۱۴ الی ۱۷ شهریور ۱۴۰۱ / دانشگاه علم و فناوری مازندران
53nd Annual Iranian Mathematics Conference
University of Science & Technology of Mazandaran
5-8 september 2022
پنجاه و سومین کنفرانس ریاضی ایران

حل مسایل بهینه‌سازی ناهموار، نامقید و محدب با روش گرادیان مزدوج اصلاح شده HS و جستجوی خطی غیریکنوا

احمد ابویی[۱,*] ، رضا قنبری[۲]، خاطره قربانی مقدم [۳]

[۱]فارغ التحصیل دکتری ریاضی، دانشگاه فردوسی مشهد

[۲]عضو هیئت علمی دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

[۳]عضو هیئت علمی موسسه تحقیقات ریاضی دکتر غلامحسین مصاحب، دانشگاه خوارزمی

چکیده. امروزه، بهینه‌سازی ناهموار یکی از ابزارهای مهم برای حل مسایلی است که در رشته‌های گوناگون با آن مواجه می شویم. در حالت کلی، بهینه‌سازی ناهموار با دسته‌ای از مسایل بهینه‌سازی سروکار دارد که در آن تابع هدف و یا قیدهای مساله به طور پیوسته مشتق پذیر نیستند. اگرچه ممکن است فضای نقاط مشتق پذیر در این مسایل چگال باشد اما مقدار بهینه در این مسایل معمولا در نقطه نامشتق‌پذیر اتفاق می افتد. این چالش موجب شده است که روش‌های حل این دسته از مسایل در دهه‌های مختلف مورد بررسی قرار بگیرند و ما در این مقاله با در نظر گرفتن فرض محدب بودن مساله و با استفاده از آزادسازی مورئو-یوسیدا مساله بهینه‌سازی ناهموار و مقید را به یک مساله بهینه‌سازی هموار و مقید که دارای جواب بهینه یکسان با مساله اصلی تبدیل خواهیم کرد و پس از آن از روش گرادیان مزدوج که از جمله روش‌های کارا به لحاظ، پیچیدگی محاسباتی، زمان و مطلوب بودن جواب بهینه است، برای حل مساله معادل استفاده می‌کنیم.

واژه‌های کلیدی: بهینه سازی ناهموار، هموارسازی، زیرگرادیان، همگرایی سراسری .

---

۱. مقدمه

مسایل بهینه‌سازی ناهموار به رده‌ای کلی از مسایل کمینه‌سازی (بیشینه‌سازی) اشاره دارند که معمولا این توابع در نقطه کمینه کننده (بیشینه کننده) مشتق‌پذیر نیستند. در حالت کلی، بهینه‌سازی ناهموار با نوعی از مسایل بهینه‌سازی سروکار دارد که در آن تابع هدف و یا قیدهای مساله به طور پیوسته مشتق‌پذیر نیستند.
بهینه‌سازی ناهموار امروزه یکی از ابزار مهم برای حل مسایل در رشته‌های گوناگون است و در زمینه‌های زیادی از جمله حذف نویز در تصویر، داده‌کاوی و یادگیری ماشین، کنترل بهینه، مهندسی و مسایل چندهدفه کاربردهای فراوانی دارد.
با توجه به اینکه فرضیات نظریه بهینه‌سازی در حالت هموار بر مبنای مشتق‌پذیر بودن و منظم بودن نقاط است [۱]، از این رو نمی‌تواند به طور مستقیم برای مسایل ناهموار مورد استفاده قرار بگیرد. در واقع بر همین اساس می‌توان گفت، تاریخچه بهینه‌سازی ناهموار به اوایل دهه ۱۹۶۰ مربوط می‌شود هنگامی که مفهوم زیرمشتق[۱] و زیرگرادیان توسط راکوفلر[۲] و فنچل[۳] معرفی شد.
مشابه مسایل بهینه‌سازی هموار، در حالت ناهموار و نامقید نیز مسایل به دو دسته محدب و نامحدب تقسیم می‌شوند. مشکل اصلی در مسایل نامحدب، عدم وجود زیرمشتق برای توابع هدف نامحدب بود. به همین دلیل دهه ۱۹۷۰ و اوایل دهه ۱۹۸۰ را می‌توان دوره مهمی برای پیشرفت نوین در تحلیل توابع ناهموار دانست. در این دوره زیرمشتق‌های گوناگونی از جمله زیرمشتق کلارک[۴] [۲]و شبه‌مشتق دیمیانوف-ربینوف[۵] [۳] معرفی شدند.

