7258

СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

ДУБНА

Р5 - 7258

Л.Александров

О РЕГУЛЯРИЗАЦИИ ИТЕРАЦИОННЫХ ПРОЦЕССОВ НЬЮТОНОВСКОГО ТИПА

1973

ЛАБОРАТОРИЯ ВЫЧИСЛИТЕЛЬНОЙ
ТЕХНИКИ И АВТОМАТИЗАЦИИ

P5 - 7258

Л.Александров

О РЕГУЛЯРИЗАЦИИ ИТЕРАЦИОННЫХ ПРОЦЕССОВ НЬЮТОНОВСКОГО ТИПА

Регуляризованные итерационные процессы (R -процессы)[I-6] все чаще применяются в вычислительной практике на ЭВМ при построении надежных алгоритмов для решения нелинейных задач[3,7-II].

Вычислительный опыт на ЭВМ показывает, что R -процессы обладают рядом преимуществ по сравнению с аналогичными методами без регуляризации[7,8,I2].

Дальнейшее изучение R -процессов способствует решению важной проблемы полной автоматизации исследования нелинейных задач при помощи ЭВМ.

В настоящей работе рассматривается общая ("аддитивно-мультипликативная") регуляризация процессов ньютоновского типа. Даются утверждения о сходимости и существовании относительно итерационной схемы, охватывающей главные подклассы, R -процессов ньютоновского типа. Приводятся следствия этих утверждений для R -процессов Шредера (третьего порядка) и R -процессов Гаусса-Ньютона.

§ I. Некоторые основные понятия

I. Пусть X и Y - банаховы пространства, а f - непрерывный нелинейный оператор, преобразующий выпуклую область $D(f)$ в $R(f) \subseteq Y$.

Пусть задано операторное уравнение

$$f(x) = y \qquad (x \in D,\ y \in R). \qquad (f)$$

Вместо решения уравнения (f) часто приходится решать вспомогательные (линейно-преобразованные) уравнения следующих двух типов

$$\varphi^R f x = \varphi^R y \qquad \text{или} \qquad (\varphi^R f)$$

$$\varphi^D f x = \varphi^D y \qquad (\varphi^D f).$$

Здесь $\varphi^R(x)$, $\varphi^D(x)$, $x \in D$ – ограниченные линейные операторы, действующие из R в R и из R в D, соответственно. $\varphi^R(x)$ и $\varphi^D(x)$ предполагаются непрерывно зависящими от x.

Пусть решение уравнения (f)отыскивается на основе одношагового итерационного процесса

$$x_0,\ x_{n+1} = x_n - a_n \varphi_n (f x_n - y)$$

$$x_n \in D \qquad (n = 0, 1, \ldots) \qquad (\mathrm{I})$$

где a_n – линейные операторы из R в D – в случае применения вспомогательного уравнения ($\varphi^D f$), или линейные операторы из R в R – в случае применения вспомогательного уравнения ($\varphi^R f$). (Операторы φ_n в (I) для этих двух случаев понимаются как $\varphi_n^R (= \varphi^R(x_n))$ и $\varphi_n^D (= \varphi^D(x_n))$, соответственно).

Отметим некоторые общие свойства процесса (I).

Будем говорить, что процесс (I) обладает свойством:

с а м о и с п р а в л е н и е (или "сходимость") в случае, когда имеет место соотношение

$$\lim_{n \to \infty} \| a_n \varphi_n (f x_n - y) \| = 0 \qquad (2)$$

(в этом случае, очевидно, в силе $\lim_{n \to \infty} \| x_{n+1} - x_n \| = 0$);

п р и т я ж е н и е – в том случае, когда для процесса (I) в силе как (2), так и соотношение

$$\lim_{n \to \infty} \| \varphi_n (f x_n - y) \| = 0 \ ; \qquad (3)$$

н а с т о я щ е е п р и т я ж е н и е – в том случае, когда для процесса (I) в силе как (2), так и соотношение

$$\lim_{n \to \infty} \| f x_n - y \| = 0 \ ;$$

л о ж н о е п р и т я ж е н и е – в том случае, когда для процесса (I) вместе с (2) и (3) имеет соотношение

$$\| f x^* - y \| \neq 0 \ ,$$

где $x^* = \lim_{n \to \infty} x_n$;

