**وزارت علوم، تحقیقات و فناوری**

**دانشگاه دامغان**

**دانشکده ریاضی و علوم کامپیوتر**

پایان‌نامه کارشناسی ارشد

ریاضی کاربردی (گرایش تحقیق در عملیات)

# توسیع روش نیوتن برای حل مسائل بهینه سازی چند هدفه

توسط:

**مریم شاه‌محمدی**

استاد راهنما:

**دکتر اکبر هاشمی برزآبادی**

استاد مشاور:

**دکتر علی عباسی ملایی**

شهریور ۱۳۹۱

بسم الله الرحمن الرحیم

**به نام خدا**

# توسیع روش نیوتن برای حل مسائل بهینه سازی چند هدفه

**توسط:**

مریم شاه‌محمدی

**پایان‌نامه**

ارائه شده به تحصیلات تکمیلی دانشگاه به عنوان بخشی از فعالیت‌های تحصیلی لازم برای اخذ درجه کارشناسی ارشد

**در رشته**

ریاضی کاربردی

از دانشگاه دامغان

ارزیابی و تأیید شده توسط کمیته پایان‌نامه با درجه: **عالی**

دکتر اکبر هاشمی برزآبادی دانشیار ریاضی کاربردی گرایش کنترل بهینه دانشکده‌ی ریاضی و علوم کامپیوتر دانشگاه دامغان (استاد راهنما)

دکتر علی عباسی ملایی استادیار ریاضی کاربردی گرایش تحقیق در عملیات دانشکده‌ی ریاضی و علوم کامپیوتر دانشگاه دامغان (استاد مشاور)

دکتر امید سلیمانی فرد دانشیار ریاضی کاربردی گرایش کنترل بهینه دانشکده‌ی ریاضی و علوم کامپیوتر دانشگاه دامغان (داور اول)

دکتر جعفر فتحعلی استادیار ریاضی کاربردی گرایش تحقیق در عملیات دانشکده‌ی ریاضی دانشگاه صنعتی شاهرود (داور دوم)

دکتر بهزاد صالحیان استادیار ریاضی محض گرایش نظریه گراف دانشکده‌ی ریاضی و علوم کامپیوتر دانشگاه دامغان (نماینده تحصیلات تکمیلی)

شهریور ۱۳۹۱

# تقدیم به

بهترین همراهان زندگیم،
پدر و مادرم

دو عزیزی که فرصت اندیشیدن و رستن را به من عطا نمودند؛

آنانکه در رویش نهال دانشم صبوری کردند و روح اندیشیدن را در من بیدار کردند؛

و پیشکش می نمایم به برادر و خواهر مهربانی که عابران همیشه ی کوچه های مهربانی اند؛

و ضرب آهنگ قدم هایشان بذر آرامش را روی چشم های خسته اما تشنه علم من پاشیدند.

الهی!
هیچگاه تنهایشان مگذار.

# سپاسگزاری

خدایا هر کس از تو نوشت واژه کم آورد، فکر کم آورد، دل، ولی تپید و تپید و ادعا کرد حتی اگر از نفس بیافتد، کم نمی آورد، دل راست می گفت که انسان قله ای دارد بنام تنهایی که در آن به تماشای غروب تمام وابستگی ها می نشیند و سپس حس می کند که تنها تو در کنار او می مانی. معبودا، در من احساس عرفانی اسیر است که از اعماق فطرتم سرچشمه گرفت و هرگاه نام تو را بر زبان می آورم، تا چشم هایم می جوشد.

پس ای نیک... که به نیکی شناخته شده ای، اینک من بر سجاده نیازم بوسه می زنم و دست های التماسم را به سویت دراز می کنم و فریاد می زنم که با یاری، توکل، امید و عشق به تو توانستم در مسیری قدم بگذارم که حتی ثانیه هایش هم با رایحه خوش علم و دانش آمیخته شده و غرقم کردی در آنچه که رسول مهربانت فرمود: «اطلبوا العلم، من المهد الی اللحد». من با کمک از تو پیمودم این مسیر سبز را؛

از دو الهه عشق و مهربانی پدر بزرگوار و مادر فداکارم بی نهایت سپاسگذارم و بر دستان پر مهرشان بوسه می زنم.

از همراهان دیروز و امروزم برادر و خواهر عزیزم که وجودشان گرمابخش لحظه لحظه زندگی ام بوده تشکر و قدردانی می کنم.

