مجله تحقیق در عملیات در کاربردهای آن، سال نوزدهم، شماره اول (پیاپی ۷۲)،بهار ۰۱، ص ص ۱۶-۱
شاپا ۷۲۸۶-۲۲۵۱

یک الگوریتم گرادیان مزدوج کارا برای حل مسایل بهینه‌سازی نامقید

فهیمه عبداللهی[1*]، سید مسعود فاطمی[2]
۱-دانشجوی دکتری، دانشگاه صنعتی خواجه نصیرالدین طوسی، گروه ریاضی کاربردی، تهران، ایران
۲-استاد، دانشگاه صنعتی خواجه نصیرالدین طوسی، گروه ریاضی کاربردی، تهران، ایران

رسید مقاله: ۲ شهریور ۱۳۹۹
پذیرش مقاله: ۳ خرداد ۱۴۰۰

چکیده
در این مقاله به معرفی یک روش گرادیان مزدوج کارآمد پرداخته می‌شود. این روش از حل یک مساله بهینه‌سازی که به کمک ویژگی‌های گرادیان مزدوج خطی و همچنین یک شرط سکانت اصلاح شده به دست آمده، حاصل می‌شود. پارامتر گرادیان مزدوج معرفی‌شده علاوه بر مقادیر گرادیان، در هر تکرار از مقادیر تابع نیز استفاده می‌کند. روش ارایه‌شده با فرضیات مناسب خاصیت همگرایی سراسری دارد. در انتها با استفاده از مسایل کتابخانه CUTEr به مقایسه روش با برخی روش‌های شناخته‌شده موجود می‌پردازیم و کارایی روش را ارزیابی می‌کنیم.

کلمات کلیدی: روش گرادیان مزدوج، شرط سکانت بهبودیافته، بهینه‌سازی نامقید.

۱ مقدمه
در دهه‌های اخیر روش‌های گرادیان مزدوج به دلیل شرایط همگرایی قوی محلی و سراسری و نیز نیاز به حافظه‌ی کم برای حل مسایل بزرگ بسیار مورد توجه محققان بوده‌اند.

مساله‌ی بهینه‌سازی نامقید را در نظر بگیرید:

$$\min_{x\in R^n} f(x) \qquad (۱)$$

که در اینجا $f:R^n\to R$ یک تابع هموار است. الگوریتم‌های گرادیان مزدوج دنباله‌ی تکراری زیر را می‌سازند:

$$x_{k+1}=x_k+\alpha_k d_k, \qquad (۲)$$

که در آن $\alpha_k>0$ طول گام تولیدشده توسط روش جستجوی خطی دقیق یا غیردقیق مانند شرایط وولف قوی (یا استاندارد) است که عبارتند از:

$$f(x_k+\alpha_k d_k)\leqslant f(x_k)+c_1\alpha_k d_k^T g_k, \qquad (۳)$$

و

* عهده‌دار مکاتبات
آدرس الکترونیکی: fabdollahi@email.kntu.ac.ir

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

$$|d_k^T g_{k+1}| \leqslant -c_2 d_k^T g_k \quad (d_k^T g_{k+1} \geq c_2 d_k^T g_k), \tag{۴}$$

که $0 < c_1 < c_2 < 1$ ثابت‌های دلخواه و $g_k := \nabla f(x_k)$ است. همچنین جهت جستجوی d_k به صورت بازگشتی از رابطه‌ی زیر به‌دست می‌آید:

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad (d_0 = -g_0). \tag{۵}$$

پارامتر گرادیان مزدوج که غالبا با β_k نمایش داده می‌شود، یکی از مهم‌ترین پارامترهای این روش است و باید با دقت تعریف شود، زیرا انتخاب‌های متفاوت β_k منجر به تولید انواع روش‌های گرادیان مزدوج می‌شود.

همان‌طور که می‌دانیم، روش‌های گرادیان مزدوج برای مسایل درجه دوم محدب قوی با جستجوی خطی دقیق در هر دو شرط مزدوج بودن

$$d_{k+1}^T y_k = 0, \tag{۶}$$

و تعامد

$$d_i^T g_k = 0 \qquad i = 0, 1, \ldots, k-1, \tag{۷}$$

صدق می‌کنند، که در اینجا $y_k := g_{k+1} - g_k$.

برخی از انتخاب‌های شناخته‌شده برای β_k عبارتند از هستن و اشتیفل(HS) [۱]، فلچر و ریوس(FR) [۲]، پولاک و ریبیر و پولیاک(PRP) [۳]، فلچر(CD) [۴]، لیو و استوری(LS) [۵] و دای و یوان(DY) [۶] که در اینجا به معرفی آنها پرداخته‌ایم

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}, \quad \beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}, \quad \beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2},$$

$$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k}, \quad \beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k}, \quad \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}$$

که $\|.\|$ نرم اقلیدسی است.

نتایج حاصل از مطالعات تئوری و عددی ارایه‌شده توسط آندری در [۷] نشان می‌دهد که برخی روش‌ها مانند HS، PRP و LS (که در صورت عبارت $g_{k+1}^T y_k$ را دارند) عملکرد عددی بهتری نسبت به روش‌های FR، CD و DY (که در صورت عبارت $\|g_{k+1}\|$ را دارند) نشان می‌دهند، در حالی‌که دسته‌ی دوم شرایط همگرایی قوی سراسری را دارند. امروزه محققان علاقمند به بهبود و معرفی پارمترهایی هستند که روش حاصل علاوه بر نتایج عددی خوب شرایط همگرایی سراسری را نیز داشته باشد. قابل ذکر است که اگر تابع f یک تابع درجه دوم محدب قوی باشد و جستجوی خطی دقیق به کار رود، روش‌های معرفی‌شده در بالا همگی معادل هستند.

دای و لیائو در [۸] پارامتری را به صورت زیر معرفی کرده‌اند

$$\beta_k^{DL} = \frac{g_{k+1}^T y_k - \tau g_{k+1}^T s_k}{y_k^T d_k}, \tag{۸}$$

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

که در آن $\tau > 0$ عدد ثابت و $s_k := x_{k+1} - x_k$ است. با تغییر τ یک خانواده از روش‌های گرادیان مزدوج معرفی می‌شود که برای توابع کلی همگرایی سراسری را دارند. جهت‌های جستجوی تولیدشده توسط این پارامتر در شرط مزدوج بودن متفاوتی به شکل $d_{k+1}^T y_k = -\tau g_{k+1}^T s_k$ صدق می‌کنند، اما شرط کاهش کافی یعنی

$$d_k^T g_k \leqslant -c\|g_k\|^2, \qquad (۹)$$

که $c > 0$ یک عدد ثابت است، لزوما برقرار نیست. انتخاب مناسب τ سبب نتایج بسیار خوبی می‌شود اما یک مقدار بهینه، که برای هر نوع مساله‌ای به خوبی عمل کند به صورت نظری برای آن ارایه نشده است. تلاش‌های بسیاری به منظور بهبود این پارامتر انجام شده است که می‌توان برخی از آنها را در مقالات [۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴] مشاهده کرد.

فاطمی در [۱۴] یک روش گرادیان مزدوج جدید معرفی کرده است که همگرایی سراسری داشته و جهت جستجوی تولید شده در آن در شرط کاهشی کافی صدق می‌کند. این روش با استفاده از یک مساله بهینه‌سازی و به‌کارگیری شرط سکانت تولیدشده و عملکرد قابل توجهی دارد.

