

Paper Type: Original Article



# Combining Non-Monotone Trust Rregion Method with a New Adaptive Radius for Unconstrained Optimization Problems

Seyed Hamzeh Mirzaei<sup>1</sup>, Ali Ashrafi<sup>1,\*</sup>

<sup>1</sup> Department of Mathematics, Semnan University, Semnan, Iran; sh\_mirzaei82@yahoo.com; a\_ashrafi@semnan.ac.ir.

### Citation:



Mirzaei, S. M., & Ashrafi, A. (2024). Combining non-Monotone trust rregion method with a new adaptive radius for unconstrained optimization problems. *Journal of decisions and operations research*, 9(1), 30-41.

Received: 08/11/2022

Reviewed: 10/12/2022

Revised: 09/01/2023

Accepted: 01/03/2023

## Abstract

**Purpose:** One of the most effective methods for solving unconstrained optimization problems is the trust region method. The strategy of determining the radius of the trust region has a significant effect on the efficiency of this method. On the other hand, imposing the monotonicity condition will decrease the convergence speed of this method. Therefore, improving and increasing the efficiency of this method is one of the most important issues and the attention of researchers.

**Methodology:** Establishing a new adaptive trust region radius as well as combining the trust region method with a non-monotone strategy to avoid the adverse effects of monotonicity.

**Findings:** A new adaptive trust region radius converged to zero is provided, and then a trust region combination is performed using a non-monotone strategy. Running the algorithm on a set of test functions shows that the new adaptive radius, along with the non-monotone strategy used, significantly improves the efficiency of the trust region method.

**Originality/Value:** The presented non-monotone adaptive algorithm has a second-order convergence rate. In addition, it significantly reduces computational costs compared to traditional algorithms. On the other hand, the new adaptive radius avoids the ineffectiveness of the trust region close to the solution.

**Keywords:** Non-monotone strategy, Unconstrained optimization, Trust region, Global convergence.



Corresponding Author: a\_ashrafi@semnan.ac.ir



Licensee. **Journal of Decisions and Operations Research**. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).



## ترکیب ناحیه اعتماد غیر یکنوا با یک شعاع وفقی جدید برای مسایل بهینه‌سازی نامقید

سید حمزه میرزایی<sup>۱</sup>، علی اشرفی<sup>۱\*</sup>  
<sup>۱</sup>گروه ریاضی دانشگاه سمنان، سمنان، ایران.

### چکیده

**هدف:** یکی از کاراترین روش‌ها برای حل مسایل بهینه‌سازی نامقید روش ناحیه اعتماد است. استراتژی تعیین شعاع ناحیه اعتماد تاثیر بسزایی بر میزان کارایی این روش دارد. از طرفی تحمیل شرط یکنوایی موجب کاهش سرعت همگرایی این روش خواهد شد؛ بنابراین بهبود و افزایش کارایی این روش یکی از مسایل بسیار مهم و مورد توجه محققان می‌باشد.

**روش‌شناسی پژوهش:** ایجاد یک شعاع ناحیه اعتماد وفقی جدید و همچنین ترکیب روش ناحیه اعتماد با استراتژی غیر یکنوا برای جلوگیری از اثرات نامطلوب یکنوایی.

**یافته‌ها:** یک شعاع ناحیه اعتماد وفقی جدید همگرا به صفر ارایه شده و سپس ترکیب ناحیه اعتماد با یک استراتژی غیر یکنوایی انجام می‌شود. اجرای الگوریتم بر روی یک مجموعه از توابع تست نشان می‌دهد که شعاع وفقی جدید به همراه استراتژی غیر یکنوایی به کار گرفته شده به صورت قابل ملاحظه‌ای کارایی روش ناحیه اعتماد را بهبود می‌بخشد.

**اصالت/ارزش افزوده علمی:** الگوریتم غیر یکنوایی وفقی ارایه شده دارای مرتبه همگرایی مرتبه دوم بوده و هزینه‌های محاسباتی را نسبت به الگوریتم‌های سنتی به مقدار قابل توجهی کاهش می‌دهد. از طرفی شعاع وفقی جدید از بی‌اثر ماندن ناحیه اعتماد در نزدیکی جواب جلوگیری می‌کند.

کلیدواژه‌ها: استراتژی غیر یکنوا، بهینه‌سازی نامقید، ناحیه اعتماد، همگرایی سراسری.

### ۱- مقدمه

مسایل بهینه‌سازی نامقید شامل گستره وسیعی از مسایل بهینه‌سازی می‌شوند که در آن‌ها متغیرها دارای هیچ نوع قیدی یا محدودیتی نیستند؛ بنابراین ارایه روش‌هایی کارا و نیرومند برای حل این نوع مسایل یک امر مهم و بسیار ضروری است. در حالت کلی یک مساله بهینه‌سازی نامقید به صورت زیر نشان داده می‌شود:

که در آن  $f: R^n \rightarrow R$  یک تابع است که دو بار به صورت پیوسته مشتق پذیر است. روش‌های متنوعی برای حل مساله (۱) ارایه شده است. دو دسته

$$\min \{f(x) | x \in R^n\}, \quad (1)$$

از معروف‌ترین روش‌ها برای حل مساله (۱) روش‌های جست‌وجوی خطی و ناحیه اعتماد است. مشابه روش ناحیه اعتماد ارایه شده توسط پاول [1] در هر مرحله برای تعیین گام آزمایشی  $d_k$  زیر مساله درجه دوم زیر حل می‌شود:

$$\min_{d \in R^n} m_k(d) = f_k + g_k^T d + \frac{1}{2} d^T B_k d, \quad (2)$$

$$s. t. \|d\| \leq \Delta_k,$$

که در آن  $B_k \in R^{n \times n}$  هسی دقیق  $H_k = \nabla^2 f(x_k)$  یا یک تقریب مقارن از آن است. همچنین  $\| \cdot \|$  نرم اقلیدسی است.  $\Delta_k > 0$  شعاع ناحیه اعتماد است که در هر مرحله به روزرسانی می شود. در روش ناحیه اعتماد نسبت ناحیه اعتماد به صورت زیر تعریف می شود:

$$r_k = \frac{f(x_k) - f(x_{k+1})}{m_k(0) - m_k(d_k)}. \quad (3)$$

صورت و معرج کسر را به ترتیب کاهش واقعی و کاهش پیش بینی شده می نامند. مقدار نسبت ناحیه اعتماد  $r_k$  نقش کلیدی در تعیین استراتژی برای افزایش یا کاهش شعاع ناحیه اعتماد و همچنین تصمیم گیری برای پذیرفتن یا عدم پذیرش گام آزمایشی دارد. اگر مقدار  $r_k$  نزدیک ۱ باشد گام آزمایشی  $d_k$  را پذیرفته و می توان مقدار شعاع ناحیه اعتماد را افزایش داد. در این صورت برای گام بعدی قرار می دهیم  $x_{k+1} = x_k + d_k$ . درحالی که اگر مقدار  $r_k$  عددی نزدیک به صفر و یا منفی باشد آنگاه، گام آزمایشی  $d_k$  رد شده، شعاع ناحیه اعتماد بایستی کاهش پیدا کرده و زیر مساله ناحیه اعتماد مجدداً حل شود.

