لازم بودن L- ایستایی برای بهینگی در بهینه‌سازی غیرخطی با قید تُنُکی

عباس خادمی، مجید سلیمانی دامنه*
دانشگاه تهران، پردیس علوم، دانشکدهٔ ریاضی، آمار و علوم کامپیوتر

دریافت ۹۷/۰۲/۲۱ پذیرش ۹۷/۰۶/۲۶

چکیده

در این مقاله، یک شرط لازم بهینگی برای مسئله‌ای خاص در بهینه‌سازی غیرخطی، تحت عنوان مسئله با قید تُنُکی، را بررسی می‌کنیم. این مسئله به کمینه کردن تابعی به‌طور پیوسته مشتق‌پذیر تحت یک محدودیت تنکی روی متغییر می‌پردازد. نشان می‌دهیم که، در حالت کلی، L-ایستایی یک شرط لازم بهینگی برای مسئله با قید تنکی است. این خاصیت در ادبیات موضوع تحت فرض لیپ شیتز بودن عملگر گرادیان اثبات شده است.

واژه‌های کلیدی: بهینه‌سازی غیرخطی، مسائل با قید تُنُکی، بهینگی، L-ایستایی.

مقدمه

مسئله بهینه‌سازی زیر را در نظر می‌گیریم:

$$(\mathrm{P}): \quad \begin{cases} \min \mathrm{f}(x) \\ \mathrm{s.t.}\ \|x\|_0 \le \mathrm{s} \end{cases}$$

که در آن $\mathrm{f}: \mathbb{R}^n \to \mathbb{R}$ به‌طور پیوسته مشتق‌پذیر است و s>0 یک عدد صحیح کوچک‌تر از n است. به‌علاوه، $\|x\|_0$ نرم صفر است که بدین‌صورت تعریف می‌شود:

$$\|x\|_0 \equiv \#\{i :\ x_i \neq 0\}.$$

نرم صفر بردار x در حقیقت بیانگر تعداد مؤلفه‌های غیرصفر بردار x است. این مسئله کاربردهای مهمی در پردازش سیگنال و سنجش فشرده دارد [۳]، [۴].

ذکر این نکته ضروری است که فرض محدب بودن روی تابع هدف را نداریم. هم‌چنین تابع محدودیت نامحدب است و لزوماً پیوسته نیست، که این مسئله را بسیار دشوار می‌سازد. هدف ما در اینجا بررسی یک شرط لازم بهینگی برای مسئله (P) است.

شرایط لازم بهینگی برای آن دسته از مسائل کمینه‌سازی که تابع هدف آنها به‌طور پیوسته مشتق‌پذیر (ممکن است نامحدب باشد) و ناحیه شدنی آنها محدب است، در ادبیات موضوع به کرّات بررسی شده است و روش‌های متنوعی برای حل این‌گونه از مسائل وجود دارد. اما نتایج موجود در بهینه‌سازی کلاسیک، مسئله (P) را پوشش نمی‌دهند، زیرا

*نویسنده مسئول soleimani@khayam.ut.ac.ir

[DOR: 20.1001.1.25882546.1399.6.3.13.6]
[DOI: 10.52547/mmr.6.3.387]

در این مسئله تابع تعریف کننده محدودیتِ تنکی نه پیوسته است و نه محدب. در ادامه، به بررسی نمونه‌هایی از شرایط لازم بهینگی برای مسئله (P) می‌پردازیم، که عبارتند از شدنی بودن پایه‌ای[۱] و L-ایستایی[۲].

برای ارائه تعاریف دو مفهوم فوق، نیازمند معرفی نمادهایی بدین‌شرح هستیم:

برای بردار داده شده $x \in \mathbb{R}^n$ تکیه‌گاه[۳] به‌صورت

$$\mathrm{I}_1(\mathrm{x}) := \{\mathrm{i} : \mathrm{x_i} \neq 0\}$$

و متمم آن به‌صورت

$$\mathrm{I}_0(\mathrm{x}) := \{\mathrm{i} : \mathrm{x_i} = 0\}$$

تعریف می‌شوند. مجموعهٔ C_s ، متشکل از بردارهایی که حداکثر s-تنک است، بدین‌صورت تعریف می‌شود:

$$\mathrm{C_s} := \{\mathrm{x} : \|\mathrm{x}\|_0 \leq \mathrm{s}\}.$$

عملگر تصویر متعامد[۴] روی یک مجموعه بسته و ناتهی $\mathrm{D} \subseteq \mathbb{R}^n$ را با نماد $P_D(\,.\,)$ نشان داده و بدین‌صورت تعریف می‌کنیم:

$$\mathrm{P_D}(\mathrm{y}) := \underset{\mathrm{x} \in \mathrm{D}}{\operatorname{argmin}} \|\mathrm{y} - \mathrm{x}\|^2$$

که در آن نماد argmin نشان‌دهندهٔ مجموعه کمینه کننده‌ها است.