در این کار با معرفی شرط سکانت اصلاح‌شده قصد داریم یک پارامتر گرادیان مزدوج کارا بر اساس حل یک مساله بهینه‌سازی معرفی کنیم که ایده اولیه این مساله بهینه‌سازی توسط فاطمی در [۱۴] معرفی شده است. ما نشان خواهیم داد استفاده از شرط سکانت اصلاح‌شده می‌تواند اقدام موثری در افزایش کارایی روش باشد. نتایج عددی ارایه‌شده در بخش آخر با استفاده از نمودار عملکرد دولان-موره [۱۵] نشان می‌دهد که روش جدید کاملا کارآمد است. در بخش آخر با استفاده از مجموعه‌ای از مسایل CUTEr و انتخاب‌های متفاوت برای پارامترهای اصلی یک مقایسه دقیق بین روش اصلی و برخی دیگر از روش‌های گرادیان مزدوج از جمله روش فاطمی و سه روش از خانواده دای-لیائو که از شرط سکانت بهبودیافته استفاده کرده‌اند، انجام شده است و نشان داده می‌شود که روش جدید عملکرد قابل توجهی دارد.

در بخش دوم خلاصه‌ای از مفاهیم اولیه بیان می‌شود. جزییات روش جدید و الگوریتم آن در بخش سوم ارایه خواهد شد. در بخش ۴ همگرایی روش مورد بررسی قرار خواهد گرفت. در بخش ۵ عملکرد عددی الگوریتم معرفی‌شده با حل مجموعه‌ای از مسایل CUTEr و انتخاب‌های گوناگون برای پارامترهای اصلی با استفاده از نمودار عملکرد دولان-موره مورد ارزیابی قرار می‌گیرد.

۲ معرفی الزامات روش

روش‌های گرادیان مزدوج معمولا از شرط سکانت

$$B_{k+1}s_k = y_k, \qquad (۱۰)$$

که در آن B_{k+1} تقریبی برای $G_{k+1} := \nabla^2 f(x_{k+1})$ است، به منظور بهره‌گیری از شرایط همگرایی زبرخطی همانند روش‌های شبه نیوتنی، استفاده می‌کنند. توجه داشته باشید که در این شرط تنها از گرادیان تابع استفاده

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

می‌شود. محققان تلاش‌های بسیاری به منظور بهبود (۱۰) انجام داده‌اند. ژنگ و همکارانش در [۱۶] و ژنگ و خی در [۱۷] شرط سکانت بهبودیافته را به صورت زیر معرفی کرده‌اند

$$B_{k+1}s_k = \hat{y}_k, \quad \hat{y}_k = y_k + \frac{\theta_k}{s_k^T u_k}u_k \tag{۱۱}$$

که

$$\theta_k = 6(f_k - f_{k+1}) + 3(g_k + g_{k+1})^T s_k \tag{۱۲}$$

برای هر بردار u_k با $s_k^T u_k \neq 0$ و در نظر گرفتن B_{k+1} به عنوان تقریبی برای G_{k+1}. در [۱۸] نشان داده شده است که برای تابع درجه دوم محدب، با جستجوی خطی دقیق برای (۱۲) داریم $\theta = 0$ و در نتیجه شرط سکانت بهبودیافته معادل است با شرط سکانت (۱۰).

در شرط سکانت بهبودیافته علاوه بر بردار گرادیان از مقادیر تابع نیز استفاده می‌شود، و در [۱۷] ثابت شده است که $\hat{y}_k$ تقریب بهتری برای $\nabla^2 f(x_{k+1})s_k$ نسبت به y_k به‌دست می‌آورد.

برخی انتخاب‌های شناخته‌شده برای u_k که در شرط $s_k^T u_k \neq 0$ صدق می‌کنند عبارتند از $u_k = s_k$، $u_k = y_k$ و $u_k = \nabla f(x_k)$.

مقالات بسیاری از جمله [۱۳، ۱۹، ۲۰، ۲۱، ۲۲، ۲۳، ۲۴، ۲۵] از شرط (۱۱) به منظور ساخت و بهبود روش‌های جدید گرادیان مزدوج استفاده کرده‌اند که از لحاظ تئوری و عددی نتایج بهتری را به‌دست آورده‌اند.

به منظور ارایه یک روش گرادیان مزدوج کارا، فاطمی در [۱۴] یک مساله بهینه‌سازی با ترکیب سه شرط (۶)، (۷) و (۹) به صورت زیر معرفی کرده است

$$\min_{\beta_k}\left[g_{k+1}^T d_{k+1} + M\left(\left(g_{k+2}^T s_k\right)^2 + \left(d_{k+1}^T y_k\right)^2\right)\right], \tag{۱۳}$$

که در آن M پارمتر جریمه‌ای است. با حل این مساله و استفاده از شرط سکانت (۱۰) یک پارامتر β_k جدید به صورت زیر حاصل شده است

$$\beta_k = \frac{-1}{2M(1+t^2)}\frac{g_{k+1}^T d_k}{(y_k^T d_k)^2} + \frac{y_k^T g_{k+1}}{y_k^T d_k} - \frac{t}{(1+t^2)}\frac{s_k^T g_{k+1}}{y_k^T d_k}. \tag{۱۴}$$

۱) این روش با انتخاب مناسب پارامترها دارای همگرایی سراسری برای توابع کلی بوده و مقایسات عددی ارایه‌شده توسط نویسنده بیانگر کارایی روش بر اساس نمودارهای دولان-مووره است.

۳ فرمول جدید برای β_k

یابه و تاکانو در [۱۳] یک توسیع از شرط سکانت اصلاح‌شده (۱۱) به صورت زیر ارایه داده‌اند

$$z_k = y_k + \rho_k \frac{\theta_k}{s_k^T u_k}u_k, \tag{۱۵}$$

که در آن $\rho_k \geq 0$ یک اسکالر است. شرط سکانت بهبودیافته را در نظر بگیرید

$$B_{k+1}s_k = z_k, \tag{۱۶}$$

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

اگر $\rho_k = ۰$ یا $\rho_k = ۱$، این شرط به ترتیب با (۱۰) و (۱۱) معادل می‌شود.

تساوی زیر با حل (۱۳) و با استفاده از $g_{k+۲} \approx tB_{k+۱}d_{k+۱} + g_{k+۱}$ به‌دست می‌آید [۱۴]:

$$\beta_k = \frac{-g_{k+۱}^T d_k + ۲Mt^۲(s_k^T B_{k+۱} g_{k+۱})(s_k^T B_{k+۱} d_k)}{۲Mt^۲(s_k^T B_{k+۱} d_k)^۲ + ۲M(y_k^T d_k)^۲} + \frac{-۲Mt(s_k^T g_{k+۱})(s_k^T B_{k+۱} d_k) + ۲M(y_k^T g_{k+۱})(y_k^T d_k)}{۲Mt^۲(s_k^T B_{k+۱} d_k)^۲ + ۲M(y_k^T d_k)^۲}, \quad (۱۷)$$

که در آن $t > ۰$ تقریبی از $\alpha_{k+۱}$ است. این پارامتر احتمال برقراری شرط‌های (۶)، (۷) و (۹) را افزایش داده و در صورتی که M به سمت بینهایت میل کند به یک پارامتر دای-لیائو تبدیل می‌شود[۱۴]. حال با جایگذاری (۱۶) در (۱۷) پارامتر جدید β_k به صورت زیر به‌دست می‌آید:

$$\beta_k = \frac{-g_{k+۱}^T d_k + ۲Mt^۲(z_k^T g_{k+۱})(z_k^T d_k)}{۲Mt^۲(z_k^T d_k)^۲ + ۲M(y_k^T d_k)^۲} + \frac{-۲t(s_k^T g_{k+۱})(z_k^T d_k) + ۲(y_k^T g_{k+۱})(y_k^T d_k)}{۲t^۲(z_k^T d_k)^۲ + ۲(y_k^T d_k)^۲} \quad (۱۸)$$

در این مرحله هدف ما معرفی یک مقدار مناسب برای پارامتر جریمه‌ای M است به‌طوری‌که شرط‌های زیر برقرار باشند:

- جهت جستجوی $d_{k+۱}$ در شرط کاهش کافی (۹) صدق کند.
- در روند تکرارها به سمت بینهایت میل کند.

