دانشگاه قم

دانشکده علوم پایه

پایان‌نامه دوره کارشناسی ارشد رشته ریاضی کاربردی

عنوان:

الگوریتم بهینه‌سازی مبتنی بر برنامه‌ریزی خطی برای حل مسائل برنامه‌ریزی غیرخطی

استاد راهنما:

دکترعلی‌اصغر فروغی

استاد مشاور:

دکترغلامحسن شیردل

نگارنده:

طیبه مقدسی‌نیا

تابستان / ۸۹

تقدیم به:

تمام سرمایه‌ام

پدر و مادرمهربانم

که با محبت‌های بی‌دریغشان همواره حامی من هستند.

لم یشکرالمخلوق لم یشکرالخالق

تشکر و قدردانی

حمد و سپاس خدای راکه توفیق کسب دانش و معرفت را به ما عطا فرمود. در اینجا بر خود لازم می‌دانم از تمامی اساتید بزرگوار، به ویژه اساتید دوره کارشناسی ارشد که مرا در تحصیل علم و معرفت و فضائل اخلاقی یاری نموده‌اند تقدیر و تشکر نمایم.

از استاد گرامی و بزرگوار جناب آقای **دکتر علی‌اصغر فروغی** که راهنمایی اینجانب را در انجام تحقیق، پژوهش و نگارش این پایان‌نامه تقبل نموده‌اند نهایت تشکر و سپاسگزاری را دارم.

از جناب آقای **دکتر غلامحسن شیردل** به عنوان مشاور که با راهنمایی خود مرا مورد لطف قرار داده‌اند کمال تشکر را دارم.

از اساتید گرامی جناب آقای **دکتر غلامرضا جهانشاهلو** و جناب آقای **دکتر مهدی احمدی‌نیا** که زحمت داوری این پایان‌نامه را متقبل شده‌اند، صمیمانه سپاسگزارم.

در پایان از پدر و مادر مهربانم و تمامی دوستان عزیزم که با محبت‌های خود مرا مورد لطف قرار داده‌اند سپاسگزارم و برای آنها آرزوی سلامتی و موفقیت روزافزون در تمام مراحل زندگی را دارم.

چکیده:

برنامه‌ریزی غیرخطی فایده و مزیت خود را به عنوان مدلی مهم برای بسیاری از پدیده‌های اقتصادی نشان داده است. گسترش دائماً فزاینده‌ی کاربردهای آن نشان‌دهنده‌ی اهمیت برنامه‌ریزی غیرخطی به عنوان یک قالب کلی برای فرمول‌بندی مسائل است. از آنجایی که نظریه‌ی مسائل خطی، پربارتر و محاسبات آن ساده‌تر از مسائل غیرخطی است، لذا استفاده از مسائل خطی در حل مسائل غیرخطی مفید به‌نظر می‌رسد. در این تحقیق یک الگوریتم بهینه سازی مبتنی بر برنامه‌ریزی خطی، برای حل مسائل برنامه‌ریزی غیرخطی که الگوریتم صفحه‌ی برشی دنباله‌ای نامیده می‌شود، ارائه می‌شود. در ابتدا چند الگوریتم مبتنی بر صفحه‌برش ارائه می‌شود تا خواننده آشنایی مختصری با نحوه‌ی عملکرد الگوریتم پیشنهادی پیدا کند، سپس الگوریتم پیشنهادی روی مسائل NLP با محدودیت‌های نامساوی ارائه و به مسائل NLP با محدودیت‌های مساوی و نامساوی بسط داده می‌شود. حل مثال‌های عددی نشان می‌دهد که این راه‌حل با چند راه‌حل موجود برای حل مسائل NLP قابل رقابت است، و لذا روشی خوب برای حل کارآمد مسائل برنامه‌ریزی غیرخطی به حساب می‌آید.

واژه‌های کلیدی : برنامه‌ریزی غیرخطی ـالگوریتم صفحه‌ی برشی ـبرنامه‌ریزی خطی دنباله‌ای.

