



دانشکده علوم ریاضی
گروه ریاضی کاربردی

پایان نامه کارشناسی ارشد

برخی روش های تکراری برای مسائل بهینه سازی چندهدفه

نرگس حسین پور

استاد راهنما

دکتر مهرداد غزنوی

استاد مشاور

دکتر مریم قرآنی

آذر ۹۴

تقدیم

تقدیم به بانوی بی نشان حضرت زهرا

تقدیم به مهربان فرشتگانی که:

خطات ناب باور بودن، لذت و غرور دانستن، حسارت

خواستن، عظمت رسیدن و تمام تجربه های یکتا و زیبای زندگیم،

مدیون حضور سبز آنهاست تقدیم به خانواده عزیزم.

سپاسگزاری

سپاس بی کران پروردگار یکتا را که، هستی مان، بخشد و به طریق علم و دانش، رهنمون مان
شده و به، هم نشینی رهروان علم و دانش مفتخر مان نمود و خوشه چینی از علم و معرفت را
روزی مان ساخت. سلام و درود بر محمد و خاندان پاک او، طاهران معصوم، هم
آنان که وجود مان و امدار وجود شان است.

سپاس بی کران از پدر و مادر عزیزم این دو معلم بزرگوارم که، همواره بر کوتاهی و درستی
من، قلم عفو کشیده و کریمانه از کنار غفلت هایم گذشته اند و در تمام عرصه های زندگی
یار و یاور بی چشم داشت برای من بوده اند.

از استاد کرامیم جناب آقای دکتر غزنوی، به دلیل یاری ها و راهنمایی های بی چشم داشت
ایشان که بسیاری از سختی ها را برایم آسان تر نمودند، بسیار سپاس گذارم چرا که بدون
راهنمایی های ایشان تا این پایان نامه بسیار مشکل می نمود.

همچنین لازم می دانم از استاد مشاورم خانم دکتر قرآنی و داوران کرامی جناب آقای

دکتر جعفر فتحعلی و جناب آقای دکتر علیرضا ناظمی که زحمت مطالعه و داوری این
پایان نامه را پذیرفته اند، تقدیر و تشکر نمایم.

نرگس حسین پور
آذر ۹۴

تعمدنامه

اینجانب نرگس حسین پور دانشجوی کارشناسی ارشد رشته ریاضی کاربردی دانشکده علوم ریاضی دانشگاه صنعتی شاهرود، نویسنده پایان نامه با عنوان برخی روش های تکراری برای مسائل بهینه سازی چندهدفه، تحت راهنمایی دکتر مهرداد غزنوی متعهد می شوم:

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

نرگس حسین پور
آذر ۹۴

مالکیت نتایج و حق نشر

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

چکیده

در این پایان نامه، ابتدا مفاهیم و مقدمات اولیه را معرفی می‌کنیم. سپس مفهوم کارایی و کارایی ضعیف را برای مسائل بهینه‌سازی چندهدفه بیان می‌کنیم. ما به حل مسائل چندهدفه با به‌کارگیری روش‌های تکراری می‌پردازیم. یک روش تکراری، یک روند محاسباتی است که دنباله‌ای تولید می‌کند و انتظار داریم تحت شرایط معقولی به نقطه کارای مساله همگرا شود. از بین روش‌های تکراری برای حل مسائل چندهدفه، روش نیوتن و روش‌های شبه‌نیوتن را مورد بحث قرار می‌دهیم.

در ادامه الگوریتم نیوتن را برای حل مسائل چندهدفه محدب بیان می‌کنیم که نیاز به محاسبه دقیق، ماتریس هسین برای هر تابع هدف دارد. در حالتی که محاسبه هسین مشکل باشد، روش‌های شبه‌نیوتن مطرح می‌شوند که مورد بحث قرار می‌دهیم. این روش‌ها با محاسبه تقریب هسین به جای هسین دقیق به بهبود روند حل مساله کمک می‌کنند. در مواردی که مساله نامحدب باشد، یک الگوریتم شبه‌نیوتن اصلاح شده، پیشنهاد می‌گردد. در این حالت، با محاسبه یک ماتریس تقریبی، مسائل چندهدفه نامحدب حل می‌شوند. در نهایت، یک الگوریتم شبه‌نیوتن برای حل مسائل چندهدفه ناهموار معرفی می‌شود.

کلمات کلیدی: بهینه‌سازی چندهدفه، جواب کارا، جواب کارای ضعیف، نقطه بحرانی، تابع محدب.

لیست مقالات مستخرج از پایان نامه

1. M. Ghaznavi, F. Soleimani, N. Hoseinpoor, Parametric analysis in fuzzy number linear programming problems, International journal of fuzzy systems, 2015, accepted.
2. N. Hoseinpoor, M. Ghaznavi, A Newton-type method for multiobjective optimization problems, The 46th Annual Iranian Mathematics Conference, 2015, Yazd, Iran.

فهرست مطالب

لیست تصاویر	ذ
لیست جداول	۱
۱ مقدمه‌ای بر بهینه‌سازی چندهدفه	۳
۱.۱ مقدمه	۳
۲.۱ تعاریف و قضایای اولیه	۴
۳.۱ مبانی بهینه‌سازی چندهدفه	۱۲
۲ روش نیوتن برای حل مسائل چندهدفه	۱۷
۱.۲ مقدمه	۱۷
۲.۲ روش نیوتن برای حل مسائل تک‌هدفه	۱۷
۳.۲ روش نیوتن برای حل مسائل چندهدفه	۱۸
۴.۲ نتایج تکنیکی	۲۳
۵.۲ همگرایی فوق‌خطی	۲۶
۶.۲ همگرایی درجه دوم	۳۲
۷.۲ مثال‌های عددی	۳۴
۳ یک روش شبه‌نیوتن برای حل مسائل چندهدفه محدب	۳۹
۱.۳ مقدمه	۳۹
۲.۳ روش شبه‌نیوتن برای حل مسائل بهینه‌سازی تک‌هدفه	۳۹
۳.۳ روش شبه‌نیوتن برای حل مسائل چندهدفه	۴۲
۴.۳ نتایج تکنیکی	۴۶
۵.۳ همگرایی	۴۸
۶.۳ مثال‌های عددی	۵۳
۴ روش شبه‌نیوتن اصلاح شده برای حل مسائل چندهدفه نامحدب	۵۷
۱.۴ مقدمه	۵۷

۵۷	مفاهیم اولیه	۲.۴
۵۹	روش شبه نیوتن اصلاح شده	۳.۴
۵۹	انتخاب یک جهت کاهنده مناسب	۱.۳.۴
۶۲	تولید ماتریس مثبت معین $B(x_k)$	۲.۳.۴
۶۴	تولید یک دنباله همگرا	۳.۳.۴
۶۸	مثال‌های عددی	۴.۴
۷۳	۵ حل مسائل بهینه‌سازی چندهدفه ناهموار با روش شبه نیوتن	
۷۳	تعاریف اولیه	۱.۵
۷۴	الگوریتم	۲.۵
۷۹	همگرایی سراسری	۳.۵
۸۱	مثال‌های عددی	۱.۳.۵
۸۵	آ	
۸۹	مراجع	
۹۳	واژه‌نامه فارسی به انگلیسی	
۹۷	واژه‌نامه انگلیسی به فارسی	

لیست تصاویر

۱۱	مجموعه شدنی مثال ۲۸.۲.۱	۱.۱
۱۴	مجموعه شدنی مثال ۴.۳.۱	۲.۱
۱۳	نقاط کارای مثال ۲.۶.۳ بر روی فضای هدف مشخص شده‌اند که طول، عرض و ارتفاع نمودار به ترتیب محور مربوط به F_2 ، F_1 و F_3 هستند.	
۵۵		
۶۹	توابع هدف مثال ۱.۴.۴	۱.۴
۷۰	توابع هدف مثال ۲.۴.۴	۲.۴
۸۲	طول و عرض محور مختصات به ترتیب مربوط به تابع F_2 و F_1 مساله ۳.۳.۵ است.	۱.۵

لیست جداول

۳۵	جدول مربوط به مثال ۱.۷.۲	۱.۲
۳۶	جدول مربوط به مثال ۲.۷.۲	۲.۲
۳۶	جدول مربوط به مثال ۳.۷.۲	۳.۲
۳۷	جدول مربوط به مثال ۴.۷.۲	۴.۲
۵۵	جدول ۱.۶.۳ مربوط به مثال	۱.۳
۵۶	جدول مربوط به زمان محاسبه دو الگوریتم	۲.۳
۷۰	جدول مربوط به مثال ۲.۴.۴	۱.۴
۷۱	جدول مربوط به مثال ۳.۴.۴	۲.۴

فصل ۱

مقدمه‌ای بر بهینه‌سازی چندهدفه

۱.۱ مقدمه

بهینه‌سازی یا برنامه‌ریزی ریاضی، در اقتصاد، ریاضیات، علوم کامپیوتر و مدیریت کاربرد فراوان دارد و به انتخاب بهترین جواب در بین وضعیت‌های موجود اشاره می‌کند. مسائل بهینه‌سازی را می‌توان از دیدگاه‌های متفاوت دسته‌بندی نمود. یکی از این دسته‌بندی‌ها مربوط به بهینه‌سازی چندهدفه است. در این پایان‌نامه ما به مسائل چندهدفه و روش حل آن می‌پردازیم.

در فصل اول تعاریف و قضایای اولیه را مطرح می‌کنیم. همچنین مفاهیم مربوط به بهینه‌سازی محدب و غیرخطی را به‌طور خلاصه بیان می‌کنیم. مساله بهینه‌سازی چندهدفه و تعاریف کارایی و کارایی ضعیف که در ادامه پایان‌نامه مورد استفاده قرار می‌گیرند را بیان می‌کنیم. اکثر مطالب این فصل از مراجع [۲۳، ۱۶، ۴] اتخاذ شده‌اند.

در فصل دوم با این فرض که محاسبه هسین به آسانی صورت می‌گیرد و تابع هدف محدب است، جهت نیوتن برای حل مسائل چندهدفه را مطرح می‌کنیم. در ادامه الگوریتم نیوتن و همگرایی آن را بیان می‌کنیم. در پایان مثال‌های عددی مربوط به الگوریتم را مطرح می‌کنیم. در این فصل بیشتر از مرجع [۱۸] استفاده شده است.

در فصل سوم با این فرض که محاسبه هسین غیرعملی است، از ماتریس تقریب هسین استفاده می‌کنیم و الگوریتم شبه‌نیوتن را برای حل مساله چندهدفه مطرح می‌کنیم. در این الگوریتم برای بروز کردن تقریب هسین از روش $BFGS$ استفاده می‌کنیم و تقریب هسین برای هر تابع هدف بروز می‌شود. در فصل چهارم الگوریتم شبه‌نیوتن اصلاح شده را برای حل مساله چندهدفه مطرح می‌کنیم. در این الگوریتم تابع هدف لزوماً محدب نیست و نیاز به محاسبه هسین برای هر تابع هدف نیست. در فصل پنجم چگونگی حل مسائل چندهدفه ناهموار را به کمک الگوریتم شبه‌نیوتن مطرح می‌کنیم و همگرایی آن را اثبات می‌کنیم. اکثر مطالب این سه فصل از مراجع [۳۳، ۳۱، ۲۱، ۱۱، ۲] اتخاذ شده‌اند.

۲.۱ تعاریف و قضایای اولیه

ابتدا برخی نمادگذاری‌های مورد نیاز برای ادامه پایان‌نامه را بیان می‌کنیم. مجموعه اعداد حقیقی را با $\mathbb{R}$ ، مجموعه اعداد حقیقی نامنفی^۱ را با $\mathbb{R}_+$ و مجموعه اعداد حقیقی مثبت اکید^۲ را با $\mathbb{R}_{++}$ نشان می‌دهیم. بدین ترتیب برای هر $x, y \in \mathbb{R}^n$ روابط زیر را داریم:

$$x < y \longleftrightarrow x_i < y_i, \quad \forall i = 1, 2, \dots, n.$$

$$x \leq y \longleftrightarrow x_i \leq y_i, \quad \forall i = 1, 2, \dots, n.$$

$$x \leq y \longleftrightarrow x_i \leq y_i, \quad \forall i = 1, 2, \dots, n \quad \text{و} \quad x_j < y_j, \quad \exists j \in \{1, 2, \dots, n\}.$$

$$\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x \geq 0\},$$

$$\mathbb{R}_{++}^n = \{x \in \mathbb{R}^n : x > 0\}.$$

در نتیجه داریم:

$$x \leq y \longleftrightarrow y - x \in \mathbb{R}_+^n,$$

$$x < y \longleftrightarrow y - x \in \mathbb{R}_{++}^n.$$

نرم اقلیدسی^۳ در $\mathbb{R}^n$ با $\|\cdot\|$ و درون^۴ یک مجموعه مانند E را با $\text{int}(E)$ نشان می‌دهیم. $\arg \min f(x)$ نقطه‌ای است که بهینگی تابع $f(x)$ در آن اتفاق می‌افتد. نرم یک ماتریس $A \in \mathbb{R}^{n \times n}$ را به صورت زیر در نظر می‌گیریم:

$$\|A\| = \max_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|}, \quad x \neq 0.$$

برای دو ماتریس $A, B \in \mathbb{R}^{n \times n}$ ، $B < A$ (به این معنی است که $A - B$ مثبت معین^۵ نیمه معین) است یعنی:

$$\forall x \neq 0 \in \mathbb{R}^n, \quad x^T(A - B)x > 0.$$

در واقع برای ماتریس A ، رابطه $A > 0$ به این معنی است که A مثبت معین است. برای یک ماتریس $M \in \mathbb{R}^{m \times n}$ برد^۶ یا فضای تصویر^۷ آن را با $R(M)$ نشان می‌دهیم که به صورت زیر تعریف می‌شود:

تعریف ۱.۲.۱. برد یا فضای تصویر یک ماتریس $M \in \mathbb{R}^{m \times n}$ ، ترکیب خطی از ستون‌های ماتریس M است.

^۱non-negative

^۲strictly positive

^۳Euclidean norm

^۴interior

^۵positive definite

^۶range

^۷image space

همچنین از نماد $F \in C^2(U, \mathbb{R}^m)$ استفاده می‌کنیم، بدین معنی که تابع $F : U \rightarrow \mathbb{R}^m$ دوبار به‌طور پیوسته مشتق‌پذیر است که $F(x) = (F_1(x), F_2(x), \dots, F_m(x))$ است. برای $x \in U \subset \mathbb{R}^n$ ژاکوبین^۸ تابع F در x را با $DF(x) \in \mathbb{R}^{m \times n}$ ، گرادیان^۹ تابع F_j در x را با $\nabla F_j(x) \in \mathbb{R}^n$ و هسین^{۱۰} تابع F_j در x را با $\nabla^2 F_j(x) \in \mathbb{R}^{n \times n}$ نشان می‌دهیم، در واقع داریم:

$$DF(x) = \begin{bmatrix} \frac{\partial F_1(x)}{\partial x_1} & \frac{\partial F_1(x)}{\partial x_2} & \dots & \frac{\partial F_1(x)}{\partial x_n} \\ \frac{\partial F_2(x)}{\partial x_1} & \frac{\partial F_2(x)}{\partial x_2} & \dots & \frac{\partial F_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m(x)}{\partial x_1} & \frac{\partial F_m(x)}{\partial x_2} & \dots & \frac{\partial F_m(x)}{\partial x_n} \end{bmatrix},$$

$$\nabla F_j(x) = \begin{pmatrix} \frac{\partial F_j(x)}{\partial x_1} \\ \frac{\partial F_j(x)}{\partial x_2} \\ \vdots \\ \frac{\partial F_j(x)}{\partial x_n} \end{pmatrix},$$

$$\nabla^2 F_j(x) = \begin{bmatrix} \frac{\partial^2 F_j(x)}{\partial x_1^2} & \frac{\partial^2 F_j(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 F_j(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F_j(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 F_j(x)}{\partial x_2^2} & \dots & \frac{\partial^2 F_j(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F_j(x)}{\partial x_n \partial x_1} & \frac{\partial^2 F_j(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 F_j(x)}{\partial x_n^2} \end{bmatrix}.$$

تعریف ۲.۲.۱. تابع $F : U \rightarrow \mathbb{R}^m$ را در نظر بگیرید. که $F(x) = (F_1(x), F_2(x), \dots, F_m(x))$ است. مشتق جهتی تابع $F_j : \mathbb{R}^n \rightarrow \mathbb{R}$ در جهت d و در نقطه x برای هر $j = 1, 2, \dots, m$ ، به‌صورت زیر تعریف می‌شود:

$$F'_j(x, d) = \lim_{t \rightarrow 0} \frac{F_j(x + td) - F_j(x)}{t}.$$

اگر مشتقات جهتی تابع F_j در نقطه x موجود و پیوسته باشند، داریم:

$$F'_j(x, d) = \nabla F_j(x)^T d.$$

توجه شود $d \in \mathbb{R}^n$ یک جهت کاهنده برای تابع F در x است، اگر برای هر $j = 1, 2, \dots, m$ ، $F'_j(x, d) < 0$ باشد. بنابراین d یک جهت کاهنده، در x است اگر $t_* > 0$ وجود داشته باشد به‌طوری‌که برای هر $t \in (0, t_*]$ ، $F(x + td) < F(x)$.

تعریف ۳.۲.۱. فرض کنید X یک مجموعه ناتهی باشد. تابع $d : X \times X \rightarrow \mathbb{R}$ را یک متر (متریک) روی X گویند هرگاه ۴ خاصیت زیر برای آن برقرار باشد:

^۸Jacobian

^۹gradient

^{۱۰}Hessian

$$۱. \quad d(x, y) \geq 0, \forall x, y \in X$$

$$۲. \quad d(x, y) = d(y, x)$$

$$۳. \quad d(x, y) = 0 \iff x = y$$

$$۴. \quad d(x, y) \leq d(x, z) + d(z, y), \forall x, y, z \in X$$

همچنین (X, d) نشان‌دهنده فضای متریک X با متر d است.

تعریف ۴.۲.۱. فرض کنید، (X, d) یک فضای متریک باشد و $x \in X$ و $r > 0$ باشد، گوی بازی به مرکز x و شعاع r را با $B(x, r)$ نمایش می‌دهیم و تعریف می‌کنیم:
 $B(x, r) = \{y \in X | d(x, y) < r\}$.

تعریف ۵.۲.۱. دنباله $\{x_n\} \subset (X, d)$ را دنباله کوشی^{۱۱} گوییم هرگاه
 $\forall \epsilon > 0, \exists k_\epsilon \in \mathbb{N} \text{ s.t. } \forall m, n \geq k_\epsilon \implies d(x_n, x_m) < \epsilon$.

تعریف ۶.۲.۱. فرض کنید (X, d) یک فضای متریک باشد و $A \subset X$. خانواده $\{G_\alpha\}_{\alpha \in I} \subset X$ از زیر مجموعه‌های باز X ، یک پوشش باز^{۱۲} A نامیم هرگاه
 $A \subset \bigcup_{\alpha \in I} G_\alpha$.

تعریف ۷.۲.۱. فرض کنید (X, d) یک فضای متریک باشد. $A \subset X$ را مجموعه فشرده^{۱۳} نامیم، هرگاه هر پوشش باز A دارای یک زیرپوشش^{۱۴} متناهی باشد. یعنی برای هر خانواده $\{G_\alpha\}_{\alpha \in I}$ از زیر مجموعه‌های باز A که $A \subset \bigcup_{\alpha \in I} G_\alpha$ داشته باشیم:
 $\exists \alpha_1, \alpha_2, \dots, \alpha_n \in I, \text{ s.t. } A \subset \bigcup_{i=1}^n G_{\alpha_i}$.

تعریف ۸.۲.۱. مجموعه V را همراه با دو عمل جمع و ضرب اسکالر، یک فضای برداری^{۱۵} روی $\mathbb{R}$ می‌نامیم هرگاه به ازای هر a و b عضو V و هر λ_1 و λ_2 عضو $\mathbb{R}$ ، عمل جمع دارای شرایط زیر باشد:

$$۱. \quad a + b = b + a$$

$$۲. \quad (a + b) + c = a + (b + c)$$

$$۳. \quad a + 0 = a, \text{ که } 0 \in V \text{ است،}$$

$$۴. \quad a + (-a) = 0, \text{ که } -a \in V \text{ است.}$$

عمل ضرب اسکالر دارای شرایط زیر باشد:

^{۱۱}Cauchy sequence

^{۱۲}open cover

^{۱۳}compact set

^{۱۴}subcover

^{۱۵}vector space

$$۱. \quad 1a = a, \text{ که } 1 \in \mathbb{R} \text{ است,}$$

$$۲. \quad \lambda_1(\lambda_2 a) = (\lambda_1 \lambda_2) a$$

$$۳. \quad \lambda_1(a + b) = \lambda_1 a + \lambda_1 b$$

$$۴. \quad (\lambda_1 + \lambda_2)a = \lambda_1 a + \lambda_2 b$$

تعریف ۹.۲.۱. یک فضای نرم‌دار^{۱۶}، یک فضای برداری V با تابع $\|\cdot\| : V \rightarrow R$ است به‌طوری‌که برای هر x و y عضو V روابط زیر برقرار باشد:

$$۱. \quad 0 \leq \|x\| < \infty$$

$$۲. \quad \|x\| = 0 \iff x = 0$$

$$۳. \quad \|\alpha x\| = |\alpha| \|x\|, \quad \forall \alpha \in \mathbb{R}$$

$$۴. \quad \|x + y\| \leq \|x\| + \|y\|$$

هر تابع $\|\cdot\| : V \rightarrow R$ با این چهار خاصیت نرم نامیده می‌شود. به زوج $(V, \|\cdot\|)$ فضای خطی نرم‌دار گویند.

تعریف ۱۰.۲.۱. فرض کنید X یک مجموعه ناتهی باشد. به ازای هر $x, y \in \mathbb{R}^n$ نامساوی کوشی شوارتز^{۱۷} به‌صورت زیر است:

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \|x\| \|y\|.$$

تعریف ۱۱.۲.۱. تابع $f : \mathbb{R}^n \rightarrow \mathbb{R}$ را پیوسته یکنواخت^{۱۸} گویند هرگاه

$$\forall \epsilon > 0, \exists \delta > 0 \text{ s.t. } \forall x, y, \quad \|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

بدیهی است اگر f پیوسته یکنواخت باشد آن‌گاه پیوسته است اما اگر پیوسته باشد لزوماً پیوسته یکنواخت نیست.

مثال ۱۲.۲.۱. توابع $f(x) = x$ و $g(x) = x^2$ را در نظر بگیرید.

با توجه به تعریف پیوسته یکنواخت واضح است که با در نظر گرفتن $\delta = \epsilon$ تابع f پیوسته یکنواخت است. همچنین با برهان خلف ثابت می‌شود تابع g پیوسته یکنواخت نیست.

تعریف ۱۳.۲.۱. تابع $f : \mathbb{R}^n \rightarrow \mathbb{R}$ را پیوسته لیپ‌شیتز^{۱۹} گویند هرگاه

$$\exists M > 0 \text{ s.t. } |f(x) - f(y)| \leq M \|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$

بدیهی است اگر f پیوسته لیپ‌شیتز باشد آن‌گاه پیوسته یکنواخت است.

^{۱۶}normed space

^{۱۷}Cauchy-Schwartz inequality

^{۱۸}uniformly continuous

^{۱۹}lipschitz continuous

تعریف ۱۴.۲.۱. یک دنباله از توابع $f_n : X \rightarrow \mathbb{R}$ را پیوسته یکسان^{۲۰} گویند هرگاه

$$\forall \epsilon > 0, \exists \delta > 0 \text{ s.t. } \forall x, y \in X, \|x - y\| < \delta \Rightarrow |f_n(x) - f_n(y)| < \epsilon, \quad \forall n \in \mathbb{N}.$$

که (X, d) یک فضای متریک است.

تعریف ۱۵.۲.۱. یک دنباله از توابع $f_n : X \rightarrow \mathbb{R}$ را کراندار یکنواخت^{۲۱} گویند اگر عدد حقیقی M وجود داشته باشد به طوری که

$$|f_n(x)| \leq M, \quad \forall n \in \mathbb{N}, \quad \forall x \in X.$$

تعریف ۱۶.۲.۱. [۴] دنباله $\{x_k\} \subset \mathbb{R}^n$ را در نظر بگیرید و فرض کنید $\bar{x} \in \mathbb{R}^n$ به گونه‌ای باشد که $x_k \rightarrow \bar{x}$ و برای هر $k, x_k \neq \bar{x}$. عدد نامنفی p را مرتبه همگرایی^{۲۲} دنباله گویند هرگاه

$$\limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|^p} = \beta < \infty. \quad (۱.۱)$$

اگر $p = ۱$ و $\beta \in (0, ۱)$ باشد، سرعت همگرایی دنباله خطی^{۲۳} است.

اگر $p = ۱$ و $\beta = 0$ باشد، سرعت همگرایی دنباله فوق خطی^{۲۴} است.

اگر $p = ۲$ و $\beta < \infty$ باشد، سرعت همگرایی دنباله درجه دوم^{۲۵} است.

اگر حد رابطه (۱.۱) موجود باشد، بزرگتر شدن مقدار p باعث می‌شود همگرایی الگوریتم سریعتر شود. همچنین برای p یکسان، کوچکتر شدن مقدار β باعث می‌شود همگرایی الگوریتم سریعتر شود. لذا سرعت همگرایی فوق خطی سریعتر از همگرایی خطی است اما همگرایی درجه دوم بسیار سریعتر از همگرایی فوق خطی است.

مثال ۱۷.۲.۱. فرض کنید $\gamma \in (0, ۱)$ باشد آن‌گاه برای توابع γ^n ، γ^{n^2} و γ^{2^n} از تعریف بالا استفاده می‌کنیم.

سرعت همگرایی γ^n خطی به 0 است اما فوق خطی نیست زیرا بنا به تعریف داریم:

$$\limsup_{n \rightarrow \infty} \frac{\|\gamma^{n+1}\|}{\|\gamma^n\|} = \gamma < ۱.$$

سرعت همگرایی γ^{n^2} فوق خطی به 0 است اما درجه دوم نیست زیرا بنا به تعریف داریم:

$$\limsup_{n \rightarrow \infty} \frac{\|\gamma^{(n+1)^2}\|}{\|\gamma^{n^2}\|} = 0.$$

و

$$\limsup_{n \rightarrow \infty} \frac{\|\gamma^{(n+1)^2}\|}{\|\gamma^{2n^2}\|} = \infty.$$

^{۲۰}equicontinuous

^{۲۱}uniformly bounded

^{۲۲}order of convergence

^{۲۳}linear

^{۲۴}superlinear

^{۲۵}quadratic

سرعت همگرایی γ^{2^n} درجه دوم به $\circ$ است زیرا بنا به تعریف داریم:

$$\limsup_{n \rightarrow \infty} \frac{\|\gamma^{2^{n+1}}\|}{\|\gamma^{2^n}\|} = 1 < \infty.$$

تعریف ۱۸.۲.۱. $C \subset \mathbb{R}^n$ را مجموعه محدب^{۲۶} گوئیم هرگاه

$$\forall x, y \in C, \forall \lambda \in [0, 1], \quad \lambda x + (1 - \lambda)y \in C.$$

تعریف ۱۹.۲.۱. تابع $f : C \rightarrow \mathbb{R}$ روی مجموعه محدب $C \subset \mathbb{R}^n$ ، یک تابع محدب^{۲۷} گوئیم هرگاه

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall \lambda \in [0, 1], \forall x, y \in C.$$

همچنین تابع f محدب اکید^{۲۸} است، اگر رابطه فوق به طور اکید برقرار باشد.

تعریف ۲۰.۲.۱. تابع $f \in C^2(U, \mathbb{R})$ را محدب قوی^{۲۹} گوئیم هرگاه

$$\exists a > 0, \quad (\nabla f(x) - \nabla f(y))^T(x - y) \geq a\|x - y\|^2, \quad \forall x, y \in U,$$

که هم‌ارز با رابطه زیر است:

$$\exists a > 0, \quad \nabla^2 f(x) \geq aI, \quad \forall x \in U.$$

باید توجه داشت اگر تابعی محدب قوی باشد، محدب اکید نیز است اما عکس آن برقرار نیست.

مثال ۲۱.۲.۱. دو تابع $f(x) = x^2$ و $g(x) = e^x$ را در نظر بگیرید. چون $f''(x) = 2$ بنا به تعریف محدب قوی است. تابع $g(x) = e^x$ محدب اکید است. اما محدب قوی نیست چون مشتق دوم می‌تواند، به اندازه دلخواه نزدیک صفر شود.

تعریف ۲۲.۲.۱. تابع $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ، $\mathbb{R}_+^m$ -محدب نامیده می‌شود هرگاه F ، مؤلفه به مؤلفه محدب باشد.

قضیه ۲۳.۲.۱. [۲۳] فرض کنید که $f : \mathbb{R}^n \rightarrow \mathbb{R}$ تابعی مشتق‌پذیر پیوسته و محدب باشد آنگاه نقطه $\bar{x} \in \mathbb{R}^n$ یک مینیمم تابع f است اگر و فقط اگر $\nabla f(\bar{x}) = 0$.

تعریف ۲۴.۲.۱. [۶] یک مساله ناهموار^{۳۰} در حالت کلی به صورت زیر تعریف می‌شود:

$$\min f(x) := \sum_{i=1}^m q_i (\max\{g_{i1}(x), \dots, g_{ik_i}(x)\}), \quad (2.1)$$

که m و k_i اعداد طبیعی ثابت هستند و توابع $q_i : \mathbb{R}^{k_i} \rightarrow \mathbb{R}$ و $g_{ij} : X \rightarrow \mathbb{R}$ به ازای $i = 1, 2, \dots, m$ و $j = 1, 2, \dots, k_i$ پیوسته و مشتق‌پذیر هستند و X فضای برداری نرم‌دار و حقیقی است. چند حالت

^{۲۶}convex set

^{۲۷}convex function

^{۲۸}strictly convexity

^{۲۹}strong convexity

^{۳۰}nonsmooth problem

خاص رابطه (۲.۱) را می‌توانیم به صورت زیر بیان کنیم:
حالت اول:

$$\min f(x) := \max\{g_1(x), g_2(x), \dots, g_s(x)\},$$

که با قراردادن $q_i(t) = \frac{2t}{s(s+1)}$ و $k_i = s$ و $g_{i1} = g_1, \dots, g_{is} = g_s$ به ازای $i = 1, 2, \dots, s$ در رابطه (۲.۱) نتیجه می‌شود.
حالت دوم:

$$\min f(x) := \sum_{i=1}^s |g_i(x)|,$$

که با قراردادن $q_i(t) = t$ و $k_i = 2$ و $g_{i1} = g_i$ و $g_{i2} = -g_i$ به ازای $i = 1, 2, \dots, s$ به دست می‌آید.
تعریف ۲.۵.۲.۱ [۴] تابع مشتق‌پذیر $f: \mathbb{R}^n \rightarrow \mathbb{R}$ و توابع مشتق‌پذیر $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ ، $i = 1, 2, \dots, s$ را در نظر بگیرید و فرض کنید $X \subset \mathbb{R}^n$ ناتهی و باز باشد. مساله برنامه‌ریزی غیرخطی زیر را در نظر بگیرید:

$$\begin{aligned} \min f(x) \\ \text{s.t. } g_j(x) \leq 0 \quad j = 1, 2, \dots, m \\ x \in X. \end{aligned} \quad (3.1)$$

شرایط KKT^{۳۱} برای مساله (۳.۱) به صورت زیر هستند:

$$1. \lambda_j \geq 0, \quad j = 1, 2, \dots, m$$

$$2. \nabla f(x) + \sum_{j=1}^m \lambda_j \nabla g_j(x) = 0, \quad j = 1, 2, \dots, m$$

$$3. \lambda_j g_j(x) = 0, \quad j = 1, 2, \dots, m$$

$$4. g_j(x) \leq 0, \quad x \in X, \quad j = 1, 2, \dots, m$$

که $j = 1, 2, \dots, m$ ضریب لاگرانژین^{۳۲} نامیده می‌شود. شرایط ۱ و ۲، شرایط شدنی دوگان^{۳۳}، شرط ۳، شرط مکمل زائد^{۳۴} و شرط ۴، شرط اولیه شدنی^{۳۵} هستند.

مساله بهینه‌سازی محدب^{۳۶} را در حالت کلی به صورت زیر بیان می‌کنیم:

$$\begin{aligned} CO \quad \min f(x) \\ \text{s.t. } g_j(x) \leq 0 \quad j = 1, 2, \dots, m \\ x \in X. \end{aligned} \quad (4.1)$$

^{۳۱}Karush-Kuhn-Tucker

^{۳۲}Lagrangian multiplier

^{۳۳}dual feasibility

^{۳۴}complementary slackness

^{۳۵}primal feasibility

^{۳۶}Convex optimization problem

که در آن X یک مجموعه محدب است و توابع $g_j(x)$ ، $j = 1, 2, \dots, m$ و $f(x)$ نیز محدب‌اند. تعریف ۲۶.۲.۱. [۲۳] تابع لاگرانژین^{۳۷} (لاگرانژ^{۳۸}) مساله CO به صورت زیر تعریف می‌شود:

$$L(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x),$$

که $x \in X$ و $\lambda \geq 0$. توجه کنید که برای λ ثابت، تابع لاگرانژین بر حسب x محدب است.

تعریف ۲۷.۲.۱. [۲۳] نقطه $\bar{x} \in \text{int}(X)$ را یک نقطه اسلاتر^{۳۹} برای مساله CO گویند هرگاه در شرایط اسلاتر زیر صدق کند:

۱. برای هر i که در آن $g_i(x)$ غیرخطی است، داشته باشیم $g_i(\bar{x}) < 0$ ،

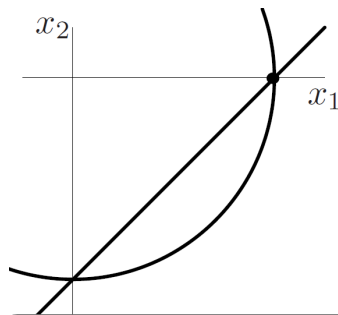
۲. برای هر i که در آن $g_i(x)$ خطی است، داشته باشیم $g_i(\bar{x}) \leq 0$.

همچنین اگر یک نقطه اسلاتر برای مساله CO وجود داشته باشد گوییم مساله منظم اسلاتر^{۴۰} است یا در شرایط اسلاتر منظم صدق می‌کند.

مثال ۲۸.۲.۱. مساله بهینه‌سازی زیر را در نظر بگیرید.

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x_1^2 + x_2^2 \leq 4 \\ & x_1 - x_2 \geq 2 \\ & x_2 \geq 0. \end{aligned}$$

ناحیه شدنی مساله تنها نقطه $(2, 0)$ است که محدودیت غیرخطی $x_1^2 + x_2^2 \leq 4$ به ازای این نقطه برابر ۴ است. بنابراین مساله منظم اسلاتر نیست.



شکل ۱۰.۱: مجموعه شدنی مثال ۲۸.۲.۱

تعریف ۲۹.۲.۱. [۲۳] فرض کنید $\psi(y) = \inf_{x \in X} \{f(x) + \sum_{j=1}^m y_j g_j(x)\}$ به مساله زیر دوگان لاگرانژ^{۴۱} مساله CO گویند:

^{۳۷}Lagrangian function

^{۳۸}Lagrange

^{۳۹}Slater point

^{۴۰}Slater regular

^{۴۱}Lagrange dual

$$LD \sup \psi(y) \\ s.t. \ y \geq 0.$$

تعریف ۳۰.۲۰.۱ [۲۳] اگر $\bar{x}$ یک جواب شدنی مساله CO و $\bar{y} \geq 0$ باشد $f(\bar{x}) - \psi(\bar{y})$ شکاف دوگان^{۴۲} $\bar{x}$ و $\bar{y}$ نامیده می‌شود.

قضیه ۳۱.۲۰.۱ [۲۳] اگر $\bar{x}$ یک جواب شدنی مساله CO و $\bar{y} \geq 0$ باشد، آنگاه

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

و تساوی برقرار است اگر و فقط اگر $\inf_{x \in X} \{f(x) + \sum_{j=1}^m \bar{y}_j g_j(x)\} = f(\bar{x})$

نتیجه ۳۲.۲۰.۱ [۲۳] اگر $\bar{x}$ جواب شدنی مساله CO و $\bar{y} \geq 0$ باشد. همچنین $\psi(\bar{y}) = f(\bar{x})$ باشد آنگاه $\bar{x}$ جواب بهینه مساله CO و $\bar{y}$ جواب بهینه مساله LD است.

قضیه ۳۳.۲۰.۱ [۴] اگر در یک مساله برنامه‌ریزی محدب (CO) شرایط KKT در نقطه $\bar{x}$ برقرار باشند آنگاه نقطه $\bar{x}$ بهینه است و برعکس.