قابل توجه است که اندازه شعاع ناحیه اعتماد تاثیر بسزایی در کارایی الگوریتم ناحیه اعتماد دارد. اگر شعاع بزرگ باشد تعداد محاسبات تابع افزایش می یابد. از طرف دیگر، اگر شعاع کوچک باشد تعداد تکرارها افزایش خواهد یافت. در الگوریتم ناحیه اعتماد سنتی شعاع ناحیه اعتماد به صورت زیر به روزرسانی می شود:

$$\Delta_{k+1} = \begin{cases} b_1 \Delta_k, & r_k < \rho_1, \\ \min\{b_2 \Delta_k, \bar{\Delta}\}, & r_k \geq \rho_2. \end{cases} \quad (4)$$

به طوری که  $0 < \rho_1 < 1 < b_2 < 1$  و  $0 < \rho_2 < 1$ . برای غلبه بر این تناقض محققان روش های بسیاری را برای به روزرسانی شعاع ناحیه اعتماد در یک رویکرد وقتی ارایه کردند [4]–[2]. به عنوان مثال با مراجعه به کار ژنگ و همکاران [5] می توان مشاهده کرد که شعاع ناحیه اعتماد ارایه شده توسط آن ها به صورت زیر تعریف می شود:

$$\Delta_k = c^p \|\hat{B}_k^{-1}\| \|g_k\|,$$

که در آن  $c \in (0, 1)$ ،  $p$  یک عدد صحیح نامنفی و  $\hat{B}_k = B_k + E_k$ . برای دوری کردن از محاسبه وارون ماتریس هسی و یا تقریبی از آن، لی [6] یک شعاع ناحیه اعتماد وقتی را به صورت زیر ارایه کرد:

$$\Delta_k = c^{p_k} \frac{\|d_{k-1}\|}{\|y_{k-1}\|} \|g_k\|,$$

به طوری که  $y_{k-1} = g_k - g_{k-1}$ . روش ناحیه اعتماد با شعاع ناحیه اعتماد همگرا به صفر نخستین بار توسط فان و یوان [8] ارایه گردید. فان و لو [7] یک الگوریتم ناحیه اعتماد اصلاح شده با شعاع ناحیه اعتماد همگرا به صفر را برای سیستم معادلات غیر خطی ارایه کردند. یک الگوریتم ناحیه اعتماد دومرحله ای اصلاح شده برای حل سیستم معادلات غیر خطی توسط بهرامی و امینی [7] ارایه گردید.

نتایج عددی بیانگر آن هستند که تحمیل شرط یک نوایی که در هر مرحله مقدار تابع هدف بایستی کاهش یابد، ممکن است سرعت همگرایی روش ناحیه اعتماد را کاهش دهد، به ویژه در حضور دره های منحنی باریک. برای غلبه بر این مشکل، گریپو و همکاران [8] یک استراتژی جست و جوی خطی غیر یکنوا را برای روش نیوتن پیشنهاد کردند. جست و جوی خطی غیر یکنوا ارایه شده توسط آن ها به شرح زیر است:

$$f(x_k + \alpha d_k) \leq f(x_k) + \sigma \alpha \nabla f(x_k)^T d_k,$$

مشروط بر این که  $\sigma \in (0, 1)$ ،  $N$  یک ثابت مثبت باشد. از آن زمان، نویسندگان زیادی به بررسی این موضوع پرداختند. ژانگ و هاگر [9] یک طرح غیر یکنوا جدید را برای استراتژی جست و جوی خطی پیشنهاد کردند. مو و همکاران [10] از استراتژی ژانگ و هاگر [9] در چارچوب ناحیه اعتماد برای پیشنهاد یک روش ناحیه اعتماد غیر یکنوا جدید استفاده کردند. گو و مو [11] یک تکنیک جست و جوی خطی غیر یکنوا را به دلیل روش ژانگ و هاگر [9] پیشنهاد کردند، به طوری که پارامتر غیر یکنوا  $D_k$  باید نابرابری زیر را برآورده کند:

$$f(x_k + \alpha d_k) \leq D_k + \sigma \alpha \nabla f(x_k)^T d_k, \quad (5)$$

که در آن  $0 < \sigma < 1$  و

$$r_k = \frac{D_k - f(x_{k+1})}{m_k(0) - m_k(d_k)}, \quad (6)$$

به طوری که  $\eta \in (0,1)$  می‌تواند ثابت یا یک متغیر به صورت  $\eta_k$  باشد. آن‌ها برای ارزیابی این‌که آیا مرحله آزمایشی پذیرفته شده است یا خیر، از  $r_k$  استفاده کردند و به جای استفاده از  $f_k - f(x_{k+1})$  به عنوان کاهش واقعی  $D_k - f(x_{k+1})$  را به عنوان کاهش واقعی انتخاب کردند به طوری که دیگر الگوریتم نیاز به کاهش مقادیر یکنوای تابع در هر تکرار نداشته باشد. در نهایت،  $r_k$  به صورت زیر تعریف می‌شود:

$$f(x_k + \alpha d_k) \leq D_k + \sigma \alpha \nabla f(x_k)^T d_k, \quad (7)$$

مشروط بر این‌که  $D_k$  توسط رابطه (۶) محاسبه شود.

امینی و آهوخوش [12] بر اساس تکنیک غیر یکنوای ژانگ و هاگر [9] و شعاع ناحیه اعتماد قابل تنظیم شی و گوا [13] یک تکنیک غیر یکنوا با شعاع ناحیه اعتماد تنظیم شده به طور خودکار را ارائه کردند. اخیراً بسیاری از نویسندگان به موضوع ناحیه اعتماد غیر یکنوا پرداخته‌اند. برای مطالعه بیشتر می‌توان به [14]–[17] مراجعه کرد.

این مطالعه با هدف ترکیب استراتژی غیر یکنوای گو و مو [11] با شعاع ناحیه اعتماد وقتی همگرا به صفر طراحی شده است. ثابت می‌کنیم که روش پیشنهادی دارای همگرایی سراسری و نرخ همگرایی درجه دوم است.

باقی مانده مقاله به صورت زیر سازمان‌دهی شده است. در بخش ۲، الگوریتم جدید خود را تشریح می‌کنیم. تحلیل همگرایی سراسری الگوریتم ارائه شده در بخش ۳ بررسی می‌شود. در بخش ۴ نشان می‌دهیم که الگوریتم جدید دارای نرخ همگرایی درجه دوم است. نتایج عددی و مقایسه مدل ارائه شده در بخش ۵ انجام می‌گیرد. در نهایت در بخش ۶، نتیجه‌گیری انجام می‌شود.

## ۲- الگوریتم جدید

در این بخش به توصیف ساختار الگوریتم جدید می‌پردازیم. ابتدا شعاع ناحیه اعتماد وقتی جدید را بر اساس کار فان و یوان [18] معرفی می‌کنیم و سپس چارچوب کلی الگوریتم جدید را شرح می‌دهیم.

