СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
Дубна

Р5 - 5137

ЛАБОРАТОРИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ
И АВТОМАТИЗАЦИИ

Л. Александров

РЕГУЛЯРИЗОВАННЫЕ ИТЕРАЦИОННЫЕ ПРОЦЕССЫ ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНЫХ ОПЕРАТОРНЫХ УРАВНЕНИЙ

1970

Р5 - 5137

Л. Александров

РЕГУЛЯРИЗОВАННЫЕ ИТЕРАЦИОННЫЕ ПРОЦЕССЫ ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНЫХ ОПЕРАТОРНЫХ УРАВНЕНИЙ

В работах /1,2,3,4/ изучается вопрос о решении нелинейного операторного уравнения

$$f(x) = \theta \tag{1}$$

методом Ньютона-Канторовича без предположения обратимости оператора $f'(x)$ в начальной точке x_0 или в некоторой сфере вокруг нее.

В работе /4/ предлагается общий подход построения итерационных процессов указанного класса и рассматривается вопрос о сходимости введенных там "регуляризованных процессов Ньютона-Канторовича" в L-суперметрических пространствах.

В настоящей работе развиваются идеи работы /4/. В частности, вводится понятие о регуляризованных итерационных процессах для решения уравнения (1). Для этой цели предварительно определяется " ϵ -квазиобратный" оператор данного оператора. В гильбертовом пространстве изучается вопрос о сходимости одного частного случая регуляризованного итерационного процесса (процесс $\Re_{\epsilon_n}$), характерного своими конструктивными свойствами.

§1. Оператор, " ϵ -квазиобратный" данному оператору

1. Пусть S_1 и S_2 - линейные метрические пространства, A - оператор, отображающий пространство S_1 в S_2.

В дальнейшем будем пользоваться следующим понятием.

Определение 1.

Пусть $\epsilon \geq 0$ - заданное действительное число. Оператор B, действующий из S_2 в S_1, для которого выполнено неравенство

$$\| A - ABA \| \le \epsilon, \tag{2}$$

называется " ϵ -квазиобратным" оператору A. Будем употреблять обозначение $B = A^{\underline{\epsilon}}$.

2. Рассмотрим некоторые примеры ϵ -квазиобратных операторов.

Пример 1. Левые обратные операторы A_{ℓ}^{-1}, правые обратные операторы A_r^{-1} или обратный оператор A^{-1} (если они существуют) являются "нуль-квазиобратными" ($\epsilon = 0$) оператору A.

Пример 2. В метрических пространствах S_1 и S_2 по аналогии с работой /6/ (стр. 39) рассмотрим " g -обратный" оператор данному линейному оператору A, как такой оператор A^- из S_2 в S_1, для которого справедливо равенство

$$A A^- A = A .$$

g - обратные операторы являются нуль-квазиобратными ($A^- = A^{\underline{0}}$).

Пример 3. По аналогии с работой /7/ в гильбертовом пространстве S рассмотрим оператор A^+ Мура-Пенроуза, псевдообратный линейному оператору A, как оператор из S в S, для которого справедливы равенства

$$AA^+A = A, \quad A^+AA^+ = A^+,$$

$$A^+A = A^* A^{+*}, \quad AA^+ = A^{+*}A^* .$$

Оператор A^+ является нуль-квазиобратным ($A^+ = A^{\underline{0}}$).

Пример 4. В гильбертовом пространстве S данному линейному оператору A сопоставим оператор $h = A^*A$. Рассмотрим оператор $d_\omega = (h + \omega)^{-1}$, где линейный оператор ω действует из $D(h)$ в $R(h)$. Очевидно, что оператор $d\omega$ всегда существует, если линейный оператор ω положительно определенный. Имеет место следующее утверждение об обратимости оператора h.

Теорема 1.

Пусть $\bar{\epsilon} > 0$ - фиксированное действительное число. Тогда для любого положительно определенного оператора $\omega_{\bar{\epsilon}}$, для которого

$$\| \omega_{\bar{\epsilon}} \| \le \bar{\epsilon}, \tag{3}$$

оператор $d_{\omega_{\bar{\epsilon}}}$ является $\bar{\epsilon}$ -квазиобратным оператору h.

