

## یک روش تندترین کاهش بدون جستجوی خطی برای حل مسائل بهینه‌سازی نامقید

نرگس بیدآبادی\*

استادیار، دانشکده علوم ریاضی، دانشگاه یزد، یزد، ایران

تاریخ ارسال مقاله: ۹۸/۰۴/۱۷ تاریخ پذیرش مقاله: ۹۹/۰۵/۱۲

### چکیده

در این مقاله به حل مسئله بهینه‌سازی نامقید با استفاده از یک روش تندترین کاهش بدون استفاده از الگوریتم‌های جستجوی خطی می‌پردازیم. ابتدا یک فرمول شبه نیوتن مقیاس‌بندی شده دو پارامتری برای محاسبه تقریبی از ماتریس هسی ارائه می‌دهیم. تقریب به دست آمده از این فرمول، یک ماتریس معین مثبت است که در رابطه سکانت استاندارد صدق می‌نماید. همچنین نشان می‌دهیم که بزرگترین مقدار ویژه این ماتریس از تعداد متغیرهای مسئله بیشتر نخواهد بود. سپس با استفاده از این فرمول شبه نیوتن مقیاس بندی شده دو پارامتری، فرمول صریحی برای محاسبه طول گام در روش تندترین کاهش ارائه می‌شود و بنابراین این روش نیازی به استفاده از روش‌های تقریبی برای محاسبه طول گام نخواهد داشت. نتایج عددی به دست آمده از اجرای الگوریتم در محیط نرم‌افزاری متلب بر روی برخی مسائل بهینه‌سازی ارائه شده است. این نتایج کارایی روش ارائه شده نسبت به سایر روش‌های موجود را نشان می‌دهد.

**واژه‌های کلیدی:** روش تندترین کاهش، جستجوی خطی، مسئله بهینه‌سازی نامقید، فرمول شبه نیوتن مقیاس‌بندی شده دو پارامتری.

۱- مقدمه

مسئله بهینه‌سازی نامقید زیر در نظر بگیرید:

$$\min_x f(x) \quad (1)$$

که در آن  $f$  تابعی به‌طور پیوسته مشتق‌پذیر از  $R^n$  به  $R$  است. الگوریتم‌های بهینه‌سازی برای حل این مسئله با شروع از یک نقطه دلخواه  $x^0$  دنباله‌ای از نقاط به صورت  $\{x^k\}$  تولید می‌کنند که معمولاً در شرط کاهشی  $f(x^{k+1}) < f(x^k)$  صدق کند. تولید نقاط دنباله تا رسیدن به دقت مطلوب ادامه می‌یابد. یک راهبرد اساسی برای حرکت الگوریتم از نقطه  $x^k$  به نقطه  $x^{k+1}$  روش جستجوی خطی است. در تکرار  $k$ ام این روش، ابتدا جهت (معمولاً کاهشی)  $p_k$  انتخاب شده و سپس طول گام  $\alpha_k$  محاسبه می‌شود و نقطه جدید برابر  $x^{k+1} = x^k + \alpha_k p_k$  در نظر گرفته می‌شود. طول گام را می‌توان با حل مسئله یک بعدی

$$\min_{\alpha} \phi(\alpha) = f(x^k + \alpha p_k) \quad (2)$$

به دست آورد و بنابراین به فرایند یافتن طول گام، جستجوی خطی گفته می‌شود. به دلیل هزینه‌بر بودن حل این مسئله، در بیشتر موارد از جستجوی خطی تقریبی برای محاسبه طول گام استفاده می‌شود، به این معنی که تقریبی از جواب مسئله (۲) را که در شرایط خاصی صدق می‌کند به عنوان طول گام در نظر می‌گیرند. از مهم‌ترین این شرایط می‌توان به شرایط جستجوی خطی ولف اشاره نمود که به صورت زیر است:

$$\begin{cases} f(x^k + \alpha_k p_k) \leq f(x^k) + c_1 \alpha_k p_k^T g_k, \\ g_k^T \nabla f(x^k + \alpha_k p_k) \geq c_2 p_k^T g_k, \end{cases}$$

که در آن  $g_k = \nabla f(x^k)$  و  $0 < c_1 < c_2 < 1$ .

