ние для оптимальной матрицы-регуляризатора в методе наименьших квадратов // Специальная радиоелектроника. М., 1993. Вып. 1-3. С. 50-52. 4. Рид М., Саймон Б. Методы современной матфизики. Т1. Функциональный анализ. М.:Мир. 1977. 358 с. 5 Тихонов А.Н., Арсенин В.Я. Методы решений некорректних уравнений. М.: Наука, 1986.

Стаття надійшла до редколегії 09.02.1996

УДК 519.6

Б.М.Голуб

Функції наповнення у глобальній оптимізації

У даній статті розглянуті деякі аспекти побудови функцій наповнення, що використовуються [1,2] в тунельних методах пошуку глобального мінімуму ліпшицевої функції $f(x)$ на компактній множині $X \subset R^n$, де R^n — n-вимірний евклідів простір.

1. Означення. Нехай $x_* \in X$ — довільна точка строгого мінімуму функції $f(x)$. Позначимо

$$A(x_*) = \{x \in X : f(x) > f(x_*)\},$$

$$B(x_*) = \{x \in X : f(x) \le f(x_*)\} \setminus \{x_*\},$$

$$B_\varepsilon(x_*) = \{x \in X: \ f(x) \le f(x_*) - \varepsilon\}, \qquad \varepsilon > 0,$$

$$H(y) = \{h \in R^n : y + \alpha h \in X, \ \alpha > 0\}, \quad y \in X.$$

Якщо L — константа Ліпшиця для функції $f(x)$, то для довільних $x \in X$, $h \in R^n$ та $\alpha > 0$ таких, що $x + \alpha h \in X$, виконується нерівність $|f(x+\alpha h) - f(x)| \le \alpha L\|h\|$. Звідси випливає існування скінченної похідної функції $f(x)$ за напрямком h у довільній точці $x \in X$:

$$|f'(x;h)| = \lim_{\alpha \to 0} \frac{|f(x+\alpha h) - f(x)|}{\alpha} \le L\|h\|.$$

Лема 1. Для того, щоб точка y була екстремальною для функції $f(x)$, необхідно і достатньо виконання нерівності $f'(y;h_1) \cdot f'(y;h_2) \ge 0$ для довільних $h_1, h_2 \in H(y)$.

Доведення леми безпосередньо випливає з означень екстремальної точки та похідної за напрямком.

Назвемо точку $y \in X$ стаціонарною для функції $f(x)$, якщо *для довільного* $h \in H(y)$ *виконується нерівність*

$$f'(y;h) \cdot f'(y;-h) \geq 0 .$$

Легко бачити, що екстремальними можуть бути лише стаціонарні точки.

Назвемо точку $y \in \operatorname{int} X$ квазістаціонарною для функції $f(x)$, якщо $f'(y;x_* - y) \cdot f'(y;y - x_*) \geq 0$.

Позначимо $Y(f) = \overline{X}(f) \setminus \{x_*\}$, де $\overline{X}(f)$ — множина квазістаціонарних точок функції $f(x)$.

Функцію $F(x)$ будемо називати наповнювальною для $f(x)$ в точці x_*, якщо

1) $F(x)$ неперервна на $X \setminus B_\varepsilon(x_*)$;

2) $Y(F) \subset B(x_*)$;

3) $Y(F) \neq \varnothing$ при $B_\varepsilon(x_*) \neq \varnothing$;

4) x_* — точка екстремуму для функції $F(x)$ на множині X.

В означенні функції наповнення точка x_* є параметром. Для спрощення подальших формул залежність від x_* в означеннях функцій та множин явно вказувати не будемо.

2. Побудова функцій наповнення. Позначимо через $T(f)$ множину функцій наповнення для функції $f(x)$.

Теорема 1. Нехай $F(x) \in T(f)$. Якщо $j(a)$ — строго монотонна неперервно диференційовна функція, визначена на множині значень функції $F(x)$, то $\varphi(F(x)) \in T(f)$.

Доведення. Позначимо $\Phi(x) = \varphi(F(x))$.