فهرست مطالب

فهرست شکل ها

فهرست علائم و اختصارات (Abreviations)

arg	Argument
BFGS	Broyden-Fletcher-Goldfarb-Shanno
card	Cardinal
dim	Dimension
ECP	Extended Cutting Plane
EMFCQ	Extended Mangasarian-Fromovitz constraint qualification
KKT	Karush-Kuhn-Tucker
LP	Linear programming
max	Maximize
min	Minimize
MINLP	Mixed-integer nonlinear programming
NLP	Nonlinear programming
SCP	Sequential cutting plane
SQP	Sequential quadratic programming
s.t	subject to

مقدمه

تصمیم‌گیرندگان چه در عرصه‌های مختلف مدیریت و چه در عرصه‌های زندگی اجتماعی و خانوادگی، همواره با مسائل عادی، بحرانی و فرصت‌های بسیاری روبه‌رو می‌شوند. در گذشته، به علت پیشرفت ناکافی اغلب علوم، برای حل مسائل مختلف، عموماً راه‌حل‌های بسیاری پذیرفتنی بودند. امّا، امروزه با پیشرفت قابل ملاحظه‌ی علوم نسبت به دهه‌های گذشته، دیگر نمی‌توان به مسائل علمی مدیریت در عرصه‌های بازرگانی و فناوری بی‌اعتنایی کرد.

اساساً برنامه‌ریزی غیرخطی که مترادف با بهینه‌سازی غیرخطی است، مسائلی را در بر می‌گیرد که متغیرهای آنها به صورت روابط غیرخطی مطرح می‌شوند. بهینه‌سازی در واقع، به حداکثر یا حداقل رساندن مقدار چند تابع هدف یا چند مقدار تابع هدف با محدودیت و یا بدون محدودیت است.

تاریخ برنامه‌ریزی غیرخطی و روش‌های بهینه‌یابی را می‌توان به روزگار ریاضیدانان گذشته از جمله بیرونی، خیام و حتی نیوتن[۱] ، لاگرانژ[۲] و کوشی [۳] مرتبط دانست. در واقع، توسعه روش‌های بهینه‌سازی با استفاده از حساب دیفرانسیل مدیون کارهای نیوتن و لایبنیتز[۴] است. لاگرانژ روش بهینه‌سازی برای توابع هدف با محدودیت را طرح نمود و کوشی برای نخستین بار روش تندترین کاهش را در حل مسائل حداقل‌سازی بدون محدودیت به کار برد. به هر حال تحول اساسی در بهینه‌سازی غیرخطی در قرن بیستم رخ داد و ظهور رایانه به این تحولات و پیشرفت‌ها سرعت بسیار بخشید. جایگاه برنامه‌ریزی غیرخطی در ارائه الگوهای تصمیم‌گیری جهت ارتقاء مهارت‌ها و توانمندی‌های مدیران برای تأمین بهینه تقاضاها مشخص می‌شود. لذا به دلیل اهمیت موضوع، همواره ریاضیدانان به دنبال ارائه راه‌حل بهتر نسبت به راه‌حل‌های قبلی

---

[۱] Newton

[۲] Lagrange

[۳] Cauchy

[۴] Leibnitz

برای حل مسائل NLP هستند.

یکی از روش‌های حل مسائل NLP، الگوریتم‌های صفحه‌ی برشی است، که این الگوریتم‌ها یک رشته برنامه‌ی خطی تقریبی اصلاحی به دست می‌دهند که جواب‌هایشان به جواب مسأله اصلی همگرا هستند.

