دانشگاه پیام نور
دانشکده علوم پایه و کشاورزی
مرکز تهران شرق

پایان نامه
برای دریافت مدرک کارشناسی ارشد
رشته ریاضی کاربردی (تحقیق در عملیات)
گروه ریاضی

عنوان پایان نامه :

یک روش گرادیان مزدوج پیش حالت ساز با BFGS مقیاس بندی شده برای بهینه سازی نامقید

اشرف بهرامی کرجی

استاد راهنما :
دکتر محمدرضا پیغامی
استاد مشاور :
دکتر فهیمه سلطانیان

اردیبهشت 1391

بسم الله الرحمن الرحیم

صورتجلسه دفاع از پایان نامه کارشناسی ارشد

جلسه دفاع ازپایان نامه کارشناسی ارشد خانم اشرف بهرامی کرجی
دانشجوی رشته ریاضی کاربردی(تحقیق در عملیات) به شماره دانشجویی: ۸۸۰۲۷۰۷۹۲
تحت عنوان:

يك روش گراديان مزدوج پيش حالت ساز با BFGS مقياس بندي شده براي بهينه سازي نامقيد

جلسه دفاع با حضور داوران نامبرده ذیل در روز شنبه مورخ: ۹۱/۰۲/۳۰ساعت : ۱۶-۱۵درمحل تهران شرق برگزار شد. و پس از بررسی پایان‌نامه مذکور با نمره به عدد.....۱۸٫۷۵ به حروف ...هیجده و هفتاد و پنج صدم..... و با درجه ارزشیابی.........عالی.........مورد قبول واقع شد ☑ نشد ☐

ردیف	نام و نام خانوادگی	هیات داوران	مرتبه دانشگاهي	دانشگاه/ موسسه	امضاء
۱	دكترمحمدرضا پيغامي	استاد راهنما	دانشیار	دانشگاه خواجه نصيرالدين طوسي	
۲	دكتر فهيمه سلطانيان	استاد مشاور	استادیار	پيام نور	
۳	دكترنظام الدين مهدوي اميري	استاد داور	استاد	دانشگاه صنعتي شريف	
۴	دكترخديجه احمدي آملي	نماينده علمي گروه / نماينده تحصيلات تكميلي	استادیار	پيام نور	

تهران، خیابان استاد نجات اللهی
خیابان شهید فلاح پور، پلاک ۲۷
تلفن: ۸۸۸۰۰۲۵۲
دورنگار: ۸۸۳۱۹۴۷۵
WWW.TPNU.AC.IR
science.agri@tpnu.ac.ir

اینجانب اشرف بهرامی کرجی دانشجوی دانشگاه پیام نور، ورودی سال 1388 مقطع کارشناسی ارشد رشته ریاضی گرایش تحقیق در عملیات، گواهی می نمایم چنانچه در پایان نامه خود از فکر، ایده و نوشته دیگری بهره گرفته ام با نقل قول مستقیم یا غیرمستقیم منبع و ماخذ آن را نیز در جای مناسب ذکر کرده ام . بدیهی است مسئولیت تمامی مطالبی که نقل قول دیگران نباشد بر عهده خویش می دانم و جوابگوی آن خواهم بود.

دانشجو تائید می نماید که مطالب مندرج در این پایان نامه (رساله) نتیجه تحقیقات خودش می باشد و در صورت استفاده از نتایج دیگران مرجع آن را ذکر نموده است.

نام و نام خانوادگی دانشجو : اشرف بهرامی کرجی

تاریخ و امضاء

اینجانب اشرف بهرامی کرجی دانشجوی دانشگاه پیام نور، ورودی سال 1388 مقطع کارشناسی ارشد رشته ریاضی گرایش تحقیق در عملیات، گواهی می نمایم چنانچه بر اساس مطالب پایان نامه خود اقدام به انتشار مقاله، کتاب و ... نمایم ضمن مطلع نمودن استاد راهنما، ب نظر ایشان نسبت به نشر مقاله، کتاب و ... و به صورت مشترک و با ذکر نام استاد راهنما مبادرت نمایم.

