بسم الله الرحمن الرحیم

دانشگاه مراغه

دانشکده علوم پایه

گروه ریاضی

پایان‌نامه:

برای دریافت درجه‌ی کارشناسی ارشد در رشته‌ی ریاضی، گرایش آنالیز عددي

عنوان:

گرادیان مزدوج مخلوط اصلاح شده

استاد راهنما:

دکتر علی شکری

استاد مشاور:

دکتر سهراب بزم

پژوهشگر:

ابراهیم اسماعیل پور

شهریور ۱۳۹۲

تقدیم به

محضر خداوند متعال

خدایا...

به من زیستنی عطا کن که در لحظه مرگ، بر بی‌ثمری لحظه‌ای که برای زیستن گذشته است، حسرت نخورم و مُردنی عطا کن که بر بیهودگیش، سوگوار نباشم. بگذار تا آن را، خود انتخاب کنم، اما آنچنان که تو دوست می‌داری.

تو می‌دانی و همه می‌دانند که شکنجه دیدن بخاطر تو، زندانی کشیدن بخاطر تو و رنج بردن به پای تو تنها لذت بزرگ زندگی من است، از شادی توست که من در دل می‌خندم، از امید رهایی توست که برق امید در چشمان خسته‌ام می‌درخشد و از خوشبختی توست که هوای پاک سعادت را در ریه‌هایم احساس می‌کنم. نمی‌توانم خوب حرف بزنم. نیروی شگفتی را که در زیر کلمات ساده و جمله‌های ضعیف و افتاده، پنهان کرده‌ام دریاب، دریاب.

تو می‌دانی و همه می‌دانند که زندگی از تحمیل لبخندی بر لبان من، از آوردن برق امیدی در نگاه من، از برانگیختن موج شعفی در دل من، عاجز است.

تو، چگونه زیستن را به من بیاموز، چگونه مردن را خود خواهم آموخت.

به من توفیق تلاش در شکست، صبر در نومیدی، رفتن بی‌همراه، جهاد بی‌سلاح، کار بی‌پاداش، فداکاری در سکوت، دین بی‌دنیا، مذهب بی‌عوام، عظمت بی‌نام، خدمت بی‌نان، ایمان بی‌ریا، خوبی بی‌نمود، گستاخی بی‌خامی، قناعت بی‌غرور، عشق بی‌هوس، تنهایی در انبوه جمعیت، و دوست داشتن بی‌آنکه دوست بداند، روزی کن.

اگر تنهاترین تنها شوم، باز خدا هست

او جانشین همه نداشتن‌هاست...

سپاس‌گزاری...

سپاس خداوندگار حکیم را که با لطف بی‌کران خود، آدمی را زیور عقل آراست.

در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ استاد راهنمای خود، جناب آقای دکتر علي شكري، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید.

از جناب آقای دکتر سهراب بزم که زحمت مطالعه و مشاوره این رساله را تقبل فرمودند و در آماده سازی این رساله، به نحو احسن اینجانب را مورد راهنمایی قرار دادند، کمال امتنان را دارم. از جناب آقای دکتر علي خاني که زحمت داوری این رساله را تقبل فرمودند تشکر می‌کنم.

و در پایان، بوسه می‌زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می‌کنم وجود مقدس‌شان را، و تشکر می‌کنم که زحمات زیادی را متحمل شدند.

ابراهیم اسماعیل پور
شهریور ۱۳۹۲

نام خانوادگی: اسماعیل پور نام: ابراهیم

عنوان پایان‌نامه: گرادیان مزدوج مخلوط اصلاح شده

استاد راهنما: دکتر علی شکری استاد مشاور: دکترسهراب بزم

مقطع تحصیلی: کارشناسی ارشد
رشته: ریاضی کاربردی گرایش: آنالیز عددی

دانشگاه: مراغه دانشکده: علوم پایه
تاریخ فارغ‌التحصیلی: مهر ۱۳۹۲ تعداد صفحه: ۸۳

کلیدواژه‌ها: گرادیان مزدوج، روشهاي تکراری، جستجوي خطی، همگرایی سراسری، برنامه ریزی غیرخطی.