о т т а л к и в а н и е п о f ("отталкивание по a " или "отталкивание по φ ") – в том случае, когда для процесса (I) в силе соотношение (2) и в то же время последовательность $\{\| f x_n - y \|\}_{n=0}^{\infty}$ ($\{\|a_n\|\}_{n=0}^{\infty}$, или последовательность $\{\| \varphi_n \|\}_{n=0}^{\infty}$) неограниченно растет, начиная с некоторого номера n .

<u>Замечание I.</u>

Отталкивание по φ встречается в практике на ЭВМ. Этот случай приводится как расширение указанных основных случаев применения вспомогательных уравнений ($\varphi^R f$) и ($\varphi^D f$).

Следующее понятие является основным в работе:

Под <u>регуляризацией процесса (I) на основе заданной числовой последовательности $A = \{A_n > 0\}_{n=0}^{\infty}$</u> понимается построение нового процесса типа (I), в котором операторы a_n $(n=0,1,\dots)$ заменяются подходяще выбранными операторами $\tilde{a}_n$ $(n=0,1,\dots)$, такими, для которых в силе ограничения

$$\|\tilde{a}_n\| \le A_n \qquad (n=0,1,\dots).$$

Число $\gamma = \max\limits_{n=0,1,\dots}\{A_n\}$ называется "уровнем регуляризации". Для регуляризованного на основе последовательности A процесса (I) в дальнейшем применяется обозначение " R_A -процесс", а для R -процесса (I), имеющего уровень регуляризации γ – применяется обозначение " R_γ -процесс".

Замечание 2.

Приведенное здесь понятие о регуляризации процесса (I) является частным случаем общего понятия о регуляризации итерационных процессов, содержащих линейные шаги, которое дается на основе " ε -квазиобратимости" /I/.

2. Пусть заданы некоторые линейные оператор-функции

$$u(x): D \to R;\quad v(x): D \to D;\quad z^R(x): D \to R;$$
$$z^D(x): D \to D;\quad w^D(x): R \to D;\quad w^R(x): R \to R$$

(в общем случае зависящие от $x \in D$).

Пусть еще предполагается существование обыкновенных обращений $(v(x))^{-1}$, $(w^R(x)\,u(x) + z^R(x))^{-1}$ и $(w^D(x)\,u(x) + z^D(x))^{-1}$.

В дальнейшем рассматриваются два основных подкласса процессов типа (I), получаемых из (I) после замены a_n операторами

$$a_n^D = v_n\,(w_n^D u_n + z_n^D)^{-1} \qquad (1a^D)$$

$$a_n^R = v_n\,(w_n^R u_n + z_n^R)^{-1}. \qquad (1a^R)$$

(Здесь используются обозначения: $v_n = v(x_n)$, $w_n^D = w^D(x_n)$, $w_n^R = w^R(x_n)$, $z_n^D = z^D(x_n)$ и $z_n^R = z^R(x_n)$).

Если в процесс ($1a^R$) подставлены следующие конкретные операторы

$$v_n = E^D,\ w_n^R = E^R,\quad z_n^R = \theta \quad \text{и} \quad u_n = f'(x_n)$$

(в этом случае в D предполагается существование производной $f'(x)$), то из него получается обыкновенный процесс Ньютона-Канторовича.

Если в процесс ($1a^D$) (в случае X и Y -координатные евклидовы пространства) подставлены операторы $v_n = E^D$, $w_n = \varphi_n = f'^T(x_n)$ (где $f'^T(x_n)$ – транспонированная матрица Якоби $f'(x_n)$), $z_n^D = \theta$ и $u_n = f'(x_n)$, то получается широко применяемый в вычислительной практике на ЭВМ процесс Гаусса-Ньютона. Далее, из обобщенных процессов ($1a^D$) и ($1a^R$) при подходящем подборе операторов v_n, w_n, φ_n, u_n и z_n можно получить все случаи рассматриваемых до сих пор R -процессов ньютоновского типа.