۳.۱ مبانی بهینه‌سازی چندهدفه

در این قسمت مبانی مربوط به بهینه‌سازی چندهدفه را به‌طور مختصر بیان می‌کنیم. در یک مساله چندهدفه، تعدادی تابع هدف که ممکن است در تضاد با هم دیگر باشند، به‌طور هم‌زمان بهینه می‌شوند. مسائل بهینه‌سازی که امروزه در عمل ظاهر می‌شوند، بیشتر چندهدفه هستند. در جهان امروز، به ندرت می‌توان مساله‌ای را پیدا کرد که فقط به یک هدف یا ارزش توجه داشته باشد.

به‌عنوان مثال در یک کارخانه یک هدف، حداکثر کردن سود است که بر حسب پول سنجش می‌شود و هدف دیگر حداقل استفاده از ساعات نیروی کار است که بر حسب ساعت سنجیده می‌شود. گاهی این اهداف در یک جهت نیستند و به‌صورت متضاد عمل می‌کنند. مثلاً تصمیم گیرنده از یک طرف تمایل دارد رضایت کارکنان را افزایش دهد که در بازدهی و سود کارخانه اثر مستقیم دارد و از طرف دیگر می‌خواهد هزینه‌های حقوق و دستمزد را حداقل کند.

بهینه‌سازی چندهدفه، یکی از زمینه‌های بسیار پرکاربرد تحقیقاتی در میان مباحث بهینه‌سازی است و اهمیت فوق‌العاده‌ای، در مسائل بهینه‌سازی دنیای واقعی دارد. برای مطالعه بیشتر کاربرد بهینه‌سازی چندهدفه می‌توانید، به مراجع [۱۵] و [۳۶] مراجعه کنید.

مشکل اصلی که در برخورد با مسائل چندهدفه مطرح است، وجود مجموعه‌ای از جواب‌ها است که هر کدام از دیدگاهی می‌توانند بهینه باشند. یک مساله چندهدفه برخلاف مسائل تک‌هدفه که دارای یک جواب بهینه هستند، معمولاً چندین جواب بهینه دارند. تصمیم گیرنده برای انتخاب این جواب‌ها می‌تواند به‌صورت دلخواه یا نظام‌مند عمل کند. این جواب‌ها در حوزه بهینه‌سازی چندهدفه، به مفهوم جواب‌های کارا هستند، که در ادامه تعریف می‌کنیم.

^{۴۲}duality gap

یک مساله بهینه‌سازی چندهدفه^{۴۳} در حالت کلی به صورت زیر تعریف می‌شود:

$$\begin{aligned} \min F(x) &= (F_1(x), \dots, F_m(x)) \\ \text{s.t. } x &\in U \subset \mathbb{R}^n. \end{aligned} \quad (۵.۱)$$

که در آن $x = (x_1, x_2, \dots, x_n)^T$ ، توابع هدف هستند، $F_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ، $i = 1, 2, \dots, m$ متغیرهای تصمیم است و U مجموعه شدنی مساله است. تصویر مجموعه شدنی، $F(U)$ را فضای هدف گویند. در ادامه U را یک مجموعه باز در نظر می‌گیریم.

تعریف ۱.۳.۱. [۱۶] نقطه $x^* \in U$ را کارای^{۴۴} سراسری (بهینه پارتو^{۴۵}) برای مساله بهینه‌سازی چندهدفه (۵.۱) گویند هرگاه

$$\nexists y \in U, \quad F(y) \leq F(x^*).$$

در واقع $x^* \in U$ کارا است اگر

$$\nexists y \in U, \quad F_i(y) \leq F_i(x^*), \forall i \text{ و } F_j(y) < F_j(x^*), \exists j \in \{1, 2, \dots, m\}.$$

به عبارت دیگر داریم:

$$\nexists y \in U, \quad F(x^*) - F(y) \in \mathbb{R}_+^m.$$

تعریف ۲.۳.۱. [۱۶] نقطه $x^* \in U$ را کارای ضعیف^{۴۶} سراسری برای مساله بهینه‌سازی چندهدفه (۵.۱) گویند هرگاه

$$\nexists y \in U, \quad F(y) < F(x^*).$$

به عبارت دیگر داریم:

$$\nexists y \in U, \quad F(x^*) - F(y) \in \mathbb{R}_{++}^m.$$

ملاحظه ۳.۳.۱. در ادامه به جای نقطه کارای سراسری از نقطه کارا استفاده می‌کنیم.

مثال ۴.۳.۱. مساله بهینه‌سازی دوهدفه زیر را در نظر بگیرید.

$$\min F(x_1, x_2) = (x_1, x_2),$$

فرض کنید مجموعه شدنی $\mathcal{X}$ به صورت زیر باشد:

$$\mathcal{X} = \left\{ (x_1, x_2) \in \mathbb{R}^2 \left| \begin{array}{ll} x_2 = 0; & x_1 = -1, \\ -\sqrt{-x_1^2 + 1} < x_2 \leq 0; & -1 < x_1 < 0, \\ -\sqrt{-x_1^2 + 1} \leq x_2 \leq 0; & 0 \leq x_1 \leq 1. \end{array} \right. \right\}$$

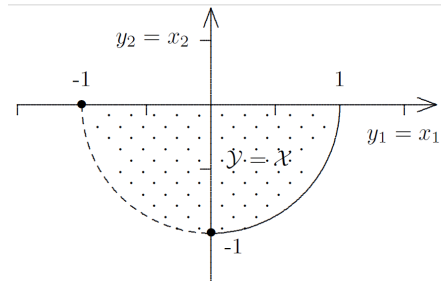
^{۴۳}multiobjective optimization problem

^{۴۴}efficient

^{۴۵}Pareto optimal

^{۴۶}weakly efficient

با توجه به مساله $\mathcal{Y} = F(\mathcal{X})$ است که $\mathcal{X} = \mathcal{Y}$ با توجه به مساله $\mathcal{Y} = F(\mathcal{X})$ است که $\mathcal{X} = \mathcal{Y}$ می‌توانید در شکل (۲.۱) ببینید، بنا به تعریف کارای ضعیف و نقاط کارا، مجموعه نقاط کارای ضعیف تهی و مجموعه نقاط $\{(-1, 0), (0, -1)\}$ کارا است.



شکل ۲.۱: مجموعه شدنی مثال ۴.۳.۱

تعریف ۵.۳.۱. نقطه $x^* \in U$ کارای موضعی (کارای موضعی ضعیف) است اگر همسایگی $V \in U$ از x^* موجود باشد به طوری که x^* ، نقطه کارا (کارای ضعیف) برای مساله بهینه‌سازی چندهدفه (۵.۱)، محدود شده توسط V باشد.

نقاط کارای موضعی، بهینه پارتو موضعی و نقاط کارای موضعی ضعیف، بهینه پارتو موضعی ضعیف نیز نامیده می‌شوند. واضح است که هر نقطه کارای موضعی، یک نقطه کارای موضعی ضعیف است.

قضیه ۶.۳.۱. [۱۸] اگر U محدب و F ، $\mathbb{R}_+^m$ -محدب باشد آن‌گاه هر نقطه کارای موضعی یک نقطه کارا است.

تعریف ۷.۳.۱. نقطه $\bar{x} \in U$ را یک نقطه بحرانی^۴ برای F گویند اگر

$$R(DF(\bar{x})) \cap (-\mathbb{R}_{++}^m) = \emptyset. \quad (6.1)$$

در واقع رابطه (۶.۱) معادل با رابطه زیر است:

$$\left(s_1 \frac{\partial F_1(\bar{x})}{\partial x_1} + \dots + s_n \frac{\partial F_1(\bar{x})}{\partial x_n}, \dots, s_1 \frac{\partial F_m(\bar{x})}{\partial x_1} + \dots + s_n \frac{\partial F_m(\bar{x})}{\partial x_n} \right) \cap (-\mathbb{R}_{++}^m) = \emptyset,$$

به ازای $m = 1$ رابطه بالا به شرط کلاسیک گرادیان F مساوی با صفر تبدیل می‌شود، زیرا بنا به این رابطه داریم:

$$\left[s_1 \frac{\partial F(\bar{x})}{\partial x_1} + s_2 \frac{\partial F(\bar{x})}{\partial x_2} + \dots + s_n \frac{\partial F(\bar{x})}{\partial x_n} \right] \cap (-\mathbb{R}_{++}) = \emptyset,$$

یعنی برای هر $s \in \mathbb{R}^n$ باید $\nabla F(\bar{x})^T s \geq 0$ ، بنابراین

اگر $\bar{x}$ نقطه بحرانی برای F باشد آن‌گاه برای هر $s \in \mathbb{R}^n$ ، اندیس $j_* = j_*(s) \in \{1, 2, \dots, m\}$ وجود دارد به طوری که

$$\nabla F_{j_*}(\bar{x})^T s \geq 0. \quad (7.1)$$

^۴critical point

توجه شود که اگر $x \in U$ غیربحرانی باشد آن‌گاه $s \in \mathbb{R}^n$ وجود دارد به‌طوری‌که $\nabla F_j(x)^T s < 0$ برای هر $j = 1, 2, \dots, m$. در این حالت چون F به‌طور پیوسته مشتق‌پذیر است، داریم:

$$\lim_{t \rightarrow 0} \frac{F_j(x + ts) - F_j(x)}{t} = \nabla F_j(x)^T s < 0, \quad j = 1, 2, \dots, m.$$

بنابراین s یک جهت نزولی برای F در نقطه x است، یعنی $t_0 > 0$ وجود دارد به‌طوری‌که

$$F(x + ts) < F(x), \quad \forall t \in (0, t_0]. \quad (۸.۱)$$

قضیه ۸.۳.۱. فرض کنید که $F \in C^1(U, \mathbb{R}^m)$ باشد.

الف. اگر $\bar{x}$ کارای موضعی ضعیف باشد آن‌گاه $\bar{x}$ یک نقطه بحرانی برای تابع F است.

ب. اگر U محدب و F یک تابع $\mathbb{R}^m$ -محدب باشد و $\bar{x} \in U$ یک نقطه بحرانی برای تابع F باشد آن‌گاه $\bar{x}$ یک نقطه کارای ضعیف است.

ج. فرض کنید U محدب باشد و تابع $F \in C^2(U, \mathbb{R}^m)$ مفروض باشد. همچنین فرض کنید برای هر $j = 1, 2, \dots, m$ و هر $x \in U$ ، $\nabla^2 F_j(x) > 0$ باشد. اگر $\bar{x} \in U$ نقطه بحرانی برای F باشد آن‌گاه $\bar{x}$ یک نقطه کارا است.

برهان. **الف.** به برهان خلف فرض کنید $\bar{x}$ بحرانی نباشد آن‌گاه رابطه (۶.۱) برقرار نیست. بنابراین $t_0 > 0$ و $s \in \mathbb{R}^n$ وجود دارند که در رابطه (۸.۱) صدق می‌کنند و این با کارای ضعیف بودن $\bar{x}$ تناقض دارد، بنابراین مورد الف برقرار است.

ب. نقطه $x \in U$ را دلخواه در نظر می‌گیریم. چون $\bar{x}$ بحرانی است لذا j_0 وجود دارد به‌طوری‌که رابطه (۷.۱) برای $s = x - \bar{x}$ برقرار است. بنا به محدب بودن F_{j_0} داریم:

$$F_{j_0}(x) \geq F_{j_0}(\bar{x}) + \nabla F_{j_0}(\bar{x})^T (x - \bar{x}) \geq F_{j_0}(\bar{x}), \quad (۹.۱)$$

در نتیجه $\bar{x}$ یک نقطه کارای ضعیف برای تابع F است.

ج. برای اثبات این مورد $x \in U$ را دلخواه در نظر می‌گیریم و با توجه به این که $\bar{x}$ نقطه بحرانی است، با استدلالی مشابه مورد ب، $j_0 \in \{1, \dots, m\}$ وجود دارد که رابطه (۹.۱) برقرار می‌شود. چون F_{j_0} محدب اکید است اگر $x \neq \bar{x}$ باشد نامساوی اول رابطه (۹.۱) اکیداً برقرار است در نتیجه $\bar{x}$ یک نقطه کارا است. $\square$

توجه شود که بنا به مورد الف، شرط لازم برای این که x کارای موضعی ضعیف باشد، این است که بحرانی باشد. همچنین وقتی $F, \mathbb{R}^m$ -محدب (در مورد ب) است بحرانی بودن، کارایی را نتیجه نمی‌دهد، تنها فرض محدب اکید بودن F ، بحرانی بودن را با کارای موضعی ضعیف معادل می‌کند. مثال زیر این مطلب را نشان می‌دهد.

مثال ۹.۳.۱. فرض کنید $n = 1, m = 2$ و $U = \mathbb{R}$. تابع $F(t) = (1, t)$ را در نظر بگیرید.

در این صورت هر $t \in \mathbb{R}$ بحرانی است، چون داریم:

$$F_1(t) = 1 \rightarrow \nabla F_1 = 0,$$

$$F_2(t) = t \rightarrow \nabla F_2 = 1,$$

$$DF(t) = \begin{bmatrix} \nabla F_1 \\ \nabla F_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$R(DF(t)) = s(0, 1),$$

$$R(DF(t)) \cap (-\mathbb{R}_{++}^m) = (0, s) \cap (-\mathbb{R}_{++}^2) = \emptyset, \quad \forall s \in \mathbb{R}.$$

اما جواب کارا برای F وجود ندارد. به برهان خلف فرض کنیم نقطه $x = (1, t_1)$ کارا برای F باشد آن‌گاه نقطه $y = (1, t_2)$ در نظر می‌گیریم به‌طوری‌که $t_2 < t_1$ ، که با کارا بودن x در تناقض است.

لذا تنها فرض محدب اکید بودن است که مجموعه نقاط کارا را بر مجموعه نقاط بحرانی منطبق می‌کند. چون بنا به مورد ج اگر $\bar{x} \in U$ یک نقطه بحرانی برای تابع محدب اکید F باشد آن‌گاه $\bar{x}$ یک نقطه کارا است و بنا به مورد الف اگر $\bar{x}$ کارا باشد آن‌گاه $\bar{x}$ یک نقطه بحرانی است.

در ادامه پایان‌نامه فرض می‌کنیم که به ازای هر $x \in U$ و هر $j = 1, 2, \dots, m$ داریم:

$$\nabla^2 F_j(x) > 0.$$

به این مفهوم که $\nabla^2 F_j(x)$ برای هر $x \in U$ و هر $j = 1, 2, \dots, m$ مثبت معین است. تحت این فرض، F روی هر زیر مجموعه محدب U ، $\mathbb{R}^m$ -محدب است. همچنین فرض می‌کنیم که $F \in C^2(U, \mathbb{R}^m)$ ، بدین معنی است که F ، دوبار به‌طور پیوسته مشتق‌پذیر روی U است. در ادامه پایان‌نامه به‌جز مواردی که صریحاً ذکر شود، این فرض‌ها را اعمال می‌کنیم.

فصل ۲

روش نیوتن برای حل مسائل چندهدفه

۱.۲ مقدمه

در این فصل به بیان روش نیوتن برای حل مسائل بهینه‌سازی می‌پردازیم. با این فرض که محاسبه و به‌کارگیری ماتریس هسین غیرعملی یا پرهزینه نیست، روش نیوتن بر استفاده از ماتریس هسین عمل می‌کند.

۲.۲ روش نیوتن برای حل مسائل تک‌هدفه

در این قسمت ما چگونگی روش نیوتن برای حل مسائل تک‌هدفه را به‌طور خلاصه بیان می‌کنیم. روش نیوتن مبتنی بر این ایده است که تابع هدف را با کمک تابع درجه دوم، بهینه می‌کند. از این رو می‌توانیم تابع $f \in C^2(\mathbb{R}^n, \mathbb{R})$ را در نزدیکی نقطه x_k با سری تیلور زیر تقریب بزنیم:

$$f(x) \simeq f(x_k) + \nabla f(x_k)(x - x_k) + \frac{1}{2}(x - x_k) \nabla^2 f(x_k)(x - x_k).$$

شرط لازم برای حداقل کردن تقریب تابع درجه دوم $f(x)$ عبارت است از:

$$\nabla f(x) = 0,$$

به عبارت دیگر $\nabla f(x_k) + \nabla^2 f(x_k)(x - x_k) = 0$. لذا در صورت وجود معکوس $\nabla^2 f$ ، نقطه جدید x_{k+1} با کمک رابطه زیر به دست می‌آید:

$$x_{k+1} = x_k - \nabla f(x_k) [\nabla^2 f(x_k)]^{-1}.$$

این روش زمانی خاتمه یافته تلقی می‌شود که روابط $\|\nabla f(x_k)\| < \varepsilon$ یا $\|x_{k+1} - x_k\| < \varepsilon$ برقرار شوند. در این روابط ε عدد از پیش تعیین شده، برای توقف حل مساله است که بسیار کوچک است.

قضیه ۱.۲.۲ [۲۶] فرض می‌کنیم $f \in C^2(\mathbb{R}^n, \mathbb{R})$ باشد و در نقطه مینیمم نسبی x^* ، ماتریس هسین $\nabla^2 f(x^*)$ مثبت معین باشد. در این صورت اگر از نقطه‌ای به حد کافی نزدیک به x^* شروع کنیم، نقاط تولید شده با کمک روش نیوتن همگرا به x^* است و مرتبه همگرایی حداقل دو است.

۳.۲ روش نیوتن برای حل مسائل چندهدفه

اکنون به تعریفی از جهت نیوتن برای مسائل چندهدفه می‌پردازیم. همانند مسائل تک‌هدفه کلاسیک، جهت نیوتن یک جواب بهینه از یک مساله است که شامل تقریب‌های درجه دوم از توابع هدف F_j باشد. علاوه بر این، همانند مسائل تک‌هدفه در یک نقطه بحرانی، گام نیوتن $\circ \in \mathbb{R}^n$ خواهد بود. برای $s(x)$ ، $x \in U$ را جهت نیوتن در x می‌نامیم که جواب بهینه‌ای از مساله زیر است:

$$\begin{aligned} \min \quad & \max_{j=1,2,\dots,m} \nabla F_j(x)^T s + \frac{1}{2} s^T \nabla^2 F_j(x) s \\ \text{s.t.} \quad & s \in \mathbb{R}^n. \end{aligned} \quad (1.2)$$

تابع $g(s) = \nabla F_j(x)^T s + \frac{1}{2} s^T \nabla^2 F_j(x) s$ برای هر $j = 1, 2, \dots, m$ محدب اکید است، چون داریم:

$$\nabla g(s) = \nabla F_j(x) + s^T \nabla^2 F_j(x), \quad \nabla^2 g(s) = \nabla^2 F_j(x).$$

که $\nabla^2 F_j(x)$ مثبت معین است، لذا مساله (۱.۲) بنا به قضیه ۲۳.۲.۱ همیشه یک مینیمم منحصر بفرد دارد.

همچنین توجه کنید که برای $m = 1$ جهت $s(x)$ ، جهت نیوتن کلاسیک برای بهینه‌سازی تک‌هدفه است. مقدار بهینه مساله (۱.۲) را با $\theta(x)$ نشان می‌دهیم و روابط زیر را داریم:

$$\theta(x) = \inf_{s \in \mathbb{R}^n} \max_{j=1,2,\dots,m} \nabla F_j(x)^T s + \frac{1}{2} s^T \nabla^2 F_j(x) s \quad (2.2)$$

و

$$s(x) = \arg \min_{s \in \mathbb{R}^n} \max_{j=1,2,\dots,m} \nabla F_j(x)^T s + \frac{1}{2} s^T \nabla^2 F_j(x) s. \quad (3.2)$$

در این جا $\max_{j=1,2,\dots,m} F_j(x+s) - F_j(x)$ را با ماکسیم مدل‌های درجه دوم موضعی در x برای هر F_j تقریب می‌زنیم. اگرچه مساله (۱.۲)، یک مساله ناهموار است اما می‌توان آن را به شکل یک مساله برنامه‌ریزی درجه دوم محدب تبدیل کرد که می‌تواند به راحتی حل شود. در واقع مساله (۱.۲)، معادل با مساله زیر است:

$$\begin{aligned} \min \quad & g(t, s) = t \\ \text{s.t.} \quad & \nabla F_j(x)^T s + \frac{1}{2} s^T \nabla^2 F_j(x) s - t \leq 0 \quad j = 1, 2, \dots, m \\ & (t, s) \in \mathbb{R} \times \mathbb{R}^n. \end{aligned} \quad (4.2)$$

لاگرانژین این مساله به صورت زیر است:

$$L((t, s), \lambda) = t + \sum_{j=1}^m \lambda_j (\nabla F_j(x)^T s + \frac{1}{2} s^T \nabla^2 F_j(x) s - t). \quad (5.2)$$

اکنون شرایط KKT را برای مساله (۴.۲) با توجه به تعریف ۲۵.۲.۱ به صورت زیر داریم:

$$\sum_{j=1}^m \lambda_j = 1, \quad (۶.۲)$$

$$\sum_{j=1}^m \lambda_j (\nabla F_j(x) + \nabla^2 F_j(x)s) = 0, \quad (۷.۲)$$

$$\nabla F_j(x)^T s + \frac{1}{p} s^T \nabla^2 F_j(x)s \leq t, \quad j = 1, 2, \dots, m, \quad (۸.۲)$$

$$\lambda_j \geq 0, \quad j = 1, 2, \dots, m, \quad (۹.۲)$$

$$\lambda_j (\nabla F_j(x)^T s + \frac{1}{p} s^T \nabla^2 F_j(x)s - t) = 0, \quad j = 1, 2, \dots, m. \quad (۱۰.۲)$$

که شرایط (۶.۲)، (۷.۲) و (۹.۲) شرایط شدنی دوگان هستند، شرط (۸.۲) شرط شدنی اولیه است و شرط (۱۰.۲)، شرط مکمل زائد است.

مساله (۴.۲) یک جواب منحصر بفرد $(\theta(x), s(x))$ دارد. قابل ذکر است که این مساله محدب است و در شرایط منظم اسلاتر صدق می کند زیرا یک نقطه مانند نقطه $(1, 0)$ وجود دارد به طوری که

$$\nabla F_j(x)^T 0 + \frac{1}{p} 0^T \nabla^2 F_j(x) 0 - 1 = -1 < 0, \quad j = 1, 2, \dots, m.$$

بنا به قضیه ۳۳.۲.۱ یک ضریب لاگرانژین، $\lambda = \lambda(x)$ وجود دارد که به همراه $s = s(x)$ و $t = \theta(x)$ در شرایط (۶.۲)-(۱۰.۲) صدق می کند. به خصوص از رابطه (۷.۲) داریم:

$$\begin{aligned} \sum_{j=1}^m \lambda_j(x) \nabla F_j(x) &= - \sum_{j=1}^m \lambda_j(x) \nabla^2 F_j(x)s(x), \\ s(x) &= - \left[\sum_{j=1}^m \lambda_j(x) \nabla^2 F_j(x) \right]^{-1} \sum_{j=1}^m \lambda_j(x) \nabla F_j(x). \end{aligned} \quad (۱۱.۲)$$

توجه شود چون $\nabla^2 F_j$ مثبت معین و حداقل یکی از $\lambda_j \neq 0$ عبارت $\sum_{j=1}^m \lambda_j(x) \nabla^2 F_j(x)$ معکوس پذیر است. وجود ضرایب لاگرانژین برای مساله محدب (۴.۲) نتیجه می دهد که شکاف دوگان وجود ندارد چون بنا به رابطه (۱۰.۲)، $L((t, s), \lambda) = t$ است. لذا $t - L((t, s), \lambda) = 0$. بنابراین جواب بهینه مساله (۴.۲) با جواب بهینه دوگان لاگرانژین آن برابر است، در واقع داریم:

$$\begin{aligned} \theta(x) &= \sup_{\lambda \geq 0} \inf_{s \in \mathbb{R}^n} \left\{ t + \sum_{j=1}^m \lambda_j (\nabla F_j^T(x)s + \frac{1}{p} s^T \nabla^2 F_j(x)s - t) \right\} \\ &= \sup_{\lambda \geq 0} \inf_{s \in \mathbb{R}^n} L((t, s), \lambda). \end{aligned}$$

لذا از رابطه (۶.۲) داریم:

$$\theta(x) = \sup_{\{\lambda \geq 0, \sum_{j=1}^m \lambda_j = 1\}} \inf_{s \in \mathbb{R}^n} \sum_{j=1}^m \lambda_j (\nabla F_j(x)^T s + \frac{1}{p} s^T \nabla^2 F_j(x)s). \quad (۱۲.۲)$$

بنابراین از رابطه (۱۱.۲) می بینیم که جهت نیوتن تعریف شده یک جهت نیوتن برای مسائل بهینه سازی تک هدفه استاندارد است، که با قرار دادن ضرایب لاگرانژین به عنوان وزن های توابع هدف به دست می آید.

لازم نیست که وزن‌ها را در ابتدا ثابت کنیم، هر نقطه $x \in U$ وزن‌ها را با کمک ضرایب لاگرانژین در جهت متناظر تعریف می‌کند.

اکنون برخی از ویژگی‌های تابع θ را بررسی می‌کنیم و رابطه آن را با $s(x)$ و نقطه بحرانی x تحلیل می‌کنیم.

لم ۱.۳.۲. تحت فرض‌های کلی بیان شده، داریم:

۱. برای هر $x \in U$ ، $\theta(x) \leq 0$ است.

۲. شرایط زیر معادل‌اند:

الف. نقطه x ، بحرانی نیست،

ب. $\theta(x) < 0$ است،

ج. $s(x) \neq 0$ است.

۳. تابع $\theta : U \rightarrow \mathbb{R}$ که با رابطه (۲.۲) نشان داده شده است، پیوسته است.

برهان. ۱. با به‌کارگیری رابطه (۱.۲) داریم:

$$\theta(x) \leq \max_{j=1,2,\dots,m} \nabla F_j^T(x) \circ + \frac{1}{p} \circ^T \nabla^2 F_j(x) \circ = 0,$$

بنابراین مورد ۱ برقرار است.

۲. ابتدا فرض کنید الف برقرار باشد لذا داریم $R(DF(x)) \cap (-\mathbb{R}_{++}^m) \neq \emptyset$. بنابراین $\tilde{s} \in \mathbb{R}^n$ وجود دارد به‌طوری‌که $\nabla F_j^T(x) \tilde{s} < 0$ برای هر $j = 1, 2, \dots, m$. با استفاده از رابطه (۲.۲) برای هر $t > 0$ داریم:

$$\begin{aligned} \theta(x) &\leq \max_{j=1,2,\dots,m} \nabla F_j^T(x) t \tilde{s} + \frac{1}{p} t \tilde{s}^T \nabla^2 F_j(x) t \tilde{s} \\ &= t \max_{j=1,2,\dots,m} \nabla F_j^T(x) \tilde{s} + \frac{t}{p} \tilde{s}^T \nabla^2 F_j(x) \tilde{s}. \end{aligned}$$

چون $\nabla F_j^T(x) \tilde{s} < 0$ است، برای $t > 0$ به اندازه کافی کوچک، سمت راست نامساوی بالا منفی است و لذا مورد ب برقرار است.

برای این‌که ثابت کنیم ب رابطه ج را نتیجه می‌دهد، توجه شود که $\theta(x)$ مقدار بهینه مساله (۱.۲) است و بنا به رابطه ب منفی است. در نتیجه جواب مساله نمی‌تواند $s(x) = 0$ باشد.

سرانجام نشان می‌دهیم ج رابطه الف را نتیجه می‌دهد، چون $\nabla^2 F_j(x)$ ، به ازای $j = 1, 2, \dots, m$ ماتریس‌های مثبت معین هستند، $s(x)^T \nabla^2 F_j(x) s(x) > 0$ است. بنا به روابط (۳.۲)–(۲.۲) برای هر $j = 1, 2, \dots, m$ داریم:

$$\nabla F_j^T(x) s(x) < \nabla F_j^T(x) s(x) + \frac{1}{p} s(x)^T \nabla^2 F_j(x) s(x) \leq \theta(x) \leq 0,$$

که نامساوی آخر یک نتیجه از مورد ۱ است. بنابراین $DF(x)s(x) \in (-\mathbb{R}_{++}^m)$ است. بدین ترتیب x بحرانی نیست، در نتیجه الف برقرار است.

۳. برای اثبات این مورد کافی است که پیوستگی θ را در مجموعه معین $W \subset U$ که W یک مجموعه فشرد و دلخواه است، نشان دهیم. بنا به مورد ۱ برای هر $y \in U$ داریم:

$$\frac{1}{\gamma} s(y)^T \nabla^2 F_j(y) s(y) \leq -\nabla F_j(y)^T s(y), \quad j = 1, 2, \dots, m. \quad (13.2)$$

چون F دوبار به طور پیوسته مشتق پذیر است و هسین ها در همه جا مثبت معین هستند، مقدار ویژه های آنها کراندار یکنواخت دور از صفر روی W هستند. بنابراین K و $\hat{\lambda} > 0$ وجود دارند به طوری که

$$K = \max_{y \in W, j=1, \dots, m} \|\nabla F_j(y)\|$$

و

$$\hat{\lambda} = \min_{x \in W, \|u\|=1, j=1, 2, \dots, m} u^T \nabla^2 F_j(x) u.$$

با ترکیب رابطه (۱۳.۲) با دو معادله بالا و استفاده از نامساوی کوشی شوارتز، به ازای هر اندیس $j = 1, 2, \dots, m$ داریم:

$$\frac{1}{\gamma} \hat{\lambda} \|s(y)\|^2 \leq \|\nabla F_j(y)\| \|s(y)\| \leq K \|s(y)\|, \quad \forall y \in W.$$

بنابراین:

$$\|s(y)\| \leq \frac{\gamma}{\hat{\lambda}} K, \quad \forall y \in W. \quad (14.2)$$

یعنی جهت های نیوتن $s(y)$ روی W کراندار یکنواخت اند.

برای $x \in W$ و $j \in \{1, \dots, m\}$ تعریف می کنیم:

$$\varphi_{x,j} : W \rightarrow \mathbb{R},$$

$$z \mapsto \nabla F_j(z)^T s(x) + \frac{1}{\gamma} s(x)^T \nabla^2 F_j(z) s(x).$$

خانواده $\{\varphi_{x,j}\}_{x \in W, j=1, 2, \dots, m}$ پیوسته یکسان است. بنابراین خانواده $\{\phi_x = \max_{j=1, 2, \dots, m} \varphi_{x,j}\}_{x \in W}$ نیز پیوسته یکسان است.

فرض کنید $\varepsilon > 0$. لذا $\delta > 0$ وجود دارد به طوری که

$$\forall y, z \in W, \quad \|y - z\| < \delta \implies |\phi_x(y) - \phi_x(z)| < \varepsilon, \quad \forall x \in W.$$

بنابراین برای $\|y - z\| < \delta$ داریم:

$$\begin{aligned} \theta(z) &\leq \max_{j=1, 2, \dots, m} \nabla F_j(z)^T s(y) + \frac{1}{\gamma} s(y)^T \nabla^2 F_j(z) s(y) \\ &= \phi_y(z) \\ &\leq \phi_y(y) + |\phi_y(z) - \phi_y(y)| \\ &< \theta(y) + \varepsilon, \end{aligned}$$

یعنی $\theta(z) - \theta(y) < \varepsilon$ است. با جابه جا کردن نقش y و z نتیجه می گیریم که $|\theta(z) - \theta(y)| < \varepsilon$. $\square$

به منظور تعیین طول گام t برای جهت $s(x)$ ، از قاعده آرمیثو استفاده می‌کنیم. تابع $f : U \rightarrow \mathbb{R}$ را در نظر بگیرید. قاعده آرمیثو کلاسیک برای جهت جستجو نیوتن $s(x)$ ، در نقطه غیربحرانی $x \in U$ به صورت زیر است:

$$f(x + ts(x)) \leq f(x) + \sigma t \nabla f(x)^T s(x),$$

که $\sigma \in (0, 1)$ است. برای پذیرش یک گام در روش نیوتن چندهدفه با توجه به این که $\nabla^2 F_j(x)$ مثبت معین است، داریم:

$$F_j(x + ts(x)) \leq F_j(x) + \sigma t (\nabla F_j(x)^T s(x) + \frac{1}{2} s(x)^T \nabla^2 F_j(x) s(x)) \leq F_j(x) + \sigma t \theta(x)$$

نامساوی بالا معیار ما برای پذیرش یک گام در جهت نیوتن چندهدفه خواهد بود.

نتیجه ۲.۳.۲. اگر $x \in U$ یک نقطه غیربحرانی برای تابع F باشد آنگاه برای هر $0 < \sigma < 1$ ، وجود دارد $\bar{t} \in (0, 1]$ به طوری که به ازای هر $j = 1, 2, \dots, m$ رابطه زیر برقرار است:

$$x + ts(x) \in U, \quad F_j(x + ts(x)) \leq F_j(x) + \sigma t \theta(x), \quad \forall t \in [0, \bar{t}].$$

برهان. چون $x \in U$ است و یک مجموعه باز است. وجود دارد $0 < t_1 \leq 1$ به طوری که به ازای $x + ts(x) \in U$ ، $t \in [0, t_1]$ است. بنابراین برای $t \in [0, t_1]$ و هر $j = 1, 2, \dots, m$ داریم:

$$F_j(x + ts(x)) = F_j(x) + t \nabla F_j(x)^T s(x) + o_j(t), \quad \lim_{t \rightarrow 0^+} o_j(t)/t = 0.$$

همچنین چون $0 < \frac{1}{2} s(x)^T \nabla^2 F_j(x) s(x) > 0$ است، بنا به تعریف $\theta(x)$ ، $\nabla F_j(x)^T s(x) \leq \theta(x)$ است. بنابراین به ازای $t \in [0, t_1]$ و $j = 1, 2, \dots, m$ داریم:

$$F_j(x + ts(x)) \leq F_j(x) + t \theta(x) + o_j(t) = F_j(x) + t \sigma \theta(x) + t[(1 - \sigma)\theta(x) + \frac{o_j(t)}{t}].$$

چون x غیربحرانی است $\theta(x) < 0$ (لم ۱.۳.۲) است. در نتیجه برای $t \in [0, t_1]$ به اندازه کافی کوچک آخرین جمله سمت راست در معادله بالا نامثبت است. لذا داریم:

$$F_j(x + ts(x)) \leq F_j(x) + \sigma t \theta(x).$$

□

اکنون الگوریتم نیوتن را برای مسائل چندهدفه به صورت خلاصه بیان می‌کنیم. در هر گام، در یک نقطه غیربحرانی ماکسیمم همه مدل‌های موضعی مساله (۱.۲) را مینیمم می‌کنیم تا جهت نیوتن را از رابطه (۳.۲) که یک جهت کاهنده است، به دست آوریم. آنگاه طول گام را به وسیله قاعده آرمیثو و با کمک نتیجه ۲.۳.۲ به همراه یک فرایند برگشت‌پذیر تعیین می‌کنیم. تحت فرض‌های موضعی، همه گام‌های نیوتن یک دنباله همگرای فوق خطی (یا درجه دوم) به یک جواب موضعی را تولید می‌کنند. برای یافتن یک نقطه کارا از الگوریتم زیر استفاده می‌کنیم.

الگوریتم نیوتن برای مسائل چندهدفه

الگوریتم ۳.۳.۲. مقداردهی اولیه $k = 0$.

۱. نقطه $x_0 \in U$ را انتخاب کنید.

۴.۲.۱. مجموعه $J = \{1/2^n \mid n = 0, 1, 2, \dots\}$ را تعریف کنید.

۴.۲.۳. $0 < \sigma < 1$ را در نظر بگیرید.

گام‌های تکراری:

k_1 . مساله (۱.۲) را حل کنید و $s(x_k)$ و $\theta(x_k)$ را به ترتیب از روابط (۳.۲) و (۲.۲) محاسبه کنید.

k_2 . اگر $\theta(x_k) = 0$ ، متوقف شوید. در غیر این صورت به گام k_3 بروید.

k_3 . نقطه t_k را بزرگترین $t \in J$ انتخاب کنید به طوری که

$$x_k + ts(x_k) \in U$$

و

$$F_j(x_k + ts(x_k)) \leq F_j(x_k) + t\sigma\theta(x_k), \quad j = 1, 2, \dots, m.$$

k_4 . تعریف کنید $x_{k+1} = x_k + t_k s(x_k)$ و $k = k + 1$ قرار دهید و به گام k_1 بروید.

خوش تعریفی الگوریتم مستقیماً از نتیجه ۲.۳.۲ به دست می‌آید. این یک روش $\mathbb{R}^m$ -کاهنده است یعنی مقادیر تابع هدف تحت یک ترتیب جزئی مؤلفه‌ای در حال کاهش است. در واقع برای x_k غیربحرانی بنا به لم ۱.۳.۲، $\theta(x_k) < 0$ است. بنابراین برای روابط k_1 - k_4 از گام‌های تکراری، $F(x_{k+1}) < F(x_k)$ است.

