جلسه دفاع از پایان‌نامه کارشناسی ارشد

یک روش گام‌وفقی اولیه‌دوگان برای حل مساله‌های بهینه‌سازی خطی

سخنران: علی هادیان

زمان: ساعت

مکان:

هیئت داوران

۱ – دکتر محمدرضا پیغامی

۲ – دکتر محمد حسن بیژن زاده

۳ – ()

۴ –

چکیده:

در حل مساله‌های بهینه‌سازی خطی به روش نقطه‌درونی، تابع‌های نزدیکی نقش مهمی را در کاهش تعداد تکرارها ایفا می‌کنند. در این پایان‌نامه تابع‌های نزدیکی خودمنظم را معرفی می‌کنیم. سپس به تحلیل الگوریتم‌های نقطه‌درونی مبتنی بر تابع‌های خودمنظم می‌پردازیم. در ادامه یک تابع غیرخودمنظم را درنظر گرفته، و یک الگوریتم گام‌وفقی مبتنی بر این تابع را ارائه می‌نماییم. هم‌چنین پیچیدگی تکرار الگوریتم را بااستفاده از این تابع به‌دست می‌آوریم.

در پایان، در یک بررسی جامع طرح کلی برای تحلیل الگوریتم‌های نقطه‌درونی اولیه‌دوگان مبتنی بر تابع‌های هسته ارائه می‌گردد و سپس چند تابع هسته با جمله مانعی نمایی را بررسی کرده، بهترین پیچیدگی تکرار را می‌یابیم.

کلمات کلیدی: نقطه‌درونی، گام‌وفقی، به‌هنگام بلند، به‌هنگام کوتاه، خودمنظم، تابع نزدیکی، پیچیدگی.

فهرست مطالب

لیست جدول‌ها

لیست شکل‌ها

پیشگفتار

یک مساله برنامه ریزی خطی به فرم $\min\{c^T x : Ax = b,\ x \geq ۰\}$ را درنظر بگیرید. تاکنون روش‌های زیادی برای حل چنین مساله‌هایی ارائه گردیده که روش سیمپلکس پیشتاز آن‌ها بوده‌است. در سال ۱۹۷۴ نشان داده شد که این روش می‌تواند برای برخی ازمساله‌ها دارای پیچیدگی ازمرتبه $\mathcal{O}(۲^n)$ باشد. لذا ارائه الگوریتم‌های بهتر برای دستیابی به جواب بهینه با پیچیدگی تکرار چند جمله‌ای مورد بررسی قرار گرفت. یکی از روش‌های حل این‌گونه از مساله‌ها، روش‌های نقطه‌درونی است که براساس حل مساله در جهت رسیدن به جواب‌های اولیه و دوگان و کاهش فاصله دوگانی استوار هستند. این روش‌ها با شروع از یک جواب اولیه، به سمت همسایگی معینی از جواب مساله‌های اولیه و دوگان حرکت می‌کنند، تازمانی‌که تقریب خوبی از جواب بهینه به‌دست آید. یکی از مفاهیمی که در این روش‌ها رایج بوده و در کاهش پیچیدگی بسیار موثر است، توابعی است که برای معرفی همسایگی ارائه می‌شوند. در این پایان‌نامه قصد داریم همسایگی‌های مبتنی بر تابع‌های خودمنظم را معرفی کرده و نشان دهیم که روش‌های نقطه‌درونی بااستفاده از تابع‌های خودمنظم دارای مرتبه پیچیدگی $\mathcal{O}(\sqrt{n}\log n \log \frac{n}{\varepsilon})$ برای روش‌های به‌هنگام بلند می‌باشند. در ادامه تحلیل پیچیدگی روش نقطه‌درونی بر اساس یک تابع غیر خودمنظم ارائه می‌شود. نشان خواهیم داد که حل مساله به روش نقطه‌درونی مبتنی بر این تابع دارای پیچیدگی تکرار از مرتبه $\mathcal{O}(q\sqrt{n\tau}\log \frac{n}{\varepsilon})$ برای روش‌های به‌هنگام بلند می‌باشد. در انتها طرح کلی برای تحلیل پیچیدگی تکرار الگوریتم‌های نقطه‌درونی مبتنی بر تابع‌های هسته ارائه شده و توسط آن، یک تابع با جمله رشد نمایی را بررسی می‌نماییم که مرتبه پیچیدگی تکرار آن $\mathcal{O}(\sqrt{n}\log \frac{n}{\varepsilon})$ می‌باشد، این مرتبه بهترین مرتبه پیچیدگی تکرار ارائه شده برای روش‌های به‌هنگام بلند است که تاکنون ارائه شده است.

این پایان‌نامه در ۵ فصل تنظیم شده‌است. در فصل اول مفاهیم و نتایج مقدماتی مورد نیاز آمده‌است. در فصل دوم ویژگی‌های عمومی تابع‌های هسته را بررسی نموده، چند دسته از تابع‌های نزدیکی را معرفی می‌کنیم. در فصل سوم به معرفی تابع‌های هسته خودمنظم پرداخته و الگوریتم نقطه‌درونی مبتنی بر این تابع‌ها را تحلیل و پیچیدگی محاسباتی این الگوریتم‌ها را به‌دست می‌آوریم. در فصل چهارم یک روش گام‌وفقی اولیه‌دوگان برای حل مساله‌های بهینه‌سازی خطی ارائه گردیده‌است. در فصل پنجم طرح کلی برای تحلیل پیچیدگی محاسباتی الگوریتم‌های نقطه درونی مبتنی بر تابع‌های هسته ارائه و چند تابع هسته با استفاده از این طرح تحلیل گردیده است.

