مجله تحقیق در عملیات در کاربردهای آن، سال شانزدهم، شماره دوم (پیاپی ۶۱)، تابستان ۹۸، ص ص ۹۵-۸۹
شاپا ۷۲۸۶-۲۲۵۱

یک روش جدید برای انتخاب پارامترهای تنظیم شعاع در روش‌های ناحیه اطمینان

احمد کمندی[۱]، کیوان امینی[۲*]
۱- استادیار، دانشگاه علم و فناوری مازندران، گروه ریاضی، مازندران، بهشهر، ایران
۲- استاد، دانشگاه رازی، گروه ریاضی، کرمانشاه، ایران

رسید مقاله: ۳۱ شهریور ۱۳۹۶
پذیرش مقاله: ۱۰ اسفند ۱۳۹۷

چکیده
خانواده روش‌های ناحیه اطمینان از جمله مهم‌ترین روش‌های کارآمد برای حل مسایل بهینه‌سازی نامقید است. کارایی این دسته از روش‌ها وابستگی زیادی به انتخاب مناسب پارامترهای اولیه، به‌خصوص پارامترهای تنظیم شعاع، دارد. در این مقاله یک روش جدید برای انتخاب پارامترهای تنظیم شعاع در روش‌های ناحیه اطمینان ارایه شده است. نتایج عددی حاصل از آزمون ایده جدید روی مجموعه‌ی مسایل کیوتست [۷] نشان می‌دهند که استفاده از روش جدید باعث افزایش کارایی روش‌های ناحیه اطمینان می‌شود.

کلمات کلیدی: بهینه‌سازی نامقید، روش‌های ناحیه اطمینان، پارمترهای تنظیم شعاع.

۱ مقدمه
مساله بهینه‌سازی نامقید زیر را در نظر بگیرید:

$$\min_{x \in \mathbb{R}^n} \; f(x) \qquad (۱)$$

که در آن $f:\mathbb{R}^n \to \mathbb{R}$ یک نگاشت پیوسته و مشتق‌پذیر است. بسیاری از مسایل دنیای واقعی را می‌توان به شکل مساله (۱) فرموله کرد. همچنین یکی از موثرترین روش‌ها برای حل دستگاه های غیرخطی استفاده از روش کم‌ترین مربعات است که مساله دستگاه غیرخطی را به یک مساله بهینه‌سازی نامقید تبدیل می‌کند[۱]. درحالت کلی حل دقیق مساله (۱) از نظر تئوری می‌تواند بسیار پیچیده و دشوار باشد، به همین دلیل از روش‌های عددی تکراری برای حل این مساله استفاده می‌شود.

* عهده‌دار مکاتبات
آدرس الکترونیکی: kamini@razi.ac.ir

از میان روش‌های عددی برای حل مساله (۱) روش‌های ناحیه اطمینان به دلیل خواصی چون همگرایی سراسری، سرعت همگرایی بالا و نتایج عددی مطلوب از اهمیت ویژه‌ای برخوردارند [۲]. ایده اصلی این دسته از روش‌ها تقریب تابع هدف با استفاده از یک مدل مجذوری و محدود کردن این مدل به ناحیه‌ای موسوم به ناحیه اطمینان، در هرتکرار، است. در تکرار kام روش‌های ناحیه اطمینان، مساله زیر که به زیرمساله ناحیه اطمینان معروف است، حل می‌شود.

$$\begin{aligned} \text{Min} \quad & M_k(s) = f_k + g_k^T s + \frac{1}{2} s^T B_k s \\ s.t. \quad & \|s\| \le \Delta_k \end{aligned} \tag{۲}$$

که در آن $f_k = f(x_k)$ ، $g_k = \nabla f(x_k)$ و B_k ماتریس دقیق هسی (یا تقریبی از آن) در نقطه x_k و Δ_k یک پارامتر مثبت است که به آن شعاع ناحیه اطمینان می‌گویند.

