دانشکده علوم ریاضی

گروه ریاضی کاربردی

پایان نامه

برای دریافت درجه کارشناسی ارشد در رشته

تحقیق در عملیات

عنوان

مروری بر سه رده از روش‌های گرادیان مزدوج سه‌جمله‌ای

نگارنده

شیوا قرایی

تابستان۱۳۹۷

فهرست مطالب

پیش گفتار ۵

۱ بهینه‌سازی نامقید و روش‌های گرادیان مزدوج ۷

۱.۱ معرفی مساله بهینه‌سازی نامقید ۷

۲.۱ تاریخچه ۸

۳.۱ نمادگزاری ۹

۴.۱ شرایط بهینگی برای مسائل نامقید ۱۱

۵.۱ جستجوی خطی ۱۲

۱.۵.۱ جستجوی خطی دقیق ۱۳

۲.۵.۱ جستجوی خطی تقریبی ۱۳

۶.۱ گرادیان مزدوج سه‌جمله ای ۱۸

۷.۱ عدد حالت ۱۹

۲ روش‌های شبه‌نیوتن ۲۱

۱.۲ بهنگام‌سازی DFP ۲۱

۲.۲ بهنگام‌سازی $BFGS$ ۲۲

۳.۲ شرایط سکانت اصلاح شده ۲۴

۱.۳.۲ شرط سکانت اصلاح شده ژنگ ۲۴

۲.۳.۲ شرایط سکانت اصلاح شده لی ۲۶
۳.۳.۲ شرایط سکانت اصلاح شده سوم(چندگامی) ۲۸

۳ روش‌های گرادیان مزدوج سه‌جمله‌ای مبتنی بر شرایط سکانت ۲۹
۱.۳ روش‌های گرادیان مزدوج دای-لیائو ۲۹
۲.۳ روش‌های گرادیان مزدوج سه‌جمله‌ای براساس شرایط سکانت ۳۲
۱.۲.۳ معرفی الگوریتم ۳۲
۳.۳ همگرایی سراسری ۳۵
۱.۳.۳ همگرایی سراسری الگوریتم $TTCG(Secant)$ با انتخاب‌های ارائه شده ۳۸

۴ روش‌های گرادیان مزدوج سه‌جمله‌ای ترکیبی ۴۴
۱.۴ بررسی روش گرادیان مزدوج سه‌جمله‌ای $NPRP$ ۴۹
۲.۴ بررسی روش گرادیان مزدوج سه‌جمله‌ای NHS ۵۵

۵ روش گرادیان مزدوج سه‌جمله‌ای مبتنی بر کاهش عدد حالت ۶۰

واژه‌نامه انگلیسی به فارسی ۸۰

واژه‌نامه فارسی به انگلیسی ۸۲

واژه‌نامه فارسی به انگلیسی ۸۵

فهرست جدول‌ها

۱.۱ فرمول‌های اساسی ارائه شده برای β_k ۱۶

۱.۳ انتخاب‌های مختلف $z_{k-۱}$ و $h_{k-۱}$ ۳۲

۲.۳ انتخاب‌های مختلف p_k و β_k ۳۴

۱.۴ فرمول های ترکیبی ارائه شده برای β_k ۴۶

فهرست شکل‌ها

۱.۱ طول گام های صادق در شرایط ولف ۱۵

پیش گفتار

روش‌های گرادیان مزدوج یک خانواده بسیار مهم از روش‌های تکراری برای حل مسائل بهینه‌سازی نامقید در ابعاد بزرگ هستند. این روش‌ها به دلیل عدم نیاز به محاسبه و ذخیرسازی ماتریس‌ها، به طور گسترده ای در مسائل ابعاد بزرگ مورد استفاده قرار می‌گیرند. یک ویژگی مهم برای این روش‌ها این است که، جهت‌های جستجوی کاهشی تولید کنند. اما جهت‌های جستجوی تولید شده توسط روش‌های استاندارد (دو جمله‌ای) گرادیان مزدوج لزوما در شرط کاهش کافی صدق نمی‌کنند. اخیرا روش‌های گرادیان مزدوجی که موسوم به روش‌های گرادیان مزدوج سه جمله‌ای هستند، بر این نقص غلبه کرده و جهت‌های جستجویی را تولید می‌کنند که به طور مستقل از جستجوی خطی استفاده شده، همواره در شرط کاهش کافی صدق می‌کنند. به طور کلی ما در این پایان‌نامه:

در فصل اول در ابتدا، به بیان مقدمات و تعاریف برای مسائل گرادیان مزدوج سه‌جمله‌ای می‌پردازیم، و شرایط بهینگی برای مسائل نامقید را بیان می‌کنیم. در ادامه روش‌های کلاسیک بهینه‌سازی نامقید و انواع جستجوی خطی را ارائه می‌دهیم، و در انتهای فصل به طور مختصر گرادیان مزدوج سه جمله‌ای و عدد حالت را توضیح می‌دهیم.

در فصل دوم در شروع فصل، فرمول‌های بهنگام‌سازی برای ماتریس تقریبی در روش‌های شبه نیوتن را بیان می‌کنیم، و در ادامه شرایط سکانت اصلاح شده را که در فصل سوم از آن‌ها استفاده می‌شود را مطرح می‌کنیم.

در فصل سوم در این فصل ابتدا به معرفی روش گرادیان مزدوج دای-لیائو و مشتقات آن می‌پردازیم.

در ادامه همگرایی سراسری برای یک روش تکراری کلی پرداخته و در نهایت شرط کافی برای همگرایی سراسری الگوریتم $TTCG(Secant)$ را مطرح می‌کنیم.

در فصل چهارم فرمول‌های ترکیبی ارائه شده برای β_k را مرور می کنیم. در نهایت به بررسی دو روش از خانواده روش‌های گرادیان مزدوج سه‌جمله‌ای ترکیبی به نام‌های $NPRP$ و NHS می پردازیم.

در فصل پنجم در آخرین فصل، به بیان روش گرادیان مزدوج سه‌جمله‌ای مبتنی بر کاهش عدد حالت می‌پردازیم. تقریب درجه دو از تابع f را در نقطه x_{k+1} نوشته و با مینیمم کردن این تقریب درجه دو، اسکالر β_k را به دست می‌آوریم و در نهایت با مینیمم کرد عدد حالت، پارامتر مجهول در جهت جستجوی تولید شده را به دست می‌آوریم.

