ми науково пояснити причини ефективності математики в природничих науках, буржуазні філософи заявляють, що неймовірна ефективність математики у природничих науках є дещо, що межує з містикою, бо ніякого раціонального пояснення цьому факту немає [2, с.23].

Діалектико-матеріалістичний підхід до аналізу процесу математизації науки має неоціненне значення для правильного розуміння закономірностей сучасної науково-технічної революції, для критики платоністських, неопозитивістських та інших ідеалістичних інтерпретацій процесу математизації і математичного знання.

Список літератури: І. Л е н і н В.І. Повне зібрання творів. 2. В и г н е р Е. Непостижимая эффективность математики в естественных науках. - В кн.: Проблемы современной математики. М., 1971. 3. К а р н а п Р. Философские основания физики. Введение в философские науки. - М., 1971. 4. С о й е р М. Путь в современную математику.-М., 1972. 5. Ш и х а н о в и ч Ю.А. Введение в современную математику. - М. , 1965. 6. Э й н ш т е й н А. Собрание научных трудов, т.IУ. - М., 1967.

УДК 518:517.948

М.Я.Бартіш, Ю.В.Нікольський

ГРАДІЄНТНО-ПАРАМЕТРИЧНИЙ МЕТОД МІНІМІЗАЦІЇ ФУНКЦІЙ

Розглянемо задачу мінімізації визначеної на множині $\Omega=\{x: f(x)\leq f(x_0)\}$ двічі неперервно диференційованої та достатньо гладкої на цій множині функції $f(x)$, ℓ змінних.

Одним з найбільш поширених методів чисельного розв"язання цієї задачі є метод найшвидшого спуску. Він має лінійну швидкість збіжності, яка низька при мінімізації функцій з погано зумовленою матрицею других похідних.

Ми пропонуємо метод, який має швидкість збіжності вищу, ніж метод найшвидшого спуску, причому прискорення досягається шляхом вибору параметра θ .

Розглянемо метод

$$\bar{x}_n = x_n - \beta_n \nabla f(x_n),$$

$$s_n = \theta \frac{\nabla f(x_n)}{\|\nabla f(x_n)\|} + (1-\theta)\frac{\nabla f(\bar{x}_n)}{\|\nabla f(\bar{x}_n)\|}, \quad 0 < \theta \leq 1, \qquad /I/$$

$$x_{n+1} = x_n - \alpha_n s_n, \quad n = 0,1,2\ldots,$$

де α_n, β_n – скалярні множники, а β_n вибирається з умови

$$f(x_n - \beta_n \nabla f(x_n)) = \min_{\beta} f(x_n - \beta \nabla f(x_n)). \qquad /2/$$

Неважко помітити, що при $\theta = 1$, метод /I/ перетворюється в метод найшвидшого спуску.

Нехай функція $f(x)$ – сильно випукла і має обмежений спектр гессіана, тобто виконуються умови

$$m\|y\|^2 \leq (\nabla^2 f(x)y, y) \leq M\|y\|^2, \quad 0 < m \leq M, \ \forall x, y \in \Omega. \qquad /3/$$

Крім того, з /2/ випливає $(\nabla f(x_n), \nabla f(\bar{x}_n)) = 0$. /4/

Враховуючи /3/, /4/, можна довести, як і в праці [I] збіжність послідовностей $\{x_n\}$ та $\{f(x_n)\}$ відповідно до x_* та $f(x_*)$ для довільного початкового наближення x_0, де x_* – точка мінімума.

Розглянемо оператор P_n, який здійснює ортогональне проектування на напрямок, визначений вектором $\nabla f(\bar{x}_n)$. Матриця оператора P_n має вигляд

$$P_n = D_n^* (D_n D_n^*)^{-1} D_n, \qquad /5/$$

де D_n – матриця, що складається з одного стовпця – компонент вектора $\frac{\nabla f(\bar{x}_n)}{\|\nabla f(\bar{x}_n)\|}$. Слід зазначити, що власні числа матриці $\lambda_1 = 1, \lambda_2 = \ldots = \lambda_\ell = 0$ [2].

З врахуванням того, що $(x_n - x_{n-1}, \nabla f(\bar{x}_n)) > 0$, та використовуючи оператор /5/, вираз для s_n набуває вигляду

$$s_n = \theta \frac{\nabla f(x_n)}{\|\nabla f(x_n)\|} + (1-\theta) \frac{P_n(x_n - x_{n-1})}{\|P_n(x_n - x_{n-1})\|}.$$

Тепер метод /I/ можна розглянути як модифікацію k -крокового градієнтного методу [3]. Використовуючи позначення праці [3], записуємо

$$\alpha_0 = 1,$$

$$T_2 = \alpha_n G_2, \quad G_2 = (1-\theta) N P_n,$$

$$T_1 = \alpha_n G_1, \quad G_1 = G_2 + \frac{\theta \nabla^2 f(x_*)}{\|\nabla f(x_n)\|},$$

де $N = \| P_n (x_n - x_{n-1}) \|$.