در روش‌های شبه نیوتن برای حل مسئله (۱)، جهت حرکت به صورت  $p_k = -B_k^{-1} g_k$  محاسبه می‌شود که در آن  $B_k \approx \nabla^2 f(x^k)$  ماتریسی متقارن و معین مثبت است. بنابراین دنباله تکرارهای روش شبه نیوتن به صورت

$$x^{k+1} = x^k - \alpha_k B_k^{-1} g_k, \quad (3)$$

است. در اکثر روش‌های شبه نیوتن، ماتریس  $B_k$  طوری محاسبه می‌شود که در رابطه سکانت استاندارد به صورت  $B_k s_{k-1} = y_{k-1}$  صدق کند که در آن  $s_{k-1} = x^k - x^{k-1}$  و  $y_{k-1} = g_k - g_{k-1}$ . واضح است که یک شرط لازم برای اینکه  $B_k$  معین مثبت صادق در رابطه سکانت موجود باشد، این است که شرط زاویه به صورت  $s_{k-1}^T y_{k-1} > 0$  برقرار باشد.

از مهم‌ترین روش‌های شبه نیوتن، روش BFGS<sup>1</sup> است که در آن ماتریس  $B_k$  با استفاده از رابطه تکراری زیر به‌روز رسانی می‌شود:

$$B_k = B_{k-1} - \frac{B_{k-1} s_{k-1} s_{k-1}^T B_{k-1}}{s_{k-1}^T B_{k-1} s_{k-1}} + \frac{y_{k-1} y_{k-1}^T}{s_{k-1}^T y_{k-1}}$$

اگر در رابطه (۳) قرار دهیم  $B_k = I$ ، تکرارهای روش تندترین کاهش حاصل می‌شود.

از آنجایی که انتخاب نامناسب تقریب اولیه از ماتریس هسی ممکن است منجر به عدم کارایی روش‌های شبه نیوتن در حل برخی مسائل (با ماتریس هسی بدحالت) شود، مقیاس‌سازی روش‌های شبه نیوتن مورد توجه قرار گرفته است. به عنوان نمونه، چنگ و لی یک فرمول BFGS مقیاس‌بندی شده طیفی به صورت زیر برای حل مسائل بهینه‌سازی نامقید ارائه نمودند [۱]:

$$B_k = B_{k-1} - \frac{B_{k-1} s_{k-1} s_{k-1}^T B_{k-1}}{s_{k-1}^T B_{k-1} s_{k-1}} + \gamma_k \frac{y_{k-1} y_{k-1}^T}{s_{k-1}^T y_{k-1}},$$

$$\gamma_k = \frac{y_k^T s_k}{\|y_k\|^2}$$

که در آن

۲- روش‌های BB و شبه نیوتن برای محاسبه طول گام

در سال ۱۹۸۸ بارزیلای و بوروین دو فرمول جدید برای محاسبه طول گام ارائه نمودند [۲] که به فرمول‌های BB<sup>2</sup> معروف شدند و به صورت زیر هستند:

1. Broyden-Fletcher-Goldfarb-Shano  
2. Barzilai-Borwein

دوپارامتری محاسبه می‌کنیم.

$$\alpha_k^{BB1} = \frac{\|s_{k-1}\|^2}{s_{k-1}^T y_{k-1}}, \quad (4)$$

$$\alpha_k^{BB2} = \frac{s_{k-1}^T y_{k-1}}{\|y_{k-1}\|^2}. \quad (5)$$

### ۳- ارائه روش جدید محاسبه طول گام

در این بخش، ابتدا به منظور برطرف نمودن خطای عددی حاصل از بد حالت بودن تقریب ماتریس هسی  $(D_k \approx \nabla^2 f(x^k))$ ، با ضرب رابطه سکانت در پارامتر  $\gamma_k$ ، رابطه سکانت مقیاس‌بندی شده به صورت زیر را به دست می‌آوریم:

$$\gamma_k B_k s_{k-1} = \gamma_k y_{k-1}. \quad (10)$$

با قرار دادن  $\bar{B}_k = \gamma_k B_k$ ، از یک فرمول BFGS مقیاس‌بندی شده با دو عامل مقیاس برای تقریب ماتریس  $\bar{B}_k$  به صورت زیر استفاده می‌کنیم:

$$\bar{B}_k = \delta_k [D_k - \frac{D_k s_{k-1} s_{k-1}^T D_k}{s_{k-1}^T D_k s_{k-1}}] + \gamma_k \frac{y_{k-1} y_{k-1}^T}{s_{k-1}^T y_{k-1}}, \quad (11)$$