فرض کنید که  $x^*$  یک مینیمم کننده رابطه (۱) است. همان‌طور که مشخص است، اگر دنباله تولید شده توسط الگوریتم ناحیه اعتماد به  $x^*$  همگرا شود، نسبت ناحیه اعتماد  $r_k$  به ۱ همگرا می‌شود. با توجه به قانون به‌روزرسانی شعاع ناحیه اعتماد، می‌دانیم که شعاع ناحیه اعتماد  $\Delta_k$  برای همه  $k \in \mathbb{N}$  به اندازه کافی بزرگ از یک ثابت مثبت، بزرگ‌تر خواهد بود. از آنجایی که  $\|x_k - x^*\|_{k \geq 0}$  به صفر همگرا می‌شود، ناحیه اعتماد هیچ نقشی در پایان نخواهد داشت؛ بنابراین فان و یوان [18] شعاع ناحیه اعتماد خود را به صورت زیر ارائه کردند:

$$\Delta_k = \alpha_k \|g_k\|, \quad (8)$$

که در آن  $\alpha_k$  با توجه به مقدار نسبت ناحیه اعتماد  $r_{k-1}$  به صورت زیر به‌روزرسانی می‌شود:

$$\alpha_{k+1} = \begin{cases} c_1 \alpha_k, & r_k < \mu_1, \\ \alpha_k, & \mu_1 < r_k < \mu_2, \\ \min\{c_2 \alpha_k, N\}, & \text{otherwise.} \end{cases} \quad (9)$$

که  $0 < c_1 < 1 < c_2$ ،  $0 < \mu_0 < \mu_1 < \mu_2 < 1$  و  $N$  یک عدد مثبت بزرگ است.

هر زمان که دنباله تکرارها به اندازه کافی به جواب نزدیک باشد، شعاع ناحیه اعتماد نباید بزرگتر از یک ثابت مثبت باشد. برای همگرایی، تنها کافی است در تکرار  $K$ th مقدار  $\Delta_k$  بزرگتر از  $O(\|x_k - x^*\|)$  باشد، به این معنی که  $\Delta_k$  می تواند همگرا به صفر شود. این مانع از بزرگ شدن گام آزمایشی در نزدیکی جواب می شود. با گرفتن ایده از روش فان و یوان [18] به جای رابطه (۸)، شعاع وفقی ناحیه اعتماد جدید خود را به صورت زیر تعریف می کنیم:

$$\Delta_k = \alpha_k \|g_k\|^{\lambda(\delta)}, \quad (10)$$

که در آن  $\lambda(\delta): (0, \delta] \rightarrow [a_0, a_1]$  یک تابع نزولی و از پایین کراندار است که  $1 \leq a_0 < a_1 < 3/2$  و  $\delta$  حداکثر شعاع ممکن است. همچنین  $\alpha_k$  توسط رابطه (۹) به روزرسانی می شود.

اکنون گام آزمایشی  $d_k$  را در هر تکرار با حل زیر مساله درجه دوم زیر به دست می آوریم:

$$\begin{aligned} \min m_k(d) &= f_k + g_k^T d + \frac{1}{2} d^T B_k d, \\ \text{s. t. } \|d\| &\leq \Delta_k = \alpha_k \|g_k\|^{\lambda(\delta_k)}. \end{aligned} \quad (11)$$

که در آن  $\Delta_k$  توسط رابطه (۱۰) به روزرسانی می شود.

**ملاحظه ۱-** اگر دنباله  $\{x_k\}$  به مینیمم کننده رابطه (۱) همگرا شود، آنگاه  $r_k$  به ۱ همگرا می شود. طبق روش به روزرسانی شعاع در ناحیه اعتماد سنتی رابطه (۴)،  $\Delta_k$  بزرگتر از یک ثابت مثبت خواهد شد.

**ملاحظه ۲-** اگر دنباله  $\{x_k\}$  به مینیمم کننده رابطه (۱) همگرا شود،  $r_k$  به ۱ همگرا می شود، از آنجایی که  $\alpha_k$  از بالا کراندار است، بنابراین به خاطر روش به روزرسانی رابطه (۹) مقدار  $\Delta_k$ ، به صفر همگرا می شود.

حال ساختار کلی الگوریتم خود را با جزییات به صورت زیر ارائه می کنیم.

**الگوریتم ۱-** روش ناحیه اعتماد غیر یکنوا با شعاع وفقی جدید.

ورودی ها: دریافت کند  $x_0 \in R^n$  ماتریس معین مثبت  $B_0 \in R^{n \times n}$ ، تابع نزولی  $\lambda(\delta)$  و ثابت های  $0 < c_1 < 1$ ،  $0 < \mu_0 < \mu_1 < \mu_2 < 1$ ،  $0 < \varepsilon$  و  $0 < \delta < c_2 \alpha_1$ ،  $N > c_2 \alpha_1$ ،  $0 < c_2$  مقدار  $f(x_0)$  را محاسبه شود و  $k=0$  قرار دهد.

**گام ۱-** اگر  $\|g_k\| < \varepsilon$  توقف شود. در غیر این صورت به گام بعدی برود.

**گام ۲-** گام آزمایشی  $d_k$  را با حل زیر رابطه (۱۱) محاسبه شود.

**گام ۳-** مقدار  $r_k$  با رابطه (۷) محاسبه کند. اگر  $r_k > \mu_0$  باشد  $x_{k+1} = x_k + d_k$  در غیر این صورت  $x_{k+1} = x_k$  قرار دهد.

**گام ۴-** مقدار  $\alpha_k$  به وسیله رابطه (۹) به روزرسانی شود و  $\Delta_{k+1} = \alpha_{k+1} \|g_{k+1}\|^{\lambda(\delta_k)}$  قرار دهد.

**گام ۵-** ماتریس جدید  $B_k$  با استفاده از فرمول اصلاح شده  $BFGS$  به روزرسانی شود.

**گام ۶-** قرار دهد  $k = k + 1$  و به گام ۱ برگردد.

## ۳- تحلیل همگرایی

در این بخش، ویژگی‌های همگرایی الگوریتم ۱ مورد بحث قرار می‌گیرد. برای تجزیه و تحلیل همگرایی الگوریتم پیشنهادی، دو فرضیه را در طول این مقاله به شرح زیر در نظر می‌گیریم.

**فرضیه ۱-** تابع هدف  $f(x)$  به صورت پیوسته مشتق پذیر می‌باشد و در مجموعه تراز  $\Omega \subset \mathbb{R}^n$   $f(x) \leq f(x_0)$  از پایین کراندار است که در آن  $\Omega \in \mathbb{R}^n$  کراندار است.

**فرضیه ۲-** ماتریس  $B_k$  به طور یکنواخت کراندار است. به عبارت دیگر ثابتی مانند  $M > 0$  وجود دارد که  $\|B_k\| \leq M$ ; for all  $k \in N \cup \{0\}$ .