در لم زیر، یک پارامتر جریمه‌ای مناسب را پیشنهاد خواهیم کرد.

لم ۱: روش گرادیان مزدوج معرفی‌شده در روابط (۲) و (۵) با جستجوی خطی وولف قوی (۳) و (۴) و پارامتر β_k ارایه‌شده در (۱۸) را در نظر بگیرید. آنگاه، برای برخی اسکالرهای مثبت $\gamma_۱$ و $\gamma_۲$ به طوری‌که $\gamma_۱ + \gamma_۲ < ۱$ ، و ρ_k که با شرط $۰ \le \rho_k < \frac{۱ - c_۲}{۳(۱ + c_۲ - ۲c_۱)}$ انتخاب شده است داریم

$$d_{k+۱}^T g_{k+۱} \leqslant -(۱ - \gamma_۱ - \gamma_۲)\|g_{k+۱}\|^۲, \quad (۱۹)$$

که در آن

$$(۱ - t) \le \frac{\gamma_۱ (y_k^T s_k)^۲}{(z_k^T s_k)\|s_k\|^۲}, \quad (۲۰)$$

و

$$M = \frac{۲\gamma_۲(t^۲(z_k^T s_k)^۲ + (y_k^T s_k)^۲)}{\left\|t^۲(z_k^T s_k) z_k + (y_k^T s_k) y_k - \lambda_k (z_k^T s_k) s_k\right\|^۲}, \quad (۲۱)$$

که $\lambda_k \le ۱$ یک اسکالر دلخواه است.

اثبات: برای اثبات از استقرای ریاضی استفاده می‌کنیم.

به وضوح، از (۵) داریم:

$$d_۰^T g_۰ = -\|g_۰\|^۲ \leqslant -(۱ - \gamma_۱ - \gamma_۲)\|g_۰\|^۲.$$

فرض می‌کنیم

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

$$d_k^T g_k \leqslant -(1-\gamma_1-\gamma_2)\|g_k\|^2 . \qquad (22)$$

ابتدا به بررسی مقدار $z_k^T s_k$ می‌پردازیم. با استفاده از (3)، (4)، (12) و (15) داریم

$$\begin{aligned} z_k^T s_k &= y_k^T s_k + \rho_k \theta_k \\ &= y_k^T s_k + 6\rho_k (f_k - f_{k+1}) + 3\rho_k (g_k + g_{k+1})^T s_k \\ &\geqslant (1+3\rho_k) g_{k+1}^T s_k + (3\rho_k - 1) g_k^T s_k - 6\rho_k c_1 g_k^T s_k \\ &= (1+3\rho_k) g_{k+1}^T s_k + (3\rho_k - 6\rho_k c_1 - 1) g_k^T s_k \\ &\geqslant (1+3\rho_k) c_2 g_k^T s_k + (3\rho_k - 6\rho_k c_1 - 1) g_k^T s_k \\ &= \left(3\rho_k (1 + c_2 - 2c_1) + (c_2 - 1)\right) g_k^T s_k . \end{aligned}$$

حال با استفاده از

$$g_k^T s_k = \alpha_k g_k^T d_k < 0,$$

و با توجه به این‌که $0 \leq \rho_k < \dfrac{1-c_2}{3(1+c_2-2c_1)}$ ، ثابت $m_1 > 0$ وجود دارد به طوری‌که

$$z_k^T s_k \geqslant -m_1 (g_k^T s_k) > 0 . \qquad (23)$$

و در نتیجه

$$z_k^T d_k > 0 . \qquad (24)$$

اکنون نشان می‌دهیم که شرط (19) برقرار است. با استفاده از (5) و (18)، داریم

$$\begin{aligned} d_{k+1}^T g_{k+1} &= -\|g_{k+1}\|^2 + \beta_k d_k^T g_{k+1} = -\|g_{k+1}\|^2 - \frac{(s_k^T g_{k+1})^2}{2M\left[t^2 (z_k^T s_k)^2 + (y_k^T s_k)^2\right]} \\ &\quad + \frac{\left[t^2 (z_k^T g_{k+1})(z_k^T s_k) + (y_k^T g_{k+1})(y_k^T s_k)\right](s_k^T g_{k+1})}{t^2 (z_k^T s_k)^2 + (y_k^T s_k)^2} - \frac{t (z_k^T s_k)(s_k^T g_{k+1})^2}{t^2 (z_k^T s_k)^2 + (y_k^T s_k)^2} . \end{aligned}$$

چون $\lambda_k \leq 1$ ، داریم

$$\begin{aligned} d_{k+1}^T g_{k+1} \pm \frac{(s_k^T g_{k+1})^2 (z_k^T s_k)}{t^2 (z_k^T s_k)^2 + (y_k^T s_k)^2} &\leq -\|g_{k+1}\|^2 - \frac{(s_k^T g_{k+1})^2}{2M\left[t^2 (z_k^T s_k)^2 + (y_k^T s_k)^2\right]} \\ &\quad + \frac{\left[t^2 (z_k^T s_k)(z_k^T g_{k+1}) + (y_k^T s_k)(y_k^T g_{k+1}) - \lambda_k (z_k^T s_k)(s_k^T g_{k+1})\right](s_k^T g_{k+1})}{t^2 (z_k^T s_k)^2 + (y_k^T s_k)^2} \\ &\quad + \frac{(1-t)(z_k^T s_k)(s_k^T g_{k+1})^2}{t^2 (z_k^T s_k)^2 + (y_k^T s_k)^2} \end{aligned}$$

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

$$\Rightarrow d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 - \frac{(s_k^T g_{k+1})^2}{2M\left[t^2(z_k^T s_k)^2 + (y_k^T s_k)^2\right]}$$
$$+ \frac{\left[t^2(z_k^T s_k)z_k + (y_k^T s_k)y_k - \lambda_k(z_k^T s_k)s_k\right]^T g_{k+1}(s_k^T g_{k+1})}{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2}$$
$$+ \frac{(1-t)(z_k^T s_k)(s_k^T g_{k+1})^2}{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2}.$$

در اینجا با استفاده از نامساوی $ab \le \frac{l}{4}a^2 + \frac{1}{l}b^2$ که در آن a، b و l اسکالرهای مثبت هستند، داریم

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 - \frac{(s_k^T g_{k+1})^2}{2M\left[t^2(z_k^T s_k)^2 + (y_k^T s_k)^2\right]} + \frac{(1-t)(z_k^T s_k)(s_k^T g_{k+1})^2}{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2}$$
$$+ \frac{l}{4}\left(\left[t^2(z_k^T s_k)z_k + (y_k^T s_k)y_k - \lambda_k(z_k^T s_k)s_k\right]^T g_{k+1}\right)^2 + \frac{1}{l}\left(\frac{(s_k^T g_{k+1})}{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2}\right)^2.$$

قرار می‌دهیم $l = \frac{2M}{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2}$،

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 + \frac{(1-t)(z_k^T s_k)(s_k^T g_{k+1})^2}{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2}$$
$$+ \frac{M}{2(t^2(z_k^T s_k)^2 + (y_k^T s_k)^2)}\left(\left[t^2(z_k^T s_k)z_k + (y_k^T s_k)y_k - \lambda_k(z_k^T s_k)s_k\right]^T g_{k+1}\right)^2.$$

در آخر، با استفاده از نامساوی کوشی-شوارتز داریم

$$d_{k+1}^T g_{k+1} \le -[1 - \frac{(1-t)(z_k^T s_k)\|s_k\|^2}{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2} -$$
$$\frac{M}{2\left[t^2(z_k^T s_k)^2 + (y_k^T s_k)^2\right]}\left\|t^2(z_k^T s_k)z_k + (y_k^T s_k)y_k - \lambda_k(z_k^T s_k)s_k\right\|^2]\|g_{k+1}\|^2.$$

با استفاده از (۲۰)، (۲۱) و (۲۴) اثبات کامل می‌شود.