---

[*]سخنران. آدرس ایمیل: ahmadaboiy@gmail.com

[۱]Subdifferential
[۲]R.T. Rockafellar
[۳]W. Fenchel
[۴]Clarke
[۵]Demyanov-Rubinov

این مفاهیم به نوعی تعمیم زیرمشتق و زیرگرادیان توابع در حالت محدب به شمار می‌آمدند، با این تفاوت که اکنون برای توابع نامحدب خوش تعریف بودند.
با توجه به فرآیند توسعه روش‌های حل مسایل بهینه‌سازی ناهموار، شاید بتوان از یک دیدگاه این روش‌ها را به دو قسمت تقسیم نمود: دسته اول روش‌هایی هستند که با توجه به تعاریف موجود در مسایل بهینه‌سازی ناهموار سعی در یافتن جواب بهینه دارند (زیرگرادیان، نیوتن بسته‌ای، نیوتن مانند روش‌های) اما دسته دوم روش‌هایی هستند که با در نظر گرفتن فرضیات ویژه‌ای روی مساله، آن را به یک مساله بهینه‌سازی هموار تبدیل می‌کند و یا اینکه با فرض اینکه مساله ناهموار روی یک زیر مجموعه‌ای چگالی از فضای جواب رفتار هموار دارد مساله را در نظر می‌گیرد. در چنین حالتی می‌توان تمام الگوریتم‌های معرفی شده برای مسایل هموار از جمله روش‌های گرادیان مزدوج و ناحیه اعتماد ناحیه اعتماد و غیره را برای حل مسایل ناهموار به کار برد. روش‌های مستقیم از زیرگرادیان و زیر مشتق استفاده می‌کنند که در بالا به آن اشاره شد و روش‌های غیرمستقیم هم از الگویتم‌های هموار استفاده می‌کنند که در ادامه به آن اشاره خواهیم کرد.
از میان روش‌های تکراری گوناگون برای حل مسایل بهینه‌سازی هموار و نامقید از جمله: روش تندترین شیب، گرادیان مزدوج، نیوتن، شبه‌نیوتن و $\cdots$، روش گرادیان مزدوج به دلیل سادگی در به روزرسانی جهت تولید شده در هر تکرار و استفاده بهینه از حافظه در حال اجرا (رم) یکی از بهترین الگوریتم‌های حل مسایل هموار و نامقید است [۱]. اگرچه روش تندترین شیب نیز همانند روش گرادیان مزدوج ساده بوده و به طور بهینه از رم استفاده می‌کند. اما به دلیل عدم بهبود قابل قبول در هر تکرار (به خصوص تکرارهای بالا) عملکرد عددی قابل قبولی ندارد. از سوی دیگر روش‌های نیوتن و شبه نیوتن از نظر بهبود مقدار تابع هدف در هر تکرار دارای عملکرد بسیار خوبی هستند اما به دلیل محاسبات ماتریسی (تقریب ماتریس هسین)، برای مسایل بزرگ ابعاد عملکرد عددی قابل قبولی نخواهند داشت. از این روش‌ها معمولا برای مسایل از بعد متوسط استفاده می‌شود. این موارد موجب شده است که روش های گرادیان مزدوج از یک سو به دلیل شباهت زیاد به روش تندترین شیب در سادگی و محاسبات برداری و از سوی دیگر به دلیل شباهت با روش‌های شبه نیوتن در میزان بهبود تابع هدف در هر تکرار بسیار کارا باشند. روش‌های گرادیان مزدوج از جمله روش‌های کارا برای حل مسایل بهینه‌سازی هموار هستند. به خصوص برای مسایل بزرگ ابعاد عملکرد قابل قبولی دارند. از جمله ویژگی‌های روش مذکور، کاهشی بودن جهت محاسبه شده در هر گام است. همچنین ثابت شده است که روش ارایه شده همگرای سراسری دارد.
با توجه به توضیحات داده شده در این مقاله قصد داریم با استفاده از روش آزادسازی مورئو–یوسیدا مساله ناهموار و نامقید زیر را،