فرض کنید تابع هدف  $f(x)$  دو بار به صورت پیوسته مشتق پذیر باشد؛ بنابراین **فرضیه ۱** حاکی از آن است که  $\|\nabla^2 f(x)\|$  پیوسته است و به طور یکنوا روی  $\Omega$  کراندار است. از این رو، یک ثابت  $L > 0$  وجود دارد به طوری که برای همه  $x \in \Omega$  داریم  $\|\nabla^2 f(x)\| \leq L$ . که در این صورت از قضیه مقدار میانگین داریم

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|; \text{ for all } x, y \in \Omega.$$

با توجه به نتایج ارائه شده در پاول [19] می‌دانیم که طول گام  $d_k$  به دست آمده از **رابطه (۱۱)** رابطه زیر را برآورده می‌کند:

$$Pred_k = m_k(0) - m_k(d_k) \geq \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}; \text{ for all } k \geq 0, \quad (12)$$

که در آن  $\gamma \in (0, 1)$  یک ثابت است.

در ادامه یک تحلیل همگرایی ارائه می‌کنیم که به **فرضیه ۱** و **فرضیه ۲** اجازه می‌دهد تا در باقی مانده این مقاله برقرار باقی بمانند.

**لم ۱-** فرض کنید دنباله  $\{x_k\}$  توسط الگوریتم ۱ ایجاد شده است، بنابراین داریم

$$|f_k - f(x_{k+1}) - Pred_k| \leq O(\|d_k\|^2). \quad (13)$$

اثبات: با استفاده از بسط تیلور همراه با **فرضیه ۲** به دست می‌آوریم

$$\begin{aligned} |f_k - f(x_{k+1}) - Pred_k| &= |-d_k^T H_k d_k + d_k^T B_k d_k| + O(\|d_k\|^2) = |d_k^T (B_k - H_k) d_k| \\ &\leq (L + M)\|d_k\|^2 + O(\|d_k\|^2) = O(\|d_k\|^2), \end{aligned}$$

بنابراین اثبات کامل می‌شود.

**لم ۲-** فرض کنید که دنباله  $\{x_k\}$  توسط الگوریتم ۱ تولید شده باشد. به علاوه، اگر ثابت  $\varepsilon > 0$  وجود داشته باشد به طوری که برای همه  $k \leq \infty$  به اندازه کافی بزرگ  $\varepsilon$   $\|g_k\| \geq \varepsilon$ ، بنابراین داریم

$$\lim_{k \rightarrow \infty} \alpha_k = 0.$$

اثبات: قرار می‌دهیم  $J = \{k | r_k \geq \mu_0\}$ ، پس با استفاده از **رابطه (۳)** و **رابطه (۱۲)** داریم

$$\infty > \sum_{k=1}^{\infty} (f(x_k) - f(x_{k+1})) \geq \sum_{k \in J} \mu_0 Pred_k \geq \sum_{k \in J} \mu_0 \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\} \geq \sum_{k \in J} \mu_0 \gamma \varepsilon \min \left\{ \Delta_k, \frac{\varepsilon}{M} \right\}. \quad (14)$$

اگر  $J$  متناهی باشد؛ بنابراین از رابطه (۹) برای  $ks$  به اندازه کافی بزرگ داریم  $\alpha_{k+1} = c_1 \alpha_k$ . از آنجایی که  $0 < c_1 < 1$  بنابراین به دست می آوریم

$$\lim_{k \rightarrow \infty} \alpha_k = 0. \tag{۱۵}$$

اگر  $J$  نامتناهی باشد؛ بنابراین از رابطه (۱۴) داریم

$$\lim_{k \in J, k \rightarrow \infty} \Delta_k = \lim_{k \in J, k \rightarrow \infty} \alpha_k \|g_k\|^{\lambda(\delta)} = 0. \tag{۱۶}$$

توجه کنید که برای همه  $ks, \varepsilon, k \|g_k\| \geq \varepsilon$ ، از این رو رابطه (۱۵) نتیجه می دهد که  $\lim_{k \in J, k \rightarrow \infty} \alpha_k = 0$ ؛ بنابراین با استفاده از رابطه (۱۵) و  $\|d_k\| \leq \Delta_k$  و

$$\alpha_k \|g_k\|^{\lambda(\delta)}$$

$$\lim_{k \rightarrow \infty} d_k = 0, \tag{۱۷}$$

و به این صورت اثبات کامل می شود.

لم ۳- فرض کنید که دنباله  $\{x_k\}$  توسط الگوریتم ۱ تولید شده باشد، بنابراین برای همه  $ks$  داریم

$$f_{k+1} \leq D_{k+1} \leq D_k.$$

اثبات: فرض کنید تکرار  $k$ th توسط الگوریتم ۱ پذیرفته شده باشد، پس  $r_k > \mu_0$ ؛ بنابراین با استفاده از رابطه (۳) می توانیم بنویسیم

$$f_{k+1} \leq D_k - \mu_0 \text{Pred}_k \geq \mu_0 \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\},$$

پس به دست می آوریم

$$f_{k+1} \leq D_k. \tag{۱۸}$$

اکنون با استفاده از رابطه (۶) و رابطه (۱۲) داریم

$$\begin{aligned} D_{k+1} &= \eta_k D_k + (1 - \eta_k) f_{k+1} \leq \eta_k D_k + (1 - \eta_k) \left( D_k - \mu_0 \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\} \right) \\ &\leq D_k - (1 - \eta_k) \left( \mu_0 \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\} \right), \end{aligned}$$

که در این صورت به دست می آوریم

$$D_{k+1} \leq D_k. \tag{۱۹}$$

از رابطه (۶) داریم  $D_{k+1} - f_{k+1} = \eta (D_k - f_{k+1})$ . حال فرض کنید تکرار  $k$ th پذیرفته شده باشد، بنابراین داریم  $r_k > \mu_0$ . پس با استفاده از رابطه

(۷) و رابطه (۱۲) به دست می آوریم

$$D_k - f_{k+1} \geq \mu_0 \text{Pred}_k \geq \mu_0 \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\},$$

بنابراین می توان نوشت

$$D_k - f_{k+1} \geq \mu_0 \text{Pred}_k \geq \mu_0 \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}, \tag{۲۰}$$

که در نهایت با استفاده از رابطه (۱۸) تا رابطه (۲۰) به دست می آوریم  $f_{k+1} \leq D_{k+1} \leq D_k$ .

نتیجه ۱- فرض کنید دنباله  $\{x_k\}$  توسط الگوریتم ۱ تولید شده باشد، آنگاه دنباله  $\{D_k\}$  همگرا است.

اثبات: لم ۳ نشان می دهد که دنباله  $\{D_k\}$  غیر افزایشی و از پایین کراندار است؛ بنابراین دنباله  $\{D_k\}$  همگرا می باشد.