که در آن

$$\gamma_k = \frac{s_{k-1}^T y_{k-1}}{\|y_{k-1}\|^2} = \alpha_k^{BB2}, \quad (12)$$

و

$$\delta_k = \frac{\|s_{k-1}\|}{\|y_{k-1}\| + \tau_k}, \quad (13)$$

که در آن  $\tau_k$  پارامتر مثبت دلخواهی است. توجه کنید که از رابطه (۱۱) داریم:

$$\bar{B}_k s_{k-1} = \delta_k [D_k s_{k-1} - \frac{D_k s_{k-1} s_{k-1}^T D_k}{s_{k-1}^T D_k s_{k-1}} s_{k-1}] + \gamma_k \frac{y_{k-1} y_{k-1}^T}{s_{k-1}^T y_{k-1}} s_{k-1} = \gamma_k y_{k-1}, \quad (14)$$

یعنی  $\bar{B}_k$  در رابطه سکانت مقیاس‌بندی شده (۱۰) صدق می‌کند.

لم ۳-۱: اگر  $s_{k-1}^T y_{k-1} > 0$ ، آنگاه ماتریس  $\bar{B}_k$  معین مثبت و متقارن است. بعلاوه  $trace(\bar{B}_k) \leq n$ .

بعد از آن، پژوهشگران طول گام‌های بسیاری بر مبنای روابط فوق و روابط سکانت اصلاح‌شده، ارائه نمودند [۳-۶].

توجه کنید که  $\alpha_k^{BB1}$  را می‌توان با حل مسئله یک بعدی

$$\min_{\alpha} \|D_k s_{k-1} - y_{k-1}\|_2^2 \quad (6)$$

که در آن  $D_k = \frac{1}{\alpha_k} I$  به دست آورد. بنابراین دنباله تولید شده در روش تندترین کاهش با استفاده از طول گام  $\alpha_k^{BB1}$  عبارتست از:

$$x^{k+1} = x^k - \alpha_k g_k = x^k - D_k^{-1} g_k. \quad (7)$$

به عبارت دیگر دنباله جوابها با استفاده از یک روش شبه نیوتن با طول گام ثابت یک، تولید شده است وقتی که  $D_k \approx \nabla^2 f(x^k)$ .

لازم به ذکر است که  $D_k = \frac{1}{\alpha_k} I = \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} I$  در رابطه سکانت استاندارد  $D_k s_{k-1} = y_{k-1}$  صدق نمی‌کند. اخیراً لی و همکارانش یک بهنگام شبه نیوتن از ماتریس  $D_k s_{k-1} = y_{k-1}$  را برای تولید طول گام مورد استفاده قرار دادند [۷] که یک حالت خاص آن فرمول BFGS به صورت زیر است:

$$B_k = D_k - \frac{D_k s_{k-1} s_{k-1}^T D_k}{s_{k-1}^T D_k s_{k-1}} + \frac{y_{k-1} y_{k-1}^T}{s_{k-1}^T y_{k-1}}, \quad (8)$$

و بنابراین فرمول زیر برای محاسبه طول گام حاصل شده است:

$$\alpha_k^{LD} = 1 / \left[ \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} \left( 1 - \frac{(g_k^T s_{k-1})^2}{\|s_{k-1}\|^2 \|g_k\|^2} \right) + \frac{\|y_{k-1}\|^2}{(s_{k-1}^T y_{k-1}) \|y_{k-1}\|^2 \|g_k\|^2} \right]. \quad (9)$$

در بخش بعدی، طول گام را بدون استفاده از فرآیند جستجوی خطی و با حل یک تقریب درجه دوم از مسئله (۲) با استفاده از یک فرمول شبه نیوتن مقیاس‌بندی شده

$$-\frac{s_{k-1}^T y_{k-1} s_{k-1}^T s_{k-1}}{\|s_{k-1}\|^2 s_{k-1}^T s_{k-1}} + 1 =$$

$$\frac{\|s_{k-1}\|}{\|y_{k-1}\|} \left[ \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} n - \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} \right] + 1 = (n -$$

$$1) \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\| \|y_{k-1}\|} + 1 \leq n - 1 + 1 = n.$$

**توجه ۱:** اگر  $s_{k-1}^T y_{k-1} \leq 0$ ، آنگاه  $\bar{B}_k$  ممکن است معین مثبت نباشد و بنابراین از رابطه  $\alpha_k = \delta \alpha_{k-1}$  برای محاسبه طول گام استفاده می‌کنیم که  $\rho \in (0, 1)$ .

