EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. IX KÖIDE
FÜÜSIKALIS-MATEMAATILISTE JA TEHNILISTE TEADUSTE SEERIA. 1960, NR. 2

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ IX
СЕРИЯ ФИЗИКО-МАТЕМАТИЧЕСКИХ И ТЕХНИЧЕСКИХ НАУК. 1960, № 2

О МЕТОДЕ НАИСКОРЕЙШЕГО СПУСКА ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

Л. КИВИСТИК

1. Пусть $P(x)$ есть нелинейный оператор из вещественного гильбертова пространства H в то же пространство. Рассмотрим уравнение

$$P(x)=0 \tag{1}$$

Ю. Лумисте [1] показал, что применение метода наискорейшего спуска для приближенного решения уравнения (1), если $P(x)$ потенциален (о потенциальности оператора см. [5]), равносильно нахождению последовательных приближений по рекуррентной формуле

$$x_{n+1}=x_n+\varepsilon_n P(x_n), \qquad n=0,1,2,\ldots \tag{2}$$

где x_0 — начальное приближение, а ε_n — вещественное решение уравнения

$$(P(x_n),\quad P(x_n+\varepsilon P(x_n)))=0\,^* \tag{3}$$

Ю. Лумисте доказал сходимость метода в предположении, что оператор $P(x)$ удовлетворяет на некотором замкнутом множестве $X\subset H$ условию

$$(P(x+h)-P(x),\ h)\geqslant m\,\|h\|^2,\quad m\geqslant 1 \tag{4}$$

Гуань Чжао-чжи [2] доказал сходимость метода для дифференцируемого оператора в предположении, что для всех $x\in X$ выполнено условие

$$(P'(x)h,h)\geqslant m\,\|h\|^2,\quad m>0,\quad h\in H \tag{5}$$

В этом же предположении, только с абсолютной величиной в левой части неравенства (5), некоторую модификацию метода наискорейшего спуска рассмотрел и М. Альтман [3] **. Эту модификацию мы рассмотрим ниже, в пп. 4—6.

Практическая проверка условий (4) и (5) для всех x целой области X часто связана с большими трудностями. Ниже мы докажем сходимость метода наискорейшего спуска в предположении, что условие (5) с абсолютной величиной в левой части выполнено лишь для x_0.

* Символом $(u,\ v)$ обозначено скалярное произведение элементов $u,\ v\ni H$.

** Отметим, что в [3] условие 5° теоремы 2 не выполнено, так как $B^2(E^2+DK)>\geqslant 1+\dfrac{\|P(x)\|\,\|P''(x)\|}{\|P'(x)\|^2}\geqslant 1$ и не может быть меньше единицы, как требуется в теореме 2. По той же причине не выполнены и условия теоремы 4.

2. Пусть оператор $P(x)$ дважды дифференцируем в смысле Фреше в некоторой замкнутой сфере $S(x_0, r)$ *, которую мы уточняем каждый раз, когда о ней идет речь. Потенциальности оператора мы не требуем. Пусть дифференциал $P'(x)h$ удовлетворяет условию

$$|(P'(x_0)h, h)| \geqslant \frac{1}{M_0} \|h\|^2, \quad M_0 > 0 \tag{6}$$

для всех ** $h \in H$, где x_0 — начальное приближение решения уравнения (1). Последовательные приближения находим из рекуррентной формулы (2), где ε_n — наименьшее по абсолютной величине решение уравнения (3). Но если ε_n — решение этого уравнения, то (3) превращается в тождество

$$(P(x_n), \quad P(x_n + \varepsilon_n P(x_n))) = 0$$

или

$$(P(x_n), \quad P(x_{n+1})) = 0 \tag{7}$$

Все следующие рассуждения остаются в силе, если уравнение (3) при $n = 0, 1, 2, \ldots$ решено приближенно, но с такой точностью, что мы можем (учитывая точность нахождения x_{n+1}) считать

$$(P(x_n), \quad P(x_{n+1})) = 0, \qquad n = 0, 1, 2, \ldots$$

Теорема 1. *Пусть выполнены условия:*

1° $\|P(x_0)\| = \delta_0 \leqslant \bar{\delta}_0$ (x_0 — *начальное приближение*);

2° $|(P'(x_0)h, h)| \geqslant \frac{1}{M_0} \|h\|^2, \quad M_0 > 0, \quad h \in H$;

3° *для всех* $x \in S(x_0, r)$, *где* $r = \frac{1}{B}\left(\frac{1}{M_0} - \frac{1}{M^*}\right)\frac{\delta_0}{\bar{\delta}_0}$, *оператор* $P(x)$ *дважды дифференцируем и*

$$\|P'(x)\| \leqslant A, \qquad \|P''(x)\| \leqslant B;$$

4° *величины* M_0, $\bar{\delta}_0$, A *и* B *такие, что последовательности* $\{M_n\}$ *и* $\{\bar{\delta}_n\}$, *вычисленные из рекуррентных соотношений*

$$M_{n+1} = \frac{M_n}{\sqrt{1 - 2M_n^2 B \bar{\delta}_n}}, \qquad n = 0, 1, 2, \ldots \tag{8}$$

$$\bar{\delta}_{n+1} = \bar{\delta}_n \sqrt{\frac{4M_n^2 A^2}{\left(1 + \sqrt{1 - 2M_n^2 B \bar{\delta}_n}\right)^2} - 1}, \qquad n = 0, 1, 2, \ldots \tag{9}$$

сходятся.

Тогда уравнение (1) *имеет в сфере* $S(x_0, r)$ *решение* x^*, *к которому сходится полученная из* (2) *последовательность* $\{x_n\}$, *и справедлива оценка*

$$\|x^* - x_n\| \leqslant M^* \delta_n \leqslant \frac{\sqrt{2}}{A} \delta_n \tag{10}$$

где $M^* = \lim\limits_{n \to \infty} M_n$ *и* $\delta_n = \|P(x_n)\|$.

---

* В дальнейшем, говоря о сфере $S(x_0, r)$, считаем ее замкнутой.