فصل ۱

بهینه‌سازی نامقید و روش‌های گرادیان مزدوج

۱.۱ معرفی مساله بهینه‌سازی نامقید

بهینه‌سازی [1] هنر یافتن بهترین جواب در بین وضعیت‌های ممکن است. بهینه‌سازی در طراحی و نگهداری بسیاری از سیستم های مهندسی، اقتصادی و حتی اجتماعی به منظور مینیمم کردن هزینه لازم و یا ماکزیمم کردن سود کاربرد دارد. به دلیل کاربرد وسیع بهینه‌سازی در علوم متفاوت، این مبحث رشد بسیاری کرده است، به طوری که در ریاضیات، مدیریت، صنایع و بسیاری از شاخه‌های علوم مورد مطالعه و بررسی قرار می‌گیرد. برای مثال در علوم مهندسی، مهندسان سعی در بهینه کردن طرح‌های خود دارند یا در سیستم‌های فیزیک تمایل به بهینه کردن انرژی است و یا در علم مدیریت، بهینه‌سازی علمی برای محاسبه بهترین استفاده ممکن از منابع است که نتیجه مطلوب نظیر کمترین هزینه یا بیشترین سود مورد استفاده قرار می‌گیرد. تاکنون پژوهشگران زیادی الگوریتم‌هایی را برای حل مسائل بهینه‌سازی پیشنهاد کرده‌اند. تعیین یک الگوریتم مناسب نقش تعیین کننده‌ای در سرعت حل و کیفیت حل مساله بهینه‌سازی دارد. بنابراین، در بهینه‌سازی باید شرایط بهینگی یک جواب، روش‌های عددی

[1] Optimization

برای محاسبه جواب بهینه، تحلیل همگرایی و کارایی عددی روش‌ها مورد بحث و بررسی قرار گیرد. در یک مساله بهینه‌سازی پیوسته اگر قیود تساوی و نامساوی موجود نباشند، آنگاه مساله، مساله بهینه سازی نامقید[1] نامیده می‌شود. به عبارت دیگر یک مساله بهینه‌سازی نامقید پیوسته به فرم کلی زیر است:

$$\min_{x\in\mathbb{R}^n} \quad f(x) \tag{۱.۱}$$

است. در این پایان‌نامه فرض بر این است که تابع هدف f، همواره مشتق‌پذیر پیوسته به دفعات مورد نیاز است.

۲.۱ تاریخچه

روش گرادیان [2](تندترین کاهش) پایه‌ای‌ترین و اولین روش برای حل مسائل بهینه‌سازی نامقید است که در سال ۱۸۴۷ توسط کوشی [3] ارائه شد [۱۸]. به دلیل رفتار همگرایی ضعیف روش گرادیان پژوهشگران به اصلاح روش گرادیان پرداختند [۷۳]. در نتیجه این تلاش‌ها، روش‌های شبه نیوتن(سکانت) و روش‌های گرادیان مزدوج ارائه شدند. شباهت تکرارها و عدم نیاز به ذخیره و محاسبه مشتقات مرتبه دوم در روش گرادیان مزدوج موجب شده است که از حافظه بسیار کمی نسبت به ابعاد مسئله استفاده شود. همگرایی سراسری و یا موضعی این روش باعث شده است که به عنوان یکی از بهترین روش‌ها برای حل مسائل بهینه‌سازی نامقید در ابعاد بزرگ به کار آید. روش‌های گرادیان مزدوج، اولین بار توسط هستنس[4] و استیفل[5] در دهه ۱۹۵۰ برای حل دستگاه‌های خطی بیان شد [۴۵]. از آنجا که حل یک دستگاه خطی معادل با حل یک مساله بهینه‌سازی نامقید با تابع هدف محدب درجه دوم است، فلچر[6] و ریوز[7] در دهه ۱۹۶۰ این روش را برای حل مسائل بهینه‌سازی نامقید

[1] Unconstraint optimization problem
[2] Gradient Method
[3] Cauchy
[4] Hestenes
[5] Stiefel
[6] Fletcher
[7] Revees

گسترش دادند [۳۶].

۳.۱ نمادگزاری

برای تابع هدف هموار f و نقطه $x_k \in D_f$، نمادهای زیر در نظر گرفته می‌شوند:

$$f_k = f(x_k), \quad g_k = \nabla f(x_k), \quad G_k = \nabla^{۲} f(x_k), \quad y_k = g_{k+۱} - g_k$$

همچنین، نماد$\| . \|$ بیانگر نرم اقلیدسی است.

روش‌های گرادیان مزدوج، بر مزدوج بودن جهت‌های جستجو استوارند. پیش از ورود به بحث روش‌های گرادیان مزدوج، لازم است به طور مختصر به مفهوم پایه‌ای مزدوج بودن و روش‌های جهت های مزدوج بپردازیم.

روش‌های جهت‌های مزدوج، اصولا براساس حل مساله درجه دوم

$$\min_{x \in \mathbb{R}^n} \quad Q(x) = \frac{۱}{۲} x^T Q x - b^T x \tag{۲.۱}$$

که در آن $Q \in \mathbb{R}^{n \times n}$ یک ماتریس متقارن معین مثبت است، ابداع شده است.

تعریف ۱.۳.۱. [۲۰] فرض کنید $f : \mathbb{R}^n \to \mathbb{R}$ و $S \subseteq \mathbb{R}^n$ یک مجموعه محدب باشد. تابع f را روی S را به طور یکنواخت محدب یا قوی محدب گویند، هرگاه داشته باشیم:

$$\exists c > ۰, \forall x_۱, x_۲ \in S, \forall a \in (۰, ۱) :$$
$$f(ax_۱ + (۱ - a)x_۲) \leq af(x_۱) + (۱ - a)f(x_۲) - \frac{۱}{۲} ca(۱ - a) \parallel x_۱ - x_۲ \parallel^{۲} .$$

تعریف زیر مفهوم مزدوج بودن را بیان می‌کند. یکی از ویژگی‌های مهم روش گرادیان مزدوج بردارهای مزدوج هستند. در ادامه تعریف بردارهای مزدوج و قضیه جهت‌های متعامد را بیان می‌کنیم.

تعریف ۲.۳.۱. [۲۰] فرض کنید $Q \in \mathbb{R}^{n \times n}$ یک ماتریس متقارن معین مثبت است. بردارهای

ناصفر $\{d_k\}_{k=1}^{m} \subset \mathbb{R}^n (m \leq n)$ را در نظر بگیرید. اگر

$$d_i^T Q d_j = 0 \quad \forall i \neq j$$

آن‌گاه بردارهای $\{d_k\}$ را بردارهاییQ-مزدوج، یا Q-متعامد و یا به طور خلاصه، بردارهای مزدوج گویند.

نکته: اگر $Q = I$، آن‌گاه بردارهای مزدوج نسبت بهم متعامدند. افزون بر این، هر مجموعه از بردارهای Q-مزدوج در$\mathbb{R}^n$ مستقل خطی هستند و همچنین، برای تولید بردارهایQ-مزدوج، می‌توان از یک روند گرام-اشمیت[1] استفاده کرد [۷۳].

تعریف ۳.۳.۱. [۲۰] برای یک جهت $0 \neq d_k \in \mathbb{R}^n$ گوییم شرط کاهشی برای مساله $\min Q(x) = \frac{1}{2} x^T Q x - b^T x$ در نقطه x_k برقرار است (یا d_k یک جهت کاهشی است) هرگاه

$$g_k^T d_k < 0$$

به‌علاوه، اگر یک عدد ثابت $c > 0$ وجود داشته باشد به‌طوری که

$$g_k^T d_k \leq -c \parallel g_k \parallel^2 \quad \forall k \geq 0 \tag{۳.۱}$$

آن‌گاه گوییم شرط کافی کاهشی برقرار است.