۴.۲ نتایج تکنیکی

به منظور تحلیل همگرایی روش نیوتن برای بهینه‌سازی چندهدفه به چند لم نیاز داریم، که در ادامه بیان می‌کنیم. ویژگی اساسی روش نیوتن، برای مینیمم‌سازی مسائل، استفاده از تقریب‌های درجه دوم و خطی است. در لم بعدی خطا تقریب‌های F و ∇F_j به ترتیب توسط مدل‌های موضعی درجه دوم و خطی به دست می‌آوریم.

لم ۱.۴.۲. فرض کنید $V \subset U$ یک مجموعه محدب باشد. $\varepsilon > 0$ و $\delta > 0$ را در نظر بگیرید به طوری که برای هر $x, y \in V$ که $\|y - x\| < \delta$ داشته باشیم:

$$\|\nabla^2 F_j(y) - \nabla^2 F_j(x)\| < \varepsilon, \quad j = 1, 2, \dots, m. \quad (15.2)$$

آنگاه تحت این فرض، برای هر $x, y \in V$ با $\|y - x\| < \delta$ و هر $j = 1, 2, \dots, m$ داریم:

$$\|\nabla F_j(y) - [\nabla F_j(x) + \nabla^2 F_j(x)(y - x)]\| < \varepsilon \|y - x\| \quad (16.2)$$

و

$$|F_j(y) - [F_j(x) + \nabla F_j(x)^T(y - x) + \frac{1}{2}(y - x)^T \nabla^2 F_j(x)(y - x)]| < \frac{\varepsilon}{2} \|y - x\|^2. \quad (17.2)$$

به علاوه، اگر $\nabla^2 F_j$ به ازای $j = 1, 2, \dots, m$ پیوسته لیپشیتز روی V با ثابت L باشد آنگاه داریم:

$$\|\nabla F_j(y) - [\nabla F_j(x) + \nabla^2 F_j(x)(y-x)]\| < \frac{L}{2} \|y-x\|^2. \quad (18.2)$$

برهان. تعریف می‌کنیم $H_t := \nabla^2 F_j(x + t(y-x)) - \nabla^2 F_j(x)$. چون ∇F_j پیوسته و مشتق‌پذیر است، داریم:

$$\begin{aligned} \int_0^1 H_t^T(y-x) dt &= \nabla F_j(x + t(y-x)) - t(y-x) \nabla^2 F_j(x) \Big|_0^1 \\ &= \nabla F_j(y) - \nabla F_j(x) - \nabla^2 F_j(x)(y-x). \end{aligned}$$

با استفاده از رابطه بالا داریم:

$$\|\nabla F_j(y) - \nabla F_j(x) - \nabla^2 F_j(x)(y-x)\| \leq \int_0^1 \|H_t\| \|y-x\| dt.$$

از آنجایی که برای $0 < t < 1$ ، $\|x + t(y-x) - x\| < t\delta$ ، بنا به رابطه (۱۵.۲) داریم:

$$\int_0^1 \|H_t\| \|y-x\| dt < \varepsilon \|y-x\|. \quad (19.2)$$

بنابراین داریم:

$$\|\nabla F_j(y) - \nabla F_j(x) - \nabla^2 F_j(x)(y-x)\| < \varepsilon \|y-x\|.$$

در نتیجه رابطه (۱۶.۲) برقرار است.

تحت فرض پیوستگی لیپشیتز، تابع زیر انتگرال در رابطه (۱۹.۲) توسط $tL\|y-x\|^2$ کراندار می‌شود، لذا داریم:

$$\begin{aligned} \|\nabla F_j(y) - \nabla F_j(x) - \nabla^2 F_j(x)(y-x)\| &\leq \int_0^1 \|H_t\| \|y-x\| dt \\ &\leq \int_0^1 tL\|y-x\| \|y-x\| dt \\ &= \frac{L}{2} \|y-x\|^2, \end{aligned}$$

لذا رابطه (۱۸.۲) نیز برقرار است.

اکنون رابطه (۱۷.۲) را اثبات می‌کنیم. بنا به بسط سری تیلور مرتبه دوم تابع F_j ، وجود دارد $\zeta \in (0, 1)$ به‌طوری‌که

$$F_j(y) - F_j(x) - \nabla F_j(x)^T(y-x) = \frac{1}{2} (y-x)^T \nabla^2 F_j(x + \zeta(y-x))(y-x). \quad (20.2)$$

از دو طرف رابطه (۲۰.۲)، $\frac{1}{2} (y-x)^T \nabla^2 F_j(x)(y-x)$ را کم می‌کنیم با استفاده از نامساوی کوشی شوارتز و $\|x + \zeta(y-x) - x\| < \delta$ داریم:

$$\begin{aligned} |F_j(y) - F_j(x) - \nabla F_j(x)^T(y-x) - \frac{1}{2} (y-x)^T \nabla^2 F_j(x)(y-x)| &< \frac{1}{2} \|H_\zeta\| \|y-x\|^2 \\ &< \frac{\varepsilon}{2} \|y-x\|^2. \end{aligned}$$

□

لذا اثبات کامل شد.

لم زیر، مقدار بهینه مساله (۱۰۲) را توسط اندازه جهت نیوتن $s(x)$ تخمین می‌زند.

لم ۲۰۴۰۲. فرض کنید $x \in U$ باشد. $a, b \in \mathbb{R}$ را در نظر بگیرید به طوری که $0 < a \leq b$ باشد. اگر

$$aI \leq \nabla^2 F_j(x) \leq bI, \quad j = 1, 2, \dots, m,$$

آنگاه

$$\frac{a}{\gamma} \|s(x)\|^2 \leq |\theta(x)| \leq \frac{b}{\gamma} \|s(x)\|^2.$$

برهان. با در نظر گرفتن $\lambda_j(x)$ به عنوان ضریب لاگرانژین، برای محدودیت j -ام مساله (۴۰۲) و برقراری روابط (۱۱۰۲)-(۶۰۲)، روابط زیر را تعریف می‌کنیم:

$$H := \sum_{j=1}^m \lambda_j(x) \nabla^2 F_j(x), \quad v := \sum_{j=1}^m \lambda_j(x) \nabla F_j(x). \quad (21.2)$$

بنا به فرض‌های مساله و رابطه (۱۱۰۲) داریم:

$$\sum_{j=1}^m \lambda_j(x) aI \leq \sum_{j=1}^m \lambda_j(x) \nabla^2 F_j(x) \leq \sum_{j=1}^m \lambda_j(x) bI,$$

$$aI \leq H \leq bI, \quad s(x) = -H^{-1}v,$$

همچنین بنا به تعریف $\theta(x)$ ، رابطه (۱۲۰۲) داریم:

$$\begin{aligned} \theta(x) &= v^T s(x) + \frac{1}{\gamma} s(x)^T H s(x) \\ &= (-H s(x))^T s(x) + \frac{1}{\gamma} s(x)^T H s(x) \\ &= -\frac{1}{\gamma} s(x)^T H s(x). \end{aligned} \quad (22.2)$$

بنا به رابطه (۲۲۰۲) و چون $aI \leq H \leq bI$ است، داریم:

$$-\frac{b}{\gamma} s(x)^T I s(x) \leq -\frac{1}{\gamma} s(x)^T H s(x) \leq -\frac{a}{\gamma} s(x)^T I s(x)$$

□

و چون $\theta(x) \leq 0$ است، حکم برقرار است.

لم ۳۰۴۰۲. فرض کنید $x \in U$ و $a > 0$ باشد. اگر به ازای هر $j = 1, 2, \dots, m$ رابطه زیر برقرار باشد:

$$aI \leq \nabla^2 F_j(x).$$

آنگاه

$$|\theta(x)| \leq \frac{1}{\gamma a} \left\| \sum_{j=1}^m \lambda_j \nabla F_j(x) \right\|^2, \quad \sum_{j=1}^m \lambda_j = 1, \quad \forall \lambda_j \geq 0, \quad j = 1, 2, \dots, m.$$

برهان. فرض کنید $\lambda_j \geq 0$ با $\sum_{j=1}^m \lambda_j = 1$ داده شده باشند. رابطه زیر را تعریف می‌کنیم:

$$w := \sum_{j=1}^m \lambda_j \nabla F_j(x).$$

آنگاه بنا به رابطه (۱۲.۲) داریم:

$$\begin{aligned}
 \theta(x) &\geq \inf_s \sum_{j=1}^m \lambda_j \left(\nabla F_j(x)^T s + \frac{1}{\gamma} s^T \nabla^2 F_j(x) s \right) \quad (23.2) \\
 &= \inf_s \left(w^T s + \sum_{j=1}^m \lambda_j \frac{1}{\gamma} s^T \nabla^2 F_j(x) s \right) \\
 &\geq \inf_s \left(w^T s + \sum_{j=1}^m \lambda_j \frac{1}{\gamma} s^T a I s \right) \\
 &= \inf_s \left(w^T s + \frac{1}{\gamma} a \|s\|^2 \right) = -\frac{\|w\|^2}{2a}, \quad (24.2)
 \end{aligned}$$

توجه شود که تابع $s \mapsto w^T s + \frac{1}{\gamma} a \|s\|^2$ یک تابع محدب است، چون اگر تابع $g(s) = w^T s + \frac{1}{\gamma} a \|s\|^2$ را در نظر بگیریم، $\nabla^2 g(s) = a > 0$ است. بنابراین مینیم آن در یک نقطه منحصر بفردی که گرادیان آن به سمت صفر می‌رود، اتفاق می‌افتد. یعنی در یک نقطه s به طوری که $w + a s = 0$ است در نتیجه $\inf_s (w^T s + \frac{1}{\gamma} a \|s\|^2) = -\frac{\|w\|^2}{2a}$.

□

بنا به لم ۱.۳.۲، $\theta(x) \leq 0$ است، لذا حکم از رابطه (۲۴.۲) نتیجه می‌شود.

۵.۲ همگرایی فوق خطی

نتایج نظری اصلی را در این قسمت ارائه می‌کنیم. ابتدا شرایط کافی برای همگرایی فوق خطی موضعی را بیان می‌کنیم.

قضیه ۱.۵.۲. یک دنباله تولید شده توسط الگوریتم ۳.۳.۲ را با $\{x_k\}$ نشان می‌دهیم. فرض کنید که $a, b, r, \delta, \varepsilon > 0$ ، $0 < \sigma < 1$ ، $V \subset U$

الف. $aI \leq \nabla^2 F_j(x) \leq bI$ ، $\forall j = 1, 2, \dots, m$ ، $\forall x \in V$

ب. $\|\nabla^2 F_j(x) - \nabla^2 F_j(y)\| \leq \varepsilon$ ، $\|x - y\| < \delta$ ، $\forall x, y \in V$

ج. $\varepsilon/a \leq 1 - \sigma$

د. $B(x_0, r) \subset V$

هـ. $\|s(x_0)\| \leq \min\{\delta, r(1 - \varepsilon/a)\}$

آنگاه برای هر k داریم:

$$1. \quad \|x_k - x_0\| \leq \|s(x_0)\| \frac{1 - (\varepsilon/a)^k}{1 - \varepsilon/a}$$

$$2. \quad \|s(x_k)\| \leq \|s(x_0)\| (\varepsilon/a)^k$$

$$t_k = ۱.۳$$

$$\|s(x_{k+۱})\| \leq \|s(x_k)\|(\varepsilon/a) \quad ۴.$$

به علاوه دنباله $\{x_k\}$ به یک نقطه کارای موضعی، $\bar{x} \in \mathbb{R}^n$ همگراست با

$$\|\bar{x} - x_0\| \leq \frac{\|s(x_0)\|}{1 - \varepsilon/a} \leq r. \quad (۲۵.۲)$$

سرعت همگرایی $\{x_k\}$ فوق خطی است.

برهان. ابتدا نشان می‌دهیم اگر مورد ۱ و ۲ برای یک k برقرار باشند، آنگاه مورد ۳ و ۴ نیز برای آن k برقراراند.

بنا به نامساوی مثلث و مورد ۱ و ۲ داریم:

$$\begin{aligned} \|(x_k + s(x_k)) - x_0\| &\leq \|x_k - x_0\| + \|s(x_k)\| \\ &\leq \|s(x_0)\| \frac{1 - (\varepsilon/a)^k}{1 - \varepsilon/a} + \|s(x_0)\|(\varepsilon/a)^k \\ &\leq \|s(x_0)\| \frac{1 - (\varepsilon/a)^{k+1}}{1 - \varepsilon/a}. \end{aligned} \quad (۲۶.۲)$$

بنابراین از فرض‌های ج و ه داریم:

$$\begin{aligned} \|x_k - x_0\| &\leq \|s(x_0)\| \frac{1 - (\varepsilon/a)^k}{1 - \varepsilon/a} \leq r(1 - \varepsilon/a) \frac{1 - (\varepsilon/a)^k}{1 - \varepsilon/a} < r, \\ \|(x_k + s(x_k)) - x_0\| &\leq \|s(x_0)\| \frac{1 - (\varepsilon/a)^{k+1}}{1 - \varepsilon/a} \leq r(1 - \varepsilon/a) \frac{1 - (\varepsilon/a)^{k+1}}{1 - \varepsilon/a} < r, \\ \|(x_k + s(x_k)) - x_k\| &= \|s(x_k)\| \leq \|s(x_0)\|(\varepsilon/a)^k \leq \delta(\varepsilon/a)^k < \delta. \end{aligned}$$

بنا به روابط فوق داریم:

$$x_k, x_k + s(x_k) \in B(x_0, r), \quad \|(x_k + s(x_k)) - x_k\| \leq \delta. \quad (۲۷.۲)$$

با استفاده از فرض ب و لم ۱۰.۴.۲ رابطه زیر را برای $j = ۱, ۲, \dots, m$ داریم:

$$\begin{aligned} F_j(x_k + s(x_k)) &\leq F_j(x_k) + \nabla F_j(x_k)^T s(x_k) + \frac{1}{4} s(x_k)^T \nabla^2 F_j(x_k) s(x_k) + \frac{\varepsilon}{4} \|s(x_k)\|^2 \\ &\leq F_j(x_k) + \theta(x_k) + \frac{\varepsilon}{4} \|s(x_k)\|^2 \\ &= F_j(x_k) + \sigma \theta(x_k) + (1 - \sigma) \theta(x_k) + \frac{\varepsilon}{4} \|s(x_k)\|^2. \end{aligned} \quad (۲۸.۲)$$

فرض الف، لم ۲۰.۴.۲ را نتیجه می‌دهد و چون بنا به لم ۱۰.۳.۲، $\theta(x) \leq 0$ است، داریم:

$$\theta(x) < -\frac{a}{4} \|s(x)\|^2,$$

لذا داریم:

$$(1 - \sigma) \theta(x_k) + \frac{\varepsilon}{4} \|s(x_k)\|^2 \leq -\frac{a}{4} (1 - \sigma) \|s(x_k)\|^2 + \frac{\varepsilon}{4} \|s(x_k)\|^2,$$

بنا به فرض ج، $\varepsilon - a(1 - \sigma) < 0$ است، بنابراین داریم:

$$(1 - \sigma)\theta(x_k) + \frac{\varepsilon}{2}\|s(x_k)\|^2 \leq (\varepsilon - a(1 - \sigma))\frac{\|s(x_k)\|^2}{2} \leq 0, \quad (29.2)$$

با ترکیب نامساوی‌های (۲۸.۲) و (۲۹.۲) نتیجه می‌گیریم:

$$F_j(x_k + s(x_k)) \leq F_j(x_k) + \sigma\theta(x_k).$$

یعنی قاعده آرمیثو (شرط k_3 الگوریتم ۳.۳.۲) بدون نیاز به حتی یک برگشت برقرار می‌شود لذا از مورد ۳ نتیجه می‌شود:

$$t_k = 1.$$

بنابراین از (۲۷.۲) داریم:

$$x_{k+1} = x_k + s(x_k), \quad x_k, x_{k+1} \in B(x_0, r), \quad \|x_{k+1} - x_k\| \leq \delta. \quad (30.2)$$

اکنون تعریف می‌کنیم:

$$\tilde{v}_{k+1} = \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_{k+1}). \quad (31.2)$$

بنا به روابط (۶.۲) و (۷.۲) و لم ۳.۴.۲ داریم:

$$|\theta(x_{k+1})| \leq \frac{1}{2a}\|\tilde{v}_{k+1}\|^2. \quad (32.2)$$

برای تخمین $\|\tilde{v}_{k+1}\|$ ، رابطه زیر را برای $x \in U$ تعریف می‌کنیم:

$$G_k(x) := \sum_{j=1}^m \lambda_j(x_k) F_j(x), \quad (33.2)$$

که $\lambda_j(x_k)$ ضرایب لاگرانژین مساله (۱.۲) در نقطه $x = x_k$ است، لذا داریم:

$$\tilde{v}_{k+1} = \nabla G_k(x_{k+1}) = \nabla G_k(x_k + s(x_k)).$$

بدین ترتیب نتیجه می‌شود:

$$\nabla G_k(x) = \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x), \quad \nabla^2 G_k(x) = \sum_{j=1}^m \lambda_j(x_k) \nabla^2 F_j(x),$$

بنابراین از رابطه (۱۱.۲) داریم:

$$s(x_k) = -[\nabla^2 G_k(x_k)]^{-1} \nabla G_k(x_k). \quad (34.2)$$

بنا به فرض ب داریم:

$$\|\nabla^2 G_k(y) - \nabla^2 G_k(x)\| \leq \varepsilon, \quad \forall x, y \in V, \quad \|x - y\| \leq \delta,$$

همچنین بنا به لم ۱.۴.۲ داریم:

$$\|\nabla G_k(x_k + s(x_k)) - [\nabla G_k(x_k) + \nabla^2 G_k(x_k)s(x_k)]\| \leq \varepsilon\|s(x_k)\|. \quad (35.2)$$

چون رابطه (۳۴.۲) هم ارز با $\nabla G_k(x_k) + \nabla^2 G_k(x_k)s(x_k) = 0$ است از رابطه (۳۵.۲) نتیجه می‌گیریم:

$$\|\tilde{v}_{k+1}\| = \|\nabla G_k(x_{k+1})\| \leq \varepsilon\|s(x_k)\|. \quad (36.2)$$

با ترکیب روابط (۳۶.۲) و (۳۲.۲) نتیجه می‌شود:

$$|\theta(x_{k+1})| \leq \frac{(\varepsilon \|s(x_k)\|)^2}{2a}.$$

بنابراین از فرض الف و لم ۲.۴.۲ داریم:

$$\frac{a}{2} \|s(x_{k+1})\|^2 \leq \frac{(\varepsilon \|s(x_k)\|)^2}{2a}.$$

لذا داریم:

$$\|s(x_{k+1})\| \leq \|s(x_k)\| \frac{\varepsilon}{a}.$$

بنابراین مورد ۴ برقرار است.

اکنون با استقرا ثابت می‌کنیم که موارد ۱ و ۲ برای هر k برقراراند. برای $k = 0$ بنا به فرض ج، موارد ۱ و ۲ برقراراند. اگر موارد ۱ و ۲ برای یک k برقرار باشند، اثبات کردیم که موارد ۳ و ۴ نیز برای این k برقراراند و این نتیجه می‌دهد که $x_{k+1} = x_k + s(x_k)$ ، بنابراین مورد ۱ برای $k+1$ بنا به رابطه (۲۶.۲) برقرار است. مورد ۲ برای $k+1$ از مورد ۴ نتیجه می‌شود:

$$\|s(x_{k+1})\| \leq \|s(x_k)\| \frac{\varepsilon}{a} \leq \|s(x_0)\| \frac{\varepsilon^{k+1}}{a}.$$

بنابراین موارد ۱ و ۲ برای هر k برقراراند، بدین ترتیب موارد ۳ و ۴ نیز برای هر k برقراراند. به‌خصوص رابطه (۳۰.۲) برای هر k برقرار است. لذا برای هر k و j با $k > j$ از نامساوی مثلث داریم:

$$\|x_k - x_j\| \leq \sum_{i=j+1}^k \|x_i - x_{i-1}\| = \sum_{i=j+1}^k \|s(x_{i-1})\|.$$

بنا به مورد ۲ داریم:

$$\|x_k - x_j\| \leq \sum_{i=j}^{k-1} \|s(x_i)\| \leq \|s(x_0)\| \sum_{i=j}^{k-1} (\varepsilon/a)^i.$$

بنا به فرض ج، $\varepsilon/a \in (0, 1)$ است. بنابراین $\{x_k\}$ یک دنباله کوشی است و چون هر دنباله کوشی در $\mathbb{R}^n$ همگراست، $\bar{x} \in \mathbb{R}^n$ وجود دارد به‌طوری‌که

$$\lim_{k \rightarrow \infty} x_k = \bar{x}. \quad (۳۷.۲)$$

به‌علاوه بنا به مورد ۲ دنباله $\{\|s(x_k)\|\}$ همگرا به صفر است، بنابراین از فرض الف و لم ۲.۴.۲ نتیجه می‌گیریم $\theta(x_k) \rightarrow 0$ ، برای $k \rightarrow \infty$. بنابراین چون θ پیوسته (مورد ۳ لم ۱.۳.۲) است از رابطه (۳۷.۲) نتیجه می‌شود که $\theta(\bar{x}) = 0$ است. لذا بنا به مورد ۲ لم ۱.۳.۲، $\bar{x}$ برای F بحرانی است و از مورد ج قضیه ۸.۳.۱ نتیجه می‌گیریم $\bar{x}$ ، کارای موضعی است.

برای اثبات این‌که همگرایی روش نیوتن فوق خطی است، تعریف می‌کنیم:

$$r_k = \|s(x_0)\| \frac{(\varepsilon/a)^k}{1 - \varepsilon/a}, \quad \delta_k = \|s(x_0)\| (\varepsilon/a)^k, \quad k = 0, 1, \dots$$

با استفاده از نامساوی مثلث، مورد ۱ و فرض‌های د و ه نتیجه می‌گیریم:

$$B(x_k, r_k) \subset B(x_0, r) \subset V. \quad (۳۸.۲)$$

رابطه بالا برقرار است چون اگر $y \in B(x_k, r_k)$ باشد، داریم:

$$\|y - x_k\| \leq r_k = \|s(x_*)\| \frac{(\varepsilon/a)^k}{1 - \varepsilon/a}. \quad (39.2)$$

لذا از رابطه (39.2) نتیجه می‌شود:

$$\|y - x_*\| \leq \|y - x_k\| + \|x_k - x_*\| \leq \|s(x_*)\| \frac{(\varepsilon/a)^k + 1 - (\varepsilon/a)^k}{1 - \varepsilon/a} \leq r.$$

برای هر $\tau > 0$ ، تعریف می‌کنیم:

$$\hat{\varepsilon} := \min \left\{ a \frac{\tau}{1 + 2\tau}, \varepsilon \right\}.$$

برای k به اندازه کافی بزرگ داریم:

$$\|\nabla^2 F_j(x) - \nabla^2 F_j(y)\| \leq \hat{\varepsilon}, \quad \forall x, y \in B(x_k, r_k), \|x - y\| \leq \delta_k. \quad (40.2)$$

بنابراین فرض‌های الف-ه برای $\hat{\varepsilon}$ ، $\hat{\delta} = \delta_k$ ، $\hat{r} = r_k$ و $\hat{x}_* = x_k$ برقراراند. در واقع فرض‌های الف و ج از این‌که برای ε ، a و σ برقراراند، نتیجه می‌شوند. فرض‌های ب و د همان روابط (40.2) و (38.2) هستند و فرض ه از تعریف‌های $\hat{\delta}$ ، $\hat{r}$ و $\hat{\varepsilon}$ نتیجه می‌شود.

در مورد ۱ از j ، $\hat{x}_*$ و $\hat{\varepsilon}$ به ترتیب به جای x_* و ε استفاده می‌کنیم، بدین ترتیب داریم:

$$\|x_j - x_k\| \leq \|s(x_k)\| \frac{1 - (\hat{\varepsilon}/a)^j}{1 - \hat{\varepsilon}/a}.$$

زمانی که $j \rightarrow \infty$ نتیجه می‌شود:

$$\|\bar{x} - x_k\| \leq \|s(x_k)\| \frac{1}{1 - \hat{\varepsilon}/a}.$$

بنابراین داریم:

$$\|\bar{x} - x_{k+1}\| \leq \|s(x_{k+1})\| \frac{1}{1 - \hat{\varepsilon}/a} \leq \|s(x_k)\| \frac{\hat{\varepsilon}/a}{1 - \hat{\varepsilon}/a}, \quad (41.2)$$

که نامساوی آخر از مورد ۴ با جایگذاری $\hat{\varepsilon}$ به جای ε نتیجه می‌شود.

با استفاده از نامساوی مثلث، تعریف x_{k+1} و رابطه (41.2) داریم:

$$\begin{aligned} \|\bar{x} - x_k\| &\geq \|x_{k+1} - x_k\| - \|\bar{x} - x_{k+1}\| \\ &\geq \|s(x_k)\| - \|s(x_k)\| \frac{\hat{\varepsilon}/a}{1 - \hat{\varepsilon}/a} \\ &= \|s(x_k)\| \frac{1 - 2\hat{\varepsilon}/a}{1 - \hat{\varepsilon}/a}, \end{aligned}$$

لذا داریم:

$$\|s(x_k)\| \leq \frac{1 - \hat{\varepsilon}/a}{1 - 2\hat{\varepsilon}/a} \|\bar{x} - x_k\|. \quad (42.2)$$

بنا به تعریف $\hat{\varepsilon}$ ، $\hat{\varepsilon} \leq a \frac{\tau}{1 + 2\tau}$ است. بنابراین رابطه زیر نتیجه می‌شود:

$$1 - \frac{2\hat{\varepsilon}}{a} \geq 1 - 2a \frac{\tau}{a(1 + 2\tau)} = \frac{1}{1 + 2\tau} > 0.$$

از روابط (۴۱.۲) و (۴۲.۲) داریم:

$$\|\bar{x} - x_{k+1}\| \leq \|\bar{x} - x_k\| \frac{\hat{\varepsilon}/a}{1 - 2\hat{\varepsilon}/a}.$$

بنا به تعریف $\hat{\varepsilon}$ ، $\frac{\hat{\varepsilon}/a}{1 - 2\hat{\varepsilon}/a} \leq \tau$ است، بنابراین داریم:

$$\|\bar{x} - x_{k+1}\| \leq \tau \|\bar{x} - x_k\|.$$

که $\tau > 0$ دلخواه است. بدین ترتیب ثابت کردیم که سرعت همگرایی $\{x_k\}$ فوق خطی است. $\square$

قضیه ۱۰.۵.۲، هیچ فرضی برای وجود جواب کارا تابع F ایجاد نمی‌کند. در واقع این قضیه نشان می‌دهد که تحت شرایط منظم، حداقل یک نقطه کارای موضعی در مجاورت نقطه شروع، یعنی در $B(x_0, r)$ وجود دارد. به علاوه این قضیه نشان می‌دهد که دنباله تولید شده توسط الگوریتم، در این همسایگی قرار دارد و همگرایی فوق خطی به جواب بهینه، است. توجه کنید که در حالت کلی تحت این شرایط منظم، نقطه حدی تنها بهینه موضعی این همسایگی نیست.

در تحلیل موضعی استاندارد روش نیوتن کلاسیک، همه نقطه‌های شروع به اندازه کافی نزدیک به یک جواب، دنباله‌هایی تولید می‌کنند که در یک همسایگی از این جواب باقی می‌مانند و به آن همگرا هستند.

برخلاف مورد تک هدفه، در روش نیوتن چندهدفه، همه نقطه‌های شروع به اندازه کافی نزدیک به یک جواب، دنباله‌هایی تولید می‌کنند که در یک همسایگی از نقطه بهینه، باقی می‌مانند و همگرا به یک جواب کارا در آن همسایگی هستند.

اگر تحت فرض‌های کلی (پیوستگی و مثبت معین بودن هسین‌ها در U و ...) یک نقطه کارای موضعی $\hat{x}$ ، وجود داشته باشد آن‌گاه بنا به مورد الف قضیه ۸.۳.۱، $\hat{x}$ یک نقطه بحرانی برای F است. بنابراین بنا به مورد ۲ لم ۱۰.۳.۲، $\theta(\hat{x}) = 0$ است. از طرف دیگر، اگر V یک همسایگی فشرده از $\hat{x}$ باشد آن‌گاه برای هر $j = 1, 2, \dots, m$ و هر $x \in V$ داریم $aI \leq \nabla^2 F_j(x) \leq bI$ (که در این رابطه $a > 0$ و $b = \max_{j=1,2,\dots,m, x \in V, \|u\|=1} u^T \nabla^2 F_j(x) u$ و $a = \min_{j=1,2,\dots,m, x \in V, \|u\|=1} u^T \nabla^2 F_j(x) u$). چون θ پیوسته (مورد ۳ لم ۱۰.۳.۲) است و $\theta(\hat{x}) = 0$ است، لذا $r > 0$ وجود دارد به طوری که $|\theta(x)|$ برای هر $x \in B(\hat{x}, r)$ ، به اندازه کافی کوچک است. بنا به لم ۲.۴.۲، $\|s(x)\|$ در آن گوی، کوچک خواهد بود. به عبارت دیگر فرض‌های الف-ه، در این حالت برقراراند. بنابراین نتیجه زیر را بیان می‌کنیم.

نتیجه ۲.۵.۲. اگر $\hat{x}$ یک نقطه کارای موضعی برای F باشد آن‌گاه $r > 0$ وجود دارد به طوری که برای هر $x_0 \in B(\hat{x}, r) \subset U$ ، الگوریتم یک دنباله تولید می‌کند که همگرایی فوق خطی به یک نقطه کارای موضعی $\bar{x}$ است.

فرض کنید که تابع F یک مجموعه تراز فشرده به صورت $\Gamma_z = \{x \in U : F(x) \leq z\}$ داشته باشد، که $z \in \mathbb{R}^m$ است. فرض کنید $x_0 \in \Gamma_z$ باشد. از این که $\{F(x_k)\}$ -کاهنده است، نتیجه می‌گیریم همه دنباله‌های $\{x_k\}$ در این مجموعه تراز فشرده، باقی خواهند ماند. به علاوه کرانداری دنباله $\{F(x_k)\}$ با ترکیب گام k_3 الگوریتم ۳.۳.۲ نتیجه می‌دهد:

$$\lim_{k \rightarrow \infty} t_k \theta(x_k) = 0.$$

فرض کنید که $\bar{x}$ حد یک زیر دنباله از $\{x_k\}$ باشد. نتیجه ۲.۵.۲ نشان می‌دهد که حول $\bar{x}$ ، زیر دنباله متناظر، $\{\theta(x_k)\}$ باید همگرا به صفر باشد. بنا به لم ۱.۳.۲، $\bar{x}$ یک بهینه موضعی است. اکنون بنا به نتیجه ۲.۵.۲، در می‌یابیم که همه دنباله‌ها، همگرا به یک بهینه هستند، بنابراین نتیجه زیر را داریم:

نتیجه ۳.۵.۲. اگر $x_0 \in U$ در یک مجموعه تراز فشرده F ، باشد آنگاه الگوریتم با شروع از x_0 یک دنباله تولید می‌کند که همگرایی فوق‌خطی به یک نقطه کارای موضعی $\bar{x}$ است.

۶.۲ همگرایی درجه دوم

اضافه کردن این فرض که $\nabla^2 F_j$ به ازای $j = 1, 2, \dots, m$ پیوسته لیپ‌شیتز هستند. تضمین می‌کند که سرعت همگرایی الگوریتم، درجه دوم است. اکنون در قضیه زیر اثبات می‌کنیم.

قضیه ۱.۶.۲. همه فرض‌های قضیه ۱.۵.۲ را در نظر بگیرید. همچنین فرض کنید که $\nabla^2 F_j$ به ازای $j = 1, 2, \dots, m$ پیوسته لیپ‌شیتز روی V با ثابت لیپ‌شیتز L هستند. فرض کنید، $\zeta \in (0, 1)$ باشد آنگاه k_0 وجود دارد، به طوری که برای هر $k \geq k_0$ داریم:

$$\tau_k := (L/a) \|s(x_k)\| < \zeta$$

و

$$\|\bar{x} - x_{k+1}\| \leq \frac{L}{a} \frac{(1 - \tau_k)}{(1 - 2\tau_k)^2} \|\bar{x} - x_k\|^2 \leq \frac{L}{a} \frac{(1 - \zeta)}{(1 - 2\zeta)^2} \|\bar{x} - x_k\|^2, \quad (43.2)$$

که $\bar{x}$ ، نقطه کارا و حد $\{x_k\}$ است. همچنین دنباله $\{x_k\}$ همگرا درجه دوم به $\bar{x}$ است.

برهان. بنا به مورد ۲ قضیه ۱.۵.۲ برای k به اندازه کافی بزرگ، $k \geq k_0$ رابطه زیر برقرار است:

$$\tau_k = (L/a) \|s(x_k)\| < \zeta.$$

برای اثبات رابطه (۴۳.۲)، از برهان قضیه ۱.۵.۲ و روابط (۳۱.۲) و (۳۳.۲) بهره می‌گیریم. توجه داشته باشید که همه $\nabla^2 F_j$ و نیز همه $\nabla^2 G_k$ پیوسته لیپ‌شیتز هستند. بنابراین با استفاده از لم ۱.۴.۲ و روابط (۳۵.۲) و (۳۶.۲) داریم:

$$\|\tilde{v}_{k+1}\| \leq L \|s(x_k)\|^2,$$

از رابطه (۳۲.۲) و رابطه بالا نتیجه می‌شود:

$$|\theta(x_{k+1})| \leq \frac{1}{4a} L^2 \|s(x_k)\|^4.$$

بنابراین از لم ۲.۴.۲ به نامساوی زیر می‌رسیم:

$$\frac{a}{4} \|s(x_{k+1})\|^2 \leq \frac{1}{4a} L^2 \|s(x_k)\|^4,$$

که به صورت زیر تبدیل می‌شود:

$$\|s(x_{k+1})\| \leq \frac{L}{a} \|s(x_k)\|^2. \quad (44.2)$$

فرض کنید $i \geq k \geq k_0$ باشد. با استفاده از نامساوی مثلث داریم:

$$\|x_i - x_{k+1}\| \leq \sum_{j=k+2}^i \|x_j - x_{j-1}\| = \sum_{j=k+2}^i \|s(x_{j-1})\|.$$

بنابراین با در نظر گرفتن $i \rightarrow \infty$ و بنا به قضیه ۱.۵.۲ داریم:

$$\|\bar{x} - x_{k+1}\| \leq \sum_{j=k+1}^{\infty} \|s(x_j)\|. \quad (۴۵.۲)$$

لذا با ترکیب روابط (۴۴.۲) و (۴۵.۲) و استفاده از $1/2 < \tau_k := (L/a)\|s(x_k)\| < 1/2$ داریم:

$$\begin{aligned} \|\bar{x} - x_{k+1}\| &\leq (L/a)\|s(x_k)\|^2 + (L/a)^3\|s(x_k)\|^4 + (L/a)^5\|s(x_k)\|^6 + \dots \\ &= (L/a)\|s(x_k)\|^2(1 + \tau_k^2 + \tau_k^4 + \dots) \\ &\leq (L/a)\|s(x_k)\|^2 \sum_{j=0}^{\infty} \tau_k^j \\ &= (L/a)\|s(x_k)\|^2 \frac{1}{1 - \tau_k}. \end{aligned} \quad (۴۶.۲)$$

بنابراین از نامساوی مثلث نتیجه می‌شود:

$$\begin{aligned} \|\bar{x} - x_k\| &\geq \|x_{k+1} - x_k\| - \|\bar{x} - x_{k+1}\| \\ &\geq \|s(x_k)\| - \|s(x_k)\|^2 \frac{L/a}{1 - \tau_k} \\ &= \|s(x_k)\| \frac{1 - 2\tau_k}{1 - \tau_k}, \end{aligned}$$

که آخرین تساوی از تعریف τ_k نتیجه می‌شود. از آنجایی که $\frac{1-2\tau_k}{1-\tau_k} > 0$ داریم:

$$\|s(x_k)\| \leq \|\bar{x} - x_k\| \frac{1 - \tau_k}{1 - 2\tau_k}. \quad (۴۷.۲)$$

اکنون از روابط (۴۶.۲) و (۴۷.۲) رابطه زیر را داریم:

$$\|\bar{x} - x_{k+1}\| \leq (L/a) \left[\|\bar{x} - x_k\| \frac{1 - \tau_k}{1 - 2\tau_k} \right]^2 \frac{1}{1 - \tau_k} = \|\bar{x} - x_k\|^2 \frac{(L/a)(1 - \tau_k)}{(1 - 2\tau_k)^2}.$$

بدین ترتیب اولین نامساوی رابطه (۴۳.۲) را اثبات کردیم. برای اثبات آخرین نامساوی مشابه اثبات بالا، روابط زیر را داریم:

$$\begin{aligned} \|\bar{x} - x_{k+1}\| &< (L/a)\|s(x_k)\|^2(1 + \zeta^2 + \zeta^4 + \dots) \\ &\leq (L/a)\|s(x_k)\|^2 \sum_{j=0}^{\infty} \zeta^j \\ &= (L/a)\|s(x_k)\|^2 \frac{1}{1 - \zeta}. \end{aligned}$$

لذا از نامساوی مثلث داریم:

$$\|\bar{x} - x_k\| \geq \|x_{k+1} - x_k\| - \|\bar{x} - x_{k+1}\| > \|s(x_k)\| - \|s(x_k)\|^2 \frac{L/a}{1 - \zeta} \geq \|s(x_k)\| \frac{1 - 2\zeta}{1 - \zeta}.$$

بدین ترتیب داریم:

$$\|\bar{x} - x_{k+1}\| < (L/a) \left[\|\bar{x} - x_k\| \frac{1-\zeta}{1-2\zeta} \right]^2 \frac{1}{1-\zeta} = \|\bar{x} - x_k\|^2 \frac{(L/a)(1-\zeta)}{(1-2\zeta)^2}.$$

لذا رابطه (۴۳.۲) را اثبات کردیم و همچنین بنا به روابط بالا دنباله $\{x_k\}$ همگرای درجه دوم به $\bar{x}$ است.