Метод /I/ в достатньо малому околі розв"язку набуває такого вигляду:

$$y_{n+1} = \tilde{T}_1 y_n + \tilde{T}_2 y_{n-1},$$

де $\tilde{T}_1 = E + T_1; \quad \tilde{T}_2 = T_2; \quad y_n = x_n - x_*;$

E – одинична матриця та враховано, що

$$\nabla f(x_n) - \nabla f(x_*) = \nabla^2 f(x_*)(x_n - x_*) + O(\|x_n - x_*\|^2).$$

Оскільки матриці операторів $G_1, G_2, \nabla^2 f(x)$ симетричні, для них правильні леми I та 2, тобто виконуються умови теореми 7 із праці [3]. Тоді при виконанні умови

$$\sup \max(|\rho_1|, |\rho_2|) \le q < 1, \qquad /6/$$
$$A \in \sigma(G_1)$$
$$B \in \sigma(G_2),$$

де ρ_1, ρ_2 – корені рівняння

$$\rho^2 = (1 - \alpha_n A)\rho + \alpha_n B,$$

існує така оцінка швидкості збіжності:

$$\|y_{n+1}\| \le C(\varepsilon)(\|y_0\|^2 + \|y_1\|^2)^{1/2}(q+\varepsilon)^n.$$

Неважко переконатись, що співвідношення /6/ виконується завжди, якщо α_n вибране з умови $0 < \alpha_n < A^{-1}$.

Цікавим вважаємо і такий результат: при виконанні методу /I/ маємо змогу отримати значення q, менше, ніж параметр δ – знаменник геометричної прогресії в оцінці швидкості збіжності методу найскорішого спуску. У цьому випадку α_n вибирається з проміжку $\left(0, \frac{\delta(1-\delta)}{\delta A - B}\right)$, причому досягти виконання цієї умови можна за рахунок підбору такого θ, при якому виконується умова

$$\delta^2 < BA^{-1} < \delta,$$

де

$$B = (1-\theta)N;$$
$$A = B + \theta\frac{M}{\|\nabla f(x_n)\|}.$$

Якщо $x_n \to x_*$, то $\alpha_n \to 0$.

Як відомо, $\delta_{min} = \frac{M-m}{M+m}$, тому застосування методу /I/ доцільне при $M \gg m$.

У таблиці наведені результати застосування методу /I/ до відшукання екстремума таких функцій:

1. $f(x) = 100(x_1^2 - x_2)^2 + (x_1-1)^2 + 90(x_3 - x_4)^2 + 10.1((x_2-1)^2 + (x_4-1)^2) + 19.8(x_2-1)(x_4-1)$ $\quad x_0 = (-1,-1,-1,-1)$.
2. $f(x) = 100(x_1 - x_2^2)^2 + (1-x_2)^2 \quad x_0 = (0.5, -0.5)$.
3. $f(x) = 100(x_1 - x_2^3)^2 + (1-x_2)^2 \quad x_0 = (0.5, -0.5)$.

Функція	θ	Точність		Кількість ітерацій	Кількість обчисл. град.
		за аргументом	за функцією		
I	0,55	10^{-6}	$0{,}6\ 10^{-8}$	716	1432
	0,6	10^{-6}	$0{,}35\ 10^{-8}$	1145	2290
	I	10^{-3}	$0{,}3\ 10^{-4}$	3001	3001
2	0,55	10^{-4}	$0{,}9\ 10^{-7}$	336	672
	0,6	10^{-4}	$0{,}9\ 10^{-7}$	239	478
	I	10^{-4}	$0{,}7\ 10^{-7}$	1800	1800
3	0,55	10^{-4}	$0{,}8\ 10^{-7}$	889	1778
	0,7	10^{-4}	$0{,}8\ 10^{-8}$	1360	2770
	0,85	10^{-4}	$0{,}16\ 10^{-7}$	881	1762
	I	10^{-3}	10^{-5}	13171	13171

Параметр θ залишався постійним на всіх ітераціях.

Список літератури: I. Пшеничный В.Н., Данилин Ю.М. Численные методы в экстремальных задачах. – М.: Наука, 1975. 2. Гантмахер Е. Теория матриц. – М.; Наука, 1967. 3. Поляк Б.Т. О некоторых способах ускорения сходимости итерационных методов. – ЖВМиМФ, 1964, т.4, № 5.