فرهنگ و اندیشهٔ ریاضی
سال ۳۸، شمارهٔ ۶۴ (بهار و تابستان ۱۳۹۸) صص. ۴۱ تا ۵۸

الگوریتم نقطهٔ پروکسیمال چیست؟

هادی خطیب‌زاده

چکیده

در حوزهٔ بهینه‌سازیِ محدب، الگوریتم‌های متعددی برای تقریب نقاط بهینهٔ یک تابع محدب وجود دارد که یکی از آنها الگوریتم نقطهٔ پروکسیمال است. چون این الگوریتم دارای بنیان نظری ژرف و زیبا و قابلیت تعمیم به فضاهای مجرد با کاربردهای متعدد به‌ویژه در بهینه‌سازی غیرهموار، مقید و بزرگ–مقیاس است، به‌طور گسترده‌ای مطالعه شده است. در این مقاله، هدف ما این است که خواننده را با مفاهیم اساسی که زیربنای این الگوریتم را تشکیل می‌دهند، آشنا کنیم.

۱. سرآغاز

کمینه‌سازیِ محدب، زیرشاخه‌ای از بهینه‌سازی است که هدف آن به‌دست آوردنِ کمینهٔ یک تابع محدب تعریف‌شده بر یک دامنهٔ محدب در فضای زمینه است. مسئلهٔ اصلی در کمینه‌سازیِ محدب، مسئلهٔ

$$\mathrm{Min}_{x\in C} f(x) \tag{۱.۱}$$

است که در آن، $f : C \to \mathbb{R}$ تابعی محدب و C مجموعه‌ای محدب متشکل از قیدهای مسئله[1] است. یکی از روش‌های کمینه‌سازیِ محدب، روش موسوم به *الگوریتم نقطهٔ پروکسیمال* است که نخستین بار در

عبارات و کلمات کلیدی. الگوریتم نقطهٔ پروکسیمال؛ تابع محدب؛ بهینه‌سازی محدب؛ عملگر یکنوای ماکسیمال؛ عملگر حلّال.

[1] constraint set

سال ۱۹۷۰ توسط مارتینه[1] ارائه شد و بعدها راکفلر[2] آن را در سال ۱۹۷۶ به‌طور سازمان یافته‌تر بررسی کرد و تعمیم داد.

در این مقاله، به معرفی الگوریتم نقطهٔ پروکسیمال در بهینه‌سازی محدب، نظریهٔ عملگرهای یکنوا و نظریهٔ نقطهٔ ثابت می‌پردازیم. پیش از ورود به بحث اصلی، در بخش ۲، پیشنیازهایی از آنالیز محدب را که برای ادامهٔ بحث ضروری هستند، معرفی می‌کنیم. بحث اصلی از بخش ۳ با معرفی نگاشت پروکسیمال و الگوریتم نقطهٔ پروکسیمال آغاز می‌شود. در این بخش، دربارهٔ چگونگیِ به‌کار بردن این الگوریتم در تقریب نقطهٔ کمینهٔ تابع محدب و دلیل همگرایی آن صحبت می‌شود. رهیافت دیگری به الگوریتم نقطهٔ پروکسیمال، شکل گسستهٔ دستگاه‌های گرادیان است که موضوع بخش ۴ است. در بخش‌های ۵ و ۶ تعمیم‌هایی از این الگوریتم برای تقریب نقطهٔ کمینهٔ مجموع دو تابع محدب و نقطهٔ کمینهٔ مشترک دو تابع محدب بیان می‌شود. کاربرد الگوریتم نقطهٔ پروکسیمال در تقریب صفرهای عملگرهای یکنوا، موضوع بخش ۷ است. در بخش ۸، به‌کاربرد این الگوریتم در تقریب نقطهٔ ثابت یک نگاشت انقباضی می‌پردازیم.

در سراسر مقاله، H یک فضای هیلبرت با ضرب داخلی $\langle \cdot,\cdot \rangle$ و نرم $\|\cdot\|$ است. خواننده‌ای که با آنالیز تابعی آشنایی ندارد، می‌تواند همه‌جا فضاهای نه لزوماً متناهی–بُعد هیلبرت را متناهی–بُعد (مثلاً $\mathbb{R}^n$ با ضرب داخلی متداول) فرض کند که در این حالت، همگرایی ضعیف (که در بخش‌های بعدی به آن اشاره خواهد شد) با همگرایی قوی یکسان خواهد بود. گذشته از این، سایر مطالب مقاله برای خواننده‌ای که آشنایی مقدماتی با آنالیز ریاضی داشته باشد، قابل درک است.

۲. مختصری از آنالیز محدب

آنالیز محدب، شاخه‌ای از آنالیز ریاضی است که به مطالعهٔ مجموعه‌ها و تابع‌های محدب می‌پردازد و در کمینه‌سازیِ محدب که شاخه‌ای از نظریهٔ بهینه‌سازی است، کاربرد دارد. ویژگی‌های زیبای تابع‌های محدب به آنها جایگاهی ویژه در بهینه‌سازی بخشیده است. در این بخش، مقدماتی از آنالیز محدب را یادآوری می‌کنیم که برای مطالعهٔ بقیهٔ مقاله، ضروری است. کتاب باوشکه و کامبته [۴] مرجعی مناسب برای دورهٔ مطالب بیان‌شده در این بخش است.

مجموعهٔ $C \subseteq H$ را محدب خوانیم اگر برای هر دو نقطهٔ $x, y \in C$ پاره‌خط واصلِ آنها نیز در C باشد. فرض کنیم $C \subseteq H$ ناتهی، بسته و محدب باشد. در این صورت برای هر $x \in H$ نقطه‌ای یکتا

[1] B. Martinet ریاضیدان فرانسوی که فقط ۱۰ مقاله از او به ثبت رسیده است و همگی به زبان فرانسوی هستند. مقاله‌ای که در آن، این الگوریتم را معرفی کرده پُر ارجاع‌ترین مقالهٔ او است.

[2] R. T. Rockafellar متولد دهم فوریهٔ سال ۱۹۳۵، یکی از پیشگامان نظریهٔ بهینه‌سازی و آنالیز محدب است. او استاد بازنشسته در دانشکده‌های ریاضی و ریاضی کاربردی دانشگاه واشینگتن و سیاتل است. کتاب او در زمینهٔ آنالیز محدب، یکی از کتاب‌های پُر ارجاع است.

در C وجود دارد که آن را با $\mathrm{Proj}_C x$ نشان می‌دهیم به‌طوری که به‌ازای هر $y \in C$،

$$\|x - \mathrm{Proj}_C x\| \leq \|x - y\|.$$

نقطهٔ $\mathrm{Proj}_C x$ را *تصویر* x روی C می‌نامیم.