□

تحت فرض پیوستگی لیپشیتز هسین‌ها، نتیجه‌های ۲.۵.۲ و ۳.۵.۲ به صورت زیر تبدیل می‌شوند.

نتیجه ۲.۶.۲. اگر همه $\nabla^2 F_j$ پیوسته لیپشیتز باشند و $\hat{x}$ نقطه کارا برای تابع F باشد آن‌گاه $r > 0$ وجود دارد به طوری که برای هر $x_0 \in B(\hat{x}, r) \subset U$ ، الگوریتم یک دنباله تولید می‌کند که همگرای درجه دوم به نقطه کارای $\bar{x}$ است.

نتیجه ۳.۶.۲. اگر $x_0 \in U$ ، در یک مجموعه تراز فشرد F ، باشد که همه $\nabla^2 F_j$ پیوسته لیپشیتز هستند آن‌گاه الگوریتم با شروع از x_0 یک دنباله تولید می‌کند که همگرای درجه دوم به نقطه کارای $\bar{x}$ است.

۷.۲ مثال‌های عددی

در این قسمت مثال‌هایی را بیان می‌کنیم که با به کارگیری الگوریتم ۳.۳.۲ در برنامه متلب (R2013a) حل می‌شوند. باید توجه داشت که جواب کارا (کارای ضعیف) برای مسائل بهینه‌سازی چندهدفه منحصریفر نیست زیرا اگر با نقطه‌های شروع متفاوت الگوریتم را آغاز کنیم، جواب‌های کارای متفاوتی را خواهیم داشت. در مثال‌های ارائه شده ما از شرط توقف $\theta(x^{(k)}) \geq -\varepsilon$ استفاده می‌کنیم که $\varepsilon > 0$ داده شده و $x^{(k)}$ نقطه پایانی است. ما در این قسمت حداکثر تعداد تکرارها را ۱۰۰ در نظر می‌گیریم.

مثال ۱.۷.۲. مساله تک‌هدفه زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}^2} F(x),$$

$$F(x) = 2x_1^2 + 2x_1x_2 + 4/5x_2^2.$$

با نقطه شروع $x^{(0)} = (-0.01, 0.01)$ و شرط توقف $\theta(x^{(k)}) \geq -10^{-8}$ مساله (۱.۲) را حل می‌کنیم.

$$\nabla F(x) = \begin{bmatrix} 4x_1 + 2x_2 \\ 2x_1 + 8x_2 \end{bmatrix}, \quad \nabla^2 F(x) = \begin{bmatrix} 4 & 2 \\ 2 & 8 \end{bmatrix}.$$

لذا این مساله محدب اکید است $s(x^k)$ و $\theta(x^k)$ را به ترتیب از روابط (۳.۲) و (۲.۲) محاسبه می‌شوند. در تکرار اول $\theta(x^{(1)}) = -10^{-4} \times 4/5 \times 14$ و $x^{(1)} = (10^{-3} \times 0.0018, 10^{-3} \times 0.0030)$ است که در شرط توقف صدق نمی‌کند. بنابراین الگوریتم را دوباره اجرا می‌کنیم. جواب مساله در دومین گام نقطه $\theta(x^{(2)}) = -10^{-9} \times 4/2113$ است با $x^{(2)} = (10^{-6} \times 0.2810, -10^{-6} \times 0.6462) \simeq (0, 0)$.

بنابراین شرط توقف برقرار است. در واقع $\theta(x^{(2)})$ به اندازه کافی نزدیک به صفر است. لذا بنا به لم ۱.۳.۲ نقطه $x^{(2)}$ بحرانی است. با توجه به این که مساله محدب اکید است، بنا به قضیه ۸.۳.۱ این نقطه بحرانی، نقطه کارا است. این مساله را با نقاط شروع تصادفی و شرط توقف بالا حل می‌کنیم، خلاصه نتایج در جدول زیر نشان داده شده است.

جدول ۱۰.۲: جدول مربوط به مثال ۱۰.۷.۲

نقطه شروع	تعداد تکرارها	$x^{(k)}$ = نقطه پایانی	$F(x^{(k)})$
(۰/۶۴۷۷, ۰/۴۵۰۹)	۱۰۰	$10^{-5} \times (0/3019, -0/1071)$	$10^{-11} \times 1/6928$
(۰/۲۹۶۳, ۰/۷۴۴۷)	۱۰۰	$10^{-6} \times (-0/8198, -0/2063)$	$10^{-12} \times 1/8741$
(۰/۶۸۶۸, ۰/۱۸۳۵)	۲	$10^{-5} \times (0/4411, 0/1527)$	$10^{-11} \times 3/5236$
(۰/۶۲۵۶, ۰/۷۸۰۲)	۱۰۰	$10^{-4} \times (0/6375, 0/0033)$	$10^{-9} \times 8/1697$
(۰/۹۲۹۴, ۰/۷۷۵۷)	۱۰۰	$10^{-6} \times (0/4223, 0/1442)$	$10^{-13} \times 5/7213$

مثال ۲.۷.۲. مساله دوهدفه زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}} F(x) = (F_1(x), F_2(x))$$

$$F_1(x) = x^2 - 4$$

$$F_2(x) = (x - 1)^2.$$

$$\nabla F_1(x) = 2x, \quad \nabla^2 F_1(x) = 2,$$

و

$$\nabla F_2(x) = 2(x - 1), \quad \nabla^2 F_2(x) = 2.$$

این مساله محدب را با نقطه شروع $x^{(0)} = 5$ و شرط توقف $\theta(x^{(k)}) \geq -10^{-20}$ حل می‌کنیم. $s(x^{(0)})$ و $\theta(x^{(0)})$ به ترتیب از روابط (۴۹.۲) و (۴۸.۲) محاسبه می‌شوند.

$$\theta(x^{(0)}) = \inf_{s \in \mathbb{R}^n} \max_{j=1,2} \nabla F_j(x^{(0)})^T s + \frac{1}{\gamma} s^T \nabla^2 F_j(x^{(0)}) s \quad (48.2)$$

و

$$s(x^{(0)}) = \arg \min_{s \in \mathbb{R}^n} \max_{j=1,2} \nabla F_j(x^{(0)})^T s + \frac{1}{\gamma} s^T \nabla^2 F_j(x^{(0)}) s. \quad (49.2)$$

در تکرار اول $x^{(1)} = 1/0000$ و $\theta(x^{(1)}) = -16/0000$ است که در شرط توقف صدق نمی‌کند. در تکرار دوم $x^{(2)} = 1/0000$ و $\theta(x^{(2)}) = -10^{-16} \times 1/1102$ است که با توجه به معیار توقف الگوریتم ادامه پیدا می‌کند تا در نقطه $x^{(6)} = 1/0000$ شرط توقف برقرار و الگوریتم متوقف می‌شود. در واقع $\theta(x^{(6)})$ به اندازه کافی نزدیک به صفر است. لذا بنا به لم ۱.۳.۲ نقطه $x^{(6)}$ بحرانی است. با توجه به این که مساله محدب اکید است، بنا به قضیه ۸.۳.۱ این نقطه بحرانی، نقطه کارا است. این مساله را با نقاط شروع تصادفی و شرط توقف بالا حل می‌کنیم، خلاصه نتایج در جدول ۲.۲ نشان داده شده است.

جدول ۲.۲: جدول مربوط به مثال ۲.۷.۲

نقطه شروع	تعداد تکرارها	$x^{(k)}$ = نقطه پایانی	$(F_1(x^{(k)}), F_2(x^{(k)}))$
۰/۳۰۶۳	۲	۰.۳۰۶۳	(-۳/۹۰۶۲, ۰/۴۸۱۲)
۰/۷۹۴۸	۲	۰.۷۹۴۸	(-۳/۳۶۸۲, ۰/۰۴۲۱)
۰/۳۷۸۶	۲	۰.۳۷۸۶	(-۳/۸۵۶۷, ۰/۳۸۶۱)
۰/۵۳۲۸	۱	۰.۵۳۲۸	(-۳/۷۱۶۱, ۰/۲۱۸۳)
۰/۱۹۴۸	۱۰۰	۰.۱۹۴۸	(-۳/۹۶۲۱, ۰/۶۴۸۴)

مثال ۳.۷.۲. مساله دوهدفه زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}^3} F(x) = (F_1(x), F_2(x))$$

$$F_1(x) = x_1^2 + x_2^2 + x_3^2$$

$$F_2(x) = \frac{1}{3}x_1^2 + \frac{1}{3}x_2^2 + \frac{1}{3}x_3^2 - \frac{4}{3}x_1 - \frac{4}{3}x_2 - \frac{4}{3}x_3 + 4.$$

$$\nabla F_1(x) = (2x_1, 2x_2, 2x_3), \quad \nabla F_2(x) = \left(\frac{2}{3}x_1 - \frac{4}{3}, \frac{2}{3}x_2 - \frac{4}{3}, \frac{2}{3}x_3 - \frac{4}{3}\right)$$

و

$$\nabla^2 F_1(x) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \nabla^2 F_2(x) = \begin{bmatrix} \frac{2}{3} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} \end{bmatrix}.$$

با نقطه شروع $x^{(0)} = (1, 0, 1)$ و شرط توقف $\theta(x^{(k)}) \geq -10^{-8}$ جواب مساله بعد از دومین گام در نقطه $x^{(2)} = (0/7321, 0/7321, 0/7321)$ با $\theta(x^{(2)}) = -10^{-10} \times 1/0799$ است. $\theta(x^{(2)})$ به اندازه کافی نزدیک به صفر است. لذا بنا به لم ۱.۳.۲، نقطه $x^{(2)}$ بحرانی است. با توجه به این که مساله محدب اکید است، بنا به قضیه ۸.۳.۱ این نقطه بحرانی، نقطه کارا است. این مساله را با نقاط شروع تصادفی و شرط توقف بالا حل می کنیم. می توانید نتایج را در جدول ۳.۲ مشاهده کنید.

جدول ۳.۲: جدول مربوط به مثال ۳.۷.۲

نقطه شروع	تعداد تکرارها	$x^{(k)}$ = نقطه پایانی	$(F_1(x^{(k)}), F_2(x^{(k)}))$
(۰/۱۷۰۷, ۰/۲۲۷۷, ۰/۴۳۵۷)	۱۰۰	(۰/۲۸۳۱, ۰/۲۸۳۱, ۰/۲۸۳۱)	(۰/۲۴۰۴, ۲/۹۴۷۸)
(۰/۹۲۳۴, ۰/۴۳۰۲, ۰/۱۸۴۸)	۲	(۰/۵۴۳۷, ۰/۵۴۳۷, ۰/۵۴۳۷)	(۰/۸۸۶۷, ۲/۱۲۰۹)
(۰/۴۰۸۷, ۰/۵۹۴۹, ۰/۲۶۲۲)	۲	(۰/۴۲۸۴, ۰/۴۲۸۴, ۰/۴۲۸۴)	(۰/۵۵۰۷, ۲/۴۶۹۸)
(۰/۳۱۸۸, ۰/۴۲۴۲, ۰/۵۰۷۹)	۲	(۰/۴۱۹۰, ۰/۴۱۹۰, ۰/۴۱۹۰)	(۰/۵۲۶۸, ۲/۴۹۹۴)
(۰/۲۶۲۵, ۰/۸۰۱۰, ۰/۰۲۹۲)	۲	(۰/۴۰۲۰, ۰/۴۰۲۰, ۰/۴۰۲۰)	(۰/۴۸۴۸, ۲/۵۵۳۶)

مثال ۴.۷.۲. مساله سه‌هدفه زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}^3} F(x) = (F_1(x), F_2(x), F_3(x))$$

$$F_1(x) = x_1^2 + 8 \times 10^{-5} x_2^2$$

$$F_2(x) = 4 \times 10^{-2} x_1^2 + x_2^2$$

$$F_3(x) = 6 \times 10^{-2} x_1^2 + 1 \times 10^{-2} x_2^2 + x_3^2.$$

$$\nabla F_1(x) = \begin{bmatrix} 2x_1 \\ 16 \times 10^{-5} x_2 \\ 0 \end{bmatrix}, \quad \nabla^2 F_1(x) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 16 \times 10^{-5} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\nabla F_2(x) = \begin{bmatrix} 8 \times 10^{-5} x_1 \\ 2x_2 \\ 0 \end{bmatrix}, \quad \nabla^2 F_2(x) = \begin{bmatrix} 8 \times 10^{-5} & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\nabla F_3(x) = \begin{bmatrix} 12 \times 10^{-2} x_1 \\ 2 \times 10^{-2} x_2 \\ 2x_3 \end{bmatrix}, \quad \nabla^2 F_3(x) = \begin{bmatrix} 12 \times 10^{-2} & 0 & 0 \\ 0 & 2 \times 10^{-2} & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

و

باید توجه داشت توابع هدف محدب اکید نیستند، لذا با حل این مساله با الگوریتم نیوتن ۳.۳.۲ اگر $x^{(k)}$ نقطه بحرانی باشد بنا به قضیه ۸.۳.۱ این نقطه بحرانی، نقطه کارا ضعیف است. با نقطه شروع $x^{(0)} = (0.1, 0.1, 0.1)$ و شرط توقف $\theta(x^{(k)}) \geq -10^{-9} \times 7$ جواب مساله بعد از سه گام در نقطه $x^{(3)} = (-0.00000, 0.00000, 0.1912) \simeq (0, 0, 0)$ است. $\theta(x^{(3)})$ به اندازه کافی نزدیک به صفر است. لذا بنا به لم ۱.۳.۲ نقطه $x^{(3)}$ بحرانی است. این مساله را با نقاط شروع تصادفی و شرط توقف بالا حل می‌کنیم. نتایج را در جدول ۴.۲ خلاصه کردیم.

جدول ۴.۲: جدول مربوط به مثال ۴.۷.۲

نقطه پایانی	تعداد تکرارها	نقطه شروع
$(-0.00000, 0.00000, 0.1904)$	۳	$(0.9133, 0.7962, 0.9887)$
$(0.00000, 0.00000, 0.2698)$	۱۰۰	$(0.3354, 0.6797, 0.1366)$
$(-0.00000, 0.00000, 0.4746)$	۱۰۰	$(0.4886, 0.5785, 0.2373)$
$(0.00000, -0.00000, 0.0755)$	۱۰۰	$(0.3674, 0.9880, 0.0377)$

فصل ۳

یک روش شبه‌نیوتن برای حل مسائل چندهدفه محدب

۱.۳ مقدمه

در این فصل به بیان روش‌های شبه‌نیوتن می‌پردازیم که تا حدودی مشابه با روش نیوتن عمل می‌کنند. با این فرض که محاسبه و به‌کارگیری ماتریس هسین غیرعملی یا پرهزینه است، ایده‌ی زیربنایی روش‌های شبه‌نیوتن بر استفاده از تقریب ماتریس هسین به‌جای هسین دقیق مورد نیاز در روش نیوتن قرار دارد. با توجه به این‌که شکل این تقریب در روش‌های متفاوت تغییر می‌کند، روش‌های شبه‌نیوتن انواع مختلفی را شامل می‌شود.

۲.۳ روش شبه‌نیوتن برای حل مسائل بهینه‌سازی تک‌هدفه

ابتدا روش شبه‌نیوتن را برای حل مسائل بهینه‌سازی تک‌هدفه بیان می‌کنیم. مساله تک‌هدفه زیر را در نظر بگیرید:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in U \subset \mathbb{R}^n. \end{aligned}$$

که $f \in C^1(U, \mathbb{R})$ و U یک مجموعه باز است.

روش شبه‌نیوتن، تعمیم یافته روش وتر^۱ است. در مسائل تک بعدی، روش وتر^۱ از تقریب زیر استفاده می‌کند،

$$f''(x_{k+1}) \simeq \frac{f'(x_{k+1}) - f'(x_k)}{x_{k+1} - x_k}. \quad (1.3)$$

^۱secant method

روش وتری با استفاده از رابطه زیر، در هر گام بروز می‌شود،

$$x_{k+1} = x_k - \frac{f'(x_k)(x_k - x_{k-1})}{f'(x_k) - f'(x_{k-1})}.$$

فرمول (۱.۳) برای مسائل چند بعدی به صورت زیر است:

$$\nabla^2 f(x_{k+1})(x_{k+1} - x_k) \simeq \nabla f(x_{k+1}) - \nabla f(x_k). \quad (2.3)$$

در روش شبه‌نیوتن $B(x_k) \in \mathbb{R}^{n \times n}$ ، تقریبی از $\nabla^2 f(x_k)$ در k -امین تکرار است، بنابراین رابطه (۲.۳) را به صورت زیر بیان می‌کنیم:

$$B(x_{k+1})(x_{k+1} - x_k) \simeq \nabla f(x_{k+1}) - \nabla f(x_k).$$

این رابطه معادله وتری^۲ نامیده می‌شود. الگوریتم شبه‌نیوتن، در حالت کلی به صورت زیر است:

الگوریتم ۱.۲.۳. [۴۱] مقداردهی اولیه $k = 0$.

i_1 . نقطه اولیه $x_0 \in \mathbb{R}^n$ را انتخاب کنید.

i_2 . ماتریس متقارن مثبت معین $B(x_0)$ و $\varepsilon > 0$ را در نظر بگیرید.

گام‌های تکراری:

k_1 . اگر نقطه x_k در رابطه $\|\nabla f(x_k)\| \leq \varepsilon$ صدق کرد، متوقف شوید.

k_2 . جهت $d(x_k)$ را با حل معادله خطی زیر به دست آورید،

$$B(x_k)d(x_k) + \nabla f(x_k) = 0.$$

k_3 . جستجو باید در امتداد جهت $d(x_k)$ باشد، طول گام α_k را پیدا کنید و قرار دهید:

$$x_{k+1} = x_k + \alpha_k d(x_k).$$

k_4 . یک ماتریس متقارن مثبت معین $B(x_{k+1})$ را انتخاب کنید. این ماتریس باید در معادله وتری زیر صدق کند:

$$B(x_{k+1})s_k = y(x_k), \quad (3.3)$$

که $s_k = x_{k+1} - x_k$ ، $y(x_k) = \nabla f(x_{k+1}) - \nabla f(x_k)$ هستند.

k_5 . قرار دهید $k = k + 1$ و به گام k_1 بروید.

قضیه ۲.۲.۳. [۲۱] اگر $B(x_k)$ ماتریس متقارن مثبت معین باشد و $B(x_{k+1})$ بروز شده آن باشد آن‌گاه $B(x_{k+1})$ مثبت معین است اگر در رابطه زیر که شرط انحنا^۳ نامیده می‌شود، صدق کند:

$$s_k^T y(x_k) > 0.$$

^۲secant equation

^۳curvature condition

اگر تابع هدف f محدب قوی باشد، شرط انحنا برقرار است. برای بررسی این موضوع می‌توانید به مرجع [۳۰] مراجعه کنید. یک فاکتور مهم در استفاده از روش شبه‌نیوتن، بروز کردن فرمولی است که یک ماتریس $B(x_{k+1})$ را از ماتریس جاری $B(x_k)$ با استفاده از اطلاعات موجود در k -امین تکرار نتیجه می‌دهد.

روش شبه‌نیوتنی که از فرمول (۴.۳) برای بروز کردن $B(x_k)$ استفاده می‌کند، روش BFGS^۴ نامیده می‌شود که معروفترین و پرکاربردترین روش شبه‌نیوتن می‌باشد. نام این روش از حروف اول اسم ارائه دهندگان این روش گرفته شده است.

$$B(x_{k+1}) = B(x_k) - \frac{B(x_k)s_k s_k^T B(x_k)}{s_k^T B(x_k)s_k} + \frac{y(x_k)y(x_k)^T}{s_k^T y(x_k)}. \quad (۴.۳)$$

همچنین روش شبه‌نیوتنی که از فرمول زیر برای بروز کردن $B(x_k)$ استفاده می‌کند، روش DFP^۵ نامیده می‌شود،

$$B(x_{k+1}) = (I - \gamma(x_k)y(x_k)s_k^T)B(x_k)(I - \gamma(x_k)s_k y(x_k)^T) + \gamma(x_k)y(x_k)y(x_k)^T, \quad (۵.۳)$$

که $\gamma(x_k) = \frac{1}{y(x_k)^T s_k}$ است. قضیه زیر همگرایی روش‌های شبه‌نیوتن را نشان می‌دهد.

قضیه ۳.۲.۳. [۲۱] فرض کنید $f: \mathbb{R}^n \rightarrow \mathbb{R}$ باشد. x_0 نقطه اولیه و x_k دنباله تولید شده با رابطه $x_{k+1} = x_k + \alpha_k d(x_k)$ باشد که $\alpha_k \geq 0$ است. فرض کنید روابط زیر برقرار باشند:

۱. مجموعه $S = \{x : f(x) \leq f(x_0)\}$ کراندار باشد،

۲. توابع f ، ∇f و $\nabla^2 f$ به ازای هر $x \in S$ پیوسته باشند،

۳. $\nabla^2 f$ به ازای هر x ، مثبت معین باشد،

۴. جهت جستجو $d(x_k)$ با رابطه زیر محاسبه شود:

$$B(x_k)d(x_k) = -\nabla f(x_k),$$

که $B(x_0) = I$ و ماتریس $B(x_k)$ از روابط (۴.۳) یا (۵.۳) به دست می‌آید،

۵. طول گام α_k در روابط زیر صدق کند:

$$f(x_k + \alpha_k d(x_k)) \leq f(x_k) + c_1 \alpha_k \nabla f(x_k)^T d(x_k),$$

$$\nabla f(x_k + \alpha_k d(x_k))^T d(x_k) \geq c_2 \nabla f(x_k)^T d(x_k),$$

که $0 < c_1 < c_2 < 1$.

آنگاه $\lim_{k \rightarrow \infty} x_k = x^*$ که x^* یک مینیمم سراسری و منحصر بفرد از f روی مجموعه S است و سرعت همگرایی x_k فوق خطی است.

^۴Broyden-Fletcher-Goldfarb-Shanno

^۵Davidon-Fletcher-Powell

۳.۳ روش شبه‌نیوتن برای حل مسائل چندهدفه

در این قسمت، ابتدا جهت شبه‌نیوتن را برای مساله چندهدفه (۵.۱) بیان می‌کنیم، با در نظر گرفتن این‌که هر تابع هدف F_j ، محدب قوی است. از تعریف‌های مشابه که در فصل ۲ برای روش نیوتن مطرح کردیم، بهره می‌گیریم. به جای استفاده از ماتریس‌های هسین دقیق، آن‌ها را با روش BFGS تقریب می‌زنیم. جهت $p(x)$ یک جهت شبه‌نیوتن برای $x \in U$ است، اگر جواب بهینه‌ای از مساله زیر باشد:

$$\begin{aligned} \min \quad & \max_{j=1,2,\dots,m} \nabla F_j(x)^T p + \frac{1}{\gamma} p^T B_j(x) p \\ \text{s.t.} \quad & p \in \mathbb{R}^n, \end{aligned} \quad (۶.۳)$$

که در آن $B_j(x)$ تقریبی از $\nabla^2 F_j(x)$ به ازای $j = 1, 2, \dots, m$ است. در این قسمت برای بروز کردن تقریب هسین از رابطه (۴.۳) استفاده می‌کنیم. جواب بهینه و مقدار بهینه مساله (۶.۳) را به ترتیب با $\tau(x)$ و $p(x)$ نشان می‌دهیم و روابط زیر را داریم:

$$\tau(x) = \min_{p \in \mathbb{R}^n} \max_{j=1,2,\dots,m} \nabla F_j(x)^T p + \frac{1}{\gamma} p^T B_j(x) p \quad (۷.۳)$$

و

$$p(x) = \arg \min_{p \in \mathbb{R}^n} \max_{j=1,2,\dots,m} \nabla F_j(x)^T p + \frac{1}{\gamma} p^T B_j(x) p. \quad (۸.۳)$$

مساله (۶.۳) محدب اکید است، زیرا $B_j(x)$ به ازای $j = 1, 2, \dots, m$ مثبت معین هستند. لذا بنا به قضیه ۲۳.۲.۱ همیشه یک مینیمم منحصر بفرد دارد. مساله (۶.۳) را می‌توان به شکل یک مساله بهینه‌سازی محدب و درجه دوم تبدیل کرد. مساله (۶.۳) معادل با مساله زیر است:

$$\begin{aligned} \min \quad & g(t, p) = t \\ \text{s.t.} \quad & \nabla F_j(x)^T p + \frac{1}{\gamma} p^T B_j(x) p - t \leq 0 \quad j = 1, 2, \dots, m \\ & (t, p) \in \mathbb{R} \times \mathbb{R}^n. \end{aligned} \quad (۹.۳)$$

لاگرانژین این مساله به صورت زیر است:

$$L((t, p), \lambda) = t + \sum_{j=1}^m \lambda_j (\nabla F_j(x)^T p + \frac{1}{\gamma} p^T B_j(x) p - t).$$

بنا به تعریف ۲۵.۲.۱ شرایط KKT برای مساله (۹.۳) به صورت زیر است:

$$\sum_{j=1}^m \lambda_j (\nabla F_j(x) + B_j(x) p) = 0, \quad (۱۰.۳)$$

$$\lambda_j \geq 0, \quad j = 1, 2, \dots, m, \quad (۱۱.۳)$$

$$\sum_{j=1}^m \lambda_j = 1, \quad (۱۲.۳)$$

$$\nabla F_j(x)^T p + \frac{1}{\gamma} p^T B_j(x) p \leq t, \quad j = 1, 2, \dots, m, \quad (۱۳.۳)$$

$$\lambda_j (\nabla F_j(x)^T p + \frac{1}{\gamma} p^T B_j(x) p - t) = 0, \quad j = 1, 2, \dots, m. \quad (۱۴.۳)$$

که شرایط (۱۰.۳)، (۱۱.۳) و (۱۲.۳) شرایط شدنی دوگان هستند، شرط (۱۳.۳) شرط شدنی اولیه است و شرط (۱۴.۳)، شرط مکمل زائد است.

جواب $(\tau(x), p(x))$ یک جواب منحصر بفرد برای مساله (۹.۳) است. همچنین این مساله محدب است و در شرایط منظم اسلاتر صدق می‌کند زیرا یک نقطه اسلاتر مانند نقطه $(1, 0)$ وجود دارد به‌طوری‌که

$$\nabla F_j(x)^T 0 + \frac{1}{p} 0^T B_j(x) 0 - 1 = -1 < 0, \quad j = 1, 2, \dots, m.$$

چون این مساله محدب، جواب بهینه دارد، لذا بنا به قضیه ۳۳.۲.۱ ضرایب لاگرانژین، $\lambda_j = \lambda_j(x)$ به ازای $j = 1, 2, \dots, m$ وجود دارند که با $p = p(x)$ و $t = \tau(x)$ در رابطه‌های (۱۴.۳)–(۱۰.۳) صدق می‌کنند.

از رابطه (۱۰.۳) نتیجه می‌شود:

$$\sum_{j=1}^m \lambda_j(x) \nabla F_j(x) = - \sum_{j=1}^m \lambda_j(x) B_j(x) p(x).$$

چون B_j مثبت معین است و حداقل یکی از λ_j مخالف صفر است، عبارت $\sum_{j=1}^m \lambda_j(x) B_j(x)$ معکوس‌پذیر است. لذا داریم:

$$p(x) = - \left[\sum_{j=1}^m \lambda_j(x) B_j(x) \right]^{-1} \sum_{j=1}^m \lambda_j(x) \nabla F_j(x). \quad (15.3)$$

تعدادی از ویژگی‌های τ را که از رابطه (۷.۳) حاصل می‌شود، در لم زیر بیان می‌کنیم.

لم ۱.۳.۳. فرض کنید به ازای هر $x \in U$ ، $B_j(x) > 0$ باشد آنگاه

۱. برای هر $x \in U$ ، $\tau(x) \leq 0$ است.

۲. شرایط زیر معادل‌اند:

الف. نقطه x بحرانی نیست،

ب. $\tau(x) < 0$ است،

ج. $p(x) \neq 0$ است.

۳. تابع τ پیوسته است.

برهان. اثبات موارد ۱ و ۲ مشابه اثبات ۱ و ۲ قضیه ۱.۳.۲ است. با توجه به این‌که ماتریس‌های هسین $\nabla^2 F_j(x)$ ، $j = 1, 2, \dots, m$ را با $B_j(x)$ ، $j = 1, 2, \dots, m$ جایگزین می‌کنیم، که مثبت معین هستند. مورد ۳ را اثبات می‌کنیم. کافی است که پیوستگی τ را در مجموعه معین $W \subset U$ که W یک مجموعه فشرده و دلخواه است، نشان دهیم. بنا به مورد ۱ برای هر $x \in U$ داریم:

$$\frac{1}{p} p(x)^T B_j(x) p(x) \leq - \nabla F_j(x)^T p(x), \quad j = 1, 2, \dots, m, \quad (16.3)$$

چون تابع F دوبار به‌طور پیوسته مشتق‌پذیر است و تقریب‌های $B_j(x)$ برای هر $j = 1, 2, \dots, m$ و $x \in W$ مثبت معین هستند، بنابراین، مقدار ویژه‌های $B_j(x)$ ، به ازای هر $j = 1, 2, \dots, m$ کراندار یکنواخت دور از صفر روی W هستند. لذا، $K, L \in \mathbb{R}_{++}$ وجود دارند به‌طوری‌که

$$K = \max_{x \in W, j=1,2,\dots,m} \|\nabla F_j(x)\|$$

و

$$L = \min_{x \in W, \|e\|=1, j=1,2,\dots,m} e^T B_j(x) e.$$

با استفاده از نامساوی کوشی شوارتز و ترکیب رابطه (۱۶.۳) با تعریف‌های K و L به ازای هر $x \in W$ و هر $j = 1, 2, \dots, m$ داریم:

$$\frac{1}{4} L \|p(x)\|^2 \leq \|\nabla F_j(x)\| \|p(x)\| \leq K \|p(x)\|.$$

بنابراین به ازای هر $x \in W$ داریم:

$$\|p(x)\| \leq 4 \frac{K}{L}.$$

لذا جهت‌های شبه‌نیوتن روی W کراندار یکنواخت هستند.

یک خانواده از توابع $\{\varphi_{x,j}\}_{x \in W, j=1,2,\dots,m}$ تعریف می‌کنیم:

$$\varphi_{x,j} : W \rightarrow \mathbb{R},$$

$$z \mapsto \nabla F_j(z)^T p(x) + \frac{1}{4} p(x)^T B_j(z) p(x).$$

نشان می‌دهیم که این خانواده از توابع، پیوسته یکسان است. ابتدا یک پوشش باز $\{B(z, \delta_z)\}_{z \in W}$ با شعاع $\delta_z \in \mathbb{R}_{++}$ تعریف می‌کنیم. برای هر $z \in W$ یک گوی $B(z, \delta_z)$ داریم که برای هر $\varepsilon_z \in \mathbb{R}_{++}$ کوچک، یک $\delta_z \in \mathbb{R}_{++}$ وجود دارد به‌طوری‌که برای هر $w \in B(z, \delta_z)$ و هر $j = 1, 2, \dots, m$ نامساوی‌های زیر را داریم:

$$\|B_j(w) - \nabla^2 F_j(z)\| < \frac{\varepsilon_z}{4}$$

و

$$\|\nabla^2 F_j(w) - \nabla^2 F_j(z)\| < \frac{\varepsilon_z}{4}.$$

باید توجه داشت چون ماتریس‌های هسین پیوسته هستند، نامساوی برقرار است.

از آنجایی که W یک فضای فشرده است، یک زیر پوشش متناهی مانند $\{B(z_i, \delta_{z_i}); i = 1, 2, \dots, l\}$ از $\{B(z, \delta_z)\}_{z \in W}$ وجود دارد. همچنین داریم:

$$\begin{aligned} \varphi_{x,j}(z) &= \nabla F_j(z)^T p(x) + \frac{1}{4} p(x)^T B_j(z) p(x) \\ &= \nabla F_j(z)^T p(x) + \frac{1}{4} p(x)^T \nabla^2 F_j(z) p(x) + \frac{1}{4} p(x)^T (B_j(z) - \nabla^2 F_j(z)) p(x). \end{aligned} \quad (17.3)$$

چون دو عبارت اول رابطه (۱۷.۳) پیوسته هستند، فقط پیوستگی عبارت آخر را بررسی می‌کنیم. دو نقطه $w_1, w_2 \in W$ را به گونه‌ای انتخاب می‌کنیم که به ازای یک $\delta \in \mathbb{R}_{++}$ کوچک، $\|w_1 - w_2\| < \delta$ باشد، آن‌گاه اندیس‌های $i_1, i_2 \in \{1, \dots, l\}$ وجود دارند به طوری که $w_i \in B(z_{i_k}, \delta_{z_{i_k}})$ ، $k = 1, 2$ بنابراین داریم:

$$\begin{aligned} & \left| \frac{1}{\gamma} p(x)^T (B_j(w_1) - \nabla^2 F_j(w_1)) p(x) - \frac{1}{\gamma} p(x)^T (B_j(w_2) - \nabla^2 F_j(w_2)) p(x) \right| \\ & \leq \frac{1}{\gamma} \|p(x)\|^2 (\|B_j(w_1) - \nabla^2 F_j(z_{i_1})\| + \|\nabla^2 F_j(z_{i_1}) - \nabla^2 F_j(w_1)\| \\ & \quad + \|B_j(w_2) - \nabla^2 F_j(z_{i_2})\| + \|\nabla^2 F_j(z_{i_2}) - \nabla^2 F_j(w_2)\|) \leq \frac{1}{\gamma} \|p(x)\|^2 (\varepsilon_{z_{i_1}} + \varepsilon_{z_{i_2}}) \end{aligned}$$

به ازای هر $x \in W$ و هر $j = 1, 2, \dots, m$ پیوسته یکنواخت است. لذا واضح است که خانواده $\{\varphi_{x,j}\}_{x \in W, j=1,2,\dots,m}$ پیوسته یکسان است. بنابراین خانواده $\{\phi_x = \max_{j=1,2,\dots,m} \varphi_{x,j}\}_{x \in W}$ نیز پیوسته یکسان است. $\varepsilon \in \mathbb{R}_{++}$ را در نظر بگیرید، $\delta \in \mathbb{R}_{++}$ وجود دارد به طوری که $\forall y, z \in W, \|y - z\| < \delta \implies |\phi_x(y) - \phi_x(z)| < \varepsilon, \forall x \in W$.

بنابراین برای $\|y - z\| < \delta$ داریم:

$$\begin{aligned} \tau(z) & \leq \max_{j=1,2,\dots,m} \nabla F_j(z)^T p(y) + \frac{1}{\gamma} p(y)^T B_j(z) p(y) \\ & = \phi_y(z) \\ & \leq \phi_y(y) + |\phi_y(z) - \phi_y(y)| < \tau(y) + \varepsilon. \end{aligned}$$

بنابراین $\tau(z) - \tau(y) < \varepsilon$. اگر نقش‌های y و z را با هم عوض کنیم، داریم $|\tau(z) - \tau(y)| < \varepsilon$. لذا پیوستگی τ اثبات می‌شود. $\square$

به منظور تعیین طول گام α برای جهت $p(x)$ ، از قاعده آرمیثو استفاده می‌کنیم. تابع $f : U \rightarrow \mathbb{R}$ را در نظر بگیرید، قاعده آرمیثو کلاسیک برای جهت شبه‌نیوتن $p(x)$ به صورت زیر است:

$$f(x + \alpha p(x)) \leq f(x) + c\alpha \nabla f(x)^T p(x),$$

که $x \in U$ ، نقطه غیربحرانی و $c \in (0, 1)$ است. برای پذیرش یک گام در روش شبه‌نیوتن چندهدفه با توجه به این‌که $B_j(x)$ مثبت معین است، داریم:

$$F_j(x + \alpha p(x)) \leq F_j(x) + c\alpha (\nabla F_j(x)^T p(x) + \frac{1}{\gamma} p(x)^T B_j(x) p(x)) \leq F_j(x) + c\alpha \tau(x).$$

نامساوی فوق معیاری برای پذیرش یک گام در جهت شبه‌نیوتن مسائل چندهدفه خواهد بود.