چکیده

روش گرادیان مزدوج یک روش مهم و کاربردی در حل مسائل بهینه سازی نامقید بویژه با مقیاس بزرگ می‌باشد، بر اساس روش گرادیان مزدوج مخلوط پیشنهاد شده توسط، جینگ و دنگ، ما روش مذکور را اصلاح کرده و همگرایی آن را با استفاده از شرط کافی برای توابع نزولی ثابت می کنیم و سپس با انجام بعضی آزمون بر روی توابع غیر خطی مختلف برتری روش اصلاح شده مخلوط را نسبت به روش فلچر - ریوز ، از نظر زمان اجرا، نشان می دهیم.

فهرست مطالب

مقدمه

روش گرادیان مزدوج یک روش تکراری برای حل مسائل بهینه سازی می باشد .این روش به دو دسته خطی و غیرخطی تقسیم می شود.روش خطی بر روی دستگاه $Qx = b$ که در آن Q یک ماتریس متقارن معین مثبت است به کار می رود.از این روش برای حل مسایل برنامه غیرخطی نامقید استفاده می شود. که در این صورت ماتریسQ همان ماتریس هسی تابع درجه دو می باشد. ولی روش غیرخطی که می توان بوسیله آن توابع درجه دو و یک را نیز حل کرد خود به دو دسته تقسیم می شود که عبارتند از اولی تقریب درجه دو که تقریبی از روش خطی می باشد ولی دومی مبتنی بر جستجوی خطی میباشد.

روش گرادیان مزدوج غیرخطی مبتنی بر جستجوی خطی پر کاربردترین و مشهورترین روش می‌باشد که ریاضیدانان و متخصصین برنامه‌ریزی غیرخطی با استفاده از نو آوری و خلاقیت خود روش‌های بهینه‌تری از آن پیشنهاد کرده‌اند که همگرایی را افزایش و زمان محاسبات را کمتر می کند یکی از این روش‌ها روش مخلوط اصلاح شده می باشد که در این رساله مورد بحث قرار خواهد گرفت.

برای آشنایی بیشتر با مبحث بهینه‌سازی،ما در فصل اول تعاریف مقدماتی را ارائه می دهیم و در فصل دو جهت های مزدوج را معرفی کرده و در فصل سه روش گرادیان مزدوج را معرفی کرده و در باره همگرایی آن بحث خواهیم کرد و در فصل آخر روش گرادیان مزدوج مخلوط اصلاح شده معرفی شده و پس از بحث در مورد

همگرایی آن نتایج عددی برای نشان دادن کارایی روش ارائه خواهد شد.

فصل ۱

مفاهیم و مباحث مقدماتی

۱.۱ تعاریف و مفاهیم اولیه بهینه سازي

تعریف ۱.۱.۱. تابع $f : \mathbb{R}^n \to \mathbb{R}$ را در نقطه x مشتق پذیر گویند ، هر گاه تابع $g : \mathbb{R}^n \to \mathbb{R}$ موجود باشد به

طوری که برای هر $\circ \neq d \in \mathbb{R}^n$ داشته باشیم

$$\lim_{\alpha \to o} \frac{f(x+\alpha d)-f(x)}{\alpha} = g^T(x)d.$$

به ازای هر بردار یکه $d = e_j$ اگر حد

$$\lim_{\alpha \to o} \frac{f(x+\alpha e_j)-f(x)}{\alpha},$$

موجود باشد آنگاه مقدار فوق را مشتق جزئیf نسبت به x_j گویند و بانماد $\frac{\partial f(x)}{\partial x_j}$ نشان می‌دهند .

اگر مشتق جزئی به ازای هر jموحود باشد آنگاه گرادیان تابع f موجود است و به صورت زیر نشان داده می‌شود:

$$\bigtriangledown f(x) = \left(\frac{\partial f(x)}{\partial x_۱}, \frac{\partial f(x)}{\partial x_۲}, ..., \frac{\partial f(x)}{\partial x_n} \right)^T,$$

و با مشتق گیری جزئی و برداری از $\bigtriangledown f(x)$ ماتریس هسیان به صورت زیر تعریف می‌شود

$$\bigtriangledown^۲ f(x) = \left(\frac{\partial^۲ f(x)}{\partial x_i \partial x_j} \right)_{n\times n},$$

تعریف ۲.۱.۱. زیر مجموعه‌ی S از فضای $\mathbb{R}^n$ را محدب گویند ،اگر

$$\forall x, y \in S, ۰ \leqslant a \leqslant ۱ \Rightarrow ax + (۱ - a)y \in S.$$