تعریف ۱.۲. فرض کنیم $C \subseteq H$ محدب باشد. تابع $f : C \to]-\infty, +\infty]$ را *تابع محدب* نامیم اگر برای هر $x, y \in C$ و هر $\lambda \in [\circ, ۱]$ داشته باشیم

$$f(\lambda x + (۱ - \lambda)y) \leq \lambda f(x) + (۱ - \lambda) f(y).$$

تابع f *اکیداً* محدب نام دارد اگر برای هر $x, y \in C$ که $x \neq y$ و هر $\lambda \in]\circ, ۱[$ داشته باشیم

$$f(\lambda x + (۱ - \lambda)y) < \lambda f(x) + (۱ - \lambda) f(y).$$

f را *قویاً* محدب می‌گوییم اگر عدد $\alpha > \circ$ موجود باشد که برای هر $x, y \in C$ و هر $\lambda \in [\circ, ۱]$،

$$f(\lambda x + (۱ - \lambda)y) \leq \lambda f(x) + (۱ - \lambda) f(y) - \alpha\lambda(۱ - \lambda)\|x - y\|^۲.$$

از این تعریف‌ها روشن است که هر تابع قویاً محدب، اکیداً محدب و هر تابع اکیداً محدب، محدب است. مجموعهٔ نقاط کمینهٔ تابع f را با $\mathrm{Argmin} f$ نشان می‌دهیم. برای یک تابع اکیداً محدب، $\mathrm{Argmin} f$ حداکثر یک نقطهٔ دارد. می‌توانیم مقدار تابع محدب f را در خارج مجموعهٔ محدب C برابر با $+\infty$ تعریف کنیم که در این صورت، تابع حاصل روی فضای هیلبرت H محدب خواهد شد.

تعریف ۲.۲. تابع $f : H \to]-\infty, +\infty]$ را سره[۱] گوییم اگر $x \in H$ موجود باشد به‌طوری که $f(x) < +\infty$. مجموعهٔ $D(f) := \{x \in H : \ f(x) < +\infty\}$ را دامنهٔ (دامنهٔ مؤثّر) f می‌خوانیم. تابع $f : H \to]-\infty, +\infty]$ را نیم‌پیوستهٔ پایینی[۲] می‌گوییم اگر در هر نقطه مانند $x \in H$ داشته باشیم

$$\liminf_{y \to x} f(y) = \sup_{r > \circ} \inf_{y \in \hat{B}_r(x)} f(y) \geq f(x)$$

که در آن، $\hat{B}_r(x) = \{y \in H : \ \circ < \|y - x\| < r\}$.

برای مجموعهٔ C در فضای هیلبرت H تعریف می‌کنیم

$$I_C(x) = \begin{cases} \circ & x \in C \\ +\infty & x \notin C \end{cases}$$

[۱] proper [۲] lower semi-continuous

تابع $I_C : H \to]-\infty, +\infty]$ را *تابع نشانگر* مجموعهٔ C می‌نامیم. اگر C ناتهی، بسته و محدب باشد، آن‌گاه I_C محدب، سره و نیم‌پیوستهٔ پایینی است.

قضیهٔ زیر، نتیجه‌ای مهم دربارهٔ وجود کمینه برای تابع‌های محدب، سره و نیم‌پیوستهٔ پایینی است. برای دیدن اثبات آن، می‌توانید نتیجهٔ ۲۳ از فصل ۳ی [۹] را بخوانید.

قضیه ۳.۲. *فرض کنیم مجموعهٔ $C \subseteq H$ بسته و محدب و تابع $f : C \to]-\infty, +\infty]$ محدب، سره و نیم‌پیوستهٔ پایینی باشد به‌طوری که یا C کراندار است و یا*

$$\lim_{\|x\|\to+\infty,\, x\in C} f(x) = +\infty.$$

در این صورت f کمینهٔ خود را روی C اختیار می‌کند؛ یعنی $x_\circ \in C$ وجود دارد که به‌ازای هر $x \in C$، $f(x_\circ) \le f(x)$.

تابع $f : H \to]-\infty, +\infty]$ را در نقطهٔ $x \in D(f)$ *مشتق‌پذیر* می‌گوییم اگر تابع خطی $L : H \to \mathbb{R}$ موجود باشد که

$$\lim_{h\to\circ} \frac{|f(x+h) - f(x) - L(h)|}{\|h\|} = \circ.$$

در این صورت، این نگاشت خطی را که با $\nabla f(x)$ نشان می‌دهیم، *گرادیان* f در x می‌نامیم. f را مشتق‌پذیر می‌نامیم اگر در هر نقطه از دامنه‌اش مشتق‌پذیر باشد. گزاره‌ای ساده در آنالیز محدب بیان می‌کند که تابع مشتق‌پذیر $f : H \to]-\infty, +\infty]$ محدب است اگر و تنها اگر

$$f(x) - f(y) \le \langle \nabla f(x), x - y\rangle, \qquad \forall x, y \in D(f).$$

یکی از کاربردهای الگوریتم نقطهٔ پروکسیمال در مسئلهٔ کمینه‌سازیِ تابع غیرهموار (مشتق‌ناپذیر) است. برای کار با این تابع‌ها، نیاز به جایگزینی مناسب برای مشتق داریم که به آن، *زیردیفرانسیل* می‌گوییم. برای یک تابع محدب نه لزوماً مشتق‌پذیر مانند f، زیردیفرانسیل f در نقطهٔ x که با $\partial f(x)$ نشان داده می‌شود، مجموعهٔ بردارهای v است که به‌جای $\nabla f(x)$ در رابطهٔ بالا می‌توانند قرار گیرند. به عبارت دیگر،

$$\partial f(x) = \{v \in H : \ f(x) - f(y) \le \langle v, x - y\rangle,\ \forall y \in H\}.$$

اگر f در x مشتق‌پذیر باشد، $\partial f(x) = \{\nabla f(x)\}$ و در نقاط مشتق‌ناپذیر، $\partial f(x)$ بیش از یک نقطه دارد. بنابراین در حالت کلی، $\partial f : H \to 2^H$ یک نگاشت *مجموعه‌مقدار*[۱] است؛ یعنی به هر نقطه از

[۱] set-valued

H یک زیرمجموعه از H را نسبت می‌دهد. برای درک بهتر مفهوم زیردیفرانسیل، مثالی ساده می‌زنیم. تابع $g : \mathbb{R} \to \mathbb{R}$ با ضابطهٔ $g(x) = |x|$ را در نظر بگیرید. یک محاسبهٔ ساده نشان می‌دهد که

$$\partial g(x) = \begin{cases} ۱ & x > ۰ \\ [-۱, ۱] & x = ۰ \\ -۱ & x < ۰ \end{cases}$$

در این مثال می‌بینیم که مقدار زیردیفرانسیل در نقاط مشتق‌پذیر، همان مقدار مشتق و در نقطهٔ مشتق‌ناپذیرِ ۰، یک مجموعه است که در واقع شیب همهٔ خط‌های زیرِ نمودار تابع در نقطهٔ صفر است. تابع نرم فضای هیلبرت، مثالی دیگر از تابع‌های محدب است که قویاً محدب و در نتیجه اکیداً محدب است. یک بررسی ساده نشان می‌دهد که برای هر $x \in H \neq ۰$ داریم

$$\partial \frac{۱}{۲}\|\cdot\|^{۲}(x) = \nabla \frac{۱}{۲}\|\cdot\|^{۲}(x) = \{x\}. \quad (۱.۲)$$

همچنین به‌راحتی از تعریف زیردیفرانسیل نتیجه می‌شود که

$$x^* \in \mathrm{Argmin} f \Leftrightarrow ۰ \in \partial f(x^*). \quad (۲.۲)$$

این واقعیت را در مثال سادهٔ زیردیفرانسیلِ قدرمطلق هم می‌توان دید.

