УДК 518:517.948

М.Я.Бартіш, Ю.В.Нікольський

ЧИСЕЛЬНИЙ МЕТОД МІНІМІЗАЦІЇ ФУНКЦІЙ З ПОБУДОВОЮ ДВОЇСТИХ БАЗИСІВ

Дослідимо один чисельний метод у роботі [1] для мінімізації достатньо гладкої сильно випуклої функції $f(x)$, визначеної на всьому просторі E_n. Він є модифікацією відомого методу двоїстих напрямків [2]. Обчислювальні формули мають вигляд

$$\tilde{x}_k = x_k - \tilde{\alpha}_k A_{k-1}^{-1} f'(x_k),$$

$$x_{k+1} = x_k - \alpha_k A_k^{-1} f'(x_k), \quad k=0,1,\ldots \qquad /1/$$

Вибір скалярних множників α_k та $\tilde{\alpha}_k$ здійснюється таким чином, щоб задовольнити виконання відповідних нерівностей

$$f(\tilde{x}_k) - f(x_k) \leq \varepsilon \tilde{\alpha}_k (f'(x_k), \tilde{p}_k), \quad \tilde{p}_k = -A_{k-1}^{-1} f'(x_k),$$

$$f(x_{k+1}) - f(x_k) \leq \varepsilon \alpha_k (f'(x_k), p_k), \quad p_k = -A_k^{-1} f'(x_k), \qquad /2/$$

де $\varepsilon \in (0, 1/2)$.

Мета досліджень – уточнення оцінок швидкості збіжності методу /1/–/2/ та побудова формул, які дають змогу обчислювати матрицю, обернену до A_k, використовуючи матриці A_{k-1}^{-1} та визначені на кожній ітерації вектори ε, ρ.

Матрицю A_k знаходимо за допомогою системи рівнянь

$$A_k \rho_{k-i} = \varepsilon_{k-i}, \quad i=0,1,\ldots,n-1, \qquad /3/$$

де

$\rho_{k-i} = x_{k-0,5i} - x_{k-0,5i-1}$, $i=0,2,\ldots,n-2$ при n парному або

$\varepsilon_{k-i} = f'(x_{k-0,5i}) - f'(x_{k-0,5i-1})$, $i=0,2,\ldots,n-1$ при n непарному;

$\rho_{k-i} = \tilde{x}_{k-0,5(i-1)} - x_{k-0,5(i-1)}$, $i=1,3,\ldots,n-1$ при n парному або

$\varepsilon_{k-i} = f'(\tilde{x}_{k-0,5(i-1)}) - f'(x_{k-0,5(i-1)})$, $i=1,3,\ldots,n-2$ при n непарному. /4/

Нескладно переконатись, що при виборі параметрів α_k та $\tilde{\alpha}_k$ з умови /2/, починаючи з деякого номеру N, для всіх $m>N$

правильне $\alpha_m = \tilde{\alpha}_m \equiv 1$. Надалі розглядатимемо послідовності точок $\{x_k\}$ та $\{\tilde{x}_k\}$, для яких виконано такий вибір скалярних множників.

Аналогічно в лемі 3.1 з праці [2] у випадку, коли послідовність векторів $\{p_k\}$ вибрана згідно з рівністю /3/, при виконанні умов

$$m_0 \|y\|^2 \leq (f''(x)y, y) \leq M_0 \|y\|^2, \; m_0 \leq M_0, \qquad /5/$$

$$(A_k^{-1} f'(x_k), f'(x_k)) > 0, \; (A_{k-1}^{-1} f'(x_k), f'(x_k)) > 0 \qquad /6/$$

виконуватиметься $\lim\limits_{k \to \infty} \|A_k - f''(x_k)\| = 0$;

при виборі параметрів α_k та $\tilde{\alpha}_k$ зі співвідношення /2/ незалежно від вибору початкової точки x_0 для послідовності $\{x_k\}$ правильні твердження

$$f(x_{k+1}) < f(x_k), \; \|x_{k+1} - x_k\| \to 0,$$

$$f(\tilde{x}_k) < f(x_k), \; \|\tilde{x}_k - x_k\| \to 0$$

чинні такі оцінки:

$$\|A_k - f''(x_k)\| \leq C \|x_{k-0,5(n-2)} - x_*\| \text{ при } n \text{ парному},$$

$$\|A_k - f''(x_k)\| \leq C \|x_{k-0,5(n-1)} - x_*\| \text{ при } n \text{ непарному}. \qquad /7/$$

Тут і далі C позначає будь-яку незалежну від номеру ітерації константу. Оцінки /7/ одержують так, як і в праці [3]. Співвідношення /7/ дають змогу довести таку теорему.

<u>Теорема.</u> Якщо $f(x)$ - двічі неперервно-диференційована функція, для якої правильна умова /5/, матриця A_k для будь-якого $k \geq n-1$ визначається системою /3/-/4/ та задовольняє нерівності /6/, скалярні множники $\alpha_k, \tilde{\alpha}_k$ шукають згідно з /2/, то швидкість збіжності надлінійна та визначена оцінками:

$$\|x_{k+1} - x_*\| \leq C \|x_k - x_*\| \|x_{k-0,5n} - x_*\| \|x_{k-0,5(n-2)} - x_*\| \text{ для } n \text{ парного},$$

$$\|x_{k+1} - x_*\| \leq C \|x_k - x_*\| \|x_{k-0,5(n+1)} - x_*\| \|x_{k-0,5(n-1)} - x_*\| \text{ для } n \text{ непарного}.$$