**توجه ۲:** اگر  $s_{k-1}^T y_{k-1} > 0$ ، از لم قبلی نتیجه می‌شود که بزرگترین مقدار ویژه ماتریس  $\bar{B}_k$  از  $n$  بیشتر نیست و این موضوع به بدحالت نشدن ماتریس  $\bar{B}_k$  کمک خواهد کرد.

حال برای به دست آوردن یک فرمول صریح برای محاسبه طول گام،  $\alpha_k$  را مینیمم‌کننده تقریب درجه دوم از تابع  $\phi$  به صورت زیر در نظر می‌گیریم:

$$\phi(\alpha) = f(x_k - \alpha g_k) \approx f(x_k) - \alpha g_k^T g_k + \frac{1}{2} \alpha^2 g_k^T B_k g_k, \quad (15)$$

یعنی  $\alpha_k$  جواب مسئله درجه دوم زیر است:

$$\min_{\alpha} f(x^k) - \alpha g_k^T g_k + \frac{1}{2} \alpha^2 g_k^T B_k g_k \quad (16)$$

با فرض معین مثبت بودن ماتریس  $B_k$ ، جواب مسئله فوق برابر است با:

$$\alpha_k^{new} = \frac{g_k^T g_k}{g_k^T B_k g_k} = \gamma_k \frac{g_k^T g_k}{g_k^T \bar{B}_k g_k}. \quad (17)$$

همچنین، از رابطه (۱۱) داریم:

$$g_k^T \bar{B}_k g_k = \delta_k g_k^T D_k g_k - \delta_k \frac{g_k^T D_k s_{k-1} s_{k-1}^T D_k g_k}{s_{k-1}^T D_k s_{k-1}} + \gamma_k \frac{g_k^T y_{k-1} y_{k-1}^T g_k}{s_{k-1}^T y_{k-1}}. \quad (18)$$

با جای گذاری  $D_k = \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} I$  به دست می‌آوریم:

$$g_k^T \bar{B}_k g_k = \delta_k \frac{g_k^T s_{k-1}^T y_{k-1} g_k}{\|s_{k-1}\|^2} - \frac{\delta_k g_k^T s_{k-1}^T y_{k-1} s_{k-1} s_{k-1}^T y_{k-1} g_k \|s_{k-1}\|^2}{\|s_{k-1}\|^4 s_{k-1}^T s_{k-1}^T y_{k-1} s_{k-1}}$$

**اثبات:** با توجه به فرض،  $D_k = \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} I$  ماتریسی متقارن و معین مثبت است که این با استفاده از رابطه (۱۱) تقارن  $\bar{B}_k$  را نتیجه می‌دهد. برای اثبات معین مثبت بودن ماتریس  $\bar{B}_k$  با استفاده از تعریف نشان می‌دهیم که برای هر  $z \in R^n, z \neq 0$  داریم  $z^T \bar{B}_k z > 0$ . با استفاده از رابطه (۱۱) داریم:

$$z^T \bar{B}_k z = \delta_k [z^T D_k z - \frac{z^T D_k s_{k-1} s_{k-1}^T D_k z}{s_{k-1}^T D_k s_{k-1}}] + \gamma_k \frac{z^T y_{k-1} y_{k-1}^T z}{s_{k-1}^T y_{k-1}}.$$

با در نظر گرفتن تجزیه چولسکی ماتریس  $D_k$  به صورت  $D_k = LL^T$  و تعریف  $a = L^T z$  و  $b = L^T s_{k-1}$  به دست می‌آوریم:

$$z^T \bar{B}_k z = \delta_k \left[ \frac{\|a\|_2^2 \|b\|_2^2 - (a^T b)^2}{\|b\|_2^2} \right] + \gamma_k \frac{\|z^T y_{k-1}\|_2^2}{s_{k-1}^T y_{k-1}}.$$