تعریف ۳.۱.۱. فرض کنید $f : \mathbb{R}^n \to \mathbb{R}$ و $S \subset \mathbb{R}^n$ یک مجموعه محدب باشد . تابع f را روی S محدب گویند،اگر داشته باشیم

$$\forall x, y \in S, o \leqslant a \leqslant ۱ \Rightarrow f(ax + (۱ - a)y) \leqslant af(x) + (۱ - a)f(y).$$

تعریف ۴.۱.۱. ماتریس n و متقارن Q را نیمه معین مثبت گویند،هرگاه برای هر $x \in \mathbb{R}^n$ داشته باشیم

$$x^T Q x \geqslant ۰,$$

و همچنین معین مثبت گویند،هرگاه برای هر $۰ \neq x \in R^n$ داشته باشیم

$$x^T Q x > ۰,$$

تعریف ۵.۱.۱. نقطه $x^* \in \Omega$ را نقطه‌ی مینیمم نسبی یا نقطه مینیمم موضعی تابع f روی Ω گویند، اگر $\varepsilon > ۰$ وجود داشته باشد به طوری که به ازای هر $x \in \Omega$ در فاصله حداکثر ε از x^* (یعنی $x \in \Omega$ و $|x - x^*| < \varepsilon$) داشته باشیم$f(x) \geqslant f(x^*)$. اگر به ازای هر $x \in \Omega$، آنگاه x^* را نقطه مینیمم نسبی اکید f روی Ω گویند.

تعریف ۶.۱.۱. نقطه $x^* \in \Omega$ را نقطه‌ی مینیمم سراسری تابع f روی Ω گویند اگر به ازای هر $x \in \Omega$ داشته باشیم $f(x) \geqslant f(x^*)$. اگر به ازای هر $x \in \Omega$، آنگاه x^* را نقطه مینیمم سراسری اکید f روی Ω گویند.

۲.۱ جستجوي خطي

تقريباً همه‌ي الگوريتم‌هايي كاهشي كه مورد بررسي قرار مي گيرند داراي يك ساختار زير بنايي اساسي هستند. هر يك از آنها از يك نقطه اوليه شروع مي‌كند، طبق قانون ثابتي، جهت حركت را مشخص مي‌كند، وسپس در آن جهت به سمت مينيمم (نسبي) تابع هدف بر روي آن خط حركت مي‌كند.در نقطه جديد، جهت جديدي تعيين مي شود،و فرايند تكرار مي‌گردد. تفاوت‌هاي عمده بين الگوريتم‌هاي(روش تندترين كاهش [۱] ، روش نيوتون و غيره) مربوط به قاعده‌ي انتخاب جهت‌هاي متوالي حركت است. پس از آنكه انتخاب انجام شد، همه‌ی الگوريتم‌ها حركت به سوی نقطه مينيمم را روی خط مربوط دنبال می‌كنند.

تعریف ۱.۲.۱. فرایند تعیین نقطه‌ی مینیمم بر روی خطی مفروض را جستجوی خطی [۲] می‌گویند.

برای توابع غیر خطی کلی که نتوان مینیمم آنها را به صورت تحلیلی به دست آورد، این فرایند در واقع جستجوی هوشمندانه برای یافتن نقطه مینیمم در طول خط است. این روش‌های جستجوی خطی، که در حقیقت رویه‌هایی برای حل مسائل مینیمم سازی توابع تك متغیره (یك بعدی) هستند ستون فقرات الگوریتم‌های برنامه‌ریزی غیرخطی را تشکیل می‌دهند، زیرا مسائل چند بعدی (چند متغیره) نهایتاً با اجرای دنباله‌ای از جستجوهای خطی متوالی حل می‌شوند.

اگر بخواهیم جستجوی خطی را به صورت ملموس‌تر بیان کنیم مسئله‌ی بهینه‌سازی نامقید زیر را در نظر می‌گیریم

$$min f(x_k), \tag{۱.۱}$$

[۱] steepest descent
[۲] linear search

که x_k از رابطه بازگشتی زیر تولید می‌شود

$$x_{k+1} = x_k + \alpha_k d_k, \qquad \alpha_k > 0, \tag{۲.۱}$$

که در آن d_k جهت تغییر x_k به x_{k+1} در جهت کاهش $f(x_k)$ می‌باشد و α_k طول گام این تغییر می‌باشد. منظور از جستجوی خطی برای مینیمم کردن $f(x_k)$، انتخاب بهترین مقدار برای α_k است، به‌طوری که با تغییر x_k به x_{k+1} در امتداد d_k ، تابع هدف $f(x_k)$ در این امتداد به مینیمم مقدار خود برسد.