قضیه ۲.۳.۳. اگر $x \in U$ یک نقطه غیربحرانی برای تابع F باشد آن‌گاه برای هر $0 < c < 1$ و هر $j = 1, 2, \dots, m$ وجود دارد $\alpha_0 \in (0, 1]$ به طوری که

$$x + \alpha p(x) \in U, \quad F_j(x + \alpha p(x)) \leq F_j(x) + c\alpha \tau(x), \quad \forall \alpha \in [0, \alpha_0].$$

$\square$

برهان. مشابه با اثبات نتیجه ۲.۳.۲ است.

اکنون الگوریتم شبه‌نیوتن را برای مسائل چندهدفه بیان می‌کنیم. در هر گام مساله (۶.۳) را حل می‌کنیم و یک جهت شبه‌نیوتن را از رابطه (۸.۳) به‌دست می‌آوریم. طول گام را با شرط آرمیثو (قضیه ۲.۳.۳) تعیین می‌کنیم. در پایان نقطه را بروز می‌کنیم و تقریب ماتریس هسین را از رابطه (۴.۳) به‌دست می‌آوریم.

الگوریتم شبه‌نیوتن برای مسائل چندهدفه

الگوریتم ۳.۳.۳. مقداردهی اولیه $k = 0$.

i_1 . نقطه $x_0 \in U$ را انتخاب کنید.

i_2 . مجموعه $L = \{1/2^n \mid n = 0, 1, 2, \dots\}$ را تعریف کنید.

i_3 . ماتریس متقارن مثبت معین $B(x_0) \in \mathbb{R}^{n \times n}$ ، $0 < c < 1$ و یک مقدار کوچک ε را در نظر بگیرید.

گام‌های تکراری:

k_1 . مساله (۶.۳) را حل کنید و $p(x_k)$ و $\tau(x_k)$ را به‌ترتیب از روابط (۸.۳) و (۷.۳) به‌دست آورید.

k_2 . اگر $\tau(x_k) > -\varepsilon$ متوقف شوید. در غیر این صورت به گام k_3 بروید.

k_3 . طول گام α_k را بزرگترین $\alpha \in L$ انتخاب کنید به‌طوری‌که

$$x_k + \alpha p(x_k) \in U$$

و

$$F_j(x_k + \alpha p(x_k)) \leq F_j(x_k) + c\alpha\tau(x_k), \quad j = 1, 2, \dots, m.$$

k_4 . تعریف کنید $x_{k+1} = x_k + \alpha_k p(x_k)$ و $B_j(x_{k+1})$ را به ازای $j = 1, 2, \dots, m$ از رابطه (۴.۳) به‌دست آورید. $k = k + 1$ قرار دهید و به گام k_1 بروید.

۴.۳ نتایج تکنیکی

پیش از آن که به تحلیل همگرایی روش شبه‌نیوتن بپردازیم، لازم است نتیجه‌ای را یادآور شویم که تأیید می‌کند، همگرایی روش BFGS فوق‌خطی است. در [۱۴] نشان داده شده است، که شرط لازم و کافی برای همگرایی فوق‌خطی رابطه (۱۸.۳) است،

$$\lim_{k \rightarrow \infty} \frac{\|(B(x_k) - \nabla^2 f(x^*))s(x_k)\|}{\|s(x_k)\|} = 0, \quad (18.3)$$

که $f: \mathbb{R}^n \rightarrow \mathbb{R}$ دوبار به‌طور پیوسته مشتق‌پذیر است، همچنین $\nabla f(x^*) = 0$ و $\nabla^2 f(x^*)$ مثبت معین است و نیز $x^* = \lim_{k \rightarrow \infty} x_k$ ، $s(x_k) = x_{k+1} - x_k = \alpha_k p(x_k)$ و $B(x_k)$ یک تقریب در

تکرار k -ام است. متأسفانه این نتیجه در این قسمت نمی‌تواند استفاده شود، چون x^* وجود ندارد که $\nabla F_j(x^*) = 0$ به ازای $j = 1, 2, \dots, m$. نتایج عددی نشان می‌دهد، وقتی که $k \rightarrow \infty$ داریم:

$$\frac{\|(B_j(x_k) - \nabla^2 F_j(x_k))p(x_k)\|}{\|p(x_k)\|} \rightarrow 0, \quad j = 1, 2, \dots, m.$$

بنابراین فرض می‌کنیم که برای $\varepsilon \in \mathbb{R}_{++}$ کوچک، $k_0 \in \mathbb{N}$ وجود دارد به طوری که به ازای هر $k \geq k_0$ داشته باشیم:

$$\frac{\|(B_j(x_k) - \nabla^2 F_j(x_k))p(x_k)\|}{\|p(x_k)\|} < \varepsilon, \quad j = 1, 2, \dots, m.$$

لم زیر همان لم ۱۰.۴.۲ هست که در آن خطای تقریب ∇F_j با مدل خطی آن و خطای تقریب F_j با مدل درجه دوم آن آورده شده است. با این تفاوت که برای ساده‌سازی نتایج بعدی $\frac{\varepsilon}{4}$ به جای ε قرار گرفته است.

لم ۱۰.۴.۳. فرض کنید $V \subset U$ یک مجموعه محدب باشد. $\delta, \varepsilon \in \mathbb{R}_{++}$ را در نظر بگیرید، به طوری که برای هر $x, y \in V$ که $\|y - x\| < \delta$ داشته باشیم:

$$\|\nabla^2 F_j(y) - \nabla^2 F_j(x)\| < \frac{\varepsilon}{4}, \quad j = 1, 2, \dots, m. \quad (۱۹.۳)$$

تحت این فرض، برای هر $x, y \in V$ با $\|y - x\| < \delta$ و هر $j = 1, 2, \dots, m$ داریم:

$$\|\nabla F_j(y) - [\nabla F_j(x) + \nabla^2 F_j(x)(y - x)]\| < \frac{\varepsilon}{4} \|y - x\| \quad (۲۰.۳)$$

و

$$|F_j(y) - [F_j(x) + \nabla F_j(x)^T(y - x) + \frac{1}{4}(y - x)^T \nabla^2 F_j(x)(y - x)]| < \frac{\varepsilon}{4} \|y - x\|^2. \quad (۲۱.۳)$$

اکنون لم ۱۰.۴.۳ را اصلاح می‌کنیم و از تقریب BFGS به جای هسین استفاده می‌کنیم.

لم ۲۰.۴.۳. فرض کنید $V \subset U$ یک مجموعه محدب باشد و $\delta, \varepsilon \in \mathbb{R}_{++}$ ثابت‌های کوچک باشند به طوری که برای هر $x, y \in V$ با $\|y - x\| < \delta$ رابطه (۱۹.۳) برقرار باشد. همچنین $\{x_k\}$ یک دنباله تولید شده توسط الگوریتم ۳.۳.۳ باشد و $\{B_j(x_k)\}$ ، $j = 1, 2, \dots, m$ دنباله‌ای از تقریب‌های BFGS باشند. فرض کنید که به ازای $\varepsilon > 0$ ، وجود داشته باشد $k_0 \in \mathbb{N}$ به طوری که برای هر $k \geq k_0$ ،

$$\frac{\|(\nabla^2 F_j(x_k) - B_j(x_k))(y - x_k)\|}{\|y - x_k\|} < \frac{\varepsilon}{4}, \quad j = 1, 2, \dots, m.$$

تحت این فرض، برای هر $\{x_k\}$ ، $k \geq k_0$ و هر $y \in V$ که $\|y - x_k\| < \delta$ داریم:

$$\|\nabla F_j(y) - (\nabla F_j(x_k) + B_j(x_k)(y - x_k))\| < \varepsilon \|y - x_k\|, \quad j = 1, 2, \dots, m \quad (۲۲.۳)$$

و

$$|F_j(y) - (F_j(x_k) + \nabla F_j(x_k)^T(y - x_k) + \frac{1}{4}(y - x_k)^T B_j(x_k)(y - x_k))| < \frac{\varepsilon}{4} \|y - x_k\|^2, \quad (۲۳.۳)$$

که $j = 1, 2, \dots, m$ است.

برهان. بنا به رابطه (۲۰.۳) داریم:

$$\begin{aligned} \|\nabla F_j(y) - (\nabla F_j(x_k) + B_j(x_k)(y - x_k))\| &\leq \|\nabla F_j(y) - \nabla F_j(x_k) - \nabla^2 F_j(x_k)(y - x_k)\| \\ &\quad + \|(\nabla^2 F_j(x_k) - B_j(x_k))(y - x_k)\| \\ &< \frac{\varepsilon}{\varphi} \|y - x_k\| + \frac{\varepsilon}{\varphi} \|y - x_k\| = \varepsilon \|y - x_k\|, \end{aligned}$$

بنابراین رابطه (۲۲.۳) برقرار است.

به‌طور مشابه، بنا به رابطه (۲۱.۳) داریم:

$$\begin{aligned} &|F_j(y) - (F_j(x_k) + \nabla F_j(x_k)^T(y - x_k) + \frac{1}{\varphi}(y - x_k)^T B_j(x_k)(y - x_k))| \\ &\leq |F_j(y) - F_j(x_k) - \nabla F_j(x_k)^T(y - x_k) - \frac{1}{\varphi}(y - x_k)^T \nabla^2 F_j(x_k)(y - x_k)| \\ &\quad + |\frac{1}{\varphi}(y - x_k)^T (\nabla^2 F_j(x_k) - B_j(x_k))(y - x_k)| \\ &< \frac{\varepsilon}{\varphi} \|y - x_k\|^2 + \frac{\varepsilon}{\varphi} \|y - x_k\|^2 = \frac{\varepsilon}{\varphi} \|y - x_k\|^2, \end{aligned}$$

□

که رابطه (۲۳.۳) را تایید می‌کند، بدین ترتیب اثبات کامل شد.

لم بعد، تخمینی برای τ بیان می‌کند.

لم ۳.۴.۳. فرض کنید $x \in U$ و $a, b \in \mathbb{R}_+$ باشند، به‌طوری‌که $a \leq b$. اگر

$$aI \leq B_j(x) \leq bI, \quad j = 1, \dots, m,$$

آنگاه دو نامساوی زیر برقراراند:

$$\frac{a}{\varphi} \|p(x)\|^2 \leq |\tau(x)| \leq \frac{b}{\varphi} \|p(x)\|^2 \quad (24.3)$$

و به ازای هر $j = 1, \dots, m$ رابطه زیر برقرار است:

$$|\tau(x)| \leq \frac{1}{\varphi a} \left\| \sum_{j=1}^m \lambda_j \nabla F_j(x) \right\|^2, \quad \forall \lambda_j \geq 0, \quad \sum_{j=1}^m \lambda_j = 1. \quad (25.3)$$

برهان. می‌توانیم با جایگزین کردن ماتریس‌های هسین $\nabla^2 F_j(x)$ با $B_j(x)$ به ازای $j = 1, 2, \dots, m$

□

در اثبات‌های لم ۲.۴.۲ و لم ۳.۴.۲ به حکم فوق برسیم.

۵.۳ همگرایی

در قضیه زیر شرایط کافی برای همگرایی موضعی را ارائه می‌دهیم.

قضیه ۱.۵.۳. فرض کنید، $\{x_k\}$ یک دنباله تولید شده توسط الگوریتم ۳.۳.۳ باشد. همچنین $V \subset U$

یک مجموعه محدب باشد، $(0, 1)$ ، $c \in (0, 1)$ ، $k_* \in \mathbb{N}$ ، $a, b, r, \delta, \varepsilon \in \mathbb{R}_{++}$ و

$$aI \leq B_j(x) \leq bI \text{ و } aI \leq \nabla^2 F_j(x) \leq bI, \quad \forall j = 1, 2, \dots, m, \quad \forall x \in V. \quad ۱.$$

۲. $\|\nabla^j F_j(x) - \nabla^j F_j(y)\| \leq \varepsilon/2, \forall j = 1, 2, \dots, m$ و $\|x - y\| < \delta$ با $\forall x, y \in V$.

۳. $\|(B_j(x_k) - \nabla^j F_j(x_k))(y - x_k)\| < \|y - x_k\| \varepsilon/2, \forall j = 1, 2, \dots, m, y \in V, \forall k \geq k_0$.

۴. $\varepsilon/a \leq 1 - c$.

۵. $B(x_{k_0}, r) \subset V$.

۶. $\|p(x_{k_0})\| \leq \min\{\delta, r(1 - \varepsilon/a)\}$.

آنگاه برای هر $k \geq k_0$ داریم:

الف. $\|x_k - x_{k_0}\| \leq \|p(x_{k_0})\| \frac{1 - (\varepsilon/a)^{k-k_0}}{(1 - \varepsilon/a)}$.

ب. $\|p(x_k)\| \leq \|p(x_{k_0})\| (\varepsilon/a)^{k-k_0}$.

ج. $\alpha_k = 1$.

د. $\|p(x_{k+1})\| \leq \|p(x_k)\| (\varepsilon/a)$.

دنباله $\{x_k\}$ همگرا به یک جواب کارای موضعی $x^* \in \mathbb{R}^n$ است و سرعت همگرایی آن فوق خطی است. برهان. فرض کنید که موارد الف و ب برای یک k برقرار باشند. اثبات می‌کنیم که موارد ج و د نیز برای آن k برقراراند. بنا به موارد الف و ب و نامساوی مثلث داریم:

$$\begin{aligned} \|(x_k + p(x_k)) - x_{k_0}\| &\leq \|x_k - x_{k_0}\| + \|p(x_k)\| \\ &\leq \|p(x_{k_0})\| \frac{1 - (\varepsilon/a)^{k-k_0}}{1 - \varepsilon/a} + \|p(x_{k_0})\| (\varepsilon/a)^{k-k_0} \\ &\leq \|p(x_{k_0})\| \frac{1 - (\varepsilon/a)^{k-k_0+1}}{1 - \varepsilon/a} \end{aligned} \quad (26.3)$$

$$\leq \|p(x_{k_0})\| \frac{1 - (\varepsilon/a)^{k+1}}{1 - \varepsilon/a}. \quad (27.3)$$

لذا، از فرض‌های ۴ و ۶ داریم:

$$\|(x_k + p(x_k)) - x_{k_0}\| \leq \|p(x_{k_0})\| \frac{1 - (\varepsilon/a)^{k+1}}{1 - \varepsilon/a} \leq r(1 - \varepsilon/a) \frac{1 - (\varepsilon/a)^{k+1}}{1 - \varepsilon/a} < r. \quad (28.3)$$

بنا به مورد الف و فرض‌های ۴ و ۶ داریم:

$$\|x_k - x_{k_0}\| \leq \|p(x_{k_0})\| \frac{1 - (\varepsilon/a)^{k-k_0}}{1 - \varepsilon/a} \leq r(1 - \varepsilon/a) \frac{1 - (\varepsilon/a)^{k-k_0}}{1 - \varepsilon/a} < r. \quad (29.3)$$

از مورد ب و فرض ۶ داریم:

$$\|(x_k + p(x_k)) - x_k\| = \|p(x_k)\| \leq \|p(x_{k_0})\| (\varepsilon/a)^{k-k_0} \leq \delta (\varepsilon/a)^{k-k_0} < \delta. \quad (30.3)$$

روابط (۲۸.۳)، (۲۹.۳) و (۳۰.۳) نتیجه می‌دهند:

$$x_k, x_k + p(x_k) \in B(x_{k_0}, r), \quad \|(x_k + p(x_k)) - x_k\| \leq \delta.$$

اکنون بنا به رابطه (۲۳.۳)، لم ۲.۴.۳ و فرض‌های ۲ و ۳ برای هر $j = 1, 2, \dots, m$ داریم:

$$\begin{aligned} F_j(x_k + p(x_k)) &\leq F_j(x_k) + \nabla F_j(x_k)^T p(x_k) + \frac{1}{4} p(x_k)^T B_j(x_k) p(x_k) + \frac{\varepsilon}{4} \|p(x_k)\|^2 \\ &\leq F_j(x_k) + \tau(x_k) + \frac{\varepsilon}{4} \|p(x_k)\|^2 \\ &= F_j(x_k) + c\tau(x_k) + (1-c)\tau(x_k) + \frac{\varepsilon}{4} \|p(x_k)\|^2. \end{aligned} \quad (31.3)$$

نامساوی $\tau(x_k) \leq 0$ (مورد ۱، لم ۱.۳.۳) همراه با رابطه (۲۴.۳)، لم ۳.۴.۳ و فرض‌های ۱ و ۴ نتیجه می‌دهند:

$$(1-c)\tau(x_k) + \frac{\varepsilon}{4} \|p(x_k)\|^2 \leq \frac{-a(1-c) + \varepsilon}{4} \|p(x_k)\|^2 \leq 0. \quad (32.3)$$

با ترکیب نامساوی‌های (۳۱.۳) و (۳۲.۳) برای هر $j = 1, 2, \dots, m$ داریم:

$$F_j(x_k + p(x_k)) \leq F_j(x_k) + c\tau(x_k).$$

بنابراین شرط آرمیثو برای $\alpha_k = 1$ برقرار است، لذا شرط ج را اثبات کردیم. قرار می‌دهیم:

$$x_{k+1} := x_k + p(x_k), \quad x_k, x_{k+1} \in B(x_{k_0}, r), \quad \|x_{k+1} - x_k\| \leq \delta.$$

همچنین تعریف می‌کنیم:

$$v_{k+1} := \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_{k+1}).$$

بنا به رابطه (۲۵.۳)، لم ۳.۴.۳ و روابط (۱۱.۳) و (۱۲.۳) داریم:

$$|\tau(x_{k+1})| \leq \frac{1}{4a} \|v_{k+1}\|^2.$$

اکنون می‌خواهیم $\|v_{k+1}\|^2$ را تخمین بزنیم. برای $x \in U$ تعریف می‌کنیم:

$$G_k(x) := \sum_{j=1}^m \lambda_j(x_k) F_j(x)$$

و

$$H_k := \sum_{j=1}^m \lambda_j(x_k) B_j(x_k).$$

که $\lambda_j(x_k)$ ضرایب لاگرانژین مساله (۹.۳) هستند. لذا داریم:

$$\nabla G_k(x) = \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x),$$

$$\nabla^2 G_k(x) = \sum_{j=1}^m \lambda_j(x_k) \nabla^2 F_j(x).$$

بنا به $v_{k+1} = \nabla G_k(x_{k+1})$ و رابطه (۱۵.۳) داریم:

$$p(x_k) = -(H_k)^{-1} \nabla G_k(x_k). \quad (33.3)$$

بنا به فرض‌های ۲ و ۳، برای هر $x, y \in V$ با $\|y - x\| \leq \delta$ و $k \geq k_0$ داریم:

$$\|\nabla^2 G_k(y) - \nabla^2 G_k(x)\| < \frac{\varepsilon}{4}$$

و

$$\|(H_k - \nabla^2 G_k(x_k))(y - x_k)\| < \frac{\varepsilon}{4} \|y - x_k\|.$$

لذا با استفاده از رابطه (۲۲.۳)، لم ۲.۴.۳ داریم:

$$\|\nabla G_k(x_k + p(x_k)) - (\nabla G_k(x_k) + H_k p(x_k))\| < \varepsilon \|p(x_k)\|.$$

باید توجه داشت که رابطه (۳۳.۳) معادل با $\nabla G_k(x_k) + H_k p(x_k) = 0$ است، که نتیجه می‌دهد:

$$\|v_{k+1}\| = \|\nabla G_k(x_{k+1})\| < \varepsilon \|p(x_k)\|. \quad (۳۴.۳)$$

بنا به رابطه (۲۵.۳)، لم ۳.۴.۳ و رابطه (۳۴.۳) داریم:

$$|\tau(x_{k+1})| \leq \frac{\|v_{k+1}\|^2}{2a} < \frac{(\varepsilon \|p(x_k)\|)^2}{2a}.$$

لذا از رابطه (۲۴.۳)، لم ۳.۴.۳ و فرض ۱ نتیجه می‌شود:

$$\frac{a}{2} \|p(x_{k+1})\|^2 < \frac{(\varepsilon \|p(x_k)\|)^2}{2a}.$$

بنابراین، $\|p(x_{k+1})\| < \frac{\varepsilon \|p(x_k)\|}{a}$ و مورد د برقرار است.

اکنون با استقرا ثابت می‌کنیم که موارد الف و ب برای هر k برقراراند، بدیهی است که برای $k = k_0$ برقراراند. نشان دادیم که اگر موارد الف و ب برای یک k برقرار باشند، آنگاه موارد ج و د نیز برای همان k برقرارند. مورد ج نتیجه می‌دهد که $x_{k+1} = x_k + p(x_k)$ و بنا به رابطه (۲۶.۳)، مورد الف برای $k+1$ برقرار است. مورد ب برای $k+1$ مستقیماً از مورد د و فرض استقرا نتیجه می‌شود:

$$\|p(x_{k+1})\| \leq \|p(x_k)\|(\varepsilon/a) \leq \|p(x_{k_0})\|(\varepsilon/a)^{k-k_0+1}.$$

لذا موارد ج و د نیز برای هر k برقراراند. بنابراین، برای هر k و l که $k > l \geq k_0$ از نامساوی مثلث داریم:

$$\|x_k - x_l\| \leq \sum_{i=l}^{k-1} \|x_{i+1} - x_i\| \leq \sum_{i=l}^{k-1} \|p(x_i)\| \leq \|p(x_{k_0})\| \sum_{i=l}^{k-1} (\varepsilon/a)^{i-k_0},$$

که نامساوی آخر از مورد ب نتیجه می‌شود. بنابراین $\{x_k\}$ یک دنباله کوشی است و چون هر دنباله کوشی در $\mathbb{R}^n$ همگراست، $x^* \in \mathbb{R}^n$ وجود دارد به طوری که $\lim_{k \rightarrow \infty} x_k = x^*$. چون بنا به مورد ب و فرض ۴، $\{\|p(x_k)\|\}$ همگرا به صفر است، از فرض ۱ و رابطه (۲۴.۳)، لم ۳.۴.۳ نتیجه می‌گیریم که $\tau(x_k)$ همگرا به صفر است، وقتی که $k \rightarrow \infty$. بنا به مورد ۳ لم ۱.۳.۳، چون τ یک تابع پیوسته است، حد و تابع را می‌توانیم، جابه‌جا کنیم. بنابراین $\tau(x^*) = 0$ است. لذا بنا به مورد ۲ لم ۱.۳.۳، x^* نقطه بحرانی برای F است. بنابراین از مورد ج قضیه ۸.۳.۱، x^* برای F محدود شده روی V ، کارا است. بنابراین کارای موضعی است.

سرانجام همگرایی فوق خطی $\{x_k\}$ را اثبات می‌کنیم. ابتدا تعریف می‌کنیم:

$$r_k = \|p(x_{k_0})\| \frac{(\varepsilon/a)^{k-k_0}}{1 - \varepsilon/a}$$

و

$$\delta_k = \|p(x_{k_0})\|(\varepsilon/a)^{k-k_0}.$$

لذا از نامساوی مثلث و فرض‌های ۵ و ۶ و مورد الف نتیجه می‌گیریم:

$$B(x_k, r_k) \subset B(x_{k_*}, r) \subset V. \quad (۳۵.۳)$$

چون اگر $y \in B(x_k, r_k)$ باشد، داریم:

$$\|y - x_k\| \leq r_k = \|p(x_{k_*})\| \frac{(\varepsilon/a)^{k-k_*}}{1 - \varepsilon/a},$$

$$\|y - x_{k_*}\| \leq \|y - x_k\| + \|x_k - x_{k_*}\| \leq \|p(x_{k_*})\| \frac{(\varepsilon/a)^{k-k_*} + 1 - (\varepsilon/a)^{k-k_*}}{1 - \varepsilon/a} \leq r.$$

نقطه $\varphi \in \mathbb{R}_{++}$ را به دلخواه انتخاب می‌کنیم، تعریف می‌کنیم:

$$\tilde{\varepsilon} := \min \left\{ a \frac{\varphi}{1 + 2\varphi}, \varepsilon \right\}.$$

برای $k \geq k_*$ و هر $j = 1, 2, \dots, m$ نامساوی زیر را داریم:

$$\|\nabla^j F_j(y) - \nabla^j F_j(x)\| \leq \frac{\tilde{\varepsilon}}{2}, \quad \forall x, y \in B(x_k, r_k), \|x - y\| \leq \delta_k. \quad (۳۶.۳)$$

برای $l \geq k$ و هر $j = 1, 2, \dots, m$ داریم:

$$\|(B_j(x_l) - \nabla^j F_j(x_l))(y - x_l)\| < \frac{\tilde{\varepsilon}}{2} \|(y - x_l)\|, \quad \forall y \in B(x_k, r_k). \quad (۳۷.۳)$$

فرض‌های ۱-۶ به ازای $\tilde{\varepsilon}$ ، r_k ، δ_k و x_k به ترتیب به جای ε ، r ، δ و x_{k_*} برقرارند. چون رابطه‌های ۱ و ۵ از رابطه (۳۵.۳) نتیجه می‌شوند و موارد ۲ و ۳ به ترتیب از نامساوی‌های (۳۶.۳) و (۳۷.۳) به دست می‌آیند. مورد ۴ نیز به ازای $\tilde{\varepsilon}$ برقرار است، چون این مورد برای ε ، a و c صحیح است. فرض ۶ از تعریف‌های δ_k ، r_k و $\tilde{\varepsilon}$ نتیجه می‌شود.

در مورد الف از اندیس l به جای k استفاده می‌کنیم، بنابراین داریم:

$$\|x_l - x_k\| \leq \|p(x_k)\| \frac{1 - (\tilde{\varepsilon}/a)^{l-k}}{1 - \tilde{\varepsilon}/a}.$$

فرض کنید $l \rightarrow \infty$ ، لذا داریم:

$$\|x^* - x_k\| \leq \|p(x_k)\| \frac{1}{1 - \tilde{\varepsilon}/a}. \quad (۳۸.۳)$$

با استفاده از نامساوی (۳۸.۳) و مورد د داریم:

$$\|x^* - x_{k+1}\| \leq \|p(x_{k+1})\| \frac{1}{1 - \tilde{\varepsilon}/a} \leq \|p(x_k)\| \frac{\tilde{\varepsilon}/a}{1 - \tilde{\varepsilon}/a}. \quad (۳۹.۳)$$

بنا به نامساوی مثلث و رابطه (۳۹.۳) داریم:

$$\begin{aligned} \|x^* - x_k\| &\geq \|x_{k+1} - x_k\| - \|x^* - x_{k+1}\| \\ &\geq \|p(x_k)\| - \|p(x_k)\| \frac{\tilde{\varepsilon}/a}{1 - \tilde{\varepsilon}/a} \\ &= \|p(x_k)\| \frac{1 - 2\tilde{\varepsilon}/a}{1 - \tilde{\varepsilon}/a}. \end{aligned} \quad (۴۰.۳)$$

بنا به تعریف $\tilde{\varepsilon}$ ، $1 - 2\tilde{\varepsilon}/a > 0$ و $1 - \tilde{\varepsilon}/a > 0$ هستند.
روابط (۳۹.۳) و (۴۰.۳) نتیجه می‌دهند:

$$\|x^* - x_{k+1}\| \leq \|x^* - x_k\| \frac{\tilde{\varepsilon}/a}{1 - 2\tilde{\varepsilon}/a}.$$

از ترکیب این نامساوی با تعریف $\tilde{\varepsilon}$ داریم:

$$\|x^* - x_{k+1}\| \leq \varphi \|x^* - x_k\|.$$

چون $\varphi \in \mathbb{R}_{++}$ ، دلخواه انتخاب شده است، دنباله $\{x_k\}$ همگرایی فوق‌خطی به x^* است. $\square$

۶.۳ مثال‌های عددی

در این قسمت با استفاده از الگوریتم ۳.۳.۳ تعدادی مساله محدب قوی را حل می‌کنیم. برای حل مسائل این قسمت $B_j(x^{(o)}) = I$ ، $j = 1, 2, \dots, m$ و شرط توقف را $\tau(x^{(k)}) \geq -\varepsilon$ در نظر می‌گیریم که $\varepsilon > 0$ داده شده و $x^{(k)}$ نقطه پایانی است. مثال‌ها با استفاده از متلب (R2013a) حل می‌شوند.

مثال ۱.۶.۳. مساله دوهدفه زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}^3} F(x) = (F_1(x), F_2(x))$$

$$F_1(x) = x_1^2 + x_2^2 + x_3^2$$

$$F_2(x) = \frac{1}{3}x_1^2 + \frac{1}{3}x_2^2 + \frac{1}{3}x_3^2 - \frac{4}{3}x_1 - \frac{4}{3}x_2 - \frac{4}{3}x_3 + 4.$$

$$\nabla F_1(x) = (2x_1, 2x_2, 2x_3), \quad \nabla F_2(x) = \left(\frac{2}{3}x_1 - \frac{4}{3}, \frac{2}{3}x_2 - \frac{4}{3}, \frac{2}{3}x_3 - \frac{4}{3}\right)$$

و

$$\nabla^2 F_1(x) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \nabla^2 F_2(x) = \begin{bmatrix} \frac{2}{3} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} \end{bmatrix}.$$

باید توجه داشت، این توابع، محدب قوی هستند. این مساله را با روش شبه‌نیوتن، الگوریتم ۳.۳.۳ حل می‌کنیم. جواب این مساله با نقطه شروع $x^{(o)} = (1, 0, 1)$ و شرط توقف $\tau(x^{(k)}) \geq -10^{-8}$ به‌دست می‌آوریم. با حل مساله (۶.۳)، $\tau(x^{(o)}) = -0.4444$ ، به‌دست آمد که در شرط توقف صدق نمی‌کند. لذا الگوریتم متوقف نمی‌شود. به گام k_3 الگوریتم می‌رویم، نقطه $x^{(1)} = (0.8889, 0.4444, 0.8889)$ حاصل شد در این نقطه ماتریس B_1 را با فرمول BFGS به‌صورت زیر محاسبه می‌کنیم:

$$y_1(x^{(o)}) = \nabla F_1(x^{(1)}) - \nabla F_1(x^{(o)}) = (-0.2222, 0.8889, -0.2222),$$

$$s_o = x^{(1)} - x^{(o)}, \quad B_1(x^{(o)}) = I,$$

$$B_1(x^{(1)}) = B_1(x^{(0)}) - \frac{B_1(x^{(0)})s_0 s_0^T B_1(x^{(0)})}{s_0^T B_1(x^{(0)})s_0} + \frac{y_1(x^{(0)})y_1(x^{(0)})^T}{s_0^T y_1(x^{(0)})}$$

$$= \begin{bmatrix} 1/0.556 & -0.2222 & 0/0.556 \\ -0.2222 & 1/8889 & -0.2222 \\ 0/0.556 & -0.2222 & 1/0.556 \end{bmatrix}.$$

ماتریس B_2 نیز به همین ترتیب محاسبه می‌شود،

$$y_2(x^{(0)}) = \nabla F_2(x^{(1)}) - \nabla F_2(x^{(0)}) = (-0.0741, 0.2963, -0.0741),$$

$$s_0 = x^{(1)} - x^{(0)}, \quad B_2(x^{(0)}) = I,$$

$$B_2(x^{(1)}) = B_2(x^{(0)}) - \frac{B_2(x^{(0)})s_0 s_0^T B_2(x^{(0)})}{s_0^T B_2(x^{(0)})s_0} + \frac{y_2(x^{(0)})y_2(x^{(0)})^T}{s_0^T y_2(x^{(0)})}$$

$$= \begin{bmatrix} 0.9815 & 0.0741 & -0.0185 \\ 0.0741 & 0.7037 & 0.0741 \\ -0.0185 & 0.0741 & 0.9815 \end{bmatrix}.$$

بدین ترتیب گام k الگوریتم با ماتریس‌های $B_1(x^{(1)})$ و $B_2(x^{(1)})$ دوباره اجرا می‌شود و مساله (۴۱.۳) را حل می‌کنیم.

$$\min \max_{j=1,2} \nabla F_j(x^{(1)})^T p + \frac{1}{p} p^T B_j(x^{(1)})p \quad (41.3)$$

$$s.t. \quad p \in \mathbb{R}^n,$$

با حل این مساله جهت $p(x^{(1)}) = (-0.1361, 0.3115, -0.1361)$ و $\tau(x^{(1)}) = -0.0758$ را به دست آوردیم که با توجه به شرط توقف -10^{-8} در $\tau(x^{(k)}) \geq -10^{-8}$ شرط توقف صدق نمی‌کند، لذا الگوریتم ادامه پیدا می‌کند و به گام k_3 الگوریتم می‌رویم، نقطه $x^{(2)}$ را محاسبه می‌کنیم در این نقطه ماتریس B_1 و B_2 را با فرمول $BFGS$ مانند قبل محاسبه می‌کنیم. این فرایند تکرار می‌شود تا زمانی که شرط توقف برقرار شود.

بعد از شش تکرار نقطه $x^{(6)} = (0.7539, 0.7539, 0.7539)$ با $\tau(x^{(6)}) = -10^{-9} \times 9.5781$ به دست آمد، که در شرط توقف صدق می‌کند. بنابراین الگوریتم در همین نقطه متوقف می‌شود. در واقع $\tau(x^{(6)})$ به اندازه کافی نزدیک به صفر است. لذا بنا به لم ۱۰.۳.۳ نقطه $x^{(6)}$ بحرانی است. با توجه به این که مساله محدب قوی است، بنا به قضیه ۸.۳.۱ این نقطه بحرانی، نقطه کارا است.

این مثال را با نقطه شروع $x^{(0)} = (1, 0, 1)$ نیز در فصل ۲ با الگوریتم ۳.۳.۲ حل کردیم که بعد از دو تکرار به نقطه $x^{(2)} = (0.7321, 0.7321, 0.7321)$ رسیدیم. باید توجه داشت برای مسائل با ابعاد بزرگ و مسائلی که محاسبه هسین آن‌ها مشکل باشد، مزیت الگوریتم شبه نیوتن ۳.۳.۳ بر الگوریتم نیوتن ۳.۳.۲ قابل مشاهده است. در این مثال که محاسبه هسین به آسانی صورت می‌گیرد و بعد مساله بزرگ نیست، الگوریتم نیوتن سریعتر از شبه نیوتن عمل می‌کند.

این مساله را با نقاط شروع تصادفی و شرط توقف بالا حل می‌کنیم. می‌توانید نتایج را در جدول ۱۰.۳ مشاهده کنید.

جدول ۱.۳: ۱.۶.۳ جدول مربوط به مثال

نقطه شروع	تعداد تکرارها	$x^{(k)}$ = نقطه پایانی	$(F_1(x^{(k)}), F_2(x^{(k)}))$
(۰/۹۵۷۵, ۰/۹۶۴۹, ۰/۱۵۷۶)	۴	(۰/۷۴۹۱, ۰/۷۴۹۱, ۰/۷۴۹۱)	(۱/۶۸۳۵, ۱/۵۶۴۸)
(۰/۹۵۷۲, ۰/۴۸۵۴, ۰/۸۰۰۳)	۴	(۰/۷۶۲۱, ۰/۷۶۲۱, ۰/۷۶۲۱)	(۱/۷۴۲۵, ۱/۵۳۲۳)
(۰/۴۲۱۸, ۰/۹۱۵۷, ۰/۷۹۲۲)	۴	(۰/۷۲۷۱, ۰/۷۲۷۱, ۰/۷۲۷۱)	(۱/۵۸۶۰, ۱/۶۲۰۳)
(۰/۶۵۵۷, ۰/۰۳۵۷, ۰/۸۴۹۱)	۵	(۰/۵۶۹۰, ۰/۵۶۹۰, ۰/۵۶۹۰)	(۰/۹۷۱۲, ۲/۰۴۷۹)
(۰/۶۷۸۷, ۰/۷۵۷۷, ۰/۷۴۳۱)	۵	(۰/۷۲۷۰, ۰/۷۲۷۰, ۰/۷۲۷۰)	(۱/۵۸۵۵, ۱/۶۲۰۶)

مثال ۲.۶.۳. مساله سه‌هدفه زیر را در نظر بگیرید:

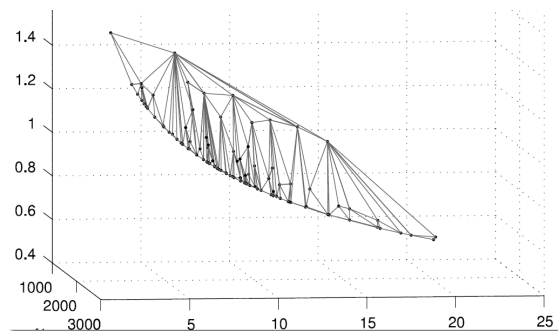
$$\min_{x \in \mathbb{R}^3} F(x) = (F_1(x), F_2(x), F_3(x))$$

$$F_1(x) = \frac{1}{n^2} \sum_{k=1}^n k(x_k - k)^2$$

$$F_2(x) = \exp\left(\sum_{k=1}^n x_k/n\right) + \|x\|^2$$

$$F_3(x) = \frac{1}{n(n+1)} \sum_{k=1}^n k(n-k+1)e^{-x_k}.$$

این توابع، محدب قوی هستند. این مساله با الگوریتم ۳.۳.۳ برای ابعاد بزرگ حل می‌شود. با توجه به این‌که محاسبه هسین این مثال مشکل است از روش شبه‌نیوتن استفاده می‌کنیم که نیاز به محاسبه دقیق ماتریس هسین ندارد و از تقریب هسین استفاده می‌کند. این مساله با روش شبه‌نیوتن، الگوریتم ۳.۳.۳ حل می‌شود. مجموعه نقاط کارای این مساله در شکل ۱.۳ با نقاط دایره‌ای شکل بر روی فضای هدف مشخص شده‌اند. این مثال با الگوریتم نیوتن ۳.۳.۲ نیز حل می‌شود و زمان حل دو الگوریتم را با هم مقایسه می‌کنیم.