با توجه به فرض  $s_{k-1}^T y_{k-1} > 0$  و مثبت بودن پارامترهای  $\delta_k$  و  $\gamma_k$  و استفاده از نامساوی کوشی-شوارتز نتیجه می‌شود که عبارت فوق همواره نامنفی است. بعلاوه جمله اول سمت راست تساوی فقط وقتی صفر می‌شود که بردار  $a$  مضربی از بردار  $b$  باشد و با توجه به وارون‌پذیر بودن ماتریس  $L$ ، این یعنی که ثابت ناصفر  $t$  موجود باشد که  $z = t s_{k-1}$  و در این صورت جمله دوم سمت راست تساوی برابر می‌شود با

$$\gamma_k \frac{\|z^T y_{k-1}\|_2^2}{s_{k-1}^T y_{k-1}} = \gamma_k t^2 \frac{\|s_{k-1}^T y_{k-1}\|_2^2}{s_{k-1}^T y_{k-1}} = \gamma_k t^2 s_{k-1}^T y_{k-1} > 0,$$

و این معین مثبت بودن ماتریس  $\bar{B}_k$  را ثابت می‌کند. همچنین، داریم:

$$\text{trace}(\bar{B}_k) = \delta_k [\text{trace}(D_k) - \frac{s_{k-1}^T D_k s_{k-1}}{s_{k-1}^T D_k s_{k-1}}] + \gamma_k \frac{y_{k-1}^T y_{k-1}}{s_{k-1}^T y_{k-1}} = \frac{\|s_{k-1}\|}{\|y_{k-1}\| + \tau_k} \left[ \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} n - \frac{s_{k-1}^T y_{k-1} s_{k-1}^T s_{k-1}}{\|s_{k-1}\|^2 s_{k-1}^T s_{k-1}} \right] + \frac{s_{k-1}^T y_{k-1} y_{k-1}^T y_{k-1}}{\|y_{k-1}\|^2 s_{k-1}^T y_{k-1}} \leq \frac{\|s_{k-1}\|}{\|y_{k-1}\|} \left[ \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} n \right]$$

$$\delta_k \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} \sin^2 \beta_k + \gamma_k \frac{\|y_{k-1}\|^2}{(s_{k-1}^T y_{k-1})} \cos^2 \varpi_k \leq \sin^2 \beta_k \cos^2 \theta_k + \cos^2 \varpi_k$$

که در آن زاویه بین  $s_{k-1}$  و  $y_{k-1}$  است. با استفاده از برخی روابط مثلثاتی می‌توان نشان داد که (اثبات لم ۳.۲ در [۷] را ببینید):

$$\sin^2 \beta_k \cos^2 \theta_k + \cos^2 \varpi_k \leq 1 - \frac{1}{2} \sin^2 \theta_k + \sqrt{\sin^2 \theta_k - \frac{3}{4} \sin^4 \theta_k}.$$

از آنجایی که ماکسیمم مقدار تابع  $h(u) = 1 - \frac{1}{2}u + \sqrt{u - \frac{3}{4}u^2}$  وقتی  $u \in [0, 1]$  برابر است با  $\frac{4}{3}$ ، نتیجه می‌شود که:

$$\sin^2 \beta_k \cos^2 \theta_k + \cos^2 \varpi_k \leq \frac{4}{3}.$$

بنابراین

$$\alpha_k^{new} \geq \frac{3}{4} \gamma_k = \frac{3}{4} \alpha_k^{BB2}.$$

الگوریتم تندترین کاهش ارائه شده به صورت زیر است:

### الگوریتم تندترین کاهش بدون جستجوی خطی

**گام ۱:** نقطه شروع  $x^0$ ، پارامتر  $\varepsilon > 0$  و  $\rho \in (0, 1)$  را در نظر بگیرید.

**گام ۲:** قرار دهید  $k = 0$  و  $g_0 = \nabla f(x^0)$  را محاسبه کنید و قرار دهید  $\alpha_k = 1 / \|g_0\|_\infty$  و به گام ۵ بروید.

**گام ۳:** مقادیر  $s_{k-1} = x^k - x^{k-1}$  و  $y_{k-1} = g_k - g_{k-1}$  را محاسبه کنید.

**گام ۴:** اگر  $s_{k-1}^T y_{k-1} \leq 0$  قرار دهید  $\alpha_k = \rho \alpha_{k-1}$ ، در غیر این صورت  $\alpha_k$  را از رابطه (۲۰) محاسبه کنید.