صادقانه ترین سپاس ها را تقدیم میکنم به استاد بزرگوارم جناب آقای دکتر اکبر هاشمی برزآبادی که شاگردی در مکتب ایشان مایه مباهات من است.

از استاد گرانقدر و بزرگوارم جناب آقای دکتر علی عباسی ملایی که قبول زحمت کردند و در امر مشاوره ی پایان نامه یاری ام کردند صمیمانه سپاس گزارم. در خاتمه از تمام دوستان و همراهان عزیزم صمیمانه سپاسگزارم.

خدایا تو را سپاس که برایم یارانی چنین نیکو قرار دادی.

مریم شاه محمدی

شهریور ۹۱

چکیده

# توسیع روش نیوتن برای حل مسائل بهینه سازی چند هدفه

به وسیله‌ی:
مریم شاه‌محمدی

هدف اصلی این پایان نامه معرفی روش نیوتن برای حل مسائل بهینه‌سازی چندهدفه و ارائه‌ی توسیع این روش به کمک روش شبه-نیوتن برای حل این‌گونه مسائل و مقایسه این دو روش می‌باشد. بدین منظور پس از بیان مفاهیم مقدماتی مورد نیاز و مروری بر روش نیوتن در حل مسائل بهینه سازی تک هدفه، راهکاری برای بکار بردن این روش بر روی مسائل چند هدفه و همچنین الگوریتمی برای این دو روش ارائه شده است. این روش از فاکتور های وزن‌دهی یا رتبه بندی کردن توابع هدف استفاده ننموده و جهت نیوتن و شبه نیوتن در هر تکرار از کمینه کردن بیشینه بردار توابع هدف بدست می‌آید. دنباله‌ی تولید شده توسط این روش‌ها برای توابع چند هدفه در صورتی که تمامی توابع هدف دوبار پیوسته مشتق‌پذیر و محدب باشد، به‌صورت فوق خطی همگرا به نقطه‌ی بهینه است. نتایج عددی حاصل از به‌کارگیری روش نیوتن و شبه-نیوتن با ارایه مسایل آزمون مقایسه شده‌اند.

# فهرست مطالب

# فهرست شکل‌ها

# پیشگفتار

بهینه‌سازی، علم یافتن بهترین پاسخ برای مسأله‌ای است که به صورت ریاضی تعریف شده است. اغلب مسائل بهینه‌سازی در طبیعت و علوم مختلف به دنبال بهینه کردن چند تابع هدف هستند. برای مثال در طراحی یک نوع پل ممکن است بطور هم‌زمان کمینه کردن توده‌های مصالح و بیشینه کردن استحکام مد نظر باشد. این گونه مسائل را بهینه‌سازی چند کرداری یا چند هدفه می‌نامند. مسأله بهینه‌سازی چند هدفه [۵۴] را می‌توان مسأله پیدا کردن یک بردار از متغیرهای تصمیم نامید که این بردار قیود را برآورده ساخته و تابع برداری که مولفه‌هایش نمایشگر تابع‌های هدف می‌باشد را بهینه می‌سازد. معمولا مدل ریاضی یک مسأله بهینه‌سازی چند هدفه **MOP** [1] به این صورت زیر بیان می‌شود:

$$\begin{aligned}
&\min_{x\in U} \quad F(x) = [f_1(x)\ ,\ f_2(x)\ ,\ \cdots\ ,\ f_m(x)\,] \\
&s.t. \\
&h_k(x) \quad = 0, \quad k \in K \\
&g_j(x) \quad \leq 0, \quad j \in J \\
&x \in U \quad , \quad U \subseteq \mathbb{R}^n.
\end{aligned}$$

که در آن $f_i$ ، $i$ امین تابع هدف و $K$ و $J$ به ترتیب مجموعه‌ی اندیس قیدهای تساوی و نامساوی و $x$ بردار متغیرهای تصمیم است. جواب مسأله بالا به‌صورت مجموعه‌ای از نقاط است که پارتو نام دارد، بنابراین در این‌گونه مسائل به جای وجود یک جواب واحد با مجموعه ای از نقاط پارتو روبه رو هستیم که ممکن است نامتناهی باشد. توجه داشته باشید که هیچ یک از جواب ها بهتر از دیگری نیست، یعنی

[1] Multi objective Programming

با حرکت از یک جواب پارتو به جواب دیگر در صورت بهبود یکی از توابع هدف، حداقل یکی از مولفه های تابع هدف بدتر می‌شود.

