EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. XVII KÕIDE
FÜÜSIKA * MATEMAATIKA. 1968, NR. 2

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ XVII
ФИЗИКА * МАТЕМАТИКА. 1968, № 2

https://doi.org/10.3176/phys.math.1968.2.01

С. УЛЬМ, В. ПОЛЛЬ

О НЕКОТОРЫХ МЕТОДАХ РЕШЕНИЯ ЗАДАЧ НА МИНИМУМ

В последнее время был разработан ряд методов для решения нелинейных операторных уравнений, использующих обобщенные разделенные разности (обзор литературы см. в [1]). Но сравнительно мало еще разработаны методы такого типа для решения задач на экстремум. По-видимому, первой в этой области была работа И. В. Шмидта [2], в которой был предложен интерполяционный аналог метода Ньютона для нахождения стационарных точек функций одного переменного, имеющий порядок сходимости 1,32 .... В работах [3, 4] изложены два независимых обобщения последнего метода на функции n переменных, различающиеся друг от друга способом построения интерполяционных полиномов. Если в [4] они были построены с помощью обобщенных разделенных разностей [1], то в [3] с помощью выражений, в некотором смысле «близких» к разделенным разностям. Отметим, что обобщение, предложенное в [4], использует на каждом итерационном шаге n значений функций меньше, чем обобщение в [3], но при этом порядки сходимости методов одинаковы. Кроме того, методы статьи [4] естественным образом обобщаются для минимизации функционалов в абстрактных пространствах.

В настоящей статье делается попытка восполнить пробел в разработке методов вышеуказанного типа для минимизации функционала $f(x)$ в гильбертовом пространстве H. Вначале приводятся определения разделенных разностей первого и второго порядка и интерполяционные формулы Ньютона для функционалов в H (пункт 1). На основании последних в пункте 2 строятся два интерполяционных аналога метода градиентов для нахождения локальной точки безусловного минимума $f(x)$. Один из них (формула (10)) использует вместо градиента $f(x)$ первую разделенную разность, другой же (формула (14)) — градиент интерполяционного полинома (11). В пункте 4 дается для нахождения точек безусловного минимума $f(x)$ метод (формула (38)), совпадающий в случае функций n переменных с одним из методов работы [4]. Основываясь на идее обобщенного метода Стеффенсена [5], в пункте 6 предлагаются две модификации (формула (49) и замечание 5) последнего метода, использующие, кроме разделенных разностей, и градиент $f(x)$, но имеющие порядки сходимости 2. В пунктах 3 и 4 даются на основании соображений статьи [6] обобщения метода хорд и метода (38) для минимизации функционала $f(x)$ при наличии ограничений (соответственно формулы (22) и (37)). Для всех вышеуказанных методов минимизации даются достаточные условия сходимости (теоремы 1—5).

Наконец, в пункте 5 рассматривается приложение метода (37) к задаче оптимального управления (44), а в пункте 7 приводится численный пример применения этого метода для решения одной задачи нелинейного программирования (53).

1. Пусть $f(x)$ — функционал, определенный в гильбертовом пространстве H.

Назовем $f(x'x'')$ разделенной разностью (разностным градиентом) функционала $f(x)$ на элементах $x', x'' \in H$, если выполнены условия

$$1^\circ \ f(x'x'') \in H \ (\text{для каждых фиксированных } x' \text{ и } x''); \tag{1}$$

$$2^\circ \ (f(x'x''), x' - x'') = f(x') - f(x''); \tag{2}$$

$$3^\circ \ f(xx) = G(x) = \operatorname{grad} f(x). \tag{3}$$

Операторы $f_1(x'x''x''')$ и $f_2(x'x''x''')$ называем разделенными разностями второго порядка функционала $f(x)$, если

$$1^\circ \ f_1(x'x''x'''), f_2(x'x''x''') \in (H \to H) \ (\text{для каждых фиксированных } x', x'', x'''); \tag{4}$$

$$2^\circ \ f_1(x'x''x''')(x'' - x''') = f(x'x'') - f(x'x'''); \tag{5}$$

$$f_2(x'x''x''')(x' - x'') = f(x'x''') - f(x''x'''); \tag{6}$$

$$3^\circ \ f_1(xxx) + f_2(xxx) = G'(x). \tag{7}$$

О построении разделенных разностей см., напр., [1].

Справедливы следующие аналоги интерполяционной формулы Ньютона:

$$\text{а)} \ f(x) = f(x') + (f(x'x''), x - x') + (f_1(x'x''x''')(x - x''), x - x') + ([f_1(x'x''x) - f_1(x'x''x''')](x - x''), x - x'); \tag{8}$$

$$\text{б)} \ f(x) = f(x') + (f(x''x'), x - x') + (f_2(x'''x''x')(x - x''), x - x') + ([f_2(xx''x') - f_2(x'''x''x')](x - x''), x - x'). \tag{9}$$

2. Рассмотрим два интерполяционных аналога метода градиентов для нахождения локальной точки минимума функционала $f(x)$. Первый из этих методов получается, если в методе градиентов $G(x_n)$ заменить на $f(x_n x_{n-1})$, т. е.

$$x_{n+1} = x_n - \varepsilon f(x_n x_{n-1}) \tag{10}$$

($\varepsilon > 0$; $n = 1, 2, \ldots$; x_0, x_1 — начальные приближения).