روش صفحه‌ی برشی کِلی[۵] برای حل مسائل برنامه‌ریزی غیرخطی (NLP) با حل یک دنباله‌ای از مسائل برنامه‌ریزی خطی (LP) در سال ۱۹۶۰ ارائه شد. تکنیک‌های صفحه‌برشی برای حل مسائل برنامه‌ریزی غیرخطی صحیح توأم ($MINLP$) با موفقیت به‌کار می‌روند. الگوریتم صفحه‌برشی تعمیم‌یافته (ECP) برای حل مسائل $MINLP$ در [۱۸] ارائه شده است. بسط الگوریتم به مسائل با محدودیت‌های محدب‌نما و تابع هدف محدب در [۱۶] و [۱۹] صورت گرفته و در ادامه الگوریتم برای حل مسائل با توابع هدف محدب‌نما گسترش یافته که در [۱۱] ارائه شده است. اگر چه روش‌های دیگری نیز مبتنی بر برنامه‌ریزی خطی از قبیل روش برنامه‌ریزی تقریبی وجود دارد [۶]، ولی با ارائه روش برنامه ریزی درجه دوم دنباله‌ای (SQP)، تکنیک‌های LP کمتر مورد توجه قرار گرفت. در این پایان‌نامه که مطالب اساسی آن برگرفته از مراجع [۱] ،[۱۴] و [۱۵] است، نشان داده می‌شود که تکنیک‌های LP می‌توانند به طور کاملاً موفق برای حل مسائل NLP به‌کار روند و نسبت به روش‌هایی که بر SQP مبتنی هستند نیز برتری دارند.

[۵] Kelley

فصل ۱

تعاریف و پیش‌نیازها

۱.۱ مفاهیم مقدماتی

تعریف ۱.۱.۱: دنباله $\{x^k\}_{k=0}^{\infty}$ را همگرا به $\bar{x}$ گوییم اگر هنگامی که $k \to \infty$، داشته باشیم:

$$|x^k - \bar{x}| \to 0$$

یا به عبارت دیگر اگر به ازای مقدار داده‌شده‌ی $\epsilon > 0$، عددی مانند N وجود داشته باشد به‌طوری‌که از $k \geq N$ نتیجه شود $|x^k - \bar{x}| < \epsilon$، گوییم $\{x^k\}_{k=0}^{\infty}$ به $\bar{x}$ همگراست و می‌نویسیم

$$x^k \to \bar{x}$$

تعریف ۲.۱.۱: نقطه‌ای چون $\bar{x}$، یک نقطه‌ی حدی دنباله‌ی $\{x^k\}$ است اگر زیردنباله‌ای از $\{x^k\}$ وجود داشته باشد که همگرا به $\bar{x}$ باشد.

تعریف ۳.۱.۱: تابعی حقیقی مقدار چون f، تعریف شده بر زیرمجموعه‌ای از $\mathbb{R}^n$، در نقطه‌ی $\bar{x}$ پیوسته نامیده می‌شود اگر از $x^k \to \bar{x}$ نتیجه شود که $f(x^k) \to f(\bar{x})$.

تعریف ۴.۱.۱: مشتق تابع $f : \mathbb{R}^n \to \mathbb{R}$ در طول یک بردار داده‌شده‌ی d در نقطه‌ی x را مشتق جهتی f در جهت d می‌نامیم و آن را به صورت زیر نمایش می‌دهیم:

$$D_d f(x) = \nabla f(x)^T d$$

تعریف ۵.۱.۱: در مورد توابع چند متغیره، به مشتق تابع نسبت به یکی از متغیرها با ثابت نگه‌داشتن سایر متغیرها مشتق جزئی تابع گفته می‌شود.

تذکر ۶.۱.۱: اگر تابع حقیقی مقدار f دارای مشتقات جزئی مرتبه pام پیوسته روی زیرمجموعه بازی از $\mathbb{R}^n$ باشد می‌نویسیم: $f \in C^p$.

تعریف ۷.۱.۱: اگر $f \in C^1$ یک تابع حقیقی مقدار بر $\mathbb{R}^n$ باشد، و داشته باشیم $f(x) = f(x_1, x_2, \cdots, x_n)$، گرادیان f را به صورت زیر تعریف می‌کنیم:

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n}\right]^T$$

تعریف ۸.۱.۱: اگر $f \in C^2$، آنگاه هسی f در نقطه‌ی x را به عنوان ماتریسی $n \times n$ تعریف می‌کنیم که با $\nabla^2 f(x)$ یا $F(x)$ نشان داده می‌شود و عبارت است از:

$$F(x) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right]_{n \times n}$$

و چون برای هر i و j، $(i, j = 1, \cdots, n)$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

به سادگی دیده می‌شود که ماتریس هسی متقارن است.