لم ۴- فرض کنید دنباله  $\{x_k\}$  توسط الگوریتم ۱ تولید شده باشد، آنگاه داریم

$$\lim_{k \rightarrow \infty} D_k = \lim_{k \rightarrow \infty} f_k.$$

اثبات: بنابر رابطه (۶) می‌توانیم بنویسیم

$$D_{k+1} = \eta_k D_k + (1 - \eta_k) f_{k+1},$$

بنابراین داریم

$$D_{k+1} - f_{k+1} = \eta (D_k - f_{k+1}),$$

$$D_{k+1} - D_k = \eta_k D_k + (1 - \eta_k) f_{k+1} - D_k = (\eta - 1) D_k + (1 - \eta) f_{k+1} = (\eta - 1) (D_k - f_{k+1}),$$

از این رو اگر  $\eta \neq 0$ ، می‌توان دریافت که

$$D_{k+1} - D_k = (\eta - 1) / \eta (D_{k+1} - f_{k+1}).$$

اکنون با استفاده از نتیجه ۱ و  $k \rightarrow \infty$  داریم

$$(\eta - 1) / \eta \lim_{k \rightarrow \infty} (D_{k+1} - f_{k+1}) = \lim_{k \rightarrow \infty} (D_{k+1} - D_k) = 0, \quad (21)$$

با توجه به این که  $\eta \in [0, 1]$  و رابطه (۲۱) اثبات کامل می‌شود.

قضیه ۱- فرض کنید دنباله  $\{x_k\}$  توسط الگوریتم ۱ تولید شده باشد، آنگاه الگوریتم ۱ تساوی زیر را برآورده می‌کند:

$$\lim_{k \rightarrow \infty} \inf \|g_k\| = 0. \quad (22)$$

اثبات: به برهان خلف فرض می‌کنیم رابطه (۲۲) برقرار نباشد؛ بنابراین یک ثابت مثبت  $\epsilon > 0$  وجود دارد به طوری که برای همه  $k$ s داریم

$$\|g_k\| \geq \epsilon, \quad (23)$$

ابتدا باید ثابت کنیم که  $r_k > \mu_0$ . فرض کنید  $d_k^p$  یک جواب زیر مساله رابطه (۱۱) مربوط به حلقه  $p$  در  $x_k$  باشد، از رابطه (۳) و رابطه (۱۲) به همراه لم ۱ داریم

$$\begin{aligned} |r_k - 1| &= \left| \frac{f(x_k) - f(x_{k+1}) - \text{Pred}_k}{\text{Pred}_k} \right| \\ &= \left| \frac{f(x_k) - f(x_{k+1}) + g_k^T d_k^p + \frac{1}{2} d_k^{pT} B d_k^p}{\text{Pred}_k} \right| \\ &\leq \frac{o(\|d_k\|^2)}{\gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}} \leq \frac{o(\|d_k\|^2)}{\gamma \epsilon \min \left\{ \Delta_k, \frac{\epsilon}{M} \right\}}. \end{aligned}$$

اکنون با استفاده از لم ۲ نتیجه می‌گیریم که چون کران بالای  $|r_k - 1|$  به صفر میل می‌کند؛ پس  $r_k \rightarrow 1$ . از این رو داریم  $r_k > \mu_0$ ؛ بنابراین با استفاده از لم ۳ به دست می‌آوریم

$$r_k = \frac{D_k - f(x_{k+1})}{\text{Pred}_k} > \frac{f(x_k) - f(x_{k+1})}{\text{Pred}_k} > \mu_0,$$

پس  $r_k > \mu_0$  که حاکی از آن است که ثابت مثبت  $\alpha^*$  وجود دارد به طوری که

$$\alpha_k \geq \alpha^*,$$

برای  $k$ s به اندازه کافی بزرگ برقرار است؛ بنابراین به یک تناقض با لم ۲ می‌رسیم؛ از این رو رابطه (۲۳) برقرار نیست و به این صورت اثبات کامل می‌شود.



قضیه ۲- فرض کنید  $\|g_k\| \neq 0$  و الگوریتم ۱ یک دنباله نامتناهی  $\{x_k\}$  را تولید کند، آنگاه الگوریتم ۱ تساوی زیر را برآورده می‌کند:

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \quad (24)$$

اثبات: در این جا فرض می‌کنیم که دنباله  $\{x_k\}_{k \geq 0}$  نامساوی  $r_k > \mu_0$  را برآورده کند. پس با استفاده از رابطه (۷)، رابطه (۱۰) و رابطه (۱۲) به دست می‌آوریم

$$\begin{aligned} D_k - f(x_{k+1}) &\geq \mu_0 \text{Pred}_k \geq \mu_0 \left( \gamma \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\} \right) = \mu_0 \left( \gamma \|g_k\| \min \left\{ \alpha_k \|g_k\|^{\lambda(\delta)}, \frac{\|g_k\|}{M} \right\} \right) \\ &\geq \mu_0 \left( \gamma \|g_k\| \min \left\{ \alpha_k \|g_k\|, \frac{\|g_k\|}{M} \right\} \right) = \mu_0 (\gamma \|g_k\|^2) \min \left\{ \alpha_k, \frac{1}{M} \right\} = \pi (\|g_k\|^2) \geq 0 \end{aligned} \quad (25)$$

که در آن  $\pi = \gamma \mu_0 \min \left\{ \alpha_k, \frac{1}{M} \right\}$ ؛ بنابراین از لم ۴ برای  $k$ های به اندازه کافی بزرگ داریم  $D_k - f(x_{k+1}) \rightarrow 0$ ، بنابراین اثبات کامل می‌شود.

#### ۴- نرخ همگرایی

در این بخش به بررسی نرخ همگرایی الگوریتم جدید می‌پردازیم. برای اثبات نرخ همگرایی درجه دوم الگوریتم ارائه شده علاوه بر فرضیه ۱ و فرضیه ۲ لازم است که فرضیه‌های زیر را نیز در نظر بگیریم.

فرضیه ۳- برخی ثابت‌های  $\vartheta_0 > 0$  و  $\theta_1 \in (0, 1)$  وجود دارند به طوری که برای همه  $x, y \in N(x^*, \theta_1)$  داریم

$$\|g(x) - g(y) + H(x-y)\| \leq \vartheta_0 \|x - y\|^2,$$

که در آن  $x^*$  یک جواب رابطه (۱) است و  $N(x^*, \theta_1) = \{x \mid \|x - x^*\| \leq \theta_1\}$ .

$$\vartheta_1 \|x - x^*\| \leq \|g(x)\| = \|g(x) - g(x^*)\|,$$

فرضیه ۴- برخی ثابت‌های  $\vartheta_1 \geq 1/\alpha_k$  و  $\theta_2 \in (0, 1)$  وجود دارند به طوری که برای همه  $x, y \in N(x^*, \theta_2)$  داریم

که در آن  $x^*$  یک جواب رابطه (۱) است و  $N(x^*, \theta_2) = \{x \mid \|x - x^*\| \leq \theta_2\}$ .

در ادامه قرار می‌دهیم،  $\theta = \min\{\theta_1, \theta_2\}$ . اگر  $g(x)$  و  $H(x)$  به ترتیب پیوسته مشتق‌پذیر و پیوسته لیپ‌شوتز باشند، آنگاه فرضیه ۲ برقرار است.