فصل ۱

مفاهیم و نتایج مقدماتی

در این فصل مفاهیم مقدماتی مورد نیاز در این پایان نامه معرفی و برخی از قضیه‌ها و نتیجه‌ها بیان خواهد شد. در ابتدا مفاهیم کلی، سپس مفاهیم اولیه در بهینه‌سازی خطی ارائه خواهد شد. در ادامه روش اولیه‌دوگان نیوتن برای بهینه سازی خطی بیان می‌گردد.

۱—۱ مفاهیم اولیه و تعاریف پایه

در این بخش تعاریف و قضیه‌های مقدماتی مورد نیاز در این پایان‌نامه را بیان می‌کنیم.

۱—۱.۱ ماتریس معین مثبت(منفی)[۱] : فضای همه ماتریس‌های مربعی $n \times n$ را با نماد $\mathcal{R}^{n\times n}$ نمایش می‌دهیم. ماتریس $A \in \mathcal{R}^{n\times n}$ را معین مثبت(منفی) گوییم اگر A متقارن بوده و همه مقادیر ویژه[۲] آن مثبت(منفی) باشد. عبارت‌های زیر درمورد هرماتریس متقارن A هم ارزاست:

(i) A معین مثبت(منفی) است

(ii) $A = C^T C$ برای بعضی از ماتریس‌های نامنفرد[۳] C

(iii) $x^T A x > 0$ $(x^T A x < 0)$ برای هربردار غیرصفر x

ماتریس $A \in \mathcal{R}^{n\times n}$ را نیمه معین[۴] مثبت گوییم هرگاه A متقارن و مقادیر ویژه آن غیرمنفی باشند. عبارت‌های زیر درمورد هرماتریس متقارن A هم ارزاست:

(i) A نیمه معین مثبت است

(ii) $A = C^T C$ برای بعضی از ماتریس‌های C

(iii) $x^T A x \geq 0$ برای هربردار x

---

۱ Positive definite, Negative definite

۲ Eigenvalue

۳ Nonsingular

۴ Semi-Definite

۱—۱.۲ ماتریس هسیان[۱] : فرض کنید f تابعی معلوم، n متغیره و دوبار مشتق‌پذیر باشد، در این‌صورت ماتریس هسیان تابع به‌صورت زیر تعریف می‌شود:

$$H_{n\times n} = \left[\frac{\partial^۲ f}{\partial x_i \partial x_j}\right]$$

شرط کافی برای آن‌که تابع f در نقطه ایستای x° ماکزیمم(مینیمم) باشد آن‌است که ماتریس هسیان f در آن نقطه معین مثبت (منفی) باشد.

۱—۱.۳ نرم بردارها و ماتریس‌ها: برای بردار n تایی $x = (x_۱, x_۲, \ldots, x_n)$، x_i رادرایه i ام x می‌نامیم. معمولا x یک بردار ستونی وترانهاده آن x^T را بردار سطری درنظر می‌گیریم. اگر همه درایه‌های x صفر باشند آن را بردار صفر نامیده و با نماد $x = \circ$ نشان می‌دهیم. اگر همه درایه‌های بردار برابر یک باشد آن را با نماد e نشان می‌دهیم. ضرب داخلی دوبردار x و $s \in \mathcal{R}^{n\times n}$ برابراست با

$$x^T s = \sum_{i=۱}^{n} x_i s_i$$

یک نرم (یا نرم برداری) در $\mathcal{R}^n$ تابعی است که به هر $x \in \mathcal{R}^n$ عدد غیرمنفی $\|x\|$ نظیر می‌کند به‌طوری که برای هر $x, s \in \mathcal{R}^n$ و $\alpha \in \mathcal{R}$:

$$\|x\| > \circ \qquad \text{اگر } x \neq \circ،$$
$$\|\alpha x\| = |\alpha| \|x\|$$
$$\|x + s\| \leq \|x\| + \|s\|$$

نرم اقلیدسی برابراست با

$$\|x\|_۲ = \sqrt{\sum_{i=۱}^{n} x_i^۲}$$

زمانی‌که منظور از نرم، نرم خاصی نباشد، $\|x\|$ را نرم اقلیدسی درنظر می‌گیریم. نامساوی کشی—شوارتز[۲] برای هر $x, s \in \mathcal{R}^n$ برقرار است، یعنی

$$x^T s \leq \|x\| \|s\|$$

وتساوی برقرار است اگر و فقط اگر x و s وابسته خطی باشند.

برای هر عدد مثبت p، p - نرم به‌صورت

$$\|x\|_p = \left(\sum_{i=۱}^{n} |x_i|^p\right)^{\frac{۱}{p}}$$

---
۱ Hessian
۲ Cauchy-Schwarz

تعریف می‌شود. اگر $p = \infty$، آن‌گاه نرم بی‌نهایت یک بردار به‌صورت زیر می‌باشد:

$$\|x\|_\infty = \lim_{p\to\infty} \|x\|_p = \max_{1\le i\le n} |x_i|$$

برای هر ماتریس معین مثبت $A_{n\times n}$ نرم برداری $\|.\|_A$ را تعریف می‌کنیم: $\|x\|_A = \sqrt{x^TAx}$

۴.۱—۱ ضرب و نامساوی هادامارد[1] : ضرب هادامارد دو بردار $x, s \in \mathcal{R}^n$ عبارت است از بردار xs که درایه iام آن برابر x_is_i می‌باشد یعنی

$$xs = (x_1s_1, x_2s_2, \ldots, x_ns_n)$$

و بنابراین ضرب داخلی دو بردار توسط ضرب هادامارد برابراست با: $x^Ts = e^T(xs)$
دترمینان ماتریس $A_{n\times n}$ با ستون‌های $a_1, a_2, \ldots, a_n$ برابراست باحجم متوازی السطوح ساخته شده توسط $a_1, a_2, \ldots, a_n$. با بسط دترمینان داریم:

$$det(A) \le \|a_1\|_2\|a_2\|_2 \ldots \|a_n\|_2$$

این نامساوی را نامساوی هادامارد می‌نامند.