تعریف ۹.۱.۱: نقطه‌ی $x^* \in S$ را نقطه‌ی مینیمم موضعی تابع f روی S گوییم اگر یک $\epsilon > 0$ وجود داشته باشد به‌طوری‌که به ازای هر $x \in S$ در فاصله حداکثر ϵ از x^* (یعنی $x \in S$ و $|x - x^*| < \epsilon$) داشته باشیم $f(x) \geq f(x^*)$.

تعریف ۱۰.۱.۱: برای یک ماتریس متقارن Q، بردارهای d_1 و d_2 را Q ـمتعامد یا مزدوج نسبت به Q، گوییم اگر $d_1^T Q d_2 = 0$.

تعریف ۱.۱.۱۱: اگر $Q_{n\times n}$ ماتریسی با درایه‌های حقیقی باشد، Q را معین مثبت گوییم هرگاه متقارن باشد و به ازای هر $x \in \mathbb{R}^n$ که $x \neq 0$ داشته باشیم $x^TQx > 0$.

تعریف ۱.۱.۱۲: تابع $f : \mathbb{R} \to \mathbb{R}$ یکنوای نزولی نامیده می‌شود اگر برای هر x و y که $x \leq y$ داشته باشیم $f(x) \geq f(y)$.

تذکر ۱.۱.۱۳: اگر g تابعی حقیقی از یک متغیر حقیقی باشد، نماد $g(x) = O(x)$ به معنای آن است که $g(x)$ حداقل به سرعت x به سمت صفر میل می‌کند. به بیان دقیقتر، این تعریف بدان معناست که $K \geq 0$ وجود دارد به‌طوری‌که

$$\left|\frac{g(x)}{x}\right| \leq K \qquad x \to 0 \quad \text{وقتی}$$

نماد $g(x) = o(x)$ به این معنی است که $g(x)$ سریعتر از x به سمت صفر میل می‌کند، یا به بیان دیگر، K در بالا برابر صفر است.

۱.۲ مفاهیم اساسی

تعریف ۱.۲.۱: هر الگوریتم A نگاشتی است تعریف شده بر فضای X که به هر نقطه‌ی $x \in X$ زیرمجموعه‌ای از X را منسوب می‌کند.

تعریف ۱.۲.۲: هر الگوریتم تکراری با مشخص کردن یک نقطه‌ی آغازی شروع می‌شود. اگر برای هر نقطه‌ی دلخواه آغازی، الگوریتم همواره دنباله‌ای از نقاط را تولید کند که همگرا به یک جواب شود، آنگاه الگوریتم را همگرای سراسری نامند.

تعریف ۱.۲.۳: فرض کنید x^* نقطه‌ای است که در قیود زیر صدق می‌کند

$$h(x^*) = 0\ , \qquad g(x^*) \leq 0$$

که $g:\mathbb{R}^n\to\mathbb{R}^{m_۱}$ ، $h:\mathbb{R}^n\to\mathbb{R}^{m_۲}$ و $h,g\in C^۱$. همچنین فرض کنید J مجموعه‌ی اندیس‌های j است به‌طوری‌که $g_j(x^*)=۰$. در این صورت x^* را یک نقطه‌ی منتظم برای قیود فوق می‌نامیم اگر بردارهای گرادیان $\nabla h_i(x^*)$، $\nabla g_j(x^*)$، $۱\leq i\leq m_۲$، $j\in J$ مستقل خطی باشند.

تعریف ۴.۲.۱: مجموعه Ω در $\mathbb{R}^n$ محدب گفته می‌شود اگر به ازای هر $x_۱,x_۲\in\Omega$ و هر عدد حقیقی λ، $۰<\lambda<۱$، نقطه‌ی $\lambda x_۱+(۱-\lambda)x_۲$ متعلق به Ω باشد.
این تعریف را می‌توان به طریق هندسی چنین تعبیر کرد: یک مجموعه، محدب است اگر با انتخاب دو نقطه‌ی دلخواه در مجموعه، هر نقطه روی خط واصل بین آن دو نقطه، عضوی از مجموعه باشد.

