Доклады Академии Наук СССР
1947. Том LVI, № 3

МАТЕМАТИКА

Л. В. КАНТОРОВИЧ

О МЕТОДЕ НАИСКОРЕЙШЕГО СПУСКА

(Представлено академиком В. И. Смирновым 23 XI 1946)

В моей заметке „Об одном эффективном методе решения задачи о минимуме квадратичных функционалов" [1] был рассмотрен общий метод решения указанных задач, а в связи с этим и различного рода линейных и, отчасти, нелинейных проблем анализа. Значительные лакуны в журнальной литературе военного времени не позволили мне своевременно познакомиться и процитировать в указанной заметке некоторые англо-американские работы, относящиеся к тому же вопросу и предшествующие ей по времени или выполненные с ней одновременно*. В частности, укажем работу H. Curry [2], где рассматривается тот же метод в применении к решению нелинейных алгебраических систем, а также J. L. Synge [3] и R. Courant [4]**. Общая идея метода восходит еще к Коши [5].

Для случая систем линейных алгебраических уравнений этот метод развит в работе G. Temple [6]. В ней содержатся, в частности, результаты первого раздела моей заметки [1], относящиеся к алгебраическим системам, включая установление сходимости процесса, но без анализа быстроты ее. Там дана и форма последовательных приближений для функциональных уравнений, но в менее общих условиях, чем в моей заметке. Temple делает также попытку применения метода к дифференциальным и интегральным уравнениям 1-го рода, но используемый им процесс сходится при условии, выполнение которого для дифференциального оператора вообще исключено, а для интегрального имеет место только в случае вырожденного ядра. В предложенной мною форме процесса сходимость имеет место и для этих случаев при довольно общих условиях.

Здесь я даю дополнение к результатам заметки [1], а также привожу доказательства некоторых результатов.

1. О порядке сходимости. Рассмотрим сперва простейший случай, когда оператор A в вещественном пространстве Гильберта самосопряженный и имеет спектр в конечном промежутке:

$$m(x,\ x)\leqslant(Ax,\ x)\leqslant M(x,\ x);\quad 0<m<M<+\infty.$$

Последовательные приближения для уравнения

$$Lx=Ax-\varphi=0 \tag{1}$$

определяем формулами

$$x_{n+1}=x_n-\frac{(z_{n+1},\ z_{n+1})}{(Az_{n+1},\ z_{n+1})}z_{n+1};\quad z_{n+1}=Lx_n=$$
$$=Ax_n-\varphi\quad (n=0,\ 1,\ 2,\ldots). \tag{2}$$

* Данная работа связана с работами автора по экстремальным проблемам 1939 г. Результаты по ней были впервые доложены в Математическом институте АН СССР в сентябре 1943 г.

** С работой R. Courant'a автор не имел возможности ознакомиться и до настоящего времени.

Покажем, что x_n сходятся к решению уравнения (1) как геометрическая прогрессия с знаменателем $\left(\frac{M-m}{M+m}\right)$. Полагая

$$H(x)=(Ax,\ x)-2(\varphi,\ x) \tag{3}$$

и обозначая x^* решение уравнения (1) и $\eta_n=x_{n-1}-x^*$, имеем

$$\Delta H=H(x_{n-1})-H(x_n)=\frac{(z_n,\ z_n)^2}{(Az_n,\ z_n)}, \tag{4}$$

$$\Delta^* H=H(x_{n-1})-H(x^*)=(A\eta_n,\ \eta_n)=(z_n,\ A^{-1}z_n), \tag{5}$$

$$\frac{\Delta H}{\Delta^* H}=\frac{(z_n,\ z_n)^2}{(Az_n,\ z_n)\ (A^{-1}z_n,\ z_n)}\cong\frac{\left[\sum\limits_{\lambda}(\Delta e_\lambda z_n,\ z_n)\right]^2}{\sum_\lambda(\Delta e_\lambda z_n,\ z_n)\sum\lambda^{-1}(\Delta e_\lambda z_n,\ z_n)}, \tag{6}$$

если применить для оператора A спектральное разложение

$$(Az,\ z)=\int\limits_m^M\lambda(de_\lambda z,\ z)\cong\sum\lambda(\Delta e_\lambda z,\ z). \tag{7}$$

Применяя к отношению (6) неравенство для конечных сумм [7]

$$\frac{(\Sigma u_k)^2}{\Sigma\gamma_k u_k\Sigma\gamma_k^{-1}u_k}\geqslant\frac{4}{\left(\sqrt{\frac{M}{m}}+\sqrt{\frac{m}{M}}\right)^2}\quad(m\leqslant\gamma_k\leqslant M,\ u_k\geqslant 0), \tag{8}$$

находим

$$\frac{H(x_{n-1})-H(x_n)}{H(x_{n-1})-H(x^*)}\geqslant\frac{4}{\left(\sqrt{\frac{M}{m}}+\sqrt{\frac{m}{M}}\right)^2}, \tag{9}$$

$$H(x_n)-H(x^*)\leqslant\left(\frac{M-m}{M+m}\right)^2[H(x_{n-1})-H(x^*)], \tag{10}$$

$$\|x_{n-1}-x^*\|^2\leqslant\frac{1}{m}(A\eta_n,\ \eta_n)=$$

$$=\frac{1}{m}[H(x_{n-1})-H(x^*)]\leqslant\frac{1}{m}\left(\frac{M-m}{M+m}\right)^{2n-2}[H(x_0)-H(x^*)], \tag{11}$$

что и дает наше утверждение. Оно было сформулировано в заметке [1] для интегральных уравнений. Случай несамосопряженного уравнения сводится к рассмотренному, если заменить (1) на $A^*Ax=A^*\varphi$.