شرط لازم بهینگی

شرایط بهینگی نقش مهمی در بررسی نظری، عددی و کاربردی مسائل بهینه‌سازی دارند. از دیدگاه عملی، شرایط بهینگی پایه بسیاری از روش‌های حل عددی هستند. بدین‌ترتیب قدم اول در مطالعه مسئلهٔ (P)، بررسی شرایط بهینگی است. برای مسائل نامقید مشتق‌پذیر، شرط لازم بهینگی این است که گرادیان تابع هدف برابر صفر شود. بنابراین طبیعی است که انتظار داشته باشیم یک شرط لازم مشابه روی مجموعه محمل $I_1(x^*)$ برای نقطه بهینه x^* برای مسئله (P) برقرار باشد. با الهام‌گیری از اصطلاحات برنامه‌ریزی‌خطی، یک بردار که این ویژگی را برآورده سازد، بردار پایه‌ای شدنی نامیده می‌شود.

تعریف ۱. [۱]، بردار $\mathrm{x}^* \in \mathrm{C_s}$ را برای مسئله (P) پایه‌ای شدنی گوییم هرگاه

$$1)\ \|\mathrm{x}^*\|_0 < s \implies \nabla \mathrm{f}(\mathrm{x}^*) = 0,$$

$$2)\ \|\mathrm{x}^*\|_0 = \mathrm{s} \implies [\nabla \mathrm{f}(\mathrm{x}^*)]_\mathrm{i} = 0, \quad \forall\, \mathrm{i} \in \mathrm{I}_1(\mathrm{x}^*).$$

فرض ۱. در سراسر این مقاله فرض می‌کنیم تابع هدف مسئله (P) روی $\mathbb{R}^n$ از پایین کراندار است.

قضیهٔ ۱ نشان می‌دهد پایه‌ای شدنی بودن یک شرط لازم بهینگی برای مسئله (P) است.

1. Basic feasibility
2. L-stationarity
3. Support
4. Orthogonal projection

[DOR: 20.1001.1.25882546.1399.6.3.13.6]
[DOI: 10.52547/mmr.6.3.387]

قضیۀ ۱. [۱]، فرض کنید x^* یک جواب بهینه برای مسئله (P) باشد. آن‌گاه x^* یک بردار پایه‌ای شدنی برای آن مسئله است.

در ادامه، مفهوم L-ایستایی را بررسی می‌کنیم که توسیعی از مفهوم ایستایی برای مسائل با محدودیت محدب است.

تعریف ۲. [۱]، بردار $x^* \in C_s$ را یک نقطه L-ایستا برای . مسئله (P) می‌نامیم هرگاه

$$x^* \in P_{C_s}(x^* - \frac{1}{L}\nabla f(x^*)).$$

توجه کنید که چون C_s مجموعه‌ای نامحدب است، عملگر تصویر متعامد $P_{C_s}(.)$ لزوما تک مقداری نیست.

فرض۲. عملگر گرادیان تابع هدف f روی $\mathbb{R}^n$ لیپ شیتز با ثابت L_f است؛ یعنی

$$\forall x, y \in \mathbb{R}^n : \quad \|\nabla f(x) - \nabla f(y)\| \leq L_f\|x - y\|.$$

در مرجع [۱] نشان داده شده است که تحت فرض ۲ هر جواب بهینه برای مسئله (P)، یک بردار L-ایستا، با $L > L_f$ است.

قضیۀ ۲. [۱]، فرض کنید f تابعی به‌طور پیوسته مشتق‌پذیر است و در فرض ۲ صدق می‌کند. به‌علاوه، فرض کنید $L > L_f$ و x^* یک جواب بهینه برای مسئله (P) است آن‌گاه:

الف) x^* یک نقطه L-ایستا است.

ب) مجموعه $P_{C_s}(x^* - \frac{1}{L}\nabla f(x^*))$ تک عضوی است.