نکته‌ای که در مورد حل زیر مساله ناحیه اطمینان باید به آن اشاره کرد این است که در برخی موارد حل دقیق زیر مساله ناحیه اطمینان بسیار پیچیده و پرهزینه است؛ لذا در بسیاری موارد این مساله به طور تقریبی حل می‌گردد. در حالتی که s_k جواب تقریبی زیر مساله (۲) باشد و در شرط:

$$M_k(0) - M_k(s_k) \ge \frac{1}{2}\|g_k\| \text{Min}\left\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right\}$$

صدق کند. آنگاه می‌توان همگرایی سراسری روش‌های ناحیه اطمینان را ثابت کرد [۲].

فرض کنید s_k جواب زیر مساله (۲) باشد. در روش‌های ناحیه اطمینان پارامتر آزمون نسبت به صورت زیر تعریف می‌شود.

$$r_k = \frac{f_k - f(x_k + s_k)}{M_k(0) - M_k(s_k)} \tag{۳}$$

به ازای یک پارامتر مثبت $\mu_0 \in (0,1)$ در حالتی که $r_k \ge \mu_0$ گام kام ناحیه اطمینان یک گام موفق نامیده می‌شود. در این حالت نقطه $x_k + s_k$ به عنوان تقریب جدید پذیرفته شده و شعاع ناحیه اطمینان برای تکرار بعدی افزایش می‌یابد. در غیر این این صورت گام ناموفق است و شعاع ناحیه اطمینان کاهش می‌یابد.

فرم کلی یک الگوریتم ناحیه اطمینان به شرح زیر است:

الگوریتم ۱

- **ورودی**: پارامترهای اولیه $\varepsilon > 0$ ، $0 < \gamma_1 < 1 < \gamma_2$ ، $0 < \mu_0 < \mu_1 < 1$ و $\bar{\Delta} > 0$.
- **گام ۰** : تقریب اولیه x_0 و ماتریس آغازین B_0 را انتخاب کن و قرار ده $k = 0$.
- **گام ۱**: اگر $\|g_k\| \le \varepsilon$، آنگاه x_k را به عنوان جواب معرفی و توقف کن.
- **گام ۲**: از حل زیر مساله (۲) جهت s_k و با استفاده از آن پارامتر نسبت r_k را از رابطه (۳) محاسبه کن.
- **گام ۳**: اگر $r_k < \mu_0$ قرار ده $\Delta_k = \gamma_1 \|s_k\|$ و به گام ۲ برو. در غیر این صورت قرار ده $x_{k+1} = x_k + s_k$.
- **گام ٤**: اگر $r_k \ge \mu_1$ قرار ده $\Delta_{k+1} = \text{Min}\left\{\gamma_2 \Delta_k, \bar{\Delta}\right\}$. در غیر این صورت قرار ده $\Delta_{k+1} = \Delta_k$.

[DOR: 20.1001.1.22517286.2019.16.2.3.4]

- **گام ۵**: ماتریس تقریب هسی B_{k+1} را بهنگام کن.
- **گام ۶**: k را یک واحد افزایش بده و به گام ۱ برو.

کارایی روش‌های ناحیه اطمینان به طور محسوسی به انتخاب پارامترهای اولیه وابسته است. پارامترهای اولیه استاندارد پیشنهاد شده به صورت زیر می‌باشند [۳، ۴]:

$$\mu_0 = ۰/۲۵,\ \mu_1 = ۰/۷۵,\ \gamma_1 = ۰/۵,\ \gamma_2 = ۲ \tag{۴}$$

درسال ۲۰۰۵، گلد و همکارانش حساسیت روش‌های ناحیه اطمینان را نسبت به پارامترهای اولیه مورد بررسی قرار دادند[۵]. پس از انجام آزمایش‌های عددی بر روی ۲۴ مساله بهینه‌سازی نامقید با در نظر گرفتن حدود ۴۰۰۰ پارامتر اولیه مختلف گلد و همکارانش پارامترهای زیر را به عنوان بهترین انتخاب معرفی کردند:

$$\mu_0 = ۰/۰۱,\ \mu_1 = ۰/۹۹,\ \gamma_1 = ۰/۲۵,\ \gamma_2 = ۳/۵ \tag{۵}$$

در این مقاله ما پارامترهای γ_1 و γ_2 را پارامترهای تنظیم شعاع می‌نامیم. آزمایش‌های عددی نشان می‌دهند که انتخاب پارامترهای تنظیم شعاع ثابت برای الگوریتم ناحیه اطمینان نمی‌تواند انتخاب هوشمندانه‌ای باشد. برای رفع این مشکل در این مقاله یک روش جدید برای انتخاب پارامترهای تنظیم شعاع پیشنهاد می‌گردد. در روش جدید به جای انتخاب پارامترهای ثابت برای تنظیم شعاع، پارامترها به صورت تابعی وابسته شعاع در هر تکرار انتخاب می‌شوند. نتایج عددی نشان می‌دهند که به کارگیری روش جدید انتخاب پارامترهای تنظیم شعاع باعث افزایش کارایی روش‌های ناحیه اطمینان می‌شود.

در ادامه مقاله، در بخش ۲ پارامترهای تنظیم شعاع جدید معرفی می‌شوند. نتایج عددی حاصل از اعمال ایده جدید بر روش‌های ناحیه اطمینان در حل مجموعه مسایل کیوتست در بخش ۳ گزارش شده است. در بخش پایانی نیز نتیجه‌گیری کار بیان شده است.

۲ پارامترهای جدید تنظیم شعاع

همان‌طور که در بخش قبل اشاره شده استفاده از پارامترهای تنظیم شعاع ثابت ممکن است باعث کاهش کارایی یک الگوریتم ناحیه اطمینان گردد؛ زیرا در برخی از مسایل، چون رفتار تابع هدف مشابه یک مدل درجه ۲ است، شعاع ناحیه اطمینان بسیار بزرگ می‌شود و اگر از پارامتر کاهش شعاع بزرگ (نزدیک به ۱) در گام‌های ناموفق استفاده شود ممکن است مجبور شویم چندین بار شعاع را کاهش دهیم و زیر مساله ناحیه اطمینان را حل کنیم تا تقریب جدید را بسازیم. همچنین استفاده از پارامتر افزایش شعاع بزرگ در گام‌های موفق هنگامی که شعاع ناحیه اطمینان بزرگ است می‌تواند باعث افزایش نامطلوب شعاع گردد.

از طرفی در برخی از مسایل به علت رفتار غیرخطی شدید تابع هدف، شعاع ناحیه اطمینان بسیار کوچک می‌شود. در این گونه از مسایل باید پارامترهای تنظیم شعاع به گونه‌ای باشند که از کاهش نامطلوب شعاع ناحیه اطمینان در حد امکان جلوگیری کنند. برای این منظور پارامترهای تنظیم شعاع باید در حد امکان بزرگ باشند.

پس از انجام آزمایش‌های عددی دقیق روی مسایل مختلف بهینه‌سازی نامقید و با توجه به مباحث بالا به این نتیجه رسیدیم که استفاده از پارامترهای وابسته شعاع می‌تواند در افزایش کارایی الگوریتم‌های ناحیه اطمینان موثر باشد؛ لذا پارامترهای جدید تنظیم شعاع را به صورت زیر پیشنهاد می‌کنیم:

$$\gamma_1=\gamma_1(\Delta_k),\ \gamma_2=\gamma_2(\Delta_k)$$

که درآن $\gamma_1(\Delta_k):[0,\bar{\Delta}]\to[c_1,c_2]$ و $\gamma_2(\Delta_k):[0,\bar{\Delta}]\to[w_1,w_2]$ توابعی نزولی هستند که $1<w_1<w_2$ و $0<c_1<c_2<1$. درحالت خاص با فرض $\bar{\Delta}=100$، $w_2=5$، $w_1=1/2$، $c_1=0/17$ و $c_2=0/9$ می‌توان توابع $\gamma_1(\Delta_k)$ و $\gamma_2(\Delta_k)$ به صورت توابع پله‌ای زیر در نظر گرفت:

$$\gamma_1(\Delta_k)=\begin{cases} 0/17 & 80<\Delta_k\le 100\\ 0/20 & 20<\Delta_k\le 80\\ 0/25 & 10^{-4}<\Delta_k\le 20\\ 0/30 & 10^{-8}<\Delta_k\le 10^{-4}\\ 0/90 & 0<\Delta_k\le 10^{-8}\end{cases} \tag{6}$$

$$\gamma_2(\Delta_k)=\begin{cases} 1/2 & 50<\Delta_k\le 100\\ 2/5 & 20<\Delta_k\le 50\\ 3 & 10<\Delta_k\le 20\\ 3/5 & 10^{-2}<\Delta_k\le 10\\ 4/5 & 10^{-8}<\Delta_k\le 10^{-2}\\ 5 & 0<\Delta_k\le 10^{-8}\end{cases} \tag{7}$$

