1. Г и л л Ф., М ю р р е й У., Р а й т М. Практическая оптимизация. М., 1985. 2. Г о л у б Б.М. Одна схема побудови квазіньютонівських алгоритмів для безумовної мінімізації функції // У цьому ж Віснику. 3. Щ е р б и н а Ю.Н., Г о л у б Б.М. Квазиньютоновские модификации метода линеаризации / Львов, 1986. Рукопись деп. в УкрНИИНТИ, № 1396 - Ук86. 4. Щ е р б и н а Ю.М., Г о л у б Б.М. Квазіньютонівська модифікація методу лінеаризації з використанням LDL^T -розкладу Холеського //Вісн. Львів. ун-ту. Сер. мех.-мат. 1988. Вип.29. С.5-8.

Стаття надійшла до редколегії 15.03.88

УДК 519.6

Б.М. Голуб

ОДНА СХЕМА ПОБУДОВИ КВАЗІНЬЮТОНІВСЬКИХ АЛГОРИТМІВ ДЛЯ БЕЗУМОВНОЇ МІНІМІЗАЦІЇ ФУНКЦІЇ

Розглянемо задачу безумовної мінімізації

$$\min f(x), \quad x \in R^n, \qquad /1/$$

де $f(x)$ - неперервно диференційована функція; R^n - n -мірний евклідів простір.

Більшість відомих ітераційних методів розв"язування задачі /1/ можна зобразити у вигляді

$$x_{k+1} = x_k + \alpha_k p_k, \qquad /2/$$

де вектор p_k є напрямком спуску; α_k - додатний кроковий множник.

Позначимо $f'_k = f'(x_k)$, $S_k = x_{k+1} - x_k$, $y_k = f'_{k+1} - f'_k$. Нехай

$$p_k = -A_k^{-1} f'_k, \qquad /3/$$

де A_k - симетрична додатно визначена матриця така, що

$$A_k S_{k-1} = y_{k-1}. \qquad /4/$$

Схема /2/ - /4/ представляє клас квазіньютонівських методів з прямою апроксимацією гесіана функції $f(x)$. Для перерахунку матриць A_k, що задовольняють квазіньютонівську умову /4/, запропоновано багато формул*. Зокрема, однією з найбільш поширених є БФГШ-формула

$$A_{k+1} = A_k + y_k y_k^T / y_k^T S_k - A_k S_k S_k^T A_k / S_k^T A_k S_k, \qquad /5/$$

* Г и л л Ф., М ю р р е й У., Р а й т М. Практическая оптимизация. М., 1985.

де T - знак транспонування.

Для забезпечення додатної визначеності матриці A_{k+1} потрібно достатньо точно розв"язувати задачу одномірної мінімізації.

Опишемо алгоритм, який має глобальну збіжність, не використовує процедур одномірної мінімізації при визначенні кроку α_k і зберігає властивості квазіньютонівських методів.

Нехай $A = LDL^T$, де нижня трикутна одинична матриця L і діагональна матриця D - фактори Холеського матриці A. Матрицю $\bar{A}$ отримують з A модифікацією вигляду $\bar{A} = A + \xi v v^T + E$, де v - вектор, $\xi \in \{-1, 1\}$, а матриця E дорівнює нульовій, якщо матриця $A + \xi v v^T$ суттєво додатно визначена.

Елементи факторів $\bar{L}$ і $\bar{D}$ матриці $\bar{A}$ можна обчислити з допомогою наступних алгоритмів /для зручності запису компоненти векторів і матриць в цих алгоритмах позначатимемо індексами знизу/.

Нехай $\sigma > 0$ - мале число.

Для $\xi = 1$:

1. Прийняти $t_0 = 1$, $v^{(1)} = v$;

2. Для $i = 1, \ldots, n$ знайти

$$z_i = v_i^{(i)}, \quad t_i = t_{i-1} + z_i^2/d_i, \quad \bar{d}_i = d_i t_i / t_{i-1}, \quad \beta_i = z_i/(d_i t_i),$$

$$v_j^{(i+1)} = v_j^{(i)} - z_i \ell_{ji}, \quad \bar{\ell}_{ji} = \ell_{ji} + \beta_i v_j^{(i+1)}, \quad j = \overline{i+1, n}.$$

Для $\xi = -1$.