Второй метод основывается на интерполяционной формуле (8)*. По приближениям x_n, x_{n-1}, x_{n-2} построим квадратичный функционал

$$f_n(x) = f(x_n) + (f(x_n x_{n-1}), x - x_n) + (f_1(x_n x_{n-1} x_{n-2})(x - x_{n-1}), x - x_n). \tag{11}$$

Новое приближение находим по формуле

$$x_{n+1} = x_n - \varepsilon \operatorname{grad} f_n(x_n). \tag{12}$$

* Второй вариант метода получается, если использовать формулу (9).

Так как

$$\operatorname{grad} f_n(x) = f(x_n x_{n-1}) + f_1(x_n x_{n-1} x_{n-2})(x - x_{n-1}) + \\ + f_1^*(x_n x_{n-1} x_{n-2})(x - x_n), \tag{13}$$

то получается следующий итерационный метод

$$x_{n+1} = x_n - \varepsilon [f(x_n x_{n-1}) + f_1(x_n x_{n-1} x_{n-2})(x_n - x_{n-1})] \tag{14}$$

($\varepsilon > 0$; $n = 2, 3, \ldots$; x_0, x_1, x_2 — начальные приближения).

Для сходимости метода (10) справедлива

Теорема 1. *Пусть*

1° $G(x^*) = f(x^* x^*) = 0$;

2° *в выпуклой области* $Q \subset H$, *содержащей последовательность* $\{x_n\}$ *и элемент* x^*, *справедливы оценки*

а) $([f_1(x' x'' x''') + f_2(x' x''' x''')] h, h) \geqslant \frac{1}{M} \|h\|^2$ $(M > 0)$;

б) $\|f_1(x'x''x''')\| \leqslant K$; $\|f_2(x'x''x''')\| \leqslant K$

(для каждого $h \in H$ *и* $x', x'', x''' \in Q)$;

3° а) $\varepsilon < \frac{1}{\sqrt{7} K}$, *если* $MK \leqslant \frac{\sqrt{7}}{4}$;

б) $\varepsilon < \frac{1}{4MK^2}$, *если* $MK \geqslant \frac{\sqrt{7}}{4}$.

Тогда x^* *является локальной точкой минимума функционала* $f(x)$, *единственной в области* Q. *Последовательность* $\{x_n\}$ *сходится к* x^*, *причем*

$$\|x_{n+1} - x^*\|^2 \leqslant \left(1 - \frac{2\varepsilon}{M} + \varepsilon^2 K^2\right) \|x_n - x^*\|^2 + \\ + \varepsilon^2 K^2 \|x_{n-1} - x^*\|^2 + 4\varepsilon^2 K^2 \|x_n - x^*\| \, \|x_{n-1} - x^*\| + \\ + 2\varepsilon^2 K^2 \|x_n - x^*\| \, \|x_{n-2} - x^*\| \qquad (n = 2, 3, \ldots). \tag{15}$$

Для сходимости метода (14) справедлива

Теорема 2. *Пусть*

1° $G(x^*) = f(x^* x^*) = 0$;

2° *в выпуклой области* $Q \subset H$, *содержащей последовательность* $\{x_n\}$ *и элемент* x^*, *справедливы оценки*

а) $([f_1(x' x'' x''') + f_2(x' x''' x''')] h, h) \geqslant \frac{1}{M} \|h\|^2$ $(M > 0)$;

б) $\|f_1(x'x''x''') + f_2(x'x^{\mathrm{IV}}x^{\mathrm{IV}})\| \leqslant K$;

в) $\|f_1(x'x''x''') - f_1(x'x''x^{\mathrm{IV}})\| \leqslant L \|x''' - x^{\mathrm{IV}}\|$;

3° $\varepsilon < \frac{2}{M(K^2+4KLm+3L^2m^2)}$,

где

$$m = \max\{\|x^*-x_0\|;\ \|x^*-x_1\|;\ \|x^*-x_2\|;\ \|x^*-x_3\|\}.$$

Тогда x^* является локальной точкой минимума функционала $f(x)$, единственной в области Q. Последовательность $\{x_n\}$ сходится к x^*, причем

$$\|x_{n+1}-x^*\|^2 \leqslant \left(1-\frac{2\varepsilon}{M}+\varepsilon^2K^2\right)\|x_n-x^*\|^2 +$$
$$+4\varepsilon^2KL\|x_n-x^*\|\,\|x_{n-1}-x^*\|\,\|x_{n-2}-x^*\| +$$
$$+2\varepsilon^2L^2\|x_n-x^*\|\,\|x_{n-2}-x^*\|^2\,\|x_{n-3}-x^*\| +$$
$$+\varepsilon^2L^2\|x_{n-1}-x^*\|^2\,\|x_{n-2}-x^*\|^2. \tag{16}$$

Приведем доказательство теоремы 2. Покажем, что x^* является точкой минимума. Поскольку по 2° а) $(G'(x)h,h) \geqslant \frac{1}{M}\|h\|^2$ и по 1° $G(x^*)=0$, то по формуле Тейлора

$$f(x)=f(x^*)+\tfrac{1}{2}(G'(x^*+\Theta(x-x^*))(x-x^*),x-x^*) \geqslant$$
$$\geqslant f(x^*)+\frac{1}{2M}\|x-x^*\|^2 \geqslant f(x^*),$$

т. е. $f(x) \geqslant f(x^*)$, если $x \in Q$.