** Можно потребовать, чтобы условие (6) было выполнено только при всех h, для которых $\|h\| = \varrho$, где ϱ — произвольное фиксированное положительное число. Но если условие выполнено для таких h, то оно выполнено во всем пространстве.

Доказательство. Из (8) и требования сходимости следует, что $\{M_n\}$ есть возрастающая последовательность, а $\lim M_n^2 B\bar{\delta}_n = 0$. Поэтому $\{\bar{\delta}_n\}$ должна (монотонно) убывать. Теперь из (9) вытекает, что $\lim M_n A \leqslant \sqrt{2}$, или $M^* \leqslant \frac{\sqrt{2}}{A}$.

Обозначим

$$f_n(\varepsilon) = (P(x_n),\ P(x_n + \varepsilon P(x_n)))$$

Тогда в силу условий теоремы

1) $$\frac{1}{|f_0'(0)|} = \frac{1}{|(P(x_0),\ P'(x_0)\,P(x_0))|} \leqslant M_0 \delta_0^{-2}$$

2) $$\frac{|f_0(0)|}{|f_0'(0)|} = \frac{\|P(x_0)\|^2}{|(P(x_0),\ P'(x_0)\,P(x_0))|} \leqslant M_0$$

3) $$|f_0''(\varepsilon)| = |(P(x_0),\ P''(x_0 + \varepsilon P(x_0))\,P^2(x_0))| \leqslant B\delta_0^3$$

для всех ε, удовлетворяющих условию

$$|\varepsilon| \leqslant \frac{1 - \sqrt{1 - 2M_0^2 B\bar{\delta}_0}}{M_0^2 B\bar{\delta}_0} M_0$$

потому что при таких ε $x_0 + \varepsilon P(x_0) \in S(x_0, r)$, что видно из неравенства

$$\|\varepsilon P(x_0)\| \leqslant \frac{1 - \sqrt{1 - 2M_0^2 B\bar{\delta}_0}}{M_0^2 B\bar{\delta}_0} M_0 \delta_0 = \frac{1}{B}\left(\frac{1}{M_0} - \frac{1}{M_1}\right)\frac{\delta_0}{\bar{\delta}_0} < r$$

4) $$M_0^2 B\bar{\delta}_0 < 1 - \frac{M_0}{M_1} = 1 - \frac{M_0 A}{M_1 A} < 1 - \frac{1}{\sqrt{2}} < \frac{1}{2}$$

так как в силу 2° и 3° $M_0 A \geqslant 1$ и, как мы показали, $M_1 A < M^* A \leqslant \sqrt{2}$.

Таким образом, выполнены все условия теоремы Л. В. Канторовича ([4], стр. 170) о сходимости метода Ньютона для уравнения

$$f_0(\varepsilon) = (P(x_0),\ P(x_0 + \varepsilon P(x_0))) = 0$$

где начальное приближение $\varepsilon_0^{(0)} = 0$. Поэтому уравнение (3) разрешимо при $n = 0$, и в силу той же теоремы

$$|\varepsilon_0| \leqslant \frac{1 - \sqrt{1 - 2M_0^2 B\bar{\delta}_0}}{M_0^2 B\bar{\delta}_0} M_0 = \frac{2M_0}{1 + \sqrt{1 - 2M_0^2 B\bar{\delta}_0}} \tag{11}$$

где ε_0 — наименьшее по абсолютной величине решение уравнения $f_0(\varepsilon) = 0$. Теперь $\|x_1 - x_0\| = |\varepsilon_0|\,\delta_0 < r$, т. е. $x_1 \in S(x_0, r)$.

Так как оператор $P(x)$ имеет в сфере $S(x_0, r)$ производную $P'(x)$ и $x_0, x_1 \in S(x_0, r)$, то справедлива формула Лагранжа ([5], стр. 55):

$$(P(x_1) - P(x_0), z) = (P'(\bar{x})\,(x_1 - x_0), z) \tag{12}$$

где z — произвольный элемент из H, а $\bar{x} = x_0 + \tau(x_1 - x_0) \in S(x_0, r)$, $0 < \tau < 1$. Из (12) вытекает, что

$$\|P(x_1) - P(x_0)\| \leqslant \|P'(\bar{x})\,(x_1 - x_0)\| \tag{13}$$

Используя (13) и учитывая (7), получим

$$\|P(x_1)\|^2+\|P(x_0)\|^2=\|P(x_1)-P(x_0)\|^2\leqslant\|P'(\bar{x})(x_1-x_0)\|^2\leqslant A^2\varepsilon_0^2\|P(x_0)\|^2$$

откуда в силу (11)

$$\delta_1\leqslant\delta_0\sqrt{\frac{4M_0^2A^2}{(1+\sqrt{1-2M_0^2B\bar{\delta}_0})^2}-1}\leqslant\bar{\delta}_0\sqrt{\frac{4M_0^2A^2}{(1+\sqrt{1-2M_0^2B\bar{\delta}_0})^2}-1}=\bar{\delta}_1 \tag{14}$$

Применяя формулу (12) к дифференциалу $P'(x)h$, получим

$$|(P'(x_1)h,h)|=|(P'(x_0)h,h)-(P'(x_0)h-P'(x_1)h,h)|\geqslant|(P'(x_0)h,h)|-$$
$$-|(P''(\bar{x})h(x_0-x_1),h)|\geqslant\frac{1}{M_0}\|h\|^2-B|\varepsilon_0|\delta_0\|h\|^2\geqslant$$
$$\geqslant\frac{1}{M_0}\left(1-\frac{1-\sqrt{1-2M_0^2B\bar{\delta}_0}}{M_0^2B\bar{\delta}_0}M_0^2B\bar{\delta}_0\right)\|h\|^2=\frac{1}{M_1}\|h\|^2$$