Как видно, операторы u_n, w_n и φ_n служат для задания основной стратегии /I3/ итерационных процессов ($1a^D$) и ($1a^R$). В то же время операторы z_n и v_n (n =0,1,....) играют вспомогательную роль – для исправления применяемых основных стратегий. Прибавление операторов z_n и v_n к классическим стратегиям обусловлено вычислительным опытом /I3-I5,3,9,I0/.

В настоящей работе операторы z_n и v_n применяются для проведения определенной выше регуляризации процессов ($1a^R$) и ($1a^D$).

В связи со способом участия операторов z_n и v_n в процессах ($1a^R$) и ($1a^D$) различаются два вида регуляризации:

"аддитивная регуляризация" ("a-регуляризация") – в случае применения ненулевых операторов z_n (n=0,I,...) и мультипликативная регуляризация ("μ-регуляризация") – в случае применения неединичных операторов v_n(n =0,I,..). a-и μ -регуляризации применяются в основном для достижения двух различных вычислительных эффектов при реализации итерационных процессов: a -регуляризация – для избежания плохой обусловленности операторов $w_n u_n$(n =0,I...); μ-регуляризация – для расширения области сходимости процесса. Однако вычислительный опыт показывает, что a – регуляризация имеет более широкое значение.

Так, например, присутствие ненулевых операторов z_n в процессах ньютоновского типа может облегчить реализацию процессов в том случае, когда операторы-производные заменены численными аппроксимациями. Обоснование сходимости R -процессов в этом случае (и нахождение законов поведения операторов z_n) рассматривается как обоснование получаемых "грубых процессов" ньютоновского типа.

Другой пример: в процессах Гаусса-Ньютона операторы z_n (подходяще выбранные) можно рассматривать как компенсацию отброшенного при введении этих процессов члена $f''(x) f'(x)(fx-y)$ [6, стр. 56].

По-видимому, нет вычислительного опыта по применению общей "a-μ- регуляризации", однако, надо предполагать, что сочетание двух регуляризаций придаст процессам ($1a^D$) и ($1a^R$) некоторую дополнительную гибкость и универсальность при реализации на ЭВМ.

Основной задачей теории общей a -μ-регуляризации является нахождение законов поведения операторов и v_n(n =0,I,...) как функций от номера итераций.

§ 2. Сходимость и притяжение R -процессов ньютоновского типа

В этом параграфе результаты относятся одновременно к процессам ($1a^D$) и ($1a^R$). Эти процессы рассматриваются формально как один процесс

$$x_0\ ,\ x_{n+1} = x_n - \bar{a}_n \varphi_n (f x_n - y)\ ,$$

$$x_n \in D \quad (n=0,1,\dots)\ , \qquad (1\bar{a})$$

где $\bar{a}_n = v_n(w_n u_n + z_n)^{-1}$. (Здесь операторы z_n , w_n и φ_n понимаются как операторы z_n^D , w_n^D и φ_n^D , или как операторы z_n^R , w_n^R и φ_n^R).

Пусть в связи с процессом ($1\bar{a}$) введены обозначения

$$\rho_n = \| w_n (f x_n - y)\|\ ,\quad \tau_n = \| w_n u_n\|\ ,$$

$$b_n = \|((w_n u_n + z_n) v_n^{-1})^{-1}\|,\ c_n = \|(w_n u_n + z_n) v_n^{-1}\|\, b_n$$

и пусть введены две вспомогательные функции

$$g_\ell (s,t)\ ;\quad s,t \in [0,\infty)\ ,\quad \ell = 1,2,$$

имеющие свойства:

(g_ℓ) если $s_1 \le s_2$ и $t_1 \le t_2$, то $g_\ell(s_1,t_1) \le g_\ell(s_2,t_2)$;

(g_1) $\| x_{n+1} - x_n\| \le g_1(b_n,\rho_n)\,\rho_n$ (n =0,I, ...)