تعریف ۵.۲.۱: تابع $f:\Omega\to\mathbb{R}$ که روی یک زیرمجموعه محدب Ω از فضای $\mathbb{R}^n$ تعریف شده است، محدب است اگر برای هر دو نقطه‌ی $x_۱,x_۲\in\Omega$ و $\lambda\in[۰,۱]$، داشته باشیم:

$$f(\lambda x_۱+(۱-\lambda)x_۲)\leq\lambda f(x_۱)+(۱-\lambda)f(x_۲).$$

تعریف ۶.۲.۱: تابع $f:\Omega\to\mathbb{R}$ که روی یک زیرمجموعه محدب Ω از فضای $\mathbb{R}^n$ تعریف شده است، شبه‌محدب است اگر برای هر دو نقطه‌ی $x_۱,x_۲\in\Omega$ و $\lambda\in[۰,۱]$، داشته باشیم:

$$f(\lambda x_۱+(۱-\lambda)x_۲)\leq\max\{f(x_۱),f(x_۲)\}.$$

تعریف ۷.۲.۱: تابع $f:\Omega\to\mathbb{R}$ که روی یک زیرمجموعه محدب Ω از فضای $\mathbb{R}^n$ تعریف شده است، محدب‌نما است اگربرای هر دو نقطه‌ی $x_۱,x_۲\in\Omega$ با شرط $\nabla f(x_۱)^T(x_۲-x_۱)\geq ۰$ داشته باشیم:

$$f(x_۲)\geq f(x_۱).$$

تعریف ۸.۲.۱: تابع f شبه‌مقعر است اگر $-f$ شبه‌محدب باشد.

تعریف ۹.۲.۱: فرض کنید a برداری غیر صفر در $\mathbb{R}^n$ و c عددی حقیقی باشد. متناظر با ابرصفحه‌ی $H = \{x : a^T x = c\}$ نیم‌فضاهای بسته‌ی مثبت و منفی زیر وجود دارند

$$H_+ = \{x : a^T x \geq c\} \ ,$$

$$H_- = \{x : a^T x \leq c\}.$$

تعریف ۱۰.۲.۱: مجموعه‌ای که قابل بیان به صورت اشتراک تعدادی متناهی از نیم‌فضاهای بسته باشد، بس‌وجهی محدب نامیده می‌شود.

تعریف ۱۱.۲.۱: ابرصفحه‌ای که شامل یک مجموعه محدب Ω در یکی از نیم‌فضاهای بسته‌اش و مشتمل بر یک نقطه‌ی مرزی Ω است، ابرصفحه‌ی محمل Ω نامیده می‌شود.

تعریف ۱۲.۲.۱: مسأله خطی در فرم استاندارد زیر را در نظر بگیرید:

$$\begin{aligned} P: \quad & \min \quad c^T x \\ & s.t. \quad Bx = b \\ & \qquad\quad x \geq ۰ \end{aligned}$$

این مسأله را مسأله اولیه P می‌نامیم. مسأله خطی زیر، دوگان مسأله اولیه نامیده می‌شود.

$$\begin{aligned} D: \quad & \max \quad b^T y \\ & s.t. \quad B^T y \leq c \\ & \qquad\quad y \ \text{نامقید} \end{aligned}$$

قضیه ۱۳.۲.۱: مسائل P و D تعریف شده در بالا را در نظر بگیرید. داریم:

(۱). قضیه دوگانی ضعیف: $c^T x \geq b^T y$ برای هر جواب شدنی x برای P و هر جواب شدنی y برای D.

(۲). روابط نشدنی ـبی‌کران: اگر P بی‌کران باشد، D نشدنی است و اگر D بی‌کران باشد، P نشدنی است .

(۳). قضیه دوگانی قوی: اگر هر دو مسأله‌ی P و D شدنی باشند، آنها هر دو جواب‌های بهینه با مقدار هدف یکسان دارند.

اثبات: به مرجع [۲] مراجعه شود.