مطالعه مسائل بهینه‌سازی چند هدفه در طول دو دهه گذشته در سه شاخه‌ی: تحقیق در عملیات، اقتصاد و روانشناسی به سرعت پیشرفت کرده است. کوهن [۲] و تاکر [۳] ابتدا این مسأله را در سال ۱۹۵۱ فرمول بندی و سپس شرایط بهینگی برای وجود جواب را بیان کردند [۳۶] . کار آنها زمینه را برای توسعه الگوریتم‌های برنامه نویسی ریاضی فراهم کرد. در سال ۱۹۵۳، اَررو [۴] و دیگران شرایط بهینگی دیگری را معرفی کردند [۳] . گاس و ساتی در سال ۱۹۵۵ اولین روش قابل اجرا برای حل مسائل بهینه‌سازی چند هدفه را ارائه دادند [۳۶] . پس از آن پژوهش‌ها در زمینه مسائل بهینه‌سازی چند هدفه ادامه یافت تا آنجا که اولین کنفرانس انحصاری برنامه نویسی بهینه‌سازی چند هدفه در دانشگاه کارولینای جنوبی [۵] در سال ۱۹۷۲ برگزار شد. مجموعه مقالات این کنفرانس در سال ۱۹۷۳ در کوکران [۶] و زلنی [۷] منتشر شد. در سال ۱۹۷۵ مجموعه مقالاتی در زمینه‌ی بهینه‌سازی چند هدفه در مؤسسه علوم مدیریت ونشست بین المللی در کیوتو ژاپن [۸] ارائه و در سال ۱۹۷۶ در زلنی منتشر شد.
در مطالعه مسائل بهینه‌سازی چند هدفه اولین چیزی که مورد توجه قرار می‌گیرد، تعریف قابل قبولی از نقاط بهینه است. اولین مفهوم بهینگی توسط اِدوس [۹] [۱۶] (۱۸۸۱) و پارتو [۱۰] [۱۶] (۱۸۹۶) بیان شد. از آن پس زمینه‌های تحقیقاتی برای بهینه‌سازی چند هدفه فراهم شد و چندین کتاب و مقاله برای معرفی بیشتر این موضوع ارائه و منتشر گردید (به عنوان مثال مراجع [۹، ۱۵، ۱۷، ۱۹، ۳۰، ۳۲، ۴۰، ۴۶، ۴۷، ۴۹، ۵۱] را ببینید.)
از آنجایی که در این گونه مسائل توابع هدف عمدتاً در تقابل با یگدیگر هستند، یافتن یک نقطه واحد که همه توابع هدف را بطور همزمان بهینه کند، تقریبا غیرممکن است، بنابراین در اینجا عمل بهینه‌سازی به معنای پیدا کردن جوابی است که بتواند به توابع هدف مقادیری القا کند که از نظر تصمیم گیرنده معقول باشد. برای یافتن جواب این گونه مسائل مفهوم بهینه را با مفهوم بهینه پارتو جایگزین می کنیم.
کاربرد این گونه مسائل را می‌توان در مهندسی [۱۸، ۳۴] (مخصوصا در تقویت بهینه سازی [۱۰،

[۲] Kuhn

[۳] Tucker

[۴] Arrow

[۵] South Carolina

[۶] cochrane

[۷] Zeleny

[۸] Institute of Management Sciences International Meeting in Kyeto

[۹] Edgewth

[۱۰] Pareto

۳۵، ۵۲] ،طراحی و فضای اکتشاف [۴۳، ۵۳، ۲۵، ۳۳، ۵۰] ) مسائل آماری [۸] ،علم مدیریت [۱۹، ۴، ۵۶، ۲۸، ۴۴، ۱] ،تجزیه وتحلیل محیطی [۲۱، ۲۰، ۳۸] ،درمان سرطان [۳۱] و غیره یافت.