قضیه زیر، که قضیه زیرفضای انبساطی [2] نامیده می‌شود، یک ویژگی اساسی روش‌های جهت‌های مزدوج را برای حل مساله (۲.۱) بیان می‌کند.

قضیه ۴.۳.۱. [۴۵] فرض کنید$\{d_k\}_{k=1}^{m}$ مجموعه ای از m بردار غیر صفر نسبت به Q مزدوج باشند، برای هر نقطه $x_0 \in \mathbb{R}^m$ دنباله نقاط $\{x_k\}$ که از فرمول تکرار

$$x_{k+1} = x_k + s_k, \quad s_k = \alpha_k d_k \tag{۴.۱}$$

به دست می‌آید با

$$\alpha_k = \frac{-g_k^T d_k}{d_k^k Q d_k}$$

[1] Gram-Schmidt

[2] Expansion subsystem theorem

از آن‌جا که $B_m = \mathbb{R}^m$ (زیر فضای تولید شده به‌وسیله مجموعه بردارهای $\{d_k\}_{k=1}^{m}$) پس از m تکرار به جواب بهینه منحصربفرد x^* همگرا می‌شود.

۴.۱ شرایط بهینگی برای مسائل نامقید

قضیه ۱.۴.۱. [۸۰] (شرط لازم مرتبه اول) فرض کنید $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ روی مجموعه باز D به طور پیوسته مشتق‌پذیر باشد. اگر $x^* \in D$ یک مینیمم کننده موضعی مساله (۱.۱) باشد، آنگاه

$$\nabla f(x^*) = 0$$

قضیه ۲.۴.۱. [۸۰] (شرط لازم مرتبه دوم) فرض کنید $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ روی مجموعه باز D دو بار مشتق‌پذیر پیوسته باشد. اگر $x^* \in D$ یک مینیمم کننده موضعی مساله (۱.۱) باشد، آنگاه $\nabla^2 f(x^*)$ نیمه معین مثبت است.

قضیه ۳.۴.۱. [۸۰] (شرط کافی مرتبه دوم) فرض کنید $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ روی مجموعه باز D دو بار مشتق‌پذیر پیوسته باشد. اگر $\nabla f(x^*) = 0$ و $\nabla^2 f(x^*)$ معین مثبت باشد، آنگاه $x^* \in D$ یک مینیمم کننده موضعی اکید مساله (۱.۱) است.

جهت‌های جستجو در روش‌های گرادیان مزدوج، برخلاف روش‌های جهت‌های مزدوج، از قبل مشخص نبوده و در هر تکرار، با استفاده از بردار گرادیان و جهت جستجوی موجود، در صورتی که شرط مزدوج بودن جهت‌های تولید شده نسبت به ماتریس هسین به نوعی حفظ شود، تولید می‌شوند. در واقع روش‌های گرادیان مزدوج را می‌توان به عنوان حد واسطی بین روش نیوتن و روش تندترین کاهش در نظر گرفت. حالت کلی یک روش گرادیان مزدوج بدین قرار است:

$$x_{k+1} = x_k + s_k, \quad s_k = \alpha_k d_k \tag{۵.۱}$$

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad d_0 = -g_0, \quad k \geq 0 \tag{۶.۱}$$

که در آن β_k اسکالری است که پارامتر روش گرادیان مزدوج نامیده می‌شود و α_k مقدار طول گام است که با انجام یک فرآیند جستجوی خطی در امتداد جهت d_k محاسبه می‌گردد.

قضیه ۴.۴.۱. [۶۳] با استفاده از فرمول تکراری (۶.۱) روش گرادیان مزدوج و با استفاده از جستجوی خطی دقیق بعد از $m < n$ تکرار شرایط زیر به ازای هر i برقرار است، $(0 \leq i \leq m)$:

$$d_i^T G d_j = 0, \quad j = 0, 1, ..., i-1 \tag{۷.۱}$$

$$g_i^T g_j = 0, \quad j = 0, 1 ..., i-1 \tag{۸.۱}$$

$$d_i^T g_i = -g_i^T g_i \tag{۹.۱}$$

روابط (۷.۱) تا (۹.۱) به ترتیب نشان‌دهنده همگرایی جهت‌ها، متعامد بودن گرادیان‌ها و شرط کاهشی است.

تعریف ۵.۴.۱. [۲۰] اگر f تابعی درجه دو و اکیدا محدب باشد و α_k با یک فرایند جستجوی خطی دقیق محاسبه شود، یعنی،

$$\alpha_k = arg \min_{\alpha \geq 0} \quad f(x_k + \alpha d_k)$$

آنگاه فرمول تکراری (۶.۱) یک روش گرادیان مزدوج خطی را تعریف می‌کند، و درغیر این صورت یک روش گرادیان مزدوج غیرخطی تعریف می‌شود.

۵.۱ جستجوی خطی

جستجوی خطی براساس بهینه‌سازی توابع تک متغیره استوار است. همان‌طور که اشاره شد، ساختار کلی روش‌های تکراری در بهینه‌سازی به صورت (۵.۱) است که در هر تکرار آن هدف یافتن بردار جهت (کاهشی) d_k و طول گام α_k در امتداد d_k است به طوری که شرط $f(x_k + \alpha_k d_k) < f(x_k)$ برقرار شود. محاسبه به دو روش جستجوی خطی دقیق و جستجوی خطی تقریبی صورت می‌گیرد.

۱.۵.۱ جستجوی خطی دقیق

در این روش α_k طوری انتخاب می‌شود که تابع f در امتداد d_k بیشترین کاهش را داشته باشد. به عبارت دیگر، α_k تابع تک متغیره زیر را مینیمم می‌کند:

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

یعنی،

$$f(x_k + \alpha_k d_k) = \min_{\alpha > 0} f(x_k + \alpha d_k)$$

یا،

$$\phi(\alpha_k) = \min_{\alpha > 0} \phi(\alpha)$$

یا،

$$\nabla f(x_k + \alpha d_k)^T d_k = 0$$

چون حل معادله اخیر به راحتی امکان‌پذیر نیست، در عمل اغلب از جستجوی خطی تقریبی استفاده می‌شود.

۲.۵.۱ جستجوی خطی تقریبی

همان‌طور که اشاره شد، دسترسی به نقطه مینیمم دقیق در الگوریتم جستجوی خطی در حالت کلی غیر ممکن است. لذا، در اعمال یک جستجوی خطی به منظور صرفه‌جویی در زمان محاسبات از دقت مطلوب صرف نظر می‌کنیم. اما، در هر تکرار باید مطمئن شویم که کاهش کافی در تابع هدف ایجاد خواهد شد.