قضیه ۳- فرض کنید دنباله  $\{x_k\}$  تولیدشده توسط الگوریتم ۱ همگرا به  $x^*$  باشد و فرضیه ۱ تا فرضیه ۴ برقرار باشند. اگر  $k$  به اندازه کافی بزرگ باشد می‌توانیم قرار دهیم

$$x_{k+1} = x_k + d_k,$$

به‌علاوه دنباله  $\{x_k\}$  به  $x^*$  دارای همگرایی درجه دوم است.

اثبات: طبق الگوریتم ۱ هنگامی که الگوریتم خاتمه نیافته است داریم  $\|g_k\| > \varepsilon$ ، بنابراین طبق اثبات قضیه ۱ داریم  $r_k > \mu_0$  این نشان می‌دهد که اگر  $k$  به اندازه کافی بزرگ باشد، یک ثابت مثبت  $\alpha^*$  وجود دارد به طوری که  $\alpha_k \geq \alpha^*$  و این بدان معنی است که گام آزمایشی  $d_k$  توسط الگوریتم ۱ پذیرفته شده است. با توجه به فرضیه ۱ واضح است که ثابت  $M_1 > 0$  وجود دارد به طوری که

$$H_k \leq M_1, \quad (26)$$

برای همه  $x \in N(x^*, \theta)$  و  $\xi \in [x_k, x^*]$ ؛ بنابراین از رابطه (۲۵) و قضیه مقدار میانگین به دست می‌آوریم

$$\|g(x_k)\| = \|g(x) - g(x^*)\| \leq \|H(\xi)\| \|x_k - x^*\| \leq M_1 \|x_k - x^*\|. \quad (27)$$

از طرفی با استفاده از رابطه (۱۰) و فرضیه ۱ داریم

$$\|d_k\| \leq \Delta_k = \alpha_k \|g(x_k)\|^{1-\delta} = \alpha_k \|g(x) - g(x^*)\|^{1-\delta} \leq \alpha_k L^{1-\delta} \|x_k - x^*\|^{1-\delta}. \quad (28)$$

برای این‌که نشان دهیم  $x_k - x^*$  یک جواب شدنی رابطه (۱۱) است، از رابطه (۱۰) به همراه فرضیه ۴ استفاده می‌کنیم که در این صورت داریم

$$\|x_k - x^*\| \leq \frac{1}{\vartheta_1} \|g(x)\| \leq \alpha_k \|g(x_k)\|^{1-\delta} = \Delta_k.$$

اکنون با استفاده از فرضیه ۳ و فرضیه ۴ به همراه رابطه (۲۸) می‌توان نشان داد که

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq \|g(x_{k+1})\| = \|g(x_k + d_k)\| \\ &\leq \|g(x_k) + H d_k\| + O(\|d_k\|^2) \\ &= \|g(x_k) - g(x^*) + H d_k\| + O(\|d_k\|^2) \\ &\leq O(\|x_k - x^*\|^2) + O(\|d_k\|^{2\lambda\delta}) = O(\|x_k - x^*\|^{2\lambda\delta}), \end{aligned}$$

بنابراین یک ثابت صحیح  $\tau$  وجود دارد به طوری که

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^{2\lambda\delta}} = \lim_{k \rightarrow \infty} \frac{O(\|x_{k+1} - x^*\|)}{\|x_k - x^*\|^{2\lambda\delta}} \leq \tau,$$

در نتیجه دنباله  $\{x_k\}$  تولیدشده توسط الگوریتم ۱ دارای همگرایی درجه دوم است.

ملاحظه ۳- با در نظر گرفتن مقادیر تابع  $\lambda(\delta)$  حتی می‌توانیم به نرخ همگرایی بیشتر از درجه دو نیز دست پیدا کنیم.

## ۵- نتایج عددی

در این بخش برخی از نتایج عددی به دست آمده توسط الگوریتم ۱ را برای نشان دادن کارایی آن ارائه می‌کنیم. کارایی الگوریتم ارائه شده با دو الگوریتم زیر مقایسه می‌شود:

۱- الگوریتم NATR برگرفته از کمندی و امینی [20].

۲- الگوریتم NTR برگرفته از فان و یوان [18].

کد الگوریتم‌ها در نرم‌افزار متلب ۲۰۱۴ با مشخصات *R 2014a (8.3.0.523)* نوشته شده و اجرای کدها توسط رایانه شخصی با مشخصات *(Intel (R) Celeron(R) CPU N3350 2.4 GHZ)* با *4GB RAM* و در ویندوز ۱۰ انجام شده است.

تابع‌های مورد تست از مور و همکاران [2] گرفته شده که در جدول ۱ نام و بعد آن‌ها آورده شده است؛ که در آن،  $n_i$  بیانگر تعداد تکرارها و  $n_f$  بیانگر تعداد محاسبات تابع توسط الگوریتم‌ها می‌باشند. علامت (-) نشان‌دهنده آن است که تعداد تکرارها بیشتر از ۲۰۰۰ می‌باشد.

جدول ۱- لیست توابع تست و نتایج عددی.

Table 1- List of test functions and numerical results.

| نام تابع              | بعد | $n_i/n_f$ |       |         |
|-----------------------|-----|-----------|-------|---------|
|                       |     | NATR      | NTR   | NTRAR   |
| Helical valley        | 3   | 24.37     | 29.55 | 26.40   |
| Biggs EXP6            | 6   | 41.44     | 38.42 | 41.54   |
| Gaussian              | 3   | 3.5       | 5.6   | 4.5     |
| Powell badly scaled   | 2   | 69.89     | 76.90 | 104.120 |
| Box three-dimensional | 3   | 25.29     | 30.38 | 14.17   |
| Variably dimensioned  | 10  | 18.26     | 16.26 | 14.25   |
|                       | 100 | 32.39     | 35.57 | 37.64   |
|                       | 500 | 46.53     | 46.76 | 49.88   |

جدول ۱- ادامه.

Table 1- Continued.

| نام تابع                               | بعد  | $n_i/n_f$ |         |          |
|--|------|-----------|---------|----------|
|  |      | NATR      | NTR     | NTRAR    |
| Watson                                 | 31   | 5371      | 58.90   | 56.78    |
| Penalty function I                     | 10   | 112.134   | 50.60   | 45.53    |
|  | 100  | 144.169   | 57.84   | 50.65    |
|  | 1000 | 88.92     | 69.87   | 149.189  |
|  | 10   | 311.375   | 192.303 | 250.424  |
| Penalty function II                    | 10   | 311.375   | 192.303 | 250.424  |
| Brown badly scaled                     | 2    | 1679.1680 | (-)     | 24.55    |
| Brown and Dennis                       | 4    | 30.37     | (-)     | 29.63    |
| Gulf research and dev<br>Trigonometric | 3    | 47.53     | 42.46   | 39.43    |
|  | 10   | 27.29     | 26.40   | 25.26    |
|  | 100  | 43.45     | 44.55   | 43.44    |
| Extended Rosenbrock                    | 1000 | 55.57     | 50.62   | 46.49    |
|  | 10   | 96.121    | 95.189  | 78.160   |
|  | 100  | 405.546   | 218.559 | 204.522  |
| Extended Powell singular               | 1000 | 748.1066  | 264.682 | 241.603  |
|  | 20   | 69.96     | 81.86   | 52.71    |
|  | 100  | 215.288   | 241.604 | 218.553  |
| Beale                                  | 100  | 1722.2368 | 354.887 | 436.1074 |
|  | 2    | 14.20     | 13.16   | 13.16    |