نام و نام خانوادگی دانشجو : اشرف بهرامی کرجی

تاریخ و امضاء

کلیه حقوق مادی مترتب از نتایج مطالعات، آزمایشات و نوآوری ناشی از تحقیق موضوع این پایان نامه متعلق به دانشگاه پیام نور می باشد.

اردیبهشت 1391

تقدیم به دو گوهر گرانبهای زندگی ام

پدر و مادر عزیزم

که محکمترین و همیشگی ترین پشتیبان های زندگی ام هستند.

تقدیم به همسرعزیزم

که با مساعدتهای خود در تمامی مراحل این تحقیق نقش بسزایی داشته و تمام این ایام مرا حمایت کرده و همراهم بوده است.

تقدیم به فرزند دلبندم کامیار

که در ایام تحصیل و تکمیل پایان نامه سختیهای فراوانی را متحمل شده و با حضور شادی بخشش انگیزه ام را برای طی این مسیر مضاعف نمود.

تشکر و قدردانی

حمد و سپاس بر خداوند متعال به جای می آورم که مرا یاری نمود تا یکی دیگر از مراحل تحصیل و پیشرفت خود را به پایان برسانم .

برخود لازم می‌دانم مراتب سپاس و قدردانی خود را از استاد گرامی جناب آقای دکتر محمدرضا پیغامی به پاس راهنمایی‌های ارزشمند و گرانبهایشان در مسیر تدوین این پژوهش ابراز نمایم.

تشکر و قدردانی خود را به استاد گرانقدر جناب آقای دکتر نظام الدین مهدوی امیری که بر من منت نهاده و زحمت داوری این پایان نامه را تقبل کردند تقدیم می دارم.

از اساتید بزرگوارخانم دکتر سلطانیان استاد مشاور و خانم دکتر احمدی نماینده گروه کمال تشکر و قدردانی را دارم و همچنین از آقای مهندس محسن سالاررضایی به جهت کمکهای بیدریغ ایشان در پیشبرد این پروژه خالصانه تشکر می نمایم.

چکیده

روش‌های گرادیان مزدوج، روش‌های تکراری برای حل مسایل بهینه‌سازی نامقید هستند و از آنجا که در تکرارهای این روش ها نیازی به محاسبه و ذخیره سازی ماتریس هسی نیست، در مقیاس بزرگ کاربرد دارند. در این پایان نامه، ما به بررسی دو الگوریتم گرادیان مزدوج ترکیبی که اخیراً در ادبیات موضوع مطرح شده‌اند؛ می‌پردازیم. در روش ترکیبی اول، جهت متناظر با روش گرادیان مزدوج از فرمولی متمایز از فرمول‌های روش های موجود قبلی بدست می‌آید. در این روش با وارد شدن پارامتری در جهت مربوط تاثیر بردار گرادیان در جهت بدست آمده، تغییر می یابد. روش ترکیبی دوم، در واقع توسعه‌ای از روش دای-یوآن[1] است، که با اصلاحات انجام شده که شرایطی به نام شرایط کاهش کافی را ،که اخیراً در ادبیات موضوع مطرح شده است، دارا است.

برای بررسی میزان کارایی الگوریتم‌های ارایه شده، این الگوریتم‌ها همراه با چند الگوریتم موجود دیگر در محیط برنامه‌نویسی ++C پیاده سازی و با اجرای برنامه ها و مقایسه نتایج محاسباتی بدست آمده کارایی الگوریتم‌ها بررسی می شوند.

کلمات کلیدی:

مینیمم‌سازی نامقید، روش گرادیان مزدوج ترکیبی، جستجوی خطی، شرایط کاهش کافی.