به طور کلی، تقریب در یک الگوریتم جستجوی خطی یعنی پایان دادن به فرآیند جستجو قبل از همگرا شدن به مینیمم. اگر α_k طوری انتخاب شود که تابع هدف کاهش قابل قبولی داشته باشد، به عبارت دیگر کاهش $f(x_k) - f(x_k + \alpha_k d_k)$ مقداری منطقی باشد، به فرآیند تعیین طول گام مربوطه جستجوی خطی تقریبی گفته می‌شود. جستجوی خطی ولف [1] [۷۳] یکی از رایج‌ترین و کاراترین روش

[1] Wolfe

های جستجوی خطی تقریبی است که در آن α_k در شرایط زیر صدق می‌کند:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \qquad (۱۰.۱)$$

$$\nabla f(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k \qquad (۱۱.۱)$$

که در آن، $0 < \delta < \sigma < 1$. نابرابری (۱۰.۱) را که یک معیار مناسب برای ایجاد کاهش قابل قبول در روند الگوریتم است، شرط آرمیژو[1] می‌نامند. به عبارت دیگر، شرط آرمیژو بیان می‌دارد که کاهش f باید با طول گام α_k و مشتق جهتی $g_k^T d_k$ سازگار باشد. همچنین نابرابری (۱۱.۱) را شرط انحنا می نامند که تضمین کننده دور شدن α_k از صفر است.

همان‌طور که در شکل ۱.۱ مشاهده می شود، یک طول گام ممکن است بدون آن‌که به مینیمم کننده نزدیک باشد، در شرایط ولف صدق کند. برای رفع این مشکل، شرایط زیر، که به شرایط قوی ولف [2] معروف است، به کار گرفته می‌شوند:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k$$

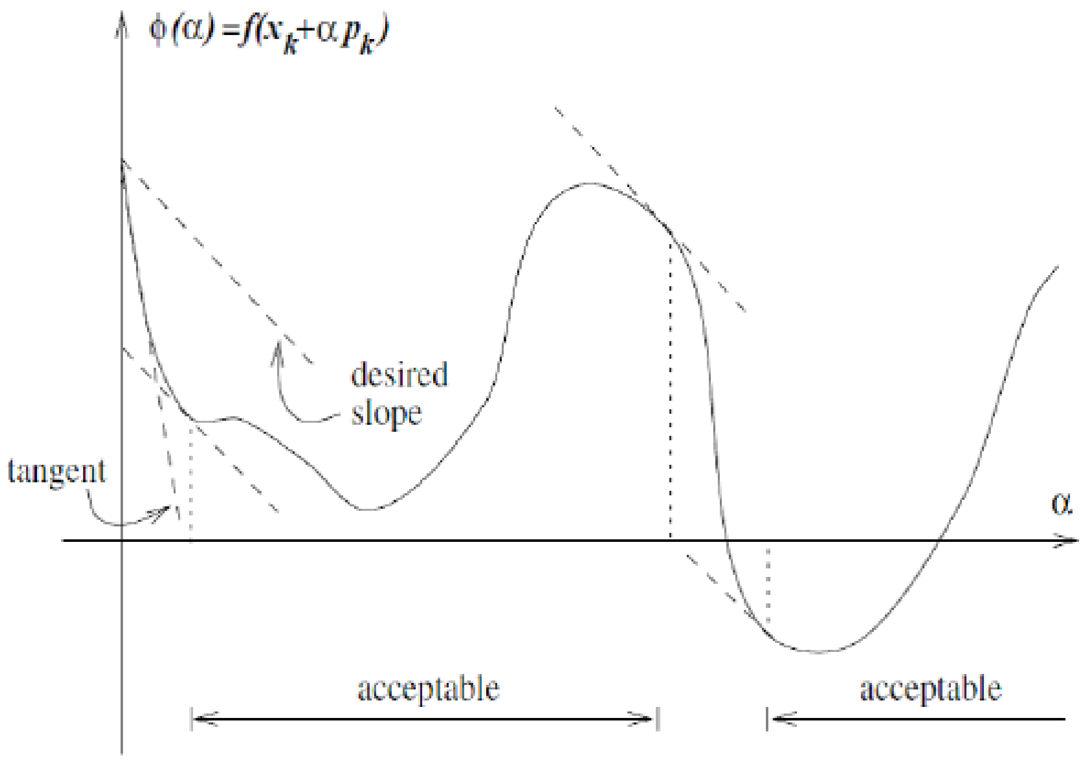
$$\mid g_{k+1}^T d_k \mid \leq -\sigma g_k^T d_k \qquad (۱۲.۱)$$

که در آن $0 < \delta < \sigma < 1$. نامساوی اخیر نشان می‌دهد که با کاهش σ طول گام حاصل به طول گام دقیق میل می‌کند.

جهت‌های جستجو در روش‌های گرادیان مزدوج خطی، طوری محاسبه می‌شوند که نسبت به ماتریس هسین مزدوج باشند [۷۳]. فرمول به دست آمده برای پارامتر یک روش گرادیان مزدوج خطی (β_k) را می‌توان برای توابع هدف دلخواه نیز به کار گرفت، زیرا در نزدیکی یک جواب بهینه، رفتار یک تابع هدف غیر درجه دو به یک تابع هدف درجه دو نزدیک است، و در صورتی که شرایط جستجوی خطی تقریبی مورد استفاده نیز به شرط جستجوی خطی دقیق $(d_k^T g(x_k + \alpha d_k) = 0)$ نزدیک باشند (که این امر با انتخاب δ نزدیک به صفر در شرایط جستجوی خطی قوی ولف امکان‌پذیر است)، آن‌گاه روش گرادیان مزدوج غیر خطی مورد استفاده به یک روش گرادیان مزدوج خطی نزدیک می‌شود و در نتیجه، به دلیل مزدوج بودن جهت‌های تولید شده، رفتار همگرایی مناسبی را از خود نشان می‌دهد.

[1] Armijo

[2] Strong Wolfe conditions



شکل ۱.۱: طول گام های صادق در شرایط ولف

با توجه به کاربرد وسیع روش‌های گرادیان مزدوج، روش‌های مختلفی ارائه شدند که تفاوت عمده آن‌ها در محاسبه پارامتر روش گرادیان مزدوج (β_k) است. سیر تاریخی ارائه فرمول‌های اساسی برای محاسبه پارامتر روش گرادیان مزدوج در جدول ۱.۱ آمده است. تحلیل رفتار عددی شمار زیادی از روش‌های گرادیان مزدوج را می‌توان در [۳] ملاحظه کرد.