برای رسیدن به این هدف تعریف می‌کنیم

$$\phi(\alpha) = f(x_k + \alpha d_k), \tag{۳.۱}$$

و با استفاده از روش‌های تندترین کاهش و نیوتون که در ادامه این فصل معرفی خواهند شد مقدار مینیمم $\phi(\alpha)$ را به دست می‌آوریم.

۳.۱ جستجوی خطی تقریبی

البته در عمل غیر ممکن است که نقطه مینیمم دقیق را ،که الگوریتم ایده آل جستجوی خطی دنبال می‌کند ،به دست آورد در واقع ،اغلب مطلوب است که در روال یك جستجوی خطی ، به منظور صرفه‌جویی در کل محاسبات، از دقت صرف نظر کنیم. به واسطه این عوامل و به منظور واقع بینی ، در هر مرحله از پیشرفت کار باید مطمئن شویم که به دلیل استفاده از جستجوی تقریبی، نظریه ما مخدوش نشود.

به طور کلی، تقریب در الگوریتم جستجوی خطی با پایان دادن به فرایند جستجو، قبل از همگرا شدن آن، معرفی می‌شود. بنابراین ماهیت دقیق تقریبی که معرفی می‌شود ممکن است که به روش جستجوی به کار گرفته شده و همچنین معیار مورد استفاده برای پایان دادن به جستجو وابسته باشد.

معیارها و آزمون‌های مختلفی وجود دارد که از آنها برای پایان دادن به تکرار روش جستجوی تقریبی برای پیدا کردن α که در معیار مذکور صدق کند، و (۳.۱) را مینیمم کند،وجود دارد که مشهور ترین آنها عبارتند از : ۱-آزمون کورواتور ۲-آزمون آرمیژور ۳- آزمون وولف [۳] ۴- آزمون گلدشتاین [۴]

که ما در ادامه بحث به معرفی آزمون‌های مذکور خواهیم پرداخت . قبل از پرداختن به آزمون‌های فوق دو خاصیت توابع نزولی را در ادامه می‌آوریم که باید همه معیارهای جستجوی تقریبی شامل این دو خاصیت باشند.

خاصیت ۱ . اگر $\bigtriangledown f(x_k) \neq 0$ ، آنگاه

$$f(x_{k+1}) < f(x_k),$$

خاصیت ۱ نشان می‌دهد که هر آزمون باید نزولی بودن تابع هدف را بررسی کند.

خاصیت ۲ . $d_k \bigtriangledown f(x_k) < 0$.

یعنی تابع

$$\phi(\alpha) = f(x_k + \alpha d_k),$$

در خاصیت

$$\phi(0) = \bigtriangledown f(x_k) d_k < 0,$$

صدق می‌کند. درنتیجه، به ازای مقادیر مثبت و به قدر کافی کوچک α، داریم:

$$f(x_k + \alpha d_k) = \phi(\alpha) < \phi(0) = f(x_k),$$

[۳] Wolfe
[۴] Goldstein

اگر خاصیت ۲ برقرار باشد و α به قدر کافی کوچك باشد.آنگاه

$$f(x_{k+۱}) < f(x_k),$$

به این معنی که خاصیت ۲ نیز برقرار است.

۱.۳.۱ آزمون وولف

آزمون وولف از ترکیب دو معیار آرمیژور و کورواتور به دست می‌آید که در فرایند جستجوی خطی اندازه α_k را بررسی می‌کند. مثلاً اگر مقدار α_k خیلی کوچك باشد. آنگاه ممکن است حرکت x_k به سمت مینیمم کننده x^* خیلی سریع نباشد؛ زیرا گام‌های متوالی خیلی کوتاه هستند. برای اجتناب از کوتاه بودن گام‌ها، معیار زیر را داریم که به معیار کورواتور مشهور است.