[1] Dai and Yuan(DY)

فهرست مطالب

فهرست شکل‌ها

فهرست جدول‌ها

پیشگفتار

امروزه، بهینه‌سازی کاربردهای گسترده ای در علوم پایه، مهندسی، اقتصاد، مدیریت و غیره دارد. در حالت کلی، یک فرایند بهینه‌سازی به یافتن بهترین جواب از میان همه جواب‌های شدنی یک مساله می-پردازد. بنابراین، در بهینه‌سازی باید شرایط بهینگی یک جواب، روش‌های عددی برای محاسبه جواب بهینه، تحلیل همگرایی و کارایی عددی روش‌ها بحث و بررسی شوند. یکی از مبانی بهینه سازی، بهینه-سازی نامقید، یعنی حل مساله

$$\min f(x)$$
$$x \in \mathbb{R}^n$$

است که در آن، f تابع هدف نامیده می‌شود. این جا ،تابع هدف، هموار مشتق‌پذیر (به دفعات مورد نیاز) فرض می‌شود.

یکی از اساسی‌ترین روش‌ها برای حل مسایل بهینه‌سازی نامقید، روش تندترین کاهش است (که اغلب روش گرادیان نامیده می‌شود) که در سال 1847 توسط کوشی[2] ارایه شد. همگرایی کند روش تندترین کاهش، پژوهشگران را به اصلاح این روش واداشت. در پی این تلاش‌ها، روش‌های شبه نیوتون و روش-های گرادیان مزدوج ارایه شدند. در این میان، روش‌های گرادیان مزدوج به دلیل استفاده کم از حافظه ماشین و همگرایی سراسری موضعی و سراسری مناسب، به روش‌هایی ایده‌آل برای حل مسایل بهینه-سازی نامقید در مقیاس بزرگ تبدیل شده‌اند. روش‌های گرادیان مزدوج، اولین بار توسط هستنس و استیفل[3] در دهه 1950 [20] برای حل دستگاه‌های خطی به کار گرفته شدند. با بهره‌گیری از این حقیقت

---

[2] Cauchy
[3] Hestenes and Stiefel

که حل یک دستگاه خطی معادل است با حل یک مساله بهینه‌سازی نامقید با تابع هدف درجه دوم، فلچر و ریوز[4] [16] در دهه 1960 این روش را برای حل مسایل بهینه‌سازی نامقید گسترش دادند. پس از آن، روش‌های گرادیان مزدوج مختلفی ارایه شدند که هر یک دارای نقاط ضعف و قوت خاص خود هستند. تحلیل رفتار عددی شمار زیادی از روش‌های گرادیان مزدوج در مقاله هگر و ژانگ[5] [18] آمده است.

در این تحقیق دو الگوریتم گرادیان مزدوج جدید که اخیراً توسط آندری[6] در مراجع [2] و [5] معرفی شده‌اند، بررسی می‌شوند. تمرکز اصلی کار ما روی مسایل بهینه‌سازی غیرخطی نامقید و الگوریتم‌های حل آن و در نهایت معرفی و تشریح دو الگوریتم یاد شده است. برای این منظور، مراحل زیر در نظر گرفته می شوند:

1. مرور ادبیات مسایل بهینه‌سازی نامقید.
2. تشریح کامل مساله بهینه سازی نامقید و بررسی انواع روش‌های حل.
3. معرفی دو الگوریتم جدید ارایه شده توسط آندری برای حل مساله و تشریح آن‌ها.
4. پیاده سازی الگوریتم‌های جدید و مقایسه نتایج محاسباتی آن‌ها با روش‌های موجود.