جدول ۱.۱: *فرمول‌های اساسی ارائه شده برای* β_k

سال ارائه	β_k	ارائه دهندگان
۱۹۵۲	$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}$	Hestenes and Stiefel [۴۵]
۱۹۶۴	$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}$	Fletcher and Revee [۳۶]
۱۹۶۹	$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}$	Polak،Ribiere [۵۲] and Polyak [۴۶]
۱۹۸۷	$\beta_k^{CD} = -\frac{\|g_{k+1}\|^2}{d_k^T g_k}$	Fletcher [۹]
۱۹۹۱	$\beta_k^{LS} = -\frac{g_{k+1}^T y_k}{g_{k+1}^T y_k}$	Liu and Storey [۲۸]
۱۹۹۹	$\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}$	Dai and Yuan [۲۶]
۲۰۰۱	$\beta_k^{DL} = \frac{g_{k+1}^T (y_k - t s_k)}{d_k^T y_k}$	Dai and Liao [۲۴]

هر یک از روش‌های ارائه شده دارای نقاط ضعف و قوت خاص خود است. به طور مثال، زوتندیک[1] [۸۰] برای توابع دلخواه نشان داد که روش گرادیان مزدوج FR با جستجوی خطی دقیق دارای ویژگی همگرایی سراسری است. به دنبال آن البعلی [2] [۱] برقراری این نتیجه را با استفاده از جستجوی خطی تقریبی نشان داد. اگر چه با وجود نتایج قوی برای همگرایی روش FR ، مشاهدات عددی نشان دهنده آن است که رفتار عددی این روش در مقایسه با روش‌های HS و پولاک [3] ، ریبیر [4] و پولیاک[5] (PRP) ضعیف است [۳]. پاول [۵۷] نیز [6] با ارائه یک مساله خاص نشان داد که روش های HS و PRP با وجود رفتار عددی مناسب ممکن است دچار دور شود و به جواب بهینه همگرا نشوند. افزون بر این، وی با پیشنهاد یک فرایند شروع دوباره با جهت تندترین کاهش (جهت منفی

[1] Zoutendijk
[2] Al-Baali
[3] Polak
[4] Ribiere
[5] Polyak
[6] Powell

گرادیان) در صورت تولید یک جهت نامناسب، توانست رفتار عددی روش FR را تا حدودی بهبود دهد [۵۵]. در نتیجه، می‌توان گفت که با وجود نتایج نظری قوی در مورد همگرایی روش کاهشی مزدوج (CD) ودای و یوان (DY) ، این روش‌ها کارایی عددی مطلوبی ندارند. از طرف دیگر، روش های HS، PRP و لیو [۱] و استوری [۲] (LS) با وجود این‌که ممکن است همیشه همگرا نباشد، ولی رفتار عددی مناسبی دارند. این ویژگی‌ها پژوهشگران را بر آن داشت تا به طراحی روش‌هایی بپردازند که با رفتار عددی مناسب همواره به جواب همگرا باشند. برای این منظور تلاش‌هایی برای استفاده ضمنی از اطلاعات مربوط به ماتریس هسین یا ترکیب روش‌های بالا صورت گرفت که در ادامه به اختصار به آن‌ها اشاره می‌کنیم.

پری[۳] [۵۲] به این نکته توجه کرد که با توجه به فرمول β_k^{HS}، جهت جستجوی روش HS را می‌توان به صورت زیر نوشت:

$$d_{k+1} = -\big(I - \frac{d_k y_k^T}{d_k^T y_k}\big) g_{k+1} = -\big(I - \frac{s_k y_k^T}{s_k^T y_k}\big) g_{k+1} = -P_{k+1} g_{k+1}$$

بنابراین، وی روش اصلاح شده HS را به صورت زیر پیشنهاد کرد:

$$d_{k+1} = -\big(I - \frac{s_k y_k^T}{s_k^T y_k} + \frac{s_k s_k^T}{s_k^T y_k}\big) g_{k+1} = -Q_{k+1} g_{k+1}$$

پری برای توجیه روش خود اشاره می‌کند که افزودن ماتریس رتبه یک $\frac{s_k s_k^T}{s_k^T y_k}$ به ماتریس P_{k+1} باعث می شود که ماتریس حاصل Q_{k+1} در شرط زیر، که به شرط سکانت معروف است، صدق کند:

$$y_k^T Q_{k+1} = s_k^T \qquad (۱۳.۱)$$

شانو[۴] [۶۰] ماتریس Q_{k+1} ارائه شده توسط پری را به صورت زیر اصلاح کرد:

[1] Liu
[2] Storey
[3] Perry
[4] Shanno

$$R_{k+1} = I - \frac{s_k y_k^T + y_k s_k^T}{s_k^T y_k} + (1 + \frac{y_k^T y_k}{s_k^T y_k}) \frac{s_k s_k^T}{s_k^T y_k}$$

شانو برای توجیه روش خود به شباهت فرمول بهنگام‌سازی ماتریس R_{k+1} با بهنگام سازی $BFGS$ اشاره می‌کند، زیرا اگر در بهنگام سازی R_{k+1} به جای ماتریس I ماتریس R_k قرار داده شود، آنگاه فرمول بهنگام‌سازی حاصل همان بهنگام‌سازی $BFGS$ است. به‌علاوه، بهنگام‌سازی R_{k+1} این مزیت را داراست که نیازی به ذخیره‌سازی هیچ ماتریسی ندارد. می‌توان نشان داد که اگر $s_k^T y_k > 0$، آنگاه ماتریس R_{k+1} همواره معین مثبت است [۷۰]، و بنابراین روش گرادیان مزدوج شانو همواره جهت هایی کاهشی تولید می‌کند. روش‌های گرادیان مزدوج دیگری نیز بر پایه استفاده از بهنگام‌سازی $BFGS$ ارائه شده‌اند (به عنوان مثال، به مراجع [۴، ۵۹] نگاه کنید).

۶.۱ گرادیان مزدوج سه‌جمله ای

از روش گرادیان مزدوج سه‌جمله ای برای حل مسائل بهینه‌سازی نامقید در ابعاد بزرگ استفاده می‌شود. این روش با استفاده از جستجوی خطی به‌دست می‌آید. برای شروع از نقطه اولیه x_0 استفاده می‌کنیم. دنباله نقاط $\{x_k\}$ به‌صورت زیر تولید می‌شوند:

$$x_{k+1} = x_k + \alpha_k d_k \qquad (۱۴.۱)$$

که $\alpha_k > 0$ و با استفاده از جستجوی خطی به دست می‌آید، و جهت‌های d_k به صورت زیر محاسبه می شوند:

$$d_{k+1} = -g_{k+1} + \delta_k s_k + \eta_k y_k \quad d_0 = -g_0 \qquad (۱۵.۱)$$

در رابطه δ_k و η_k پارامترهای گرادیان مزدوج هستند، که $s_k = x_{k+1} - x_k$، $y_k = g_{k+1} - g_k$ و $g_k = \bigtriangledown f(x_k)$. همان‌طور که مشاهده می‌کنیم جهت جستجوی d_{k+1}، متشکل از ترکیب خطی $-g_{k+1}$، s_k و y_k است. جستجوی خطی در الگوریتم‌های گرادیان مزدوج اغلب به‌وسیله شرایط قوی

ولف انجام می‌شود، که برای اطمینان از همگرایی و افزایش ثبات مورد نیاز است، که در مطالب قبل به آن اشاره شده است. با توجه به وجود انتخاب‌های متفاوت برای پارامترهای δ_k و η_k الگوریتم‌های گرادیان مزدوج سه‌جمله ای مختلفی به‌دست آمده است. مقالات بیل [1] [۹، ۱۵]، نسخه‌های مختلفی از الگوریتم‌های گرادیان مزدوج سه‌جمله ای را همراه با خواص آن‌ها، همگرایی سراسری و عملکرد عددی ارائه می‌دهند. تمام این الگوریتم‌های گرادیان مزدوج سه‌جمله ای با اصلاح الگوریتم‌های گرادیان مزدوج کلاسیک به دنبال آن هستند تا شرایط همگرایی و کاهشی بودن را به‌دست آورند.