Доказательство.

Рассмотрим тождество

$$\omega = A^*A + \omega - A^*A = (A^*A + \omega)(E - (A^*A + \omega)^{-1} A^*A). \tag{4}$$

Так как операторы ω и $A^*A + \omega$ положительно определенные, то из (4) следует, что оператор $E - (A^*A + \omega)^{-1} A^*A$ неотрицательно определенный. Из этого следует неравенство

$$\| A^*A\,(E - (A^*A + \omega)^{-1} A^*A) \| \le$$

$$\le \| (A^*A + \omega)(E - (A^*A + \omega)^{-1} A^*A) \| . \tag{5}$$

Используя соотношения (3), (4), (5) и определение I, получаем утверждение теоремы 1.

§2. Регуляризованные итерационные процессы

1. Уравнение (1) рассматриваем в полном L -суперметрическом /5/ (стр. 263) пространстве S. Пусть приближенное решение уравнения (1) находится итерационным процессом $\mathcal{P}$, в n -той итерации которого выполняются k_n обращений k_n операторов $A_{n1}, \dots, A_{nk_n}$.

$$\mathcal{P}[x_0; f; A_{n1}^{-1}, \dots, A_{nk_n}^{-1}] :$$

$$x_0, x_{n+1} = F_n [x_n; A_{n1}^{-1}, \dots, A_{nk_n}^{-1}],$$

где $x_n \in S$, а операторы F_n; $A_{n1},\dots,A_{nk_n}$ действуют из S в S $(n=0,1,2,\dots)$.

Введем следующее основное понятие.

Определение 2.

Итерационный процесс $\mathfrak{R}[x_0; f; A_{n1}^{\epsilon_{n1}},\dots,A_{nk_n}^{\epsilon_{nk_n}}]$, полученный из процесса $\mathcal{P}$, после замены операторов $A_{n1}^{-1},\dots,A_{nk_n}^{-1}$ ϵ-квази-обратными $A_{n1}^{\epsilon_{n1}},\dots,A_{nk_n}^{\epsilon_{nk_n}}$ называется "регуляризованным итерационным процессом" для решения уравнения (1).

Очевидно, процесс $\mathfrak{R}$ может иметь смысл и в случае, когда хотя бы один из операторов $A_{n1},\dots,A_{nk_n}$ необратим, т.е. тогда, когда первоначальный процесс $\mathcal{P}$ не имеет смысла.

2. Исходя из классических итерационных процессов типа $\mathcal{P}$ можно получить ряд важных регуляризованных процессов для решения уравнения (1). Так, например, в работе /4/ в L-суперметрических пространствах изучаются "регуляризованные вычислительные процессы Ньютона-Канторовича", полученные на основе обычного метода Ньютона. В работе /8/ (в конечномерном евклидовом пространстве) используется регуляризованный итерационный процесс для анализа экспоненциальных зависимостей. Этот процесс может служить примером регуляризованного процесса, где в каждой итерации осуществляется больше, чем одно ϵ-квазиобращение линейных операторов.

3. Пусть S -гильбертово пространство; f -оператор, отображающий выпуклое множество $X \subseteq S$ в S, и пусть существуют первая и вторая производные Фреше $f'(x)$ и $f''(x)$ в X.

Дальнейшей целью настоящей работы является изучение следующего итерационного процесса

$$\mathfrak{R}_{\epsilon_n}: x_0,\ x_{n+1} = x_n - (f'^*(x_n) f'(x_n) + \omega_n)^{-1} f'^*(x_n) f(x_n),$$
$$x_n \in X \quad (n = 0,1,2,\dots),$$

где линейные операторы ω_n, действующие из X в S, - положительно определенные (для всех конечных n) и такие, что при заданных действительных числах $\epsilon_n > 0$ (для всех конечных n) имеют место неравенства

$$\|\omega_n\| \le \epsilon_n. \qquad (6)$$

На основе теоремы 1 и неравенств (6) заключаем, что для всех конечных n имеют место равенства

$$(f'^*(x_n) f'(x_n) + \omega_n)^{-1} = (f'^*(x_n) f'(x_n))^{\epsilon_n}$$

и, следовательно, любая конечная часть процесса $\mathfrak{R}_{\epsilon_n}$ является регуляризованным итерационным процессом.