۵.۱—۱ میانگین: فرض کنید $\{x_1, x_2, \cdots, x_n\}$ مجموعه ای از n عدد حقیقی باشد، دراین‌صورت:

i. میانگین حسابی[2] این اعداد عبارت‌است از

$$A = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

ii. میانگین هندسی[3] این اعداد از رابطه زیر به‌دست می آید

$$G = \sqrt[n]{x_1x_2\cdots x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

iii. میانگین هارمونیک[4] این اعداد به‌صورت زیر است

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

بنابراین با یک محاسبه ساده رابطه سه میانگین فوق به‌صورت زیر خواهد بود[۱۱].

$$G^n = AH$$

لازم به ذکر است، میانگین یک مجموعه می‌تواند در مجموعه نباشد.

---

[1] Hadamard
[2] Arithmatic mean
[3] Geometric mean
[4] Harmonic mean

تعریف ۱—۶.۱ تابع $f : \mathcal{R}^n \to \mathcal{R}$ را یک تابع محدب[1] نامیم هرگاه به‌ازای هر دو بردار n تایی x_1 و x_2 داشته باشیم:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \quad \forall \lambda \in [0, 1]$$

درصورتی‌که داشته باشیم:

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2) \quad \forall \lambda \in [0, 1]$$

آن‌گاه f را مقعر[2] گوییم. به‌عبارت دیگر f مقعر است هرگاه $-f$ محدب باشد.

یادآوری می‌شود که در حساب دیفرانسیل مقدماتی ثابت شده‌است که اگر تابع f دوبار مشتق پذیر و پیوسته باشد، آن‌گاه شرط $f'' > 0$ تحدب و شرط $f'' < 0$ تقعر تابع f در یک بازه خاص را نتیجه می‌دهد. تعریف دیگر تابع‌های محدب را می‌توان به‌صورت لم زیر بیان کرد.

لم ۱—۷.۱ تابع $f : \mathcal{R} \to \mathcal{R}$ یک تابع محدب است، هرگاه به‌ازای هر $x_1, x_2 \in \mathcal{R}$، شرط $x_1 < x_2$ نتیجه دهد $f'(x_1) \leq f'(x_2)$. و درصورتی‌که $f'(x_1) \geq f'(x_2)$ آن‌گاه f مقعر است.

تعریف ۱—۸.۱ مخروط محدب[3] C، مجموعه‌ای است که:

i. محدب می‌باشد، یعنی برای هر $x_1, x_2 \in C$ و هر $\lambda \in [0, 1]$ داریم: $\lambda x_1 + (1-\lambda)x_2 \in C$.

ii. برای هر $x \in C$، داریم: $\lambda x \in C$.

۱—۹.۱ مرتبه اجرایی: فرض کنید f و g دوتابع از $\mathcal{R}_+$به $\mathcal{R}_+$ باشند در این‌صورت:

الف) گوییم f از مرتبه $\mathcal{O}(g)$ است هرگاه عدد مثبت c چنان موجود باشد که $f(x) \leq cg(x)$، برای هر $x > 0$. در این‌صورت می‌نویسیم: $f(x) = \mathcal{O}(g(x))$.

ب) گوییم f از مرتبه $\Omega(g)$ است هرگاه عدد مثبت c چنان موجود باشد که $f(x) \geq cg(x)$، برای هر $x > 0$. در این‌صورت می‌نویسیم: $f(x) = \Omega(g(x))$.

ج) گوییم f از مرتبه $\Theta(g)$ است هرگاه اعداد مثبت c_1 و c_2 چنان موجود باشند که $c_1 g(x) \leq f(x) \leq c_2 g(x)$، برای هر $x > 0$. در این‌صورت می‌نویسیم: $f(x) = \Theta(g(x))$.

مفهوم مرتبه اجرایی برای تحلیل الگوریتم‌ها به‌کار می‌رود. یک الگوریتم را کاراتر(بهتر) از الگوریتم دیگری گوییم هرگاه دارای مرتبه اجرایی $\mathcal{O}$ کمتری نسبت به دیگری باشد. به بیان ساده، الگوریتمی که دارای مرتبه اجرایی $\mathcal{O}$ کوچک‌تری باشد، سرعت رسیدن به جواب آن بیشتر است. در بسیاری از کاربردها، شرط $x \geq 0$ ممکن است اتفاق نیافتد و این رابطه‌ها برای $x > x_0$ برقرار باشند. در این‌صورت با تغییر

[1] Convex
[2] Concave
[3] Convex cone

متغیر $t = x - x_0 > 0$ به تعاریف فوق می‌رسید. بنابراین در بسیاری از منبع‌های تحلیل الگوریتم شرط وجودی را برای دو ثابت در تعاریف فوق درنظر می‌گیرند. به‌عنوان مثال، تعریف $\mathcal{O}(g)$ را به‌صورت زیر بیان می‌کنند:(بقیه تعاریف نیز مشابه است.)

گوییم f از مرتبه $\mathcal{O}(g)$ است هرگاه عدد ثابت c و x_0 چنان موجودباشند که $f(x) \leq cg(x)$ برای هر $x > x_0$.

شکل ۱–۱ مفهوم هندسی مرتبه اجرایی را بیان می‌کند.