Очевидно, що $\Phi(x)$ означена і неперервна на $X \setminus B_\varepsilon(x_*)$.

Розглянемо похідну функції $\Phi(x)$ за напрямком $h \in H(x)$ в точках $x \in X \setminus B_\varepsilon(x_*)$:

$$\Phi'(x;h) = \lim_{\alpha \to 0} \frac{1}{\alpha}\left[\Phi(x + \alpha h) - \Phi(x)\right] =$$
$$= \lim_{\alpha \to 0} \frac{\varphi(F(x + \alpha h)) - \varphi(F(x))}{F(x + \alpha h) - F(x)} \frac{F(x + \alpha h) - F(x)}{\alpha}$$

або

$$\Phi'(x;h) = \frac{\partial \varphi(F(x))}{\partial F(x)} F'(x;h) . \tag{1}$$

З цієї рівності отримаємо

$$\Phi'(x;h)\cdot\Phi'(x;-h)=\left[\frac{\partial\varphi(F(x))}{\partial F(x)}\right]^2 F'(x;h)\cdot F'(x;-h)\,. \qquad (2)$$

З означення функції $\varphi(\alpha)$ випливає, що $\partial\varphi(F(x))/\partial F(x)\neq 0$. Тому з (2) отримаємо: $Y(F)=Y(\Phi)\subset B(x_*)$.

Оскільки x_* — екстремальна точка для функції $F(x)$, то $F'(x_*,h_1)\cdot F'(x_*,h_2)\geq 0$ для довільних $h_1,h_2\in H(x_*)$. Тоді з (1) отримаємо: $\Phi'(x_*,h_1)\cdot\Phi'(x_*,h_2)\geq 0$. Звідси за лемою 1 випливає екстремальність точки x_* для функції $\Phi(x)$, що завершує доведення теореми.

Означимо множини:

$$\overline{B}_\varepsilon(x_*)=\{x\in B_\varepsilon(x_*)\ :\ x+\alpha(x_*-x)\notin B_\varepsilon(x_*),\ 0<\alpha\leq 1\},$$

$$T_+(f)=\{F(x)\in T(f)\ :\ x_*=\arg\max F(x),$$
$$\lim_{\alpha\to 0}F(\overline{x}+\alpha(x_*-\overline{x}))=\infty\ \ \forall\overline{x}\in\overline{B}_\varepsilon(x_*)\}\ ,$$

$$T_-(f)=\{F(x)\in T(f)\ :\ x_*=\arg\min F(x),$$
$$\lim_{\alpha\to 0}F(\overline{x}+\alpha(x_*-\overline{x}))=-\infty\ \ \forall\overline{x}\in\overline{B}_\varepsilon(x_*)\}\ .$$

Очевидно, що $-F(x)\in T_-(f)$, якщо $F(x)\in T_+(f)$.

Теорема 2. Якщо $F_1(x),\ F_2(x)\in T_+(f)$, то для $a>0,\ b>0$

$$F(x)=aF_1(x)+bF_2(x)\in T_+(f)\,.$$

Доведення. Очевидно, що $F(x)$ означена і неперервна на $X\setminus B_\varepsilon(x_*)$.

З означень $T_+(f)$ та функції $F(x)$ випливає, що для довільного $x\in A(x_*)$ буде $F'(x;x-x_*)\cdot F'(x;x_*-x)<0$, тобто $Y(F)\subset B(x_*)$.

Нехай $B_\varepsilon(x_*)\neq\varnothing$. Тоді $\varphi(\alpha)=F(x+\alpha(x_*-x))\to\infty$ при $\alpha\to 0$.

З іншого боку, $j(a)$ зростає і при $\alpha\to 1$, бо x_* — точка максимуму. $j(a)$ неперервна на $(0,1]$, тому існує $\overline{\alpha}=\arg\min\limits_{0<\alpha\leq 1}\varphi(\alpha)$. Тоді за лемою 1 для точки $\overline{x}=x+\overline{\alpha}(x_*-x)$ отримаємо:

$$F'(\overline{x};x_*-\overline{x})\cdot F'(\overline{x};\overline{x}-x_*)\geq 0\,,$$

тобто $\overline{x}\in Y(F)$ і $Y(F)\neq\varnothing$, що завершує доведення теореми.