اگر چه در مرجع [۱] برای اثبات لازم بودن L-ایستایی از فرض ۲ (لیپ شیتز بودن گرادیان) استفاده شده است، در قضیۀ زیر نشان می‌دهیم L-ایستایی بدون برقراری فرض ۲ نیز یک شرط لازم برای بهینگی برای مسئله (P) است. قضیۀ زیر نتیجه اصلی این مقاله است.

قضیۀ ۳. فرض کنید f تابعی به‌طور پیوسته مشتق‌پذیر باشد. اگر x^* یک جواب بهینه برای مسئله (P) باشد، آن‌گاه L>0ای وجود دارد به‌طوری‌که x^* نقطه‌ای L-ایستا برای این مسئله است.

برهان. فرض کنید x^* یک جواب بهینه برای مسئله (P) باشد. بنا به قضیۀ ۱، x^* یک بردار پایه‌ای شدنی برای مسئله (P) است. اگر $\|x^*\|_0 < s$ داریم $\nabla f(x^*) = 0$ و لذا $x^* \in P_{C_s}(x^* - \frac{1}{L}\nabla f(x^*))$ برای هر L>0. پس فرض کنید $\|x^*\|_0 = s$. داریم

$$\forall x' \in C_s: \ \left[x' - x^* + \frac{1}{L}\nabla f(x^*)\right]_i = \begin{cases} x'_i - x^*_i\,, & i \in I_1(x^*) \\ x'_i + 1/L[\nabla f(x^*)]_i\,, & i \in I_0(x^*) \end{cases}$$

حال بردار $d = d_{x'}$ را به‌صورت

$$d_i := \begin{cases} x'_i - x^*_i\,, & i \in I_1(x^*) \\ 0\,, & i \in I_0(x^*) \end{cases}$$

تعریف می‌کنیم. دو حالت داریم

[DOR: 20.1001.1.25882546.1399.6.3.13.6]
[DOI: 10.52547/mmr.6.3.387]

حالت اول. برای هر $i \in I_1(x^*)$، $x_i' \neq 0$ با توجه به $x' \in C_s$ و $\|x^*\|_0 = s$ داریم $x_i' = 0$ برای هر $i \in I_o(x^*)$. پس برای هر L>0 داریم:

$$\left\|x' - x^* + \frac{1}{L}\nabla f(x^*)\right\|^2 = \sum_{i\in I_1(x^*)} (x_i' - x_i^*)^2 + \sum_{i\in I_o(x^*)} (\frac{1}{L}[\nabla f(x^*)]_i)^2$$
$$= \|x' - x^*\|^2 + \frac{1}{L^2}\|\nabla f(x^*)\|^2$$
$$\geq \frac{1}{L^2}\|\nabla f(x^*)\|^2$$
$$= \left\|x^* - x^* + \frac{1}{L}\nabla f(x^*)\right\|^2.$$

حالت دوم. وجود داشته باشد $k \in I_1(x^*)$ به‌طوری‌که $x_k' = 0$. در این صورت $d_k = -\, x_k^*$ و لذا

$$\|d\| \geq |d_k| = |x_k^*|.$$

قرار می‌دهیم:

$$\theta := \min\{|x_k^*|:\ k \in I_1(x^*),\ x_k' = 0\ \}.$$

در نتیجه $\|d\| \geq \theta > 0$. حال L^* را طوری در نظر می‌گیریم که $\left\|\frac{1}{L^*}\nabla f(x^*)\right\| \leq \theta$. داریم

$$\left\|x' - x^* + \frac{1}{L^*}\nabla f(x^*)\right\| = \sqrt{\sum_{i\in I_1(x^*)}(x_i' - x_i^*)^2 + \sum_{i\in I_o(x^*)}\left(x_i' + \frac{1}{L^*}[\nabla f(x^*)]_i\right)^2}$$

$$= \sqrt{(x_k^*)^2 + \sum_{i\in I_1(x^*)\setminus\{k\}}(x_i' - x_i^*)^2 + \sum_{i\in I_o(x^*)}\left(x_i' + \frac{1}{L^*}[\nabla f(x^*)]_i\right)^2}$$
$$\geq |x_k^*|.$$

به‌علاوه،

$$\left\|x^* - x^* + \frac{1}{L^*}\nabla f(x^*)\right\| = \left\|\frac{1}{L^*}\nabla f(x^*)\right\| \leq \theta \leq |x_k^*|.$$