لم ۱–۱۰.۱ [?]: برای تابع‌های f, g, h، داریم:

(i) $f = \Theta(g)$ اگر و فقط اگر $f = \mathcal{O}(g)$ و $f = \Omega(g)$.

(ii) اگر $f = \Theta(g)$ و $g = \Theta(h)$ آن‌گاه $f = \Theta(h)$.

(iii) $f = \Theta(g)$ اگر و فقط اگر $g = \Theta(f)$.

(iv) $f = \Omega(g)$ اگر و فقط اگر $g = \Omega(f)$.

(v) $f = \mathcal{O}(g)$ اگر و فقط اگر $g = \Omega(f)$

(vi) اگر $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \alpha$ آن‌گاه

$$\begin{cases} f = \Theta(g) & \text{اگر}\quad 0 < \alpha < \infty \text{،} \\ f = \mathcal{O}(g) \text{ و } f \neq \Theta(g) & \text{اگر}\quad \alpha = 0 \text{،} \\ f = \Omega(g) \text{ و } f \neq \Theta(g) & \text{اگر}\quad \alpha = \infty \text{،} \end{cases}$$

لم ۱–۱۱.۱ [?]: فرض کنید $f(x) = a_n x_n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ در این‌صورت $f = \mathcal{O}(x^n)$.

لم ۱–۱۲.۱ [?]: اگر $1 < a < b$ و $2 < j < k$ اعدا ثابتی باشند، آن‌گاه

$$\mathcal{O}(\log n) < \mathcal{O}(\sqrt{n}) < \mathcal{O}(n) < \mathcal{O}(n \log n) < \mathcal{O}(n^2) < \mathcal{O}(n^j) < \mathcal{O}(n^k) < \mathcal{O}(a^n) < \mathcal{O}(b^n) < \mathcal{O}(n!)$$

در حقیقت این لم بیانگر این‌است که با رشد تعداد داده‌های یک مساله، یعنی n، تعداد دستورهای لازم برای حل مساله، با چه سرعتی رشد می‌کند.

لم ۱–۱۳.۱ [?]:

i. اکر $a, b > 1$ آن‌گاه $\log_a n = \Theta(\log_b n)$

یعنی پیچیدگی همه تابع‌های لگاریتمی در یک دسته پیچیدگی قرار دارند.

ii. اگر $0 < a < b$ آن‌گاه $a^n = \mathcal{O}(b^n)$

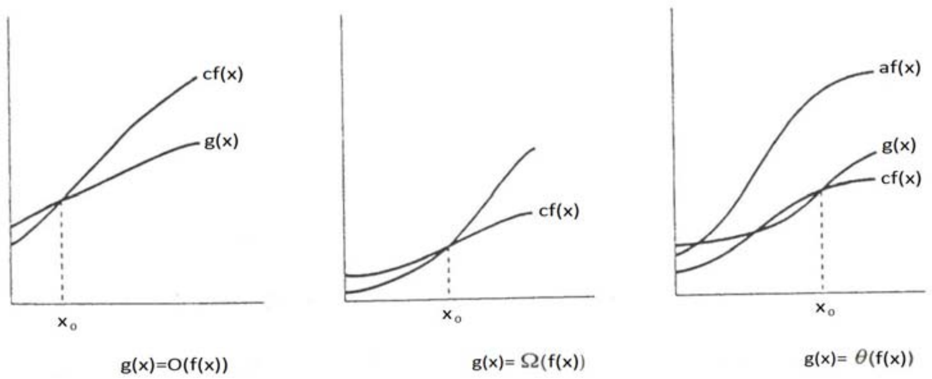
یعنی همه تابع‌های با پیچیدگی نمایی در یک دسته پیچیدگی قرار دارند.

iii. به‌ازای هر $a > 0$ داریم $a^n = \mathcal{O}(n!)$

یعنی $n!$ از هر تابع پیچیدگی نمایی بدتر است.

iv. اگر $d, c > 0$ و $f = \mathcal{O}(g)$ و $h = \Theta(g)$ آن‌گاه $cf + dh = \Theta(g)$.

در بسیاری از مساله‌ها، ارائه الگوریتم برای رسیدن به جواب اجتناب ناپذیر است و شاید الگوریتم‌های مختلفی برای حل یک مساله خاص ارائه شود. بنابراین، بررسی این‌که کدام الگوریتم از دیگر الگوریتم‌ها بهتر است، امری ضروری می‌باشد. از این موضوع با عنوان تحلیل الگوریتم یاد می‌شود.
به‌طور کلی الگوریتم‌ها را به دو روش تحلیل می‌کنند. روش اول، تحلیل الگوریتم بر حسب میزان حافظه مورد نیاز برای الگوریتم و روش دوم، تحلیل الگوریتم براساس سرعت رسیدن به جواب. در حالت کلی، تحلیل الگوریتم به روش اول از اهمیت کم‌تری نسبت به روش دوم برخوردار است.



