УДК 517.9

Т. А. Войтова, С. В. Денисов, В. В. Семенов

Альтернуючий проксимальний алгоритм для задачі дворівневої опуклої мінімізації

(Представлено членом-кореспондентом НАН України С. І. Ляшком)

Розглянуто питання розв'язання дворівневої опуклої задачі мінімізації за допомогою альтернуючого проксимального алгоритму. При деяких метричних умовах на функціонал задачі першого рівня доведено теореми про сильну та слабку збіжність.

В оптимізації та теорії некоректних задач популярним є такий підхід до розв'язання задач з неєдиним розв'язком [1]: задачі ставлять у відповідність до родини збурених задач, однозначно та коректно розв'язних. Частинний розв'язок початкової задачі одержують як границю розв'язків збурених задач при зменшенні збурень. Знайдені так частинні розв'язки задовольняють певні додаткові умови, наприклад, мінімальність норми нормального розв'язку оптимізаційної задачі, отриманого методом тіхоновської регуляризації.

Іншим джерелом задач вигляду

$$f_2(x) \to \min, \qquad x \in \operatorname{argmin} f_1$$

є метод штрафів та задачі оптимізації за послідовно заданими критеріями (лексикографічна, послідовна або багаторівнева оптимізація) [2, 3]. Також задачу оптимального керування [4]

$$F(y,u) \to \min, \qquad Ly = Bu$$

можна переформулювати у вигляді

$$F(y,u) \to \min, \qquad (y,u) \in \operatorname*{argmin}_{(\xi,\eta)} \|L\xi - B\eta\|^2.$$

У даному повідомленні розглядається дворівнева задача опуклої мінімізації в гільбертовому просторі. Метою є дослідження збіжності схем вигляду

$$\begin{cases} y_n = \operatorname*{argmin}_{y\in H}\left\{\lambda_n f_2(y) + \dfrac{1}{2}\|y - x_n\|^2\right\}, \\ x_{n+1} = \operatorname*{argmin}_{y\in H}\left\{\lambda_n \alpha_n f_1(y) + \dfrac{1}{2}\|y - y_n\|^2\right\}. \end{cases}$$

При $\alpha_n = 1$ маємо альтернуючий метод проксимальної декомпозиції для задачі $f_1 + f_2 \to \min$ [5–7].

Основний результат такий: для опуклих напівнеперервних знизу функціоналів f_2 та опуклих напівнеперервних знизу функціоналів f_1, що задовольняють деяку метричну умову, доведено теореми збіжності (сильної та слабкої) наведеної схеми. У роботі використана техніка, розвинута в [5–9]. Усі необхідні відомості з нелінійного та опуклого аналізу є в роботах [10–12].

Постановка задачі. Нехай H — дійсний простір Гільберта з скалярним добутком $(\cdot,\cdot)$ та нормою $\|\cdot\|$. Нехай f_1, $f_2\colon H\to\mathbb{R}\bigcup\{+\infty\}$ — власні опуклі напівнеперервні знизу функціонали. Будемо вважати, що

$$\operatorname{argmin} f_1 \neq \varnothing \qquad \text{та} \qquad \min f_1 = 0.$$

Розглянемо задачу

$$f_2(x)\to\min, \qquad x\in \operatorname{argmin} f_1. \tag{1}$$

Припустимо, що

$$0\in \operatorname{int}(\operatorname{dom} f_2 - \operatorname{argmin} f_1).$$

Тоді задача (1) еквівалентна включенню

$$\text{знайти} \quad x\in H\colon \quad 0\in\partial f_2(x) + N_{\operatorname{argmin} f_1}x,$$

де $N_M x$ — нормальный конус замкненої опуклої множини $M\subseteq H$ в точці $x\in H$, тобто

$$N_M x = \begin{cases} \{z\in H\colon (z, y-x)\leqslant 0 \ \forall\, y\in M\}, & \text{якщо} \qquad x\in M,\\ \varnothing, & \text{інакше}. \end{cases}$$

Позначимо через C множину $\operatorname{argmin} f_1$, а через S — множину розв'язків задачі (1).