به منظور سرعت بخشیدن به نرخ رشد M که دومین هدف ما است، مینیمم‌سازی مخرج کسر در رابطه (۲۱) نتیجه می‌دهد که

$$\lambda_k = \min\left\{1, \frac{t^2(z_k^T s_k)^2 + (y_k^T s_k)^2}{(z_k^T s_k)\|s_k\|^2}\right\} \qquad (۲۵)$$

توجه۱: بر خلاف روش نیوتن که $\alpha_k = 1$ یک همگرایی درجه دوم را نتیجه می‌دهد، در الگوریتم‌های گرادیان مزدوج یک مقدار طول گام مناسب، قابل پیش‌بینی نیست و روند زیر توصیه می‌شود:

$$t = \min\{t_1, T\} \qquad (۲۶)$$

که در آن

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

$$t_1=\begin{cases}\alpha_k & if\ (1-\alpha_k)\leqslant\dfrac{\gamma_1(y_k^Ts_k)^2}{(z_k^Ts_k)\|s_k\|^2}\\ \left|1-\dfrac{\gamma_1(y_k^Ts_k)^2}{(z_k^Ts_k)\|s_k\|^2}\right| & o.w.\end{cases}$$

و $T>0$ یک عدد ثابت دلخواه بزرگ است.

در نهایت، با جایگذاری (۲۱) در (۱۸) داریم:

$$\beta_k=\frac{(y_k^Tg_{k+1})(y_k^Td_k)}{t^2(z_k^Td_k)^2+(y_k^Td_k)^2}-\frac{t(s_k^Tg_{k+1})(z_k^Td_k)}{t^2(z_k^Td_k)^2+(y_k^Td_k)^2}+\frac{t^2(z_k^Tg_{k+1})(z_k^Td_k)}{t^2(z_k^Td_k)^2+(y_k^Td_k)^2}$$
$$-\frac{\left\|t^2(z_k^Ts_k)z_k+(y_k^Ts_k)y_k-\lambda_k(z_k^Ts_k)s_k\right\|^2}{4\gamma_2\left[t^2(z_k^Ts_k)^2+(y_k^Ts_k)^2\right]}\frac{g_{k+1}^Td_k}{t^2(z_k^Td_k)^2+(y_k^Td_k)^2} \qquad (۲۷)$$

حال با جمع بندی آنچه گفته شد الگوریتم گرادیان مزدوج جدید خود را معرفی می‌کنیم.

الگوریتم ۱: الگوریتم گرادیان مزدوج سکانت بهبودیافته (MSCG)

گام ۱: نقطه اولیه $x_0\in R^n$ و مقادیر مناسب برای پارامترهای نامنفی ρ_k، T، $0<c_1<c_2<1$ و $\gamma_1+\gamma_2<1$ را انتخاب کنید. $g_0=\nabla f(x_0)$ را محاسبه کرده و قرار دهید $d_0=-g_0$ و $k=0$.

گام ۲: شرط توقف را بررسی‌کرده و در صورت برقراری متوقف شوید، در غیراین‌صورت به گام ۳ بروید.

گام ۳: با استفاده از شرایط وولف قوی (۳) و (۴) طول گام α_k را محاسبه کنید.

گام ٤: قرار دهید $x_{k+1}=x_k+\alpha_kd_k$، $f_{k+1}=f(x_{k+1})$، $g_{k+1}=\nabla f(x_{k+1})$، $s_k=x_{k+1}-x_k$ و $y_k=g_{k+1}-g_k$.

گام ٥: بردار مناسب $u_k\in R^n$ را به گونه‌ای انتخاب کرده که $s_k^Tu_k\neq0$ و سپس θ_k و z_k را به ترتیب با استفاده از روابط (۱۲) و (۱۵) به‌دست آورید.

گام ٦: λ_k و t را با استفاده از (۲۵) و (۲۶) حساب کنید.

گام ۷: پارامتر گرادیان مزدوج β_k را با استفاده از (۲۷) به‌دست آورید.

گام ۸: جهت جستجوی $d_{k+1}=-g_{k+1}+\beta_kd_k$ را حساب کنید.

گام ۹: قرار دهید $k=k+1$ و به گام ۲ بروید.

هدف نهایی ما بررسی شرایط همگرایی سراسری برای الگوریتم ۱ (MSCG) است.

٤ همگرایی سراسری

برای تضمین همگرایی نیاز است مطابق آنچه در ادبیات موضوع مرسوم است تغییراتی جزیی را به صورت زیر در پارامتر β_k ایجاد کنیم، و پارامتر را به صورت زیر تعریف می‌کنیم

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

$$\beta_k^N = \max(\beta_k, \chi_k) \quad (۲۸)$$

که در آن

- χ_k یک تابع حقیقی مقدار است،
- $|\chi_k| \|d_k\|$ از بالا کراندار است،
- برای برخی $\varepsilon < ۱$ و $g_{k+۱}^T d_k < ۰$ داریم $\chi_k \leqslant \frac{\varepsilon \|g_{k+۱}\|^۲}{g_{k+۱}^T d_k}$.

قبل از بررسی همگرایی قضیه زیر را بیان می‌کنیم.

قضیه ۱: فرض کنید، f یک تابع از پایین کراندار است و تابع گرادیان $g(x)$ بر روی مجموعه کراندار زیر پیوسته‌ی لیپ شیتز می‌باشد،

$$V = \{x \in R^n : f(x) \leq f(x_0)\}, \quad (۲۹)$$

یعنی، $L > ۰$ وجود دارد به طوری‌که

$$\|g(x) - g(y)\| \leq L \|x - y\| \quad (۳۰)$$

به‌علاوه، دنباله‌ی حاصل از تکرارهای الگوریتم یعنی x_k، کراندار است. برای هر روش کاهشی مانند (۲) که در آن α_k با استفاده از جستجوی خطی وولف قوی محاسبه می‌شود، داریم:

$$\sum_{k=۱}^{\infty} \frac{(g_k^T d_k)^۲}{\|d_k\|^۲} < \infty \quad (۳۱)$$

اثبات: به مقاله[۶] مراجعه شود.

حال با استفاده از قضیه ۱ همگرایی سراسری الگوریتم ۱ را اثبات می‌کنیم.