شکل ۱—۱: نمودار مفهوم مرتبه اجرایی

در تحلیل الگوریتم‌ها به روش دوم، برای بررسی این‌که کدام الگوریتم سریع‌تر به جواب هم‌گرا می‌شود، تعداد دستورهای کلیدی الگوریتم را شمارش می‌کنند، که برای n داده اطلاعاتی، تابعی بر حسب n (برای مثال $f(n)$) خواهد بود. تابع به‌دست آمده را تابع زمانی الگوریتم می‌نامند. تابع زمانی یک الگوریتم بیان می‌کند که برای محاسبه جواب با این الگوریتم بایستی $f(n)$ عمل صورت پذیرد. این تابع با شمارش تعداد تکرارها (در الگوریتم‌های تکراری) و یا تعداد فراخوانی‌ها (در الگوریتم‌های بازگشتی) به‌دست می‌آید. از این تابع با عنوان تابع پیچیدگی محاسباتی[۱] نیز یاد می‌شود. هر اندازه این تابع به چندجمله‌ای نزدیک‌تر باشد تحلیل آن راحت‌تر خواهد بود. الگوریتمی که تابع $f(n)$ چندجمله‌ای باشد را الگوریتم با پیچیدگی چندجمله‌ای یا مرتبه اجرایی چندجمله‌ای گویند. در عمل همیشه این تابع یک چندجمله‌ای نیست بلکه حتی می‌تواند از درجه نمایی باشد. لذا یافتن الگوریتم‌ها با تابع پیچیدگی چندجمله‌ای همیشه مدنظر متخصصان علوم مهندسی و پایه بوده‌است.

---
۱ Complexity function

نکته دیگری که مهم است، این است که منظور از دستورهای کلیدی، دستورهایی است که حل مساله مبتنی بر آن است و بستگی به الگوریتم ندارد. یعنی برای رسیدن به جواب با هر الگوریتم، ناگزیر به اجرای این دستورها خواهیم بود، لذا تحلیل الگوریتم می‌تواند بر اساس تعداد کمتر این دستورها باشد. به‌عنوان مثال، برای محاسبه وارون یک ماتریس مربعی مجبور به محاسبه دترمینان خواهیم بود(در روش‌های سطری مقدماتی نیز به محاسبه وارون می‌پردازیم!). لذا می‌توان دستور کلیدی را محاسبه دترمینان قرار دهیم و تحلیل الگوریتم‌های مختلف را بر این اساس قرار دهیم که: کدام یک بامحاسبه تعداد کمتری وارون به جواب می‌رسد.

به‌عنوان مثالی دیگر، در روش‌های نقطه‌درونی که در این پایان‌نامه به تحلیل برخی از الگوریتم‌های آن خواهیم پرداخت، خواهیم دید که همه این الگوریتم‌ها باید یک دستگاه نیوتن را حل کنند. لذا دستور کلیدی که مبنای تحلیل الگوریتم‌های این روش است، حل یک دستگاه نیوتن می‌باشد. در این پایان‌نامه با این مفهوم بیشتر آشنا خواهیم شد.

نمادگذاری: در این پایان‌نامه از I برای ماتریس همانی استفاده می‌کنیم که بُعد آن وابسته به مفهوم مورد بحث می‌باشد. I_n ماتریس همانی $n \times n$ است که قطر اصلی آن یک وبقیه درایه‌ها صفر می‌باشد. به همین صورت، نماد $\circ_{n\times n}$ ماتریسی است که همه درایه‌های آن صفر می‌باشد. برای مجموعه بردارهای n تایی مثبت (نامنفی) از نماد $\mathcal{R}^n_{++}$ $(\mathcal{R}^n_{+})$ استفاده می‌کنیم. ماتریس قطری بردار $d \in \mathcal{R}^n$ را ماتریس $D = diag(d)$ درنظر می‌گیریم که در آن عناصر روی قطر اصلی درایه‌های بردار d می‌باشند و بقیه عناصر آن همگی صفر هستند، یعنی

$$D = diag(d) = \begin{cases} \circ & i \neq j \\ d_i & i = j \end{cases} \quad \Longrightarrow \quad D^{-1} = \begin{cases} \circ & i \neq j \\ \dfrac{1}{d_i} & i = j \end{cases}$$

برای تابع مفروض $f : \mathcal{R} \to \mathcal{R}$، نماد $\mathcal{C}^n$ را برای مجموعه همه تابع‌های پیوسته و n بار به‌طور پیوسته مشتق‌پذیر f درنظر می‌گیریم.

در انتهای این بخش قضیه‌ای را از آنالیز مقدماتی که در این پایان‌نامه از آن به‌صورت پیشفرض استفاده می‌شود بیان می‌کنیم.

قضیه ۱—۱۴.۱ [۹]:

(i) اگر f و g تابع‌های مختلط بر فضای متری X باشند، در این‌صورت، تابع‌های $f+g$، fg و $\frac{f}{g}$ (با $g(x) \neq \circ$) بر x پیوسته خواهند بود.

(ii) فرض کنید $f_1, \cdots, f_k$ توابعی حقیقی بر فضای متریک X و f نگاشتی از X به‌توی $\mathcal{R}^k$ باشد که با ضابطه $f = (f_1(x), \cdots, f_k(x)) \quad (x \in X)$ تعریف می‌شود. در این‌صورت، f پیوسته است، اگر و فقط اگر هر یک از تابع‌های $f_1, \cdots, f_k$ پیوسته باشند.

۱—۲ مفاهیم اولیه در بهینه‌سازی خطی

یک مساله برنامه ریزی خطی به فرم $\min\{c^Tx : Ax = b, x \geq 0\}$ را درنظر می‌گیریم. تاکنون روش‌های زیادی برای حل این‌گونه مساله‌ها ارائه شده‌است که به‌طور خلاصه تحت عنوان تاریخچه بیان می‌کنیم:

تاریخچه

روش‌هایی‌که برای حل مساله‌های برنامه ریزی خطی بیان شده‌اند در دو دسته مقدماتی و نقطه‌درونی تقسیم بندی می‌شوند، که در زیر به ترتیب و به‌طور خلاصه بیان می‌گردند.

روش‌های مقدماتی:

۱۹۴۷: روش سیمپلکس[۱] که توسط دانتزیگ[۲] ابداع شد و اولین روش برای حل مساله‌های بهینه‌سازی خطی است که در نوع خود بی‌نظیر است.