ثابت می‌شود که اگر f محدب، سره و نیم‌پیوستهٔ پایینی باشد، آن‌گاه نگاشت

$$(I + \lambda \partial f)^{-۱} : H \to H$$

برای هر $\lambda > ۰$ تک‌مقداری و دامنهٔ آن کل فضای هیلبرت H است. عملگر $(I + \lambda \partial f)^{-۱}$ را *عملگر حلّال* نگاشت ∂f می‌نامند. در مورد مثال قدرمطلق در بالا، می‌توان دید که

$$(I + \lambda \partial g)^{-۱}(x) = \begin{cases} x - \lambda & x > \lambda \\ ۰ & x \in [-\lambda, \lambda] \\ x + \lambda & x < -\lambda \end{cases}$$

می‌گوییم دنبالهٔ $\{x_n\}$ در فضای هیلبرت H *همگرای قوی* به $x \in H$ است اگر $\|x_n - x\| \to ۰$ و *همگرای ضعیف* به x است اگر برای هر $y \in H$ داشته باشیم $\lim_{n\to+\infty}\langle x_n - x, y\rangle = ۰$. خواننده‌ای که با آنالیز تابعی و همگرایی ضعیف آشنا نیست، می‌تواند فضای هیلبرت را متناهی‌–بُعد بگیرد که در این حالت، همگراییِ ضعیف و قوی یکسان خواهند بود. اما در حالت کلی، همگراییِ قوی، همگراییِ ضعیف را نتیجه می‌دهد.

۳. الگوریتم نقطهٔ پروکسیمال

روش نقطهٔ پروکسیمال بر اساس منظم‌سازیِ مسئلهٔ اصلی (۱.۱) بنا شده است به این معنی که به‌جای مسئلهٔ (۱.۱) برای یک تابع محدب f، توجه خود را معطوف می‌کنیم به دنباله‌ای از مسائل کمینه‌سازی برای تابع‌های هدفی که حاصل اغتشاشی از تابع هدف اصلی با تابعی از نرم فضا هستند. فرض کنیم $x_0 \in H$ نقطهٔ شروع باشد. اگر x_{n-1} داده شده باشد، برای $\lambda > 0$، x_n جواب مسئلهٔ کمینه‌سازیِ

$$\mathrm{Min}_{x\in H}\{f(x) + \frac{1}{2\lambda}\|x - x_{n-1}\|^2\} \tag{۱.۳}$$

است. به عبارت دیگر،

$$x_n = \mathrm{Argmin}\{f(x) + \frac{1}{2\lambda}\|x - x_{n-1}\|^2 :\ x \in H\}. \tag{۲.۳}$$

مارتینه ثابت کرد که اگر $\mathrm{Argmin} f \neq \varnothing$، آن‌گاه دنبالهٔ $\{x_n\}$ تولیدشده توسط (۲.۳) همگرایِ ضعیف به یک نقطهٔ کمینهٔ تابع f است. اینکه چرا زیرمسئله‌های (۱.۳) دارای جواب یکتا هستند، نتیجه‌ای از کاربرد قضیهٔ ۳.۲ و ویژگی نرم فضای هیلبرت است. در واقع تابع $x \mapsto f(x) + \frac{1}{2\lambda}\|x - x_{n-1}\|^2$ با توجه به اینکه $\mathrm{Argmin} f \neq \varnothing$ در شرایط قضیهٔ ۳.۲ صدق می‌کند و لذا دارای کمینهٔ یکتا است، زیرا نرم فضای هیلبرت اکیداً محدب است. برای تابع محدب f با مجموعهٔ نقاط کمینهٔ ناتهی، عملگر غیرخطی $\mathrm{Prox}_f : H \to H$ را به‌صورت زیر تعریف می‌کنیم و به آن *نگاشت پروکسیمال* می‌گوییم:

$$\mathrm{Prox}_f(v) := \mathrm{Argmin}\{f(x) + \frac{1}{2}\|x - v\|^2 :\ x \in H\}. \tag{۳.۳}$$

با استفاده از نگاشت پروکسیمال، الگوریتم مارتینه به‌صورت زیر بازنویسی می‌شود:

$$x_n = \mathrm{Prox}_{\lambda f} x_{n-1} \tag{۴.۳}$$

تعبیری زیبا از نگاشت Prox_f برحسب عملگر حلّال ∂f وجود دارد. در واقع به‌راحتی می‌توان دید که

$$\mathrm{Prox}_{\lambda f}(x) = (I + \lambda\partial f)^{-1}(x), \tag{۵.۳}$$

زیرا بنابر تعریف عملگر حلّال، (۱.۲) و (۲.۲)، داریم

$$\begin{aligned} z = (I + \lambda\partial f)^{-1}(x) &\Leftrightarrow x \in (I + \lambda\partial f)(z) = z + \lambda\partial f(z) \\ &\Leftrightarrow 0 \in \partial f(z) + \frac{1}{\lambda}(z - x) \\ &\Leftrightarrow 0 \in \partial(f(z) + \frac{1}{2\lambda}\|z - x\|^2). \end{aligned}$$

همچنین با توجه به تعریفِ زیردیفرانسیل و (۵.۳)، داریم

$$x \in \operatorname{Argmin} f \Leftrightarrow \circ \in \partial f(x)$$
$$\Leftrightarrow x \in x + \lambda \partial f(x)$$
$$\Leftrightarrow x = (I + \lambda \partial f)^{-۱}(x).$$

بنابراین نقطهٔ کمینهٔ f در صورت وجود، نقطهٔ ثابت نگاشت پروکسیمال (۳.۳) است و به‌عکس. همچنین می‌توان دید که به‌ازای هر x, yH،

$$\langle \operatorname{Prox}_f x - \operatorname{Prox}_f y, x - y \rangle \geq \|\operatorname{Prox}_f x - \operatorname{Prox}_f y\|^{۲}.$$

نگاشتی با این ویژگی را *قویاً غیرانبساطی*[۱] می‌گوییم [۴]. در نظریهٔ عملگرهای غیرخطی ثابت می‌شود که دنبالهٔ تکرارهای یک نگاشت قویاً غیرانبساطی تعریف‌شده بر یک فضای هیلبرت با مجموعهٔ نقاط ثابت ناتهی، همگرایِ ضعیف به یک نقطهٔ ثابت این نگاشت است [۴]. بنابراین اگر $\operatorname{Argmin} f \neq \varnothing$، همگراییِ ضعیفِ دنبالهٔ تولیدشده در (۴.۳) به یک نقطهٔ ثابت $\operatorname{Prox}_{\lambda f}$ یا یک نقطهٔ کمینه برای f ثابت می‌شود. به‌طور طبیعی این پرسش پیش می‌آید که آیا همگرایی قوی را نیز برای این الگوریتم می‌توان ثابت کرد؟ ما پاسخ این پرسش را به بخش ۷ موکول می‌کنیم که در آنجا این الگوریتم را برای عملگرهای یکنوا مطالعه خواهیم کرد.

