stochastic programming problem with random RHS obeying a discrete probability distribution", **Zeitscrift fur Operations Research**, 34, 441–461.

[78] A. Prekopa, (1993), "Programming under probabilistic constraint and maximizing a probability under constraints". center for Operations Research, Rutgers university, P.O.Box 5062, New Brunswick, New Jersey, 35-93.

[79] F. Hillier, (1967), "Chance-constrained programming with 0–1 or bounded continuous decision variables", **Management Science**,14, 34–57.

[80] Y. Seppala, (1971), "Constructing sets of uniformly tighter linear approximations for a chanceconstraint", **Management Science**, 17, 736–749.

[81] Y. Seppala, T. Orpana, (1984), "Experimental study on the efficiency and accuracy of a chanceconstrained programming algorithm", **European Journal of Operational Research**, 16, 345–357.

[82] D. Olson, S. Swenseth, (1987), "A linear approximation for chance constrained programming", **Journal of Operational Research Society**, 38, 261–267.

[83] M. Yahia, A. Daneshmand, (1995), "A linear approximation method for solving a special class of the chance constrained programming problem", **European Journal of Operational Research**, 80, 213–225.

[84] A. Weintraub, J. Vera, (1991), "A cutting plane approach for chance constrained linear program", **Operations Research**, 39, (5), 770–776.

[85] D. Dentcheva, A. Prekopa, A. Ruszczynski, (2000), "Concavity and efficient points of discrete distributions in probabilistic programming", **Mathematical Programming**, 89, 235–263.

[86] D. Dentcheva, A. Prekopa, A. Ruszczynski, (2001), "On convex probabilistic programming with discrete distributions", **Nonlinear Analysis**, 47, 1997–2009.

[87] P. Beraldi, A. Ruszczynski, (2002), "A branch and bound method for stochastic integer problems under probabilistic constraints", **Optimization Methods and Software**, 359–382.

[88] A. Ruszczynski, (2002), "Probabilistic programs with discrete distributions and precedence constrained knapsack polyhedra", **Mathematical Programming**, 93, 195–215.

[89] R. Aringheri, (2004), "Lecture Notes in Computer Science, Solving Chance-Constrained Program Combining Tabu Search and Simulation", Springer-Verlag, Berlin, Heidelberg, 30–41.

[90] G. Calafiore, M. Campi, (2005), "Uncertain convex programs: randomized solutions and confidence levels", **Mathematical Programming**, 102, 1, 25–46.

[91] D. de Farias, B. V. Roy, (2004), "On constraint sampling for the linear programming approach to approximate dynamic programming", **Mathematics of Operations Research**, 29, (3), 462–478.

[92] E. Erdogan, G. Iyengar, (2004), "Ambiguous chance constrained problems and robust optimization", CORC Technical Report TR-2004-10,IEOR Department, Columbia University.

[93] K. Iwamura, B. Liu, (1996), "A genetic algorithm for chance constrained programming", **Journal of Information and Optimization Sciences**, 17, (2), 409-422.

[94] C. A. Poojari, B. Varghese,(2008), "Genetic Algorithm based technique for solving Chance Constrained Problems", **European Journal of Operational Research**, 185, 1128-1154.

[95] A. Rodriguez-Vazquez, R. Dominguez-Castro, A. Rueda, J. L. Huertas and E. S anchez- Sinencio, (1990), "Nonlinear switched-capacitor neural networks for optimization problems", **IEEE Transactions on Circuits and Systems**, 37, 384–397.

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

[97] T. L. Friesz, D. H. Bernstein, N. J. Mehta, R. L. Tobin, and S. Ganjlizadeh,(1994), "Day-to-day dynamic network disequilibria and idealized traveler information systems", **Operation Research**, 42, 1120-1136.

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

[99] X. Gao, (2004), "A novel neural network for nonlinear convex programming", **IEEE Transactions on Neural Networks**, 15, 613–621.

[100] C. H. Ko, J. S. Chen, C. Y. Yang, (2011), "Recurrent neural networks for solving second-order cone programs", **Neurocomputing**, 74, 3646-3653.

[101] Y. Leung, K. Chen, X. Gao, (2003), "A high-performance feedback neural network for solving convex nonlinear programming problems", **IEEE Transantions on Neural Networks**, 14, 1469–1477.