۷.۱ عدد حالت

در فصل پنجم این پایان‌نامه به روشی اشاره می‌شود، که به منظور به‌دست آوردن پارامتر جهت جستجو نیاز به مینیمم کردن عدد حالت دارد. که در این بخش، به معرفی عدد حالت ماتریس A می‌پردازیم. عدد حالت یک ابزار ساده و مفید برای اندازه‌گیری حساسیت دستگاه خطی $Ax = b$ می‌باشد.

دستگاه خطی $Ax = b$ را در نظر بگیرید، که در آن A نامنفرد و b غیر صفر است. در این صورت یک جواب منحصر به فرد x وجود دارد که غیر صفر می‌باشد. حال، فرض کنید که یک بردار کوچک δb به b اضافه شود و دستگاه پریشیده شده $A\hat{x} = b + \delta b$ را در نظر بگیرید. این دستگاه نیز یک جواب منحصر به فرد $\hat{x}$ دارد، که امیدواریم خیلی دور از x نباشد. فرض کنید δx نشان دهنده اختلاف بین x و $\hat{x}$ باشد، یعنی $\hat{x} = x + \delta x$. انتظار داریم که اگر δb کوچک است، آن‌گاه δx نیز کوچک باشد. عبارت دقیق‌تری برای این مطلب شامل خطاهای نسبی است، وقتی می‌گوییم δb کوچک است منظورمان این است که δb در مقایسه با b کوچک می‌باشد، به طور مشابه، δx کوچک است، یعنی این که در مقایسه با x کوچک می‌باشد. برای سنجش اندازه بردارها، از یک نرم برداری مانند $\| \cdot \|$ استفاده می‌کنیم. در این صورت اندازه δb نسبت به b عبارت است از $\| \delta b \| / \| b \|$، و اندازه δx نسبت به xعبارت است از $\| \delta x \| / \| x \|$. دوست داریم که اگر $\| \delta b \| / \| b \|$ کوچک است، آن‌گاه $\| \delta x \| / \| x \|$ نیز کوچک باشد.

[1] Beale

معادلات $Ax = b$ و $A(x + \delta x) = b + \delta b$ نتیجه می‌دهند که $A\delta x = \delta b$ یا $\delta x = A^{-1}\delta b$. نرم برداری انتخاب شده هر چه که باشد، از نرم ماتریسی ناشی شده از آن برای سنجش اندازه ماتریس‌ها استفاده می‌کنیم. از معادله $\delta x = A^{-1}\delta b$ نتیجه می‌شود که

$$\| \delta x \| \leq \| A^{-1} \| \| \delta b \| . \tag{۱.۱۶}$$

به طور مشابه، معادله $Ax + b$ نتیجه می‌دهد که $\| b \| \leq \| A \| \| x \|$، یا به طور هم ارز

$$\frac{1}{\| x \|} \leq \| A \| \frac{1}{\| b \|} \tag{۱.۱۷}$$

با ضرب نامساوی‌های (۱.۱۶) و (۱.۱۷) داریم

$$\frac{\| \delta x \|}{\| x \|} \leq \| A \| \| A^{-1} \| \frac{\| \delta b \|}{\| b \|} \tag{۱.۱۸}$$

که یک کران برای $\| \delta x \| / \| x \|$ برحسب $\| \delta b \| / \| b \|$ به دست می‌دهند. به $\| A \| \| A^{-1} \|$ عدد حالت A می‌گوییم و آن را با $\kappa(A)$ نشان می‌دهیم [۵۰]. یعنی

$$\kappa(A) = \| A \| \| A^{-1} \|$$

در محاسبات علمی نشان داده می‌شود که خطاهای محاسبات(خطای گرد کردن) با عدد حالت متناسب هستند پس هر چه عدد حالت کوچکتر باشد نتایج قابل اطمینان‌تر است.

مراجع

[1] Al-Baali, Mehiddin. Descent property and global convergence of the fletcher—reeves method with inexact line search. *IMA Journal of Numerical Analysis*, 5(1):121–124, 1985.

[2] Andrei, N. Hybrid conjugate gradient algorithm for unconstrained optimization. *Journal of Optimization Theory and Applications*, 141(2):249–264, 2009.

[3] Andrei, Neculai. Numerical comparison of conjugate gradient algorithms for unconstrained optimization. *Studies in Informatics and Control*, 16(4):333–352, 2007.

[4] Andrei, Neculai. A scaled bfgs preconditioned conjugate gradient algorithm for unconstrained optimization. *Applied Mathematics Letters*, 20(6):645–650, 2007.

[5] Andrei, Neculai. Another hybrid conjugate gradient algorithm for unconstrained optimization. *Numerical Algorithms*, 47(2):143–156, 2008.

[6] Andrei, Neculai. Accelerated conjugate gradient algorithm with finite difference hessian/vector product approximation for unconstrained

optimization. *Journal of Computational and Applied Mathematics*, 230(2):570–582, 2009.

[7] Andrei, Neculai. Accelerated hybrid conjugate gradient algorithm with modified secant condition for unconstrained optimization. *Numerical Algorithms*, 54(1):23–46, 2010.

[8] Andrei, Neculai. On three-term conjugate gradient algorithms for unconstrained optimization. *Applied Mathematics and Computation*, 219(11):6316–6327, 2013.

[9] Andrei, Neculai. A simple three-term conjugate gradient algorithm for unconstrained optimization. *Journal of Computational and Applied Mathematics*, 241:19–29, 2013.

[10] Babaie-Kafaki, Saman. On optimality of two adaptive choices for the parameter of dai–liao method. *Optimization Letters*, 10(8):1789–1797, 2016.

[11] Babaie-Kafaki, Saman, Fatemi, Masoud, and Mahdavi-Amiri, Nezam. Two effective hybrid conjugate gradient algorithms based on modified bfgs updates. *Numerical Algorithms*, 58(3):315–331, 2011.

[12] Babaie-Kafaki, Saman and Ghanbari, Reza. The dai–liao nonlinear conjugate gradient method with optimal parameter choices. *European Journal of Operational Research*, 234(3):625–630, 2014.

[13] Babaie-Kafaki, Saman and Ghanbari, Reza. A descent family of dai–liao conjugate gradient methods. *Optimization Methods and Software*, 29(3):583–591, 2014.

[14] Babaie-Kafaki, Saman and Mahdavi-Amiri, Nezam. Two modified hybrid conjugate gradient methods based on a hybrid secant equation. *Mathematical Modelling and Analysis*, 18(1):32–52, 2013.

[15] Beale, EML and Lootsma, FA. Numerical methods for nonlinear optimization, 1972.