روش‌های بسیاری برای حل مسائل بهینه سازی چند هدفه در دسترس است. در بسیاری از این روش ها از تبدیل $MOP$ به یک مسأله‌ی بهینه سازی تک هدفه استفاده می‌شود به این ترتیب که با استفاده از روشی مناسب تابع هدف مسأله اولیه را از فرم برداری به اسکالر تبدیل می‌کنند. یکی از استراتژی های اصلی برای حل مسائل بهینه سازی چند کرداری، تقریب عددی است [۲۶، ۵۷، ۲۷] . در این روش‌ها مسأله بهینه سازی تک هدفه با یک یا چند پارامتر حل شده که نتایج همان نقاط بهینه پارتو برای مسأله چند هدفه اولیه است. در این تقریب پارامترها در ابتدا شناخته شده نیستند و انتخاب آنها بر عهده تصمیم گیرنده است. روش هایی با عنوان، روش وزنی و روش قیدی و روش کمینه بیشینه وزنی از این دسته اند که در فصل مقدمه بطور مختصر به ساختار آن‌ها اشاره شده است. در بعضی از مسائل انتخاب پارامترها ممکن است مشکل ساز شود که در ادامه به یکی از آنها که در آن اغلب تمام پارامترهای انتخاب شده باعث بیکران شدن مسأله می‌شوند، اشاره خواهیم کرد. اخیراً در تکنیک‌های عددی تطبیقی، پارامترهای عددی که به طور خودکار در طول الگوریتم پیشنهاد می‌شوند، به گونه‌ای هستند که کیفیت معینی از تقریب برقرار باشد [۱۶، ۲۲، ۲۹، ۵۵] . امروزه الگوریتم‌های بهینه‌سازی چند هدفه‌ای که از روش‌های عددی استفاده نمی‌شود، توسعه یافته ( [۶، ۷] برای مروری بر این موضوع مشاهده نمایید). برخی از این تکنیک ها بسطی از الگوریتم بهینه‌سازی عددی هستند (مانند الگوریتم تندترین شیب [۲۴، ۱۴، ۱۳] با حداکثر همگرایی خطی). با این وجود برای یافتن جواب مسائل بهینه‌سازی چند هدفه برخی ایده‌های ابتکاری قوی‌تر در بهینه سازی مانند الگوریتم‌های تکاملی ارائه شده است. [۳۷، ۴۱] . تاکنون هیچ اثبات همگرایی برای این روش‌ها شناخته نشده و نتایج تجربی نشان می‌دهد که سرعت همگرایی عمومی در این روش‌ها بسیار کند است [۵۸] . تکنیک‌های بهینه‌سازی چند هدفه‌ی بدون پارامتر براساس معیارهای دیگری تقسیم‌بندی می‌شوند، به عنوان مثال مرتب کردن از لحاظ اهمیت مؤلفه‌های بردار تابع هدف. در این مورد اولویت‌ها باید از قبل تعیین شده باشد. بعلاوه فرایندهای بهینه‌سازی متناظر معمولا توسط رویه‌های متقابل تکمیل می‌گردد [۳۹] .

در این پایان نامه در فصل یک ابتدا مفاهیم و تعاریف مقدماتی مورد نیاز بیان شده، در فصل دوم روش نیوتن برای حل مسائل بهینه سازی چندهدفه که الگوریتم آن برای اولین بار توسط فلینگ ارائه شده [۲۳] را بیان می‌کنیم. در نهایت در فصل سوم الگوریتم بیان شده در فصل دو را به کمک روش شبه نیوتن برای حل این‌گونه مسائل تعمیم داده‌ایم.

# فصل ۱

# تعریف و مفاهیم مقدماتی

## ۱-۱ مقدمه

در این بخش مفاهیم مقدماتی آنالیزی مورد نیاز و همچنین مفاهیمی در خصوص بهینه سازی و برخی روش‌های بهینه سازی برای حل مسائل بهینه سازی چند هدفه را بیان خواهیم کرد. مفاهیم و روش‌های این بخش از مراجع [۵، ۴۵، ۴۲] ذکر گردیده است.

## ۱-۲ برخی مفاهیم مقدماتی از آنالیز و جبر خطی

**تعریف ۱.۲.۱.** **نرم اقلیدس** یا طول بردار $X \in \mathbb{R}^n$ به صورت $\|X\|_2 = (\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}}$ که در آن $x_i$ ها مؤلفه‌های بردار $X$ هستند، تعریف می‌شود.

**تعریف ۲.۲.۱.** تابع $f$ روی مجموعه‌ی $X$ از $\mathbb{R}^n$ ، **پیوسته‌ی لیب‌شیتس** [۱] (انقباضی) است اگر $0 < \gamma < 1$ وجود داشته باشد به طوری که به ازای $x, y \in X$ آن‌گاه