(g_2) $\rho_{n+1} \le g_2(b_n,\rho_n)\,\rho_n^k$, (n =0,I, ...)

где $k \geqslant 1$.

Функции g_1 и g_2 и число k в дальнейшем называются "функция сходимости", "функция притяжения" и "порядок притяжения", соответственно.

Пусть для решения уравнения (f) применяется процесс ($1\bar{a}$).

В связи с рассмотрением вопроса о сходимости и существовании решения предварительно формулируются следующие гипотезы:

$h_1)$ существуют функции g_1, g_2 и число $K \geq 1$, обладающие свойствами (g_e), (g_1) и (g_2);

$h_2)$ операторы φ_n, u_n и v_n ($n = 0, 1, \dots$) подчиняются условию

$$\| w_{n+1} u_{n+1} v_{n+1}^{-1} - w_n u_n v_n^{-1} \| \leq 2M \| x_{n+1} - x_n \| \qquad (4)$$

($n = 0, 1, \dots$)

(где $M \in [0, \infty)$ – постоянная);

$h_3)$ существует ограниченный оператор $\bar{a}_0 = (w_0 u_0 v_0^{-1} + \tau_0 v_0^{-1})^{-1}$, задано действительное число

$$\gamma > b_0 \qquad (5)$$

и введены в рассмотрение связанные с ним числа

$\beta = g_2(\gamma, \rho_0)\rho_0^{K-1}$, $\sigma = g_1(\gamma, \rho_0)$ и $\varphi_0 = (2M\sigma + N\gamma) b_0 \rho_0$;

$h_4)$ для начальных x_0, τ_0 и v_0 и некоторой постоянной $N \in [0, \infty)$ в силе соотношения

$$\frac{\| \tau_0 \|}{b_0 \rho_0} \leq N , \qquad (6)$$

$$\beta < 1 , \quad \varphi_0 < 1 - \beta , \qquad (7)$$

$$\beta + \frac{\varphi_0}{1 - \varphi_0} \frac{\gamma^2}{(\gamma - b_0) b_0} \leq 1 ; \qquad (8)$$

$h_5)$ операторы регуляризации τ_n и v_n подчиняются ограничениям

$$\| \tau_n v_n^{-1} \| \leq N b_n \rho_n \quad (n = 1, 2, \dots), \qquad (9)$$

$$\| \tau_{n+1} v_{n+1}^{-1} - \tau_n v_n^{-1} \| \leq \| \tau_n v_n^{-1} \| \quad (n = 0, 1, \dots) \qquad (10)$$

$h_6)$ сфера $W = x \in X / \| x_1 - x \| \leq \frac{\sigma \beta \rho_0}{1 - \beta^K}$ лежит в области D.

Имеет место следующее утверждение о сходимости, притяжении и регуляризованности применяемого процесса ($1\bar{a}$).

Теорема I.

Пусть при решении (f) на основе ($1\bar{a}$) в силе предположения $h_1)$ – $h_6)$. Тогда процесс ($1\bar{a}$) является сходящимся R_γ-процессом (с уровнем регуляризованного числа γ из условия h_3)).

Предельная точка $x_\tau = \lim_{n \to \infty} x_n$ лежит в сфере W и удовлетворяет вспомогательному уравнению

$$\varphi(x)(fx - y) = \theta . \qquad (\varphi f)$$

Имеет место оценка

$$\| x_n - x_\tau \| \leq \frac{\sigma \rho_0 B_n}{1 - \beta^{K^n}} \quad (n = 1, 2, \dots) , \qquad (11)$$

где $B_n = \beta^{1 + K + \dots K^{n-1}}$, $B_0 = 1$.

Доказательство. (Проводится по схеме /16, стр.302/ с необходимыми изменениями).

1. По индукции доказывается справедливость утверждений

а) $\rho_n \leq \rho_0 B_n < \rho_0$;

б) оператор $\bar{a}_n$ существует и ограничен;

в) $\varphi_n \le \varphi_0 < 1$, где $\varphi_n = (2M\sigma + N\gamma)\beta_n\rho_n$;

г) $\beta_n = \gamma - (\gamma - \beta_0)B_n \le \gamma$;

д) $x_{n+1} \in D$.