همه الگوریتم‌ها هر زمان که شرط  $\|g_k\| < 10^5$  را برآورده کنند یا حداکثر تکرارها از ۲۰۰۰ تجاوز کند خاتمه می‌یابند. برای داشتن شرایط مشابه برای مقایسه الگوریتم‌ها، از الگوریتم اشتیهاگ [21] برای حل زیر مساله درجه دوم ناحیه اعتماد استفاده می‌کنیم. در آزمایش‌های عددی، مقادیر اولیه زیر را در نظر می‌گیریم:

$$\mu_0 = 0.1, \mu_1 = 0.25, \mu_2 = 0.75, c_1 = 0.25, c_2 = 14, N = 10^5, \alpha_0 = 0.138, \delta = 1000.$$

با توجه به آزمایش‌های عددی انجام‌شده برای دست‌یابی به بهترین نتیجه، تابع  $\lambda(\delta)$  در الگوریتم NTRAR را به صورت زیر تعریف می‌کنیم:

$$\eta_k = \begin{cases} \eta_0/2, & k = 1, \\ \eta_{k-1} + \eta_{k-2}/2, & k = 2. \end{cases}$$

سایر مقادیر مربوط به پارامترهای الگوریتم NATR مانند کمندی و امینی [11] می‌باشد. در الگوریتم NTRAR قرار داده شده  $\eta_0 = 0.85$  و  $\eta_k$  به صورت زیر به روزرسانی می‌شود:

$$\lambda(\delta) = \begin{cases} 1.07, & 100 \leq \delta < 100, \\ 1.2, & 100 \leq \delta < 10, \\ 1.279, & 10 \leq \delta < 1, \\ 1.299, & 1 \leq \delta < 10^{-5}, \\ 1.34, & \text{otherwise.} \end{cases}$$

در همه الگوریتم‌ها برای به روزرسانی ماتریس هسی از فرمول به روزرسانی BFGS که به صورت زیر بیان شده است استفاده می‌شود:

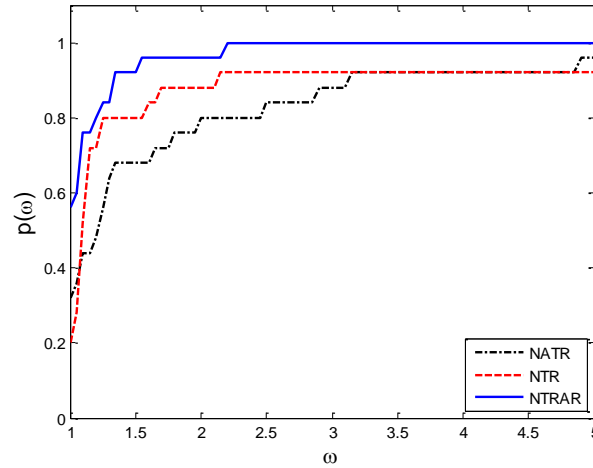
$$B_{k+1} = B_k - \frac{B_k d_k d_k^T B_k}{d_k^T B_k d_k} + \frac{y_k y_k^T}{y_k^T d_k}.$$

در الگوریتم NTR مقدار  $\alpha_k$  به صورت زیر به روزرسانی می‌شود:

$$\alpha_{k+1} = \begin{cases} c_1 \alpha_k, & r_k < \mu_1, \\ c_2 \alpha_k, & \mu_1 < r_k \text{ and } \|d_k\| > 1/2\Delta_k, \\ \alpha_k, & \text{otherwise.} \end{cases}$$

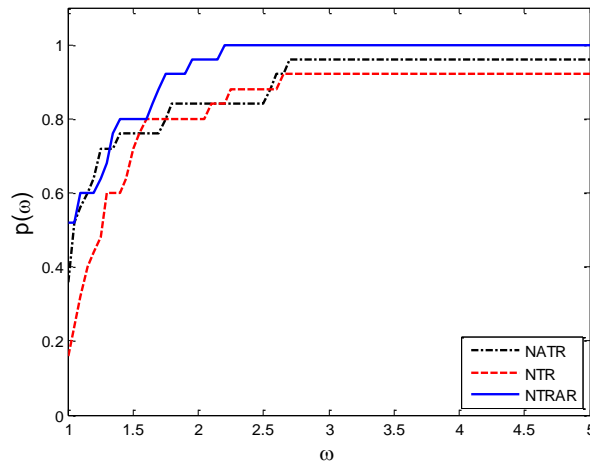
با توجه به مقادیر جدول ۱، مشخص است که الگوریتم NTRAR در تعداد تکرارها در ۶۵٪ موارد بهتر از الگوریتم NATR است و در ۶۹٪ موارد الگوریتم NTRAR از الگوریتم NATR نیز بهتر است. الگوریتم NTRAR در ۵۳٪ موارد از نظر تعداد محاسبات تابع بهتر از الگوریتم NATR است و در ۶۹٪ موارد الگوریتم NTRAR بهتر از الگوریتم NTR است.

ما برای نمایش میزان کارایی الگوریتم ارایه شده و مقایسه آن با سایر الگوریتم‌ها به صورت بصری از روش دولان و موره [21] استفاده می‌کنیم. شکل ۱ تا شکل ۳ نتایج مقایسه سه الگوریتم را به تصویر کشیده است. در شکل ۱، تعداد تکرارهای الگوریتم‌ها مقایسه شده و همان‌طور که مشاهده می‌شود، عملکرد الگوریتم *NTRAR* بهترین است. در شکل ۲، مقایسه تعداد ارزیابی‌های تابع برای الگوریتم‌ها نشان داده شده و به روشنی مشاهده می‌شود که عملکرد الگوریتم *NTRAR* بهتر از عملکرد دو الگوریتم دیگر می‌باشد. در شکل ۳، زمان اجرا توسط الگوریتم‌ها به تصویر کشیده شده و واضح است که الگوریتم *NTRAR* بهتر از سایرین است.



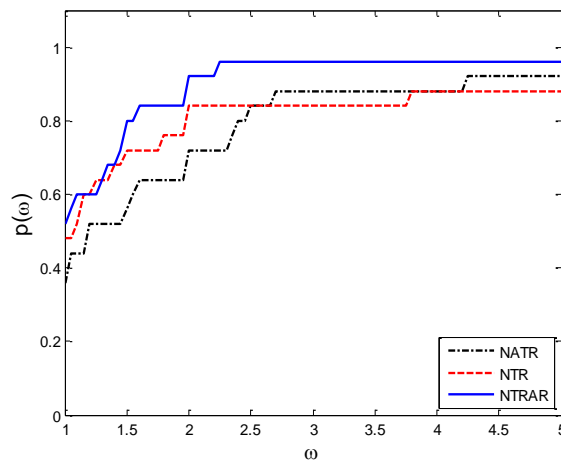
شکل ۱- مشخصات عملکرد برای تعداد تکرارها.