برای اجرای الگوریتم (۴.۳) و به‌دست آوردنِ x_n، در هر مرحله باید یک مسئلهٔ کمینه‌سازی را حل کنیم. ممکن است برای خواننده این پرسش پیش بیاید که تبدیل یک مسئلهٔ کمینه‌سازی به خانواده‌ای از مسائل با ظاهری حتی پیچیده‌تر چه مزیتی می‌تواند داشته باشد و چگونه می‌تواند کار حل مسئلهٔ اول را ساده‌تر کند؟ برای پاسخ به این پرسش، باید نگاهی به چگونگی عملکرد نگاشت پروکسیمال بیندازیم. در واقع $\operatorname{Prox}_f(v)$ نقطه‌ای است که نسبت به تابع f کمترین فاصله را تا بردار v دارد. به همین دلیل به آن نگاشت پروکسیمال می‌گویند.[۲] بنابراین $\operatorname{Prox}_f(v)$ نقطه‌ای بین کمینهٔ f و v است. در (۴.۳) عدد λ پارامتر کنترل است. برای λی بزرگ، چون جملهٔ $\frac{۱}{۲\lambda}\|x_{n-۱} - v\|^{۲}$ کوچکتر می‌شود، x_n به کمینهٔ f نزدیک‌تر و از $x_{n-۱}$ دورتر است. به‌عکس، برای λی کوچک، چون جملهٔ $\frac{۱}{۲\lambda}\|x_{n-۱} - v\|^{۲}$ بزرگتر می‌شود، x_n به $x_{n-۱}$ نزدیک‌تر و از کمینهٔ f دورتر است.

فرض کنیم f هموار (مشتق‌پذیر) باشد و از یکی از روش‌های صریح برای تقریب نقطهٔ کمینه در زیرمسئلهٔ (۱.۳) یا معادلاً (۲.۳) استفاده کنیم؛ مثلاً از روش گرادیانِ کاهشی که روشی متداول در بهینه‌سازی

---

[۲] proximal در لغت به معنی نزدیکی به مبدأ است.

[۱] firmly nonexpansive

محدب است. در این صورت با کوچکتر گرفتنِ پارامتر λ، جملهٔ $\frac{1}{2\lambda}\|x_{n-1}-v\|^2$ بزرگتر می‌شود و تأثیر این جمله در تعیین کمینهٔ تابع $f(x)+\frac{1}{2\lambda}\|x-x_{n-1}\|^2$ بیشتر خواهد شد. چون جملهٔ $\frac{1}{2\lambda}\|x_{n-1}-v\|^2$ یک تابع قویاً محدب است، هرچه ضریب $\frac{1}{\lambda}$ بزرگتر باشد، جواب بهتری برای زیرمسئلهٔ (۱.۳) یا معادلاً (۲.۳) با استفاده از روش گرادیان با تعداد گام‌های کمتر یا نرخ همگراییِ بهتر، به‌دست می‌آید. بنابراین اگر در الگوریتم نقطهٔ پروکسیمال، پارامتر λ را بتوانیم تا جایی که ممکن است کوچک بگیریم، زیرمسئله‌های به‌دست آمده آسان‌تر حل می‌شوند و جواب آنها با نرخ همگرایی بهتری به‌دست می‌آید؛ هرچند آهنگ همگرایی به نقطهٔ کمینهٔ تابع هدفِ اصلی، کندتر می‌شود. به این منظور، الگوریتم نقطهٔ پروکسیمال با جایگزینیِ پارامتر λ با دنباله‌ای از پارامترهای λ_n و با استفاده از عملگر حلّال به‌صورت

$$x_{n+1}=(I+\lambda_n\partial f)^{-1}(x_n) \tag{۶.۳}$$

تبدیل می‌شود. در [۲۲] مطالعه و نشان داده شده است که اگر $\liminf_{n\to+\infty}\lambda_n>0$، آن‌گاه باز هم همگراییِ ضعیف دنباله به یک نقطهٔ کمینهٔ تابع f به‌دست می‌آید. در واقع جمله‌های دنبالهٔ کنترلی را در هر مرحله از تکرار می‌توان کوچکتر گرفت به این شرط که از یک مقدار مثبت کوچک، به صفر نزدیک‌تر نشوند. این شرط بعدها توسط برزیس و لیونز [۱۰] بهتر شد و آنها نشان دادند که شرط $\sum_{n=1}^{+\infty}\lambda_n=+\infty$ برای همگرایی الگوریتم (۶.۳) کافی است. این شرط به‌روشنی بهتر از شرط راکفلر دربارهٔ دنبالهٔ $\{\lambda_n\}$ است. برای مثال، دنبالهٔ $\{\frac{1}{n}\}$ در شرط برزیس و لیونز صدق می‌کند اما در شرط راکفلر صادق نیست. به‌ویژه در این شرط، جمله‌های دنبالهٔ $\{\lambda_n\}$ می‌تواند خیلی کوچک انتخاب شود و حتی می‌تواند همگرا به صفر باشد.

۴. الگوریتم نقطهٔ پروکسیمال و دستگاه گرادیان

در این بخش، نشان می‌دهیم که الگوریتم نقطهٔ پروکسیمال را می‌توان شکل گسستهٔ دستگاه گرادیان تعبیر کرد. مسئلهٔ مقدار اولیهٔ

$$\begin{cases} -x'(t)=\nabla f(x(t)) \\ x(0)=x_0 \end{cases} \tag{۱.۴}$$

را در نظر بگیرید که در آن، f تابعی مشتق‌پذیر با مشتق پیوسته روی فضای هیلبرت H است. (۱.۴) را دستگاه گرادیان *کاهشی* می‌نامیم، زیرا جهت حرکت خم $x(t)$ خلاف جهت گرادیان f در هر نقطه است که جهت بیشترین کاهش مقدار f در آن نقطه را نشان می‌دهد. اگر تابع f محدب و مجموعهٔ نقاط کمینهٔ آن ناتهی باشد، نتیجه‌ای کلاسیک حاکی است که خم جواب (۱.۴) همگرایِ ضعیف به یک نقطهٔ کمینهٔ f است. بنابراین می‌توان از این ویژگی برای کمینه‌سازیِ تابع‌های محدب استفاده کرد.

دستگاه دینامیکی (۱.۴) برای حالتی که f لزوماً مشتق‌پذیر نیست نیز تعمیم داده و مطالعه شده است. در این حالت، باید گرادیان تابع محدب را با زیردیفرانسیل معرفی‌شده در بخش ۲ عوض کنیم. فرض کنیم $f\,:\,H\to]-\infty,+\infty]$ یک تابع محدب، سره و نیم‌پیوستهٔ پایینی باشد. ثابت می‌شود که شمول دیفرانسیلی[1]

$$\begin{cases} -x'(t)\in\partial f(x(t)) \\ x(\circ)=x_\circ \end{cases} \tag{۲.۴}$$

دارای جواب یکتا در بازهٔ $[\circ,+\infty[$ است [۱۱، ۱۹]. براک [۱۱] ثابت کرد که اگر مجموعهٔ نقاط کمینهٔ f ناتهی باشد، آن‌گاه جواب (۲.۴) همگرای ضعیف به یک نقطهٔ کمینهٔ f است. از این ویژگی که تعمیم حالت هموار است، می‌توان برای تقریب جواب مسئلهٔ کمینه‌سازیِ غیرهموار استفاده کرد. اما چون رایانه قادر به درک خم‌های پیوسته نیست، برای به‌دست آوردن الگوریتمی برای تقریب نقطهٔ کمینه، ناچار به گسسته‌سازی (۲.۴) هستیم.