معیار۱ .(معیار کورواتور) عددی مانند $c_۲$ ، $۰ < c_۲ < ۱$ وجود دارد به طوری که

$$d_k \bigtriangledown f(x_{k+۱}) \geqslant c_۲ d_k \bigtriangledown f(x_k).$$

دلیل اینکه چرا خاصیت ۲ و معیار ۱ از انتخاب‌های کوچك α_k جلوگیری می‌کنند به صورت زیر می‌باشد. فرض کنید $c_۲$ یك عدد ثابت با $۰ < c_۲ < ۱$ باشد و α_k در معیار ۱ صدق کند. در این صورت

$$d_k \bigtriangledown f(x_k + \alpha_k d_k) \geqslant c_۲ d_k \bigtriangledown f(x_k) > d_k \bigtriangledown f(x_k),$$

چون $c_۲ < ۱$ و بنابر خاصیت ۲

$$d_k \bigtriangledown f(x_k) < ۰,$$

به طور متناظر،

$$d_k \bigtriangledown f(x_k+\alpha_k d_k) - d_k \bigtriangledown f(x_k) \geqslant (c_2 - 1) d_k \bigtriangledown f(x_k) - d_k \bigtriangledown f(x_k) > 0,$$

که معنی آن این است که α_kنمی تواند به دلخواه کوچك بوده و یا صفر باشد چون اگر در نامعادله فوق α_k به سمت صفر میل کند، آنگاه سمت چپ آن صفر می‌شود در حالی که سمت راست آن در مقدار $(c_2 - 1) d_k \bigtriangledown f(x_k) > 0,$ ثابت باقی می‌ماندکه ناممکن است. بنابر این، خاصیت ۲ و معیار ۱ باهم از انتخاب مقادیر به دلخواه کوچك برای α_k جلوگیری می کنند.

اما در مورد بزرگی α_k چه می‌توان گفت؟ یك مقدار بزرگ برای α_k باعث یك گام بزرگ از x_k به x_{k+1} می‌شود. یکی از مشکلات روش تندترین کاهش این است که به دنبال مقادیر بزرگ α_k است حتی اگر یك مقدار کوچکتر برای آن نیز می‌تواند باعث همان تقلیل در $f(x)$ شود. این الزام ما را به معیار زیر که به آزمون آرمیژور مشهور است رهنمون می کند.

معیار ۲ . (معیار آرمیژور) عددی مانند c_1 ، $0 < c_1 < c_2 < 1$ ، وجود دارد به طوری که

$$f(x_{k+1}) \leqslant f(x_k) + c_1 \alpha_k d_k \bigtriangledown f(x_k).$$

برای مشاهده اینکه چرا معیار فوق نتیجه مطلوب را به دست می‌دهد، توجه می کنیم که نامساوی بالا را می‌توان به صورت

$$\frac{f(x_k) - f(x_{k+1})}{\alpha_k} \geqslant c_1 [-d_k \bigtriangledown f(x_k)],$$

نوشت. سمت چپ این نامساوی تقلیل نسبی مقدار $f(x)$ را نسبت به افزایش مقدار α بین x_k و x_{k+1} نشان می‌دهد. جمله

$$[-d_k \bigtriangledown f(x_k)],$$

در سمت راست معیار ۲ که بنابر خاصیت ۲ مثبت می باشد، مضربی از اندازه نرخ تقلیل $f(x)$ در x_k و در جهت d_k است. بنابر این معنی معیار ۲ این است که تقلیل نسبی مقدار تابع $f(x)$ نسبت به α_k باید از یك مضرب از پیش تعیین شده اندازه تقلیل $f(x)$ در x_k در جهتی که به x_k می‌رسد، بیشتر باشد.

پس آزمون وولف را می‌توان به صورت زیر خلاصه کرد: اعداد $c_۱$ و $c_۲$ با $۱ > c_۲ > c_۱ > ۰$ مفروضند. عدد α_k را باید طوری باشد که در

$$d_k \bigtriangledown f(x_{k+۱}) \geqslant c_۲ d_k \bigtriangledown f(x_k). \tag{۴.۱}$$

و

$$f(x_{k+۱}) \leqslant f(x_k) + c_۱ \alpha_k d_k \bigtriangledown f(x_k). \tag{۵.۱}$$

صدق کند که ما این مطلب را در قضیه زیر بیان می‌کنیم.