Справедливость а)-д) для $n=0$ следует из $h_1) - h_6)$. Пусть а)-д) верны для некоторого n . Доказывается, что они верны и для $n+1$.

а) Приближение x_{n+1} существует согласно б) и в силу д) принадлежит D . Следовательно, вектор $f x_{n+1} - y \in R$ определен. На основе предположения $h_1)$ и утверждений а) и г) получаются неравенства

$$\rho_{n+1} \le g_2(\beta_n, \rho_n)\rho_n^K \le g_2(\gamma, \rho_0)\rho_n^K \le$$
$$\le g_2(\gamma, \rho_0)(\rho_0 B_n)^K = \beta\rho_0 B_n^K = \rho_0 B_{n+1} . \tag{12}$$

Поскольку $\beta < 1$ (см. (7)), в силе $B_{n+1} < 1$ и,таким образом, $\rho_{n+1} \le \rho_0 B_{n+1} \le \rho_0$.

б) Согласно б), вводится оператор

$$\ell_n = \bar{a}_n((\bar{a}_n)^{-1} + (\bar{a}_{n+1})^{-1}) = E^D - z_n ,$$

где E^D - единичный оператор в D , а $z_n = \bar{a}_n(\bar{a}_{n+1})^{-1}$.

Для нормы $\|\ell_n\|$ с помощью неравенств (4), (9) и (10) получается оценка

$$\|\ell_n\| = \beta_n \|w_n u_n v_n^{-1} - w_{n+1}u_{n+1}v_{n+1}^{-1}\| +$$
$$+\|z_n v_n^{-1} - z_{n+1}v_{n+1}^{-1}\| \le 2M\beta_n\|x_n - x_{n+1}\| + N\beta_n\rho_n . \tag{13}$$

Далее, в силу (g_e) и (g_1) имеет место

$$\|x_n - x_{n+1}\| \le g_1(\beta_n, \rho_n)\rho_n \le g_1(\gamma, \rho_0)\rho_n = \sigma\rho_n .$$

Таким образом

$$\|\ell_n\| \le (2M\sigma + N\gamma)\beta_n\rho_n = \varphi_n < 1 .$$

Следовательно, ряд $\sum_{i=0}^{\infty} \ell_n^i x \ (x \in D)$ сходится, и оператор $z_n = E^D - \ell_n$ имеет ограниченный обратный, для которого

$$\|z_n^{-1}\| \le \frac{1}{1-\|\ell_n\|} \le \frac{1}{1-\varphi_n} .$$

Вместе с операторами $\bar{a}_n$ и z_n^{-1} существует ограниченный оператор $\bar{a}_{n+1}((w_{n+1}u_{n+1} + z_{n+1})v_{n+1}^{-1} = \bar{a}_n z_n)$. Для его нормы справедлива оценка

$$\beta_{n+1} = \|z_n^{-1}(\bar{a}_n)^{-1}\| \le \beta_n \frac{1}{1-\varphi_n} . \tag{14}$$

в) Из соотношения (12) и (14) получается

$$\varphi_{n+1} = (2M\sigma + N\gamma)\beta_{n+1}\rho_{n+1} \le (2M\sigma + N\gamma)\frac{\beta_n}{1-\varphi_n} g_2(\gamma, \rho_0)\rho_n^K =$$
$$= \frac{\varphi_n}{1-\varphi_n} g_2(\gamma, \rho_0)\rho_n^{K-1}$$

откуда с учетом а) и в) и того факта, что $\beta < 1 - \varphi_0$ (см. (7)) получается

$$\varphi_{n+1} \le \frac{\varphi_n}{1-\varphi_0}\beta \le \varphi_n .$$

г) Согласно в) имеют место

$$1 - \varphi_n > 0 , \quad 1 - \varphi_0 > 0, \quad \varphi_n(\varphi_0 - \varphi_n) \ge 0$$