где

$$M_1=\frac{M_0}{\sqrt{1-2M_0^2B\bar{\delta}_0}}$$

Покажем, что $S(x_1,r_1)\subset S(x_0,r)$, где $r_1=\frac{1}{B}\left(\frac{1}{M_1}-\frac{1}{M^*}\right)\frac{\delta_1}{\bar{\delta}_1}$. Действительно, пусть $x\in S(x_1,r_1)$; тогда

$$\|x-x_0\|\leqslant\|x-x_1\|+\|x_1-x_0\|\leqslant\frac{1}{B}\left(\frac{1}{M_1}-\frac{1}{M^*}\right)\frac{\delta_1}{\bar{\delta}_1}+$$
$$+\frac{1}{B}\left(\frac{1}{M_0}-\frac{1}{M_1}\right)\frac{\delta_0}{\bar{\delta}_0}\leqslant\frac{1}{B}\left(\frac{1}{M_0}-\frac{1}{M^*}\right)\frac{\delta_0}{\bar{\delta}_0}=r$$

ибо в силу (14) $\frac{\delta_1}{\bar{\delta}_1}\leqslant\frac{\delta_0}{\bar{\delta}_0}$. Таким образом, $x\in S(x_0,r)$, т. е. $S(x_1,r_1)\subset S(x_0,r)$. Поэтому оценки, данные условием 3°, сохраняются для всех $x\in S(x_1,r_1)$. Так как мы получили также постоянные $\bar{\delta}_1$, M_1, которые удовлетворяют условиям 1° и 2° для x_1, и 4° тоже остается в силе, то все условия теоремы выполнены для x_1. Мы можем все рассуждения повторить, заменяя x_0 везде на x_1. Таким образом, нетрудно индуктивно доказать, что условия теоремы выполнены для всех x_n. Попутно получим рекуррентные соотношения (8), (9) и неравенства

$$|\varepsilon_n|\leqslant\frac{1-\sqrt{1-2M_n^2B\bar{\delta}_n}}{M_n^2B\bar{\delta}_n}M_n$$

$$\delta_{n+1}\leqslant\delta_n\sqrt{\frac{4M_n^2A^2}{(1+\sqrt{1-2M_n^2B\bar{\delta}_n})^2}-1}\leqslant\bar{\delta}_n\sqrt{\frac{4M_n^2A^2}{(1+\sqrt{1-2M_n^2B\bar{\delta}_n})^2}-1}=$$
$$=\bar{\delta}_{n+1}$$

а также включение $x_n\in S(x_0,r)$.

Покажем, что последовательность $\{x_n\}$ сходится к решению x^* уравнения (1), лежащему в сфере $S(x_0, r)$. Для любого κ имеем

$$\|x_{k+1}-x_k\| = |\varepsilon_k|\delta_k \leqslant \frac{1-\sqrt{1-2M_k^2 B\bar{\delta}_k}}{M_k^2 B\bar{\delta}_k} M_k\delta_k = \frac{1}{B}\left(\frac{1}{M_k}-\frac{1}{M_{k+1}}\right)\frac{\delta_k}{\bar{\delta}_k},$$

$$\frac{\delta_{k+1}}{\bar{\delta}_{k+1}} \leqslant \frac{\delta_k}{\bar{\delta}_k}$$

Поэтому

$$\|x_{n+p}-x_n\| \leqslant \|x_{n+p}-x_{n+p-1}\| + \ldots + \|x_{n+2}-x_{n+1}\| + \|x_{n+1}-x_n\| \leqslant$$
$$\leqslant \frac{1}{B}\left(\frac{1}{M_{n+p-1}}-\frac{1}{M_{n+p}}\right)\frac{\delta_{n+p-1}}{\bar{\delta}_{n+p-1}} + \ldots + \frac{1}{B}\left(\frac{1}{M_{n+1}}-\frac{1}{M_{n+2}}\right)\frac{\delta_{n+1}}{\bar{\delta}_{n+1}} +$$
$$+\frac{1}{B}\left(\frac{1}{M_n}-\frac{1}{M_{n+1}}\right)\frac{\delta_n}{\bar{\delta}_n} \leqslant \frac{1}{B}\left(\frac{1}{M_n}-\frac{1}{M_{n+p}}\right)\frac{\delta_n}{\bar{\delta}_n} \tag{15}$$

и в силу полноты пространства H существует $x^* = \lim x_n$. Переход к пределу $(p\to\infty)$ дает

$$\|x^*-x_n\| \leqslant \frac{1}{B}\left(\frac{1}{M_n}-\frac{1}{M^*}\right)\frac{\delta_n}{\bar{\delta}_n}$$

Из последнего неравенства следует, что $x^* \in S(x_0, r)$. Так как в силу 3° оператор $P(x)$ непрерывен, то

$$\|P(x^*)\| = \lim\|P(x_n)\| \leqslant \lim\bar{\delta}_n = 0$$

откуда следует, что x^* — решение уравнения (1).

Для всех $\bar{x} \in S(x_0, r)$

$$|(P'(\bar{x})h, h)| \geqslant \frac{1}{M_0}\left(1-M_0B\|\bar{x}-x_0\|\right)\|h\|^2 \geqslant \frac{1}{M_0}\left(1-M_0Br\right)\|h\|^2$$

Так как $x^* \in S(x_0, r)$ и $x_n \in S(x_0, r)$, то, пользуясь формулой (12), получим

$$\|P(x_n)\|\,\|x_n-x^*\| \geqslant |(P(x_n)-P(x^*),\ x_n-x^*)| =$$
$$= |(P'(\bar{x})(x_n-x^*), x_n-x^*)| \geqslant \frac{1}{M_0}\left(1-M_0Br\right)\|x_n-x^*\|^2$$

откуда

$$\|x^*-x_n\| \leqslant \frac{M_0}{1-M_0Br}\delta_n \leqslant M^*\delta_n \leqslant \frac{\sqrt{2}}{A}\delta_n$$

Теорема полностью доказана.