قضیه ۲: الگوریتم MSCG با β_k^N معرفی‌شده در (۲۸) را در نظر بگیرید، آنگاه برای تابع دلخواه f با جستجوی خطی وولف قوی با شرایط (۳) و (۴) و $۰ \leq \rho_k < \min\left\{\frac{۱-c_۲}{۳(۱+c_۲-۲c_۱)}, \mathrm{m}\, y_k^T s_k\right\}$ که m ثابت بزرگ و مثبتی است، داریم

$$\liminf_{k\to\infty} \|g_k\| = ۰$$

اثبات: از لم ۱ و با روشی مشابه اثبات لم ۴ در [۱۴] داریم

$$d_{k+۱}^T g_{k+۱} \leqslant -\max(۱-\varepsilon, ۱-\gamma_۱-\gamma_۲)\|g_{k+۱}\|^۲ . \quad (۳۲)$$

اگر برای یک زیر دنباله k_j، داشته باشیم $\beta_{k_j}^N = \chi_{k_j}$، آنگاه با توجه به کرانداری $|\chi_k| \|d_k\|$ برای $\varepsilon > ۰$ داریم

$$\|d_{k_j+۱}\| = \|-g_{k_j+۱} + \beta_{k_j}^N d_{k_j}\| \leq \|g_{k_j+۱}\| + |\chi_{k_j}| \|d_{k_j}\| \leq \varepsilon$$

کران حاصل برای $d_{k_j+۱}$ و (۳۲) با استفاده از (۳۱) سبب به وجود آمدن تناقض شده و اثبات کامل می‌شود.

اگر برای k به اندازه کافی بزرگ، داشته باشیم $\beta_k^N = \beta_k$، آنگاه یک کران روی β_k به‌صورت زیر حاصل می‌شود.

ابتدا، با استفاده از (۴) و (۱۹) داریم

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

$$d_k^T y_k = d_k^T g_{k+1} - d_k^T g_k \geq (c_2 - 1) d_k^T g_k > 0, \tag{۳۳}$$

فرض کنید $\eta_1 > 0$ وجود دارد به طوری که $\|g_k\| > \eta_1$، آنگاه با استفاده از (۳۲) و (۳۳) داریم

$$d_k^T y_k \geq (1 - c_2) \max(1 - \varepsilon, 1 - \gamma_1 - \gamma_2) \eta_1^2, \tag{۳۴}$$

و

$$\frac{|d_k^T g_{k+1}|}{|d_k^T y_k|} \leq \max\left(\frac{c_2}{1 - c_2}, 1\right), \tag{۳۵}$$

(به [۲۵] مراجعه شود). از طرف دیگر، (۹) و (۲۳) نتیجه می‌دهند که

$$d_k^T z_k \geq -m_1 (d_k^T g_k) \geq m_1 c \|g_k\|^2 \geq m_1 c \eta_1^2. \tag{۳۶}$$

سپس، نشان می‌دهیم که ثابت $R > 0$ وجود دارد به طوری که

$$|\beta_k| \leq R \|s_k\|. \tag{۳۷}$$

به وضوح داریم

$$t^2 (z_k^T d_k)^2 + (y_k^T d_k)^2 \geq t^2 (z_k^T d_k)^2 \tag{۳۸}$$

و

$$t^2 (z_k^T d_k)^2 + (y_k^T d_k)^2 \geq (y_k^T d_k)^2 \tag{۳۹}$$

با استفاده از (۲۷)، (۳۸)، (۳۹) و $s_k = \alpha_k d_k$ داریم

$$\begin{aligned} |\beta_k| \leq{}& \frac{|y_k^T g_{k+1}|\,|y_k^T d_k|}{(y_k^T d_k)^2} + \frac{|t|\,\|s_k^T g_{k+1}\|\,|z_k^T d_k|}{(y_k^T d_k)^2} + \frac{t^2\,|z_k^T g_{k+1}|\,|z_k^T d_k|}{t^2 (z_k^T d_k)^2} \\ &+ \frac{|g_{k+1}^T d_k|}{4\gamma_2 (t^2 (z_k^T d_k)^2 + (y_k^T d_k)^2)} \Big(\frac{\|t^2 (z_k^T s_k) z_k\|^2}{t^2 (z_k^T s_k)^2} + \frac{\|(y_k^T s_k) y_k\|^2}{(y_k^T s_k)^2} + \frac{\|(z_k^T s_k) s_k\|^2}{(y_k^T s_k)^2} \\ &+ \frac{2\|t^2 (z_k^T s_k) z_k\| \cdot \|(y_k^T s_k) y_k\|}{t^2 (z_k^T s_k)^2} + \frac{2\|t^2 (z_k^T s_k) z_k\| \cdot \|(z_k^T s_k) s_k\|}{t^2 (z_k^T s_k)^2} + \frac{2\|(y_k^T s_k) y_k\| \cdot \|(z_k^T s_k) s_k\|}{(y_k^T s_k)^2} \Big) \end{aligned}$$

$$\begin{aligned} \Rightarrow |\beta_k| \leq{}& \frac{|y_k^T g_{k+1}|}{y_k^T d_k} + \frac{|t|\,\|s_k^T g_{k+1}\|\,|z_k^T d_k|}{(y_k^T d_k)^2} + \frac{|z_k^T g_{k+1}|}{z_k^T d_k} \\ &+ \frac{1}{4\gamma_2} \Big(\|z_k\|^2 \times \frac{|g_{k+1}^T d_k|}{(z_k^T d_k)^2} + \|y_k\|^2 \times \frac{|g_{k+1}^T d_k|}{(y_k^T d_k)^2} + \frac{(z_k^T s_k)^2 \|s_k\|^2}{(y_k^T s_k)^2} \times \frac{|g_{k+1}^T d_k|}{(y_k^T d_k)^2} \\ &+ \frac{2\|z_k\| \cdot \|y_k\|}{(z_k^T d_k)} \times \frac{|g_{k+1}^T d_k|}{(y_k^T d_k)} + 2\|z_k\| \cdot \|s_k\| \times \frac{|g_{k+1}^T d_k|}{(y_k^T d_k)^2} + \frac{2(z_k^T s_k) \cdot \|y_k\| \cdot \|s_k\|}{(y_k^T s_k)} \times \frac{|g_{k+1}^T d_k|}{(y_k^T d_k)^2} \Big) \end{aligned}$$

[Downloaded from jorlu.ir on 2026-11-28]

[DOR: 20.1001.1.22517286.2022.19.1.5.5]

[DOI: 10.52547/jamlu.19.1.1]

$$\Rightarrow |\beta_k| \le \frac{1}{y_k^T d_k}\Big(|y_k^T g_{k+1}| + |t|\,\|s_k^T g_{k+1}|\,\frac{|z_k^T s_k|}{(y_k^T s_k)} + \frac{\|y_k\|^2}{4\gamma_2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k}$$
$$+ \frac{\|s_k\|^2}{4\gamma_2}\times\frac{(z_k^T s_k)^2}{(y_k^T s_k)^2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k} + \frac{\|z_k\|\cdot\|s_k\|}{2\gamma_2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k} + \frac{\|y_k\|\cdot\|s_k\|}{2\gamma_2}\times\frac{|z_k^T s_k|}{(y_k^T s_k)}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k}\Big) \quad (40)$$
$$+ \frac{1}{z_k^T d_k}\Big(|z_k^T g_{k+1}| + \frac{\|z_k\|^2}{4\gamma_2}\times\frac{|g_{k+1}^T d_k|}{z_k^T d_k} + \frac{\|z_k\|\cdot\|y_k\|}{2\gamma_2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k}\Big)$$

اکنون با استفاده از قضیه مقدار میانگین کران بالای θ_k در رابطه (۱۵) را به‌دست می‌آوریم

$$\begin{aligned}\theta_k &= 6(f_k - f_{k+1}) + 3(\nabla f_k + \nabla f_{k+1})^T s_k \\ &= 6\nabla f(\eta_k)^T (x_k - x_{k+1}) + 3(\nabla f_k + \nabla f_{k+1})^T s_k \\ &= -6\nabla f(\eta_k)^T s_k + 3(\nabla f_k + \nabla f_{k+1})^T s_k \\ &= 3\{\nabla f_k - \nabla f(\eta_k) + \nabla f_{k+1} - \nabla f(\eta_k)\}^T s_k\end{aligned}$$