Покажем, что x^* является единственной точкой минимума в Q. Пусть $x^{**} \in Q$ и $G(x^{**})=0$ $(x^{**} \neq x^*)$. Тогда

$$0=(G(x^{**})-G(x^*),x^{**}-x^*)=$$
$$=\int_0^1 (G'(x^*+t(x^{**}-x^*))(x^{**}-x^*),x^{**}-x^*)\,dt \geqslant$$
$$\geqslant \frac{1}{M}\|x^{**}-x^*\|^2, \text{ т. е. } x^*=x^{**}.$$

Дальше получим:

$$\|x_{n+1}-x^*\|^2=\|x_n-x^*\|^2-2\varepsilon(x_n-x^*,f(x_n\,x_{n-1})+$$
$$+f_1(x_n\,x_{n-1}\,x_{n-2})(x_n-x_{n-1}))+\varepsilon^2\|f(x_n\,x_{n-1})+f_1(x_n\,x_{n-1}\,x_{n-2})(x_n-x_{n-1})\|^2; \tag{17}$$

$$(x_n-x^*;f(x_n\,x_{n-1})+f_1(x_n\,x_{n-1}\,x_{n-2})(x_n-x_{n-1}))=$$
$$=(x_n-x^*,[f_1(x_n\,x_{n-1}\,x^*)+f_2(x_n\,x^*\,x^*)](x_n-x^*))-$$
$$-(x_n-x^*,[f_1(x_n\,x_{n-1}\,x^*)-f_1(x_n\,x_{n-1}\,x_{n-2})](x_n-x_{n-1})) \geqslant$$
$$\geqslant \frac{1}{M}\|x_n-x^*\|^2-L\|x_n-x^*\|\,\|x_{n-2}-x^*\|\,\|x_n-x_{n-1}\|; \tag{18}$$

$$\|x_n-x_{n-1}\|=\|\varepsilon\{[f_1(x_{n-1}\,x_{n-2}\,x^*)-f_1(x_{n-1}\,x_{n-2}\,x_{n-3})](x_{n-2}-x^*)+$$
$$+[f_1(x_{n-1}\,x_{n-2}\,x_{n-3})+f_2(x_{n-1}\,x^*\,x^*)](x_{n-1}-x^*)\}\| \leqslant$$
$$\leqslant \varepsilon(L\|x_{n-2}-x^*\|\,\|x_{n-3}-x^*\|+K\|x_{n-1}-x^*\|); \tag{19}$$

$$\|f(x_n x_{n-1}) + f_1(x_n x_{n-1} x_{n-2})(x_n - x_{n-1})\| \leqslant$$
$$\leqslant K\|x_n - x^*\| + L\|x_{n-1} - x^*\|\,\|x_{n-2} - x^*\|. \tag{20}$$

Из (17)—(20) следуют оценки (16). Условие 3° является достаточным для сходимости x_n к x^*. Отметим, что в качестве Q можно взять сферу $\|x - x^*\| \leqslant m$.

3. Используя соображения статьи [6], дадим обобщение метода хорд для решения задачи

$$\min_{x \in R} f(x), \tag{21}$$

где R — замкнутое выпуклое множество в гильбертовом пространстве H. Предположим двукратную непрерывную дифференцируемость функционала $f(x)$. На каждом итерационном шаге определим x_{n+1} ($n = 1, 2, \ldots$) из условий

$$x_{n+1} \in R;\ f_n(x_{n+1}) \leqslant f_n(x) \text{ для всех } x \in R;$$
$$f_n(x) = (G(x_n), x - x_n) + \tfrac{1}{2}(G(x_n x_{n-1})(x - x_n), x - x_n) \tag{22}$$

(x_0, x_1 — начальные приближения; $G(x' x'')$ — разделенная разность оператора $G(x)$).

Отметим, что при задачах без ограничений из (22) получим алгоритм

$$x_{n+1} = x_n - 2[G(x_n x_{n-1}) + G^*(x_n x_{n-1})]^{-1} G(x_n),$$

который несколько отличается от обычного метода хорд для решения уравнения $G(x) = 0$ (ср. [8]).