توجه شود که توابع پله‌ای بالا توابعی نزولی، نسبت به شعاع ناحیه اطمینان، هستند. استفاده از این توابع برای تنظیم شعاع ناحیه اطمینان باعث می‌شود که تا حدودی از افزایش یا کاهش نامطلوب شعاع ناحیه اطمینان جلوگیری شود و در نتیجه کارایی روش ناحیه اطمینان افزایش یابد.

در این مقاله از ارایه قضایای مربوط به همگرایی سراسری الگوریتم ناحیه اطمینان با پارامترهای تنظیم شعاع پله‌ای صرف‌نظر می‌شود؛ زیرا پارامترهای پله‌ای اشاره شده در بالا کراندارند و همه شرایط همگرایی سراسری الگوریتم ناحیه اطمینان تحت آن‌ها حفظ می‌شوند؛ لذا تحت فرض‌های استاندارد الگوریتم ناحیه اطمینان با پارامترهای تنظیم شعاع پله‌ای نیز به نقطه بهینه موضعی مساله (۱) همگرا خواهد بود.

۳ نتایج عددی

برای نشان دادن کارایی الگوریتم ناحیه اطمینان با پارامترهای تنظیم شعاع جدید در حل مسایل بهینه‌سازی نامقید در مقایسه با روش ناحیه اطمینان با پارامترهای تنظیم شعاع ثابت، الگوریتم ۱ را در محیط نرم افزار متلب سیستم عامل لینوکس پیاده‌سازی کردیم.

در ادامه این بخش از نمادهای زیر استفاده خواهیم کرد:

[DOR: 20.1001.1.22517286.2019.16.2.3.4]

- BTR0 – الگوریتم ۱ با پارامترهای استاندارد (۴).
- BTR1 – الگوریتم ۱ با بهترین پارامترهای معرفی شده توسط گلد و همکارانش (۵).
- NEW0 – الگوریتم ۱ با پارامترهای تنظیم شعاع پله‌ای (۶) و (۷) و $\mu_1 = ۰/۷۵,\ \mu_0 = ۰/۲۵$.
- NEW1 – الگوریتم ۱ با پارامترهای تنظیم شعاع پله‌ای (۶) و (۷) و $\mu_1 = ۰/۹۹,\ \mu_0 = ۰/۰۱$.

برای این الگوریتم‌ها ماکزیمم شعاع ناحیه اطمینان به صورت $\bar{\Delta} = ۱۰۰$ ، شرط توقف به صورت: $\|g_k\| \le \varepsilon$ با $\varepsilon = ۱۰^{-۶}\|g_0\|$ و عدد ۴۰۰۰ به عنوان ماکزیمم تعداد تکرار در نظر گرفته شده است. زیر مساله ناحیه اطمینان در الگوریتم‌های پیاده‌سازی شده با استفاده از روش تکراری اشتیهاگ-توینت [۲] حل شده است. همچنین در هر تکرار برای محاسبه تقریب ماتریس هسی روش BFGS مورد استفاده قرار گرفته است [۶].

نتایج عددی از به کارگیری الگوریتم‌های اشاره شده برای حل ۱۵۷ مساله آزمون از [۷] که ابعاد آن‌ها بین ۵۰ تا ۱۰۰۰ بود، حاصل شده‌اند. برای هر مساله تقریب اولیه x_0، نقطه شروع استاندارد همان مساله در نظر گرفته شده است. تعداد کل مسایل آزمون ۱۷۲ مساله بودکه در ۱۵ تای آن‌ها هر چهار الگوریتم شکست خورده‌اند که این مسایل برای در مقایسه الگوریتم‌ها در نظر گرفته نشده‌اند. از میان ۱۵۷ مساله باقیمانده که حداقل یکی از چهار الگوریتم آن‌ها را حل کرده بودند، الگوریتم BTR0 در حل ۵ مساله و الگوریتم‌های BTR1 و NEW0 در حل ۲ مساله و الگوریتم NEW1 در حل ۳ مساله شکست خورده‌اند.