3. Рассмотрим вопрос о сходимости последовательностей $\{M_n\}$, $\{\bar{\delta}_n\}$. Введем обозначения

$$a_n = 2M_n^2 B\bar{\delta}_n, \qquad b_n = M_n^2 A^2 \tag{16}$$

Тогда, в силу (8) и (9), получим соотношения

$$a_{n+1}=\frac{a_n}{1-a_n}\sqrt{\frac{4b_n}{(1+\sqrt{1-a_n})^2}-1} \tag{17}$$

$$b_{n+1}=\frac{b_n}{1-a_n} \tag{18}$$

Имеет место следующая

Теорема 2. *Если $a_0<1$ и существует такое вещественное число $k>0$, что $\alpha=k^2a_0<2$, а также выполнены неравенства*:

$$1^\circ\ 4b_0\leqslant(2-a_0)^2\left[1+\left(\frac{k}{k+1}\right)^4(1-a_0)^2\right],$$

$$2^\circ\ (2-a_0)^2\left[1+\left(\frac{k}{k+1}\right)^4(1-a_0)^2\right]\leqslant$$
$$\leqslant(1-a_0)\left(2-\frac{k^2a_0}{(k+1)^2}\right)^2\left[1+\left(\frac{k+1}{k+2}\right)^4\left(1-\frac{k^2a_0}{(k+1)^2}\right)^2\right],$$

то последовательности $\{a_n\}$, $\{b_n\}$ (как и $\{M_n\}$ и $\{\bar\delta_n\}$) сходятся. При этом

$$\left.\begin{aligned} a_n&<\left(\frac{k}{k+n}\right)^2a_0\\ b_n&<\frac{1}{4}\left(2-\frac{k^2a_0}{(k+n)^2}\right)^2\left[1+\left(\frac{k+n}{k+n+1}\right)^4\left(1-\frac{k^2a_0}{(k+n)^2}\right)^2\right]\end{aligned}\right\} \tag{19}$$

Доказательство. Обозначим

$$F(u,\alpha)=\left(1-\frac{\alpha}{u^2}\right)\left(2-\frac{\alpha}{(u+1)^2}\right)^2\left[1+\left(\frac{u+1}{u+2}\right)^4\left(1-\frac{\alpha}{(u+1)^2}\right)^2\right]-$$
$$-\left(2-\frac{\alpha}{u^2}\right)^2\left[1+\left(\frac{u}{u+1}\right)^4\left(1-\frac{\alpha}{u^2}\right)^2\right]$$

где $\alpha=k^2a_0$. Теперь условие 2° можно записать в виде

$$F(k,\alpha)\geqslant 0$$

Покажем, что из этого неравенства следуют неравенства

$$F(k+n,\alpha)>0\qquad n=1,\,2,\,3,\ldots \tag{20}$$

Для доказательства воспользуемся методом математической индукции. Достаточно показать, что из неравенства $F(u,\alpha)\geqslant 0$, где $u\geqslant k>0$, $0<\alpha<2$, следует неравенство $F(u+1,\alpha)>0$. Для того, чтобы это доказать, приведем в выражении $F(u,\alpha)$ все члены к общему знаменателю

$$F(u,\alpha)=\frac{f(u,\alpha)}{\varphi(u)}$$

где

$$f(u,\alpha)=\left\{(u+2)^4+[(u+1)^2-\alpha]^2\right\}[2(u+1)^2-\alpha]^2\ (u^2-\alpha)u^2-$$
$$-\left\{(u+1)^4+(u^2-\alpha)^2\right\}(2u^2-\alpha)^2\ (u+2)^4$$

и

$$\varphi(u)=[u\ (u+1)\ (u+2)]^4>0$$

Таким образом, $F(u,\alpha)$ и $f(u,\alpha)$ имеют один и тот же знак, и поэтому достаточно показать, что из неравенства $f(u, \alpha) \geqslant 0$ следует неравенство $f(u+1, \alpha) > 0$.

После разложения $f(u, \alpha)$ и $f(u+1, \alpha)$ по степеням $u+1$ можно непосредственно проверить, что выражение

$$\psi(u, \alpha) = f(u+1, \alpha) - \Big[1 + \frac{10}{u+1} + \frac{49}{(u+1)^2} + \frac{156+4\alpha}{(u+1)^3} + \\ + \frac{330+31\alpha}{(u+1)^4} + \frac{334+90\alpha+\alpha^2}{(u+1)^5}\Big] f(u, \alpha)$$

положительно при всех $u > 0$, $0 < \alpha < 2$. Отсюда следует, что $f(u+1, \alpha) > 0$, а также $F(u+1, \alpha) > 0$. Но тогда выполнено и (20).

В силу условия 1° (и учитывая, что $a_0 < 1$) имеем

$$\frac{4b_0}{(1+\sqrt{1-a_0})^2} - 1 < \frac{4b_0}{(2-a_0)^2} - 1 \leqslant \left(\frac{k}{k+1}\right)^4 (1-a_0)^2$$

В силу (17) получим отсюда

$$a_1 < \left(\frac{k}{k+1}\right)^2 a_0 \tag{21}$$

Учитывая (18), получим, в силу 1° и 2°

$$4b_1 = \frac{4b_0}{1-a_0} \leqslant \frac{(2-a_0)^2}{1-a_0}\left[1 + \left(\frac{k}{k+1}\right)^4 (1-a_0)^2\right] \leqslant \\ \leqslant \left(2 - \frac{k^2 a_0}{(k+1)^2}\right)^2 \left[1 + \left(\frac{k+1}{k+2}\right)^4 \left(1 - \frac{k^2 a_0}{(k+1)^2}\right)^2\right] \tag{22}$$