که در اینجا $\eta_k = \omega x_k + (1-\omega)x_{k+1}$ و $\omega \in (0,1)$. در نتیجه با استفاده از (۳۰) داریم

$$\begin{aligned}|\theta_k| &\le 3(\|\nabla f_k - \nabla f(\eta_k)\| + \|\nabla f_{k+1} - \nabla f(\eta_k)\|)\|s_k\| \\ &\le 3(L\|x_k - \eta_k\| + L\|x_{k+1} - \eta_k\|)\|s_k\| \\ &= 3(L(1-\omega)\|x_k - x_{k+1}\| + L\omega\|x_{k+1} - x_k\|)\|s_k\| \\ &= 3L\|s_k\|^2.\end{aligned} \quad (41)$$

حال با استفاده از فرض قضیه و همچنین روابط (۱۵)، (۳۰) و (۴۱) داریم

$$\left|\frac{z_k^T s_k}{y_k^T s_k}\right| = \left|\frac{y_k^T s_k + \rho_k \frac{\theta_k}{s_k^T u_k} u_k^T s_k}{y_k^T s_k}\right| = \left|1 + \frac{\rho_k \theta_k}{y_k^T s_k}\right| < 1 + |m\theta_k| < 1 + 3Lm\|s_k\|^2. \quad (42)$$

و

$$\|z_k\| = \left\|y_k + \rho\frac{\theta_k}{s_k^T u_k}u_k\right\| \le \|y_k\| + m|\theta_k|\|u_k\| \le L\|s_k\| + 3Lm\|s_k\|^2\|u_k\|,$$

در نامساوی فوق با توجه به کراندار بودن $\|s_k\|$ و $\|y_k\|$ با هر یک از انتخاب‌های $u_k = s_k$ یا $u_k = y_k$ می‌توان فرض کرد ثابت $L' > 0$ وجود دارد به طوری‌که

$$\|z_k\| \le L'\|s_k\| \quad (43)$$

در ادامه با استفاده از (۲۶)، (۳۰)، (۴۲)، (۴۳)، نامساوی کوشی-شوارتز و این فرض که η_2 یک کران بالا برای $\|g_k\|$ است، داریم

$$|\beta_k| \le \frac{1}{y_k^T d_k}\Big(L\eta_2\|s_k\| + T\eta_2(1+3Lm\|s_k\|^2)\|s_k\| + \frac{L^2\|s_k\|^2}{4\gamma_2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k}$$
$$+ \frac{\|s_k\|^2}{4\gamma_2}\times(1+3Lm\|s_k\|^2)^2\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k} + \frac{L'\|s_k\|^2}{2\gamma_2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k} + \frac{L\|s_k\|^2}{2\gamma_2}\times(1+3Lm\|s_k\|^2)\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k}\Big) \quad (44)$$
$$+ \frac{1}{z_k^T d_k}\Big(L'\eta_2\|s_k\| + \frac{L^2\|s_k\|^2}{4\gamma_2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k} + \frac{LL'\|s_k\|^2}{2\gamma_2}\times\frac{|g_{k+1}^T d_k|}{y_k^T d_k}\Big).$$

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

از روابط (۳۴)، (۳۵)، (۳۶)،

$$D=(1-c_2)\max(1-\varepsilon,1-\gamma_1-\gamma_2)\eta_1^2,\ W=\max\left(\frac{c_2}{1-c_2},1\right)$$

و کرانداری دنباله تکرارها می‌توان ثابت $R>0$ در رابطه (۳۷) را به صورت زیر معرفی کرد

$$R=\frac{1}{D}\Big(L\eta_2+T\eta_2(1+3Lm\|s_k\|^2)+\frac{L^2\|s_k\|}{4\gamma_2}\times W$$
$$+\frac{\|s_k\|}{4\gamma_2}\times(1+3Lm\|s_k\|^2)^2\times W+\frac{L'\|s_k\|}{2\gamma_2}\times W+\frac{L\|s_k\|}{2\gamma_2}\times(1+3Lm\|s_k\|^2)\times W\Big)$$
$$+\frac{1}{m_1c\eta_1^2}\Big(L'\eta_2+\frac{L^4\|s_k\|}{4\gamma_2}\times W+\frac{LL'\|s_k\|}{2\gamma_2}\times W\Big).$$

ادامه اثبات مشابه قضیه ۲-۳ در [۲۶] است.

۵ نتایج عددی و مقایسه‌ها

در این بخش نتایج عددی به‌دست آمده را گزارش می‌دهیم. در ادامه روند تکراری (۲) و (۵) را به ازای پنج مقدار متفاوت برای β_k در نظر می‌گیریم:

β_k^{New}: پارامتر ارایه‌شده در الگوریتم ۱ (MSCG) به ازای

$$\rho_k=0,0/1,0/2,0/3,0/4,0/5,0/6,0/7,0/8,0/9,1$$

β_k^F: پارامتر ارایه‌شده توسط فاطمی در [۱۴].

β_k^{YT}: پارامتر ارایه‌شده توسط یابه و تاکانو در [۱۳] به ازای ($\rho=0/5,t=0/3$) و ($\rho=0/7,t=0/5$).

β_k^{BG1}: پارامتر (۲۱. ۳) ارایه‌شده توسط بابایی کفاکی و همکارانش در [۲۲] با $t=0/05005$.

β_k^{BG2}: پارامتر (۵۳. ۳) ارایه‌شده توسط بابایی کفاکی و همکارانش در [۲۲] با $t=0/05010$.

تمام پارامترهای گرادیان مزدوج در نظرگرفته‌شده متعلق به خانواده‌ی دای-لیائو هستند و از شرط سکانت اصلاح شده استفاده کرده‌اند به جز پارامتر β_k^F.

همچنین موارد زیر در پیاده‌سازی الگوریتم‌ها در نظر گرفته شده است:

- از روند جستجوی خطی CG-DESCENT با پارامترهای اولیه ارایه‌شده در [۲۷] استفاده شده است.
- دو انتخاب برای u_k در نظر گرفته شده است: s_k و y_k.
- شرط توقف الگوریتم‌ها عبارت است از

$$\|\nabla f(x_k)\|_\infty\le\max\left(10^{-6},10^{-12}\|\nabla f(x_1)\|_\infty\right)$$

و یا تعداد استفاده بیش از ۳۰۰۰ بار از مقدار تابع.

- انتخاب $\gamma_1=0/01$ و $\gamma_2=0/98$.

به منظور بررسی و مقایسه کارایی الگوریتم‌ها، ۸۰ مساله از مجموعه CUTEr با ابعاد مختلف در نظر گرفته شده است که در جدول ۱ معرفی شده‌اند (به منظور مشاهده مسایل به [۲۸] مراجعه شود). از آنجایی که روش‌های

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

گرادیان مزدوج برای حل مسایل اندازه بزرگ بسیار کارآمد هستند، لذا سعی کرده‌ایم که بیشتر از این مسایل استفاده کنیم. علاوه بر آن برخی مسایل اندازه کوچک که حل آنها توسط بسیاری از الگوریتم‌ها دشوار بوده و ممکن است بدوضع باشند نیز در نظر گرفته شده‌اند.