Сходимость метода хорд к решению задачи (21) характеризует

Теорема 3. *Пусть*

1° а) $([G(x' x'') + G^*(x' x'')]h, h) \geqslant \frac{1}{M}\|h\|^2 \quad (M > 0)$;

б) $$\int_0^1 \|G'(x'' + t(x' - x'')) - \tfrac{1}{2}[G(x'' x''') + G^*(x'' x''')]\|\,dt \leqslant$$
$$\leqslant K(\|x' - x''\| + \|x'' - x'''\|)$$

(для каждого $h \in H$ и $x', x'', x''' \in R$);

2° $\|x_2 - x_1\| \leqslant \|x_1 - x_0\|$;

3° $d_0 = 2MK\|x_1 - x_0\| < \frac{1}{2}$.

Тогда функционал $f(x)$ имеет на множестве R единственную точку минимума x^*, к которой последовательность $\{x_n\}$ сходится с быстротой

$$\|x^* - x_n\| \leqslant \frac{1}{4MK}\sum_{i=n}^{\infty}(2d_0)^{\alpha_i} \qquad (n = 0, 1, \ldots), \tag{23}$$

где α_i — числа Фибоначчи ($\alpha_0 = \alpha_1 = 1$).

Доказательство. Так как

$$G_n(x) = \operatorname{grad} f_n(x) = \\ = G(x_n) + \tfrac{1}{2}[G(x_n x_{n-1}) + G^*(x_n x_{n-1})](x - x_n) \tag{24}$$

и

$$G'_n(x) = \tfrac{1}{2}[G(x_n x_{n-1}) + G^*(x_n x_{n-1})], \tag{25}$$

то по условию 1°а) функционалы $f_n(x)$ являются сильно выпуклыми и. следовательно, достигают единственного минимума на множестве R.

Итак, для всех $x \in R$

$$(\operatorname{grad} f_n(x_{n+1}), x - x_{n+1}) \geqslant 0 \tag{26}$$

(см. [6], формула (1.1)).

По (22), (24), (26) и 1°а)

$$f_n(x_{n+1}) = (\operatorname{grad} f_n(x_{n+1}), x_{n+1} - x_n) - \\ - \tfrac{1}{4}([G(x_n x_{n-1}) + G^*(x_n x_{n-1})](x_{n+1} - x_n), x_{n+1} - x_n) \leqslant \\ \leqslant -\frac{1}{4M}\|x_{n+1} - x_n\|^2,$$

т. е.

$$\|x_{n+1} - x_n\|^2 \leqslant -4M f_n(x_{n+1}). \tag{27}$$

С другой стороны,

$$-f_n(x_{n+1}) \leqslant -(G(x_n), x_{n+1} - x_n) - \frac{1}{4M}\|x_{n+1} - x_n\|^2 \tag{28}$$

и поскольку

$$G(x_n) = G(x_{n-1}) + \tfrac{1}{2}[G(x_{n-1} x_{n-2}) + G^*(x_{n-1} x_{n-2})](x_n - x_{n-1}) + R_n, \tag{29}$$

где

$$R_n = \int_0^1 \{G'(x_{n-1} + t(x_n - x_{n-1})) - \tfrac{1}{2}[G(x_{n-1} x_{n-2}) + \\ + G^*(x_{n-1} x_{n-2})]\}(x_n - x_{n-1})\,dt,$$

то по (26) и 1° б)

$$-f_n(x_{n+1}) \leqslant -(\operatorname{grad} f_{n-1}(x_n), x_{n+1} - x_n) - (R_n, x_{n+1} - x_n) - \\ - \frac{1}{4M}\|x_{n+1} - x_n\|^2 \leqslant \\ \leqslant K(\|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\|)\|x_n - x_{n-1}\|\|x_{n+1} - x_n\| - \\ - \frac{1}{4M}\|x_{n+1} - x_n\|^2. \tag{30}$$

На основании (27) и (30) получим оценки:

$$\|x_{n+1} - x_n\| \leqslant 2MK(\|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\|)\|x_n - x_{n-1}\| \\ (n = 2, 3, \ldots). \tag{31}$$

Отсюда по индукции легко следуют существование предела x^* последовательности $\{x_n\}$ и справедливость оценок (23).

Так как $G'(x) = G'^*(x)$ (см. [7], теорема 5.3), то по 1° а) $(x' = x'' = x)$ функционал $f(x)$ является сильно выпуклым. Поскольку выполнено и необходимое условие минимума

$$(G(x^*), x - x^*) + \tfrac{1}{2}(G'(x^*)(x - x^*), x - x^*) \geqslant 0 \tag{32}$$

(см. [6], формула (1.3)), то x^* является единственной точкой минимума.

Теорема доказана.

Замечание 1. Пусть $G_1(x' x'' x''')$ и $G_2(x' x'' x''')$ — вторые разделенные разности оператора $G(x)$ (см. [1]); G_1^* и G_2^* — их сопряженные*. Пусть выполнены оценки

$$\|G_i(x' x'' x''')\| \leqslant K; \qquad \|G_i^*(x' x'' x''')\| \leqslant K \qquad (i = 1, 2). \tag{33}$$

Тогда легко видеть, что выполнена оценка 1°б). Из доказательства теоремы ясно, что достаточно требовать выполнения (33) (или 1°б) лишь на множестве $R \cap S$,

$$\text{где } S = \left\{x : \|x - x_0\| \leqslant \frac{1}{4MK} \sum_{i=0}^{\infty} (2d_0)^{\alpha_i}\right\}. \tag{34}$$

Замечание 2. В практике представляет трудность проверка условия 1°а). Но если известно, что для каждого $x \in R$ и $h \in H$ выполнено

$$(G'(x)h, h) \geqslant \frac{1}{M'} \|h\|^2, \tag{35}$$

то достаточно проверить справедливость оценки 1а) только для $x', x'' \in R \cap S$ (ср. замечание 1). При этом можно взять

$$\frac{1}{M} = \frac{2}{M'} - 4Kr, \tag{36}$$

где r — радиус сферы S.