Повторяя таким образом все рассуждения и учитывая при этом неравенство (21), а вместо 1° и 2° соответственно (22) и (20), получим неравенства (19) для $n = 2$. При помощи индукции получим неравенства (19) для всех n. Из них вытекает сходимость последовательностей $\{a_n\}$ и $\{b_n\}$; вместе с тем получим:

$$\lim a_n = 0, \qquad \lim b_n \leqslant 2$$

или

$$\lim \bar{\delta}_n = 0, \qquad \lim M_n \leqslant \frac{\sqrt{2}}{A}$$

4. Рассмотрим теперь одну модификацию метода наискорейшего спуска, практическое применение которой значительно проще, но точность почти та же самая, что и у метода {(2), (3)}. Постоянные ε_n находим приближенно следующим образом. Уравнение (3), которое может быть записано в виде

$$(P(x_n),\ P(x_n) + P'(x_n)\,\varepsilon P(x_n) + R(x_n, \varepsilon P(x_n))) = 0 \tag{23}$$

где

$$R(x_n, \varepsilon P(x_n)) = \int_0^1 (1-\lambda)\, P''(x_n + \lambda\varepsilon P(x_n))\, \varepsilon^2 P^2(x_n)\, d\lambda,$$

заменим приближенным

$$(P(x_n),\ P(x_n) + P'(x_n)\,\varepsilon P(x_n)) = 0 \tag{24}$$

Решение ε_n уравнения (24) выражается просто:

$$\varepsilon_n = -\frac{\|P(x_n)\|^2}{(P(x_n),\ P'(x_n)\,P(x_n))} \tag{25}$$

Отсюда получим следующую рекуррентную формулу для нахождения приближений уравнения (1):

$$x_{n+1} = x_n - \frac{\|P(x_n)\|^2}{(P(x_n),\ P'(x_n)\,P(x_n))}\,P(x_n) \tag{26}$$

Этот метод рассмотрен и в статьях [2] и [3].

Отметим, что если в (1) $P(x)$ — вещественная функция вещественной переменной, то метод (26) дает метод Ньютона, рассмотренный в [4].

Справедлива следующая теорема:

Теорема 3. *Пусть выполнены условия:*

1° $\|P(x_0)\| = \delta_0 \leqslant \bar{\delta}_0$ *(x_0 — начальное приближение);*

2° $|(P'(x_0)h, h)| \geqslant \frac{1}{M_0}\|h\|^2$, $M_0 > 0$, $h \in H$;

3° *для всех* $x \in S(x_0, r)$, *где* $r = \frac{1}{B}\left(\frac{1}{M_0} - \frac{1}{M^*}\right)\frac{\delta_0}{\bar{\delta}_0}$, *оператор* $P(x)$ *дважды дифференцируем, и справедливы оценки*

$$\|P'(x)\| \leqslant A, \quad \|P''(x)\| \leqslant B;$$

4° *величины* M_0, $\bar{\delta}_0$, A *и* B *такие, что последовательности* $\{M_n\}$, $\{\bar{\delta}_n\}$, *вычисленные из рекуррентных соотношений*

$$M_{n+1} = \frac{M_n}{1 - M_n^2 B\bar{\delta}_n}, \qquad n = 0, 1, 2, \ldots \tag{27}$$

$$\bar{\delta}_{n+1} = \bar{\delta}_n \sqrt{M_n^2 A^2 + M_n^2 B\bar{\delta}_n - 1}, \qquad n = 0, 1, 2, \ldots \tag{28}$$

сходятся (так что $M^2{}_n B\bar{\delta}_n < 1$ *для всех* n*).*

Тогда уравнение (1) *имеет в сфере* $S(x_0, r)$ *решение* x^*, *к которому сходится полученная из* (26) *последовательность* $\{x_n\}$, *и справедлива оценка*

$$\|x^* - x_n\| \leqslant M^*\delta_n \leqslant \frac{\sqrt{2}}{A}\,\delta_n \tag{29}$$

где $M^* = \lim M_n$, $\delta_n = \|P(x_n)\|$.

Доказательство аналогично доказательству теоремы 1. Как и раньше, получим, что $M_0 A \geqslant 1$ и $M_{n+1} > M_n$. Из требования сходимости последовательностей (27) и (28) следует, что $\lim \bar{\delta}_n = 0$ и $M^* = \lim M_n \leqslant \frac{\sqrt{2}}{A}$.

Пользуясь оценками 1°, 2° и соотношением (27), получим

$$\|x_1 - x_0\| \leqslant M_0\delta_0 = \frac{1}{B}\left(\frac{1}{M_0} - \frac{1}{M_1}\right)\frac{\delta_0}{\bar{\delta}_0} < \frac{1}{B}\left(\frac{1}{M_0} - \frac{1}{M^*}\right)\frac{\delta_0}{\bar{\delta}_0} = r$$

т. е. $x_1 \in S(x_0, r)$. В силу соотношений (23) и (25)

$$(P(x_{n+1}),\ P(x_n)) = (P(x_n),\ R(x_n,\ \varepsilon_n P(x_n)))$$

Так как справедлива оценка

$$\|R(x_n, \varepsilon_n P(x_n))\| \leqslant \frac{1}{2} \varepsilon_n^2 \|P(x_n)\|^2 \max_{0 \leqslant \lambda \leqslant 1} \|P''(x_n + \lambda \varepsilon_n P(x_n))\|$$

$x_0 + \lambda \varepsilon_0 P(x_0) \in S(x_0, r)$ для всех $\lambda \in [0,1]$, а в силу 2° и (25) $|\varepsilon_0| \leqslant M_0$, то имеем

$$|(P(x_1), P(x_0))| \leqslant \frac{1}{2} M_0^2 B \bar{\delta}_0 \|P(x_0)\|^2$$

Теперь

$$\|P(x_1)\|^2 - 2(P(x_1), P(x_0)) + \|P(x_0)\|^2 = \|P(x_1) - P(x_0)\|^2 \leqslant$$
$$\leqslant \|P'(\bar{x})(x_1 - x_0)\|^2 \leqslant A^2 M_0^2 \|P(x_0)\|^2$$

откуда

$$\delta_1^2 \leqslant (M_0^2 A^2 + M_0^2 B \bar{\delta}_0 - 1) \delta_0^2 \leqslant (M_0^2 A^2 + M_0^2 B \bar{\delta}_0 - 1) \bar{\delta}_0^2 = \bar{\delta}_1^2$$

Как и в доказательстве теоремы 1, получим

$$|(P'(x_1) h, h)| \geqslant \frac{1}{M_1} \|h\|^2, \quad h \in H$$

где

$$M_1 = \frac{M_0}{1 - M_0^2 B \bar{\delta}_0}$$

И здесь $S(x_1, r_1) \subset S(x_0, r)$, где $r_1 = \frac{1}{B}\left(\frac{1}{M_1} - \frac{1}{M^*}\right) \frac{\delta_1}{\bar{\delta}_1}$, поэтому оценки для $\|P'(x)\|$ и $\|P''(x)\|$ остаются в силе при переходе от x_0 к x_1.