$$\min_{x\in\mathbb{R}^n} \quad f(x). \tag{۱}$$

را به یک مساله بهینه‌سازی هموار تبدیل کنیم (با فرض محدب بودن تابع) و با توجه به کارا بودن روش گرادیان مزدوج از این روش برای حل مساله معادل استفاده می‌شود.

۲. روش گرادیان مزدوج اصلاح شده HS برای حل مساله (۱)

در این بخش با اعمال اصلاحاتی روی β_k^{HS}، روش گرادیان مزدوج اصلاح شده HS را برای حل مساله (۱) ارایه می‌دهیم. اکنون اولین پرسشی که به ذهن می‌رسد دلیل انتخاب روش گرادیان مزدوج HS و اعمال اصلاحات روی آن است. در بخش ۱.۲ ساختار اصلی روش پیشنهاد شده را بیان می‌کنیم.

۱.۲. ساختار اصلی الگوریتم. روش گرادیان مزدوج HS اگرچه از میان روش‌های ارایه شده برای به‌روزرسانی مقدار β_k در هر تکرار، دارای عملکرد عددی بسیار خوبی است اما دارای همگرایی سراسری نیست. با اعمال اصلاحاتی بر روی نحوه به‌روزرسانی β_k در روش HS همگرایی سراسری آن تضمین خواهد شد. اما پیش از اعمال اصلاحات بر روی روش گرادیان مزدوج HS، به ایده اصلی که منجر به اصلاح این روش شده است، اشاره می‌کنیم.

۱.۱.۲. ایده اصلی برای اصلاح روش گرادیان مزدوج HS. ژنگ و همکارانش[6] [۴]، برای تولید جهت کاهشی در روش گرادیان مزدوج FR، روش اصلاح شده‌ای را ارایه دادند که در آن، جهت d_k در هر تکرار به صورت زیر محاسبه می‌شود:

$$d_{k+1} = \begin{cases} -g_{k+1} & if \;\; k=-1, \\ -\theta_k^{(2)} g_{k+1} + \beta_k^{FR} d_k & if \;\; k \geq 1, \end{cases} \tag{۲}$$

که در آن $\theta_k^{(2)} = \frac{d_{k-1}^\intercal y_{k-1}}{||g_{k-1}||^2}$ و $\beta_k^{FR} = \frac{||g_{k+1}||^2}{||g_k||^2}$.
به راحتی می‌توان نشان داد که $g_k^\intercal d_k = -||g_k||$ ، d_k در هر تکرار یک جهت کاهشی است .