4. Приведенное в пункте 3 обобщение метода хорд для решения задачи на экстремум при наличии ограничений требует вычисления градиента $G(x)$ функционала $f(x)$. Дадим для решения задачи (21) обобщение метода, предложенного в [4], которое требует при реализации лишь вычисления значений разделенных разностей функционала $f(x)$.

Предположим, что для $f(x)$ существуют непрерывные разделенные разности первого и второго порядка $f(x' x'')$, $f_1(x' x'' x''')$, $f_2(x' x'' x''')$. Для разностей второго порядка потребуем, чтобы выполнялись (вместо пункта 3° в (7)) следующие условия:

$$f_1(x x x) + f_1^*(x x x) = G'(x);$$

$$f_2(x x x) + f_2^*(x x x) = G'(x),$$

где f_1^* и f_2^* — операторы, сопряженные соответственно с f_1 и f_2.

---

* Сопряженный A^* к билинейному оператору A можно определить по формуле $(Ahk, l) = (k, A^*hl)$ для всех $h, k, l \in H$.

Для решения проблемы (21) предлагаем следующий итерационный метод:

пусть известны начальные приближения x_0, x_1, x_2 к точке минимума (21);

если построены приближения до $x_n \in R$, то x_{n+1} $(n=2, 3, \ldots)$ найдутся из условий

$$x_{n+1} \in R;\ f_n(x_{n+1}) \leqslant f_n(x) \text{ для всех } x \in R,$$

где

$$f_n(x) = (f(x_n x_{n-1}), x - x_n) + (f_1(x_n x_{n-1} x_{n-2})(x - x_{n-1}), x - x_n) \tag{37}$$

(ср. (11))*.

Отметим, что для минимизации $f(x)$ без ограничений из (37) получим алгоритм

$$x_{n+1} = x_n - [f_1(x_n x_{n-1} x_{n-2}) + f_1^*(x_n x_{n-1} x_{n-2})]^{-1} \times$$
$$\times [f(x_n x_{n-1}) + f_1(x_n x_{n-1} x_{n-2})(x_n - x_{n-1})] \tag{38}$$

($n = 2, 3, \ldots$; x_0, x_1, x_2 — начальные приближения).

Метод (38), примененный для нахождения стационарных точек функций нескольких переменных, при $f_1^* = f_2$ совпадает с одним из методов, приведенных в [4].

Для сходимости метода (37) имеет место

Теорема 4. *Пусть*

1° а) $([f_1(x' x'' x''') + f_1^*(x' x'' x''')]h, h) \geqslant \frac{1}{M}\|h\|^2$ $(M > 0)$;

б) $\|f_1(x' x'' x''') - f_1(x'' x''' x^{IV})\| \leqslant$
$\leqslant a\|x' - x''\| + b\|x'' - x'''\| + c\|x''' - x^{IV}\|$;

в) $\|f_2(x' x'' x''') - f_1^*(x'' x''' x^{IV})\| \leqslant$
$\leqslant d\|x' - x''\| + e\|x'' - x'''\| + f\|x''' - x^{IV}\|$

(для каждого $h \in H$ и $x', x'', x''', x^{IV} \in R$);

2° $\|x_3 - x_2\| \leqslant \|x_2 - x_1\| \leqslant \|x_1 - x_0\|$;

3° $d_0 = M\{[2(a+b) + d + e]\|x_2 - x_1\| + (2c + f)\|x_1 - x_0\|\} < 1$.

Тогда функционал $f(x)$ имеет на множестве R единственную точку минимума x^*, к которой последовательность $\{x_n\}$ сходится с быстротой

$$\|x^* - x_n\| \leqslant \|x_2 - x_1\| \sum_{i=n}^{\infty} d_0^{\alpha_i} \quad (n = 1, 2, \ldots), \tag{39}$$

где

$$\alpha_i = \alpha_{i-2} + \alpha_{i-3} + 1, \quad (\alpha_0 = \alpha_1 = \alpha_2 = 0).$$

* Другой алгоритм можно получить на основании интерполяционной формулы (9).

Доказательство. Так как доказательство данной теоремы в основном аналогично доказательству теоремы 3, то приведем только некоторые более важные его моменты.