[16] Byrd, Richard H and Nocedal, Jorge. A tool for the analysis of quasi-newton methods with application to unconstrained minimization. *SIAM Journal on Numerical Analysis*, 26(3):727–739, 1989.

[17] Byrd, Richard H, Nocedal, Jorge, and Yuan, Ya-Xiang. Global convergence of a cass of quasi-newton methods on convex problems. *SIAM Journal on Numerical Analysis*, 24(5):1171–1190, 1987.

[18] Cauchy, Augustin. Méthode générale pour la résolution des systemes d'équations simultanées. *Comp. Rend. Sci. Paris*, 25(1847):536–538, 1847.

[19] Cheng, Wanyou. A two-term prp-based descent method. *Numerical Functional Analysis and Optimization*, 28(11-12):1217–1230, 2007.

[20] Cheng, Wanyou and Liu, Qunfeng. Sufficient descent nonlinear conjugate gradient methods with conjugacy condition. *Numerical Algorithms*, 53(1):113, 2010.

[21] Cheng, Wanyou and Liu, Qunfeng. Sufficient descent nonlinear conjugate gradient methods with conjugacy condition. *Numerical Algorithms*, 53(1):113, 2010.

[22] Dai, Y-H and Liao, L-Z. New conjugacy conditions and related nonlinear conjugate gradient methods. *Applied Mathematics and Optimization*, 43(1):87–101, 2001.

[23] Dai, Y-H and Liao, L-Z. New conjugacy conditions and related nonlinear conjugate gradient methods. *Applied Mathematics and Optimization*, 43(1):87–101, 2001.

[24] Dai, YH and Yuan, Yaxiang. An efficient hybrid conjugate gradient method for unconstrained optimization. *Annals of Operations Research*, 103(1-4):33–47, 2001.

[25] Dai, Yu-Hong and Kou, Cai-Xia. A nonlinear conjugate gradient algorithm with an optimal property and an improved wolfe line search. *SIAM Journal on Optimization*, 23(1):296–320, 2013.

[26] Dai, Yu-Hong and Yuan, Yaxiang. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, 10(1):177–182, 1999.

[27] Dai, Yu-Hong and Yuan, Yaxiang. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, 10(1):177–182, 1999.

[28] Deng, Naiyang and Li, Zhengfeng. Global convergence of three terms conjugate gradient methods. *Optimization Methods and Software*, 4(4):273–282, 1994.

[29] Dennis, John E and Moré, Jorge J. A characterization of superlinear convergence and its application to quasi-newton methods. *Mathematics*

of computation, 28(126):549–560, 1974.

[30] Dong, Xiao-Liang, Han, De-Ren, Ghanbari, Reza, Li, Xiang-Li, and Dai, Zhi-Feng. Some new three-term hestenes–stiefel conjugate gradient methods with affine combination. *Optimization*, 66(5):759–776, 2017.

[31] Dong, Xiao Liang, Liu, Hong Wei, He, Yu Bo, and Yang, Xi Mei. A modified hestenes–stiefel conjugate gradient method with sufficient descent condition and conjugacy condition. *Journal of Computational and Applied Mathematics*, 281:239–249, 2015.

[32] Dong, XiaoLiang, Liu, Hongwei, and He, Yubo. A self-adjusting conjugate gradient method with sufficient descent condition and conjugacy condition. *Journal of Optimization Theory and Applications*, 165(1):225–241, 2015.

[33] Dong, XiaoLiang, Liu, HongWei, He, YuBo, Babaie-Kafaki, Saman, and Ghanbari, Reza. A new three–term conjugate gradient method with descent direction for unconstrained optimization. *Mathematical Modelling and Analysis*, 21(3):399–411, 2016.

[34] Fatemi, M. An optimal parameter for dai–liao family of conjugate gradient methods. *Journal of Optimization Theory and Applications*, 169(2):587–605, 2016.

[35] Fatemi, Masoud. A new efficient conjugate gradient method for unconstrained optimization. *Journal of Computational and Applied Mathematics*, 300:207–216, 2016.

[36] Fletcher, Reeves and Reeves, Colin M. Function minimization by conjugate gradients. *The computer journal*, 7(2):149–154, 1964.

[37] Ford, JA and Moghrabi, IA. Alternative parameter choices for multi-step quasi-newton methods. *Optimization Methods and Software*, 2(3-4):357–370, 1993.

[38] Ford, JA and Moghrabi, IA. Multi-step quasi-newton methods for optimization. *Journal of Computational and Applied Mathematics*, 50(1-3):305–323, 1994.

[39] Ford, John A, Narushima, Yasushi, and Yabe, Hiroshi. Multi-step nonlinear conjugate gradient methods for unconstrained minimization. *Computational Optimization and Applications*, 40(2):191–216, 2008.

[40] Gilbert, Jean Charles and Nocedal, Jorge. Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on optimization*, 2(1):21–42, 1992.

[41] Gilbert, Jean Charles and Nocedal, Jorge. Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on optimization*, 2(1):21–42, 1992.

[42] Guo, Qiang, Liu, Jian-Guo, and Wang, Dan-Hong. A modified bfgs method and its superlinear convergence in nonconvex minimization with general line search rule. *Journal of Applied Mathematics and Computing*, 28(1-2):435, 2008.

[43] Hager, William W and Zhang, Hongchao. A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM*

Journal on optimization, 16(1):170–192, 2005.

[44] Hager, William W and Zhang, Hongchao. A survey of nonlinear conjugate gradient methods. *Pacific journal of Optimization*, 2(1):35–58, 2006.

[45] Hestenes, Magnus Rudolph and Stiefel, Eduard. *Methods of conjugate gradients for solving linear systems*, vol. 49. NBS Washington, DC, 1952.

[46] Hestenes, Magnus Rudolph and Stiefel, Eduard. *Methods of conjugate gradients for solving linear systems*, vol. 49. NBS Washington, DC, 1952.

[47] Hu, YF and Storey, C. Global convergence result for conjugate gradient methods. *Journal of Optimization Theory and Applications*, 71(2):399–405, 1991.

[48] Li, Dong-Hui and Fukushima, Masao. A modified bfgs method and its global convergence in nonconvex minimization. *Journal of Computational and Applied Mathematics*, 129(1-2):15–35, 2001.

[49] Li, Dong-Hui and Fukushima, Masao. On the global convergence of the bfgs method for nonconvex unconstrained optimization problems. *SIAM Journal on Optimization*, 11(4):1054–1064, 2001.

[50] Luenberger, David G, Ye, Yinyu, et al. *Linear and nonlinear programming*, vol. 2. Springer, 1984.

[51] Narushima, Yasushi and Yabe, Hiroshi. A new three term conjugate gradient method for unconstrained optimization. *FRONTIERS SCIENCE SERIES*, 49:109, 2007.

[52] Perry, Avinoam. A modified conjugate gradient algorithm. *Operations Research*, 26(6):1073–1078, 1978.

[53] Perry, Avinoam. A modified conjugate gradient algorithm. *Operations Research*, 26(6):1073–1078, 1978.