$$\|f(x) - f(y)\| \leq \gamma \|x - y\|.$$

**تعریف ۳.۲.۱.** مجموعه‌ای که هر پوشش آن زیر پوشش باز متناهی داشته باشد **مجموعه‌ی فشرده** نامیده می‌شود.

**تعریف ۴.۲.۱.** توابعی که مشتق آن از هر مرتبه موجود و پیوسته باشد را **تابع هموار** می‌نامند.

[۱] Lipschitz

**تعریف ۵.۲.۱.** هر تابعی که هموار نباشد را **تابع ناهموار** می‌نامیم.

**قضیه ۶.۲.۱.** *دنباله‌ی* $x_1, x_2, x_3, \cdots$ *به* $x$ ***همگرا** است اگر و فقط اگر*

$$\lim_{k\to\infty} \|x_k - x\| = 0.$$

**تعریف ۷.۲.۱.** گوییم دنباله‌ی $\{x_k|k \geq 0\}$ همگرا به $\alpha$ با مرتبه همگرایی $p > 0$ است اگر

$$\|x_{k+1} - \alpha\| \leq c\|x_k - \alpha\|^p, k \geq 0$$

که در آن $c$ کمیتی ثابت است و $c \geq 0$ . بعضی از انواع همگرایی را به صورت زیر می‌توان دسته‌بندی کرد:

(۱) اگر $p = 1$ و $0 < c < 1$ ، همگرایی به دست آمده **همگرایی خطی** نامیده می‌شود،

(۲) اگر $p = 2$ ، **همگرای درجه دوم** داریم،

(۳) اگر $\dfrac{\|x_{k+1} - \alpha\|}{\|x_k - \alpha\|} \to 0$ ، برای زمانی که $k \to \infty$ ، آن گاه **همگرایی فوق خطی** [2] به دست می‌آید.

**تعریف ۸.۲.۱.** ماتریس مربعی حقیقی $A = (a_{ij})$ از مرتبه‌ی $n$ را **معین مثبت** گوییم هرگاه:

(۱) $A$ متقارن باشد،

(۲) برای هر بردار ناصفر $x$ در $\mathbb{R}^n$ داشته باشیم $x^TAx > 0$ .

**تعریف ۹.۲.۱.** یک **مجموعه‌ی تراز** [3] تابع حقیقی مقدار و $n$ متغیره‌ی $f$ را به صورت زیر تعریف می‌کنیم

$$L_c(f) = \{(x_1, x_2, \cdots x_n)|f(x_1, x_2, \cdots x_n) = c\},$$

که $c$ در آن مقدار ثابتی است. به عبارت دیگر مجموعه‌ای از نقاط است که به ازای آن‌ها $f$ مقدار ثابت $c$ را به خود می‌گیرد.

---

[2] Superlinear convergence

[3] level set

## ۱ـ۳ مفاهیم مقدماتی مربوط به بهینه‌سازی

**تعریف ۱.۳.۱.** نقطه $x \in D_F$ را **نقطه بحرانی** تابع $F$ گویند هرگاه $F'(x) = 0$ یا $F'(x)$ موجود نباشد.

### ۱ـ۳ـ۱ تابع محدب

یکی از مفاهیم مهم در بهینگی، محدب بودن یا مقعر بودن تابع $f$ می‌باشد که تعریف آن را در زیر ارئه می‌کنیم.

**تعریف ۲.۳.۱.** به **مجموعه** $S \subseteq \mathbb{R}^n$ **محدب** گفته می‌شود، اگر برای هر دو نقطه‌ی $x, y \in S$ و برای هر $\lambda \in [0, 1]$ داشته باشیم $\lambda x_1 + (1 - \lambda)x_2 \in S$ .

**تعریف ۳.۳.۱.** فرض کنید $S$ یک مجموعه محدب ناتهی در $\mathbb{R}^n$ و تابع $f : S \to \mathbb{R}$ را داشته باشیم. تابع $f$ در $S$ **محدب** است اگر به ازای هر $x, y \in S$ و $\lambda \in (0, 1)$ داشته باشیم

$$f[\lambda x + (1 - \lambda)y] \leq \lambda f(x) + (1 - \lambda)f(y).$$

اگر در رابطه‌ی فوق اگر $\leq$ به $<$ تبدیل شود تابع را **محدب اکید** گوییم.

در بعضی مواقع شرط محدب بودن تابع ممکن است خیلی قوی باشد و به شرط ضعیف‌تری مثل محدب بودن در یک نقطه نیاز داشته باشیم. این مفهوم را به صورت زیر بیان می‌کنیم.