Сильная выпуклость $f_n(x)$ вытекает из условия 1°а), если учесть, что дифференцирование выражения (13) дает

$$G_n'(x) = f_1(x_n x_{n-1} x_{n-2}) + \overset{*}{f_1}(x_n x_{n-1} x_{n-2}).$$

По (37), (13), 1°а), (26) имеем:

$$f_n(x_{n+1}) \leqslant -\frac{1}{2M}\|x_{n+1} - x_n\|^2$$

или

$$\|x_{n+1} - x_n\|^2 \leqslant -2Mf_n(x_{n+1}). \tag{40}$$

С другой стороны, учитывая (26), 1°б), в) и то, что

$$f(x_n x_{n-1}) = f(x_{n-1} x_{n-2}) + f_2(x_n x_{n-1} x_{n-2})(x_n - x_{n-1}) +$$
$$+ f_1(x_n x_{n-1} x_{n-2})(x_{n-1} - x_{n-2}),$$

получим

$$-f_n(x_{n+1}) \leqslant [(a\|x_n - x_{n-1}\| + b\|x_{n-1} - x_{n-2}\| + c\|x_{n-2} - x_{n-3}\|) \times$$
$$\times \|x_n - x_{n-2}\| + (d\|x_n - x_{n-1}\| + e\|x_{n-1} - x_{n-2}\| + f\|x_{n-2} - x_{n-3}\|) \times$$
$$\times \|x_n - x_{n-1}\|]\,\|x_{n+1} - x_n\| - \frac{1}{2M}\|x_{n+1} - x_n\|^2. \tag{41}$$

На основании (40) и (41)

$$\|x_{n+1} - x_n\| \leqslant M[(a\|x_n - x_{n-1}\| + b\|x_{n-1} - x_{n-2}\| + c\|x_{n-2} - x_{n-3}\|) \times$$
$$\times \|x_n - x_{n-2}\| + (d\|x_n - x_{n-1}\| + e\|x_{n-1} - x_{n-2}\| + f\|x_{n-2} - x_{n-3}\|) \times$$
$$\times \|x_n - x_{n-1}\|] \qquad (n = 2, 3, \ldots). \tag{42}$$

По индукции из (42) легко показать существование $x^* = \lim\limits_{n\to\infty} x_n$ и справедливость (39).

Сильная выпуклость $f(x)$ вытекает из 1° а) $(x' = x'' = x''' = x)$, а необходимое условие минимума (32) следует из (37), переходя к пределу при $n \to \infty$.

Теорема доказана.

Замечание 3. Из доказательства теоремы 4 ясно, что достаточно требовать выполнения 1° б) и в) лишь на множестве $R \cap S'$, где

$$S' = \{x : \|x - x_0\| \leqslant r'\}, \tag{43}$$

где

$$r' = \|x_2 - x_1\| \sum_{i=1}^{\infty} d_0^{\alpha_i} + \|x_1 - x_0\|.$$

Замечание 4. Если для каждого $x \in R$ и $h \in H$ выполнено (35), то достаточно проверить справедливость оценки 1° а) теоремы 4 только для $x', x'', x''' \in R \cap S'$. При этом можно взять, учитывая 1° б) и в),

$$\frac{1}{M} = \frac{1}{M'} - 2r'(b + c + e + f).$$

5. Рассмотрим приложение метода (37) к задаче оптимального управления: минимизировать по $u \in U$

$$f(u) = \int_0^T F(x(t), u(t), t)\, dt + \Phi(x(T)),$$

если

$$\begin{cases} \dfrac{dx}{dt} = A(t)\, x(t) + B(t)\, u(t) \\ x(0) = c. \end{cases} \tag{44}$$

Для простоты рассмотрим случай скалярных x и u.

В данном случае можно положить

$$(f(u'\, u''), \Delta u) = \int_0^T [F(x'\, x'';\ u',\ t)\, \Delta x + F(x'';\ u'\, u'';\ t)\, \Delta u]\, dt + \\ + \Phi(x'(T)\, x''(T))\, \Delta x(T) \tag{45}$$

и

$$(f_1(u'\, u''\, u''')\, \Delta u,\ \Delta \bar{u}) = \int_0^T [F(x'\, x''\, x''';\ u',\ t)\, \Delta x\, \Delta \overline{x} + \\ + F(x''\, x''';\ u'\, u'';\ t)\, \Delta x\, \Delta \bar{u} + F(x''';\ u'\, u''\, u''';\ t)\, \Delta u\, \Delta \bar{u}]\, dt + \\ + \Phi(x'(T)\, x''(T)\, x'''(T))\, \Delta x(T)\, \Delta \overline{x}(T). \tag{46}$$

Здесь x', x'', x''', Δx, $\Delta \overline{x}$ — решения уравнения (44), если управлениями являются соответственно u', u'', u''', Δu, $\Delta \bar{u}$, причем $x'(0) = x''(0) = x'''(0) = c$; $\Delta x(0) = \Delta \overline{x}(0) = 0$.

Нетрудно показать, что в данном случае $f_1^* = f_2$; $f(u\, u) = G(u)$ (см. [6]); $f_1(u\, u\, u) + f_1^*(u\, u\, u) = G'(u)$.