[54] Polyak, Boris Teodorovich. The conjugate gradient method in extremal problems. *USSR Computational Mathematics and Mathematical Physics*, 9(4):94–112, 1969.

[55] Powell, Michael James David. Restart procedures for the conjugate gradient method. *Mathematical programming*, 12(1):241–254, 1977.

[56] Powell, Michael JD. Some global convergence properties of a variable metric algorithm for minimization without exact line searches. *Nonlinear programming*, 9(1):53–72, 1976.

[57] Powell, Michael JD. Nonconvex minimization calculations and the conjugate gradient method. in *Numerical analysis*, pp. 122–141. Springer, 1984.

[58] Powell, MJD. On the convergence of the variable metric algorithm. *IMA Journal of Applied Mathematics*, 7(1):21–36, 1971.

[59] Powell, MJD. Updating conjugate directions by the bfgs formula. *Mathematical Programming*, 38(1):29, 1987.

[60] Shanno, David F. Conjugate gradient methods with inexact searches. *Mathematics of operations research*, 3(3):244–256, 1978.

[61] Shanno, DF. On the convergence of a new conjugate gradient algorithm. *SIAM Journal on Numerical Analysis*, 15(6):1247–1257, 1978.

[62] Sugiki, Kaori, Narushima, Yasushi, and Yabe, Hiroshi. Globally convergent three-term conjugate gradient methods that use secant conditions and generate descent search directions for unconstrained optimization. *Journal of Optimization Theory and Applications*, 153(3):733–757, 2012.

[63] Sun, Wenyu and Yuan, Ya-Xiang. *Optimization theory and methods: nonlinear programming*, vol. 1. Springer Science & Business Media, 2006.

[64] Sun, Wenyu and Yuan, Ya-Xiang. *Optimization theory and methods: nonlinear programming*, vol. 1. Springer Science & Business Media, 2006.

[65] Toint, Ph L. Global convergence of the partitioned bfgs algorithm for convex partially separable optimization. *Mathematical Programming*, 36(3):290–306, 1986.

[66] Touati-Ahmed, D and Storey, C. Efficient hybrid conjugate gradient techniques. *Journal of Optimization Theory and Applications*, 64(2):379–397, 1990.

[67] Watkins, David S. Fundamentals of matrix computations. 2002.

[68] Wei, Zengxin, Li, Guoyin, and Qi, Liqun. New quasi-newton methods for unconstrained optimization problems. *Applied Mathematics and Computation*, 175(2):1156–1188, 2006.

[69] Wei, Zengxin, Li, Guoyin, and Qi, Liqun. New quasi-newton methods for unconstrained optimization problems. *Applied Mathematics and Computation*, 175(2):1156–1188, 2006.

[70] Wright, Stephen and Nocedal, Jorge. Numerical optimization. *Springer Science*, 35(67-68):7, 1999.

[71] Yabe, Hiroshi and Takano, Masahiro. Global convergence properties of nonlinear conjugate gradient methods with modified secant condition. *Computational Optimization and Applications*, 28(2):203–225, 2004.

[72] Yuan, Ya-xiang. A modified bfgs algorithm for unconstrained optimization. *IMA Journal of Numerical Analysis*, 11(3):325–332, 1991.

[73] Yuan, Ya-xiang and Byrd, Richard H. Non-quasi-newton updates for unconstrained optimization. *Journal of Computational Mathematics*, pp. 95–107, 1995.

[74] Zhang, Jianguo, Xiao, Yunhai, and Wei, Zengxin. Nonlinear conjugate gradient methods with sufficient descent condition for large-scale unconstrained optimization. *Mathematical Problems in Engineering*, 2009, 2009.

[75] Zhang, JZ, Deng, NY, and Chen, LH. New quasi-newton equation and related methods for unconstrained optimization. *Journal of Optimization Theory and Applications*, 102(1):147–167, 1999.

[76] Zhang, Li, Zhou, Weijun, and Li, Dong-Hui. A descent modified polak–ribière–polyak conjugate gradient method and its global convergence. *IMA Journal of Numerical Analysis*, 26(4):629–640, 2006.

[77] Zhang, Li, Zhou, Weijun, and Li, Donghui. Some descent three-term conjugate gradient methods and their global convergence. *Optimisation Methods and Software*, 22(4):697–711, 2007.

[78] Zhang, Li and Zhou, Youhua. A note on the convergence properties of the original three-term hestenes-stiefel method. *AMO—Advanced Modeling and Optimization*, 14:159–163, 2012.

[79] Zhou, Weijun and Zhang, Li. A nonlinear conjugate gradient method based on the mbfgs secant condition. *Optimisation Methods and Software*, 21(5):707–714, 2006.

[80] Zoutendijk, G. Nonlinear programming, computational methods. *Integer and nonlinear programming*, pp. 37–86, 1970.

واژه‌نامه انگلیسی به فارسی

A

فرض .... Assumption

الگوریتم .... Algorithm

C

سراسری .... Convergence

قید .... Constant

E

آزمایش .... Experiment

F

تابع .... Function

I

تکراری .... Iterative

اطلاعات .... Information

L

لم .... Lemma

M

روش .... Method

Matrix .... ماتریس

N

Numerical .... عددی

Nonlinear .... غرخطی

O

Optimization .... بهینه سازی

P

Problem .... مسئله

Properties .... شرایط

Purpose .... هدف

Positive .... مثبت

Proof .... اثبات

Q

Quasi-Newton .... شبه نیوتن

Quadratic .... درجه دو

R

Results .... نتایج

S

Sequence .... دنباله

Symmetrization .... متقارن

W

Wolfe conditions .... شرایط ولف

واژه‌نامه فارسی به انگلیسی

آ

الگوریتم .... Algorithm

آرمیژو .... Armijo

اندازه گیری .... Measurement

استاندارد .... Standard

ب

بهینه سازی .... Optimization

بردار .... Victor

بهنگام سازی .... Updating

پ

پارامتر .... Parameter

پیوسته .... Continuous

ت

تقریبی .... Approximaton

تیلور .... Taylor

تناقض .... Contradiction

ترکیب محدب .... Convex composition

ث

ثابت .... Constant

ج

جهت .... Direction

د

دترمینال .... Deteminal

ر

روش .... Method

ز

زیرفضا .... Subsystem

زوتندیک .... Zotendijk

س

سکانت .... Secant

سه جمله ای .... Three term

ع

عدد حالت .... Condition number

عدد ثابت .... Constant number

ق

قضیه .... Theorem

قضیه مقدار میانگین .... The mean value theorem

ک

کاهشی .... Descent

کراندار .... Bounded

گ

گرادیان .... Gradient

گرام-اشمیت .... Gram-Schmidt

م

محدب .... Convex

متقارن .... Symmetrical

ماتریس .... Matrix

ن

نیوتن .... Newton

نرمال .... Normal

ه

همگرایی .... Convergence

هسین .... Hessian

و

وارون پذیر ....

ولف .... Wolfe

ی

یکنواخت .... Uniform