Будемо розглядати функціонали f_1, що задовольняють умову

(A1) $\exists\, k>0 : f_1(x)\geqslant k\cdot d_C^2(x) = k\cdot\min\limits_{y\in C}\|x-y\|^2 \ \forall\, x\in H$.

Має місце

Лема 1. *Нехай для f_1 виконується (A1). Тоді для $z\in C$ і $w\in N_C z$ має місце нерівність*

$$(w,x) - f_1(x) - (w,z)\leqslant \frac{1}{4k}\|w\|^2 \qquad \forall\, x\in H.$$

Доведення. Оскільки $(w,x)-f_1(x)-(w,z)\leqslant f_1^*(w)-\sigma_C(w)$, то достатньо довести оцінку $f_1^*(\cdot)-\sigma_C(\cdot)\leqslant \dfrac{1}{4k}\|\cdot\|^2$. Ця оцінка випливає з однорідності опорної функції σ_C, нерівності

$$f_1^*(\cdot)\leqslant (k\cdot d_C^2)^*(\cdot) = 2k\left(\frac{d_C^2}{2}\right)^*\left(\frac{\cdot}{2k}\right)$$

та рівності

$$\left(\frac{d_C^2}{2}\right)^* = \left(\frac{\|\cdot\|^2}{2}\oplus\chi_C\right)^* = \frac{\|\cdot\|^2}{2} + \sigma_C,$$

де χ_C — індикаторна функція множини C; $\oplus$ — інфімальна конволюція.

Зауваження 1. Якщо припустити існування $k>0$, $p>1$, таких, що

$$f_1(x)\geqslant k d_C^p(x) = k\min_{y\in C}\|x-y\|^p \qquad \forall\, x\in H,$$

то для $z \in C$ і $w \in N_C z$ можна довести оцінку

$$(w, x) - f_1(x) - (w, z) \leqslant \frac{1}{q(pk)^{q-1}} \|w\|^q \qquad \forall\, x \in H,$$

де $q > 1$ та $1/q + 1/p = 1$.

Допоміжні факти. Нехай $g\colon H \to \mathbb{R} \cup \{+\infty\}$ — власний опуклий напівнеперервний знизу функціонал. Проксимальним оператором, асоційованим з g, називають оператор $H \ni \ni x \mapsto \mathtt{prox}_g x = \operatorname*{argmin}_{y \in H} \left\{ g(y) + \frac{1}{2}\|y - x\|^2 \right\}$.

Для доведення збіжності алгоритму будемо використовувати такі факти.

Лема 2. *Нехай обмежена знизу послідовність* (a_n) *та послідовності невід'ємних чисел* (b_n) *і* (c_n) *такі, що* $a_{n+1} - a_n + b_n \leqslant c_n$ $(n \in \mathbb{N})$, $\sum_{n=1}^{\infty} c_n < +\infty$. *Тоді існує* $\lim_{n\to\infty} a_n \in \mathbb{R}$ *і* $\sum_{n=1}^{\infty} b_n < +\infty$.

Лема 3 (Z. Opial, [13]). *Нехай* H *— гільбертів простір;* $F \subseteq H$ *— непорожня множина;* (x_n) *— послідовність точок* H. *Припустимо, що:*

1) усі слабкі часткові границі послідовності (x_n) *належать* F*;*

2) для всіх $y \in F$ *існує* $\lim_{n\to\infty} \|x_n - y\| \in \mathbb{R}$.

Тоді (x_n) *слабко збігається до деякої точки* $x \in F$.

Лема 4 (G. B. Passty, [5]). *Нехай* H *— гільбертів простір;* $F \subseteq H$ *— непорожня множина;* (x_n) *— послідовність точок* H *і* $\overline{x}_n = \sum_{k=1}^{n} \lambda_k x_k \big/ \sum_{k=1}^{n} \lambda_k$, *де* (λ_n) *— послідовність додатних чисел, така, що* $\sum_{n=1}^{\infty} \lambda_n = +\infty$. *Припустимо, що:*

1) усі слабкі часткові границі послідовності $(\overline{x}_n)$ *належать* F*;*

2) для всіх $y \in F$ *існує* $\lim_{n\to\infty} \|x_n - y\| \in \mathbb{R}$.