2. Сходимость в случае $m=0$. Если $m=0$, но 0 — изолированное собственное значение, то легко показать, что, если решение уравнения (1) существует, последовательность x_n сходится к некоторому его решению. Если $m=0$ не изолировано, верен более слабый результат, что если решение существует, то сходимость к нему имеет место в том смысле, что $\lim\limits_{n\to\infty}(A\eta_n,\ \eta_n)=0$, точнее $(A\eta_n,\ \eta_n)=O(1/n)$.

Действительно, исходя из геометрической интерпретации, данной в [1], а также у Synge [3], имеем $\|x_n-x^*\|\leqslant\|x_{n-1}-x^*\|$. Далее,

$$\|z_n\|\geqslant\frac{(z_n,\eta_n)}{\|\eta_n\|}\geqslant\frac{H(x_{n-1})-H(x^*)}{\|\eta_0\|},$$

$$H(x_{n-1})-H(x_n)=\frac{(z_n,\ z_n)^2}{(Az_n,\ z_n)}\geqslant\frac{\|z_n\|^4}{M\|z_n\|^2}\geqslant$$

$$\geqslant\frac{[H(x_{n-1})-H(x^*)]}{M\|x_0-x^*\|^2}\geqslant q[H(x_{n-1})-H(x^*)]^2;\quad q>0. \tag{12}$$

Или, если положить $r_n = H(x_{n-1}) - H(x^*) = (A\eta_n, \eta_n)$, то

$$r_n - r_{n+1} \geqslant q r_n^2, \qquad (13)$$

откуда $r_n \to 0$ $(r_n = O(1/n))$, что доказывает наше утверждение.

Этот результат устанавливает, в частности, сходимость данного метода при решении интегральных уравнений первого рода, если решение существует. При этом последовательные приближения сходятся в смысле метрики, определяемой ядром уравнения.

3. Случай неограниченного оператора. Пусть A — неограниченный оператор. Здесь процесс видоизменяется ([1]), вместо единичного принимается за основной некоторый „известный" полуограниченный оператор B; расстояние определяется, исходя от него.

Пусть B — оператор, заданный на плотном множестве Ω_0, и

$$xBx = (x, \mathrm{B}x) \geqslant (x, x). \qquad (14)$$

Форма $xBx' = (x, Bx')$ (на Ω_0) распространяется единственным образом, по Friedrichs'у ([7]), на область Ω_B элементов вида $x = \lim\limits_{n\to\infty} x_n$, где

$$\lim_{n,\, m\to\infty} (x_n - x_m) B (x_n - x_m) = 0 \quad (x_n \in \Omega_0).$$

Оператор B распространяется с сохранением самосопряженности на область Ω_B^* $(\Omega_0 \subset \Omega_\mathrm{B}^* \subset \Omega_\mathrm{B})$, причем совокупность его значений в Ω_B^* есть $\mathfrak{H}$, т. е. B^{-1} определено в $\mathfrak{H}$.

Относительно оператора A предположим, что он самосопряженный и ограничен относительно B в Ω_0, именно

$$m(\mathrm{B}x, x) \leqslant (Ax, x) \leqslant M(\mathrm{B}x, x) \quad (0 < m < M < +\infty). \qquad (15)$$

Определяя приближения процессом наискорейшего спуска, исходя из расстояния $\|x\|' = (x, Bx)$, получаем их в виде

$$x_n = x_{n-1} - \frac{z_n B z_n}{z_n A z_n} z_n, \qquad (16)$$

где z_n есть обобщенное решение уравнения $\mathrm{B}z_n = Ax_{n-1} - \varphi$, т. е. элемент Ω_B, для которого выполнено условие

$$uBz_n = uAx_{n-1} - (u, \varphi) \text{ для } u \in \Omega_0. \qquad (17)$$

Определенные этими формулами приближения x_n сходятся к решению x^*, причем не только $\|x_n - x^*\| \to \infty$, но и $\|x_n - x^*\| \to 0$.