به دو روش متفاوت می‌توان (۲.۴) را گسسته‌سازی کرد. روش گسسته‌سازی پیشرو[2] و روش گسسته‌سازی پسرو[3]. گسسته‌سازی پسرو به شمول

$$x_{k-1}-x_k\in\lambda_k\partial f(x_k)$$

منجر می‌شود که می‌توان با استفاده از عملگر حلّال، آن را به‌صورت

$$x_k=(I+\lambda_k\partial f)^{-1}(x_{k-1})$$

نوشت که همان الگوریتم نقطهٔ پروکسیمال است. بنابراین الگوریتم نقطهٔ پروکسیمال به تعبیر دیگر، شکل گسستهٔ دستگاه گرادیان کاهشیِ غیرهموار با روش گسسته‌سازی پسرو است. ویژگی همگرایی دستگاه گرادیان (۲.۴) همان‌طور که دیدیم، به شکل گسسته (الگوریتم نقطهٔ پروکسیمال) به‌ازای هر انتخاب پارامترهای λ_k که در شرط $\sum_{k=1}^{+\infty}\lambda_k=+\infty$ صدق کند، منتقل می‌شود.

۵. الگوریتم گرادیان پروکسیمال

مسئلهٔ کمینه‌سازیِ

$$\mathrm{Min}\{f(x)+g(x):\ x\in H\} \tag{۱.۵}$$

---

[1] differential inclusion با توجه به مطالب بخش ۲، ∂f احتمالاً چندمقداری است. بنابراین در اینجا یک شمول دیفرانسیلی داریم و نه لزوماً یک معادلهٔ دیفرانسیلی.

[2] forward discretization [3] backward discretization

را در نظر بگیرید که در آن، تابع $f : H \to \mathbb{R}$ محدب و مشتق‌پذیر و تابع $g : H \to]-\infty, +\infty]$ محدب، سره و نیم‌پیوستهٔ پایینی است. نمونه‌ای از این مسائل وقتی رخ می‌دهد که تابع هدف[1] را بتوان به مجموع دو تابع که یکی از آنها مشتق‌پذیر است (در اینجا f) تجزیه کرد. برای مثال، مسئلهٔ کمینه‌سازی مقید

$$\mathrm{Min}\{f(x) : \ x \in C\} \tag{۲.۵}$$

اگر مجموعهٔ C بسته و محدب باشد، معادل با مسئلهٔ کمینه‌سازیِ نامقید

$$\mathrm{Min}\{f(x) + I_C(x) : \ x \in H\} \tag{۳.۵}$$

است که در آن، I_C تابع نشانگر مجموعهٔ C است. الگوریتم گرادیان پروکسیمال برای مسئلهٔ (۱.۵) که از ترکیب الگوریتم‌های تندترین کاهش[2] و پروکسیمال به‌ترتیب، برای f و g حاصل می‌شود، به‌صورت

$$x_{k+1} = \mathrm{Prox}_{\lambda_k g}(x_k - \lambda_k \nabla f(x_k)) \tag{۴.۵}$$

است که در آن، $\lambda_k > 0$ طول گام این الگوریتم را نشان می‌دهد. انتخاب مناسب پارامترهای λ_k به همگراییِ ضعیف دنبالهٔ حاصل به یک نقطهٔ کمینهٔ $f + g$، منتهی می‌شود [۲۰]. به‌سادگی می‌توان دید که

$$\mathrm{Prox}_{I_C}(x) = \mathrm{Proj}_C x. \tag{۵.۵}$$

بنابراین الگوریتم گرادیان پروکسیمال برای مسئلهٔ (۳.۵) به‌شکل

$$x_{k+1} = \mathrm{Proj}_C(x_k - \lambda_k \nabla f(x_k)) \tag{۶.۵}$$

است. الگوریتم (۴.۵) را از این نظر که تابع هدف به‌صورت جمع دو تابع محدب نوشته می‌شود، *الگوریتم شکافت*[3] هم می‌نامند.

الگوریتم گرادیان پروکسیمال را می‌توان یک روش تکراری برای تقریب یک نقطهٔ ثابت نیز دانست. در واقع x^* جواب مسئلهٔ (۱.۵) است اگر و تنها اگر نقطهٔ ثابت عملگر

$$\mathrm{Prox}_{\lambda g}(I - \lambda \nabla f)$$

باشد که در آن، I عملگر همانی است [۲۰]. روش گرادیان پروکسیمال با تکرار عملگر

$$(I + \lambda \partial g)^{-1}(I - \lambda \nabla f),$$

[2] روشی معمول در بهینه‌سازی محدب که بر اساس الگوریتم تکراری $x_{k+1} = x_k - \lambda_k \nabla f(x_k)$ بنا شده است و در واقع گسسته‌سازی پیشرو در دستگاه گرادیان ۱.۴ است.

[1] objective function [3] splitting algorithm

یک نقطهٔ ثابت این عملگر و بنابراین یک جواب مسئلهٔ (۱.۵) را تقریب می‌زند.

۶. ترکیب نگاشت‌های پروکسیمال

یکی از مسائل آنالیز غیرخطی در حوزهٔ بهینه‌سازی، مسئلهٔ تقریب زدنِ نقاط مشترک دو مجموعهٔ محدب و بستهٔ دلخواه با اشتراک ناتهی است. فون‌نویمان [۲۴] در سال ۱۹۳۳ به این مسئله در حالتی که دو مجموعهٔ $C_۱$ و $C_۲$ زیرفضاهایی بسته از فضای هیلبرت H باشند، توجه کرد و نشان داد که با تصویر متوالی یک نقطهٔ دلخواه روی این زیرفضاها، دنباله‌ای به‌دست می‌آید که همگرای قوی به نقطه‌ای در این مجموعه است که نزدیکترین نقطه به نقطهٔ اولیه است. این نتیجه را برگمن [۸] به مجموعه‌های بسته و محدب تعمیم داد. البته در این حالت با جایگزینیِ زیرفضاهای بسته با مجموعه‌های بسته و محدب، فقط همگرایی ضعیف به‌دست می‌آید. هاندال [۱۴] ثابت کرد که همگرایی قوی در حالت کلی با این روش به‌دست نمی‌آید.

در اینجا می‌خواهیم این نتیجه را به شکلی دیگر تعمیم دهیم. چنان‌که قبلاً گفتیم، اگر $C_۱$ و $C_۲$ مجموعه‌های محدب و بسته باشند، تابع‌های نشانگر آنها تابع‌های محدب و نیم‌پیوستهٔ پایینی هستند. قبلاً عملگر حلّال را برای آنها محاسبه کردیم. در واقع عملگر حلّال برای تابع‌های نشانگر $C_۱$ و $C_۲$ بنابر (۵.۵)، همان عملگر تصویر روی این مجموعه‌های محدب و بسته است. بنابراین روش فون‌نویمان در تصویر متوالی روی مجموعه‌های بسته و محدب، اثر نگاشت حلّال زیردیفرانسیل تابع نشانگر روی یک نقطهٔ دلخواه است. حال به‌جای این نگاشت حلّال خاص، نگاشت‌های حلّال متناظر برای دو تابع محدب f و g را به‌کار می‌بریم. در [۵] با فرض‌های مناسب روی دنبالهٔ پارامترهای $\{\lambda_n\}$ و $\{\mu_n\}$ و بافرضِ ناتهی بودنِ مجموعهٔ نقاط ثابت $\mathrm{Prox}_f\mathrm{Prox}_g$ نشان داده شده است که دنبالهٔ حاصل از روش تکراریِ

$$x_{n+۱} = \mathrm{Prox}_{\mu_n g}\mathrm{Prox}_{\lambda_n f}(x_n)$$

همگرایِ ضعیف به یک نقطهٔ ثابت $\mathrm{Prox}_f\mathrm{Prox}_g$، می‌شود.