§3. Сходимость процесса $\mathfrak{R}_{\epsilon_n}$

1. Введем обозначения: $g_n = f'^*(x_n) f(x_n)$,
$\rho_n = \|g_n\|$, $h_n = f'^*(x_n) f'(x_n)$, $\tau_n = \|f'(x_n)\|$,
$b_n = \|(h_n + \omega_n)^{-1}\|$.

Предварительно сформулируем некоторые условия, связанные с процессом $\mathfrak{R}_{\epsilon_n}$.

А) Первая и вторая производные Фреше $f'(x)$, $f''(x)$ ограничены в X:

$$\|f'(x)\| \le T < \infty, \quad \frac{1}{2}\|f''(x)\| \le M < \infty. \qquad (A1)$$

Существует константа $P \in [0,\infty)$ [или константа $Q \in [0,1)$] такая, что для любых точек $x, y \in X$ и любой точки $z \in \{tx + (1-t)y / t \in [0,1]\}$ (X - выпуклое) выполнено неравенство

$$\|(f'^*(x) - f'^*(y)) f(z)\| \le u(x,y), \qquad (A2)$$

где для функционала u имеем
$u(x,y) = P\|x-y\|^2$ [или $u(x,y) = Q\|f'^*(z) f(z)\|$].

Б) Для начальной точки x_0, начального оператора ω_0 регуляризованного процесса $\Re_{\epsilon_n}$ выполнены неравенства (начальные условия)

$$0 \le \phi_0 \le 1-\beta < 1, \quad \beta < 1, \qquad (Б1)$$

$$\beta + \frac{\phi_0}{1-\phi_0}\,\frac{\gamma^2}{(\gamma-b_0)b_0} \le 1, \qquad (Б2)$$

где $\gamma > b_0$ – заданное действительное число,

$$\nu = (MT+P+N_0)\gamma^2 \quad \left[\text{ или } \nu = \frac{MT+N_0}{1-Q}\gamma^2 \right] \qquad (Б3)$$

$$\beta = \nu\rho_0, \; \phi_0 = (2MT+N_0)\gamma b_0\rho_0,$$

$N_0 = \frac{a_1}{\rho_0}(\|\omega_0\|^2 + 2r_0\|\omega_0\|)$ и $a_1 \in [0,1]$ – заданное действительное число.

В) Для операторов ω_n процесса $\Re_{\epsilon_n}$ выполняются неравенства

$$\|\omega_n - \omega_{n+1}\| \le \|\omega_n\| \le \epsilon_n \quad (n=0,1,2,\dots), \qquad (В1)$$

где

$$\epsilon_n = a_2(\sqrt{r_n^2 + N_0\rho_n} - r_n)$$

и $a_2 \in (0,1]$ – заданное действительное число.

Г) $W = \{x / \|x_1 - x\| \le \frac{\gamma\beta\rho_0}{1-\beta^2}\} \subseteq X$.

Очевидно, что для достаточно малых по величине ρ_0 и $\|\omega_0\|$ всегда существует оператор $\omega_0 \ne \theta$, для которого начальные условия (неравенства (Б1) и (Б2)) выполняются.

Имеет место следующее основное утверждение о сходимости процесса $\Re_{\epsilon_n}$.

Теорема 2

Пусть для процесса $\Re_{\epsilon_n}$ выполнены условия А)– Г). Тогда процесс $\Re_{\epsilon_n}$ сходится, $\lim_{n\to\infty} x_n = x_\omega \in W$. Точка x_ω является решением уравнения

$$f'^*(x)f(x) = \theta \qquad (7)$$

в области X. Справедлива оценка

$$\|x_\omega - x_n\| < \frac{\gamma\rho_0\beta^{2^n-1}}{1-\beta^{2^n}} \quad (n=1,2,\dots).$$

Доказательство.