۲.۱.۲. محاسبه جهت جستجو. ایده ژنگ و همکارانش [۴] ما را بر آن داشت تا با اصلاح مشابه بر روی روش گرادیان مزدوج HS، در هر تکرار از الگوریتم یک جهت کاهشی تولید نماییم و در نتیجه مقدمات همگرایی سراسری روش اصلاح شده را فراهم کنیم. محاسبه گرادیان تابع هدف مساله (۵) به صورت دقیق مقرون به صرفه نخواهد بود از این رو از مقدار تقریبی g_k یعنی $g^a(x_k,\varepsilon_k)$ در هر تکرار استفاده خواهیم نمود. بر این اساس، جهت جستجو در هر تکرار به صورت زیر محاسبه می‌شود:

$$d_{k+1} = \begin{cases} -g^a(x_{k+1},\varepsilon_{k+1}) & if \;\; k=-1, \\ -\theta_k g^a(x_{k+1},\varepsilon_{k+1}) + \beta_k^{HS} d_k & if \;\; k \geq 0, \end{cases} \tag{۳}$$

که در آن

$$\theta_k = 1 + \frac{g^a(x_{k+1},\varepsilon_{k+1})^\intercal y_k \times g^a(x_{k+1},\varepsilon_{k+1})^\intercal d_k}{||g^a(x_{k+1},\varepsilon_{k+1})||^2 (d_k^\intercal y_k)}, \tag{۴}$$

و

$$\beta_k^{HS} = \frac{g^a(x_{k+1},\varepsilon_{k+1})^\intercal y_k}{d_k^\intercal y_k}, \qquad y_k = g^a(x_{k+1},\varepsilon_{k+1}) - g^a(x_k,\varepsilon_k).$$

لم بعد کاهشی بودن جهت تولید شده در هر گام را نشان می‌دهد.

لم ۲٫۱ (جهت کاهشی). برای هر $k \in \mathbb{Z}^{\geq -1}$، d_{k+1} یک جهت کاهشی است.

۳.۱.۲. محاسبه طول گام α_k در تکرار kام. اکنون پس از تعیین جهت جستجو، نیاز است طول گام حرکت در جهت d_k در تکرار kام را محاسبه کنیم. حل این مساله به دلیل پیچیدگی‌های موجود در تابع هدف ممکن است مقرون به صرفه نباشد. از این رو روش‌های گوناگونی برای تقریب مقدار α در گام kام ارایه شده است. جستجوی خطی یکنوا آرمیژو یکی از ساده‌ترین و پرکاربردترین روش‌های تقریب مقدار α_k است و روی تقریب تابع هدف مساله (۵)به صورت زیر تعریف می‌شود [۱]. مساله

$$\min_{x \in \mathbb{R}^n} F(x). \tag{۵}$$

معادل است[7]. بر اساس روش پیشنهاد شده توسط امینی و همکاران [۵] ، برای تقریب تابع هدف مساله (۵)، α_k پذیرفتنی است، اگر

$$F^a(x_k + \alpha_k d_k, \varepsilon_{k+1}) \leq R_k + \delta \alpha_k g^a(x_k,\varepsilon_k)^\intercal d_k, \tag{۶}$$

که در آن

$$R_k = \eta_k F_{l(k)} + (1-\eta_k) F(x_k,\varepsilon_k),$$

و $0 < \delta < 1$ ، $0 \leq \eta_{\min} \leq \eta_{\max} \leq 1$ ، $\eta_k \in [\eta_{\min}, \eta_{\max}]$ و $F_{l(k)}$ به صورت زیر تعریف می‌شود:

$$F_{l(k)} \max_{0 \leq j \leq m_k} \{F^a(x_{k-j},\varepsilon_{k-j})\}, \qquad k = 0,1,2,\ldots,$$

و در آن $m_0 = 0$ ، $0 \leq m_k \leq \min\{m_{k-1}+1, M\}$ و $M \geq 0$.

[6] Zhang et al.

[7] معادل بودن به این معنی است که هر دو مساله دارای یک جواب بهینه و مقدار تابع هدف یکسان هستند.

ملاحظه ۲٫۲. امینی و همکاران [۵] با اعمال شرایطی بر روی الگوریتم از جمله کاهشی بودن جهت d و از پایین کران‌دار بودن تابع هدف خوش تعریفی α_k را اثبات کرده‌اند.