Окончания доказательств теорем 3 и 1 полностью совпадают, и мы не будем здесь их повторять.

5. Обозначим

$$a_n = M_n^2 B \bar{\delta}_n, \qquad b_n = M_n^2 A^2 \tag{30}$$

Тогда в силу (27) и (28) получим соотношения

$$a_{n+1} = \frac{\sqrt{b_n + a_n - 1}}{(1 - a_n)^2} \, a_n \tag{31}$$

$$b_{n+1} = \frac{b_n}{(1 - a_n)^2} \tag{32}$$

О сходимости последовательностей $\{a_n\}$, $\{b_n\}$ имеем место

Теорема 4. *Если $a_0 < 1$ и существует такое вещественное число $k > 0$, что $\alpha = k^2 a_0 < 1$, а также выполнены неравенства:*

1° $b_0 + a_0 - 1 \leqslant \left(\frac{k}{k+1}\right)^4 (1 - a_0)^4$,

2° $\frac{k^4}{(k^2 - \alpha)^2}\left[\frac{k^2 - \alpha}{k^2} + \left(\frac{k^2 - \alpha}{k(k+1)}\right)^4\right] \leqslant \frac{(k+1)^2 - \alpha}{(k+1)^2} + \left(\frac{(k+1)^2 - \alpha}{(k+1)(k+2)}\right)^4$,

то последовательности $\{a_n\}$, $\{b_n\}$, вычисленные из (31), (32) (*как и последовательности* $\{M_n\}$, $\{\bar{\delta}_n\}$, *вычисленные из* (27), (28)), *сходятся.*

Доказательство аналогично доказательству теоремы 2.

6. Если в теоремах 1 и 3 условие 2° заменить более строгим условием (ср. [2], [3] и [1])

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

то имеют место теоремы 5 и 6.

Теорема 5. *Пусть выполнены условия:*

1° *для всех* $x \in S(x_0, r)$, *где* $r = \frac{M\delta_0}{1-q}$, *оператор* $P(x)$ *дважды дифференцируем и*

а) $\|P'(x)\| \leqslant A$, б) $\|P''(x)\| \leqslant B$,

в) $|(P'(x)h, h)| \geqslant \frac{1}{M}\|h\|^2$, $M > 0$, $h \in H$;

2° $q = \sqrt{M^2A^2 - 1} < 1$;

3° $M^2B\,\delta_0 \leqslant \frac{1}{2}$, $1 - \sqrt{1 - 2M^2B\,\delta_0} < 2\sqrt{M^2A^2 - 1}$.

Тогда уравнение (1) *имеет в сфере* $S(x_0, r)$ *решение* x^*, *к которому сходится последовательность* $\{x_n\}$, *полученная методом* $\{(2), (3)\}$. *Имеет место оценка*

$$\|x^* - x_n\| \leqslant M\,\delta_n \leqslant M\,\delta_0\,q^n$$

где $\delta_n \geqslant \|P(x_n)\|$ $(n = 0, 1, 2, \ldots)$.

Эту теорему в несколько ином виде, без требования дифференцируемости оператора, дал Ю. Лумисте [1]. Здесь уточнены область $S(x_0, r)$ и оценка погрешности. Эти уточнения легко получить, если иметь в виду доказательство теоремы 1.

Теорема 6. *Пусть выполнены условия:*

1° *для всех* $x \in S(x_0, r)$, *где* $r = \frac{M\delta_0}{1-q}$, *оператор* $P(x)$ *дважды дифференцируем и*

а) $\|P'(x)\| \leqslant A$, б) $\|P''(x)\| \leqslant B$,

в) $|(P'(x)h, h)| \geqslant \frac{1}{M}\|h\|^2$, $M > 0$, $h \in H$;

2° $q = \sqrt{M^2A^2 + M^2B\,\delta_0 - 1} < 1$.

Тогда уравнение (1) *имеет в сфере* $S(x_0, r)$ *решение* x^*, *к которому сходится последовательность* $\{x_n\}$, *полученная методом* (26). *Имеет место оценка*

$$\|x^* - x_n\| \leqslant M\,\delta_n \leqslant M\,\delta_0\,q^n$$

где $\delta_n \geqslant \|P(x_n)\|$ $(n = 0, 1, 2, \ldots)$.

7. Рассмотрим вопрос о единственности решения уравнения (1). Покажем, что имеет место

Теорема 7. *Пусть*

1° $|(P'(x_0)h, h)| \geqslant \frac{1}{M_0}\|h\|^2$, $M_0 > 0$, $h \in H$;

2° *в сфере* $S(x_0, r)$

$$\|P''(x)\| \leqslant B;$$

3° $M_0 B r < 1$.

Тогда уравнение (1) *имеет в сфере* $S(x_0, r)$ *не более одного решения.*

Доказательство. Пусть $x^* \in S(x_0, r)$ есть решение уравнения (1). Допустим, что уравнение (1) имеет в этой сфере еще другое решение x^{**}. Пользуясь условием 1°, получим, как в доказательстве теоремы 1,

$$|(P'(\bar{x})h, h)| \geqslant \frac{1}{M_0}(1 - M_0 B r)\|h\|^2 \tag{33}$$

если $\bar{x} \in S(x_0, r)$. Возьмем $h = x^* - x^{**}$. Тогда в силу (33)

$$0 = |(P(x^*) - P(x^{**}),\, x^* - x^{**})| =$$

$$= |(P'(\bar{x})(x^* - x^{**}), x^* - x^{**})| \geqslant \frac{1}{M_0}(1 - M_0 B r)\|x^* - x^{**}\|^2$$

где $\bar{x} = x^{**} + \tau(x^* - x^{**}) \in S(x_0, r)$, $0 < \tau < 1$. Так как $M_0 B r < 1$, то отсюда следует, что

$$x^* = x^{**}$$

Теорема доказана.