**گام ۵:** قرار دهید  $x^{k+1} = x^k - \alpha_k g_k$  قرار دهید  $k = k + 1$ .

**گام ۶:** اگر  $\|g_k\| \leq \varepsilon$  توقف کنید و  $x^k$  را به عنوان تقریبی از جواب مسئله در نظر بگیرید ( $x^k$ ).

$$+ \gamma_k \frac{g_k^T y_{k-1} y_{k-1}^T g_k}{s_{k-1}^T y_{k-1}}$$

پس داریم:

$$\begin{aligned} g_k^T \bar{B}_k g_k &= \delta_k \frac{s_{k-1}^T y_{k-1} \|g_k\|^2}{\|s_{k-1}\|^2} \\ &- \delta_k \frac{(s_{k-1}^T y_{k-1})^2 (g_k^T s_{k-1})^2 \|s_{k-1}\|^2}{\|s_{k-1}\|^6 s_{k-1}^T y_{k-1}} \\ &+ \gamma_k \frac{(g_k^T y_{k-1})^2}{s_{k-1}^T y_{k-1}} \\ &= \delta_k \frac{s_{k-1}^T y_{k-1} \|g_k\|^2}{\|s_{k-1}\|^2} \\ &- \delta_k \frac{(s_{k-1}^T y_{k-1})(g_k^T s_{k-1})^2}{\|s_{k-1}\|^4} \\ &+ \gamma_k \frac{(g_k^T y_{k-1})^2}{s_{k-1}^T y_{k-1}}, \end{aligned} \quad (19)$$

که نتیجه می‌دهد:

$$\begin{aligned} \alpha_k^{new} &= \gamma_k \frac{g_k^T g_k}{g_k^T \bar{B}_k g_k} \gamma_k / \\ &\left[ \delta_k \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} \left( 1 - \frac{(g_k^T s_{k-1})^2}{\|s_{k-1}\|^2 \|g_k\|^2} \right) + \gamma_k \frac{\|y_{k-1}\|^2}{(s_{k-1}^T y_{k-1})} \frac{(g_k^T y_{k-1})^2}{\|y_{k-1}\|^2 \|g_k\|^2} \right]. \end{aligned} \quad (20)$$

بنابراین رابطه فوق را برای محاسبه طول گام مورد استفاده قرار می‌دهیم.

**لم ۳-۲:** برای طول گام  $\alpha_k^{new}$  داریم:

$$\frac{3}{4} \alpha_k^{BB2} \leq \alpha_k^{new}. \quad (21)$$

**اثبات:** داریم:

$$\alpha_k^{new} = \gamma_k / \left[ \delta_k \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\|^2} \sin^2 \beta_k + \gamma_k \frac{\|y_{k-1}\|^2}{(s_{k-1}^T y_{k-1})} \cos^2 \varpi_k \right],$$

که در آن زاویه بین  $s_{k-1}$  و  $g_k$  و  $\varpi_k$  زاویه بین  $g_k$  و  $y_{k-1}$  است. بنابراین با استفاده از تعریف پارامترهای  $\delta_k$  و  $\gamma_k$  نتیجه می‌شود که

که در آن  $a = 20$ ،  $b = 0.2$ ،  $c = 2\pi$ . برای این مسئله  $n = 5$ ، و نقطه شروع الگوریتم را به صورت  $x_i^0 = -2$ ،  $i = 1, \dots, 5$

در نظر گرفتیم.

#### مسئله ۴:

$$\min f(x) = 12x_1^2 - 6.3x_1^4 + x_1^6 - 6x_1x_2 + 6x_2^2$$

نقطه شروع الگوریتم را به صورت  $x^0 = [-10, -10]$  در نظر گرفتیم.

#### مسئله ۵:

$$\min f(x) = (10x_1 + x_2 - 7)^2 + (x_1 - 1)^2$$

نقطه شروع الگوریتم را به صورت  $x^0 = [10, 10]$  در نظر گرفتیم.