Для процесса $\Re_{\epsilon_n}$ имеем

$$\|x_n - x_{n+1}\| = \|(h_n+\omega_n)^{-1}g_n\| \le b_n\rho_n. \quad (n=0,1,2,\dots). \qquad (8)$$

Из неравенств (В1) следуют неравенства

$$(\|\omega_n\| - \epsilon_n)(\|\omega_n\| + a_2(\sqrt{r_n^2 + N_0\rho_n} + r_n)) \le 0,$$

$$2r_n a_2 \|\omega_n\| + \|\omega_n\|^2 \le a_2^2 N_0 \rho_n, \qquad (9)$$

$$\|\omega_n\|\|h_n+\omega_n\| \le N_0\rho_n \le N_0 c_n\rho_n \quad (n=0,1,2,\dots),$$

где

$$c_n = \|(h_n+\omega_n)^{-1}\|\,\|h_n+\omega_n\| \ge 1$$

– число обусловленности оператора $h_n + \omega_n$ /5/ (стр. 105).

Из неравенства (9) получаем окончательно

$$\|\omega_n\| \le N_0 b_n\rho_n \quad (n=0,1,2,\dots). \qquad (10)$$

Исходя из этих неравенств, находим

$$\|\omega_n\| \, \|(h_n+\omega_n)^{-1} g_n\| \le N_0 b_n^2 \rho_n^2$$

и, используя обозначения

$$\delta_n = x_{n+1} - x_n = -(h_n+\omega_n)^{-1} g_n, \tag{11}$$

можем записать

$$\|\omega_n \delta_n\| \le N_0 b_n^2 \rho_n^2 \qquad (n=0,1,2,\dots). \tag{12}$$

Из (11) следуют соотношения

$$\rho_{n+1} = \|g_{n+1} - g_n - (h_n+\omega_n)\delta_n\| \le$$

$$\|(f'^*(x_{n+1}) - f'^*(x_n)) f(x_{n+1})\| +$$

$$\|f'^*(x_n)\| \, \|f(x_{n+1}) - f(x_n) - f'(x_n) r_n\| + \|\omega_n \delta_n\|.$$

На основе формулы Тейлора и неравенства (А2) и (8) получаем неравенства

$$\rho_{n+1} \le (MT + P + N_0) b_n^2 \rho_n^2 \quad (n=0,1,2,\dots) \tag{13}$$

[или неравенства

$$\rho_{n+1} \le \frac{MT+N_0}{1-Q} b_n^2 \rho_n^2 \quad (n=0,1,2\dots) \tag{14}$$].

С помощью полной математической индукции доказываются следующие утверждения

a) $\rho_n \le \rho_0 \beta^{2^n-1} < \rho_0$;

b) $\phi_n \le \phi_0 < 1$, где $\phi_n = (2MT + N_0)\gamma b_0 \rho_0$;

c) $b_n \le 1 - (\gamma - b_0)\beta^{2^n-1} \le \gamma$;

d) $x_{n+1} \in X$.

Дальнейшее доказательство теоремы с необходимыми изменениями проводится по схеме доказательства теоремы 3 в работе /4/. Приведем здесь только доказательство утверждения b).

Для $n=0$ b) верно в силу (Б1).

Пусть b) верно для n . Рассматриваем линейный оператор

$$p = E - (h_n+\omega_n)^{-1}(h_{n+1}+\omega_{n+1}).$$

Для его нормы имеем

$$\|p\| \le b_n \|h_n - h_{n+1}\| + b_n \|\omega_n - \omega_{n+1}\| \le$$

$$\le b_n \|f'^*(x_n) - f'^*(x_{n+1})\| \, \|f'(x_n)\| +$$

$$+ b_n \|f'(x_n) - f'(x_{n+1})\| \, \|f'^*(x_{n+1})\| + b_n \|\omega_n - \omega_{n+1}\|.$$

На основе формулы Тейлора и неравенства (10), (В1) и b) получаем

$$\|p\| \le (2MT + N_0)\gamma b_n \rho_n = \phi_n < 1,$$

и, следовательно, ряд $\sum_{i=0}^{\infty} p^i v$ сходится для всех $v \in X$. Отсюда заключаем, что $\|(E-p)^{-1}\| < \frac{1}{1-\phi_n}$. В результате имеем

$$b_{n+1} = \|(E-p)^{-1}(h_n+\omega_n)^{-1}\| \le b_n \frac{1}{1-\phi_n}.$$

Из соотношений (13) [или (14)] и обозначений (Б3) получаем

$$\phi_{n+1} = (2MT+N_0)\gamma b_{n+1} \rho_{n+1} \le$$

$$\le (2MT+N_0)\gamma \frac{b_n}{1-\phi_n} \nu \rho_n^2 = \frac{\phi_n}{1-\phi_n} \nu \rho_n ,$$