Тоді $(\overline{x}_n)$ *слабко збігається до деякої точки* $x \in F$.

Зауваження 2. Леми 3 та 4 дозволяють доводити слабку збіжність послідовностей без апріорного знання границі.

Альтернуючий проксимальний алгоритм. Нехай (λ_n), (α_n) — послідовності додатних чисел.

Алгоритм 1. Обираємо $x_1 \in H$ та генеруємо послідовність елементів (x_n) за допомогою ітераційної схеми

$$\begin{cases} y_n = \mathtt{prox}_{\lambda_n f_2} x_n, \\ x_{n+1} = \mathtt{prox}_{\lambda_n \alpha_n f_1} y_n. \end{cases}$$

Для породжених алгоритмом 1 послідовностей (x_n) та (y_n) мають місце такі твердження.

Лема 5. *Нехай* $z \in C$, $v \in \partial f_2(z) + N_C z$, *а точки* $w^* \in \partial f_2(z)$, $w^{**} \in N_C z$, *такі, що* $v = w^* + w^{**}$. *Тоді виконується нерівність*

$$\|x_{n+1} - z\|^2 - \|x_n - z\|^2 + \|x_n - y_n\|^2 + \frac{1}{2}\|x_{n+1} - y_n\|^2 + \lambda_n \alpha_n f_1(x_{n+1}) \leqslant$$

$$\leqslant 2\lambda_n (v, z - x_{n+1}) + \frac{\lambda_n}{\alpha_n} \frac{1}{k} \|w^{**}\|^2 + 2\lambda_n^2 \|w^*\|^2. \qquad (2)$$

Доведення. Оскільки

$$\begin{cases} y_n = \operatorname*{argmin}_{y \in H} \left\{ \lambda_n f_2(y) + \dfrac{1}{2}\|y - x_n\|^2 \right\}, \\ x_{n+1} = \operatorname*{argmin}_{y \in H} \left\{ \lambda_n \alpha_n f_1(y) + \dfrac{1}{2}\|y - y_n\|^2 \right\}, \end{cases}$$

то $x_n - y_n \in \lambda_n \partial f_2(y_n)$, $y_n - x_{n+1} \in \lambda_n \alpha_n \partial f_1(x_{n+1})$. З монотонності оператора ∂f_2 випливає оцінка

$$\|x_n - z\|^2 - \|y_n - z\|^2 - \|x_n - y_n\|^2 = 2(x_n - y_n, y_n - z) \geqslant 2\lambda_n (w^*, y_n - z). \tag{3}$$

З нерівності

$$-f_1(x_{n+1}) = f_1(z) - f_1(x_{n+1}) \geqslant \frac{1}{\lambda_n \alpha_n}(y_n - x_{n+1}, z - x_{n+1})$$

випливає оцінка

$$\|y_n - z\|^2 - \|x_{n+1} - z\|^2 - \|x_{n+1} - y_n\|^2 = 2(y_n - x_{n+1}, x_{n+1} - z) \geqslant 2\lambda_n \alpha_n f_1(x_{n+1}). \tag{4}$$

Додавши (3) до (4), одержимо

$$\|x_n - z\|^2 - \|x_{n+1} - z\|^2 - \|x_n - y_n\|^2 - \|x_{n+1} - y_n\|^2 \geqslant$$
$$\geqslant 2\lambda_n \alpha_n f_1(x_{n+1}) + 2\lambda_n (w^*, y_n - z).$$

Маємо

$$2\lambda_n (w^*, y_n - z) = 2\lambda_n (w^*, x_{n+1} - z) + 2\lambda_n (w^*, y_n - x_{n+1}).$$