از ترکیب نگاشت‌های پروکسیمال برای حل مسئلهٔ (۱.۵) با این فرض که مجموعهٔ جواب مسئلهٔ ناتهی باشد، استفاده می‌شود. پستی[۱] در [۲۱] نشان داد که با قرار دادنِ فرض مناسب روی دنباله‌های $\{\lambda_n\}$ و $\{\mu_n\}$، میانگین وزنداری از دنبالهٔ تولیدشده، به نقطهٔ کمینهٔ $f+g$ همگرا خواهد شد. این روش به این دلیل به کار گرفته می‌شود که به‌دست آوردنِ حلّال یک تابع محدب (در اینجا $f+g$) غالباً آسان نیست و همان‌طور که در بخش ۳ گفتیم، لازمهٔ آن، حل یک مسئلهٔ بهینه‌سازی دیگر است. اما در صورت شکافتن تابع به مجموع دو تابع محدب، ممکن است به‌دست آوردن حلّال تک‌تک آنها آسان‌تر باشد. این

[۱] Passty

روش علاوه بر کاربردهای متعدد، اخیراً در به‌دست آوردن نقاط میانگین و میانه برای مجموعهٔ نقاط دلخواه در فضاهای متری خاصی موسوم به *فضاهای آدامار*[۱] به‌طور موفقیت‌آمیزی به‌کار رفته است [۲].

۷. الگوریتم نقطهٔ پروکسیمال و عملگرهای یکنوا

تعمیمی سودمند از زیردیفرانسیل تابع‌های محدب، عملگر یکنوا است. این عملگرها که ابتدا توسط مینتی [۱۸] معرفی و وارد آنالیز غیرخطی شدند، تعمیم غیرخطیِ عملگرهای مثبت در نظریهٔ عملگرهای خطی هستند. عملگر مجموعه‌مقدار $A : H \to 2^H$ را در نظر بگیرید. دامنهٔ A مجموعهٔ نقاطی از فضای H است که مجموعهٔ تهی به آنها منسوب نمی‌شود. بنابراین

$$D(A) := \{x \in H : A(x) \neq \varnothing\}.$$

بُرد و نمودار عملگر غیرخطی A را به‌ترتیب، به‌صورت

$$R(A) := \bigcup_{x \in D(A)} A(x)$$

و

$$G(A) := \{(x, y) :\ x \in D(A),\ \ y \in A(x)\}$$

تعریف می‌کنیم. به‌روشنی یک تناظر یک‌به‌یک بین عملگرهای غیرخطی و نمودار آنها که زیرمجموعه‌هایی از $H \times H$ هستند، وجود دارد. بنابراین می‌توان یک عملگر غیرخطی را زیرمجموعه‌ای از $H \times H$ دانست.

عملگر A را *یکنوا* می‌نامیم اگر به‌ازای هر $x, y \in D(A)$، هر $x^* \in A(x)$ و هر $y^* \in A(y)$ داشته باشیم $\langle x^* - y^*, x - y \rangle \geq 0$. عملگر یکنوای A را *یکنوای ماکسیمال* می‌نامیم اگر به‌عنوان زیرمجموعه‌ای از $H \times H$ ماکسیمال باشد، یعنی هر عملگر غیرخطی دیگری که نمودار آن به‌طور سره شامل نمودار A باشد، یکنوا نباشد. در فضای هیلبرت این، معادل است با اینکه $R(I + A) = H$ که در اینجا $R(A)$ بُرد عملگر A و $I : H \to H$ نگاشت همانی است. به عبارت دیگر، عملگر یکنوای A ماکسیمال است اگر برای هر $y \in H$، $x \in D(A)$ وجود داشته باشد که $y \in x + A(x)$. چنان‌که گفتیم، اگر A تک‌مقداری و خطی باشد، عملگر یکنوا همان عملگر مثبت است که در فضای متناهی–بُعد، همان ماتریس‌های معین مثبت و نیم‌معین مثبت هستند. زیردیفرانسیل یک تابع محدب و نگاشت تصویر روی یک مجموعهٔ محدب و بسته، نمونه‌هایی از عملگرهای یکنوا هستند که قبلاً به آنها اشاره کردیم. ثابت می‌شود که این عملگرها در واقع یکنوای ماکسیمال هستند.

[۱] Hadamard spaces

عملگرهای یکنوا در دو زمینهٔ متفاوت از ریاضیات کاربردی اهمیت بسیار پیدا کرده‌اند: یکی در نظریهٔ بهینه‌سازی و نامساوی‌های تغییراتی[۱] و دیگری در نظریهٔ نیمگروه‌های غیرخطی[۲] برای معادلات تحوّلی[۳] و معادلات دیفرانسیل با مشتقات جزئی. بسیاری از معادلات دیفرانسیل با مشتقات جزئی در ریاضی–فیزیک مانند معادلهٔ موج، معادلهٔ حرارت و معادلهٔ شرودینگر مثال‌هایی ملموس از معادلات تحوّلی (وابسته به زمان) شامل عملگرهای یکنوا هستند. برای نمونه، مسئلهٔ کمینه‌سازی مقید برای تابع محدب $f : \mathbb{R}^n \to \mathbb{R}$ با قید $Ax = b$ را در نظر بگیرید که A ماتریسی $m \times n$ است و $b \in \mathbb{R}^m$. با استفاده از روش KKT[۴] که یکی از روش‌های متداول در بهینه‌سازی مقید و غیرخطی است، ثابت می‌شود که جواب این مسئلهٔ کمینه‌سازیِ مقید، صفرِ عملگر یکنوای

$$F(x,y) = \begin{bmatrix} \partial f(x) + A^T y \\ b - Ax \end{bmatrix}$$

است.

یکی از مسائل مهم در نظریهٔ عملگرهای یکنوا، پیدا کردنِ (تقریب‌زدنِ) صفر یک عملگر یکنوا است. مثال بالا اهمیتِ پیدا کردن صفرِ عملگر یکنوا در بهینه‌سازی را روشن می‌کند. اگر عملگر یکنوا، زیردیفرانسیل یک تابع محدب باشد، صفرِ عملگر در صورت وجود، نقطهٔ کمینهٔ تابع محدب خواهد بود. این نتیجه به‌راحتی از نامساوی زیردیفرانسیل به‌دست می‌آید. بنابراین یکی از مهم‌ترین بحث‌ها در نظریهٔ عملگرهای یکنوا، پیدا کردن صفرِ این عملگرها و تقریب آن است. برای به‌دست آوردن صفرِ یک عملگر یکنوا، یکی از مهم‌ترین روش‌ها استفاده از الگوریتم نقطهٔ پروکسیمال است. در این حالت باید از این الگوریتم در قالب (۶.۳) البته با جایگزینی زیردیفرانسیل تابع محدب f با یک عملگر یکنوا، استفاده کرد. به این ترتیب، الگوریتم نقطهٔ پروکسیمال برای عملگر یکنوای ماکسیمال A با لحاظ کردنِ خطای محاسبهٔ عملگر حلّال در هر گام، به‌صورت

$$x_n = (I + \lambda_n A)^{-۱} x_{n-۱} + e_n \tag{۱.۷}$$

درمی‌آید که $\{e_n\}$ دنبالهٔ خطاها است. نخستین کسی که این الگوریتم را برای عملگرهای یکنوا مطالعه کرد، راکفلر [۲۲] بود. او نشان داد که (۱.۷) برای پارامترهای λ_n با شرط $\liminf_{n\to+\infty} \lambda_n > ۰$ و دنبالهٔ خطاهای $\{e_n\}$ با شرط $\sum_{n=۱}^{\infty} \|e_n\| < +\infty$ همگرای ضعیف است. راکفلر دو پرسش مهم را نیز مطرح کرد: یکی اینکه آیا در (۱.۷) همگرایی قوی نیز به‌دست می‌آید؟ و دوم اینکه آیا شرط جمع‌پذیری[۵]