В частности, для рассмотренного в ([1]) случая, когда

$$Au = (au_x)_x + (bu_y)_y - cu, \qquad (18)$$

а $\mathrm{B}u = \Delta u$, при решении уравнения $Lu = Au - f = 0$ ($u = 0$ на контуре), последовательные приближения определяются так:

$$u_n = u_{n-1} - \frac{\iint\limits_D [(z_n)_x^2 + (z_n)_y^2]\, dx\, dy}{\iint\limits_D [a(z_n)_x^2 + b(z_n)_y^2 + cz_n^2]\, dx\, dy} z_n(x, y), \qquad (19)$$

где $z_n(x, y)$ есть решение уравнения $\Delta z_n = L(u_{n-1})$, причем имеем в виду обобщенное его решение в смысле, аналогичном тому, как вводилось это понятие в работах С. Л. Соболева ([8]), К. Friedrichs'а ([9]) и H. Weyl'я ([10]). Точнее, z_n — функция, удовлетворяющая условию

$$\iint\limits_D [z_x v_x + z_y v_y]\, dx\, dy = \iint\limits_D [au_x v_x + bu_y v_y + cuv + fv]\, dx\, dy, \qquad (20)$$

где v — любая дважды непрерывно дифференцируемая функция, равная нулю на контуре. Из указанного выше результата вытекает возможность построения всех $u_n(uz_n)$ в соответствующем пространстве Ω_B, а также сходимость их к решению в смысле метрики, определяемой интегралом Дирихле.

4. Повышение быстроты сходимости процесса. Если определять приближения по формулам (2), то $(p+1)$-е имеет вид

$$x_{p+1} = x_0 + \alpha_0 z_1 + \alpha_1 A z_1 + \ldots + \alpha_p A^p z_1. \tag{21}$$

Если x разыскивать прямо в форме (21) с неопределенными $\alpha_0, \alpha_1, \ldots, \alpha_p$, находя их из условия минимума функционала H (3), то получим для него значение, не превосходящее $H(x_{p+1})$. При этом для определения $\alpha_0, \alpha_1, \ldots, \alpha_p$ придется решать систему уравнений

$$(A^k z_1, z_1) + (A^k z_1, A z_1)\alpha_0 + \ldots + \\ + (A^k z_1, A^{p+1} z_1)\alpha_p = 0 \quad (k = 0, 1, \ldots, p). \tag{22}$$

Один шаг такого процесса заменяет $(p+1)$ шагов первоначального. Применение такого видоизменения метода обеспечивает большую практическую эффективность его.

5. Нахождение собственных значений. Если A — ограниченный снизу оператор, то наименьшее собственное значение его или нижняя граница спектра) есть минимум формы

$$L(x) = \frac{(Ax, x)}{(x, x)}. \tag{23}$$

Для нахождения последнего применяем метод наискорейшего спуска. Начав с значения x_0, можем считать $(x_0, x_0) = 1$, рассматриваем варьированное значение $L = L_0 + \Delta L = L(x_0 + \varepsilon z)$. Находим:

$$\frac{d}{d\varepsilon}(L_0 + \Delta L)|_{\varepsilon=0} = 2\,[(Ax_0, z) - (Ax_0, x_0)(x_0, z)]. \tag{24}$$

Отсюда ясно, что значение для z, соответствующее направлению градиента, есть $z_1 = k[Ax_0 - (Ax_0, x_0)x_0]$, где k определяется условием $(z_1, z_1) = 1$. При таком выборе z оказывается, что наивыгоднейшее значение ε есть ε_1 — отрицательный корень уравнения

$$(Ax_0, z_1) + \varepsilon\,[(Az_1, z_1) - (Ax_0, x_0)] - \varepsilon^2 (Ax_0, z_1) = 0. \tag{25}$$

Тогда новым приближением для собственного элемента будет $x_1 = x_0 + \varepsilon_1 z_1$, а для собственного значения

$$L_1 = L_0 + \Delta L = L_0 + \varepsilon_1 (Ax_0, z_1). \tag{26}$$

Если оператор A — положительно определенный, нижняя граница спектра есть изолированное собственное значение λ_1 и нижняя граница остального спектра есть λ_2, то быстрота приближения к минимуму характеризуется прогрессией с знаменателем $q = \left(\frac{\lambda_1}{2\lambda_2 - \lambda_1}\right)^2$.

Математический институт
Академии Наук СССР им. В. А. Стеклова
Ленинградское отделение

Поступило
23 XI 1946

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Л. В. Канторович, ДАН, **48**, № 7 (1945). [2] H. Curry, Quarterly of Appl. Mathem., **2**, No. 3, 251 (1944). [3] J. L. Synge, ibid. **2**, No. I, 87 (1944). [4] R. Courant, Bull. Am. Math. Soc., 49, 1 (1943). [5] A. Cauchy, C. R., **25**, 536 (1847). [6] G. Temple, Proc. Roy. Soc. London (A), **169**, 476 (1939). [7] K. Friedrichs, Math. Ann., **109**. [8] С. Соболев, Тр. Матем. ин-та им. В. А. Стеклова, **9** (1934). [9] K. Friedrichs, Am. J. Math., **61**, No. 2 (1939). [10] H. Weyl, Duke Mathem. J., **7**, 411 (1940).