۱۹۵۰ تا ۱۹۶۰: در این سال‌ها روش‌های نقطه‌درونی اولین مراحل معرفی خود را سپری می‌کرد که سه الگوریتم از این روش‌ها برای حل مساله‌های LO ارائه شد: یکی توسط فریچ[۳] [۱۳، ۱۴] (سال۱۹۵۵) که تابع مانعی‌لگاریتمی را معرفی‌نمود. دیگری هوارد[۴] [۱۵] که روش مرکزی را بیان نمود. و دیکین[۵] [۱۶] که روش اولیه‌دوگان با مقیاس‌بندی آفین را بیان نمود.

۱۹۵۶: مدل خوددوگان متجانس[۶] که توسط تاکر[۷] [۲۴] ابداع شد.

۱۹۷۲: در این سال کلی و مینتی[۸] [۱۷] نشان دادند که روش سیمپلکس در بدترین حالت از مرتبه $\mathcal{O}(2^n)$ می‌باشد.

۱۹۷۸: روش بیضوی[۹] که توسط خاچیان[۱۰] [۱۸] ابداع شد. این الگوریتم اولین الگوریتم از مرتبه چندجمله‌ای بود که برای مساله‌های بهینه‌سازی خطی[۱۱] ارائه می‌گردید. تعداد تکرارها در این روش $\mathcal{O}(n^2L)$ می‌باشد که L اندازه مساله ورودی می‌باشد.

روش‌های نقطه‌درونی:

۱۹۸۴: روش نقطه‌درونی[۱۲] کاراتر در تئوری و عمل برای LO توسط کارمارکار[۱۳] [۲۱] ابداع شد. این روش دارای مرتبه اجرایی $\mathcal{O}(nL)$ بود.

۱۹۸۶: روش تحلیل مرکزی که توسط مگیدو[۱۴] [۱۹] ارائه گردید.

---

۱ Simplex
۲ Dantzig
۳ Frisch
۴ Huard
۵ Dikin
۶ Homogeneous self-dual model
۷ Tucker
۸ Minty , Klee
۹ Ellipsoid
۱۰ Khchayan
۱۱ Linear Optimization (LO)
۱۲ Interior Point Method (IPM)
۱۳ Karmarkar
۱۴ Megido

۱۹۸۹: روش اولیه‌دوگان تعقیب مسیر[۱] توسط کوجیما[۲] [۲۲] ابداع شد. کوجیما ثابت کرد که مسیر مرکزی یک مساله اولیه‌دوگان خطی به سمت جواب بهینه مساله هم‌گرا است.

۱۹۸۹ و ۱۹۹۲: روش پیشگو—تصحیح‌کننده که مبتنی بر روش‌های تعقیب مسیر بود توسط مهروترا[۳] [۲۹] ارائه گردید.

۱۹۹۴: مدل خوددوگان جاسازی شده[۴] توسط تاد و میزونو[۵] [۲۰] ارائه گردید.

۲۰۰۰: روش‌های خودمنظم توسط پنگ، راس و ترلاکلی[۶] [۳] ارائه شدند که مبتنی بر تابع‌های خودمنظم هستند. این روش بهترین الگوریتم ارائه شده برای حل مساله‌های بهینه‌سازی خطی به روش نقطه‌درونی به‌هنگام بلند است که از مرتبه اجرایی $\mathcal{O}(\sqrt{n}\log n \log \frac{n}{\varepsilon})$ می‌باشند. در این پایان‌نامه قصد معرفی وتوضیح تفصیلی در مورد این روش را خواهیم داشت.

۲۰۰۶: در این سال توسط دزا، نعمت‌اللهی، پیغامی و ترلاکی[۷] [۲۵] نشان داده شد که مسیر مرکزی می‌تواند رفتار بدی از خود نشان دهد، به‌طوری که حل مساله مبتنی بر این روش‌ها با تعداد تکرار زیادی به سمت جواب بهینه حرکت کند.

۲۰۰۴ تا ۲۰۱۰: روش‌های گام‌وفقی برای حل مساله‌ها ارائه شده که در این پایان‌نامه با برخی از آن‌ها آشنا خواهیم شد. بهترین الگوریتم ارائه شده مبتنی براین روش از مرتبه اجرایی $\mathcal{O}(q\sqrt{n\tau}\log \frac{n}{\varepsilon})$ می‌باشد.

نکته قابل ذکر در این‌جا این‌است که باوجود این‌که روش‌های نقطه‌درونی ارائه شده برروی حل با تعداد تکرار کمتر تمرکز دارند ولی روش سیمپلکس علی‌رغم مرتبه اجرایی نمایی (در بدترین حالت) آن، بخاطر وجود تحلیل حساسیت[۸] هم‌چنان به‌عنوان یکی از بهترین روش‌های عملی مورد استفاده قرار می‌گیرد.

در این پایان‌نامه، قصد داریم روش‌های نقطه‌درونی مبتنی بر تابع‌های هسته را برای مساله‌های بهینه‌سازی خطی معرفی نموده و نقش این تابع‌ها را در تحلیل مرتبه اجرایی الگوریتم‌ها مورد بررسی قرار داده و چند تابع هسته را تحلیل نماییم.