قضیه ۱.۳.۱. *(وولف) فرض کنید تابع $f(x)$ مشتق‌های نسبی مرتبه اول پیوسته داشته و از پایین بر $\mathbb{R}^n$ کراندار باشد.اعداد ثابت $c_۱$ و $c_۲$ با $۱ > c_۲ > c_۱ > ۰$ را در نظر بگیرید. اگر d_k و x_k بردارهایی در $\mathbb{R}^n$ با*

$$d_k \bigtriangledown f(x_k) < ۰,$$

باشند، آنگاه اعداد a_k و b_k با $۰ \leqslant a_k < b_k$ وجود دارند به طوری که

الف) به ازای هر α_k در بازه $(۰, b_k)$ معیار ۲ برقرار است.

ب) به ازای هر α_k در بازه (a_k, b_k)، معیار ۱ برقرار است.

بنابر این، به ازای هر α_k در بازه (a_k, b_k) هر دو معیار ۱ و ۲ برقرار می‌باشند.

اثبات. *به ازای هر α قرار دهید*

$$\phi(\alpha) = f(x_k + \alpha d_k).$$

همانطور که قبلاً دیدیم

$$0 > d_k \nabla f(x_k) = \phi'(0) = \lim_{\alpha \to 0} \frac{f(x_k + \alpha d_k) - f(x_k)}{\alpha}.$$

چون $d_k \nabla f(x_k) < 0$ و $c_1 < 1$ پس

$$c_1 d_k \nabla f(x_k) > d_k \nabla f(x_k) = \lim_{\alpha \to 0} \frac{f(x_k + \alpha d_k) - f(x_k)}{\alpha}.$$

بنابر این عددی مانند $0 < \varepsilon$ وجود دارد به طوری که به ازای هر $0 < \alpha < \varepsilon$

$$c_1 d_k \nabla f(x_k) > \frac{f(x_k + \alpha d_k) - f(x_k)}{\alpha}.$$

از این رو به ازای هر $\varepsilon > \alpha > 0$ داریم:

$$f(x_k + \alpha d_k) < f(x_k) + c_1 \alpha d_k \nabla f(x_k). \tag{۶.۱}$$

فرض کنید α^* بزرگترین مقدار ممکن برای ε باشد. توجه داریم که α^* وجود دارد چون در غیر این صورت به ازای هر $0 < \alpha$ که

$$f(x_k + \alpha d_k) < f(x_k) + c_1 \alpha d_k \nabla f(x_k),$$

غیر ممکن است؛ زیرا $f(x)$ از پایین کراندار است و وقتی که $\alpha \to \infty$،

$$f(x_k) + c_1 \alpha d_k \nabla f(x_k),$$

به سمت $-\infty$ میل می‌کند. و کران پایینی وجود ندارد و این متناقض با کراندار بودن $f(x)$ است. این نشان می‌دهد که α^* یك عدد حقیقی متناهی است و معیار ۲ به ازای هر α در بازه $(0, \alpha^*)$ برقرار می‌باشد.

توجه داریم که

$$f(x_k + \alpha^* d_k) = f(x_k) + c_1 \alpha^* d_k \bigtriangledown f(x_k), \tag{۷.۱}$$

چون در غیر این صورت، از پیوستگی $f(x)$ نتیجه می‌گیریم که $\delta > 0$ وجود دارد به طوری که به ازای هر

$\alpha^* - \delta < \alpha < \alpha^* + \delta$

$$f(x_k + \alpha d_k) < f(x_k) + c_1 \alpha d_k \bigtriangledown f(x_k).$$

بنابر این نامساوی (۶.۱) برای

$$\varepsilon = \alpha^* + \delta > \alpha^*,$$

برقرار است و این متناقض با تعریف α^* می‌باشد.

رابطه (۷.۱) را می‌توان به صورت

$$f(x_k + \alpha^* d_k) - f(x_k) = c_1 \alpha^* d_k \bigtriangledown f(x_k), \tag{۸.۱}$$

نیز نوشت. با به کار بردن قضیه‌ی مقدار میانگین برای تابع

$$\phi(\alpha) = f(x_k + \alpha d_k),$$

بر بازه $[0, \alpha^*]$، داریم:

$$f(x_k + \alpha^* d_k) - f(x_k) = \phi(\alpha^*) - \phi(0) = \phi'(\alpha^{**})(\alpha^* - 0), \tag{۹.۱}$$

به ازای بعضی $\alpha^{**} \in (0, \alpha^*)$. از طرفی دیگر، چون

$$\phi'(\alpha^{**}) = \bigtriangledown f(x_k + \alpha^{**} d_k),$$