۴.۱.۲. الگوریتم گرادیان مزدوج اصلاح شدهHS با جستجوی خطی غیریکنوا (A(MHS)). اکنون که نحوه‌ی تعیین جهت جستجو و طول گام مشخص شده‌اند. می‌توان الگوریتم را به صورت زیر در نظر گرفت:

الگوریتم ۱ گرادیان مزدوج اصلاح شده HS با جستجوی خطی غیریکنوا (A(MHS))

گام ۰. تعیین کن: $x_0 \in \mathbb{R}^n$، $0 < \rho < 1$، $0 < \delta < \frac{1}{2}$، $s > 0$، $0 \leq \eta_{\min} \leq \eta_0 \leq \eta_{\max}$، $M \geq 0$، $\lambda > 0$ و $0 < \epsilon < 1$. همچنین قرار بده: $m_0 = 0$، $R_0 = F^\alpha(x_0, \varepsilon_0)$، $d_0 = -g^\alpha(x_0, \varepsilon_0)$ و $k = 0$.

گام ۱. اگر $||g^\alpha(x_k, \varepsilon_k)|| < \varepsilon$ یا $|F^a(x_k, \varepsilon_k) - F^a(x_{k-1}, \varepsilon_{k-1})| < \varepsilon$، در این صورت متوقف شو.

گام ۲. عدد حقیقی ε_{k+1} را به گونه‌ای محاسبه کن که $0 < \varepsilon_{k+1} < \varepsilon_k$ و $\sum_{k=0}^{\infty} \varepsilon_k < +\infty$. همچنین مقدار طول گام α_k را از طریق (۶) و بر اساس عملیات عقبگرد $(\alpha_k = \rho^i s, \quad i \in \{0, 1, 2, \ldots\})$ محاسبه کن.

گام ۳. قرار بده: $x_{k+1} = x_k + \alpha_k d_k$.

گام ۴. جهت جستجوی d_{k+1} را بر اساس (۳) محاسبه کن.

گام ۵. قرار بده: $k := k + 1$ و به گام ۱ بازگرد.

همگرایی سراسری الگوریتم را با در نظرگرفتن فرضیات معینی اثبات کرده‌ایم.

۳. نتیجه‌گیری

در این مقاله با استفاده از تکنیک منظم‌سازی مورئو–یوسیدا مساله (۱) را به مساله (۵) که یک مساله هموار و نامقید است تبدیل کردیم. شرط اولیه برای تبدیل، محدب بودن تابع هدف مساله (۱) است. در ادامه از روش گرادیان مزدوج برای حل مساله (۵) استفاده کردیم. از میان شاخه‌های گوناگون روش گرادیان مزدوج، روش HS به دلیل عملکرد عددی مناسب انتخاب شد. به دلیل عدم همگرایی سراسری در این روش، اصلاحاتی روی آن صورت گرفت. با ایده گرفتن از روش ژنگ و همکاران [۴] یک جهت جستجوی کاهشی در هر تکرار توسط روش گرادیان مزدوج اصلاح شده HS تولید شد. روش پیشنهاد شده روی توابع محدب قوی همگرای سراسری است.

مراجع

1. Sun, W. and Yuan, Y. X.: Optimization Theory and Methods: Nonlinear Programming. Springer, New York (2006)
2. Bagirov, A., Karmitsa, N. and Makela, M. M. , *Introduction to Nonsmooth Optimization: Theory, Practice and Software*, Springer International Publishing (2014).
3. Hager, W., and Zhang, H., A survey of nonlinear conjugate gradient method, SIAM Journal on Optimization (2006) 1-21.
4. Zhang, L., Zhou, W. and Li, D.: Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search, Numerische Mathematik, 104, 561-572 (2006).
5. Amini, K., Ahookhosh, M. and Nosratipour, H.: An inexact line search approach using modified nonmonotone strategy for unconstrained optimization, Numerical Algorithms, 66, 49-78 (2014).