شکل ۱.۳: نقاط کارای مثال ۲.۶.۳ بر روی فضای هدف مشخص شده‌اند که طول، عرض و ارتفاع نمودار به‌ترتیب محور مربوط به F_1 ، F_2 و F_3 هستند.

این مثال با استفاده از دو الگوریتم نیوتن ۳.۳.۳ و شبه‌نیوتن ۳.۳.۲ حل شده است و زمان محاسبه این دو الگوریتم در جدول ۲.۳ نشان داده شده است. با توجه به جدول ۲.۳ می‌بینیم که زمان محاسبه

الگوریتم شبه‌نیوتن کمتر از نیوتن است. در واقع همان طور که در مثال قبل اشاره کردیم، برای مسائل با ابعاد بزرگ مزیت الگوریتم شبه‌نیوتن به الگوریتم نیوتن مشخص می‌شود و با توجه به جدول ۲.۳ هر چه بعد مساله بزرگتر می‌شود الگوریتم ۳.۳.۳ سریعتر از الگوریتم ۳.۳.۲ عمل می‌کند.

جدول ۲.۳: جدول مربوط به زمان محاسبه دو الگوریتم

زمان محاسبه الگوریتم ۳.۳.۲	زمان محاسبه الگوریتم ۳.۳.۳	بعد مساله
۰.۰۰۲۱۹	۰.۰۰۱۲۷	۵۰
۰.۰۰۴۸۴	۰.۰۰۴۵۳	۱۰۰
۰.۰۲۲۵۲	۰.۰۱۰۲۸	۱۵۰
۰.۰۶۲۰۲	۰.۰۲۱۳۷	۲۰۰
۰.۳۰۴۳۹	۰.۰۸۸۱۶	۳۰۰
۰.۷۸۷۵۷	۰.۱۸۴۰۱	۴۰۰
۱.۵۷۷۲۷	۰.۳۴۱۰۲	۵۰۰

فصل ۴

روش شبه نیوتن اصلاح شده برای حل مسائل چند هدفه نامحدب

۱.۴ مقدمه

در این فصل الگوریتم روش شبه نیوتن اصلاح شده^۱ را بیان می‌کنیم که در آن تابع هدف الزاماً محدب نیست. همچنین الگوریتم ارائه شده به جای محاسبه هسین یا تقریب هسین هر تابع هدف، یک ماتریس مثبت معین تولید می‌کند.

۲.۴ مفاهیم اولیه

لم زیر را بیان می‌کنیم، تا با استفاده از آن لم بعد را اثبات کنیم.

لم ۱.۲.۴. [۲۰] برای یک ماتریس ناتهی^۲ A یکی از دو سیستم زیر جواب دارد:

الف. $Ax > 0$ ،

ب. $A^T y = 0$ ، $y \geq 0$.

لم ۲.۲.۴. فرض کنید $x^* \in \mathbb{R}^n$ ، به گونه‌ای باشد که برای حداقل یک $j \in \{1, 2, \dots, m\}$ ، $\nabla F_j(x^*) \neq 0$ باشد و یک ماتریس $c(x^*)$ از مرتبه $n \times m$ که ستون‌های آن $j = 1, 2, \dots, m$ ، $-\nabla F_j(x^*)$ هستند را در نظر بگیرید. اگر $y \in \mathbb{R}^m$ وجود داشته باشد که در سیستم $c(x^*)y = 0$ ، $y \geq 0$ صدق کند آن‌گاه x^* یک جواب کارای موضعی ضعیف از مساله (۵.۱) است.

^۱modified Quasi-Newton method

^۲non-vacuous

برهان. فرض کنید، $x^* \in \mathbb{R}^n$ و $y \geq 0$ وجود دارند به‌طوری‌که $c(x^*)y = 0$ است اما x^* یک جواب کارای موضعی ضعیف از مساله (۵.۱) نباشد. چون برای حداقل یک $j \in \{1, 2, \dots, m\}$ ، $\nabla F_j(x^*) \neq 0$ است، $c(x^*)$ ناتهی است. بنا به لم ۱.۲.۴ سیستم $d \in \mathbb{R}^n$ ، $c(x^*)^T d > 0$ جواب ندارد. لذا سیستم $\nabla F_j(x^*)^T d < 0$ برای هر $j = 1, 2, \dots, m$ در $\mathbb{R}^n$ جواب ندارد. به عبارت دیگر $d \in \mathbb{R}^n$ وجود ندارد، به‌طوری‌که برای هر نقطه $x^* + \lambda d$ با $\lambda > 0$ در همسایگی از x^* در نامعادله $F(x^* + \lambda d) < F(x^*)$ صدق کند. بنابراین، x^* یک جواب کارای موضعی ضعیف برای مساله (۵.۱) است. $\square$

لم ۳.۲.۴. اگر x^* یک نقطه بحرانی باشد و برای حداقل یک $j \in \{1, 2, \dots, m\}$ ، $\nabla F_j(x^*) \neq 0$ باشد و $j_0 \in \{1, 2, \dots, m\}$ وجود داشته باشد به‌طوری‌که $\nabla F_{j_0}(x^*) = 0$ باشد آنگاه x^* یک نقطه کارای موضعی ضعیف برای مساله (۵.۱) است.

برهان. نقطه x^* که در فرض‌های قضیه صدق می‌کند را در نظر بگیرید. $y \in \mathbb{R}^m$ را به گونه‌ای انتخاب کنید که در رابطه زیر صدق کند:

$$y_j = \begin{cases} 0, & j \neq j_0, \\ 1, & j = j_0. \end{cases}$$

لذا $y \geq 0$ در سیستم $c(x^*)^T y = 0$ صدق می‌کند که $c(x^*)$ یک ماتریس از مرتبه $n \times m$ است که ستون‌های آن $\nabla F_j(x^*)$ ، $j = 1, 2, \dots, m$ هستند. لذا بنا به لم ۲.۲.۴ نتیجه می‌گیریم که x^* یک جواب کارای موضعی ضعیف برای مساله (۵.۱) است. $\square$

مثال ۴.۲.۴. مساله زیر را در نظر بگیرید.

$$P : \min F(x) = (F_1(x), F_2(x), F_3(x))$$

$$F_1(x_1, x_2) = \frac{1}{4}[(x_1 - 1)^4 + 2(x_2 - 2)^4]$$

$$F_2(x_1, x_2) = e^{\left(\frac{x_1 + x_2}{2}\right)} + x_1^2 + x_2^2$$

$$F_3(x_1, x_2) = \frac{1}{e}(e^{-x_1} + 2e^{-x_2}).$$

با استفاده از لم‌های ۲.۲.۴ و ۳.۲.۴ کارای ضعیف بودن نقطه $x^* = (1, 2)^T$ را بررسی می‌کنیم.

$$\nabla F_1(x_1, x_2) = \begin{bmatrix} (x_1 - 1)^3 \\ 2(x_2 - 2)^3 \end{bmatrix}, \quad \nabla F_2(x_1, x_2) = \begin{bmatrix} \frac{1}{2}e^{\left(\frac{x_1 + x_2}{2}\right)} + 2x_1 \\ \frac{1}{2}e^{\left(\frac{x_1 + x_2}{2}\right)} + 2x_2 \end{bmatrix},$$

$$\nabla F_3(x_1, x_2) = \begin{bmatrix} -\frac{1}{e}e^{-x_1} \\ -\frac{1}{e}e^{-x_2} \end{bmatrix}.$$

بنابراین داریم:

$$c(x^*) = \begin{bmatrix} 0 & -4/2408 & 0/0613 \\ 0 & -6/2408 & 0/0451 \end{bmatrix}.$$

یک جواب سیستم $y \geq 0$ و $c(x^*)y = 0$ ، $y = (1, 0, 0)$ است. لذا بنا به لم ۲.۲.۴ نقطه x^* یک جواب کارای موضعی ضعیف مساله P است. اکنون لم ۳.۲.۴ را بررسی می‌کنیم،

$$DF(x^*) = \begin{bmatrix} 0 & 0 \\ \frac{1}{4}e^{\frac{1}{2}} + 2 & \frac{1}{4}e^{\frac{1}{2}} + 4 \\ -\frac{1}{6}e^{-1} & -\frac{1}{4}e^{-2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 4/2408 & 6/2408 \\ -0/0613 & -0/0451 \end{bmatrix}.$$

نقطه x^* بحرانی است، زیرا برای هر $d \in \mathbb{R}^n$ ، $\nabla F_1(x^*)^T d = 0$ ، لذا $R(DF(x^*)) \cap (-\mathbb{R}_{++}^m) = \emptyset$ است.

$$\nabla F_1(x^*) = (0, 0)^T, \quad \nabla F_2(x^*) = (\frac{1}{4}e^{\frac{1}{2}} + 2, \frac{1}{4}e^{\frac{1}{2}} + 4)^T \neq 0.$$

لذا شرایط لم ۳.۲.۴، برقرار است و نقطه x^* یک جواب کارای موضعی ضعیف برای مساله P است.

۳.۴ روش شبه‌نیوتن اصلاح شده

در این قسمت یک روش تقریب عددی را، برای حل مسائل بهینه‌سازی چندهدفه بیان می‌کنیم. برای رسیدن به این مقصود، در ادامه سه گام زیر را به‌ترتیب توضیح می‌دهیم:

الف. انتخاب یک جهت کاهنده مناسب،

ب. تولید یک دنباله از ماتریس‌های مثبت معین،

ج. تولید یک دنباله از نقاط همگرا به نقطه بحرانی یا نقطه کارای ضعیف.

در پایان، یک الگوریتم برای این فرآیند معرفی می‌کنیم.

۱.۳.۴ انتخاب یک جهت کاهنده مناسب

مساله بهینه‌سازی چندهدفه (۵.۱) را در نظر بگیرید. جهت نیوتن برای مساله (۵.۱)، جواب بهینه‌ای از مساله (۱.۲) است که تقریب درجه دوم، از هر تابع هدف F_j در $x \in \mathbb{R}^n$ را شامل می‌شود. در فصل ۲ بیان کردیم، که اگر هر F_j محدب اکید باشد آنگاه مساله (۱.۲) یک جواب منحصر بفرد دارد و یک مساله برنامه‌ریزی درجه دوم محدب است. همچنین با استفاده از شرایط بهینگی KKT، جواب مساله (۱.۲) به‌صورت زیر تعیین می‌شد:

$$s(x) = - \left[\sum_{j=1}^m \lambda_j(x) \nabla^2 F_j(x) \right]^{-1} \sum_{j=1}^m \lambda_j(x) \nabla F_j(x).$$

که $\lambda_j(x)$ ضرایب لاگرانژ بودند. در الگوریتم ۳.۳.۲، نیاز داشتیم که هسین هر تابع هدف را محاسبه کنیم و همچنین فرض کنیم که هر تابع هدف محدب است. هدف از این فصل بهبود جهت نیوتن $s(x)$ به یک جهت جدید $d_{QN}(x)$ است به‌طوری‌که

۱. به‌جای محاسبه هسین هر تابع هدف، یک ماتریس مثبت معین تولید می‌شود،

۲. لزومی ندارد، هر تابع هدف F_j محدب باشد،

۳. هر نقطه حدی از دنباله $\{x_{k+1}\}$ تولید شده با رابطه $x_{k+1} = x_k + \alpha_k d_{QN}(x_k)$ برای α_k مناسب، یک نقطه بحرانی یا نقطه کارای موضعی ضعیف از مساله (۵.۱) است.

برای تعیین $d_{QN}(x)$ ، یک دنباله از ماتریس‌های متقارن مثبت معین $\{B(x_k)\}$ تولید می‌کنیم که با یک ماتریس اولیه متقارن مثبت معین $\{B(x_0)\}$ در یک نقطه اولیه x_0 شروع می‌شود. برای یک ماتریس متقارن مثبت معین داده شده، $B(x)$ در x ، تعریف می‌کنیم:

$$Q_j(x) = F_j(x) - \frac{1}{p} d^T B(x) d, \quad j = 1, 2, \dots, m.$$

که به ازای هر j ، $Q_j(x)$ یک تقریب مناسبی از F_j در x است. این رابطه برای حل مساله بهینه‌سازی $\inf_{d \in \mathbb{R}^n} \max_{j=1,2,\dots,m} [F_j(x+d) - Q_j(x)]$ ضروری است که این مساله به‌صورت زیر تبدیل می‌شود:

$$\min_{d \in \mathbb{R}^n} \max_{j=1,2,\dots,m} [\nabla F_j(x)^T d + \frac{1}{p} d^T B(x) d].$$

یک جهت شبه‌نیوتن اصلاح شده برای مساله (۵.۱)، در نقطه x جواب بهینه‌ای از مساله زیر است:

$$\begin{aligned} \min \quad & t(x) \\ \text{s.t.} \quad & \nabla F_j(x)^T d + \frac{1}{p} d^T B(x) d - t(x) \leq 0 \quad j = 1, 2, \dots, m \\ & (t(x), d) \in \mathbb{R} \times \mathbb{R}^n. \end{aligned} \quad (1.4)$$

بنابراین داریم:

$$t(x) = \inf_{d \in \mathbb{R}^n} \max_{j=1,2,\dots,m} \nabla F_j(x)^T d + \frac{1}{p} d^T B(x) d. \quad (2.4)$$

چون $B(x)$ مثبت معین است، مساله (۱.۴) یک برنامه‌ریزی محدب است. بنابراین یک جواب منحصر بفرد دارد، که با استفاده از شرایط بهینگی KKT به‌دست می‌آید. لاگرانژین این مساله به‌صورت زیر است:

$$L((t, d), \lambda) = t + \sum_{j=1}^m \lambda_j (\nabla F_j(x)^T d + \frac{1}{p} d^T B(x) d - t),$$

که $\lambda_j \geq 0, j = 1, 2, \dots, m$ ضرایب لاگرانژین هستند. شرایط بهینگی KKT برای مساله (۱.۴) به صورت زیر است:

$$\sum_{j=1}^m \lambda_j = 1, \quad (3.4)$$

$$\sum_{j=1}^m \lambda_j (\nabla F_j(x) + B(x)d) = 0, \quad (4.4)$$

$$\nabla F_j(x)^T d + \frac{1}{\gamma} d^T B(x) d \leq t, \quad j = 1, 2, \dots, m, \quad (5.4)$$

$$\lambda_j \geq 0, \quad j = 1, 2, \dots, m, \quad (6.4)$$

$$\lambda_j (\nabla F_j(x)^T d + \frac{1}{\gamma} d^T B(x) d - t) = 0, \quad j = 1, 2, \dots, m. \quad (7.4)$$

فرض کنید $d_{QN}(x)$ در روابط (۳.۴)–(۷.۴) صدق کند و مقدار بهینه مساله (۱.۴)، $t(x)$ باشد. لذا از رابطه (۴.۴) داریم:

$$d_{QN}(x) = -B^{-1}(x) \sum_{j=1}^m \lambda_j(x) \nabla F_j(x). \quad (8.4)$$

قضیه ۱.۳.۴. اگر x نقطه غیر بحرانی باشد آن گاه $d_{QN}(x)$ تعریف شده در رابطه (۸.۴) یک جهت کاهنده در x است.

برهان. فرض کنید x یک نقطه غیر بحرانی از F باشد. بنابراین $\bar{d} \in \mathbb{R}^n$ وجود دارد، به طوری که برای هر $j = 1, 2, \dots, m$ ، $\nabla F_j(x)^T \bar{d} < 0$ ، در رابطه (۲.۴)، d را با $\beta \bar{d}$ برای هر $\beta > 0$ جایگزین می کنیم. لذا داریم:

$$t(x) \leq \max_{j=1,2,\dots,m} [\nabla F_j(x)^T \beta \bar{d} + \frac{1}{\gamma} (\beta \bar{d})^T B(x) (\beta \bar{d})] = \beta \max_{j=1,2,\dots,m} [\nabla F_j(x)^T \bar{d} + \frac{\beta}{\gamma} (\bar{d})^T B(x) \bar{d}].$$

بنابراین برای $\beta > 0$ ، به اندازه کافی کوچک، $t(x) < 0$ است. لذا داریم:

$$d_{QN}(x) = \arg \min_{d \in \mathbb{R}^n} \max_{j=1,2,\dots,m} [\nabla F_j(x)^T d + \frac{1}{\gamma} d^T B(x) d] \neq 0.$$

چون $B(x)$ ماتریس مثبت معین است و $d_{QN}(x) \neq 0$ است، داریم:

$$\nabla F_j(x)^T d_{QN}(x) < \nabla F_j(x)^T d_{QN}(x) + \frac{1}{\gamma} d_{QN}(x)^T B(x) d_{QN}(x) \leq t(x) < 0.$$

لذا برای هر $j = 1, 2, \dots, m$ ، $\nabla F_j(x)^T d_{QN}(x) < 0$ ، بنابراین $d_{QN}(x)$ یک جهت کاهنده در x است. $\square$

$t(x)$ و $d_{QN}(x)$ در ویژگی های لم ۱.۳.۲ صدق می کنند که این ویژگی ها را در قضیه ۲.۳.۴ بیان می کنیم.

قضیه ۲.۳.۴. اگر F روی U به طور پیوسته مشتق پذیر باشد آن گاه

۱. برای هر $x \in U$ ، $t(x) \leq 0$ است.

۲. شرایط زیر معادل‌اند:

الف. نقطه x ، غیربحرانی است،

ب. $t(x) < 0$ است،

ج. $d_{QN} \neq 0$ است.

۳. تابع $d_{QN} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ داده شده با رابطه (۸.۴) روی زیر مجموعه فشرده از $\mathbb{R}^n$ ، کراندار است. همچنین تابع $t : \mathbb{R}^n \rightarrow \mathbb{R}$ داده شده با رابطه (۲.۴) در هر زیر مجموعه فشرده از $\mathbb{R}^n$ ، پیوسته است.

گزاره ۳.۳.۴. بنا به موارد ۱ و ۲ قضیه ۲.۳.۴، x^* نقطه بحرانی F است اگر و فقط اگر $t(x^*) = 0$ یا $d_{QN}(x^*) = 0$ یا هر دو هم‌زمان برقرار باشند.

۲.۳.۴ تولید ماتریس مثبت معین $B(x_k)$

فرض کنید به ازای یک $\alpha_k > 0$ داشته باشیم $x_{k+1} = x_k + \alpha_k d_{QN}(x_k)$ روابط درجه دوم زیر را در نقطه x ، به ازای $j = 1, 2, \dots, m$ در نظر بگیرید:

$$F_j(x_k) + (x - x_k)^T \nabla F_j(x_k) + \frac{1}{2}(x - x_k)^T B(x_k)(x - x_k) \quad (9.4)$$

و

$$F_j(x_{k+1}) + (x - x_{k+1})^T \nabla F_j(x_{k+1}) + \frac{1}{2}(x - x_{k+1})^T B(x_{k+1})(x - x_{k+1}). \quad (10.4)$$

از رابطه (۳.۴) داریم $\sum_{j=1}^m \lambda_j(x_k) = 1$ روابط (۹.۴) و (۱۰.۴) را روی j جمع می‌کنیم و به ترتیب با $m^k(x)$ و $m^{k+1}(x)$ نشان می‌دهیم. لذا داریم:

$$m^k(x) = \sum_{j=1}^m \lambda_j(x_k) (F_j(x_k) + (x - x_k)^T \nabla F_j(x_k)) + \frac{1}{2}(x - x_k)^T B(x_k)(x - x_k)$$

و

$$m^{k+1}(x) = \sum_{j=1}^m \lambda_j(x_k) (F_j(x_{k+1}) + (x - x_{k+1})^T \nabla F_j(x_{k+1})) + \frac{1}{2}(x - x_{k+1})^T B(x_{k+1})(x - x_{k+1}).$$

ماتریس بروز شده $B(x_{k+1})$ تقریب مناسبی از $\sum_{j=1}^m \lambda_j(x_k) \nabla^2 F_j(x_{k+1})$ است، اگر شرط زیر برقرار باشد:

$$\nabla m^k(x_k) = \nabla m^{k+1}(x_k). \quad (11.4)$$

با ساده کردن رابطه (۱۱.۴) داریم:

$$\sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k) = \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_{k+1}) + B(x_{k+1})(x_k - x_{k+1}).$$

بنابراین داریم:

$$B(x_{k+1})(x_{k+1} - x_k) = \sum_{j=1}^m \lambda_j(x_k) (\nabla F_j(x_{k+1}) - \nabla F_j(x_k)). \quad (12.4)$$

روابط زیر را تعریف می‌کنیم:

$$\begin{aligned} s_k &= x_{k+1} - x_k, \\ y_j(x_k) &= \nabla F_j(x_{k+1}) - \nabla F_j(x_k), \quad j = 1, 2, \dots, m, \\ u_k &= \sum_{j=1}^m \lambda_j(x_k) y_j(x_k). \end{aligned}$$

بدین ترتیب رابطه (۱۲.۴) که یک شرط وتری اصلاح شده است، به صورت زیر تبدیل می‌شود:

$$B(x_{k+1})s_k = u_k.$$

لم ۴.۳.۴. اگر برای یک $c < 1$ و هر $j = 1, 2, \dots, m$ رابطه زیر برای نقطه غیربحرانی x_k برقرار باشد،

$$\nabla F_j(x_{k+1})^T d_{QN}(x_k) \geq c \nabla F_j(x_k)^T d_{QN}(x_k), \quad (13.4)$$

آنگاه $u_k^T s_k > 0$ است.

برهان. فرض کنید نقطه x_{k+1} به گونه‌ای انتخاب شود که رابطه (۱۳.۴)، برای یک $c < 1$ برقرار باشد. چون $d_{QN}(x_k)$ یک جهت کاهنده است، برای هر $j = 1, 2, \dots, m$ $\nabla F_j(x_k)^T d_{QN}(x_k) < 0$ است. بنابراین داریم:

$$u_k^T s_k = \sum_{j=1}^m \lambda_j(x_k) (\nabla F_j(x_{k+1}) - \nabla F_j(x_k))^T (x_{k+1} - x_k).$$

چون $x_{k+1} - x_k = \alpha_k d_{QN}(x_k)$ داریم:

$$u_k^T s_k = \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_{k+1})^T \alpha_k d_{QN}(x_k) - \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k)^T \alpha_k d_{QN}(x_k).$$

بنا به فرض داریم:

$$\begin{aligned} u_k^T s_k &\geq c \alpha_k \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k)^T d_{QN}(x_k) - \alpha_k \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k)^T d_{QN}(x_k) \\ &= \alpha_k \sum_{j=1}^m \lambda_j(x_k) (c - 1) \nabla F_j(x_k)^T d_{QN}(x_k) > 0. \end{aligned}$$

□

در این جا، مشابه روش شبه‌نیوتن برای مسائل بهینه‌سازی تک‌هدفه عمل می‌کنیم. چون $u_k^T s_k > 0$ برقرار است، بنابراین متناظر با فرمول DFP تک‌هدفه، $\{B(x_k)\}$ تولید شده برای مساله (۵.۱) به صورت زیر بیان می‌شود:

$$B(x_{k+1}) = (I - \gamma(x_k) u_k s_k^T) B(x_k) (I - \gamma(x_k) s_k u_k^T) + \gamma(x_k) u_k u_k^T, \quad (14.4)$$

که در آن

$$\gamma(x_k) = \frac{1}{u_k^T s_k} = \frac{1}{\sum_{j=1}^m \lambda_j(x_k) y_j(x_k)^T s_k}, \quad u_k^T s_k > 0.$$

همچنین با استفاده از فرمول بروز شده BFGS برای روش شبه‌نیوتن تک‌هدفه، فرمول بروز شده BFGS برای مساله (۵.۱) به‌صورت زیر است:

$$B(x_{k+1}) = B(x_k) + \frac{u_k u_k^T}{u_k^T s_k} - \frac{B(x_k) s_k s_k^T B(x_k)}{s_k^T B(x_k) s_k}. \quad (15.4)$$

گزاره ۵.۳.۴. باید توجه داشت که در روش شبه‌نیوتن تک‌هدفه، شرط انحنا به‌صورت $s_k^T y(x_k) > 0$ است، که در آن $y(x_k) = \nabla f(x_{k+1}) - \nabla f(x_k)$. در مسائل بهینه‌سازی چندهدفه، $y(x_k)$ با u_k جایگزین می‌شود، که $u_k = \sum_{j=1}^m \lambda_j(x_k) (\nabla F_j(x_{k+1}) - \nabla F_j(x_k))$. بنابراین $\{B(x_k)\}$ تولید شده در بالا یک دنباله از ماتریس‌های مثبت معین است، مشروط به این‌که $B(x_0)$ یک ماتریس مثبت معین باشد. اثبات مثبت معین بودن $\{B(x_k)\}$ مشابه اثبات مثبت معین بودن ماتریس تولید شده در روش شبه‌نیوتن برای مسائل بهینه‌سازی تک‌هدفه است، زیرا در هر دو مورد شرط انحنا برقرار است. لذا از بیان اثبات آن صرف نظر می‌کنیم.

گزاره ۶.۳.۴. اگر هر تابع F_j محدب باشد آن‌گاه

$$\begin{aligned} u_k^T s_k &= \sum_{j=1}^m \lambda_j(x_k) (\nabla F_j(x_{k+1}) - \nabla F_j(x_k))^T s_k \\ &= \sum_{j=1}^m \lambda_j(x_k) s_k^T \frac{(\nabla F_j(x_{k+1}) - \nabla F_j(x_k))^T}{s_k^T} s_k \\ &\simeq \sum_{j=1}^m \lambda_j(x_k) s_k^T \nabla^2 F_j(x_k + \theta s_k) s_k > 0, \end{aligned}$$

که $0 < \theta < 1$. بدین ترتیب دنباله $\{B(x_k)\}$ داده شده در رابطه بالا تولید می‌شود. اما بنا به لم ۴.۳.۴ دنباله $\{B(x_k)\}$ می‌تواند، بدون وجود خاصیت تحدب F_j نیز تولید شود.

۳.۳.۴ تولید یک دنباله همگرا

قضیه ۷.۳.۴. فرض کنید برای هر F_j ، $j = 1, 2, \dots, m$ روی یک مجموعه فشرد $U \subset \mathbb{R}^n$ به‌طور پیوسته مشتق‌پذیر باشد و $\{x_k\}$ دنباله تولید شده با رابطه $x_{k+1} = x_k + \alpha_k d_{QN}(x_k)$ ، $\alpha_k > 0$ باشد که $d_{QN}(x_k)$ با رابطه (۸.۴) تولید می‌شود و α_k در رابطه زیر صدق می‌کند:

$$F_j(x_{k+1}) \leq F_j(x_k) + c_1 \alpha_k \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k)^T d_{QN}(x_k), \quad j = 1, 2, \dots, m, \quad (16.4)$$

که $0 < c_1 < 1$. به‌علاوه فرض کنید $L_0 = \{x \in U : F(x) < F(x_0)\}$ مجموعه‌ای کراندار و محدب باشد که در آن $x_0 \in U$ نقطه اولیه باشد و $F_j(x)$ به ازای حداقل یک $j \in \{1, 2, \dots, m\}$ از پایین کراندار باشد. آن‌گاه هر نقطه حدی دنباله $\{x_k\}$ یک نقطه بحرانی از F است.

برهان. چون $\sum_{j=1}^m \lambda_j(x_k) = 1$ و $\lambda_j(x_k) \geq 0$ است از رابطه (۱۶.۴) داریم:

$$\begin{aligned} F_j(x_{k+1}) &\leq F_j(x_k) + c_1 \alpha_k \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k)^T d_{QN}(x_k) \\ &\leq F_j(x_k) + c_1 \alpha_k \max_{j=1,2,\dots,m} (\nabla F_j(x_k)^T d_{QN}(x_k)). \end{aligned} \quad (۱۷.۴)$$

با توجه به این‌که $B(x_k)$ مثبت معین است، داریم:

$$\begin{aligned} F_j(x_{k+1}) &< F_j(x_k) + c_1 \alpha_k \max_{j=1,2,\dots,m} \{ \nabla F_j(x_k)^T d_{QN}(x_k) + \frac{1}{2} d_{QN}(x_k)^T B(x_k) d_{QN}(x_k) \} \\ &= F_j(x_k) + c_1 \alpha_k t(x_k). \end{aligned} \quad (۱۸.۴)$$

با استفاده از رابطه (۱۸.۴) برای هر $j = 1, 2, \dots, m$ داریم:

$$\begin{aligned} F_j(x_{k+1}) &< F_j(x_{k-1}) + c_1 \alpha_{k-1} t(x_{k-1}) + c_1 \alpha_k t(x_k) \\ &< F_j(x_0) + c_1 \alpha_0 t(x_0) + \dots + c_1 \alpha_k t(x_k) \\ &< F_j(x_0) + c_1 \sum_{i=0}^k \alpha_i t(x_i). \end{aligned} \quad (۱۹.۴)$$

چون $F_j(x)$ برای حداقل یک $j \in \{1, 2, \dots, m\}$ از پایین کراندار است، $j_1 \in \{1, 2, \dots, m\}$ وجود دارد به‌طوری‌که برای هر $x \in U$ ، $F_{j_1}(x) > -\infty$ است. همچنین $\{F_{j_1}(x_k)\}$ دنباله‌ای به‌طور یکنوا نزولی و از پایین کراندار است، لذا $k \rightarrow \infty$ همگرا به یک دنباله $F_{j_1}(x^*) > -\infty$ است. بنابراین از رابطه (۱۹.۴) داریم:

$$F_{j_1}(x_0) - F_{j_1}(x_{k+1}) > -c_1 \sum_{i=0}^k \alpha_i t(x_i) = \sum_{i=0}^k c_1 \alpha_i (-t(x_i)).$$

با گرفتن $k \rightarrow \infty$ از نامساوی فوق داریم:

$$\infty > F_{j_1}(x_0) - F_{j_1}(x^*) > \sum_{i=0}^{\infty} c_1 \alpha_i (-t(x_i)). \quad (۲۰.۴)$$

چون به ازای هر i ، $t(x_i) \leq 0$ است، بنابراین به ازای هر i ، $c_1 \alpha_i (-t(x_i)) > 0$ است. لذا جمع سمت راست رابطه (۲۰.۴)، متناهی است. بنابراین $k \rightarrow \infty$ ، $c_1 \alpha_k (-t(x_k)) \rightarrow 0$.

طول گام α_k کراندار است زیرا اگر k موجود باشد که $\alpha_k \rightarrow \infty$ ، بی‌کران می‌شود و این با فرض مساله تناقض دارد. به‌علاوه اگر β موجود باشد که به ازای هر k ، $\alpha_k \geq \beta > 0$ باشد، آنگاه $k \rightarrow \infty$ ، $-t(x_k) \rightarrow 0$. چون L_0 مجموعه‌ای کراندار است بنابراین $\{x_k\}$ یک دنباله کراندار است و حداقل یک نقطه حدی دارد. فرض کنید $\{p_1^*, \dots, p_r^*\}$ مجموعه‌ای از نقاط حدی دنباله $\{x_k\}$ باشد. از آنجایی که برای هر $s \in \{1, 2, \dots, r\}$ ، p_s^* یک نقطه حدی است و t یک تابع پیوسته است، بنابراین $t(p_s^*) = 0$. بنا به موارد ۱ و ۲ قضیه ۲.۳.۴، برای هر $s \in \{1, 2, \dots, r\}$ یک نقطه بحرانی از F است. $\square$

قضیه ۸.۳.۴. فرض‌های قضیه ۷.۳.۴ را در نظر بگیرید. همچنین فرض کنید، تابع ∇F_j برای هر $j = 1, 2, \dots, m$ پیوسته لیپشیتز باشد و یک $\delta > 0$ وجود داشته باشد به‌طوری‌که به ازای هر

(۲۱.۴) برقرار باشد آن‌گاه هر نقطه حدی دنباله $\{x_k\}$ ، یک جواب کارای موضعی ضعیف برای مساله (۵.۱) است.

$$\sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_{k+1})^T d_{QN}(x_k) \geq c_2 \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k)^T d_{QN}(x_k). \quad (21.4)$$

که $1 < c_2 < c_1$ است.

برهان. چون فرض‌های قضیه ۷.۳.۴ برقراراند، بنابراین هر نقطه حدی دنباله $\{x_k\}$ یک نقطه بحرانی برای F است. فرض کنید x^* یک نقطه حدی برای دنباله $\{x_k\}$ باشد. بنا به لم ۳.۲.۴ اگر حداقل یک $j_0 \in \{1, 2, \dots, m\}$ موجود باشد به‌طوری‌که $\nabla F_{j_0}(x^*) = 0$ باشد آن‌گاه x^* یک جواب کارای ضعیف است.