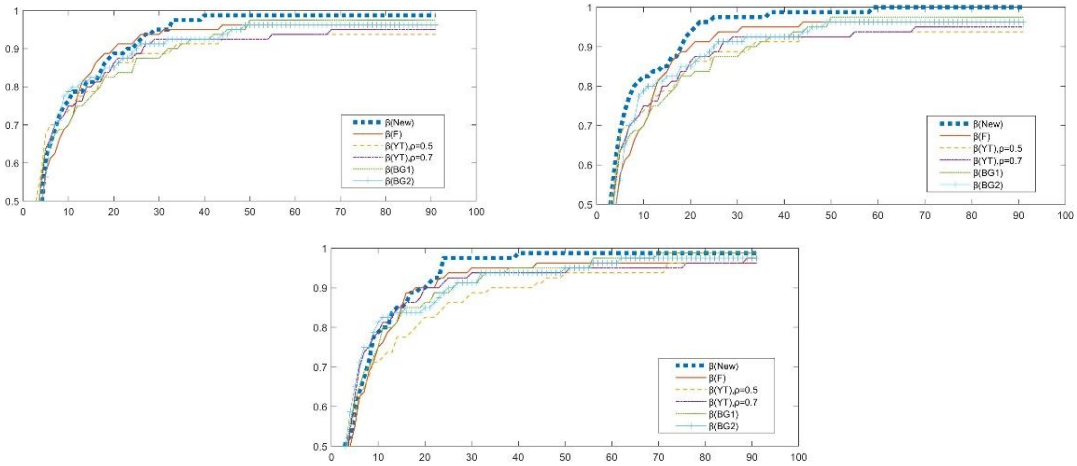
نتایج با استفاده از یک کامپیوتر با پردازنده 2.4 Intel Core(TM) i7-5500 و رم 8GB و سیستم عامل لینوکس در محیط برنامه برنامه‌نویسی MATLAB R2015a به‌دست آمده‌اند.

الگوریتم پیشنهادی ما تمام مسایل را با موفقیت حل کرده، در حالی که پارامترهای β_k^{YT}، β_k^{BG1} و β_k^{BG2} برای مساله DANWOODLS بدون جواب بوده‌اند. با مشاهده نتایج و بررسی نمودار عملکرد معرفی‌شده در [۱۵] برای تعداد محاسبات تابع، تعداد محاسبات گرادیان و زمان پردازنده در اجرای مسایل (CPU time)، در می‌یابیم که الگوریتم ۱ با پارامتر β_k^{New} برای بسیاری از انتخاب‌های ρ_k از سایر الگوریتم‌ها کارآمدتر است، به ویژه برای $u_k = s_k$. ما مقادیر $\rho_k = ۰/۷$ و $\rho_k = ۰/۸$ را به عنوان بهترین انتخاب‌ها برای الگوریتم معرفی می‌کنیم. تصاویر ۱ تا ۳، نمودارهای عملکرد را برای تمام الگوریتم‌ها نشان می‌دهند.

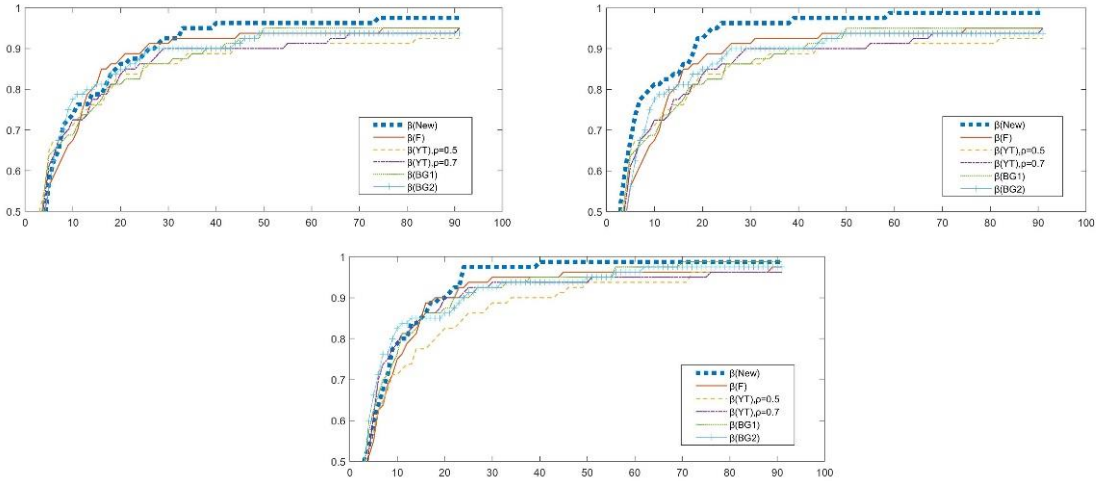
جدول ۱. مسایل

تابع	بعد	تابع	بعد	تابع	بعد	تابع	بعد
AKIVA	2	DANWOODLS	2	HILBERTA	2	QUARTC	5000
ALLINITU	4	DENSCHA	2	HILBERTB	10	ROSENBR	2
ARGLBLE	200	DENSCHB	2	HUMPS	2	S308	2
ARGLCLE	200	DIXMAANA	3000	INTEQNELS	12	SBRYBND	5000
ARGLINA	200	DIXMAANB	3000	JENSMP	2	SCHMVETT	5000
ARGLINB	200	DIXMAANC	3000	KOWOSB	4	SENSORS	100
ARGLINC	200	DIXMAAND	3000	KOWOSBNE	4	SINEVAL	2
ARWHEAD	5000	DIXMAANE	3000	LIARWHD	5000	SISSER	2
BARD	3	DIXMAANF	3000	LOGHAIRY	2	SPARSOUR	10000
BEALE	2	DIXMAANG	3000	LSCILS	3	SPMSRTLS	4999
BIGGS6	6	DIXMAANH	3000	MANCINO	100	SROSENBR	5000
BOX3	3	DJTL	2	MARATOSB	2	TOINTGOR	50
BROWNAL	200	DQDRTIC	5000	MGH09LS	4	TOINTGSS	5000
BROYDN3DLS	10	ECKERLE4LS	3	MGH10LS	3	TOINTQOR	50
BRYBND	500	EG2	1000	MODBEALE	2000	TQUARTIC	5000
CHAINWOO	4000	ENSOLS	9	MOREBV	5000	VARDIM	200
CHNROSNB	50	EXTROSNB	1000	NONDIA	5000	VAREIGVL	50
CHNRSNBM	50	FMINSURF	5625	PENALTY1	1000	WOODS	4000
COSINE	10000	GENROSE	500	POWELLSG	5000	YFITU	3
DANIWOODLS	2	HAIRY	2	POWER	10000	ZANGWIL2	2

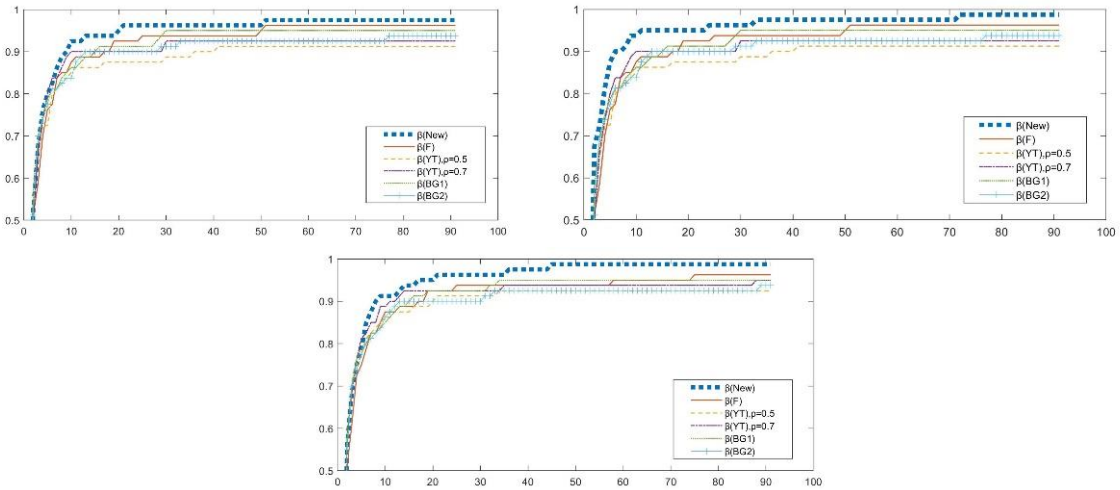
[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]