برای مقایسه هر چه بهتر نتایج عددی حاصل، مشابه [۸] ، از نمودار دولان و موری [۹] استفاده شده است. بدین منظور، مدت زمان اجرای الگوریتم، تعداد مقداردهی تابع و تعداد تکرارهای آن، به عنوان شاخص‌های کارایی در نظر گرفته شده است. از آنجا که در هر تکرار روش ناحیه اطمینان فقط یک بار گرادیان تابع محاسبه می‌شود، تعداد تکرارها با تعداد مقداردهی گرادیان برابر است.

شکل ۱ نمودار دولان و موری الگوریتم‌های بالا با شاخص کارایی تعداد مقدار دهی تابع هدف را نشان می‌دهد. همان‌گونه که از این شکل پیداست الگوریتم NEW1 در مقایسه با دیگر الگوریتم‌ها بهترین کارایی را دارد. الگوریتم‌های BTR1 و NEW0 کارایی تقریبا یکسانی دارند هرچند BTR1 در ۴۹ درصد از کل مسایل بهترین الگوریتم بوده در حالی که NEW0 در حدود ۲۵ درصد از مسایل کم‌ترین تعداد مقداردهی تابع را داشته است. همچنین، با شاخص تعداد مقداردهی تابع هدف الگوریتم BTR0 بدترین الگوریتم بوده و ضعیف‌ترین کارایی را از خود نشان داده است.

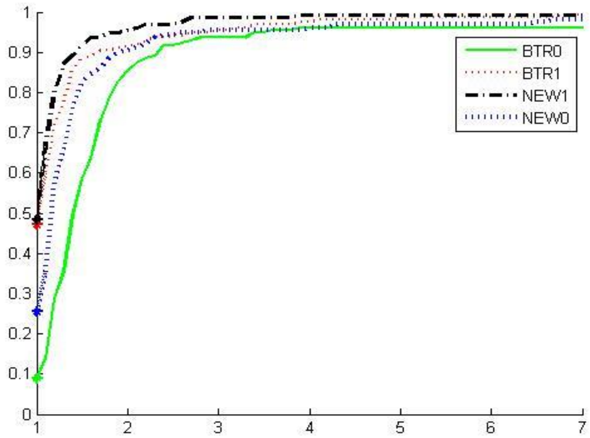
در شکل ۲ نمودار دولان و موری الگوریتم‌ها با شاخص کارایی تعداد تکرار رسم شده است. در این شکل نیز نمودار الگوریتمNEW1 در نهایت بالای نمودار دیگر الگوریتم‌ها قرار گرفته است که نشان از برتری این الگوریتم نسبت به سایر الگوریتم‌ها دارد. همچنین الگوریتم‌های BTR1، NEW0 و BTR0 به ترتیب در جایگاه دوم، سوم و چهارم قرار می‌گیرند.

[DOR: 20.1001.1.22517286.2019.16.2.3.4]

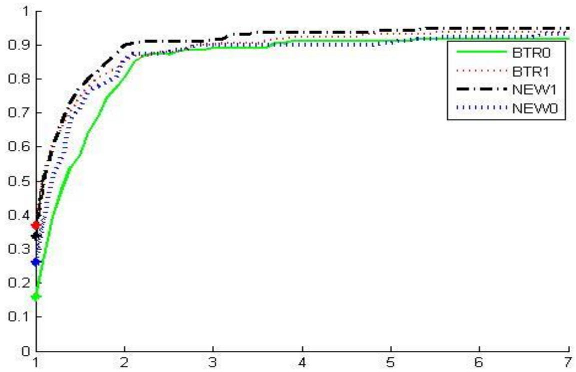
شاخص کارایی مهم دیگری که معمولاً در مقایسه الگوریتم‌ها استفاده می‌شود، مدت زمان اجرای الگوریتم‌ها در حل مسایل است. نمودار دولان و موری با این شاخص کارایی در شکل ۳ نمایش داده شده است. در این شکل مشاهده می‌شود که الگوریتم NEW1 از نظر مدت زمان اجرای الگوریتم‌ها نیز بهترین الگوریتم است. الگوریتم‌های BTR1 و NEW0 کارایی تقریبا یکسانی دارند و الگوریتم BTR0 بدترین کارایی را نسبت سایر الگوریتم‌ها دارد.