Еще легче доказать следующую теорему:

Теорема 8. *Если для всех* $x \in S(x_0, r)$

$$|(P'(x)h, h)| \geqslant \frac{1}{M}\|h\|^2, \qquad 0 < M < \infty\,, \quad h \in H \tag{34}$$

то уравнение (1) *имеет в сфере* $S(x_0, r)$ *не более одного решения.*

Следствие. Если существует такая постоянная M, $0 < M < \infty$, что условие (34) выполнено для всех $x \in H$, то уравнение (1) не может иметь в H более одного решения.

8. В качестве примера рассмотрим интегральное уравнение

$$P(x) = x(s) - \frac{1}{5}\int_0^{\frac{1}{2}} \sqrt{s+t}\cos x(t)\,dt = 0 \tag{35}$$

к которому применяем метод (26). Возьмем за начальное приближение $x_0(s) = 0$. В качестве пространства H выберем множество функций $L^2(0, \frac{1}{2})$ с интегрируемым квадратом. В рассматриваемом случае легко получить следующие оценки:

$$\|P'(x)\| \leqslant 1 + 0{,}05\sqrt{2} < 1{,}071 = A$$

$$\|P''(x)\| \leqslant 0{,}05\sqrt{2} < 0{,}07072 = B$$

Так как $(P'(x_0)h, h) = \|h\|^2$, то можем взять $M_0 = 1$. Вычисление $\|P(x_0)\|$ дает

$$\|P(x_0)\| < 0{,}049372 < 0{,}04938 = \bar{\delta}_0$$

Теперь

$$a_0 = M_0^2 B\,\bar{\delta}_0 = 0{,}003421, \qquad b_0 = M_0^2 A^2 = 1{,}147$$

Если выбрать $k=2$, то условия теоремы 4 выполнены. Поэтому последовательности $\{M_n\}$ и $\{\overline{\delta}_n\}$, полученные из (27) и (28), сходятся. Тогда выполнены и условия теоремы 3, откуда следует, что последовательность приближений $\{x_n\}$, вычисленная методом (26), сходится к решению уравнения (35).

Применение метода (26) дает следующие результаты:

$$\varepsilon_0 = -1$$

$$x_1(s) = -P(x_0(s)) = \frac{2}{15}\left(\sqrt{\left(s+\frac{1}{2}\right)^3} - \sqrt{s^3}\right)$$

$$\varepsilon_1 = -0{,}9949385$$

$$x_2(s) = 0{,}00067487\left(\sqrt{\left(s+\frac{1}{2}\right)^3} - \sqrt{s^3}\right) + 0{,}01177164\sqrt{s+0{,}02345504} +$$

$$+0{,}02376868\sqrt{s+0{,}11538267} + 0{,}02823129\sqrt{s+0{,}25000000} +$$

$$+0{,}02373579\sqrt{s+0{,}38461733} + 0{,}01174402\sqrt{s+0{,}47654496}.$$

Формула (29) дает следующие оценки погрешностей:

$$\|x^* - x_0\| < 6{,}6\cdot 10^{-2}$$
$$\|x^* - x_1\| < 1{,}7\cdot 10^{-4}$$
$$\|x^* - x_2\| < 2\cdot 10^{-8}$$

Здесь при вычислении $\varepsilon_1, x_2(s)$ и оценок погрешностей для x_1 и x_2 использована квадратурная формула Гаусса для случая пяти ординат. Погрешность, возникшая из-за пользования квадратурной формулой, не принята в расчет. Впрочем, она незначительна.

В рассматриваемом случае применима и теорема 6, так как (даже независимо от x)

$$|(P'(x)h, h)| \geqslant (1-0{,}05\sqrt{2})\|h\|^2$$

и мы можем взять $M = \dfrac{1}{1-0{,}05\sqrt{2}}$. Теперь

$$q = \sqrt{M^2A^2 + M^2B\delta_0 - 1} < 0{,}577 < 1$$

значит, условия теоремы 6 выполнены. Апостериорная оценка $\|x^* - x_n\| \leqslant M\,\delta_n$ теоремы 6 дает практически те же самые результаты, что и формула (29), но априорная оценка $\|x^* - x_n\| \leqslant M\,\delta_0\,q^n$ значительно грубее.

Теорема 7 гарантирует, что решение уравнения (35) единственно в каждой сфере $S(0,r)$, радиус r которой меньше $10\sqrt{2}$. Но, так как найдется такая постоянная $M = \dfrac{1}{1-0{,}05\sqrt{2}}$, для которой условие (34) теоремы 8 выполнено независимо от x, то уравнение (35) имеет во всем множестве $L^2\left(0,\dfrac{1}{2}\right)$ единственное решение.

ЛИТЕРАТУРА

1. Ю. Лумисте, Метод наискорейшего спуска при нелинейных уравнениях, Уч. зап. Тартуского гос. ун-та, № 37, 1955, 106—111.
2. Гуань Чжао-чжи, О методе наискорейшего спуска для решения нелинейных операторных уравнений, Шусюэ сюэбао (Acta math. sinica), **6**, **№ 4**, 1956, 638—650.
3. M. Altman, Connection between the method of steepest descent and Newton's method, Bull. Acad. Polon. Sci., Cl. III, **5**, No. 11, 1957, 1031—1036.
4. Л. В. Канторович, Функциональный анализ и прикладная математика, Успехи матем. наук, **3**, вып. 6, 1948, 89—185.
5. М. М. Вайнберг, Вариационные методы исследования нелинейных операторов, М., 1956.