**تعریف ۴.۳.۱.** فرض کنید $S$ یک مجموعه‌ی محدب ناتهی در $\mathbb{R}^n$ و تابع $f : S \to \mathbb{R}$ تعریف شود، تابع $f$ محدب در $\bar{x} \in S$ گفته می‌شود اگر به ازای هر $x \in S$ و $\lambda \in (0, 1)$ داشته باشیم

$$f[\lambda \bar{x} + (1 - \lambda)x] \leq \lambda f(\bar{x}) + (1 - \lambda)f(x).$$

**قضیه ۵.۳.۱.** [۵] *فرض کنید $f$ یک تابع حقیقی مقدار مشتق‌پذیر تعریف شده روی یک مجموعه‌ی غیر تهی باز و محدب $S$ از $\mathbb{R}^n$ باشد. آن‌گاه، $f$ محدب در $\bar{x} \in S$ است اگر و تنها اگر به ازای هر $x \in S$ داشته باشیم*

$$f(x) - f(\bar{x}) \geq \nabla f(\bar{x})^T (x - \bar{x}).$$

### ۱-۳-۲ شرایط بهینگی کروش-کان-تاکر (KKT)

مسأله‌ی برنامه‌ریزی غیر خطی زیر را در نظر بگیرید

$$\begin{aligned} &\min f(x) \\ &s.t \quad x \in S. \end{aligned} \tag{۱.۱}$$

که در آن ناحیه شدنی $S$ به‌صورت زیر است

$$S = \{x \in X : g_i(x) \leq 0 \, for \, i = 1, \cdots, m\},$$

که در آن $X$ یک مجموعه باز ناتهی در $\mathbb{R}^n$ است.
اکنون شرایط بهینه کروش-کان-تاکر [۴] برای مسأله بالا در قضیه زیر مطرح می‌شود.

**قضیه ۱.۳.۶.** [۵] ***(شرایط لازم کروش-کان-تاکر)***
*فرض کنید $X$ یک مجموعه‌ی باز ناتهی در $\mathbb{R}^n$ باشد و $f : \mathbb{R}^n \to \mathbb{R}$ و $g_i : \mathbb{R}^n \to \mathbb{R}$، برای $i = 1, \cdots, m$ تعریف شده باشد. مسأله‌ی (۱.۱) را در نظر بگیرید. فرض کنید $\bar{x}$ یک جواب شدنی باشد، در این صورت تعریف می‌کنیم $I = \{i : g_i(\bar{x}) = 0\}$ . با فرض این که $f$ و $g_i$ ها برای $i \in I$ مشتق پذیر در $\bar{x}$ و $g_i$ برای $i \notin I$ ، در $\bar{x}$ پیوسته هستند و $\nabla g_i(\bar{x})$ به ازای $i \in I$ مستقل خطی است. اگر $\bar{x}$ جواب بهینه‌ی موضعی مسأله باشد، آن‌گاه اسکالرهای $\lambda_i$ برای $i \in I$ وجود دارد به‌طوری‌که*

$$\nabla f(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) = 0, \tag{۲.۱}$$

$$\lambda_i \geq 0 \quad i \in I. \tag{۳.۱}$$

*در مجموع با مفروضات بالا اگر $g_i$ به‌ازای هر $i \notin I$ نیز مشتق‌پذیر در $\bar{x}$ باشد، در این صورت شرایط فوق را می‌توان به صورت معادلات زیر نوشت:*

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) = 0, \tag{۴.۱}$$

$$\lambda_i g_i(\bar{x}) = 0, \quad i = 1, \cdots, m. \tag{۵.۱}$$

$$\lambda_i \geq 0, \qquad i = 1, \cdots, m. \tag{۶.۱}$$

**تعریف ۱.۳.۷.** [۵] در شرایط (۲.۱) تا (۶.۱) اسکالرهای $\lambda_i$ **ضرایب لاگرانژ** [۵] نامیده می‌شوند.

---
[۴] Karush-Kuhn-Tuker

[۵] Lagrange multipliers

### ۱-۳-۳ مسأله بهینه سازی محدب

در حالت کلی یک مسأله بهینه سازی را می‌توان به صورت زیر نوشت

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} \quad f(x) \\ s.t. \quad & h_k(x) \quad = 0, \quad k\in K, \\ & g_j(x) \quad \leq 0, \quad j\in J. \end{aligned} \tag{۱.۷}$$

مسأله بهینه سازی بالا را **محدب** [۶] $(CO)$ گوییم هرگاه در شرایط زیر صدق کند

- تابع هدف محدب باشد،
- قیدهای تساوی $h_k$ ، $k\in K$ ، خطی باشند و
- قیدهای نامساوی $g_j$ ، $j\in J$ ، محدب باشند.