Наслідок 1. Якщо $F_i(x) \in T_+(f),\ a_i > 0,\ i = 1,\dots,m$, то

$$F(x) = \sum_{i=1}^{m} a_i F_i(x) \in T_+(f).$$

Наслідок 2. Якщо $F_i(x) \in T_-(f),\ a_i > 0,\ i = 1,\dots,m$, то

$$F(x) = \sum_{i=1}^{m} a_i F_i(x) \in T_-(f).$$

Теорема 3. Нехай $F(x) \in T_+(f)$, а $r(x)$ — неперервна функція така, що $0 \le r(x) < r(x + \alpha(x_* - x)) < r(x_*)$ для всіх $x \in X$, $0<\alpha<1$. Тоді $F(x) + r(x) \in T_+(f)$.

Доведення. Очевидно, що $\Phi(x) = F(x) + r(x)$ неперервна на $X \setminus B_\varepsilon(x_*)$, $x_* = \arg\max \Phi(x)$ і $\lim\limits_{\alpha \to 0} \Phi(\bar{x} + \alpha(x_* - \bar{x})) = \infty$ для $\bar{x} \in \bar{B}_\varepsilon(x_*)$.

З означення $r(x)$ легко бачити, що для довільного $x \in \operatorname{int} X$ виконується нерівність $r'(x; x_* - x) \cdot r'(x; x - x_*) \le 0$. Тому $Y(\Phi) \subset B_\varepsilon(x_*)$.

Нерівність $Y(\Phi) \ne \varnothing$ при $B_\varepsilon(x_*) \ne \varnothing$ доводиться аналогічно теоремі 2.

Теореми 1-3 та наслідки до них дають механізм побудови нових функцій наповнення на основі відомих.

Наведемо приклади функцій з множини $T(f)$.

Позначимо

$$S(x_*) = \{x \in X:\ f(x_*) \le f(x_* + \alpha(x - x_*)) \le f(x_* + \beta(x - x_*)) \le f(x)\ \ 0 < \alpha < \beta < 1\},$$

$$D(x_*) = \inf_{x \in X \setminus S(x_*)} \|x - x_*\|.$$

Легко бачити, що $S(x_*) \ne \varnothing$ і $D(x_*) > 0$.

Теорема 4. Якщо $\rho^2 < 2\varepsilon D(x_*) / L$, де $\varepsilon > 0$, то

$$F(x) = \|x - x_*\|^2 / \rho^2 + \ln[f(x) - f(x_*) + \varepsilon] \in T_-(f).$$

Доведення. Легко бачити, що функція $F(x)$ означена і обмежена на $A(x_*)$, не має квазістаціонарних точок на множині $S(x_*) \setminus \{x_*\}$ і $x_* = \arg\min F(x)$.

Нехай $B_\varepsilon(x_*) \neq \varnothing$. Тоді $f(x) - f(x_*) + \varepsilon = 0$ для $x \in \overline{B}_\varepsilon(x_*)$ і $F(x + \alpha(x_* - x)) \to -\infty$ при $\alpha \to 0$. Міркування, аналогічні доведенню теореми 2, дають: $Y(F) \neq \varnothing$.

Нехай $x \in A(x_*) \setminus S(x_*)$, $h \in H(x)$. Позначимо

$$g(\alpha) = f(x + \alpha h) - f(x_*) + \varepsilon,$$

де $\alpha > 0$. Очевидно, що $g(\alpha) > 0$ для достатньо малих a.

Розглянемо різницю $F(x + \alpha h) - F(x)$:

$$F(x + \alpha h) - F(x) = \frac{\|x + \alpha h - x_*\|^2 - \|x - x_*\|^2}{\rho^2} +$$

$$+ \frac{g(\alpha) - g(0)}{g(0)} \ln\left(1 + \frac{g(\alpha) - g(0)}{g(0)}\right)^{\frac{g(0)}{g(\alpha) - g(0)}}$$

Розділимо ліву та праву частини на a та спрямуємо a до 0. У результаті отримаємо

$$F'(x;h) = 2\langle x - x_*, h\rangle / \rho^2 + f'(x;h)/(f(x) - f(x_*) + \varepsilon), \tag{3}$$

де $\langle \cdot, \cdot \rangle$ — скалярний добуток векторів.