[102] X. B. Liang, J.Wang, (2000), "A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bounded constraints", **IEEE Transactions on Neural Networks**, 11, 1251–1262.

[103] L. Z. Liao, H. D. Qi, (1999), "A neural network for the linear complementarity problem", **Mathematical and Computer Modeling**, 29, (3), 9-18.

[104] W. E. Lillo, M. H. Loh, S. Hui and S. H. Zak, (1993), "On solving constrained optimization problems with neural networks: A penalty method approach", **IEEE Trans. Neural Networks**, 4, 931-939.

[105] A. Malek, N. Hosseinipour-Mahani, S. Ezazipour, (2010), "Efficient recurrent neural network model for the solution of general nonlinear optimization problems", **Optimization Methods and Software**, 25, 1–18.

[106] X. Mu, S. Liu, Y. Zhang, (2005), "A neural network algorithm for second-order conic programming", in: Proceedings of the **Second International Symposiumon Neural Networks**, Chongqing, China, PartII, 718724.

[107] A. R. Nazemi, (2012), "A dynamic system model for solving convex nonlinear optimization problems", **Communications in Nonlinear Science and Numerical Simulation**, 17, 1696–1705.

[108] A. R. Nazemi, (2011), "A dynamical model for solving degenerate quadratic minimax problems with constraints", **Journal of Computational and Applied Mathematics**, 236, 1282-1295.

[109] A. R. Nazemi, F. Omidi, (2012), "A capable neural network model for solving the maximum flow problem", **Journal of Computational and Applied Mathematics**,236, 14, 3498-3513.

[110] A. R. Nazemi, F. Omidi, (2012), "An efficient dynamic model for solving the shortest path problem", **Transportation Research Part C: Emerging Technologies** ,(In press).

[111] A. R. Nazemi, E. Sharifi, (2012), "Solving a class of geometric programming problems by an efficient dynamic model", **Communications in Nonlinear Science and Numerical Simulation**,(In press).

[112] J. Sun, J. S. Chen, C. H. Ko, "Neural networks for solving second-order cone constrained variational inequality problem", **Computational Optimization and Applications**, (First online).

[113] J. Sun, L. Zhang, (2009), "A globally convergent method based on Fischer–Burmeister operators for solving second-order cone constrained variational inequality problems", **Computers and Mathematics with Applications**, 58, 1936–1946.

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

[115] J. Wang, Q. Hu, D. Jiang,(1999), "A Lagrangian neural network for kinematic control of redundant robot manipulators", **IEEE Transactions on Neural Networks**, 10, (5), 1123-1132.

[116] Y. Xia, J. Wang, (2005), "A recurrent neural network for solving nonlinear convex programs subject to linear constraints", **IEEE Transactions on Neural Networks**, 16, 379-386.

[117] Y. Xia, J. Wang, (2004), "A recurrent neural network for nonlinear convex optimization subject to nonlinear inequality constraints", **IEEE Transactions on Circuits and Systems I:Regular Papers 51**, (7), 1385-1394.

[118] Y. Xia, G. Feng, (2007), "A new neural network for solving nonlinear projection equations", **Neural Networks**, 20, 577-589.

[119] Y. Xia, J. Wang, L. M. Fok, (2004), "Grasping-force optimization formultifingered robotic hands using a recurrent neural network", **IEEE Transactions on Robotics and Automation**, 20, (3), 549-554.

[120] X. Xue, W. Bian, (2007), "A project neural network for solving degenerate convex quadratic program", **Neural Networks**, 70, 2449-2459.

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

[122] Y. Yanga, J. Cao, (2010), "The optimization technique for solving a class of non-differentiable programming based on neural network method", **Nonlinear Analysis**,11, 1108–1114.

[123] Y. Yang, X. Xu, (2007), "The projection neural network for solving convex nonlinear programming, D. S. Huang, L. Heutte, and M. Loog (Eds.): ICIC 2007, LNAI 4682, Springer-Verlag Berlin Heidelberg, 174-181.

[124] Y. Yang, J. Cao, (2008), "A feedback neural network for solving convex constraint optimization problems", **Applied Mathematics and Computation**, 201, 340-350.