قضیه ۱۴.۲.۱: (شرایط مکمل کمبود و تشخیص بهینگی)

برنامه‌ی خطی اولیه و دوگان P و D را در نظر بگیرید و فرض کنید $\bar{x}$ جواب شدنی اولیه و $\bar{y}$ جواب شدنی دوگان باشد. پس $\bar{x}$ و $\bar{y}$ به ترتیب جواب‌های بهینه برای P و D هستند اگر و تنها اگر $\bar{v}_j\bar{x}_j = \circ$ برای $j = ۱, \ldots, n$ که $\bar{V} = (\bar{v}_۱, \ldots, \bar{v}_n)^T = c - B^T\bar{y}$، بردار متغیرهای کمبود در محدودیت‌های دوگان برای جواب دوگان $\bar{y}$ است.

اثبات: به مرجع [۲] مراجعه شود.

قضیه ۱۵.۲.۱: فرض کنید که $f : \mathbb{R}^n \to \mathbb{R}$ در $\bar{x}$ مشتق‌پذیر باشد. اگر بردار d موجود باشد که $\nabla f(\bar{x})^T d < \circ$، پس $\delta > \circ$ موجود است به‌طوری‌که $f(\bar{x} + \lambda d) < f(\bar{x})$ برای هر $\lambda \in (\circ, \delta)$ و بنابراین d یک جهت نزولی برای f در $\bar{x}$ است.

اثبات: با توجه به مشتق‌پذیر بودن f در $\bar{x}$ داریم:

$$f(\bar{x} + \lambda d) = f(\bar{x}) + \lambda \nabla f(\bar{x})^T d + \lambda \|d\| \alpha(\bar{x}; \lambda d)$$

که $\alpha(\bar{x}; \lambda d) \to \circ$ ، هنگامی‌که $\lambda \to \circ$. با تقسیم بر $\lambda \neq \circ$ داریم:

$$\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} = \nabla f(\bar{x})^T d + \|d\| \alpha(\bar{x}; \lambda d)$$

چون $\nabla f(\bar{x})^T d < \circ$ و $\alpha(\bar{x}; \lambda d) \to \circ$ هنگامی‌که $\lambda \to \circ$، پس $\delta > \circ$ وجود دارد به‌طوری‌که برای همه $\lambda \in (\circ, \delta)$

$$\nabla f(\bar{x})^T d + \|d\| \alpha(\bar{x}; \lambda d) < \circ$$

و بنابراین برای همه $\lambda \in (0, \delta)$

$$f(\bar{x} + \lambda d) - f(\bar{x}) < 0$$

و لذا نتیجه‌ی مورد نظر حاصل شد.□

فرض کنید S یک مجموعه‌ی باز ناتهی در $\mathbb{R}^n$ باشد و $f : \mathbb{R}^n \to \mathbb{R}$ ، $g_i : \mathbb{R}^n \to \mathbb{R}$ برای $i = 1, \ldots, m_1$ و $h_j : \mathbb{R}^n \to \mathbb{R}$ برای $j = 1, \ldots, m_2$. مسأله P را به صورت زیر در نظر بگیرید.

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & g_i(x) \le 0 \ , \qquad i = 1, \ldots, m_1 \\ & h_j(x) = 0 \ , \qquad j = 1, \ldots, m_2 \\ & x \in S \end{aligned} \tag{1.1}$$

قضیه ۱.۲.۱۶: مسأله‌ی P تعریف شده در(۱.۱) را در نظر بگیرید. فرض کنید $\bar{x}$ یک جواب شدنی باشد و $I = \{i : \ g_i(\bar{x}) = 0\}$. علاوه‌براین فرض کنید که g_i برای هر $i \notin I$ در $\bar{x}$ پیوسته، f و g_i برای $i \in I$ در $\bar{x}$ مشتق‌پذیر، و اینکه h_j برای هر $j = 1, \ldots, m_2$ در $\bar{x}$ به‌طور پیوسته مشتق‌پذیرند. اگر $\bar{x}$ مسأله P را به‌طور موضعی حل کند، اسکالرهای u_0 و u_i برای $i \in I$ و v_j برای $j = 1, \ldots, m_2$ موجودند به‌طوری‌که

$$u_0 \nabla f(\bar{x}) + \sum_{i \in I} u_i \nabla g_i(\bar{x}) + \sum_{j=1}^{m_2} v_j \nabla h_j(\bar{x}) = 0 \ ,$$

$$u_0, u_i \ge 0 \ , \qquad i \in I$$

$$(u_0, U_I, V) \ne (0, 0, 0)$$

که U_I برداری است با مؤلفه‌های u_i که $i \in I$، و $V = (v_1, \ldots, v_{m_2})^T$.