این پایان نامه شامل چهار فصل است و به شکل زیر سامان دهی شده است:

درفصل اول پس از بیان مفاهیم و تعاریف مقدماتی، انواع روش‌های حل مساله بهینه‌سازی نامقید معرفی می‌شوند. الگوریتم‌های کاهشی برای حل این مسایل معرفی و تشریح می‌شوند. الگوریتم گرادیان و

[4] Fletcher and Reeves
[5] Hager and Zhang
[6] Andrei

نیوتون و هم چنین الگوریتم‌های شبه‌نیوتون که از اصلی ترین رده الگوریتم‌های کاهشی هستند در این فصل تشریح خواهند شد. سرانجام در پایان این فصل، مروری بر شرایط جستجوی خطی،که در همه الگوریتم‌های بهینه‌سازی نامقید نیاز است، پرداخته می‌شود.

درفصل دوم، ابتدا برخی از روش‌های گرادیان مزدوج مطرح شده معرفی می‌شوند و سپس ترکیب روش-های گرادیان مزدوج با روش‌های شبه نیوتون می آید.

فصل سوم به دو الگوریتم گرادیان مزدوج جدید ارایه شده توسط آندری می پردازد و در خصوص همگرایی آنها بحث می‌کند.

فصل چهارم به نتایج بدست آمده از اجرای پیاده سازی ها و مقایسه الگوریتم‌های مختلف می پردازد.

فصل1 تعاریف و مفاهیم اولیه

در این فصل، با ذکر برخی از تعاریف و مفاهیم اولیه بهینه‌سازی، روش جهت‌های مزدوج برای حل مساله بهینه‌سازی درجه دوم را که اساس روش‌های گرادیان مزدوج است، بیان و تحلیل می‌کنیم. سپس، روش گرادیان مزدوج خطی و غیر خطی را مورد بررسی قرار می‌دهیم. سرانجام، در بخش پایانی فصل نیز برخی از مفاهیم مربوط به جستجوی خطی ارایه می‌شوند.

1-1 تعاریف اولیه

تعریف 1-1. تابع $f:\mathbb{R}^n \to \mathbb{R}$ را در نقطه x پیوسته گویند، اگر برای هر $\varepsilon > 0$ داده شده، $\delta > 0$ وجود داشته باشد به طوری که برای هر y با $|x-y| < \delta$، داشته باشیم:

$$|f(x)-f(y)| < \epsilon.$$

اگر f روی همه نقاط $x\epsilon\mathbb{R}^n$ پیوسته باشد، آنگاه f را روی $\mathbb{R}^n$ پیوسته گویند.

تعریف 1-2. (تابع پیوسته لیپ شیتس[7]) پیوستگی لیپ شیتس تابع $f:\mathbb{R}^n \to \mathbb{R}$ به صورت زیر تعریف می‌شود:

a) فرض کنید $B \subseteq \mathbb{R}^n$ مجموعه بازی باشد. گوییم f دارای پیوستگی لیپ شیتس روی مجموعه B است هرگاه عدد ثابت $\Lambda \in \mathbb{R}^+$، معروف به ثابت لیپ شیتس، وجود داشته باشد به طوری که:

$$\|f(x)-f(y)\| \le \Lambda\|x-y\|, \qquad \forall x,y \in B$$

[7] Lipschitz-continuous function

b) تابع f پیوسته لیپ‌شیتس موضعی است، هرگاه برای هر $z \in \mathbb{R}^n$ ، $L > 0$ وجود داشته باشد به طوری که f روی گوی باز به مرکز z و شعاع L، یعنی مجموعه $B_L(z) = \{y \in \mathbb{R}^n : \|y - z\| < L\}$، پیوسته لیپ شیتس باشد.

c) تابع f پیوسته لیپ شیتس سراسری است، هرگاه f روی $B = \mathbb{R}^n$ پیوسته لیپ شیتس باشد.