[125] M. S. Lobo, L. Vandenberghe, S. Boyd, H. Lebret, (1998), "Applications of second-Order cone programming", **Linear Algebra and its Applications**, 284, 193–228.

[126] H. Y. Benson, R. J. Vanderbei (2003), "Solving Problems with Semidefinite and Related Constraints Using Interior-Point Methods for Nonlinear Programming", **Mathematical Programming Series B**, 95, 279–302.

[127] E. Chinlar,(1975), **"Interoduction to stochactic processes"**, Northwestern university.

[128] S. Effati, M. M. Moghadam, (2006), "Conversion of some classes of fractional programming to second-order cone programming and solving it by potential reduction interior point method", **Applied Mathematics and Computation**, 181, 563–578.

[129] B.D. Craven, (1988), "Fractional Programming", **Sigma Series in Applied Mathematics**, 4, Helderman Verlag, Berlin, Germany, 144-146.

[130] S. Schaible, T. Ibaraki, (1983), "Fractional programming", **European Journal of Operational Research**, 12, 325–338.

[131] S. Schaible, (1995), **"Fractional programming"**, Handbook of Global Optimization, Kluwer Academic Publishers, Dordrecht, Netherlands, 495–608.

[132] I.M. Stancu-Minasian, (1997), **"Fractional Programming: Theory Methods and Applications"**, Kluwer Academic Publishers, Dordrecht, Netherlands.

[133] I.M. Stancu-Minasian,(1980), "Applications of the fractional programming, theory, methods and applications", **Economic Comput. Economic Cyber. Studies Res**,14, 69–86.

[134] I.M. Stancu-Minasian,(1999), " A fifth bibliography of fractional programming", **Optimization**, 45, 343–367.

[135] I.M. Stancu-Minasian,(2006), " A sixth bibliography of fractional programming", **Optimization**, 55, 405–428.

[136] G. B. Dantzing, (1955), **"Linear Programming under Uncertaintly"**, Magmt. Sci, 197-206.

[137] R. E. Bellman, (1957), **"Daynamic Programming"**, Princeton University Press.

[138] A. Ben-Tal, A. Nemirovski, (1998), "Robust convex optimization", **Mathematics of Operations Research**, 23, 769-805.

[139] A. Ben-Tal, A. Nemirovski, (1995), "Robust solution to uncertain linear problems", Report No 6/95, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion- the Israel Institute of Technology, Technion City, Haifa 32000, Israel.

[140] A. Ben-Tal, L. EL Ghaoui, A. Nemirovski, (2009), **"Robust optimization"**, Princeton Unversity, Princeton Series in Applied Mathematics.

[141] A. Ben-Tal, A. Nemirovski, (1999), "Robust solutions to uncertain programs", **Operations Research Letters**, 25, 1-13.

[142] A. Ben-Tal, A. Nemirovski, (2000), "Robust solutions of linear programming problems contaminated with uncertain data", **Mathimatical Programming**, 88, 411-424.

[143] A. Ben-Tal, L. EL Ghaoui, A. Nemirovski, (2000), **"Robust semidefinite programming"**, Handbook on Semidefinite Programming and application, Kluwer Academic Publishers, 139-162.

[144] A. Ben-Tal, A. Nemirovski, C. Roos, (2002), "Robust solutions to uncertain quadratic and conic-quadratic problems", **SIAM Journal on Optimization**, 13, 535-560.

[145] A. Ben-Tal, A. Goryashko, E. Guslitzer, A. Nemirovski, (2004), "Adjustable robust solutions of uncertain linear programs", **Mathimatical Programming**, 99, 351-376.

[146] L. El-Ghaoui, H. Lebret, (1997), "Robust solutions to least-squares problems with uncertain data matrices", **SIAM Journal on Matrix Analysis Application**, 18, 1035-1064.

[147] G. Cornuejols, R. H. Tutuncu, (2007), **"Optimizatiom method in finance"**, 295-301.

[148] A. Ben-Tal, T. Margalit, A. N. Nemirovski, (2002), "Robust modeling of multi-stage portfolio problems", In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, High Performance OptimizatioNn, 303–328.

[149] D. Goldfarb, G. Iyengar, (2003), "Robust portfolio selection problems", **Mathematics of Operations Research**, 28, 1-38.