اثبات: به مرجع [۲] مراجعه شود.

قضیه ۱.۲.۱۷: (شرایط لازم مرتبه اول KKT)

مسأله P تعریف شده در (۱.۱) را در نظر بگیرید. فرض کنید $\bar{x}$ یک جواب شدنی باشد

و $I = \{i:\ g_i(\bar{x}) = ۰\}$. فرض کنید f و g_i برای $i \in I$ در $\bar{x}$ مشتق‌پذیر و هر g_i برای $i \notin I$ در $\bar{x}$ پیوسته است و اینکه هر h_j برای $j = ۱,\ldots,m_۲$ در $\bar{x}$ به‌طور پیوسته مشتق‌پذیر است. علاوه‌براین فرض کنید که $\nabla g_i(\bar{x})$، $\nabla h_j(\bar{x})$ برای $i \in I$ ، $j = ۱,\ldots,m_۲$ مستقل خطی‌اند (به‌طوری‌که $\bar{x}$، منتظم نامیده می‌شود).

اگر $\bar{x}$ مسأله را به طور موضعی حل کند، اسکالرهای منحصربه‌فرد λ_i برای $i \in I$ و μ_j برای $j = ۱,\ldots,m_۲$ موجودند به‌طوری‌که

$$
\begin{aligned}
&\nabla f(\bar{x}) + \sum_{i\in I} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=۱}^{m_۲} \mu_j \nabla h_j(\bar{x}) = ۰ \ , \\
&\lambda_i \geq ۰ \ , \qquad\qquad i \in I
\end{aligned}
$$

به‌علاوه اگر هر g_i برای $i \notin I$ در $\bar{x}$ مشتق‌پذیر باشد شرایط KKT می‌تواند به فرم معادل زیر نوشته شود.

$$
\begin{aligned}
&\nabla f(\bar{x}) + \sum_{i=۱}^{m_۱} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=۱}^{m_۲} \mu_j \nabla h_j(\bar{x}) = ۰ \ , \\
&\lambda_i g_i(\bar{x}) = ۰ \ , \qquad\qquad i = ۱,\ldots,m_۱ \\
&\lambda_i \geq ۰ \ , \qquad\qquad i = ۱,\ldots,m_۱
\end{aligned}
$$

اثبات: با توجه به قضیه‌ی قبل اسکالرهای $u_۰$ و u_i برای $i \in I$ و v_j برای $j = ۱,\ldots,m_۲$ که همگی صفر نیستند وجود دارند به‌طوری‌که

$$
\begin{aligned}
&u_۰ \nabla f(\bar{x}) + \sum_{i\in I} u_i \nabla g_i(\bar{x}) + \sum_{j=۱}^{m_۲} v_j \nabla h_j(\bar{x}) = ۰ \ , \\
&u_۰, u_i \geq ۰ \ , \qquad\qquad i \in I
\end{aligned}
\tag{۲.۱}
$$

نشان می‌دهیم که $u_۰ > ۰$، زیرا اگر $u_۰ = ۰$، (۲.۱) با فرض مستقل خطی بودن $\nabla g_i(\bar{x})$، $\nabla h_j(\bar{x})$ برای $i \in I$ ، $j = ۱,\ldots,m_۲$ متناقض است. بنابراین نتیجه‌ی اول با قراردادن $\lambda_i = \frac{u_i}{u_۰}$ برای $i \in I$ و $\mu_j = \frac{v_j}{u_۰}$ برای $j = ۱,\ldots,m_۲$ به دست می‌آید، و فرض مستقل خطی بودن، منحصربه‌فرد بودن این ضرایب لاگرانژ را نتیجه می‌دهد. همچنین فرم معادل شرایط لازم با قرار دادن $\lambda_i = ۰$ برای $i \notin I$ نتیجه می‌شود، و بدین ترتیب اثبات کامل می‌شود.□