Беручи до уваги умову теореми, обмеженість $f'(x;h)$ і те, що $f(x) > f(x_*)$, з (3) при $h = x_* - x$ отримаємо:

$$F'(x; x_* - x) < -L\|x - x_*\|^2 / \varepsilon D(x_*) + L\|x - x_*\|/\varepsilon \leq 0.$$

Припустимо, що $F'(x; x - x_*) \leq 0$. Тоді при $h = x - x_*$ з (3) випливає, що

$$2\|x - x_*\|^2 / \rho^2 \leq L\|x - x_*\|/(f(x) - f(x_*) + \varepsilon),$$

або

$$\|x - x_*\| < D(x_*),$$

що суперечить означенню $D(x_*)$.

Отже, $F'(x; x - x_*) \cdot F'(x; x_* - x) < 0$ і точка x не є квазістаціонарною точкою функції $F(x)$. Теорему доведено.

Інші приклади функцій наповнення:

$$F_1(x) = e^{F(x)} = [f(x) - f(x_*) + \varepsilon]\, e^{\|x - x_*\|^2/\rho^2} \in T_-(f),$$

$$F_2(x) = e^{-F(x)} = \frac{1}{f(x) - f(x_*) + \varepsilon}\, e^{-\|x - x_*\|^2/\rho^2} \in T_+(f),$$

$$F_3(x) = F_2(x) + \frac{1}{\|x - x_*\| + \varepsilon} \in T_+(f).$$

Зауважимо, що нові функції набувають, взагалі кажучи, нових властивостей Зокрема, $F_1(x)$ означена і неперервна на X, $F_2(x)$ — від'ємна лише на $B_\varepsilon(x_*)$, $F_3(x)$ — істотно різниться від машинного нуля на $\operatorname{int} X$ тощо.

3. Алгоритм глобальної мінімізації. Нехай x^* — точка глобального мінімуму функції $f(x)$ на X і $f_* = f(x^*)$. Будемо вважати задачу глобальної мінімізації розв'язаною, якщо знайдена точка $x_\varepsilon^* \in X$, така, що $f(x_\varepsilon^*) \le f_* + \varepsilon$.

Загальну схему алгоритму глобальної мінімізації з використанням функції наповнення можна описати таким чином.

Нехай $x_k \in X$.

1. Оберемо x_k за початкову точку і за допомогою деякого методу локального пошуку знайдемо $x_k^* = \arg\min\limits_{x \in X} f(x)$.

2. Застосуємо алгоритм локального пошуку до функції наповнення $F_k(x)$, побудованої для $f(x)$ в точці x_k^*. Якщо всі траєкторії пошуку виходять за межі X, то вважаємо x_k^* точкою глобального мінімуму. У протилежному випадку отримаємо точку $x_{k+1} \in B(x_k^*)$, приймемо $k=k+1$ і повернемося до кроку 1.

Зважаючи на властивості функції наповнення, легко бачити, що за скінченну кількість кроків алгоритм згенерує точку x_k^* таку, що $f(x_k^*) \le f_* + \varepsilon$.

Зазначимо, що наявність точок розриву чи областей неозначеності у наповнювальної функції принципово не утруднює застосування методів локальної оптимізації, оскільки ці точки (області) власне і є шуканими.

1. Голуб Б.М., Оліярник Ю.П. Тунельний алгоритм пошуку глобального мінімуму неперервної функції // Вісн. Львів. ун-ту. Сер. мех.- мат. 1995. Вип.41. С.39-42. 2. Ge R.P. The theory of filled function method for finding global minimizers of nonlinearly constrained minimization problems // J. of Computational Mathe-matics. 1987. Vol.5, N1.

Стаття надійшла до редколегії 04.03.96