**قضیه ۱.۳.۸.** [۵] *مسأله‌ی $P$ به صورت کمینه سازی $f(x)$ مشروط بر این‌که*

$$x\in X=\{x\in\mathbb{R}^n : g_j(x)\leq 0, j=1,\cdots,m\}$$

*در نظر بگیرید و فرض کنید یک مسأله‌ی محدب باشد و $\bar{x}\in X$ . همچنین فرض کنید $f$ و $g_j$ ، $j=1,\cdots,m$ در $\bar{x}$ به طور پیوسته مشتق‌پذیر باشند. اگر ضرایب لاگرانژ $\lambda_j\geq 0$ ، $j=1,\cdots,m$ وجود داشته باشند، به طوری که*

$$\nabla f(\bar{x})+\sum_{j=1}^{m}\lambda_j\nabla g_j(\bar{x})=0,$$

*و همچنین*

$$\lambda_j g_j(\bar{x})=0, \quad \forall j=1,\cdots,m.$$

*آن‌گاه $\bar{x}$ یک جواب بهینه‌ی مسأله‌ی $P$ است.*

### ۱-۳-۴ شرایط اسلَتر

مسأله‌ی بهینه سازی محدب زیر را در نظر بگیرید

$$\begin{aligned} (CO) \quad & \min f(x) \\ s.t. \quad & g_j(x)\leq 0, \quad j=1,\cdots,m \\ x \quad & \in C. \end{aligned} \tag{۱.۸}$$

[۶] convex optimization

که در آن $C \subseteq \mathbb{R}^n$ ، یک مجموعه‌ی محدب و $f, g_1, \cdots, g_m$ توابع محدب روی $C$ (یا روی یک مجموعه‌ی باز شامل $C$ ). فرض می‌کنیم توابع، $f$ و $g_1, \cdots, g_m$ مشتق‌پذیرند. اگر مجموعه‌ی نقاط درونی $C$ را $C^\circ$ و مجموعه اندیس‌های $\{1, \cdots, m\}$ را $J$ بنامیم، می‌توان مجموعه جواب‌های شدنی را به صورت زیر تعریف کرد

$$F = \{x \in C | g_j \leq 0, j \in J\}.$$

**تعریف ۹.۳.۱.** یک بردار (نقطه) $x^\circ \in C^\circ$ را **نقطه‌ی اسلَتر** [۷] می‌نامند هرگاه

- به ازای $j$ هایی که $g_j$ غیرخطی باشد داشته باشیم، $g_j < 0$ ،
- به ازای $j$ هایی که $g_j$ خطی باشد داشته باشیم، $g_j \leq 0$ .

اگر نقطه‌ی اسلَتر وجود داشته باشد، گوییم مسأله‌ی $(CO)$ در شرایط اسلَتر صدق می‌کند.

## ۱-۳-۵ خلاصه‌ای از روش نیوتن

روش نیوتن یک روش جستجوی چند بعدی است که از مشتقات در تعیین جهت‌های جستجو استفاده می‌کند. این روش از حاصل‌ضرب بردار گرادیان در معکوس ماتریس هسیان به عنوان جهت کاهشی استفاده می‌کند. در این روش از تقریب درجه دوم تابع بجای تقریب خطی برای تابع استفاده می‌شود. در زیر $q$ یک تقریب درجه دو از تابع $f$ در نقطه‌ی $x_k$ است

$$q(x) = f(x_k) + \nabla f(x_k)^t (x - x_k) + \frac{1}{2}(x - x_k)^t H(x_k)(x - x_k).$$

که $H(x_k)$ ماتریس هسیان $f$ در نقطه‌ی $x_k$ است. شرط لازم برای کمینه تقریب درجه دوم $q$ عبارت است از: $\nabla q(x) = 0$ یا $\nabla f(x_k)^t + H(x_k)(x - x_k) = 0$ . با فرض این که معکوس $H(x_k)$ وجود دارد، اگر قرار دهیم $p_k = H(x_k)^{-1} \nabla f(x_k)$ آن‌گاه نقطه‌ی بعدی $x_{k+1}$ به صورت زیر مشخص می‌گردد

$$x_{k+1} = x_k - \alpha_k p_k,$$

که $\alpha_k$ همان طول گام در جهت $p_k$ است. فرض کنید $\nabla f(\bar{x}) = 0$ و $H(\bar{x})$ در کمینه محلی $\bar{x}$ معین مثبت است و از این رو، نقطه بعدی $x_{k+1}$ خوش تعریف می‌باشد.