#### مسئله ۶:

$$\min f(x) = \sum_{i=1}^n (\exp(x_i) - \sqrt{i}x_i)$$

در این مسئله که یک مسئله محدب است، وقتی که  $n = 5$ ، نقطه شروع الگوریتم را به صورت  $x_i^0 = 4i$ ،  $i = 1, \dots, 5$

در نظر گرفتیم که ماتریس هسی تابع هدف در آن ماتریسی بد حالت ( $\text{cond}(\nabla^2 f(x^0)) \approx 10^6$ ) است. همچنین، وقتی که  $n = 10$ ، نقطه شروع الگوریتم را به صورت

$$x_i^0 = 2i, \quad i = 1, \dots, 10$$

در نظر گرفتیم که ماتریس هسی تابع هدف در آن ماتریسی بد حالت ( $\text{cond}(\nabla^2 f(x^0)) \approx 10^7$ ) است.

نتایج به دست آمده از اجرای روش تندترین کاهش برای حل این مسائل به ازای مقادیر مختلف برای طول گام، در جدول ۱ آمده است. در این جدول Time نشان‌دهنده

خروجی الگوریتم است). در غیر این صورت به گام ۳ بروید.

در بخش بعدی نتایج عددی حاصل از اجرای الگوریتم ارائه شده است.

#### ۴- نتایج عددی

این الگوریتم را در محیط نرم افزاری MATLAB 2013 پیاده‌سازی و برای حل چندین مسئله بهینه‌سازی به صورت زیر آزمون کردیم.

#### مسئله ۱:

$$\min f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

نقطه شروع الگوریتم برای حل این مسئله را به صورت  $x^0 = (0, -20)$  در نظر گرفتیم.

#### مسئله ۲:

$$\min f(x) = \frac{588600}{(3r_0^2 - 4\cos(\theta)r_0^2 - 2(\sin^2(\theta)\cos(x - \frac{2\pi}{3}) - \cos^2(\theta))r_0^2)^6} - \frac{1079.1}{(3r_0^2 - 4\cos(\theta)r_0^2 - 2(\sin^2(\theta)\cos(x - \frac{2\pi}{3}) - \cos^2(\theta))r_0^2)^3} + \frac{600800}{(3r_0^2 - 4\cos(\theta)r_0^2 - 2(\sin^2(\theta)\cos(x) - \cos^2(\theta))r_0^2)^6} - \frac{1071.5}{(3r_0^2 - 4\cos(\theta)r_0^2 - 2(\sin^2(\theta)\cos(x) - \cos^2(\theta))r_0^2)^3} + \frac{481300}{(3r_0^2 - 4\cos(\theta)r_0^2 - 2(\sin^2(\theta)\cos(x + \frac{2\pi}{3}) - \cos^2(\theta))r_0^2)^6} - \frac{1064.6}{(3r_0^2 - 4\cos(\theta)r_0^2 - 2(\sin^2(\theta)\cos(x + \frac{2\pi}{3}) - \cos^2(\theta))r_0^2)^3},$$

$$0 \leq x \leq 2\pi,$$

که در آن  $\theta = 1.9111$  و  $r_0 = 1.54$  این مسئله کاربردی توسط ماراناز و فلوداز در سال ۱۹۹۴ مطرح شده است [۸] و دارای چندین بهینه موضعی است. نقطه شروع الگوریتم برای این مسئله  $x^0 = 1$  در نظر گرفته شده است.

#### مسئله ۳:

$$\min f(x) = -a \exp\left(-b \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(cx_i)\right) + a + \exp(1)$$

دولان-موره برای مقایسه زمان اجرای الگوریتم‌ها در شکل ۱ آورده شده است که کارایی روش ارائه شده را نشان می‌دهد.

**نتیجه‌گیری**

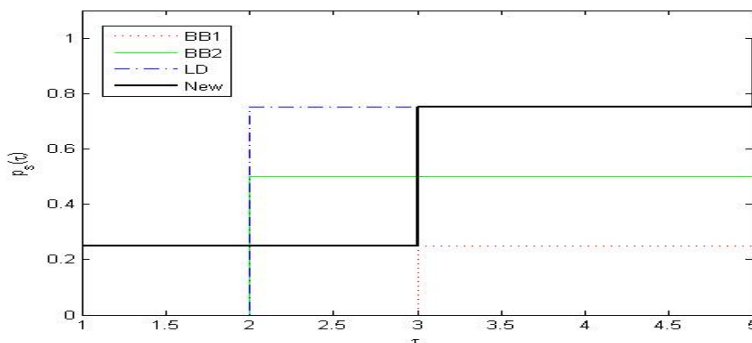
در این مقاله به ارائه یک روش تندترین کاهش بدون جستجوی خطی برای حل مسائل بهینه‌سازی نامقید پرداختیم. با استفاده از یک فرمول شبه نیوتن مقیاس‌بندی‌شده با دو عامل مقیاس، یک رابطه صریح برای محاسبه طول گام به دست آوردیم. کارایی روش ارائه شده را نیز با بررسی نتایج عددی به دست آمده از الگوریتم به ازای مقادیر مختلف طول گام نشان دادیم.