и отсюда

$$\frac{1}{1-\varphi_n} = \frac{1-\varphi_0}{(1-\varphi_n)(1-\varphi_0)} < \frac{1-\varphi_0 + \varphi_n(\varphi_0 - \varphi_n)}{(1-\varphi_n)(1-\varphi_0)} =$$
$$= 1 + \frac{\varphi_n}{1-\varphi_0} = 1 + \frac{(2M\sigma + N\gamma)\beta_n\rho_n}{1-\varphi_0} .$$

Таким образом, из (14) следует, что

$$\beta_{n+1} \le \beta_n\left(1 + \frac{(2M\sigma + N\gamma)\beta_n\rho_n}{1-\varphi_0}\right) .$$

При помощи г), а) и при обозначении

$$q = \frac{(2M\sigma + N\gamma)\rho_0}{1-\varphi_0} = \frac{\varphi_0}{1-\varphi_0}\,\frac{1}{b_0}$$

получается $b_{n+1} \le b_n(1+\gamma q B_n)$ и в силу г) находится окончательно

$$b_{n+1} \le (\gamma - (\gamma - b_0)B_n)(1+\gamma q B_n) = \gamma - (\gamma - b_0)B_n P, \qquad (I5)$$

где

$$P = 1 + \gamma q B_n - \frac{\gamma^2 q}{\gamma - b_0} \geqslant 1 - \frac{\gamma q}{\gamma - b_0}.$$

Знаменатель $\gamma - b_0 > 0$ в силу (5). Если подставить q, то с учетом $\varphi_0 < 1-\beta$ из (7) получается

$$P \geqslant 1 - \frac{\gamma^2\varphi_0}{(\gamma-b_0)(1-\varphi_0)b_0} \geqslant \beta \geqslant \beta^{K^n}.$$

Отсюда, в силу $B_n P \geqslant B_{n+1}$ неравенство (I5) принимает вид

$$b_{n+1} \le \gamma - (\gamma - b_0)B_{n+1}$$

Из $(\gamma - b_0)B_{n+1} > 0$ следует $b_{n+1} \le \gamma$.

д) В неравенстве $\|x_{n+2}-x_1\| \le \sum_{i=1}^{n+1}\|x_i - x_{i+1}\|$ каждый член суммы оценивается в отдельности с помощью (g_1):

$$\|x_{i+1}-x_i\| \le g_1(b_i,\rho_i)\rho_i \le g_1(\gamma,\rho_0)\rho_0 B_i =$$
$$= \sigma\rho_0 B_i \qquad (i=1,2,\dots,n+1). \qquad (I6)$$

Следовательно,

$$\|x_{n+2}-x_1\| \le \sigma\rho_0\sum_{i=1}^{n+1} B_i =$$
$$= \sigma\rho_0(\beta+\beta\beta^K+\beta\beta^K\beta^{K^2}+\dots+\beta\beta^K\dots\beta^{K^n}).$$

Поскольку $\beta<1$ и $K\geqslant 1$ получается

$\beta^{K^i}\le\beta^K\le\beta<1$ для всех $i=1,2,\dots,n+1$.

Следовательно,

$$\|x_{n+2}-x_1\| \le \sigma\rho_0(\beta+\beta(\beta^K)+\beta(\beta^K)^2+\dots+\beta(\beta^K)^n) =$$
$$= \sigma\rho_0\beta\,\frac{1-(\beta^K)^{n+1}}{1-\beta^K} < \sigma\rho_0\frac{\beta}{1-\beta^K}.$$

Отсюда получается, что $x_{n+2}\in W$ и $x_{n+2}\in D$.

Этим доказано, что утверждения а)-д) выполнены для всех n=I,2, ...