Оскільки

$$2\lambda_n (w^*, y_n - x_{n+1}) = (2\lambda_n w^*, y_n - x_{n+1}) \geqslant -2\lambda_n^2 \|w^*\|^2 - \frac{1}{2}\|y_n - x_{n+1}\|^2,$$

то

$$\|x_n - z\|^2 - \|x_{n+1} - z\|^2 - \|x_n - y_n\|^2 - \frac{1}{2}\|x_{n+1} - y_n\|^2 \geqslant$$
$$\geqslant 2\lambda_n \alpha_n f_1(x_{n+1}) + 2\lambda_n (w^*, x_{n+1} - z) - 2\lambda_n^2 \|w^*\|^2. \tag{5}$$

Нерівність (5) перепишимо у вигляді (використали рівність $v = w^* + w^{**}$)

$$\|x_{n+1} - z\|^2 - \|x_n - z\|^2 + \|x_n - y_n\|^2 + \frac{1}{2}\|x_{n+1} - y_n\|^2 + \lambda_n \alpha_n f_1(x_{n+1}) \leqslant$$
$$\leqslant -\lambda_n \alpha_n f_1(x_{n+1}) + 2\lambda_n (w^*, z - x_{n+1}) + 2\lambda_n^2 \|w^*\|^2 =$$
$$= 2\lambda_n (v, z - x_{n+1}) + \lambda_n \alpha_n \left\{ \left(\frac{2w^{**}}{\alpha_n}, x_{n+1} \right) - f_1(x_{n+1}) - \left(\frac{2w^{**}}{\alpha_n}, z \right) \right\} + 2\lambda_n^2 \|w^*\|^2.$$

Враховуючи лему 1, отримуємо нерівність (2).

Лема 6. *Для точок z, v, w^* та w^{**} з леми 5 виконується нерівність*

$$-\frac{\|x_1 - z\|^2}{2\sum\limits_{i=1}^{n}\lambda_i} \leqslant (v, z - \overline{x}_{n+1}) + \frac{\left(\sum\limits_{i=1}^{n}\frac{\lambda_i}{\alpha_i}\frac{1}{2k}\|w^{**}\|^2 + \sum\limits_{i=1}^{n}\lambda_i^2\|w^*\|^2\right)}{\sum\limits_{i=1}^{n}\lambda_i},$$

де $\overline{x}_n = \sum\limits_{i=1}^{n-1}\lambda_i x_{i+1}\Big/\sum\limits_{i=1}^{n-1}\lambda_i$.

Лема 7. *Нехай функціонал f_2 сильно опуклий з сталою $c > 0$. Тоді для єдиного розв'язку задачі* (1) $z \in H$ *виконується нерівність*

$$2c\lambda_n\|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 - \|x_{n+1} - z\|^2 + \left(\frac{\lambda_n}{\alpha_n}\frac{1}{k} + 2\lambda_n^2\right)\|w^*\|^2, \tag{6}$$

де $w^* \in -\partial f_2(z)\bigcap N_C z$.

Доведення. Сильна монотонність оператора ∂f_2 [11] замість (3) дає оцінку

$$\|x_n - z\|^2 - \|y_n - z\|^2 - \|x_n - y_n\|^2 = 2(x_n - y_n, y_n - z) \geqslant$$
$$\geqslant 2\lambda_n(-w^*, y_n - z) + 2c\lambda_n\|y_n - z\|^2.$$

З нерівності (4) випливає

$$\|x_{n+1} - z\|^2 \leqslant \|y_n - z\|^2.$$

Тому

$$\|x_n - z\|^2 - \|y_n - z\|^2 - \|x_n - y_n\|^2 \geqslant 2\lambda_n(-w^*, y_n - z) + 2c\lambda_n\|x_{n+1} - z\|^2.$$

Повторивши міркування доведення леми 5, отримаємо нерівність (6).

Теореми збіжності. Стосовно послідовностей (λ_n) та (α_n) зробимо такі припущення:

(A2) $\sum\limits_{n=1}^{\infty}\lambda_n = +\infty$;

(A3) $\sum\limits_{n=1}^{\infty}\frac{\lambda_n}{\alpha_n} < +\infty$;

(A4) $\sum\limits_{n=1}^{\infty}\lambda_n^2 < +\infty$.