شکل ۱. نمودار عملکرد تعداد محاسبات تابع (بالا سمت راست: $\rho_k = ۰/۷$ و $u_k = s_k$، بالا سمت چپ: $\rho_k = ۰/۸$ و $u_k = s_k$، پایین: $\rho_k = ۰/۸$ و $u_k = y_k$)



شکل ۲. نمودار عملکرد تعداد محاسبات گرادیان (بالا سمت راست: $\rho_k = ۰/۷$ و $u_k = s_k$، بالا سمت چپ: $\rho_k = ۰/۸$ و $u_k = s_k$، پایین: $\rho_k = ۰/۸$ و $u_k = y_k$)



شکل ۳. نمودار عملکرد زمان پردازنده (بالا سمت راست: $\rho_k = ۰/۷$ و $u_k = s_k$، بالا سمت چپ: $\rho_k = ۰/۸$ و $u_k = s_k$، پایین: $\rho_k = ۰/۸$ و $u_k = y_k$)

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

٦ نتیجه‌گیری

نتایج قابل قبول حاصل از به‌کارگیری شرط سکانت بهبودیافته، ما را ترغیب کرد تا به ارایه یک روش گرادیان مزدوج کارا مبتنی بر این شرط بپردازیم. این روش علاوه بر همگرایی سراسری در شرط کاهش کافی صدق می‌کند. نتایج عددی نشان می‌دهند که با انتخاب پارامترهای مناسب این روش بسیار قابل قبول بوده و عملکرد بهتری نسبت به روش‌های دیگر به ویژه روش ارایه‌شده توسط فاطمی در [۱۴] و روش یابه و تاکانو در [۱۳] دارد. در ادامه پیشنهاداتی برای کارهای آتی ارایه می‌شود.

- بررسی کارایی روش در حل مسایل در حوزه‌های مهندسی و به‌ویژه پردازش تصویر.
- امکان‌سنجی تاثیر پارامترهای جدید در بهبود عملکرد روش و معرفی پارامترهای بهینه.
- بررسی تغییرات پارامترهای روش به جهت طراحی یک الگوریتم کارا برای حل مسایل ناهموار.

منابع

[1] Hestenes, M. R., Stiefel, E., (1952). Methods of conjugate gradients for solving linear systems. NBS Washington, DC, 49(1).

[2] Fletcher, R., Reeves, C. M., (1964). Function minimization by conjugate gradients. Comput. J., 7, 149-154.

[3] Polyak, B. T., (1969). The conjugate gradient method in extremal problems. USSR Comput. Math. Math. Phys., 9(4), 94-112.

[4] Fletcher, R., (2013). Practical methods of optimization. John Wiley & Sons.

[5] Liu, Y., Storey, C., (1991). Efficient generalized conjugate gradient algorithms, part 1: theory. 69(1), 129-137.

[6] Dai, Y. H., Yuan, Y., (1999). A nonlinear conjugate gradient method with a strong global convergence property. 10(1), 177-182.

[7] Andrei, N., (2007). Numerical comparison of conjugate gradient algorithms for unconstrained optimization. Stud. Informatics Control, 16(4), 333-352.

[8] Dai, Y.H., Liao, L.Z., (2001). New conjugacy conditions and related nonlinear conjugate gradient methods. Appl. Math. Optim., 43(1), 87-101.

[9] Babaie-Kafaki, S., Ghanbari, R., (2014). The Dai--Liao nonlinear conjugate gradient method with optimal parameter choices. Eur. J. Oper. Res., 234(3), 625-630.

[10] Babaie-Kafaki, S., Ghanbari, R., (2014). A descent family of Dai--Liao conjugate gradient methods. Optim. Methods Softw., 29(3), 583-591.

[11] Andrei, N., (2018). A Dai-Liao conjugate gradient algorithm with clustering of eigenvalues. Numer. Algorithms, 77(4), 1273-1282.

[12] Dai, Y. H., Kou, C. X.,(2013). A nonlinear conjugate gradient algorithm with an optimal property and an improved Wolfe line search. SIAM J. Optim., 23(1), 296-320.

[13] Yabe, H., Takano, M.,(2004). Global convergence properties of nonlinear conjugate gradient methods with modified secant condition. Comput. Optim. Appl., 28(2), 203-225.

[14] Fatemi, M., (2015). A new efficient conjugate gradient method for unconstrained optimization. J. Comput. Appl. Math., 300, 207-216.

[15] Dolan, E. D., Moré, J. J.,(2002). Benchmarking optimization software with performance profiles. Math. Program., 91(2), 201-213.

[16] Zhang, J. Z., Deng, N. Y., Chen, L. H., (1999). New quasi-Newton equation and related methods for unconstrained optimization. J. Optim. Theory Appl., 102(1), 147-167.

[17] Zhang, J., Xu, C., (2001). Properties and numerical performance of quasi-{N}ewton methods with modified quasi-{N}ewton equations. J. Comput. Appl. Math., 137(2), 269-278.

[18] Li, G., Tang, C., Wei, Z., (2007). New conjugacy condition and related new conjugate gradient methods for unconstrained optimization. J. Comput. Appl. Math., 202(2), 523-539.

[19] Babaei, S., Arazm, M., (2018). An extension of a three-term conjugate gradient method based on the objective function values with guaranteeing global convergence without convexity

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]

assumption. Journal of Operational Research and Its Applications., 15 (1), 19-28.
[20] Andrei, N., (2009). Accelerated conjugate gradient algorithm with modified secant condition for unconstrained optimization. Stud. Informatics Control, 18(3), 211-232.
[21] Babaie-Kafaki, S., (2014). Two modified scaled nonlinear conjugate gradient methods. J. Comput. Appl. Math., 261, 172-182.
[22] Babaie-Kafaki, S., Ghanbari, R., Mahdavi-Amiri, N., (2010). Two new conjugate gradient methods based on modified secant equations. J. Comput. Appl. Math., 234(5), 1374-1386.
[23] Babaie-Kafaki, S., Fatemi, M., Mahdavi-Amiri, M.,(2011). Two effective hybrid conjugate gradient algorithms based on modified BFGS updates. Numer. Algorithms, 58(3), 315-331.
[24] Narushima, Y., Yabe, H.,(2012). Conjugate gradient methods based on secant conditions that generate descent search directions for unconstrained optimization. J. Comput. Appl. Math., 236(17), 4303-4317.
[25] Akbari, Z., (2017). A Conic Trust Region Method for Locally Lipschitz Optimization Problems. Journal of Operational Research and Its Applications. 14(2), 95-109.
[26] Hager, W. W., Zhang, H.,(2005). A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim., 16(1), 170-192.
[27] Hager, W. W., Zhang, H.,(2006). Algorithm 851: {CG}_{DESCENT}, a conjugate gradient method with guaranteed descent. ACM Trans. Math. Softw., 32(1), 113-137.
[28] Gould, N. I. M., Orban, D. M., Toint, P. L.,(2003). CUTEr and SifDec: A constrained and unconstrained testing environment, revisited. ACM Trans. Math. Softw., 29(4), 373-394.

[Downloaded from jorlu.ir on 2026-11-28]
[DOR: 20.1001.1.22517286.2022.19.1.5.5]
[DOI: 10.52547/jamlu.19.1.1]