[۷] Slater point

### ۱-۳-۶ خلاصه‌ای از روش شبه-نیوتن

این روش همانند روش نیوتن است با این تفاوت که در روش شبه-نیوتن جهت جستجو از رابطه زیر بدست می‌آید

$$p_k = -B_k^{-1} \nabla f(x_k),$$

یعنی $x_k$ در هر تکرار به صورت زیر به‌هنگام می‌شود

$$x_{k+1} = x_k - \alpha_k B_k^{-1} \nabla f(x_k).$$

ماتریس $B_k$ تقریبی از ماتریس هسیان است. این ماتریس متقارن و معین مثبت است و در هر تکرار توسط یکی از فرمول‌های شبه-نیوتن به‌هنگام می‌شود. یکی از فرمول‌هایی که معمولا از آن برای به‌هنگام کردن $B_k$ استفاده می‌شود، فرمول $BFGS$ [8] است که به صورت زیر نوشته می‌شود،

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k},$$

که در آن $s_k = x_{k+1} - x_k$ و $y_k = \nabla f_{k+1} - \nabla f_k$ است.

### ۱-۳-۷ شرایط آرمیجو

به طور کلی در روش‌های جستجوی خطی، در ابتدا الگوریتم به دنبال یافتن جهت $p_k$ است و سپس در راستای این جهت از نقطه‌ی جاری $x_k$ حرکت می‌کند تا به نقطه‌ای جدیدی که در آن مقدار تابع هدف بهبود می‌یابد، برسد. میزان طول گامی که در مسأله‌ی کمینه سازی باعث کاهش اساسی در تابع هدف شود را می‌توان از حل تابع زیر بدست آورد،

$$\min_{\alpha > 0} \phi(\alpha) = \min_{\alpha > 0} f(x_k + \alpha p_k).$$

با وجود این که با یافتن جواب مسأله‌ی کمینه سازی بالا بهترین انتخاب برای طول گام مشخص می‌شود، اما این روش بسیار هزینه بر و وقت‌گیر است. بنابراین، به دنبال یافتن استراتژی کاربردی هستیم که با هزینه‌ی کم، طول گام مورد نظر را تعیین کند. راهکاری که در روش‌های جستجوی خطی بکار می‌رود امتحان کردن دنباله‌ای از مقادیر کاندید برای طول گام در شرایط مشخص که در صورت صدق کردن در آن شرط توقف کرده و به این صورت طول گام مورد نظر بدست می‌آید. یکی از این

[8] Broyden, Fletcher, Goldfarb, Shanno

شرایط که کاهش تابع هدف را در الگوریتم‌های جستجوی خطی تضمین می‌کند شرط آرمیجو [۹] است که به صورت زیر بیان می‌شود،

$$f(x_k + \alpha p_k) \leq f(x_k) + c_۱ \alpha \nabla f_k^T p_k,$$

که در آن ثابت $c_۱ \in (۰, ۱)$ . به عبارت دیگر، کاهش تابع $f$ به دو عامل طول گام $\alpha$ و مشتق سویی $\nabla f_k^T p_k$ بستگی دارد.

**تعریف ۱۰.۳.۱.** فرض کنید، داشته باشیم

$$\psi(y) = \inf_{x \in C} \{ f(x) + \sum_{j=۱}^{m} y_j g_j(x) \}.$$

در این صورت، مسأله‌ی

$$\begin{array}{lll} \textbf{(LD)} & \sup & \psi(y) \\ & y & \geq ۰, \end{array}$$

را **دوگان لاگرانژ** [۱۰] از مسأله بهینه سازی محدب $(CO)$ می‌نامند.

**تعریف ۱۱.۳.۱.** اگر $\bar{x}$ یک نقطه‌ی شدنی از مسأله بهینه سازی محدب با تابع هدف $f(x)$ باشد و $\bar{y} \geq ۰$ ، آن‌گاه مقدار

$$f(\bar{x}) - \psi(\bar{y})$$

**شکاف دوگان** [۱۱] در $\bar{x}$ و $\bar{y}$ گویند.

**قضیه ۱۲.۳.۱.** [۱۱] *اگر $\bar{x}$ یک نقطه‌ی شدنی از مسأله‌ی $(CO)$ باشد و $\bar{y} \geq ۰$ و داشته باشیم*

$$\psi(\bar{y}) = f(\bar{x})$$

*، آن‌گاه بردار $\bar{x}$ جواب بهینه‌ی مسأله‌ی $(CO)$ و $\bar{y}$ بهینه برای $(LD)$ خواهد بود.*

---

[۹] Armijo condition

[۱۰] Lagrange dual

[۱۱] duality gap