چون به ازای هر $j = 1, 2, \dots, m$ ∇F_j پیوسته لیپ‌شیتز است، بنابراین $L_j > 0$ وجود دارد، به‌طوری‌که به ازای هر $j = 1, 2, \dots, m$ $\|\nabla F_j(x) - \nabla F_j(y)\| \leq L_j \|x - y\|$ ، بنا به نامساوی کوشی شوارتز به ازای هر $j = 1, 2, \dots, m$ داریم:

$$\begin{aligned} (\nabla F_j(x_{k+1}) - \nabla F_j(x_k))^T d_{QN}(x_k) &\leq \|\nabla F_j(x_{k+1}) - \nabla F_j(x_k)\| \|d_{QN}(x_k)\| \\ &\leq L_j \|x_{k+1} - x_k\| \|d_{QN}(x_k)\| \\ &\leq L \alpha_k \|d_{QN}(x_k)\|^2, \end{aligned}$$

که $L = \max_{j=1,2,\dots,m} L_j$ است. از آنجایی که $\sum_{j=1}^m \lambda_j(x_k) = 1$ و بنا به رابطه (۲۱.۴) داریم:

$$\begin{aligned} L \alpha_k \|d_{QN}(x_k)\|^2 &\geq \max_{j=1,2,\dots,m} (\nabla F_j(x_{k+1}) - \nabla F_j(x_k))^T d_{QN}(x_k) \\ &\geq \sum_{j=1}^m \lambda_j(x_k) (\nabla F_j(x_{k+1}) - \nabla F_j(x_k))^T d_{QN}(x_k) \\ &\geq (c_2 - 1) \sum_{j=1}^m \lambda_j(x_k) \nabla F_j(x_k)^T d_{QN}(x_k) \\ &\geq (c_2 - 1) \max_{j=1,2,\dots,m} \nabla F_j(x_k)^T d_{QN}(x_k). \end{aligned}$$

نامساوی آخر برقرار است چون $c_2 < 1$ است. از رابطه بالا نتیجه می‌شود:

$$\alpha_k \geq \frac{c_2 - 1}{L \|d_{QN}(x_k)\|^2} \max_{j=1,2,\dots,m} \nabla F_j(x_k)^T d_{QN}(x_k).$$

چون $\max_{j=1,2,\dots,m} \nabla F_j(x_k)^T d_{QN}(x_k) < 0$ داریم:

$$\begin{aligned} \alpha_k \left(\max_{j=1,2,\dots,m} \nabla F_j(x_k)^T d_{QN}(x_k) \right) &\leq \frac{c_2 - 1}{L \|d_{QN}(x_k)\|^2} \left(\max_{j=1,2,\dots,m} \nabla F_j(x_k)^T d_{QN}(x_k) \right)^2 \\ &\leq \frac{c_2 - 1}{L \|d_{QN}(x_k)\|^2} \min_{j=1,2,\dots,m} (\nabla F_j(x_k)^T d_{QN}(x_k))^2. \end{aligned}$$

با ضرب $-c_1$ در رابطه بالا داریم:

$$\begin{aligned} -c_1 \alpha_k \left(\max_{j=1,2,\dots,m} \nabla F_j(x_k)^T d_{QN}(x_k) \right) &\geq \frac{c_1(1-c_2)}{L \|d_{QN}(x_k)\|^2} \min_{j=1,2,\dots,m} (\nabla F_j(x_k)^T d_{QN}(x_k))^2 \\ &= \frac{c_1(1-c_2)}{L \|d_{QN}(x_k)\|^2} \min_{j=1,2,\dots,m} [\|\nabla F_j(x_k)\|^2 \|d_{QN}(x_k)\|^2 \cos^2 \theta_j^k] \\ &= \frac{c_1(1-c_2)}{L} \min_{j=1,2,\dots,m} [\|\nabla F_j(x_k)\|^2 \cos^2 \theta_j^k], \end{aligned} \quad (22.4)$$

که θ_j^k زاویه بین $\nabla F_j(x_k)$ و $d_{QN}(x_k)$ است. چون فرض‌های قضیه ۷.۳.۴ برقراراند، لذا بنا به رابطه (۱۷.۴)، قضیه ۷.۳.۴ داریم:

$$\begin{aligned} \infty > F_{j_1}(x_0) - F_{j_1}(x_{k+1}) &\geq -c_1 \sum_{i=0}^k \alpha_i \max_{j=1,2,\dots,m} \nabla F_j(x_i)^T d_{QN}(x_i) \\ &= \sum_{i=0}^k \alpha_i \left(-c_1 \max_{j=1,2,\dots,m} \nabla F_j(x_i)^T d_{QN}(x_i) \right). \end{aligned}$$

اگر $k \rightarrow \infty$ داریم:

$$\infty > F_{j_1}(x_0) - F_{j_1}(x^*) \geq \sum_{i=0}^{\infty} \alpha_i \left(-c_1 \max_{j=1,2,\dots,m} \nabla F_j(x_i)^T d_{QN}(x_i) \right). \quad (23.4)$$

چون $-c_1 \max_{j=1,2,\dots,m} \nabla F_j(x_i)^T d_{QN}(x_i) > 0$ ، سمت راست رابطه (۲۳.۴) مجموعی از اعداد مثبت است. بنابراین $k \rightarrow \infty$ ، $\alpha_k (-c_1 \max_{j=1,2,\dots,m} \nabla F_j(x_k)^T d_{QN}(x_k)) \rightarrow 0$ ، از رابطه (۲۲.۴) داریم:

$$\frac{c_1(1-c_2)}{L} \min_{j=1,2,\dots,m} [\|\nabla F_j(x_k)\|^2 \cos^2 \theta_j^k] \rightarrow 0.$$

چون به ازای هر $j = 1, 2, \dots, m$ ، $\cos^2 \theta_j^k > \delta$ است، بنابراین با گرفتن $k \rightarrow \infty$ داریم:

$$\min_{j=1,2,\dots,m} \|\nabla F_j(x_k)\|^2 \rightarrow 0.$$

لذا $j_0 \in \{1, 2, \dots, m\}$ وجود دارد به گونه‌ای که $k \rightarrow \infty$ ، رابطه زیر برقرار است:

$$\|\nabla F_{j_0}(x_k)\|^2 = \min_{j=1,2,\dots,m} \|\nabla F_j(x_k)\|^2 \rightarrow 0.$$

چون $\|\nabla F_{j_0}(x)\|$ یک تابع پیوسته است و $k \rightarrow \infty$ ، $\|\nabla F_{j_0}(x_k)\| \rightarrow 0$ ، بنابراین به ازای هر نقطه حدی x^* از $\{x_k\}$ ، $\nabla F_{j_0}(x^*) = 0$ است. بنابراین بنا به لم ۳.۲.۴، یک جواب کارای موضعی ضعیف برای مساله (۵.۱) است. $\square$

گزاره ۹.۳.۴. باید توجه داشت، که قضایا ۷.۳.۴ و ۸.۳.۴ روی یک زیر مجموعه فشرده برقراراند. لذا زیر مجموعه فشرده به اندازه کافی بزرگ از $\mathbb{R}^n$ را به عنوان دامنه‌ای از (۵.۱) در نظر می‌گیریم، بدین ترتیب این قضایا برای مساله (۵.۱) برقراراند. این نتایج را در الگوریتم زیر خلاصه می‌کنیم.

الگوریتم شبه‌نیوتن اصلاح شده برای مسائل چندهدفه

الگوریتم ۱۰.۳.۴. مقداردهی اولیه $k = 0$.

i_1 . نقطه $x_0 \in U$ را انتخاب کنید که U یک زیر مجموعه فشرده از $\mathbb{R}^n$ است.

i_2 . ماتریس مثبت معین $B(x_0)$ ، ε_1 ، ε_2 ($\varepsilon_1 < \varepsilon_2$) و عدد مثبت کوچک δ را در نظر بگیرید.

گام‌های تکراری:

k_1 . روابط (۷.۴)–(۳.۴) را به ازای $\lambda_j(x_k)$ حل کنید و جهت کاهنده $d_{QN}(x_k)$ را پیدا کنید.

k_2 . اگر به ازای هر $j = 1, 2, \dots, m$ ، $\cos^2(\theta_j^k) > \delta$ باشد آنگاه طول گام α_k را به گونه‌ای انتخاب کنید که $x_k + \alpha_k d_{QN}(x_k) \in U$ باشد و در روابط (۱۶.۴) و (۲۱.۴) صدق کند. در غیر این صورت، طول گام α_k را به گونه‌ای انتخاب کنید که $x_k + \alpha_k d_{QN}(x_k) \in U$ باشد و در رابطه (۱۶.۴) صدق کند. $x_{k+1} = x_k + \alpha_k d_{QN}(x_k)$ قرار دهید.

k_3 . اگر $|t(x_k)| < \varepsilon_1$ یا $\min_{j=1,2,\dots,m} \|\nabla F_j(x_k)\| < \varepsilon_2$ متوقف شوید. در غیر این صورت به گام بعد بروید.

k_4 . ماتریس مثبت معین $B(x_{k+1})$ را با استفاده از فرمول (۱۴.۴) تولید کنید. قرار دهید $k = k + 1$ و به گام k_1 بروید.

الگوریتم شبه‌نیوتن اصلاح شده برای مسائل چندهدفه نامحدب نیز قابل استفاده است اما دو الگوریتم نیوتن و شبه‌نیوتن که در فصل‌های قبل بیان کردیم تنها برای مسائل محدب، جواب کارا دارند. همچنین این الگوریتم برای تمامی توابع هدف تنها یک ماتریس تقریب هسین محاسبه می‌کند.

۴.۴ مثال‌های عددی

در این قسمت تعدادی مثال را به کمک الگوریتم ارائه شده، حل می‌کنیم. اگر با نقطه‌های شروع متفاوت الگوریتم را آغاز کنیم، جواب‌های کارای متفاوتی را خواهیم داشت. اگر فرض‌های قضیه ۸.۳.۴ برقرار باشند آنگاه هر نقطه حدی دنباله $\{x_k\}$ یک نقطه کارای ضعیف از مساله (۵.۱) است. لذا باید فرض $\cos^2(\theta_j^k) > \delta$ نیز برقرار باشد تا دنباله $\{x_k\}$ یک نقطه کارای ضعیف از مساله باشد. در غیر این صورت فرض‌های قضیه ۷.۳.۴ بررسی می‌شوند و نقطه حدی یک نقطه بحرانی از مساله (۵.۱) می‌شود. هر دو مورد در مثال‌های زیر در بعدهای مختلف بررسی می‌شوند.

مثال ۱.۴.۴. مساله دوهدفه زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}} F(x) = (F_1(x), F_2(x))$$

$$F_1(x) = x^2 - 4$$

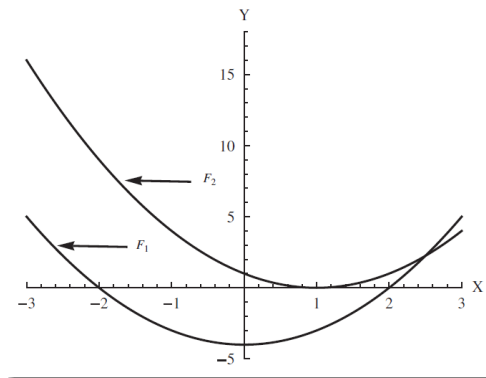
$$F_2(x) = (x - 1)^2.$$

$$\nabla F_1(x) = 2x, \quad \nabla^2 F_1(x) = 2,$$

و

$$\nabla F_2(x) = 2(x-1), \quad \nabla^2 F_2(x) = 2.$$

با استفاده از الگوریتم ۱۰.۳.۴ این مساله حل می‌شود. $\varepsilon_1 = 10^{-7}$, $\varepsilon_2 = 10^{-11}$ و $\delta = 10^{-4}$ در نظر بگیرید. با نقطه اولیه $x^{(0)} = 10$ جواب این مساله بعد از ۲۹ تکرار، نقطه $x^{(29)} = 1.000240356 \simeq 1$ است. توجه داشته باشید که این مساله محدب است و با استفاده از الگوریتم نیوتن ۳.۳.۲ در مثال ۲.۷.۲ حل کردیم. توابع هدف F_1 و F_2 در شکل ۱.۴ رسم کردیم.



شکل ۱.۴: توابع هدف مثال ۱.۴.۴

مثال ۲.۴.۴. مساله دوهدفه زیر را در نظر بگیرید:

$$\begin{aligned} \min_{x \in \mathbb{R}} F(x) &= (F_1(x), F_2(x)) \\ F_1(x) &= \frac{1}{4}[(x_1 - 1)^4 + 2(x_2 - 2)^4] \\ F_2(x) &= (x_2 - x_1^2)^2 + (1 - x_1)^2. \end{aligned}$$

$$\nabla F_1(x) = \begin{bmatrix} (x_1 - 1)^3 \\ 2(x_2 - 2)^3 \end{bmatrix}, \quad \nabla^2 F_1(x) = \begin{bmatrix} 3(x_1 - 1)^2 & 0 \\ 0 & 6(x_2 - 2)^2 \end{bmatrix},$$

و

$$\nabla F_2(x) = \begin{bmatrix} 4x_1^2 - 4x_2x_1 - 2(1 - x_1) \\ 2(x_2 - x_1^2) \end{bmatrix}, \quad \nabla^2 F_2(x) = \begin{bmatrix} -4x_2 + 12x_1^2 + 2 & -4x_1 \\ -4x_1 & 2 \end{bmatrix}.$$

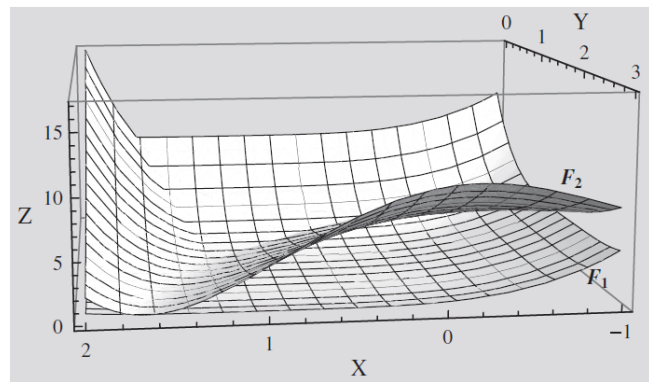
تابع $F_2(x)$ محدب نیست چون تابع $\nabla^2 F_2(x)$ مثبت معین نیست. بنابراین این مساله با استفاده از دو الگوریتم نیوتن و شبه‌نیوتن قابل حل نیست. این مساله با استفاده از الگوریتم ۱۰.۳.۴ حل می‌شود. شرط توقف $\min\{\|\nabla F_1(x^{(k)})\|, \|\nabla F_2(x^{(k)})\|\} < 10^{-5}$ ، نقطه اولیه $x^{(0)} = (0, 3)$ ، $\delta = 10^{-4}$ و $B(x^{(0)}) = I$ را در نظر بگیرید. با حل روابط (۷.۴)-(۳.۴)، $\lambda_1 = 1$ و $\lambda_2 = 0$ جهت $d_{QN}(x^{(k)}) = (1, -2)$ به‌دست آمد. شرایط قضیه ۸.۳.۴ برای آن برقرار است. جواب این مساله نقطه

$(\nabla F_1(x^{(k)}))^T = (0, 0)^T$ چون به ازای این نقطه $x^{(16)} = (0.9930, 2.0143) \simeq (1, 2)$ است که در شرط توقف صدق می‌کند. جدول ۱۰۴ به ازای نقطه اولیه $x^{(0)} = (0, 3)$ نقاط به‌دست آمده در هر گام را تا برقراری شرط توقف نشان می‌دهد.

جدول ۱۰۴: جدول مربوط به مثال ۲.۴.۴

$\min_{j=1,2} \{\ \nabla F_j(x^{(k)})\ \}$	$\min_{j=1,2} \{\cos^2(\theta_j^k)\}$	$d_{QN}(x^{(k)})$	$x^{(k)}$	k
	۰.۹۸۹۹	$(1, -2)$	$(0, 3)$	۰
۰.۲۳۹۶	۰.۶۵۳۷	$(0.6124, 2.4880)$	$(0.25, 2.5)$	۱
۰.۱۴۵۱	۰.۹۰۲۱	$(0.0621, -0.0750)$	$(0.3124, 2.4880)$	۲
۰.۰۵۴۶	۰.۹۰۴۳	$(0.0930, -0.1147)$	$(0.6745, 2.4131)$	۳
...	...	...	...	...
$10^{-6} \times 5.9128$	۰.۸۳۶۳	$(0.0035, -0.0045)$	$(0.9895, 2.0188)$	۱۵

می‌توانید در شکل ۲.۴ تغییرات توابع $F_1(x)$ و $F_2(x)$ را در نقطه $(1, 2)$ مشاهده کنید. همان طور که در شکل می‌بینید در این همسایگی F_1 و F_2 به‌ترتیب افزایش و کاهش پیدا می‌کنند.



شکل ۲.۴: توابع هدف مثال ۲.۴.۴

مثال ۳.۴.۴. مساله سه‌هدفه زیر را در نظر بگیرید:

$$\begin{aligned}
 \min_{x \in \mathbb{R}} \quad & F(x) = (F_1(x), F_2(x), F_3(x)) \\
 F_1(x) = & \frac{1}{9} [(x_1 - 1)^4 + 2(x_2 - 2)^4 + 3(x_3 - 3)^4] \\
 F_2(x) = & e^{\frac{x_1 + x_2 + x_3}{3}} + x_1^2 + x_2^2 + x_3^2 \\
 F_3(x) = & \frac{1}{12} [3e^{-x_1} + 4e^{-x_2} + 3e^{-x_3}].
 \end{aligned}$$

نقطه اولیه $x^{(0)} = (0, 0, 0)$ ، ماتریس اولیه $B(x^{(0)}) = I$ و شرط توقف $|t(x^{(k)})| < 10^{-10}$ را در نظر بگیرید. در اولین تکرار $\cos^2(\theta_1^0) = 10^{-7} \times 8.6226 < 10^{-5} = \delta$ بنابراین قضیه ۷.۳.۴ برای آن برقرار است. نتایج به‌صورت خلاصه در جدول ۲.۴ مشخص شده است. سرانجام بعد

از ۲۹ تکرار نقطه $x^{(29)} = (-0.02199, 0.02157, -0.000390)$ به دست می‌آید که تقریبی از نقطه بحرانی است.

جدول ۲.۴: جدول مربوط به مثال ۳.۴.۴

$ t(x^{(k)}) $	$(d_{QN}(x^{(k)}))^T$	$x^{(k)}$	k
$10^{-4} \times 5.80630$	$\begin{pmatrix} -0.02337 \\ 0.02439 \\ -0.000450 \end{pmatrix}$	$(0, 0, 0)$	۰
$10^{-4} \times 3.94187$	$\begin{pmatrix} -0.01858 \\ 0.01901 \\ -0.000348 \end{pmatrix}$	$(-0.02337, 0.02439, -0.000450)$	۱
$10^{-4} \times 2.21905$	$\begin{pmatrix} -0.01417 \\ 0.01406 \\ -0.000255 \end{pmatrix}$	$(-0.000757, 0.000780, -0.000143)$	۲
...	...	...	...
$10^{-11} \times 7.04398$	$\begin{pmatrix} -10^{-5} \times 0.83423 \\ 10^{-5} \times 0.75591 \\ -10^{-5} \times 0.13302 \end{pmatrix}$	$(-0.02199, 0.02157, -0.000390)$	۲۸

فصل ۵

حل مسائل بهینه‌سازی چندهدفه ناهموار با روش شبه‌نیوتن

۱.۵ تعاریف اولیه

در این فصل الگوریتمی ارائه می‌دهیم که نقطه بحرانی برای مسائل چندهدفه ناهموار و نامحدب را محاسبه کند. برای این منظور از روش شبه‌نیوتن استفاده می‌کنیم. در هر تکرار، جهت کاهنده با حل مساله برنامه‌ریزی خطی با محدودیت‌های درجه دوم محدب به‌دست می‌آید. تحت فرض‌های مناسب همگرایی الگوریتم را نیز نشان می‌دهیم.

مساله برنامه‌ریزی چندهدفه ناهموار زیر را در نظر بگیرید:

$$\begin{aligned} \min \quad & F(x) = (F_1(x), \dots, F_m(x)) \\ \text{s.t.} \quad & x \in X \subset \mathbb{R}^n. \end{aligned} \quad (1.5)$$

که $F_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 1, 2, \dots, m$ توابع هدف هستند و X یک مجموعه باز است.

تعریف ۱.۱.۵. اگر $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ لیپ‌شیتز موضعی باشد، مشتق جهتی تابع F در جهت d به‌صورت زیر تعریف می‌شود:

$$F'(x, d) = \lim_{\alpha \downarrow 0} \frac{F(x + \alpha d) - F(x)}{\alpha}. \quad (2.5)$$

اگر مشتقات جهتی تابع F در نقطه x موجود باشد، ماتریس ژاکوبین در نقطه x را با نماد $DF(x)$ نشان می‌دهیم.

تعریف ۲.۱.۵. پوسته محدب مجموعه S ، کوچکترین مجموعه محدب شامل S است که با نماد $\text{conv} S$ نشان می‌دهیم.

تعریف ۳.۱.۵. مشتق کلارک، تابع F در نقطه x به‌صورت زیر تعریف می‌شود:

$$\partial F(x) = \text{conv} \left\{ \lim_{k \rightarrow \infty} DF(x_k) \mid x_k \rightarrow x \right\}. \quad (3.5)$$

تعریف ۴.۱.۵. نقطه $x^* \in X$ یک نقطه بحرانی برای F است اگر

$$R(\partial F(x^*)) \cap (-\mathbb{R}_{++}^m) = \emptyset. \quad (۴.۵)$$

تعریف ۵.۱.۵. فرض کنید یک مشتق جهتی برای تابع F_j به ازای هر $j = 1, 2, \dots, m$ در جهت d وجود داشته باشد. آن‌گاه $d \in \mathbb{R}^n$ یک جهت کاهنده برای F در x است اگر برای هر $j = 1, 2, \dots, m$ مشتق جهتی F_j در جهت d در شرط زیر صدق کند:

$$F'_j(x, d) < 0. \quad (۵.۵)$$

۲.۵ الگوریتم

در این قسمت، نقطه بحرانی را برای مساله چندهدفه ناهموار (۱.۵) با استفاده از روش شبه‌نیوتن ناهموار بررسی می‌کنیم. با این فرض که، $F : X \rightarrow \mathbb{R}^m$ لیپ‌شیتز موضعی است.

فرض ۱.۲.۵. فرض کنید برای هر $x \in X$ و $d \in \mathbb{R}^n$ به ازای هر $j = 1, 2, \dots, m$ موجود باشد.

گزاره ۱.۲.۵ [۳۲] فرض کنید $x \in \mathbb{R}^n$ باشد و برای هر $d \in \mathbb{R}^n$ حد زیر موجود باشد،

$$\lim_{U \in \partial F(x+\alpha d), \alpha \downarrow 0} \{U^T d\}.$$

آن‌گاه مشتق جهتی F وجود دارد و $F'(x, d) = \lim_{U \in \partial F(x+\alpha d), \alpha \downarrow 0} \{U^T d\}$.

بنا به تعریف ۴.۱.۵، اگر $x^* \in X$ نقطه بحرانی F باشد آن‌گاه در x^* جهت کاهنده وجود ندارد. فرض ۱.۲.۵ وجود مشتق جهتی را تضمین می‌کند.

لم ۲.۲.۵. تحت فرض ۱.۲.۵، اگر $x^* \in X$ نقطه بحرانی F باشد آن‌گاه جهت کاهنده برای یک تابع هدف، بدون بدتر کردن حداقل یک تابع هدف دیگر وجود ندارد. یعنی اگر $\bar{d}$ یک جهت کاهنده برای F_j در x^* باشد آن‌گاه حداقل یک $j_1 = j_1(\bar{d})$ وجود دارد به‌طوری‌که $F'_{j_1}(x^*, \bar{d}) \geq 0$.

برهان. بنا به فرض ۱.۲.۵ و گزاره ۱.۲.۵، مشتق جهتی $F'_j(x, d)$ وجود دارد و مساوی با حد فرض ۱.۲.۵ است. یعنی برای هر $j = 1, 2, \dots, m$ $F'_j(x, d) = \lim_{U_j \in \partial F_j(x+\alpha d), \alpha \downarrow 0} \{U_j^T d\}$. لذا از رابطه (۴.۵) داریم، اگر $x^* \in X$ نقطه بحرانی F و $\bar{d}$ یک جهت کاهنده برای F_j در x^* باشد آن‌گاه حداقل یک $j_1 = j_1(\bar{d})$ وجود دارد به‌طوری‌که $F'_{j_1}(x^*, \bar{d}) \geq 0$. در غیر این صورت، اگر $\bar{d}$ یک جهت کاهنده برای هر تابع F_j در x^* باشد آن‌گاه بنا به فرض ۱.۲.۵ برای هر $j = 1, 2, \dots, m$ $F'_j(x, \bar{d}) < 0$ و لذا برای هر $j = 1, 2, \dots, m$ $\lim_{U_j \in \partial F_j(x+\alpha d), \alpha \downarrow 0} \{U_j^T d\} < 0$ و این یعنی $R(\partial F(x^*)) \cap (-\mathbb{R}_{++}^m) \neq \emptyset$ بنا براین رابطه (۴.۵) برقرار نیست و این تناقض با بحرانی بودن $x^* \in X$ دارد. لذا حکم برقرار است. $\square$

بنا به لم ۲.۲.۵، در یک نقطه بحرانی جهت کاهنده برای تمام F_j ، $j = 1, 2, \dots, m$ وجود ندارد. این استدلال را برای حالت کلی تعمیم می‌دهیم و لم ۳.۲.۵ را به‌صورت زیر بیان می‌کنیم.

لم ۳.۲.۵. تحت فرض ۱.۲.۵، $x^* \in X$ روی F بحرانی است اگر و فقط اگر در یکی از دو شرط زیر صدق کند:

الف. حداقل یک $j_0 = j_0(d) \in \{1, 2, \dots, m\}$ وجود داشته باشد به طوری که جهت کاهنده در x^* برای F_{j_0} وجود نداشته باشد. یعنی برای هر d ، $F'_{j_0}(x^*, d) \geq 0$.

ب. در حالت خاص، حداقل یک $j_0 \in \{1, 2, \dots, m\}$ وجود داشته باشد به طوری که $0 \in \partial F_{j_0}(x^*)$.

برهان. فرض کنید $x^* \in X$ یک نقطه بحرانی باشد. مشابه با اثبات لم ۲.۲.۵، برای هر $j = 1, 2, \dots, m$ ، $F'_j(x, d) = \lim_{U_j \in \partial F_j(x + \alpha d), \alpha \downarrow 0} \{U_j^T d\}$. بنا به رابطه (۴.۵) اگر $x^* \in X$ بحرانی باشد، حداقل یک $j_0 = j_0(d) \in \{1, \dots, m\}$ وجود دارد به طوری که $F'_{j_0}(x^*, d) \geq 0$. اگر الف یا ب یا هر دو برقرار باشند آن گاه از تعریف ۴.۱.۵ و فرض این لم داریم که x^* نقطه بحرانی F است. $\square$

ملاحظه ۴.۲.۵. الف. بنا به لم ۳.۲.۵، در یک نقطه بحرانی، جهت کاهنده برای حداقل یک تابع هدف F_j وجود ندارد یا در حالت خاص، آن یک نقطه بحرانی برای حداقل یک تابع هدف F_j است. در این فصل ما یک جواب ارائه می دهیم که در مورد الف لم ۳.۲.۵ صدق می کند. یعنی جستجو زمانی متوقف می شود که جهت کاهنده پیدا نشود.

ب. بنا به لم ۳.۲.۵، اگر x نقطه غیر بحرانی برای F باشد آن گاه $d \in \mathbb{R}^n$ وجود دارد به طوری که برای هر $j = 1, 2, \dots, m$ ، $F'_j(x, d) < 0$. بدین معنی است که d یک جهت کاهنده برای F در نقطه x است یعنی $0 < \alpha_0$ وجود دارد به طوری که برای هر $\alpha \in (0, \alpha_0]$ و هر $j = 1, 2, \dots, m$ داریم:

$$F_j(x + \alpha d) < F_j(x). \quad (۶.۵)$$

قضیه ۵.۲.۵. رابطه بین نقاط کارا و نقاط بحرانی را بیان می کند.

قضیه ۵.۲.۵. تحت فرض ۱.۲.۵ موارد زیر را داریم:

الف. اگر $x^* \in X$ ، کارای موضعی ضعیف باشد آن گاه x^* یک نقطه بحرانی F است.

ب. اگر X محدب باشد، $F: \mathbb{R}^m \rightarrow \mathbb{R}$ محدب و $x^* \in X$ نقطه بحرانی F باشد آن گاه x^* کارای ضعیف است.

ج. اگر X محدب باشد، $F: \mathbb{R}^m \rightarrow \mathbb{R}$ محدب اکید و $x^* \in X$ نقطه بحرانی F باشد آن گاه x^* کارا است.

برهان. الف. فرض کنید x^* کارای ضعیف و غیر بحرانی باشد. لذا $0 < \alpha_0$ و $d \in \mathbb{R}^n$ وجود دارند به طوری که (۶.۵) برقرار است، این با کارای ضعیف بودن x^* در تناقض است. بنابراین x^* نقطه بحرانی F است.

ب. چون $x^* \in X$ نقطه بحرانی F است، برای هر $x \in X$ ، j_* وجود دارد به‌طوری‌که $F'_{j_*}(x^*, d) = \lim_{U_j \in \partial F_{j_*}(x^* + \alpha d), \alpha \downarrow 0} \{U_j^T d\}$ مثبت است. بنا به تحدب F_{j_*} به ازای $d = x - x^*$ رابطه زیر را داریم:

$$F_{j_*}(x) - F_{j_*}(x^* + \alpha d) \geq (1 - \alpha)U_j^T(x - x^*), \quad \forall U_j \in \partial F_{j_*}(x^* + \alpha d). \quad (۷.۵)$$

فرض کنید در رابطه (۷.۵)، $\alpha \downarrow 0$ آن‌گاه از پیوستگی F_{j_*} داریم:

$$F_{j_*}(x) - F_{j_*}(x^*) \geq F'_{j_*}(x, x - x^*) \geq 0. \quad (۸.۵)$$

این بدین معنی است که x^* کارای ضعیف است.

ج. چون F محدب اکید است، رابطه (۸.۵) به ازای هر $x \neq x^*$ ، اکید است. لذا حکم برقرار است. $\square$

تعریف ۶.۲.۵. فرض کنید x_* یک نقطه در X باشد و $f : X \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$ یک تابع حقیقی مقدار توسعه یافته باشد. گوئیم f در x_* نیم‌پیوسته پایینی^۱ است اگر $f(x_*) < \infty$ باشد و برای هر $\epsilon > 0$ یک همسایگی V از x_* وجود داشته باشد به‌طوری‌که برای هر $x \in V$ ، $f(x) \geq f(x_*) - \epsilon$ باشد. اگر $f(x_*) = \infty$ باشد، $f(x)$ به سمت بی‌نهایت میل کند. به طور معادل برای فضای متریک $\lim_{x \rightarrow x_*} \inf f(x) \geq f(x_*)$.

برای $x \in X$ ، جهت $d(x)$ مساله (۱.۵) را جواب بهینه‌ای از مساله ناهموار $NSP(x)$ تعریف می‌کنیم.

$$NSP(x) \begin{cases} \min & \max_{j=1,2,\dots,m} \phi_j(x, d) + \frac{1}{2} d^T B_j(x) d \\ s.t. & d \in \mathbb{R}^n. \end{cases} \quad (۹.۵)$$

که $\phi_j(x, d)$ و $\frac{1}{2} d^T B_j(x) d$ ، به‌ترتیب اطلاعات مربوط به مشتق مرتبه اول و مشتق مرتبه دوم از $F_j(x)$ هستند. گرچه، در حالت کلی ممکن است مشتق مرتبه اول و مرتبه دوم $F_j(x)$ موجود نباشند. در این‌جا ما فرض می‌کنیم که $-\phi(\cdot, d)$ برای هر d ، نیم‌پیوسته پایینی است و برای هر $x \in L$ ، $\phi(x, \alpha d) \leq \alpha \phi(x, d)$ است که $0 \leq \alpha \leq 1$ و $L = \{x \in \mathbb{R}^n : F(x) \leq F(x_*)\}$ مجموعه تراز است. ما فرض می‌کنیم که برای هر $x \in X$ و هر j ، $B_j(x)$ مثبت معین است.

فرض ۲.۲.۵. برای هر $x \in X$ ، $\phi(x, 0) = 0$ است و $\phi(x, \cdot)$ نیم‌پیوسته پایینی است.

با در نظر گرفتن $\epsilon \geq 0$ ، مساله $NSP(x)$ را می‌توان به‌صورت تقریبی به فرم مساله هموار محدب زیر بیان کرد:

$$NSP_\epsilon(x) \begin{cases} \min & t \\ s.t. & \phi_j(x, d) + \frac{1}{2} d^T B_j(x) d \leq t \quad j = 1, 2, \dots, m \\ & \|d\| \leq 1, t \leq -\epsilon, d \in \mathbb{R}^n. \end{cases} \quad (۱۰.۵)$$

^۱lower semicontinuous

قید $\|d\| \leq 1$ برای بهبود اجرای الگوریتم قرار داده شده است. زیرا $\|d\| \leq 1$ باعث می‌شود که حالت $\|d\| \rightarrow \infty$ اتفاق نیفتد. اگر $\epsilon = 0$ باشد، $NSP_\epsilon(x)$ با $NSP(x)$ معادل است و مجموعه شدنی آن همواره ناتهی است و یک جواب بهینه منحصر بفرد دارد. به علاوه در این مورد، اگر تابع F_j دوبار به طور پیوسته مشتق پذیر باشد، $\phi_j(x, d) = \nabla F_j(x)^T d$ و $B_j(x) = \nabla^2 F_j(x)$ هستند آن‌گاه مساله (۱۰.۵) به مساله (۱۰.۲) فصل ۲ تبدیل می‌شود. تحت فرض مناسب، اگر مجموعه شدنی $NSP_\epsilon(x)$ ، برای یک اسکالر مثبت به اندازه کافی کوچک $\epsilon > 0$ تهی باشد آن‌گاه جهت کاهنده در x وجود ندارد و تخمین مناسبی از یک نقطه بحرانی به دست می‌آید. علت انتخاب اسکالر مثبت به اندازه کافی کوچک ϵ در مساله $NSP_\epsilon(x)$ این است که $t \leq 0$ با $t \leq -\epsilon$ جایگزین می‌شود. می‌توانیم، جواب بدیهی $(d^*, t^*) = (0, 0)$ را مستثنی کنیم. اگر ϵ ناصفر باشد، ممکن است $NSP_\epsilon(x)$ نشدنی شود. در این حالت نقطه جاری x یک تقریب مناسب برای نقطه بحرانی F است. این موارد را تحت فرض‌های ۳.۲.۵ و ۲.۲.۵ در قضیه ۷.۲.۵ بیان می‌کنیم.

فرض ۳.۲.۵. برای هر $x \in X$ و $d \in \mathbb{R}^n$ داریم:

$$\liminf_{t \downarrow 0} \frac{\phi(x, td)}{t} \leq F'(x, d),$$

$$\phi = (\phi_1, \dots, \phi_m)^T \text{ که}$$

قضیه ۷.۲.۵. فرض کنید فرض‌های ۱.۲.۵، ۲.۲.۵ و ۳.۲.۵ برقرار باشند. اگر برای یک اسکالر مثبت به اندازه کافی کوچک ϵ و نقطه $x \in X$ داده شده، مجموعه شدنی $NSP_\epsilon(x)$ ناتهی باشد آن‌گاه x غیر بحرانی است و هر جواب شدنی $d_\epsilon(x)$ یک جهت کاهنده برای F است. در غیر این صورت، x تخمین مناسبی از نقطه بحرانی F است.

برهان. با برهان خلف نشان می‌دهیم که اگر مجموعه شدنی $NSP_\epsilon(x)$ ، به ازای ϵ به اندازه کافی کوچک، تهی باشد آن‌گاه جهت کاهنده وجود ندارد. در واقع، فرض کنید یک جهت $\bar{d} \in \mathbb{R}^n$ وجود داشته باشد، به طوری که به ازای $j = 1, 2, \dots, m$ رابطه زیر برقرار باشد،

$$F'_j(x, \bar{d}) < 0. \quad (11.5)$$

نامساوی بالا به همراه فرض‌های ۳.۲.۵ و ۲.۲.۵ نتیجه می‌دهند که یک اسکالر مثبت $\bar{\alpha}$ وجود دارد به طوری که برای هر $\alpha \in (0, \bar{\alpha}]$

$$\frac{\phi_j(x, \alpha \bar{d})}{\alpha} + \frac{1}{4} \alpha \bar{d}^T B_j(x) \bar{d} \leq F'_j(x, \bar{d}) + \frac{1}{4} \alpha \bar{d}^T B_j(x) \bar{d} < 0. \quad (12.5)$$

برای هر $\alpha \in (0, \bar{\alpha}]$ ، تعریف می‌کنیم $-\epsilon = \alpha F'_j(x, \bar{d}) + \frac{1}{4} \alpha \bar{d}^T B_j(x) \bar{d}$. آن‌گاه می‌بینیم که $\alpha \bar{d}$ یک جواب شدنی $NSP_\epsilon(x)$ است. این تناقض دارد با این که مجموعه شدنی $NSP_\epsilon(x)$ تهی است. بنابراین برای ϵ به اندازه کافی کوچک، x یک تخمین مناسبی از نقطه بحرانی F است. $\square$

مجموعه زیر را تعریف می‌کنیم:

$$S(x, \epsilon) = \{(d, t) \in \mathbb{R}^{n+1} | \phi_j(x, d) + \frac{1}{4} d^T B_j(x) d \leq t, \forall j, \|d\| \leq 1, t \leq -\epsilon\}. \quad (13.5)$$

بدین ترتیب جواب بهینه مساله $NSP_\epsilon(x)$ ، $\bar{d}_\epsilon(x) = (d_\epsilon(x), t_\epsilon(x))$ و مقدار بهینه متناظرش $v_\epsilon(x)$ است، که به صورت زیر بیان می‌شوند:

$$v_\epsilon(x) = \min_{(d,t) \in S(x,\epsilon)} t \quad (14.5)$$

و

$$\bar{d}_\epsilon(x) = (d_\epsilon(x), t_\epsilon(x)) = \arg \min_{(d,t) \in S(x,\epsilon)} t, \quad (15.5)$$

که $d_\epsilon(x)$ جهت کاهنده است. اگر $\epsilon = 0$ ، تعریف می‌کنیم:

$$v(x) = v_0(x) = \inf_{\|d\| \leq 1} \max_{j=1,2,\dots,m} \phi_j(x,d) + \frac{1}{p} d^T B_j(x) d \quad (16.5)$$

و

$$d(x) = d_0(x) = \arg \min_{\|d\| \leq 1} \max_{j=1,2,\dots,m} \phi_j(x,d) + \frac{1}{p} d^T B_j(x) d. \quad (17.5)$$

لم ۸.۲.۵. اگر فرض‌های ۱.۲.۵، ۲.۲.۵ و ۳.۲.۵ برقرار باشند و $x \in X$ نقطه غیربحرانی F آن‌گاه برای اسکالر به اندازه کافی کوچک و نامنفی ϵ ، مجموعه جواب $S(x, \epsilon)$ ناتهی است و $0 \notin S(x, \epsilon)$.