زمان اجرای الگوریتم به ثانیه، F نشان‌دهنده مقدار بهینه حاصل‌شده از الگوریتم، و OPT نشان‌دهنده مقدار بهینه مسئله است. در گام ۱ الگوریتم مقادیر پارامترها را  $\rho = 0.2$  و  $\varepsilon = 10^{-10}$  قرار دادیم. همچنین، در رابطه (۱۳)،  $\beta_k = |s_k^T g_k|$  در نظر گرفته شده است. حداکثر تکرار مجاز الگوریتم ۱۰۰۰ بوده است.

با توجه به جدول ۱ مشخص است که روش ارائه شده در حل مسائل بد حالت نسبت به سایر روش‌ها برتری قابل ملاحظه‌ای داشته است، به ویژه برای مسائل ۴ و ۶ روش ارائه‌شده تقریب بسیار مناسبی از مقدار بهینه را به دست آورده است و این در حالی است که سایر روش‌ها در حل این دو مسئله با شکست مواجه شده‌اند. همچنین نمودار

جدول ۱: نتایج عددی حاصل از اجرای الگوریتم به ازای مقادیر مختلف طول گام ۱

شماره مسئله	طول گام استفاده شده		$\alpha_k^{BB1}$		$\alpha_k^{BB2}$		$\alpha_k^{LD}$		$\alpha_k^{new}$	
	OPT	Time	F	Time	F	Time	F	Time	F	
۱	0	0.061	32.570	0.038	1.6e-04	0.036	1.03e-04	0.038	8.6e-10	
۲	-1.071	0.026	-0.79	4e-04	-0.79	0.002	-0.79	0.046	-1.0709	
۳	0	0.028	3.574	0.002	3.574	0.002	3.574	0.111	0.0427	
۴	0	0.025	1.791	0.003	1.791	0.005	1.791	0.003	1.2e-12	
۵	0	0.028	7e-31	0.002	1.5e-16	0.002	7.8e-31	0.002	0	
۶ (n = 5)	3.755	0.061	7e+70	0.043	3.2e+02	0.007	3.7556	0.008	3.7551	
۶ (n = 10)	3.195	0.060	5e+14	0.043	Inf	0.049	1.1e+92	0.007	3.195	



شکل ۱: نمودار دولان-موره برای مقایسه کارایی الگوریتم‌ها از نظر زمان اجرا

فهرست منابع

- [1] W. Cheng, D. Li. Spectral scaling bfgs method. *Journal of Optimization Theory and Applications* 146(2):305–319 (2010)
- [2] J. Barzilai, J.M. Borwein. Two-point step size gradient methods. *IMA Journal of Numerical Analysis* 8 :141–148 (1988)
- [3] F. Biglari, M.A. Hassan, W.J. Leong. New quasi-Newton methods via higher order tensor models. *Journal of Computational and Applied Mathematics* 235(8):2412–2422 (2011)
- [4] F. Biglari, M. Solimanpur. Scaling on the spectral gradient method. *Journal of Optimization Theory and Applications* 158(2):626–635 (2013)
- [5] Y.H. Xiao, Q.Y. Wang, D. Wang. Notes on the Dai-Yuan-Yuan modified spectral gradient method. *Journal of computational and Applied Mathematics* 234(10): 2986–2992 (2010)
- [6] B. Zhou, L. Gao, Y.H. Dai. Gradient methods with adaptive step sizes. *Computational Optimization and Applications* 35(1): 69–86 (2006)
- [7] H. Liu, Z. Liu, X. Dong. A new adaptive Barzilai and Borwein method for unconstrained optimization. *Optimization Letters* 12: 845–873(2018)
- [8] C.D. Maranas, C.A. Floudas. A deterministic global optimization approach for molecular structure determination, *Journal of Chemical Physics* 100(2): 1247–1261 (1994)