Для нахождения поправки управления $\Delta u_n = u_{n+1} - u_n$ ($u_{n+1} \in U$) приходится на каждом итерационном шаге (ср. [6]) минимизировать квадратичный функционал

$$\int_0^T [F(x_n\, x_{n-1};\ u_n,\ t)\, \Delta x_n + F(x_{n-1};\ u_n\, u_{n-1};\ t)\, \Delta u_n + \\ + F(x_n\, x_{n-1}\, x_{n-2};\ u_n,\ t)\, (\Delta x_n + \Delta x_{n-1})\, \Delta x_n + \\ + F(x_{n-1}\, x_{n-2};\ u_n\, u_{n-1};\ t)\, (\Delta x_n + \Delta x_{n-1})\, \Delta u_n + \\ + F(x_{n-2};\ u_n\, u_{n-1}\, u_{n-2};\ t)\, (\Delta u_n + \Delta u_{n-1})\, \Delta u_n]\, dt + \\ + \Phi(x_n(T)\, x_{n-1}(T))\, \Delta x_n(T) + \Phi(x_n(T)\, x_{n-1}(T)\, x_{n-2}(T))\, (\Delta x_n(T) + \\ + \Delta x_{n-1}(T))\, \Delta x_n(T),$$

причем

$$\frac{d\Delta x_n}{dt} = A(t)\, \Delta x_n(t) + B(t)\, \Delta u_n(t);\ \Delta x_n = 0. \tag{47}$$

В формулах (45), (46) и (47) используются обозначения, приведенные в [1].

6. Предлагаем модификацию метода (38) для нахождения точек безусловного минимума, использующую, кроме разделенных разностей, и градиент функционала $f(x)$, но имеющую уже порядок сходимости 2.

Предположим, что уравнение $\operatorname{grad} f(x) = G(x) = 0$ представлено в виде

$$x = H(x). \tag{48}$$

В методе (38) можно в таком случае (используя идею метода Стеффенсена [5]) x_{n-1} и x_{n-2} заменить соответственно на $y_n = \alpha x_n + (1-\alpha) z_n$ $(0 < \alpha < 1)$ и $z_n = H(x_n)$. Получим следующий метод:

$$x_{n+1} = x_n - [f_1(x_n y_n z_n) + \overset{*}{f_1}(x_n y_n z_n)]^{-1} \times$$
$$\times [f(x_n y_n) + f_1(x_n y_n z_n)(x_n - y_n)] \qquad (n = 0, 1, \ldots). \tag{49}$$

Для метода (49) справедлива

Теорема 5. *Пусть*

1° $G(x^*) = f(x^* x^*) = 0$, *причем* $\|x^* - x_0\| \leqslant m$;

2° *для всех* $h \in H$ и $x', x'', x''', x^{\mathrm{IV}} \in S(x^*; \max\{Nm; m\}) \subset H$ *справедливы оценки*

а) $([f_1(x' x'' x''') + \overset{*}{f_1}(x' x'' x''')]h, h) \geqslant \frac{1}{M}\|h\|^2$ $(M > 0)$;

б) $\|f_1(x' x' x'') - f_1(x' x' x''')\| \leqslant c\|x'' - x'''\|$;

в) $\|f_2(x' x'' x'') - \overset{*}{f_1}(x' x''' x^{\mathrm{IV}})\| \leqslant e\|x'' - x'''\| + f\|x'' - x^{\mathrm{IV}}\|$;

г) $\|H(x') - H(x'')\| \leqslant N\|x' - x''\|$;

3° $Fm = M\{\alpha e + [f + \alpha c + (1-\alpha)e]N + (1-\alpha)cN^2\}m < 1$.

Тогда x^* является локальной точкой минимума функционала $f(x)$, единственной в сфере S. Последовательность $\{x_n\}$ сходится к x^*, причем

$$\|x^* - x_n\| \leqslant F^{-1}(Fm)^{2^n} \qquad (n = 0, 1, \ldots). \tag{50}$$

Доказательство. Точно таким же образом, как в теореме 2, доказывается (используя условия 1° и 2°а) утверждение, что x^* является единственной точкой минимума $f(x)$.

Далее, из 2°а) следует, что в S существует $[f_1(x' x'' x''') + \overset{*}{f_1}(x' x'' x''')]^{-1}$, причем

$$\|[f_1(x' x'' x''') + \overset{*}{f_1}(x' x'' x''')]^{-1}\| \leqslant M. \tag{51}$$

Поскольку в силу 2°б), 2°в), 2°г) и (51)

$$\|x^* - x_n\| \leqslant M\{\alpha e + [f + \alpha c + (1-\alpha)e]N + (1-\alpha)cN^2\} \times$$
$$\times \|x^* - x_{n-1}\|^2 \qquad (n = 1, 2, \ldots), \tag{52}$$

то, учитывая 1°, легко получить по индукции из (52) оценки (50). Условие 3° является достаточным для сходимости x_n к x^*.

Теорема доказана.

Замечание 5. Если $f(x)$ — функция n переменных, то при $n \geqslant 4$ в методе (49) на каждом итерационном шаге требуется вычислить суммарно меньше значений функции $f(x)$ и компонентов градиента $G(x)$, чем в методе Стеффенсена [5] для решения уравнения $G(x) = 0$.