[150] V. S. Bawa, S. J. Brown, and R.W. Klein,(1976), **"Estimation Risk and Optimal Portfolio Choice"**, (Amsterdam: North-Holland).

[151] R. O. Michaud, (1989), "he Markowitz optimization enigma: is optimized optimal", **Financial Analysts Journal**, 45, 31-42.

[152] R. O. Michaud,(1998), **"Efficient Asset Management"**, (Boston, MA: Harvard Business School Press).

[153] B. Halldorsson, R. H. Tutuncu, (2003), "An interior-point method for a class of saddle point problems", **Journal of Optimization Theory and Applications**, 116, 3, 559-590.

[154] S. Ceria, R. Stubbs, (2006), "Incorporating estimation errors into portfolio selection. Robust portfolio selection", **Journal of Asset Management**.

واژه‌نامه انگلیسی به فارسی

انحراف مطلق . . . . . Absolute Divation
تابع تحریک . . . . . Activation Function
تطبیق پذیر . . . . . Adaptive
قابل پیش بینی . . . . . Anticipative
دارایی . . . . . Asset
اریبی . . . . . Bias
محک . . . . . Benchmark
بازده ناشی از تغییر قیمت . . . . . Capital Gain
قیود شانس . . . . . Chance Constrained
تابع ترکیب . . . . . Combination Function
مکمل . . . . . Complementarity
محدب . . . . . Convex
پوسته محدب . . . . . Convex Hull
همبستگی . . . . . Corrolation
ریسک نامطلوب . . . . . Downside Risk
مشتقات . . . . . Dreivatives
پویا . . . . . Dynamic
مرز کارایی . . . . . Efficent Frontier
نقطه تعادل . . . . . Equilibrium Point
بازده مورد انتظار . . . . . Expected Return
برنامه‌ریزی کسری . . . . . Fractional Programming
پایداری مجانبی سراسری . . . . . Globally Asymptotically Stable
همگرای سراسری . . . . . Global Convergance
کمینه سراسری . . . . . Global Minimum
برنامه‌ریزی آرمانی . . . . . Goal Programming
مجموعه پایدار . . . . . Invariant Set
کمینه موضعی . . . . . Local Minimum
سلول عصبی . . . . . Neuron
مقادیر اسمی . . . . . Nominal Values
چند سقفی . . . . . Polytopic

سبد سرمایه . . . . . Portfolio
بازده . . . . . Return
سطح پشیمانی . . . . . Regret Level
بازده تحقق یافته . . . . . Realized Return
ریسک گریزی . . . . . Risk-Averse
مقاوم(استوار) . . . . . Robust
همتای مقاوم . . . . . Robust Conterpart
مقاوم بودن . . . . . Robustness
مخروط مرتبه دوم . . . . . Second Order Cone
نیمه معین . . . . . Semi Definite
نیم واریانس . . . . . Semi Variance
تحلیل حساسیت . . . . . Sensitivity Analysis
سناریو(طرح) . . . . . Scenario
فروش استقراضی . . . . . Short-Selling
هموار . . . . . Smooth
تصادفی(احتمالی) . . . . . Stochastic
اکیدا یکنوا . . . . . Strictly Monotonic
قابل تحمل . . . . . Tolerable
خطای پیگردی((ردیابی) . . . . . Tracking Error
تابع انتقال . . . . . Transfor Function
ارزش در معرض ریسک . . . . . Value-at-Risk
عدم‌قطعیت . . . . . Uncertainty
مطلوبیت . . . . . Utility
بازده ناشی از دریافت سود . . . . . yeild

Abstract

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

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

In this thesis, we tried two different models of recursive neural network is presented to solve optimization problems in the traces. Analysis of uniqueness, stability and convergence of global solutions are examined and the performance of the proposed method using several examples of stochastic programming problems, Fractional programming, optimization and robust portfolio optimization is shown.

Finally, we provide conclusions and recommendations for future work.

Keywords: *Neural networks, stability, the mean-variance, portfolio selection problem, second order cone programming, stochastic programming, Fractional programming, Robust optimization*

Shahrood University of Technology

Facualty of Methematical

Department of Mathematics

M.S Thesis

An application of neural networks models for solving portfolio optimization problem

By:

Narges Tahmasbi

Supervisor:
P.hD. Alireza Nazemi

Date
Sep 2012