Зауваження 3. Наприклад, $\lambda_n = 1/n^p$, $1/2 < p \leqslant 1$, $\alpha_n = n^q$, $q > 1 - p$. Якщо $c_1/\alpha_n \leqslant \lambda_n \leqslant c_2/\alpha_n$, де c_1, $c_2 > 0$, то умова (A3) рівносильна умові (A4).

Лема 8. *Нехай $S \neq \varnothing$, виконуються (A1)–(A4). Тоді*

1) $\forall\, z \in S\ \exists \lim\limits_{n\to\infty}\|x_n - z\| \in \mathbb{R}$;

2) $\sum\limits_{n=1}^{\infty}\|x_n - y_n\|^2 < +\infty$;

3) $\sum\limits_{n=1}^{\infty}\|x_{n+1} - y_n\|^2 < +\infty$;

4) $\displaystyle\sum_{n=1}^{\infty}\lambda_n\alpha_n f_1(x_{n+1}) \ < \ +\infty.$

У випадку сильної опуклості f_2 алгоритм 1 сильно збігається до єдиного розв'язку задачі (1).

Теорема 1. *Нехай функціонал f_2 сильно опуклий, виконуються (A1)–(A4). Тоді породжена алгоритмом 1 послідовність (x_n) сильно збігається до єдиного розв'язку задачі* (1).

Доведення. Нехай z — розв'язок задачі (1), $w^* \in -\partial f_2(z)\bigcap N_C z$, $c > 0$ — стала сильної опуклості f_2. За лемою 7, для $n = \overline{1, N}$ маємо

$$2c\lambda_n\|x_{n+1}-z\|^2 \leqslant \|x_n - z\|^2 - \|x_{n+1}-z\|^2 + \left(\frac{\lambda_n}{\alpha_n}\frac{1}{k} + 2\lambda_n^2\right)\|w^*\|^2.$$

Просумувавши нерівності, одержимо

$$2c\sum_{n=1}^{N}\lambda_n\|x_{n+1}-z\|^2 \leqslant \|x_1 - z\|^2 - \|x_{N+1}-z\|^2 + \left(\frac{1}{k}\sum_{n=1}^{N}\frac{\lambda_n}{\alpha_n} + 2\sum_{n=1}^{N}\lambda_n^2\right)\|w^*\|^2 \leqslant$$

$$\leqslant \|x_1 - z\|^2 + \left(\frac{1}{k}\sum_{n=1}^{N}\frac{\lambda_n}{\alpha_n} + 2\sum_{n=1}^{N}\lambda_n^2\right)\|w^*\|^2.$$

Після граничного переходу при $N \to \infty$ маємо

$$2c\sum_{n=1}^{\infty}\lambda_n\|x_{n+1}-z\|^2 \leqslant \|x_1 - z\|^2 + \left(\frac{1}{k}\sum_{n=1}^{\infty}\frac{\lambda_n}{\alpha_n} + 2\sum_{n=1}^{\infty}\lambda_n^2\right)\|w^*\|^2 < +\infty.$$

З умови (A2) та існування $\lim\limits_{n\to\infty}\|x_n - z\|$ випливає $\lim\limits_{n\to\infty}\|x_n - z\| = 0$.

Для не сильно опуклих функціоналів f_2 встановлено факт слабкої збіжності алгоритму 1.

З лем 4, 6, 8 та міркувань роботи [5] випливає слабка збіжність чезарівських середніх послідовності (x_n).

Теорема 2. *Нехай виконуються умови (A1)–(A4). Тоді справедливі твердження:*

1) якщо $S \neq \varnothing$, то послідовність чезарівських середніх $(\overline{x}_n)$ слабко збігається до точки з множини S;

2) якщо $S = \varnothing$, то $\|\overline{x}_n\| \to +\infty$.