Институт энергетики
Академии наук Эстонской ССР

Поступила в редакцию
2. VI 1959

KIIREIMA LANGUSE MEETODIST MITTELINEAARSETE VÕRRANDITE LAHENDAMISEKS

L. Kivistik

Resümee

Olgu $P(x)$ mittelineaarne, Fréchet' mõttes kaks korda diferentseeruv operaator Hilberti ruumist H samasse ruumi. Artiklis käsitletakse võrrandi

$$P(x)=0 \tag{1}$$

ligikaudset lahendamist kiireima languse meetodil, mille rakendamine seisneb võrrandi lähislahendite leidmises rekurrentsest seosest

$$x_{n+1}=x_n+\varepsilon_n P(x_n) \qquad (n=0,1,2\ldots) \tag{2}$$

kus x_0 on antud alglähend ja ε_n võrrandi

$$(P(x_n),\ P(x_n+\varepsilon P(x_n)))=0 \tag{3}$$

vähima absoluutväärtusega lahend. Tõestatakse teoreem:

Kui

1° $\| P(x_0)\| = \delta_0 \leqslant \bar{\delta}_0$;

2° $|(P'(x_0)h,\ h)| \geqslant \frac{1}{M_0}\|h\|^2, \qquad M_0>0,\ \ h\in H;$

3° $\|P'(x)\|\leqslant A,\ \|P''(x)\|\leqslant B$ *iga x puhul kinnisest sfäärist* $S(x_0,\ r)$, *kus*

$$r=\frac{1}{B}\left(\frac{1}{M_0}-\frac{1}{M^*}\right)\frac{\delta_0}{\bar{\delta}_0};$$

4° *konstandid* M_0, $\bar{\delta}_0$, A, B *on sellised, et koonduvad rekurrentsetest seostest*

$$M_{n+1}=\frac{M_n}{\sqrt{1-2M_n^2B\bar{\delta}_n)}},\qquad \bar{\delta}_{n+1}=\bar{\delta}_n\sqrt{\frac{4M_n^2A^2}{(1+\sqrt{1-2M_n^2B\bar{\delta}_n})^2}-1}$$

arvutatud jadad $\{M_n\}$ *ja* $\{\bar{\delta}_n\}$;

siis võrrandil (1) *on sfääris* $S(x_0,\ r)$ *lahend* x^*, *milleks koondub seosest* (2) *arvutatud lähislahendite jada* $\{x_n\}$ *kiirusega*

$$\|x^*-x_n\|\leqslant M^*\delta_n\leqslant\frac{\sqrt{2}}{A}\delta_n$$

kus $M^*=\lim M_n$ *ja* $\delta_n=\|P(x_n)\|$.

Edasi vaadeldakse meetodi üht modifikatsiooni, mille puhul ε_n arvutatakse valemi järgi

$$\varepsilon_n = -\frac{\| P(x_n) \|^2}{(P'(x_n)P(x_n),\ P(x_n))}$$

Artikli lõpuosas tuuakse kaks teoreemi lahendi ainsuse kohta ja lahendatakse üks näide.

Eesti NSV Teaduste Akadeemia
Energeetika Instituut

Saabus toimetusse
2. VI 1959

ON THE METHOD OF STEEPEST DESCENT FOR SOLVING NON-LINEAR EQUATIONS

L. Kivistik

Summary

Let H be a real Hilbert space and let $P(x)$ be a non-linear operator from H into H which is twice differentiable in the sense of Fréchet. The present paper considers the method of the steepest descent for solving the equation

$$P(x) = 0 \tag{1}$$

The method may be defined by the following formula:

$$x_{n+1} = x_n + \varepsilon_n P(x_n) \qquad (n=0,\ 1,\ 2,\ \ldots) \tag{2}$$

where x_0 is the initial approximate solution and ε_n is the solution (which has the smallest absolute value) of the equation

$$(P(x_n),\ P(x_n + \varepsilon P(x_n))) = 0 \tag{3}$$

The following theorem is proved:

If

1° $\| P(x_0) \| = \delta_0 \leqslant \overline{\delta}_0$;

2° $| (P'(x_0)h,\ h) | \geqslant \frac{1}{M_0} \| h \|^2, \quad M_0 > 0, \quad h \in H$;

3° $\| P'(x) \| \leqslant A,\ \| P''(x) \| \leqslant B$ *for every* $x \in S(x_0,\ r)$, *where* $S(x_0,\ r)$ *is the closed sphere with the radius* $r = \frac{1}{B}\left(\frac{1}{M_0} - \frac{1}{M^*}\right)\frac{\delta_0}{\overline{\delta}_0}$;

4° *the constants* M_0, $\overline{\delta}_0$, A *and* B *are such that the sequences* $\{M_n\}$ *and* $\{\overline{\delta}_n\}$, *defined by the recurrence formulae*

$$M_{n+1} = \frac{M_n}{\sqrt{1 - 2M_n^2 B\overline{\delta}_n}}, \qquad \overline{\delta}_{n+1} = \overline{\delta}_n \sqrt{\frac{4M_n^2A^2}{(1+\sqrt{1-2M_n^2B\overline{\delta}_n})^2} - 1}$$

converge;

then equation (1) has a solution x^* in the sphere $S(x_0, r)$ and the sequence of approximate solutions x_n, defined by $\{(2), (3)\}$, converges to x^*. For the error estimate we have the formula

$$\| x^* - x_n \| \leqslant M^* \delta_n \leqslant \frac{\sqrt{2}}{A} \delta_n$$

where $M^* = \lim M_n$ *and* $\delta_n = \| P(x_n) \|$.

A modification of the method of steepest descent is also considered. In this case ε_n is defined by the formula

$$\varepsilon_n = - \frac{\| P(x_n) \|^2}{(P'(x_n) P(x_n), P(x_n))}$$

At the end of the paper two theorems are given on the uniqueness of the solution, and an example is solved.

Academy of Sciences of the Estonian S.S.R.,
Institute of Energetics

Received
June 2nd, 1959