Figure 1- Performance profiles for the number of iterations.



شکل ۲- مشخصات عملکرد تعداد برای ارزیابی‌های تابع.

Figure 2- Performance profiles for the number of function evaluations.



شکل ۳- مشخصات عملکرد برای زمان اجرا.

Figure 3- Performance profiles for the running times.

## ۶- نتیجه گیری

در این مطالعه، یک الگوریتم ناحیه اعتماد غیر یکنوا با شعاع وقتی ارزیابی شده برای جلوگیری از نادیده گرفتن تاثیر شعاع در نزدیکی جواب، از یک شعاع وقتی همگرا به صفر که از یک تابع کاهش می‌ربوط به شعاع بهره می‌برد استفاده کرده است. به علاوه برای غلبه بر اثر نامطلوب یک‌نوایی، از پارامتر غیر یکنوا ارزیابی شده توسط شی و ژو [13] استفاده کرده‌ایم. همچنین در شرایط مناسبی اثبات همگرایی سراسری و نرخ همگرایی درجه دوم الگوریتم جدید انجام شده است. در نهایت، عملکرد الگوریتم ارزیابی شده را با یکی از جدیدترین الگوریتم‌های معرفی شده توسط کمندی و امینی [20] و همچنین الگوریتم ارزیابی شده توسط فان و یوان [18] مقایسه کرده‌ایم. نتایج عددی نشان داده است که الگوریتم *NTRAR* از کارایی و نیرومندی بهتری نسبت به هر دو الگوریتم قبلی برخوردار است.

## تعارض با منافع

نویسندگان این مقاله اعلام می‌دارند که هیچ تضادی در منافع در مورد انتشار این نسخه وجود ندارد، همچنین نویسندگان، نسخه نهایی ارسال شده را مشاهده و تایید می‌کنند. نویسندگان تضمین می‌کنند که مقاله، اثر اصلی آن‌ها بوده، قبلاً چاپ نشده و در حال حاضر تحت انتشار نمی‌باشد.

## منابع

- [1] Zang, I. (1978). A new arc algorithm for unconstrained optimization. In *Mathematical programming* (Vol. 15, pp. 36–52). Elsevier. DOI: 10.1007/BF01608998
- [2] Moré, J. J., Garbow, B. S., & Hillstom, K. E. (1981). Testing unconstrained optimization software. *ACM transactions on mathematical software (TOMS)*, 7(1), 17–41. DOI:10.1145/355934.355936
- [3] Rezaee, S., & Babaie-Kafaki, S. (2019). An adaptive nonmonotone trust region algorithm. *Optimization methods and software*, 34(2), 264–277. DOI:10.1080/10556788.2017.1364738
- [4] Esmaeili, H., & Kimiaei, M. (2014). An improved adaptive trust-region method for unconstrained optimization. *Mathematical modelling and analysis*, 19(4), 469–490. DOI:10.3846/13926292.2014.956237
- [5] Zhang, X., Zhang, J., & Liao, L. (2002). An adaptive trust region method and its convergence. *Science in china, series a: mathematics, physics, astronomy*, 45(5), 620–631. DOI:10.1360/02ys9067
- [6] Li, G. D. (2006). A trust region method with automatic determination of the trust region radius. *Chinese journal of engineering mathematics*, 23, 843–848.
- [7] Bahrami, S., & Amini, K. (2020). An efficient two-step trust-region algorithm for exactly determined consistent systems of nonlinear equations. *Journal of computational and applied mathematics*, 367, 112470. DOI:10.1016/j.cam.2019.112470
- [8] Grippo, L., Lampariello, F., & Lucidi, S. (1986). Nonmonotone line search technique for newton's method. *Siam journal on numerical analysis*, 23(4), 707–716. DOI:10.1137/0723046
- [9] Zhang, H., & Hager, W. W. (2004). A nonmonotone line search technique and its application to unconstrained optimization. *Siam journal on optimization*, 14(4), 1043–1056. DOI:10.1137/S1052623403428208
- [10] Mo, J., Liu, C., & Yan, S. (2007). A nonmonotone trust region method based on nonincreasing technique of weighted average of the successive function values. *Journal of computational and applied mathematics*, 209(1), 97–108. DOI:10.1016/j.cam.2006.10.070
- [11] Gu, N. zhu, & Mo, J. tao. (2008). Incorporating nonmonotone strategies into the trust region method for unconstrained optimization. *Computers and mathematics with applications*, 55(9), 2158–2172. DOI:10.1016/j.camwa.2007.08.038
- [12] Amini, K., & Ahookhosh, M. (2014). A hybrid of adjustable trust-region and nonmonotone algorithms for unconstrained optimization. *Applied mathematical modelling*, 38(9–10), 2601–2612. DOI:10.1016/j.apm.2013.10.062
- [13] Yingliang, Z., & Chengxian, X. (2000). A new trust region dogleg method for unconstrained optimization. *Applied mathematics*, 15(1), 83–92. DOI:10.1007/s11766-000-0013-8
- [14] Niri, T. D., Heydari, M., & Hosseini, M. M. (2020). Two nonmonotone trust region algorithms based on an improved Newton method. *Journal of applied mathematics and computing*, 64(1–2), 179–194. DOI:10.1007/s12190-020-01350-7
- [15] Qu, Q., Ding, X., & Wang, X. (2020). A filter and nonmonotone adaptive trust region line search method for unconstrained optimization. *Symmetry*, 12(4), 1–12. DOI:10.3390/SYM12040656
- [16] Akbari, Z. (2020). A new trust region method for minimizing locally lipschitz functions. *Journal of decisions and operations research*, 5(2), 204–220. (In Persian). [https://www.journal-dmor.ir/article\\_120044\\_en.html](https://www.journal-dmor.ir/article_120044_en.html)
- [17] Liu, J., & Ma, C. (2013). A nonmonotone trust region method with new inexact line search for unconstrained optimization. *Numerical algorithms*, 64(1), 1–20. DOI:10.1007/s11075-012-9652-0
- [18] Fan, J., Pan, J., & Song, H. (2016). A retrospective trust region algorithm with trust region converging to zero. *Journal of computational mathematics*, 34(4), 421–436. DOI:10.4208/jcm.1601-m2015-0399
- [19] Powell, M. J. D. (1975). Convergence properties of a class of minimization algorithms. In *Nonlinear programming 2* (pp. 1–27). Elsevier. DOI: 10.1016/b978-0-12-468650-2.50005-5
- [20] Kamandi, A., & Amini, K. (2022). A new nonmonotone adaptive trust region algorithm. *Applications of mathematics*, 67(2), 233–250. DOI:10.21136/AM.2021.0122-20
- [21] Dolan, E. D., & Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematical programming, series b*, 91(2), 201–213. DOI:10.1007/s101070100263