Більш тонкий аналіз дозволяє довести слабку збіжність послідовності (x_n). На основі леми 3 та твердження 1 леми 8 для цього досить показати, що всі слабкі часткові границі послідовності (x_n) належать множині S. Оскільки функціонали f_1 та f_2 слабко напівнеперервні знизу, $\lim\limits_{n\to\infty}\|x_{n+1} - y_n\| = 0$, то останнє випливатиме з такої асимптотичної поведінки числових послідовностей $(f_1(x_n))$, $(f_2(y_n))$:

$$\lim_{n\to\infty} f_1(x_n) = 0,$$

$$\limsup_{n\to\infty} f_2(y_n) \leqslant f_2(z) \qquad \forall\, z \in S.$$

Має місце

Теорема 3. *Нехай $S \neq \varnothing$, виконуються умови (A1)–(A4) та*

$$\liminf_{n\to\infty}\alpha_n\lambda_n > 0.$$

Тоді породжена алгоритмом 1 послідовність (x_n) слабко збігається до розв'язку задачі (1).

1. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. – Москва: Наука, 1979. – 288 с.
2. *Еремин И. И.* О задачах последовательного программирования // Сиб. мат. журн. – 1973. – **14**, № 1. – С. 53–63.
3. *Подиновский В. В., Гаврилов В. Н.* Оптимизация по последовательно применяемым критериям. – Москва: Сов. радио, 1975. – 192 с.
4. *Лионс Ж.-Л.* Оптимальное управление системами, описываемыми уравнениями с частными производными. – Москва: Мир, 1972. – 414 с.
5. *Passty G. B.* Ergodic convergence to a zero of the sum of monotone operators in Hilbert spaces // J. Math. Anal. Appl. – 1979. – **72**. – P. 383–390.
6. *Attouch H., Redont P., Soubeyran A.* A new class of alternating proximal minimization algorithms with costs-to-move // SIAM J. Optim. – 2007. – **18**, No 3. – P. 1061–1081.
7. *Семенов В. В.* О методе параллельной проксимальной декомпозиции для решения задач выпуклой оптимизации // Пробл. управления и информатики. – 2010. – № 2. – С. 42–46.
8. *Войтова Т. А., Семенов В. В.* Метод решения двухэтапных операторных включений // Журн. обчислюв. та прикл. математики. – 2010. – № 3 (102). – С. 34–39.
9. *Денисов С. В.* Параллельная схема декомпозиции для поиска седловой точки и равновесия Нэша // Там само. – 2010. – № 3 (102). – С. 40–48.
10. *Обен Ж.-П., Экланд И.* Прикладной нелинейный анализ. – Москва: Мир, 1988. – 510 с.
11. *Гольштейн Е. Г., Третьяков Н. В.* Модифицированные функции Лагранжа. Теория и методы оптимизации. – Москва: Наука, 1989. – 400 с.
12. *Магарил-Ильяев Г. Г., Тихомиров В. М.* Выпуклый анализ и его приложения. – Москва: ЛИБРОКОМ, 2011. – 176 с.
13. *Opial Z.* Weak convergence of the sequence of successive approximations for nonexpansive mappings // Bull. Amer. Math. Soc. – 1967. – **73**. – P. 591–597.

Київський національний університет
ім. Тараса Шевченка

Надійшло до редакції 30.06.2011

Т. А. Войтова, С. В. Денисов, В. В. Семенов

Альтернирующий проксимальный алгоритм для задачи двухуровневой выпуклой минимизации

Рассмотрен вопрос решения двухуровневой выпуклой задачи минимизации при помощи альтернирующего проксимального алгоритма. При некоторых метрических условиях на функционал задачи первого уровня доказаны теоремы сильной и слабой сходимости.

T. A. Voitova, S. V. Denisov, V. V. Semenov

Alternating proximal algorithm for the problem of bilevel convex minimization

We consider a solution of the bilevel convex minimization problem by the alternating proximal algorithm. Under certain metric conditions for the functional of the first-level problem, the strong and weak convergence theorems are proved.