[۴] مخفف Karush–Kuhn–Tucker

[۱] variational inequalities [۲] nonlinear semigroups [۳] evolution equations [۵] summability

روی دنبالهٔ خطاها (یعنی $\sum_{n=1}^{\infty} \|e_n\| < +\infty$) را می‌توان حذف کرد یا با شرطی ضعیف‌تر جایگزین کرد؟

در سال ۱۹۹۱ به سؤال اول راکفلر پاسخ منفی داده شد و با یک مثال نقض در فضای l^2 نشان داده شد که الگوریتم نقطهٔ پروکسیمال حتی در حالت خاصی که عملگر یکنوا، زیردیفرانسیل یک تابع محدب باشد نیز همگرای قوی نیست [۱۲]. بنابراین برای رسیدن به همگراییِ قوی که البته از نظر عددی بسیار مهم است،[۱] باید چاره‌ای دیگر اندیشید. دو روش اصلی برای اصلاح الگوریتم نقطهٔ پروکسیمال راکفلر به‌منظور رسیدن به همگرایی قوی، متداول است. اولی استفاده از یک منظم‌سازی مناسب است. این کاری بود که قبلاً در نظریهٔ نقطهٔ ثابت برای اثبات همگرایی قویِ دنبالهٔ تکرارهای یک نگاشت غیرانبساطی انجام شده بود [۱۳]. در آنجا طرحی برای به‌دست آوردن همگراییِ قوی به نقطهٔ ثابت یک نگاشت غیرانبساطی، پیشنهاد شده بود که بعداً در [۲۶، ۲۷] به‌صورت کلی‌تر بیان شد و برای به‌دست آوردن همگراییِ قوی در الگوریتم نقطهٔ پروکسیمال به‌کار رفت. برای مطالعهٔ بیشتر در این زمینه، [۱۶، ۲۵] و مراجع موجود در آنها را ببینند. روش دومِ اصلاح الگوریتم راکفلر برای دست یافتن به همگرایی قوی، بر اساس تصویر جمله به‌دست آمده در هر مرحله روی مجموعه‌های مناسب است که در [۲۳] سازمان یافت.

مسئلهٔ دوم راکفلر یعنی بهینه کردن شرط روی خطاها نیز در [۶، ۱۵] مورد توجه قرار گرفته است.

۸. کاربرد در نظریهٔ نقطهٔ ثابت

در این بخش پایانی، کاربردی از الگوریتم نقطهٔ پروکسیمال را در نظریهٔ نقطهٔ ثابت مورد توجه قرار می‌دهیم. نظریهٔ نقطهٔ ثابت متری، یکی از شاخه‌های فعال در آنالیز غیرخطی است که با قضیهٔ مشهور نقطهٔ ثابت باناخ آغاز شد و کاربردهایی فراوان به‌ویژه در اثبات قضیه‌های وجودی در آنالیز ریاضی و معادلات دیفرانسیل و معادلات انتگرال پیدا کرد. در این شاخه از آنالیز ریاضی، به‌دنبال یافتنِ نقطهٔ ثابت نگاشت‌های غیرخطی هستیم که غالباً تعمیم‌هایی از انقباض‌ها هستند.

نقطهٔ $x \in H$ را *نقطهٔ ثابت* نگاشت $T : H \to H$ می‌نامیم اگر $Tx = x$. مجموعهٔ نقاط ثابت T را با $\mathrm{Fix}(T)$ نشان می‌دهیم. عملگر $T : H \to H$ را *انقباضی* می‌خوانیم اگر عددی مانند $0 < k < 1$ موجود باشد که به‌ازای هر $x, y \in H$،

$$\|Tx - Ty\| \leq k\|x - y\|$$

[۱] در آنالیز تابعی ثابت می‌شود که مجموعه‌های باز در توپولوژی ضعیف در یک فضای نامتناهی–بُعد، کراندار نیستند. بنابراین در صورت همگراییِ ضعیف یک دنباله به یک نقطه، چون آن دنباله از جایی به بعد در همسایگی بازِ ضعیف نقطهٔ حدی قرار می‌گیرد و همسایگی‌ها بی‌کران هستند، این دنباله ممکن است تقریب خوبی برای نقطهٔ کمینه نباشد.

و آن را *غیرانبساطی*[۱] می‌نامیم اگر به‌ازای هر $x, y \in H$،

$$\|Tx - Ty\| \le \|x - y\|.$$

یکی از مسائل مورد توجه در کنار مسئلهٔ وجود نقطهٔ ثابت برای یک نگاشت غیرخطی (غالبا غیرانبساطی یا تعمیم‌هایی از آن)، مسئلهٔ تقریب این نقطهٔ ثابت است. بنابر قضیهٔ مشهور نقطهٔ ثابت باناخ، هر نگاشت انقباضی روی یک فضای متری کامل، دارای نقطهٔ ثابت یکتا است. در واقع تکرارهای اثر آن نگاشت انقباضی روی یک نقطهٔ دلخواه از فضای متری کامل، دنباله‌ای تشکیل می‌دهد که به نقطهٔ ثابت یکتای آن همگرا است. اما در مورد نگاشت‌های غیرانبساطی، این مطلب صادق نیست. مثالی ساده، تابع $T : H \to H$ با ضابطه $Tx = -x$ است. در [۳] نشان داده شده است که میانگین تکرارهای یک نگاشت غیرانبساطی T با مجموعهٔ نقاط ثابت ناتهی در یک فضای هیلبرت، همگرایِ ضعیف به یک نقطهٔ ثابت آن است؛ یعنی اگر $\mathrm{Fix}(T) \neq \varnothing$، آن‌گاه برای هر $x \in H$، $\frac{۱}{n}\sum_{i=۰}^{n-۱} T^i x$ همگرایِ ضعیف به یک نقطهٔ ثابت T است. روش‌های متعددی برای به‌دست آوردن همگراییِ ضعیف و قویِ تکرارهای یک نگاشت غیرانبساطی وجود دارد. در اینجا می‌خواهیم به کاربرد الگوریتم نقطهٔ پروکسیمال برای تقریب نقطهٔ ثابت یک نگاشت غیرانبساطی اشاره کنیم.