بنابراین

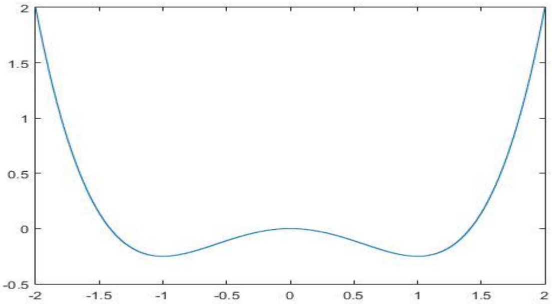
$$\forall x' \in C_s:\ \ \left\|x^* - x^* + \frac{1}{L^*}\nabla f(x^*)\right\| \leq \left\|x' - x^* + \frac{1}{L^*}\nabla f(x^*)\right\|,$$

و در نتیجه $x^* \in P_{C_s}(x^* - \frac{1}{L^*}\nabla f(x^*))$. پس x^*، نقطه‌ای L^*-ایستاست. □

هر چند L-ایستایی یک شرط لازم بهینگی برای مسئله (P) است، اما شرط کافی نیست. به این معنی که هر نقطه L-ایستا لزوماً بهینه نیست و در حقیقت می‌تواند از جواب بهینه سراسری دور باشد. دو مثال ۱ و ۲، این موضوع را روشن می‌سازند.

[DOR: 20.1001.1.25882546.1399.6.3.13.6]
[DOI: 10.52547/mmr.6.3.387]

مثال ۱. فرض کنید $f(x)=\frac{x^4}{4}-\frac{x^2}{2}$ و s=1. در این صورت $x^*=0$ نقطه‌ای L-ایستا است، زیرا $\nabla f(x^*)=0$ و در نتیجه $x^*\in P_{C_s}(x^*-\frac{1}{L}\nabla f(x^*))$، در حالی که $x^*=0$ بهینه نیست (شکل زیر را ببینید). دقت کنید که $C_1=\mathbb{R}$.



شکل ۱. نمودار تابع $\boldsymbol{f(x)=\frac{x^4}{4}-\frac{x^2}{2}}$

در مثال۲ تابع هدف را محدب در نظر می‌گیریم ولی باز هم L -ایستایی شرط کافی برای بهینگی نیست.

مثال ۲. فرض کنید $f(x,y)=(x-10)^2+(y-1)^2$ و s=1. در این صورت $x^*=(0,1)$ نقطه‌ای L-ایستا است زیرا

$$x^*-\frac{1}{L}\nabla f(x^*)=\begin{pmatrix}0\\1\end{pmatrix}-\frac{1}{L}\begin{pmatrix}-20\\0\end{pmatrix}=\begin{pmatrix}\frac{20}{L}\\1\end{pmatrix}.$$

برای L به قدر کافی بزرگ $x^*=P_{C_s}\left(x^*-\frac{1}{L}\nabla f(x^*)\right)$، از طرفی $f(0,1)=100>1=f(10,0)$ بنابراین، L>0ای وجود دارد به‌طوری‌که $x^*=(0,1)$ یک نقطه L-ایستاست ولی این نقطه، بهینه نیست.

در این مقاله، یک شرط لازم بهینگی موجود در ادبیات موضوع را تحت یک فرض ضعیف‌تر اثبات کردیم. به‌علاوه، نشان دادیم که این شرط کافی نیست. تعدادی شرایط کافی بهینگی برای مسائل با قید تنکی در مراجع [۲]، [۵]،[۶] دیده می‌شود.

منابع

1. Beck A., Eldar Y.C., "Sparsity Constrained Nonlinear Optimization: Optimality Conditions and Algorithms", SIAM J. Optim., 23 (3) (2013) 1480-1509.
2. Bucher M., Schwartz A., "Second-Order Optimality Conditions and Improved Convergence Results for Regularization Methods for Cardinality-Constrained Optimization Problems", J. Optim. Theory Appl., 178 (2018) 383-410.
3. Donoho D. L., "Compressed sensing", IEEE Transactions on Information Theory, 52 (4), (2006) 1289-1306.
4. Foucart S., Rauhut H., "A Mathematical Introduction to Compressive Sensing", Springer, NY (2013).

[DOR: 20.1001.1.25882546.1399.6.3.13.6]
[DOI: 10.52547/mmr.6.3.387]

5. Luo Lili Pan Z., Xiu N., "Restricted Robinson Constraint Qualification and Optimality for Cardinality-Constrained Cone Programming", J. Optim. Theory Appl., 175, 104-118 (2017).

6. Movahedian N., Nobakhtian S., Sarabadan M., "Nonsmooth Sparsity Constrained Optimization Problems: Optimality Conditions", Optimization Letters, In press; DOI: 10.1007/s11590-018-1310-6.