برهان. فرض کنید $x \in X$ نقطه غیربحرانی F باشد. لذا، برای اسکالر به اندازه کافی کوچک $\epsilon > 0$ ، بنا به قضیه ۷.۲.۵ همیشه یک جهت کاهنده وجود دارد که در محدودیت‌های $NSP_\epsilon(x)$ صدق می‌کند، یعنی $S(x, \epsilon)$ ناتهی است و بنا به فرض ۲.۲.۵، $0 \notin S(x, \epsilon)$. لذا حکم برقرار است. $\square$

بر اساس بحث‌های ذکر شده، در نقطه جاری $x \in X$ با حل مساله $NSP_\epsilon(x)$ ، می‌توانیم با گام‌های تکراری، جهت‌های کاهنده را پیدا کنیم و با استفاده از جهت کاهنده به دست آمده، تکرار جاری را بروز کنیم. این نتایج را در الگوریتم زیر خلاصه می‌کنیم.

الگوریتم ۹.۲.۵. مقداردهی اولیه $k = 0$.

i_1 . نقطه اولیه $x_0 \in X$ را انتخاب کنید.

i_2 . یک اسکالر مثبت به اندازه کافی کوچک ϵ و یک ماتریس مثبت معین اولیه $B_j(x_0)$ را در نظر بگیرید.

گام‌های تکراری:

k_1 . مساله $NSP_\epsilon(x_k)$ را حل کنید. اگر مساله نشدنی باشد، الگوریتم خاتمه می‌یابد. در غیر این صورت به گام k_2 بروید.

k_2 . طول گام تقریبی α_k را انتخاب کنید به طوری که $x_k + \alpha_k d_\epsilon^*(x_k) \in X$ و قرار دهید $x_{k+1} = x_k + \alpha_k d_\epsilon^*(x_k)$ که $d_\epsilon^*(x_k)$ جواب بهینه $NSP_\epsilon(x_k)$ است.

k_3 . به ازای هر $j = 1, \dots, m$ ، ماتریس مثبت معین $B_j(x_{k+1})$ را تولید کنید. قرار دهید $k = k + 1$ و به گام k_1 بروید.

ملاحظه ۱۰.۲.۵. الف. در الگوریتم بالا، به دست آوردن یک جهت کاهنده برای حل مساله $NSP_\epsilon(x_k)$ ، به راحتی انجام می شود. چون مساله، یک برنامه ریزی خطی با محدودیت های درجه دوم محدب است که می تواند به آسانی با به کار بردن روش های نقطه درونی محاسبه شود.

ب. اگر مساله $NSP_\epsilon(x_k)$ نشدنی باشد آنگاه نتیجه می گیریم که تکرار جاری، تقریب مناسبی از نقطه بحرانی است.

ج. در گام k_3 ، از روش BFGS برای بروز کردن $B_j(x_{k+1})$ استفاده می کنیم. برای هر $j = 1, 2, \dots, m$ ، $B_j(x_{k+1})$ مثبت معین بودن را از $B_j(x_k)$ به ارث می برد، تا زمانی که x_k غیر بحرانی است.

۳.۵ همگرایی سراسری

برای اثبات همگرایی سراسری به تعدادی فرض نیاز داریم.

فرض ۱.۳.۵. مجموعه تراز $L_\circ = \{x \in \mathbb{R}^n : F(x) \leq F(x_\circ)\}$ کراندار است.

فرض ۲.۳.۵. برای هر زیر دنباله همگرای $\{x_k\}_{k \in K}$ ، اگر $d_k \rightarrow \circ$ آنگاه

$$\max_{j=1,2,\dots,m} (F_j(x_k + d_k) - F_j(x_k)) \leq \max_{j=1,2,\dots,m} \phi_j(x_k, d_k) + o(\|d_k\|).$$

رابطه زیر را تعریف می کنیم:

$$\psi(x) = \sup_{\|d\| \leq 1} \max_{j=1,2,\dots,m} \{-\phi_j(x, d)\}. \quad (18.5)$$

از فرض ۲.۲.۵ نتیجه می شود که $\psi(x) \geq \circ$. به علاوه تحت فرض های ۱.۲.۵، ۲.۲.۵ و ۳.۲.۵ حکم زیر در مورد ψ واضح است.

لم ۱.۳.۵. اگر فرض های ۱.۲.۵، ۲.۲.۵ و ۳.۲.۵ برقرار باشد و $\psi(x^*) = \circ$ باشد آنگاه x^* نقطه بحرانی برای F است.

برای اثبات همگرایی سراسری، به فرض زیر نیز نیاز داریم.

فرض ۳.۳.۵. فرض کنید که برای k به اندازه کافی بزرگ، طول گام $\alpha_k = 1$ قابل پذیرش است.

قضیه ۲.۳.۵. فرض کنید که برای هر $x \in L_\circ$ و هر $j = 1, 2, \dots, m$ ، یک ثابت c وجود داشته باشد، به طوری که $\|B_j(x)\| \leq c$ آنگاه تحت فرض های ۱.۲.۵، ۲.۲.۵، ۳.۲.۵، ۱.۳.۵، ۲.۳.۵ و ۳.۳.۵ هر نقطه حدی دنباله $\{x_k\}$ ، یک نقطه بحرانی F است.

برهان. بنا به لم ۱.۳.۵، کافی است نشان دهیم که هر نقطه حدی یک جواب از $\psi(\cdot) = \circ$ است. فرض کنید برای $x = x_k$ ، $\bar{d}_k$ در رابطه (۱۸.۵) ماکسیمم باشد. لذا برای هر $\alpha \in [0, 1]$ ، بنا به رابطه (۱۶.۵)

و فرض در مورد ϕ_j ، به ازای هر $j = 1, 2, \dots, m$ داریم:

$$\begin{aligned} -v(x_k) &\geq \max_{j=1,2,\dots,m} [-\phi_j(x_k, \bar{d}_k) - \frac{1}{\gamma} \bar{d}_k^T B_j(x_k) \bar{d}_k] \\ &\geq \max_{0 \leq \alpha \leq 1} \max_{j=1,2,\dots,m} [-\alpha \phi_j(x_k, \bar{d}_k) - \frac{1}{\gamma} \alpha^2 \bar{d}_k^T B_j(x_k) \bar{d}_k] \\ &\geq \max_{0 \leq \alpha \leq 1} [\alpha \psi(x_k) - \frac{1}{\gamma} \alpha^2 c] \geq \frac{1}{\gamma} \psi(x_k) \min\{1, \frac{\psi(x_k)}{c}\}. \end{aligned} \quad (19.5)$$

بدون از دست دادن کلیت، فرض می‌کنیم زیر دنباله $\{x_k\}_{k \in K}$ همگرا به x^* است. آنگاه بنا به فرض ۳.۳.۵، $\bar{k}$ وجود دارد که برای هر $k \geq \bar{k}$ به اندازه کافی بزرگ و $k \in K$ ، طول گام $\alpha_k = 1$ است. از کران‌داری مجموعه L و فرض ۲.۳.۵ نتیجه می‌شود:

$$\begin{aligned} \frac{1}{\gamma} \sum_{k \in K, k \geq \bar{k}} \psi(x_k) \min\{1, \frac{\psi(x_k)}{c}\} &\leq \sum_{k \in K, k \geq \bar{k}} \max_{j=1,2,\dots,m} [F_j(x_k) - F_j(x_{k+1})] \\ &\leq \max_{j=1,2,\dots,m} [F_j(x_0) - F_j(x^*)] < +\infty. \end{aligned} \quad (20.5)$$

این رابطه نتیجه می‌دهد:

$$\sum_{k \in K} \psi(x_k) \min\{1, \frac{\psi(x_k)}{c}\} < +\infty. \quad (21.5)$$

اکنون اثبات می‌کنیم، $\psi(x^*) = 0$ است. به برهان خلف، فرض کنید $\psi(x^*) > 0$ است. بنا به فرض ۳.۳.۵، $\beta > 0$ و $\epsilon_0 > 0$ وجود دارند به طوری که برای هر $0 < \epsilon \leq \epsilon_0$ و $\|x_k - x^*\|_{k \in K} \leq \epsilon$ ،

$$\psi(x_k) \geq \beta > 0. \quad (22.5)$$

لذا داریم:

$$\sum_{k \in K} \psi(x_k) \min\{1, \frac{\psi(x_k)}{c}\} \geq \sum_{\{k | \|x_k - x^*\|_{k \in K} \leq \epsilon\}} \psi(x_k) \min\{1, \frac{\beta}{c}\} = +\infty. \quad (23.5)$$

که با رابطه (۲۱.۵) در تناقض است. بنابراین $\psi(x^*) = 0$ است. بنا به لم ۱.۳.۵، x^* نقطه بحرانی F است. $\square$

به کمک الگوریتم بیان شده، مساله ناهموار چندهدفه (۱۰.۵) را حل می‌کنیم. در هنگام استفاده از الگوریتم، حل مساله $NSP_\epsilon(x)$ اهمیت دارد. گرچه $\phi_j(\cdot, \cdot)$ در $NSP_\epsilon(x)$ به طور ضمنی تعریف شده است، ما باید فرمول صریحی از $\phi_j(\cdot, \cdot)$ برای یک خانواده از توابع لیپ‌شیتز به ازای هر $j = 1, 2, \dots, m$ بیان کنیم. وقتی تابع F_j لیپ‌شیتز قطعه‌ای C^1 و یک خانواده از توابع مینیم-ماکسیم جدا از هم باشد، ما تعریف صریحی از $\phi_j(\cdot, \cdot)$ ارائه می‌دهیم. باید توجه داشت $\phi_j(\cdot, \cdot)$ در فرض‌های ۳.۲.۵، ۲.۲.۵، ۱.۳.۵ و ۲.۳.۵ صدق می‌کند. بنابراین الگوریتم با $\phi_j(\cdot, \cdot)$ تعریف شده همگرای سراسری است.

حالت اول. ابتدا حالتی را بررسی می‌کنیم که F_j لیپ‌شیتز موضعی C^1 است. برای هر $x \in X$ ، مجموعه $\{x | F_j(x) < \infty\}$ می‌تواند به صورت اجتماعی از چند وجهی‌های محدب بیان شود، یعنی یک

مجموعه از توابع $\{F_{ji}\}_{i=1}^{M_i}$ ، برای هر تابع داده شده با فرمول C^1 وجود دارد. که ازای هر $i = 1, 2, \dots, M_i$ یک تابع C^1 است به طوری که

$$\{x | F_j(x) < \infty\} = \{x | F_{ji}(x) < \infty, i = 1, 2, \dots, M_i\}.$$

بنابراین برای هر $j = 1, 2, \dots, m$ ، زیرمشتق تابع F_j در $x \in X$ به صورت زیر تعریف می شود:

$$\partial F_j(x) = \text{conv}\{\nabla F_{ji}(x) | F_j(x) \text{ is given by } F_{ji}(x) \text{ at } x, i = 1, 2, \dots, M_i\}.$$

لذا در این مورد برای هر $j = 1, 2, \dots, m$ ، $\phi_j(.,.)$ به صورت زیر تعریف می شود:

$$\phi_j(x, d) = \max_{v \in \partial F_j(x)} v^T d.$$

با توجه به این تعریف ϕ_j ، باید زیرمشتق ∂F_j محاسبه شود. اگر F_j به طریق جبری به آسانی قابل حل نباشد، به دست آوردن ∂F_j یک مساله باز است.

حالت دوم. در این حالت برای هر $j = 1, 2, \dots, m$ ، F_j به صورت مینیم-ماکسیم مجزا از یک خانواده توابع بیان می شود یعنی برای هر $x \in X$ ، $F_j(x) = \max_{i=1,2,\dots,M_i} \{F_{ji}(x)\}$ ، با $F_{ji} \in C^1$ برای هر $i = 1, 2, \dots, M_i$ است. آنگاه برای هر $j = 1, 2, \dots, m$ ، $\phi_j(.,.)$ به صورت زیر تعریف می شود:

$$\phi_j(x, d) = \max_{i=1,2,\dots,M_i} \{F_{ji}(x) + \nabla F_{ji}(x)^T d\} - F_j(x). \quad (24.5)$$

۱.۳.۵ مثال های عددی

یک مثال عددی ارائه می دهیم که هر تابع هدف، به صورت مینیم-ماکسیم مجزا از یک خانواده توابع بیان می شود. $\phi_j(.,.)$ در فرض های ۲.۲.۵، ۳.۲.۵، ۲.۳.۵ و ۲.۳.۵ صدق می کند و همگرایی الگوریتم سراسری است. در این مثال $\epsilon = 10^{-6}$ انتخاب کنید. در این مثال از شرط توقف $v_\epsilon(x^{(k)}) > -\epsilon$ استفاده می شود.

مثال ۳.۳.۵. مساله زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}^3} F(x) = (F_1(x), F_2(x)) = (\max\{F_{11}(x), F_{12}(x)\}, \max\{F_{21}(x), F_{22}(x)\})$$

$$F_{11}(x) = -1/0 + 8x_1 + 8x_2 - 3x_1x_2$$

$$F_{12}(x) = 3/6 - 12x_1 - 4x_2 + 4x_1x_2 + 10x_1^2 + 2x_2^2$$

$$F_{21}(x) = x_1 + x_2^2 + x_3^2 + \sin^2(x_1 + x_2 + x_3)$$

$$F_{22}(x) = \cos(x_3)(0/1 + x_2)(\exp(-\frac{x_1}{0/1 + x_3}))$$

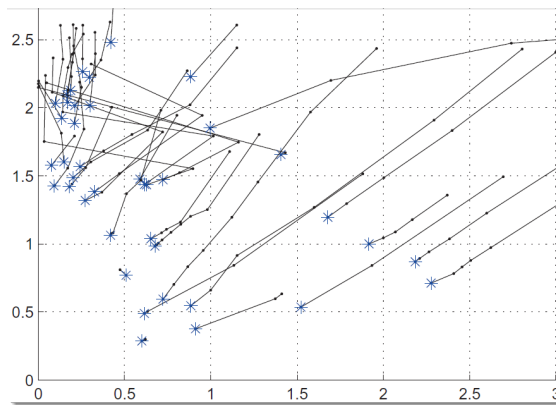
$$-10^{-4} < x_i < 1 + 10^{-4}, \quad i = 1, 2, 3.$$

تابع F لیپشیتز موضعی است و در فرض ۱.۲.۵ صدق می کند. فرض کنید $i_1(x) = \arg \max_{k=1,2} F_{1k}(x)$ و $i_2(x) = \arg \max_{k=1,2} F_{2k}(x)$ باشند. توابع $\phi_1(x, d) = \nabla F_{i_1(x)}(x)^T d$ و $\phi_2(x, d) =$

را تعریف می‌کنیم. بنابراین ϕ_1 و ϕ_2 تعریف شده در بالا، در فرض‌های ۳.۲.۵، ۲.۲.۵، ۱.۳.۵ و ۲.۳.۵ صدق می‌کنند. در گام k_3 الگوریتم، با کمک روش BFGS ماتریس $B_j(x_{k+1})$ را بروز می‌کنیم. ابتدا مجموعه اندیس $\tilde{K}_j = \{k : y_j(x_k)^T s_k \geq \epsilon \min\{-v_\epsilon(x_k), 1\}\}$ که $y_j(x_k) = \nabla F_{j i_j(x_k)}(x_{k+1}) - \nabla F_{j i_j(x_k)}(x_k)$ و $s_k = x_{k+1} - x_k$ هستند را تعریف می‌کنیم. ماتریس شبه‌نیوتن $B_j(x_{k+1})$ به صورت زیر تعیین می‌شود:

$$B_j(x_{k+1}) = \begin{cases} B_j(x_k) - \frac{B_j(x_k)s_k s_k^T B_j(x_k)}{s_k^T B_j(x_k)s_k} + \frac{y_j(x_k)y_j(x_k)^T}{s_k^T y_j(x_k)}, & \text{if } k \in \tilde{K}_j, \\ B_j(x_k), & \text{otherwise.} \end{cases} \quad (25.5)$$

با توجه به این فرمول برای هر j در نقطه غیربحرانی x_k ، $B_j(x_{k+1})$ مثبت معین بودن را از $B_j(x_k)$ به ارث می‌برد. این مساله را به کمک الگوریتم ارائه شده با ۴۰ نقطه شروع تصادفی حل می‌کنیم. نقاط بحرانی مساله در شکل ۱۰۵ نشان داده شده‌اند. نقاط دایره‌ای شکل، نقاط بحرانی مساله به ازای ۴۰ نقطه تصادفی را نشان می‌دهد. میانگین مینیم تکرارها، میانگین ماکسیم تکرارها و میانگین تعداد محاسبات را به ازای ۴۰ نقطه شروع تصادفی به ترتیب ۱، ۱۱ و ۲۳۴.۰۰ هستند.



شکل ۱۰۵: طول و عرض محور مختصات به ترتیب مربوط به تابع F_1 و F_2 مساله ۳.۳.۵ است.

نتیجه‌گیری و پیشنهادات

نتیجه‌گیری

افزایش دامنه کاربرد بهینه‌سازی چندهدفه در علوم مختلف، نیاز به شیوه‌های حل مناسب را گسترش داده است. روش‌ها و الگوریتم‌های متفاوتی برای حل مسائل چندهدفه مطرح شده است که هر کدام دارای ویژگی‌های خاصی هستند. در این پایان‌نامه برای حل مسائل چندهدفه از روش‌های تکراری از جمله روش نیوتن و شبه‌نیوتن استفاده شده است. همان‌طور که در مثال‌های عددی مطرح کردیم، روش شبه‌نیوتن برای مسائل محدب با ابعاد بزرگ در مقایسه با روش نیوتن سریع‌تر عمل می‌کند و ما را به جواب می‌رساند. لذا برای مسائلی با ابعاد کوچک و مسائلی که محاسبه هسین به راحتی صورت می‌گیرد روش نیوتن به روش شبه‌نیوتن ترجیح داده می‌شود. در الگوریتم شبه‌نیوتن تابع هدف باید محدب باشد لذا برای مسائلی که تابع هدف نامحدب است، این روش قابل استفاده نیست. بنابراین در چنین مواردی می‌توان از روش شبه‌نیوتن اصلاح شده استفاده کرد که برای توابع نامحدب نیز کاربرد دارد. در الگوریتم‌های نیوتن، شبه‌نیوتن و شبه‌نیوتن اصلاح شده باید توابع هدف به‌طور پیوسته مشتق‌پذیر باشند. لذا این الگوریتم برای مسائل ناهموار قابل استفاده نیست. مسائل ناهموار با کمک الگوریتم ارائه شده در فصل آخر قابل حل هستند.

پیشنهادهای

موارد زیر از جمله پیشنهادات آتی برای کار در زمینه حل مسائل چندهدفه با کمک روش‌های تکراری است:

۱. ارائه الگوریتم نیوتن برای حل مسائل چندهدفه نامحدب
۲. ارائه الگوریتم نیوتن برای حل مسائل چندهدفه ناهموار
۳. ارائه الگوریتم شبه‌نیوتن فازی برای حل مسائل چندهدفه فازی

پوست آ

کد متلب

```
function f=fminmulti(x)
f=x(2);
end
*****
function [c,ceq]=multicon(x,xk)

gf1=@(t)2*t;
gf2=@(t)2*(t-1);
gff1=@(t)2;
gff2=@(t)2;

c1=gf1(xk(1))*x(1)+0.5*gff1(xk(1))*x(1)^2-x(2);
c2=gf2(xk(1))*x(1)+0.5*gff2(xk(1))*x(1)^2-x(2);
c=[c1;c2];
ceq=[];
end
*****
clc
clear all
n=0:100;
J=1./2.^n;
sigma=0.1;

%xk=[5;1];
xk=rand(1,2)

%x0=[.1;.1];
x(2)=-10^12;
```

```

f1=@(t)t^2-4;
f2=@(t)(t-1)^2;
%while x(2)~=0
    %for k=1:5
        s=0;
        while x(2)<-10^(-20) & s<100

[x]=fmincon(@(x)fminmulti(x),xk,[],[],[],[],[],[],@(x)multicon(x,xk));
for i=1:100
    if (f1(xk(1)+J(i)*x(1)))<=(f1(xk(1))+J(i)*sigma*x(2))
        && (f2(xk(1)+J(i)*x(1)))<=(f2(xk(1))+J(i)*sigma*x(2))
            t=J(i);break
        end
    end
end
x(2)
xk=[xk(1)+t*x(1);x(2)]
s=s+1
end
disp(xk(1))
s
*****
Quasi-Newton
*****
function f=fminmulti3(x)
f=x(4);
end
*****
function [c,ceq]=multicon3(x,xk,B1,B2)

gf1=@(x1,x2,x3)[2*x1 2*x2 2*x3];
gf2=@(x1,x2,x3)[0.6667*x1-1.3333 .6667*x2-1.3333 .6667*x3-1.3333 ];
c1=gf1(xk(1),xk(2),xk(3))*[x(1);x(2);x(3)]
+0.5*[x(1) x(2) x(3)]*B1*[x(1);x(2);x(3)]-x(4);
c2=gf2(xk(1),xk(2),xk(3))*[x(1);x(2);x(3)]

```

```

+0.5*[x(1) x(2) x(3)]*B2*[x(1);x(2);x(3)]-x(4);
c=[c1;c2];
ceq=[];
end

*****

clc
clear all
n=0:100;
J=1./2.^n;
sigma=0.1;

%xk=[1;0;1;1];
xk=rand(1,4)

sk=[1;1;1];
x(4)=-10^12;
f1=@(x1,x2,x3) x1.^2+x2.^2+x3.^2;
f2=@(x1,x2,x3) 0.3333*x1.^2+0.3333*x2.^2+0.3333*x3.^2-1.3333*x1-1.3333*x2-1.3333*x3;
gf1=@(x1,x2,x3) [2*x1 2*x2 2*x3];
gf2=@(x1,x2,x3) [0.6667*x1-1.3333 0.6667*x2-1.3333 0.6667*x3-1.3333 ];

s=0;
B1=eye(3);
B2=eye(3);

while ((x(4))<-10^-8 && s<100)
    x(4)
    [x]=fmincon(@(x)fminmulti3(x),xk,[],[],[],[],[],[],@(x)multicon3(x,xk,B1,B2));
    for i=1:100
        if (f1(xk(1)+J(i)*x(1),xk(2)+J(i)*x(2),xk(3)+J(i)*x(3)))
            <=(f1(xk(1),xk(2),xk(3))+J(i)*sigma*x(4))
            && (f2(xk(1)+J(i)*x(1),xk(2)+J(i)*x(2),xk(3)+J(i)*x(3)))
            <=(f2(xk(1),xk(2),xk(3))+J(i)*sigma*x(4))

```

```

        t=J(i);break
    end
end
%
yk=[xk(1);xk(2);xk(3);x(4)];
xk=[xk(1)+t*x(1);xk(2)+t*x(2);xk(3)+t*x(3);x(4)]
sk=[t*x(1);t*x(2);t*x(3)];

s=s+1
yy1=(gf1(xk(1),xk(2),xk(3))-gf1(yk(1),yk(2),yk(3)))
B1=B1-(B1*(sk)*(sk)'*B1)/((sk)'*B1*(sk))+((yy1)'*(yy1))/((sk)'*(yy1)')
gff1=[2 0 0;0 2 0;0 0 2];

yy2=(gf2(xk(1),xk(2),xk(3))-gf2(yk(1),yk(2),yk(3)))
B2=B2-(B2*(sk)*(sk)'*B2)/((sk)'*B2*(sk))+((yy2)'*(yy2))/((sk)'*(yy2)')
gff2=[0.6667 0 0;0 0.6667 0;0 0 0.6667];
end
disp('y=')
disp([xk(1) xk(2) xk(3)])
s
x(4)

```

مراجع

- [1] Ansari, Q. H., Yao, J. C. (2011). Recent developments in vector optimization. Springer Science and Business Media.
- [2] Ansary, M. A., Panda, G. (2014). A modified Quasi-Newton method for vector optimization problem. Optimization, 1-18.
- [3] Ball, J. M. (1980). Strict convexity, strong ellipticity, and regularity in the calculus of variations. In Mathematical Proceedings of the Cambridge Philosophical Society. Cambridge University Press, 87(3), 501-513.
- [4] Bazaraa, M. S., Sherali, H. D., Shetty, C. M. (2013). Nonlinear programming: theory and algorithms. John Wiley and Sons.
- [5] Bazaraa, M. S., Jarvis, J. J., Sherali, H. D. (2011). Linear programming and network flows. John Wiley and Sons.
- [6] Ben-Tal, A., Zowe, J. (1982). Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems. Mathematical Programming, 24(1), 70-91.
- [7] Bento, G. C., Ferreira, O. P., Oliveira, P. R. (2012). Unconstrained steepest descent method for multicriteria optimization on Riemannian manifolds. Journal of Optimization Theory and Applications, 154(1), 88-107.
- [8] Blasco, F., Cuchillo-Ibáñez, E., Morón, M. A., Romero, C. (1999). On the monotonicity of the compromise set in multicriteria problems. Journal of optimization theory and applications, 102(1), 69-82.
- [9] Broyden, C. G. (1965). A class of methods for solving nonlinear simultaneous equations. Mathematics of computation, 577-593.
- [10] Chamberlain, G., Rothschild, M. (1982). Arbitrage, factor structure, and mean-variance analysis on large asset markets.
- [11] Clarke, F. H. (1990). Optimization and nonsmooth analysis. Siam.

- [12] Dantzig, G. B., Thapa, M. N. (2006). Linear programming 2: theory and extensions. Springer Science and Business Media.
- [13] Das, I., Dennis, J. E. (1998). Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization*, 8(3), 631-657.
- [14] Dennis, J., Moré, J. (1977). Quasi-Newton methods, motivation and theory. *Siam*, 19(1), 46-89.
- [15] Ehrgott, M., Klamroth, K., Schwehm, C. (2004). An MCDM approach to portfolio optimization. *European Journal of Operational Research*, 155(3), 752-770.
- [16] Ehrgott, M. (2006). Multicriteria optimization. Springer Science and Business Media.
- [17] Fliege, J., Svaiter, B. F. (2000). Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, 51(3), 479-494.
- [18] Fliege, J., Drummond, L. G., Svaiter, B. F. (2009). Newton's method for multiobjective optimization. *Siam*, 20(2), 602-626.
- [19] Geoffrion, A. M. (1968). Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications*, 22(3), 618-630.
- [20] Gordan, P. (1873). Ueber die Auflösung linearer Gleichungen mit reellen Coefficienten. *Mathematische Annalen*, 6(1), 23-28.
- [21] Griva, I., Nash, S. G., Sofer, A. (2009). Linear and nonlinear optimization. *Siam*.
- [22] Jin, Y., Olhofer, M., Sendhoff, B. (2001). Dynamic weighted aggregation for evolutionary multi-objective optimization: Why does it work and how?.
- [23] Klerk, E., Roos, C., Terlaky, T. (2005). Optimization (WI4 207). Delft University of Technology.
- [24] Krahnke, A. (2001). The Clarke Derivative and Set-Valued Mappings in the Numerical Optimization of Non-Smooth, Noisy Functions.
- [25] Lewis, A. S., Overton, M. L. (2013). Nonsmooth optimization via quasi-Newton methods. *Mathematical Programming*, 141(2), 135-163.
- [26] Luenberger, D. G., Ye, Y. (2008). Linear and nonlinear programming. Springer Science and Business Media.
- [27] Miettinen, K. (2012). Nonlinear multiobjective optimization. Springer Science and Business Media.

- [28] Monjezi, N. H. (2015). Quasi-Newton methods for solving non-smooth multiobjective optimization. *Journal of Mathematical Extension*, 8, 9-21.
- [29] Nesterov, Y. (2005). Smooth minimization of non-smooth functions. *Mathematical programming*, 103(1), 127-152.
- [30] Nocedal, J., Wright, S. (2006). *Numerical optimization*. Springer Science and Business Media.
- [31] Povalej, Ž. (2014). Quasi-Newton's method for multiobjective optimization. *Journal of Computational and Applied Mathematics*, 255, 765-777.
- [32] Qi, L., Sun, J. (1993). A nonsmooth version of Newton's method. *Mathematical programming*, 58(3), 353-367.
- [33] Qu, S., Liu, C., Goh, M., Li, Y., Ji, Y. (2014). Nonsmooth multiobjective programming with quasi-Newton methods. *European Journal of Operational Research*, 235(3), 503-510.
- [34] Qu, S., Goh, M., Chan, F. T. (2011). Quasi-Newton methods for solving multiobjective optimization. *Operations Research Letters*, 39(5), 397-399.
- [35] Schutze, O., Talbi, E. G., Coello, C. C., Santana-Quintero, L. V., Pulido, G. T. (2007). A memetic PSO algorithm for scalar optimization problems. In *Swarm Intelligence Symposium*, 128-134.
- [36] Steuer, R. E., Na, P. (2003). Multiple criteria decision making combined with finance: A categorized bibliographic study. *European Journal of operational research*, 150(3), 496-515.
- [37] Tammer, C., Göpfert, A. (2002). Theory of vector optimization. In *Multiple Criteria Optimization: State of the Art Annotated Bibliographic Surveys*. Springer US, 1-70.
- [38] Torokhti, A., Howlett, P. (1985). *Theory of multiobjective optimization*. Elsevier.
- [39] Vial, J. P. (1982). Strong convexity of sets and functions. *Journal of Mathematical Economics*, 9(1), 187-205.
- [40] Vieira, D. A., Takahashi, R. H., Saldanha, R. R. (2012). Multicriteria optimization with a multiobjective golden section line search. *Mathematical programming*, 131(1), 131-161.
- [41] Wei, Z., Li, G., Qi, L. (2006). New quasi-Newton methods for unconstrained optimization problems. *Applied Mathematics and Computation*, 175(2), 1156-1188.
- [42] Wojtaszczyk, P. (1996). *Banach spaces for analysts*. Cambridge University Press.

-
- [43] Zitzler, E. (1999). Evolutionary algorithms for multiobjective optimization: Methods and applications (Vol. 63). Ithaca: Shaker.

واژه‌نامه فارسی به انگلیسی

Modified	اصلاح شده
Algorithm	الگوریتم
Primal	اولیه
Range	برد
Optimal	بهینه
Pareto optimal	بهینه پارتو
Biobjective optimization	بهینه‌سازی دوهدفه
Convex hull	پوسته محدب
Open cover	پوشش باز
Lipschitz continuous	پیوسته لیپ‌شیتز
Equicontinuous	پیوسته یکسان
Uniformly continuous	پیوسته یکنواخت
Lagrangian function	تابع لاگرانژین
Convex function	تابع محدب
Approximation	تقریب
Single objective	تک‌هدفه
Optimal solution	جواب بهینه
Efficient solution	جواب کارا
Weakly efficient solution	جواب کارای ضعیف
Descent direction	جهت کاهنده
Multiobjective	چندهدفه
Approximation error	خطای تقریب
Linear	خطی
Quadratic	درجه دوم
Interior	درون

Cauchy sequence	دنباله کوچی
Dual	دوگان
Lagrange dual	دوگان لاگرانژ
Subcover	زیرپوشش
Jacobian	ژاکوبین
Quasi-newton	شبه نیوتن
Dual feasibility	شدنی دوگان
Curvature condition	شرط انحنا
Primal feasibility	شرط اولیه شدنی
Duality gap	شکاف دوگان
Lagrangian multiplier	ضریب لاگرانژین
Vector space	فضای برداری
Image space	فضای تصویر
Feasible space	فضای شدنی
Normed space	فضای نرم‌دار
Objective space	فضای هدف
Superlinear	فوق خطی
Efficient	کارا
Weakly efficient	کارای ضعیف
Bounded	کراندار
Uniformly bounded	کراندار یکنواخت
Iteration steps	گام‌های تکرار
Gradient	گرادیان
Lagrange	لاگرانژ
Maximization	ماکسیم‌سازی
Strictly positive	مثبت اکید
Positive definite	مثبت معین
Compact Set	مجموعه فشرده
Convex set	مجموعه محدب
Convex	محدب
Strictly convex	محدب اکید
Strong convex	محدب قوی

Order of convergence.....	مرتبه همگرایی.....
Multiobjective optimization problem.....	مساله بهینه‌سازی چندهدفه.....
Convex optimization problem.....	مساله بهینه‌سازی محدب.....
Secant equation.....	معادله وتری.....
Value.....	مقدار.....
Complementary slackness.....	مکمل زائد.....
Slater regular.....	منظم اسلاتر.....
Minimization.....	مینیم‌سازی.....
Non-vacuous.....	ناتهی.....
Cauchy-Schwartz inequality.....	نامساوی کوشی شوارتز.....
Inequality.....	نامعادله.....
Non-negative.....	نامنفی.....
Nonsmooth.....	ناهموار.....
Euclidean norm.....	نرم اقلیدسی.....
Slater point.....	نقطه اسلاتر.....
Newton.....	نیوتن.....
Hessian.....	هسین.....

واژه‌نامه انگلیسی به فارسی

Algorithm	الگوریتم
Approximation	تقریب
Approximation error	خطای تقریب
Biobjective optimization	بهینه‌سازی دوهدفه
Bounded	کراندار
Cauchy-Schwartz inequality	نامساوی کوشی شوارتز
Cauchy sequence	دنباله کوشی
Compact Set	مجموعه فشرده
Complementary slackness	مکمل زائد
Convex	محدب
Convex function	تابع محدب
Convex optimization problem	مساله بهینه‌سازی محدب
Convex hull	پوسته محدب
Convex set	مجموعه محدب
Curvature condition	شرط انحنا
Descent direction	جهت کاهنده
Dual	دوگان
Dual feasibility	شدنی دوگان
Duality gap	شکاف دوگان
Efficient	کارا
Efficient solution	جواب کارا
Equicontinuous	پیوسته یکسان
Euclidean norm	نرم اقلیدسی
Feasible space	فضای شدنی
Gradient	گرادیان

Hessian	هسین
Image space	فضای تصویر
Inequality	نامعادله
Iteration steps	گام‌های تکرار
Interior	درون
Jacobian	ژاکوبین
Lagrange	لاگرانژ
Lagrange dual	دوگان لاگرانژ
Lagrangian function	تابع لاگرانژین
Lagrangian multiplier	ضریب لاگرانژین
Linear	خطی
Lipschitz continuous	پیوسته لیپ‌شیتز
Maximization	ماکسیم‌سازی
Minimization	مینیم‌سازی
Modified	اصلاح شده
Multiobjective	چندهدفه
Multiobjective optimization problem	مساله بهینه‌سازی چندهدفه
Newton	نیوتن
Nonsmooth	ناهموار
Non-negative	نامنفی
Non-vacuous	ناتهی
Normed space	فضای نرم‌دار
Objective space	فضای هدف
Open cover	پوشش باز
Optimal	بهینه
Optimal solution	جواب بهینه
Order of convergence	مرتبه همگرایی
Pareto optimal	بهینه پارتو
Positive definite	مثبت معین
Primal	اولیه
Primal feasibility	شرط اولیه شدنی
Quadratic	درجه دوم

Quasi-newton	شبه‌نیوتن
Range	برد
Secant equation	معادله وتری
Single objective	تک‌هدفه
Slater point	نقطه اسلاتر
Slater regular	منظم اسلاتر
Strictly convex	محدب اکید
Strictly positive	مثبت اکید
Strong convex	محدب قوی
Subcover	زیرپوشش
Superlinear	فوق خطی
Uniformly bounded	کراندار یکنواخت
Uniformly continuous	پیوسته یکنواخت
Value	مقدار
Vector space	فضای برداری
Weakly efficient	کارای ضعیف
Weakly efficient solution	جواب کارای ضعیف

Aabstract

In this thesis, we first introduce the basic concepts and preliminaries. Then, we express the concepts of the efficiency and weak efficiency for multiobjective optimization problem. We deal to solve multiobjective optimization problem by applying iterative methods. An iterative method is a computational process that generates a sequence and we expect that under reasonable conditions converges to the efficient solution of the problem. From the iterative methods for multiobjective optimization problem, we discuss, the Newton method and the quasi-Newton methods.

Following, we express the Newton algorithm for solving convex multiobjective optimization problem that requires precise calculation of the Hessian matrix for each objective function. In the case that calculation of the Hessian be difficult, quasi-Newton methods are introduced and are discussed. These methods calculate the Hessian approximate instead of the exact Hessian and help us to improve the solving process. In cases where the problem is nonconvex, a modified quasi-Newton algorithm is suggested. In this case, by calculating an approximate matrix, nonconvex multiobjective problems are solved. Finally, a quasi-Newton method for solving nonsmooth multiobjective problem is introduced.

keywords: Multiobjective optimization, Efficient solution, Weakly efficient solution, Critical point, Convex function.



Shahrood University of Technology

Faculty Of Mathematical Sciences

Some Iterative Methods For Multiobjective Optimization Problems

Narges Hoseinpoor

Supervisor

Dr. Mehrdad Ghaznavi

Advisor

Dr. Maryam Ghorani

December 2015