откуда с учетом a) и b) и того факта, что $\beta \le 1 - \phi_0$, находим окончательно

$$\phi_{n+1} \le \frac{\phi_n}{1-\phi_n} \nu \rho_0 = \frac{\phi_n}{1-\phi_n} \beta \le \phi_n < 1,$$

т.е. b) верно и для $n+1$.

§ 4. О применении и реализации регуляризованных процессов

1. Процесс $\Re_{\epsilon_n}$ можно использовать для построения приближенного решения уравнения (1); если линейный оператор $f'^*(x_n)$ – мономорфизм, то уравнение (1) эквивалентно уравнению (7). В общем случае все решения уравнения (1) являются решениями (7), что также позволяет находить решения уравнения (1) посредством процесса $\Re_{\epsilon_n}$

2. Опишем одну из возможных схем применения процесса $\Re_{\epsilon_n}$ и теоремы 2:

при заданных x_0 и ω_0 вычисляется константа N_0 и проверяются начальные условия (Б1) и (Б2) при фиксированных числах a_1 и a_2. В случае выполняемости неравенств (Б1), (Б2), а также условия $W \subseteq X$ строится процесс $\Re_{\epsilon_n}$ с операторами $\omega_n = \epsilon_n E$; затем проверяется выполнение неравенства $\|\omega_n - \omega_{n+1}\| \le \|\omega_n\|$. Очевидно, что неравенства $\|\omega_n\| \le \epsilon_n$ выполнены. Другая возможность заключается в выборе операторов ω_n в виде $\omega_n = a e^{-bn} E$, $a > 0$, $b \ge 0$. Тогда проверяется второе из неравенств (В1), а первое всегда выполнено. Процесс $\Re_{\epsilon_n}$ останавливается по критерию минимума невязки. В некоторых случаях процесс $\Re_{\epsilon_n}$ нельзя продолжать слишком долго, так как возможно сильное увеличение числа c_n. Величину числа c_n тоже можно использовать как характеристику поведения регуляризованного процесса $\Re_{\epsilon_n}$.

3. В работах /9,10/ для решения реальных задач (нахождения скрытых гауссианов из экспериментально измеренных кривых) используется введенный автором процесс $\Re_{\epsilon_n}$. В этих работах приводятся некоторые эвристические алгоритмы для реализации процесса $\Re_{\epsilon_n}$, а также численные примеры. В работе /8/ используется специфический регуляризованный итерационный процесс для анализа экспоненциальных зависимостей. На основе этого процесса развита программа SHE для ЭВМ. Приводятся численные примеры для решения модельных и реальных задач (например, определения времени термализации кристаллических модераторов из анализа экспериментальной кривой).

Литература

1. A. Ben-Israel. J. Math. Anal. and Appl., 1966, 15, N2, 243-252.
2. M. Altman. Bull. Acad. Polon. Sci., 1967, 5, N8, 789-795.
3. Л. Раковщик. ЖВМ и МФ, т. 8, №6, 1202-1217, 1968.
4. Л. Александров. Препринт ОИЯИ Р5-5136, Дубна, 1970.
5. Л. Коллатец. Функциональный анализ и вычислительная математика. М., 1968.
6. С.Р. Рао. Линейные статистические методы и их применения. М., 1968.
7. R. Penrose. A Generalized of Matrices. Proc. Comb. Soc., 51, 3, 1965, 406-413.
8. Л. Александров. Совещание по программированию и вычислительные методы. ОИЯИ, Дубна, май, 1969.
9. В. Гаджоков, Л. Александров, Ч. Стоянов. Препринт ОИЯИ Р10-5043, Дубна, 1970.
10. В. Гаджоков. Препринт ОИЯИ Р10-5035, Дубна, 1970.

Рукопись поступила в издательский отдел
25 мая 1970 года.