٤ نتیجه‌گیری

در این مقاله یک روش جدید برای انتخاب پارامترهای تنظیم شعاع در روش‌های ناحیه اطمینان ارایه شده است. در روش جدید به جای استفاده از پارامترهای تنظیم شعاع ثابت استفاده از پارامترهای تنظیم شعاع وابسته به شعاع پیشنهاد می‌گردد که یک نمونه ساده و موثر از آن‌ها با عنوان پارامترهای تنظیم شعاع پله‌ای مطرح شده است. نتایج حاصل از آزمایش‌های عددی در حل مجموعه‌ای از مسایل آزمون بهینه‌سازی نامقید از کیوتست نشان می‌دهد که استفاده از پارامترهای تنظیم شعاع پله‌ای کارایی الگوریتم ناحیه اطمینان را افزایش می‌دهد.

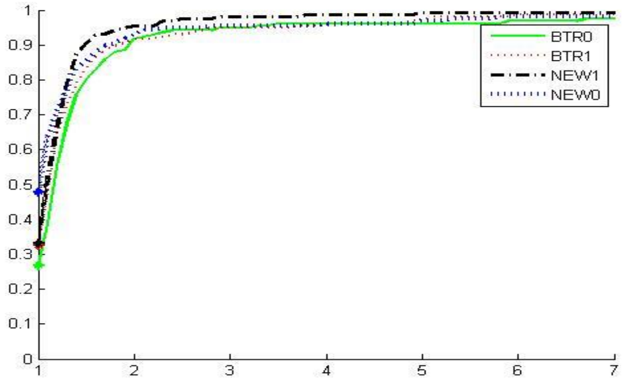


شکل ۱. نمودار دولان موری با شاخص تعداد مقداردهی تابع



شکل ۲. نمودار دولان موری با شاخص تعداد تکرار الگوریتم

[DOR: 20.1001.1.22517286.2019.16.2.3.4]



شکل۳. نمودار دولان موری با شاخص مدت زمان اجرای الگوریتم

منابع

[۱] امینی، ک، رستمی، ف، (۱۳۹۴). یک روش لونبرگ-مارکوارت اصلاح شده با یک جستجوی خطی جدید غیریکنوا برای حل دستگاه معادلات غیرخطی. مجله تحقیق در عملیات در کاربردهای آن، ۱۲ (۴)، ۱-۱۴.

[۸] مدرس خیابانی، ف، دانشیان، ب، (۱۳۹۷). روش به روز رسانی متقارن از مرتبه اول برای حل مسایل بهینه سازی مقیاس بزرگ. مجله تحقیق در عملیات در کاربردهای آن، ۱۵ (۴)، ۱۵۱-۱۷۰.

[2] A. R. Conn, N. I. Gould and P. L. Toint, (2000). Trust region methods, SIAM.

[3] T. F. Coleman and Y. Li, (1996). An interior trust region approach for nonlinear minimization subject to bounds, SIAM Journal on optimization, 6, 418-445.

[4] A. R. Conn, N. I. Gould and P. L. Toint, (1988). Testing a class of methods for solving minimization problems with simple bounds on the variables, Mathematics of Computation, 50, 399-430.

[5] N. I. Gould, D. Orban, A. Sartenaer and P. L. Toint, (2005). Sensitivity of trust-region algorithms to their parameters, 4OR: A Quarterly Journal of Operations Research, 3, 227-241.

[6] S. J. Wright and J. Nocedal (1999). Numerical optimization, Springer Science.

[7] N. I. Gould, D. Orban and P. L. Toint, (2015). CUTEst: a constrained and unconstrained testing environment with safe threads for mathematical optimization, Computational Optimization and Applications, 60, 545-557.

[9] E. D. Dolan and J. J. More, (2002). Benchmarking optimization software with performance profiles, Mathematical programming, 91, 201-213.

[DOR: 20.1001.1.22517286.2019.16.2.3.4]