خواننده می‌تواند به‌راحتی بررسی کند که اگر T یک نگاشت غیرانبساطی روی فضای هیلبرت H باشد، $I - T$ یک عملگر یکنوا خواهد بود و بدیهی است که نقاط ثابت T، صفرهای عملگر یکنوای حاصل است. بنابر مطالب بخش قبل، با به‌کار بردنِ الگوریتم نقطهٔ پروکسیمال برای عملگر یکنوای $I - T$، می‌توان نقطهٔ ثابت T را تقریب زد اما مشکلی که هنگام استفاده از این الگوریتم به وجود می‌آید، محاسبه حلّال عملگر یکنوا است. الگوریتم (۱.۷) را برای عملگر یکنوای A در حالتی که $e_n = ۰$ می‌توان به‌صورت

$$x_{n-۱} - x_n \in \lambda_n A x_n$$

نیز نوشت که با فرضِ $A = I - T$ و غیرانبساطی بودنِ T، به‌صورت معادلهٔ

$$x_n = \frac{۱}{۱+\lambda_n} x_{n-۱} + \frac{\lambda_n}{۱+\lambda_n} T x_n$$

درمی‌آید. بنابراین با فرضِ در اختیار داشتنِ λ_n و $x_{n-۱}$، نقطهٔ x_n نقطهٔ ثابت نگاشت

$$x \mapsto \frac{۱}{۱+\lambda_n} x_{n-۱} + \frac{\lambda_n}{۱+\lambda_n} T x$$

[۱] nonexpansive

است. چون T غیرانبساطی است، این نگاشت یک نگاشت انقباضی با ثابت انقباضِ $\frac{\lambda_n}{۱+\lambda_n}$ است و بنابر قضیهٔ نقطهٔ ثابت باناخ، یک نقطهٔ ثابت یکتا دارد که همان x_n، یعنی گام بعدی در الگوریتم نقطهٔ پروکسیمال است. بنابراین در حالتی که عملگر یکنوای A به‌شکل $I-T$ است، داریم

$$(I+\lambda A)^{-۱}x = \mathrm{Fix}(T_x^\lambda)$$

که در آن،

$$T_x^\lambda u = \frac{۱}{۱+\lambda}x + \frac{\lambda}{۱+\lambda}Tu$$

یک نگاشت انقباضی با ثابت انقباضِ $\frac{\lambda}{۱+\lambda}$ است و $\mathrm{Fix}(T_x^\lambda)$ یک مجموعهٔ تک‌عضوی است. بنابراین در این حالت، نگاشت حلّال در هر نقطه، نقطهٔ ثابت یک انقباض وابسته به آن نقطه است.

برای محاسبهٔ هر گام از الگوریتم نقطهٔ پروکسیمال برای تقریب یک نقطهٔ ثابت نگاشت T، باید یک مسئلهٔ نقطهٔ ثابت را برای نگاشت $T_{x_{n-۱}}^{\lambda_n}$ حل کرد. اما چون این نگاشت، انقباضی است، می‌توان با چند بار تکرار این نگاشت روی یک نقطهٔ دلخواه، تقریب خوبی برای x_n به‌دست آورد. نگاهی به برهان همگراییِ تکرارها در قضیهٔ نقطهٔ ثابت باناخ، نشان می‌دهد که هرچه ثابت انقباض، یعنی $\frac{\lambda_n}{۱+\lambda_n}$ کوچکتر باشد، با تعداد تکرارهای کمتر و با آهنگ همگرایی بالاتر، می‌توانیم تقریبی بهتر برای نقطهٔ ثابت $T_{x_{n-۱}}^{\lambda_n}$، یعنی x_n به‌دست آوریم. این مجدداً اهمیت اثبات همگرایی الگوریتم نقطهٔ پروکسیمال برای پارامترهای کوچکتر λ_n را نشان می‌دهد.

مراجع

[1] Aubin, J. P., Ekeland, I., *Applied Nonlinear Analysis*, Reprint of the 1984 original, Dover Publications Inc., Mineola, NY, 2006.

[2] Bacak, M., *Convex Analysis and Optimization in Hadamard Spaces*, De Gruyter Series in Nonlinear Analysis and Applications, **22**, De Gruyter, Berlin, 2014.

[3] Baillon, J. B., Un théorème the type ergodique pour les contraction non linéaires dans un espace de Hilbert, *C. R. Acad. Sci. Paris*, **280** (1975), A1511–A1514.

[4] Bauschke, H. H., Combettes, P. L., *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC., Springer-Verlag, New York, 2011.

[5] Bauschke, H. H., Combettes, P. L., Reich, S., The asymptotic behavior of the composition of two resolvents, *Nonlinear Anal.*, **60** (2005), 283–301.

[6] Boikanyo, O. A., Morosanu, G., A proximal point algorithm converging strongly for general errors, *Optim. Lett.*, **4** (2010), 635–641

[8] Bregman, L. M., The method of successive projection for finding a common point of convex sets, *Sov. Math. Dokl.*, **6** (1965), 688–692.

[9] Brézis, H., *Functional analysis, Sobolev Spaces and Partial Differential Equations*, Universitext, Springer-Verlag, New York, 2011.

[10] Brézis, H., Lions, P. L., Produits infinis de résolvantes, *Israel J. Math.*, **29** (1978), 329–345.

[11] Bruck, R. E., Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, *J. Func. Anal.*, **18** (1975), 15–26.

[12] Güler, O., On the convergence of the proximal point algorithm for convex minimization, *SIAM J. Control Optim.*, **29** (1991), 403–419.

[13] Halpern, B., Fixed points of nonexpanding maps, *Bull. Amer. Math. Soc.*, **73** (1967), 957–961.

[14] Hundal, H., An alternating projection that does not converge in norm, *Nonlinear Anal.*, **57** (2004), 35–61.

[15] Khatibzadeh, H., Some remarks on the proximal point algorithm, *J. Optim. Theory Appl.*, **153** (2012), 769–778.

[16] Khatibzadeh, H., Ranjbar, S., On the strong convergence of Halpern type proximal point algorithm, *J. Optim. Theory Appl.*, **158** (2013), 385–396.

[17] Martinet, B., Régularisation d'inéquations variationnelles par approximations successives, *Rev. Francaise Informat. Recherche Operationnelle*, **4** (1970), 154–158.

[18] Minty, G. J., Monotone (nonlinear) operators in Hilbert space, *Duke Math. J.*, **29** (1962), 341–346.

[19] Morosanu, G., *Nonlinear Evolution Equations and Applications*, vol. **26**, Reidel, Dordrech, 1988.

[20] Parikh, N., Boyd, S., Proximal algorithms, *Fundations and Trends in Optimization*, **1** (2013), 123–231.

[21] Passty, G. B., Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, *J. Math. Anal. Appl.*, **72** (1979), 383–390.

[22] Rockafellar, R. T., Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.*, **14** (1976), 877–898.

[23] Solodov, M. V., Svaiter, B. F., Forcing strong convergence of proximal point iterations in a Hilbert space, *Math. Program. Ser. A*, **87** (2000), 189–202.

[24] von Neumann, J., *Functional Operators*, Vol. II, Princeton University Press, 1950.

[25] Wang, F., Cui, H., On the contraction proximal point algorithms with multi parameters, *J. Global Optim.*, **54** (2012), 485–491.

[25] Wang, F., Cui, H., On the contraction proximal point algorithms with multi parameters, *J. Global Optim.*, **54** (2012), 485–491.

[26] Xu, H. K., Iterative algorithms for nonlinear operators, J. *Lond. Math. Soc.*, **66** (2002), 240–256.

[27] Xu, H. K., A regularization method for the proximal point algorithm, *J. Glob. Optim.*, **36** (2006), 115–125.

تاریخ ارسال: ۹۵/۱۱/۲۰؛ تاریخ بازنگری: ۹۶/۸/۱۰؛ تاریخ پذیرش: ۹۶/۸/۱۵

هادی خطیب‌زاده: دانشگاه زنجان، دانشکدهٔ علوم، گروه ریاضی

رایانامه: hkhatibzadeh@znu.ac.ir