تعریف ۱—۱.۲ مساله‌های اولیه‌دوگان را به فرم استاندارد زیر درنظر بگیرید:

$$\begin{aligned} LP: &\quad \min\{c^Tx : Ax = b, x \geq 0\} \\ LD: &\quad \max\{b^Ty : A^Ty + s = c, s \geq 0\} \end{aligned} \tag{۱.۱}$$

۱ Primal-dual path flowing
۲ Kojima
۳ Mehrotra
۴ Self-dual Embedding model
۵ Mizuno , Todd
۶ Peng , Roos , Terlaky
۷ Deza,Nematollahi,Peyghami and Terlaky
۸ Sensitivity analysis

که در آن $A \in \mathcal{R}^{m\times n}$ و $b \in \mathcal{R}^m$ و $c \in \mathcal{R}^n$. نقطه $(x^\circ, y^\circ, s^\circ)$ را یک نقطه شدنی اولیه‌دوگان گوییم، هرگاه x° و (y°, s°) به ترتیب شدنی برای اولیه و دوگان باشند.

لم ۱—۲.۲ لم فارکاس[۱] [۲]: فرض کنید $Q_{k\times l}$ ماتریس و $q_{l\times 1}$ بردار ثابتی باشند، در این‌صورت دو عبارت زیر معادل‌اند:

i. برای هر $x \in \mathcal{R}^l$ شرط $pQx \geq \circ$ ایجاب می‌کند که $q^T x \geq \circ$.

ii. بردار نامنفی $u \in \mathcal{R}^k$ وجود دارد به‌طوری که در دستگاه $u^T Q = q^T$ (یا $Q^T u = q$) صدق می‌کند.

قضیه **۱—۳.۲ قضیه ضعیف دوگانی**[۲] [۴]: فرض کنید x شدنی برای LP و y شدنی برای LD باشد. در این‌صورت $b^T y \leq c^T x$.

به‌عبارت دیگر هر جواب شدنی مساله مینیمم‌سازی، کران بالایی برای مساله ماکزیمم‌سازی است. از این قضیه بی‌درنگ نتیجه زیر حاصل می‌شود.

نتیجه **۱—۴.۲** اگر x و y به ترتیب شدنی برای اولیه و دوگان باشند، به‌طوری که $c^T x = b^T y$، آن‌گاه x بهینه برای اولیه و y بهینه برای دوگان است.

از قضیه ضعیف و نتیجه آن قضیه مهم زیر به‌دست می‌آید.

قضیه **۱—۵.۲ قضیه قوی دوگانی**[۳] [۴]: برای مساله‌های اولیه‌دوگان یکی‌از چهار حالت زیر برقرار است:

(i) هردو مساله LP و LD شدنی است و $(x^\star, y^\star, s^\star)$ موجود است به‌طوری که $c^T x^\star = b^T y^\star$

$LP(ii)$ نشدنی است و LD بی‌کران است.

$LD(iii)$ نشدنی و LP بی‌کران است.

(iv) هر دومساله LP و LD نشدنی است.

تعریف ۱—۶.۲ برای جواب x شدنی اولیه و y شدنی دوگان، عبارت $c^T x - b^T y$ را فاصله دوگانی[۴] می‌نامند.

باتوجه به قضیه قوی دوگانی اگر $(x^\star, y^\star, s^\star)$ جواب بهینه اولیه‌دوگان باشد، آن‌گاه

$$c^T x^\star = b^T y^\star \Rightarrow c^T x^\star - b^T y^\star = \circ \Rightarrow c^T x^\star - Ay^\star x^\star = \circ \Rightarrow (c^T - Ay^\star)x^\star = \circ \Rightarrow s^{\star T} x^\star = \circ$$

[۱] Farkas lemma
[۲] Weak duality
[۳] Strong duality
[۴] Duality gap

لذا درصورتی‌که دستگاه زیر دارای جواب باشد، جواب حاصل یک جواب اولیه‌دوگان است.

$$\begin{cases} Ax = b & x \geq 0 \\ A^T y + s = c & s \geq 0 \\ xs = 0 \end{cases} \qquad (۲.۱)$$

تساوی آخر به شرط مکمل زاید معروف است وبیانگر آن‌است که اگر در مساله اولیه متغیری صفر شود، قید متناظر آن در مساله دوگان فعال نیست(نامساوی اکید است).

قضیه ۱—۷.۲ [۳]: مساله‌های(۱.۱) دارای جواب شدنی $(x^\star, y^\star, s^\star)$ هستند به‌طوری که $x^{\star T} s^\star = 0$ و $x^{\star T} + s^\star > 0$. جواب $(x^\star, s^\star)$ با این خاصیت را جواب بهینه مکمل اکید[۱] می‌نامند.

درصورتی‌که مجموعه جواب شدنی ناتهی باشد، می‌توان تعریف مهم زیر را که پایه روش‌های نقطه‌درونی است بیان کرد:

تعریف ۱—۸.۲ شرط نقطه‌درونی[۲] : یک دستگاه (خطی) معادله‌ها و نامعادله‌ها در شرط نقطه‌درونی صدق می‌کند، هرگاه یک جواب شدنی موجودباشد که در همه معادله‌ها یا نامعادله‌ها به‌طور اکید صدق کند.

به‌عبارت دیگر، دستگاه(۲.۱) در شرط نقطه‌درونی صدق می‌کند هرگاه جواب $(x^\circ, y^\circ, s^\circ)$ موجود باشد به‌طوری که

$$\begin{cases} Ax^\circ = b & x^\circ > 0 \\ A^T y^\circ + s^\circ = c & s^\circ > 0 \end{cases}$$

برای یافتن جواب دستگاه (۲.۱) دستگاه پارامتریک زیر را درنظر می‌گیریم.

$$LS(\mu) : \begin{cases} Ax = b & x \geq 0 \\ A^T y + s = c & s \geq 0 \\ xs = \mu e \end{cases} \qquad (۳.۱)$$

در این دستگاه μ همان فاصله دوگانی است، یعنی $\mu = \dfrac{x^T s}{n}$.