2. На основе (I6) получаются неравенства

$$\|x_{n+m+1}-x_n\| \le \sum_{i=n}^{n+m}\|x_i-x_{i+1}\| \le$$
$$\le \sum_{i=n}^{n+m}\sigma\rho_0 B_i = \sigma\rho_0 B_n\frac{1-(\beta^{K^n})^{m+1}}{1-\beta^{K^n}} < \frac{\sigma\rho_0 B_n}{1-\beta^{K^n}}. \qquad (I7)$$

При $n\to\infty$ $1-\beta^{K^n}$ остается ограниченным, а $\sigma\rho_0 B_n$ вместе с B_n стремится к 0. Все x_n лежат в W. Для всех m=I,2, ... $\|x_{n+m+1}-x_n\|\to 0$ при $n\to\infty$.

Таким образом, существует предел $x_z = \lim_{n\to\infty} x_n$ и $x_z\in W$.

Далее, при $n\to\infty$ вместе с B_n стремится к нулю и $\rho_n\le\rho_0 B_n$, а последовательность $\{\varphi_n(fx_n-y)\}_{n=0}^{\infty}$ в силу непрерывной зависимости функций $\varphi(x)$ и $fx-y$ от x стремится к $\varphi(x_z)(fx_z-y)$. Таким образом, имеет место соотношение $\|\varphi(x_z)(fx_z-y)\|=0$, т.е. x_z является решением вспомогательного уравнения (φf).

Приближение x_{n+m+1} при $m\to\infty$ стремится к x_z и (I7) при фиксированном n дает оценку (II). Утверждение, что ($1\bar{a}$) является R_γ-процессом, следует из справедливости неравенств г) (для всех n). Этим теорема I доказана.

Гипотезы $h_3)$, $h_4)$ и $h_5)$ обеспечивают априори свойство процесса ($1\bar{a}$) быть R_γ-процессом. С другой стороны, ЭВМ и программирование предоставляют большие возможности для реализации итерационных процессов типа ($1\bar{a}$), которые приобретают свойства регуляризованности во время выполнения на ЭВМ. Такие процессы, как правило, являются более сложными при реализации на ЭВМ, но зато обладают более простой теорией сходимости и притяжения. В § 5 настоящей работы приводятся некоторые утверждения о сходимости и притяжении для процессов, регуляризованных во время выполнения .

§ 3. Явный вид условия регуляризации

Определение операторов регуляризации τ_n и v_n прямо из неравенств (9) может оказаться затруднительным по причине зависимости b_n от самих операторов τ_n и v_n. По этой причине ставится задача о нахождении эквивалентных условий (9), связывающих операторы τ_n и v_n проще, чем неравенства (9). В связи с этим имеет место утверждение.

<u>Теорема 2</u>.

Выполнение неравенств

$$\|\tau_n v_n^{-1}\| \le \frac{1}{2}\left(\sqrt{\|w_n u_n v_n^{-1}\|^2 + 4N\rho_n} - \|w_n u_n v_n^{-1}\|\right) \qquad \text{(I8)}$$

влечет за собой выполнение неравенств (9).

<u>Доказательство</u>.

Из (I8) следуют

$$\left(\|\tau_n v_n^{-1}\| - \frac{1}{2}\sqrt{\|w_n u_n v_n^{-1}\|^2 + 4N\rho_n} - \frac{1}{2}\|w_n u_n v_n^{-1}\|\right) *$$

$$\left(\|\tau_n v_n^{-1}\| + \frac{1}{2}\sqrt{\|w_n u_n v_n^{-1}\|^2 + 4N\rho_n} + \frac{1}{2}\|w_n u_n v_n^{-1}\|\right) \le 0$$

(n =0,I, ...).

Отсюда получаются неравенства

$$\|\tau_n v_n^{-1}\| \left(\|\tau_n v_n^{-1}\| + \|w_n u_n v_n^{-1}\|^2\right) \le N\rho_n \qquad (n=0,1,\dots)$$

и неравенства

$$\|\tau_n v_n^{-1}\| \, \|(w_n u_n + \tau_n) v_n^{-1}\| \le N b_n \rho_n \ . \qquad (n=0,1,\dots)$$

Из свойства числа обусловленности $C_n \geqslant 1$ /I6/ можно заключить, что окончательно в силе

$$\|\tau_n v_n^{-1}\| \le N b_n \rho_n \qquad (n=0,\mathrm{I},\ \dots).$$

Для важного частного случая α-регуляризации ($v_n = E^D$, $\tau_n \ne \theta$) неравенства (I8) являются условиями для операторов регуляризации τ_n (n =0,I, ...) уже в явном виде.