Если в (38) x_{n-1} и x_{n-2} заменить соответственно на $y_n = H(x_n)$ и $z_n = H(H(x_n))$, то получится также метод, имеющий порядок сходимости 2. В этом случае преимущество перед методом Стеффенсена в вышеуказанном смысле проявляется при $n \geqslant 6$.

7. Приведем численный пример применения метода (37).

Рассмотрим задачу

$$\min_{Ax \leqslant b} f(x), \tag{53}$$

где

$$f(x) = f(x_1, x_2) = (x_1 - 2)^4 + (x_2 - 3)^4 + 3$$

$$A = \begin{pmatrix} 2 & 3 \\ -2 & 3 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad b = \begin{pmatrix} 12 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Для решения задач квадратичного программирования

$$\min \{ f_n(x) \,/\, Ax \leqslant b \}$$

пользовались методом Хилдрета и д'Эзопо [9].

Результаты расчета (53) на ЭВМ «Минск-2» приведены в нижеследующей таблице:

n	x_1	x_2	$f(x_1, x_2)$
0	4,5	3,0	42,0625
1	155	72	$0,5706484 \cdot 10^9$
2	1,0	2,0	5,0
3	2,830722	1,887148	5,009968
4	1,996051	1,330700	10,76491
5	3,000004	2,000003	5,000006
6	2,919685	1,946456	4,947409
7	2,912570	1,941713	4,947863
8	2,920011	1,946674	4,947407
9	2,920109	1,946739	4,947407

Точные значения переменных x_1, x_2, при которых достигается минимум $f(x_1, x_2)$ ($\cong 4,947407857$):

$$x_1 = \frac{3}{2} x_2 \ (\cong 2,92021662),$$

$$x_2 = \frac{3 + 2\sqrt[3]{\frac{3}{2}}}{1 + \frac{3}{2}\sqrt[3]{\frac{3}{2}}} \ (\cong 1,94681108).$$

Отметим, что при решении примера (53) разделенные разности первого и второго порядка функции $f(x_1, x_2)$ были вычислены по следующим формулам:

если $$x' = (x_1', x_2'),\ x'' = (x_1'', x_2''),\ x''' = (x_1''', x_2'''),$$

то $$f(x'x'') = (f(x_1'x_1''; x_2'),\ f(x_1''; x_2'x_2''));$$

$$f_1(x'x''x''') = \begin{pmatrix} f(x_1'x_1''x_1'''; x_2') & 0 \\ f(x_1''x_1'''; x_2'x_2'') & f(x_1'''; x_2'x_2''x_2''') \end{pmatrix}$$

(ср. [1, 4]).

ЛИТЕРАТУРА

1. Ульм С., Изв. АН ЭССР. Физика * Математика, **16**, № 1, 13—26 (1967).
2. Schmidt J. W., Wiss. Z. Techn. Univ. Dresden, **12**, № 6, 1601—1605 (1963).
3. Schmidt J. W., Trinkaus H.-F., Computing, **1**, fasc. 3, 224—232 (1966).
4. Полль В., Изв. АН ЭССР. Физика * Математика, **16**, № 1, 35—44 (1967).
5. Ульм С., Ж. вычисл. матем. и матем. физ., **4**, № 6, 1093—1097 (1964).
6. Левитин Е. С., Поляк Б. Т., Ж. вычисл. матем. и матем. физ., **6**, № 5, 787—823 (1966).
7. Вайнберг М. М., Вариационные методы исследования нелинейных операторов, М., 1956.
8. Schmidt J. W., Z. angew. Math. und Mech., **41**, Sonderheft, 61—63 (1961).
9. Кюнци Г. П., Крелле В., Нелинейное программирование, М., 1965.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
18/XII 1967

S. ULM, V. POLL

MÕNEDEST MIINIMUMI LEIDMISE MEETODITEST

Artiklis antakse esimest ja teist järku diferentssuhete definitsioonid ning Newtoni interpolatsioonivalemid funktsionaalide puhul Hilberti ruumis. Vaadeldakse diferentssuhteid kasutavaid algoritme tõketeta (valemid (10), (14) ja (49)) ja tõketega miinimumülesannete (valemid (22) ja (37)) lahendamiseks. Antakse piisavad tingimused nende meetodite koonduvuseks (teoreemid 1—5). Vaadeldakse meetodi (37) rakendamist optimaaljuhtimisülesande (44) ja mittelineaarse programmeerimisülesande (53) lahendamiseks.

S. ULM, V. POLL

ON SOME METHODS FOR SOLVING MINIMUM PROBLEMS

In this paper definitions of divided differences and Newton's interpolation formulae for the functionals in the Hilbert space are presented. Algorithms are considered for solving unconstrained minimum problems (formulae (10), (14) and (49)) and the same with constraints (formulae (22) and (37)) with divided differences. Sufficient conditions for the convergence of these methods (theorems 1—5) are described.

The application (47) of method (37) to optimal control problem (44) and a numerical illustration of the application of the same method in the solution of the nonlinear programming problem (53) are given.