قضیه ۱—۹.۲ [۷]: درصورتی‌که دستگاه (۲.۱) در شرط نقطه‌درونی صدق کند، دستگاه (۳.۱) به‌ازای هر $\mu > 0$ دارای جواب یکتاست.

از دستگاه (۳.۱) بی‌درنگ تعریف زیر را داریم.

---
۱ Strictly complementary
۲ Interior Point Condition. (IPC)

تعریف ۱—۱۰.۲ مجموعه $C=\{(x(\mu),y(\mu),s(\mu))|\mu>0\}$ را مسیرمرکزی[۱] مساله‌های اولیه‌دوگان می گویند. $x(\mu)$ را μ ـ مرکز LP و $(y(\mu),s(\mu))$ را μ ـ مرکز LD می‌نامند.

وقتی μ به سمت بی‌نهایت میل می‌کند، آن‌گاه جواب دستگاه (۳.۱) به سمت جواب دستگاه (۲.۱) میل می‌کند که همان جواب اولیه‌دوگان است. لذا کافی‌است دستگاه پارامتریک غیرخطی (۳.۱) را حل نماییم. یکی از روش‌هایی که برای حل دستگاه (۳.۱) به کار می‌رود، روش نیوتن رافسون می‌باشد. یعنی در هر مرحله مقدار دستگاه زیر محاسبه می‌شود تا به یک تقریب خوب از جواب μ ـ مرکز نزدیک شویم:

$$\begin{cases} A\Delta x=0 \\ A^T\Delta y+\Delta s=0 \\ s\Delta x+x\Delta s=\mu e-xs \end{cases} \tag{۴.۱}$$

تعریف ۱—۱۱.۲ مجموعه

$$\mathcal{F}=\{(x,s):Ax=b,A^Ty+s=c,x\geq 0,s\geq 0,x^Ts=0\}$$

را مجموعه جواب‌های بهینه اولیه‌دوگان درنظر می گیریم. قرار دهید:

$$\mathcal{B}_\mathcal{F}=\{i:\exists(x,s)\in\mathcal{F},\quad s.t\quad x_i>0\}$$

$$\mathcal{N}_\mathcal{F}=\{i:\exists(x,s)\in\mathcal{F},\quad s.t\quad s_i>0\}$$

در این‌صورت مرکز تحلیلی[۲] $\mathcal{F}$ از رابطه زیر به‌دست می آید:

$$(x^\star,s^\star)=arg\max_{(x,s)\in\mathcal{F}}\prod_{i\in\mathcal{B}_\mathcal{F}}x_i\prod_{j\in\mathcal{N}_\mathcal{F}}s_j$$

قضیه ۱—۱۲.۲ [۷]: اگر $\mu\to 0$ آن‌گاه مسیر مرکزی به مرکز تحلیلی جواب بهینه مجموعه $\mathcal{F}$ هم‌گراست.

۱—۳ روش اولیه‌دوگان نیوتن برای بهینه‌سازی خطی وتابع‌های نزدیکی

همان‌گونه که بیان شد، برای حل دستگاه (۲.۱) می‌توان از روش نیوتن استفاده کرد. برای این منظور الگوریتم نیوتن برای حل دستگاه‌های غیر خطی را بیان کرده و برهان قضیه‌های هم‌گرایی این روش را به [۳۰] ارجاع می‌دهیم.

---
۱ Centeral Path (C.P)
۲ Analytic center

شکل کلی یک دستگاه معادله‌های غیر خطی را به‌صورت زیر درنظر بگیرید:

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0 \\ f_2(x_1, x_2, \cdots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \cdots, x_n) = 0 \end{cases}$$

با نمادگذاری $F(x_1, x_2, \cdots, x_n) = (f_1(x_1, x_2, \cdots, x_n), f_2(x_1, x_2, \cdots, x_n), \cdots, f_n(x_1, x_2, \cdots, x_n))$ دستگاه فوق به‌شکل $F(x) = 0$ خواهد شد. برای پیدا کردن تقریبی از جواب دستگاه غیرخطی $F(x) = 0$ یک تقریب اولیه x^0 انتخاب می‌کنیم. حال الگوریتم زیر را برای به‌دست آوردن جواب، که به الگوریتم نیوتن معروف است داریم [۳۰]:

الگوریتم روش نیوتن

ورودی :

◇ مقدار اولیه x^0 و پارامتر دقت $\tau > 0$

شروع کن :

◇ قرار بده $k = 0$ و $x = x^0$

◇ تازمانی‌که $\|F(x^k)\| \geq \tau$ انجام بده

◇ $F(x^k)$ و $J(x^k)$ را محاسبه کن

$$J(x^k) = \begin{vmatrix} \frac{\partial f_1(x^k)}{\partial x_1} & \frac{\partial f_1(x^k)}{\partial x_2} & \cdots & \frac{\partial f_1(x^k)}{\partial x_n} \\ \frac{\partial f_2(x^k)}{\partial x_1} & \frac{\partial f_2(x^k)}{\partial x_2} & \cdots & \frac{\partial f_2(x^k)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(x^k)}{\partial x_1} & \frac{\partial f_n(x^k)}{\partial x_2} & \cdots & \frac{\partial f_n(x^k)}{\partial x_n} \end{vmatrix}$$

◇ دستگاه خطی $n \times n$، $J(x^k)y^k = -F(x^k)$ را برای y^k حل کن.

◇ قرار بده $x^{k+1} = x^k + y^k$

◇ $k + 1$ را در k قرار بده

◇ پایان تکرار

توقف کن. جواب حاضر یک τ−تقریب از جواب دستگاه می‌باشد.