1. Розв"язати рівняння $Lz = v$ і прийняти

$$t_{n+1} = max\{\sigma, 1 - z^T D^{-1} z\}.$$

2. Для $i = n, \ldots, 1$ обчислити

$$t_i = t_{i+1} + z_i^2/d_i, \quad \bar{d}_i = max\{\sigma, d_i t_{i+1}/t_i\}, \quad \beta_i = -z_i/(d_i t_{i+1}),$$

$$v_i^{(i)} = z_i, \quad \bar{\ell}_{ji} = \ell_{ji} + \beta_i v_j^{(i+1)}, \quad v_j^{(i)} = v_j^{(i+1)} + z_i \ell_{ji}, \quad j = \overline{i+1, n}.$$

Використовуючи наведені алгоритми, /5/ можна замінити формулою

$$L_{k+1} D_{k+1} L_{k+1}^T = (L_k D_k L_k^T + sign(S_k^T y_k) v_k v_k^T) - w_k w_k^T + E_{k+1}, \qquad /6/$$

де $v_k = y_k / |S_k^T y_k|^{1/2}$, $w_k = -f_k' / |p_k^T f_k'|^{1/2}$

/тут враховано, що $S_k = \alpha_k p_k$, $A_k S_k = -\alpha_k f_k'$ /.

Сформулюємо тепер алгоритм розв"язування задачі /1/.

Нехай підібрано числа $S>1$, $\sigma>0$, $C_0>0$, $0<\gamma<1$, $0<\varepsilon<1/2$ і вектор x_0. Приймемо $L_0=I, D_0=I$ /I - одинична матриця/.

Опишемо загальний крок алгоритму. Нехай точка x_k, число C_k, матриці L_k і D_k вже обчислені.

1. Розв"язати систему рівнянь $L_k D_k L_k^T p_k = -f_k'$.

2. Якщо $\| p_k \| > C_k$ або $f(x_k+p_k) > f(x_0)$, то прийняти $C_{k+1}=C_k$ і перейти до кроку 3. Інакше прийняти $x_{k+1}=x_k+p_k$, $C_{k+1}=\gamma \| p_k \|$ і перейти до кроку 4.

3. Починаючи з $\alpha=1$, дробити α шляхом поділу навпіл до першого виконання нерівності

$$f(x_k+\alpha p_k) \leqslant f(x_k)+\alpha \varepsilon p_k^T f_k'.$$

Прийняти $x_{k+1}=x_k+\alpha_k p_k$.

4. Перерахувати матриці L_{k+1} і D_{k+1} за формулою /6/ і обчислити діагональні елементи матриці $A_{k+1}=L_{k+1} D_{k+1} L_{k+1}^T$ за формулою

$$a_{k+1}^i = d_{k+1}^i + \sum_{j=1}^{i-1} (l_{k+1}^{ij})^2 d_{k+1}^j, \quad i=\overline{1,n}.$$

5. Якщо $\max_i a_{k+1}^i / \min_i d_{k+1}^i \leqslant S$, то перейти до кроку 1. Інакше прийняти $L_{k+1}=I$, $D_{k+1}=I$ і перейти до кроку 1.

Процес обчислень припинити, якщо $\| f_k' \| \leqslant \omega$, де ω - задана точність.

<u>Теорема</u>. Нехай множина $\Omega=\{x: f(x) \leqslant f(x_0)\}$ компактна і градієнт $f'(x)$ в Ω задовольняє умову Ліпшиця. Тоді в довільній граничній точці x_* послідовності $\{x_k\}$ виконується рівність $f'(x_*)=0$.

Коли, крім того, $f \in C^2$, x_* - єдина точка, в якій виконується необхідна умова мінімуму, матриця $f''(x_*)$ невироджена і

$$\| [A_k - f''(x_*)] p_k \| / \| p_k \| \to 0 \quad \text{при } k \to \infty,$$

то послідовність $\{x_k\}$ збігається до x_* з надлінійною швидкістю.

Стаття надійшла до редколегії 15.03.88