تعریف 3-1. تابع $f: \mathbb{R}^n \to \mathbb{R}$ را در نقطه x مشتق‌پذیر گویند، هرگاه یک تابع $g: \mathbb{R}^n \to \mathbb{R}$ وجود داشته باشد به طوری که برای هر $0 \neq d \in \mathbb{R}^n$، داشته باشیم:

$$\lim_{\tau \to 0} \frac{f(x + \tau d) - f(x)}{\tau} = g^T(x) d$$

برای بردار یکه $d = e_j$، اگر حد

$$\lim_{\tau \to 0} \frac{f\left(x + \tau e_j\right) - f(x)}{\tau}$$

وجود داشته باشد، آنگاه آن را مشتق پاره‌ای f نسبت به x_j در نقطه x گویند و با نماد $\frac{\partial f(x)}{\partial x_j}$ نشان می‌دهند.

اگر مشتق پاره‌ای f در نقطه x به ازای هر j موجود باشد، آنگاه گرادیان تابع f موجود است و به صورت زیر نمایش داده می‌شود:

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_j}, \dots, \frac{\partial f(x)}{\partial x_n},\right)^T.$$

تابع f را در ناحیه D مشتق‌پذیر گویند هرگاه برای هر $x \in D$، $\nabla f(x)$ موجود باشد و آن را مشتق‌پذیر پیوسته و یا به طور پیوسته مشتق‌پذیر گویند، هرگاه $\nabla f(x)$ تابعی پیوسته از x باشد.

تعریف 1-4. ماتریس هسی تابع $f: \mathbb{R}^n \to \mathbb{R}$، در صورت وجود، به صورت زیر تعریف می‌شود:

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

تابع f را دوبار مشتق‌پذیر روی D گویند، هرگاه ماتریس $\nabla^2 f(x)$ برای هر $x \in D$ موجود باشد و آن را دوبار مشتق‌پذیر پیوسته گویند، هرگاه $\nabla^2 f(x)$ تابعی پیوسته روی D باشد.

نمادگذاری 1-1: در این پایان نامه به جای $\nabla f(x_k)$ از نماد g_k و به جای $\nabla^2 f(x_k)$ از نماد G_k استفاده می‌کنیم.

تعریف 1-5. زیرمجموعه S از فضای $\mathbb{R}^n$ را محدب گویند، هرگاه

$$\forall x, y \in S, 0 \leq \theta \leq 1;\ \theta x + (1-\theta) y \epsilon S.$$

تعریف 1-6. فرض کنید $f: \mathbb{R}^n \to \mathbb{R}$ و $S \subseteq \mathbb{R}^n$ یک مجموعه محدب باشد؛ تابع f را روی S محدب گویند، هرگاه:

$$\forall x, y \in S, 0 \leq \theta \leq 1;\ f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y).$$

در تعریف 1-6 اگر نامساوی به ازای $\theta \epsilon (0,1)$ و $x \neq y$ اکید باشد، آنگاه f را اکیداً محدب گویند.

گزاره زیر شرط معادلی را برای محدب بودن یک تابع بیان می‌کند. اثباتی از این گزاره را می‌توان در (Nocedal and Wright, 1999) یافت.

گزاره 1-1. تابع به طور پیوسته مشتق‌پذیر f روی مجموعه S محدب است اگر و تنها اگر برای هر $x, y \in S$ داشته باشیم:

$$f(y) \geq f(x) + \nabla^T f(x)(y - x)$$

یا به طور معادل:

$$\left(\nabla f(y) - \nabla f(x)\right)^T (y - x) \geq 0.$$

تعریف 1-7. ماتریس مربعی و متقارن $H_{n\times n}$ را نیمه معین مثبت گویند، هرگاه برای هر $x \in \mathbb{R}^n$ داشته باشیم:

$$x^T H x \geq 0.$$

تعریف 1-8. ماتریس مربعی و متقارن $H_{n\times n}$ را معین مثبت گویند، هرگاه برای هر $0 \neq x \in \mathbb{R}^n$ داشته باشیم:

$$x^T H x > 0.$$

تعریف 1-9. نقطه x^* را یک مینیمم کننده سراسری $f: \mathbb{R}^n \to \mathbb{R}$ گویند، اگر برای هر $x \in D_f$ داشته باشیم:

$$f(x) \geq f(x^*).$$