Доведення. Спосіб побудови матриці A_K дає змогу переконатися у справедливості такого співвідношення:

$$A_K P_{K-1} = \mathcal{E}_{K-1},$$

звідси

$$x_K - A_K^{-1} f'(x_K) = \tilde{x}_K - A_K^{-1} f'(\tilde{x}_K),$$

а друга рівність в /I/ має вигляд

$$x_{K+1} = \tilde{x}_K - A_K^{-1} f'(\tilde{x}_K).$$

Тоді

$$\| x_{K+1} - x_* \| \le \| A_K^{-1} \| \| A_K - f''(x_{KC}) \| \| \tilde{x}_K - x_* \|,$$

$$x_{KC} = \tilde{x}_K + \theta_1 (\tilde{x}_K - x_*), \quad \theta_1 \in [0,1].$$

Оцінку норми $\| \tilde{x}_K - x_* \|$ можна одержати аналогічно

$$\| \tilde{x}_K - x_* \| \le \| A_{K-1}^{-1} \| \| A_{K-1} - f''(x_{KC1}) \| \| x_K - x_* \|,$$

$$x_{KC1} = x_K + \theta_2 (x_K - x_*), \quad \theta_2 \in [0,1].$$

Останні співвідношення дають змогу одержати оцінку швидкості /8/.

Оскільки система /3/ зображається у вигляді

$$A_K^{-1} = R_K E_K^{-1},$$

причому матриці R_{K+1} та E_{K+1} дістають з R_K та E_K відновленням двох стовпців у кожній з них, то нову матрицю E_{K+1}^{-1} можна записати за допомогою рекурентних співвідношень [4]. Нехай матриця E_K^{-1} вже побудована, тобто наявний базис $S_K, S_{K-1}, \dots, S_{K-n+1}$, який є рядками цієї матриці. Тоді система векторів $\bar{S}_{K+2}, \bar{S}_{K+1}, \dots, \bar{S}_{K-n+3}$, двоїстих до базису матриці E_{K+1}, будують таким чином:

$$\bar{S}_{K+2} = \frac{(S_{K-n+1}, \mathcal{E}_{K+1}) S_{K-n+2} - (S_{K-n+2}, \mathcal{E}_{K+1}) S_{K-n+1}}{\Delta},$$

$$\bar{S}_{K+1} = \frac{(S_{K-n+2}, \mathcal{E}_{K+2}) S_{K-n+1} - (S_{K-n+1}, \mathcal{E}_{K+2}) S_{K-n+2}}{\Delta},$$

$$\Delta = (S_{K-n+1}, \mathcal{E}_{K+1})(S_{K-n+2}, \mathcal{E}_{K+2}) - (S_{K-n+2}, \mathcal{E}_{K+1})(\mathcal{E}_{K-n+1}, \mathcal{E}_{K+2}),$$

$$\bar{S}_{K+1-j} = S_{K+1-j} - [(S_{K+1-j}, \mathcal{E}_{K+1}) \bar{S}_{K+1} + (S_{K+1-j}, \mathcal{E}_{K+2}) \bar{S}_{K+2}], \quad j = \overline{1, n-2},$$

$$\bar{S}_{K+2-j} = S_{K+2-j} - [(S_{K+2-j}, \mathcal{E}_{K+1}) \bar{S}_{K+1} + (S_{K+2-j}, \mathcal{E}_{K+2}) \bar{S}_{K+2}], \quad j = \overline{2, n-1}.$$

З огляду на двоїстість базисів $\mathcal{E}_k, \mathcal{E}_{k-1}, \ldots, \mathcal{E}_{k-n+1}$ та $S_k, S_{k-1}, \ldots, S_{k-n+1}$ правильні співвідношення

$$(\bar{S}_{k+1-j}, \mathcal{E}_{k+2-m}) = \delta_{jm} \quad j = 0,1,\ldots,n-1 \quad m = 1,2,\ldots,n,$$

$$(\bar{S}_{k+2-j}, \mathcal{E}_{k+2-m}) = \delta_{jm} \quad j, m = 0,1,\ldots,n-1,$$

де δ_{jm} – символ Кронекера $(\delta_{jj} = 1, \delta_{jm} = 0,\ j \neq m)$.

У цьому випадку для визначення напрямків ρ_k та $\tilde{\rho}_k$ беруть формули

$$\rho_k = -\sum_{i=0}^{n-1} \rho_{k-i} S^*_{k-i} f'(x_k) = -\sum_{i=0}^{n-1} (S_{k-i}, f'(x_k))\rho_{k-i},$$

$$\tilde{\rho}_k = -\sum_{i=0}^{n-1} \rho_{k-i-2} S^*_{k-i-2} f'(x_k) = -\sum_{i=0}^{n-1} (S_{k-i-2}, f'(x_k))\rho_{k-i-2}.$$

Список літератури: І. Бартиш М.Я., Никольский Ю.В. Метод двойственных направлений с ускоренной сходимостью. – В кн.: Оптимальное управление в механических системах. К., 1979, т.І, с.60. 2. Данилин Ю.М., Пшеничный Б.Н. О методах минимизации с ускоренной сходимостью. Журн. вычислительной мат. и мат. физики. 1970, 10, № 6, с. 1341–1354. 3. Пшеничный Б.Н., Данилин Ю.М. Численные методы в экстремальных задачах. – М.: Наука, 1975. 4. Фадеев Д.К., Фадеева В.Н. Вычислительные методы линейной алгебры. – М.: Физматгиз, 1970. – 656 с.

Стаття надійшла до редколегії 14.04.82