Аналогично приведенной теореме 2 можно доказать утверждение, что выполнение неравенств

$$\|\tau_n\| \ge \frac{1}{2}\left(\sqrt{\tau_n^2 + \frac{4N}{\|v_n^{-1}\|^2}\rho_n} - \tau_n\right) \qquad (n=0,\mathrm{I},2,\ \dots). \qquad \text{(I9)}$$

влечет за собой выполнение (9).

Теперь очевидно следующее следствие теоремы I.

<u>Теорема 3</u>.

Пусть выполнены условия $h_1)$- $h_6)$ теоремы I, где вместо неравенства (6) в силе

$$\|\tau_0\| \left(\|\tau_0\| + \tau_0^2\right) \le N\rho_0 \ , \qquad \text{(20)}$$

а вместо неравенств (9) в силе (I8) (или (I9)). Тогда имеют место утверждения теоремы I.

При построении R -процессов с операторами регуляризации, удовлетворяющими законам (I9) или (20), существуют две возможности определения операторов $\mathcal{V}_n$ и τ_n. Первая возможность - когда сначала определяются (на каждый шаг) операторы $\mathcal{V}_n$, а потом (на основе (I9) или (20)) определяются операторы τ_n. Вторая возможность - когда сначала находятся τ_n(с $\|\tau_n\| \neq 0$ и не используя для этой цели (I9) или (20)), а $\mathcal{V}_n$ определяются из условий

$$\|\mathcal{V}_n^{-1}\| \leq 2\sqrt{\frac{N\rho_n}{(2\|\tau_n\|+\tau_n)^2-\tau_n^2}}$$

$$(n=0,I,\ \dots). \tag{2I}$$

При применении неравенств (2I) для построения R -процессов типа (I0) можно доказать утверждения, аналогичные приведенным теоремам 2 и 3.

Замечание 3.

Теоремы 2 и 3 дают возможность ответить положительно на вопрос о существовании операторов $\tau_n \neq 0$, $\tau_n \neq E^D$ и числа $N \neq 0$ таких, для которых соотношения (6), (7), и (8) выполняются (аналогично теореме 5 из/2/).

§ 4. Конкретные R -процессы ньютоновского типа

Без указания конкретных функций сходимости и притяжения общие теоремы о сходимости I и 3 имеют формальный характер.

В настоящем параграфе рассматривается построение функций g_1 и g_2 для некоторых имеющих значение в практике на ЭВМ R -процессов ньютоновского типа и приводятся соответствующие следствия основных теорем о сходимости I и 3.

R -процессы Шредера.

Пусть в D существует вторая произвольная Фреше $f''(x)$ (которая как билинейный оператор действует над парами (ρ,ρ) $(\rho\in D)$.

Пусть задан следующий процесс типа

$$R^S:\ x_0,\ x_{n+1}=x_n-\mathcal{V}_n^S(f'(x_n)+\tau_n)^{-1}(fx_n-y)$$

$$x_n\in D \qquad (n=0,I,\ \dots), \tag{22}$$

где операторы $\mathcal{V}_n^S$ задаются формулой

$$\mathcal{V}_n^S\rho=\rho+\tfrac{1}{2}(f'(x_n)+\tau_n)^{-1}f''(x_n)(\rho,\rho)$$

и являются обратимыми для любого конечного n. Процесс R^S можно рассматривать как обобщение процесса Шредера/I7, стр.270/ (при $K=3$).

Имеет место следующее следствие теоремы I.

Теорема 4.

Пусть уравнение (f) решается на основе процесса R^S и пусть выполнено условие:

h_1') В D существуют ограниченные производные Фреше $f'(x)$, $f''(x)$ и $f'''(x)$

$$(\|f'(x)\|\leq T<\infty,\ \|f''(x)\|\leq 2M<\infty,\ \|f'''(x)\|\leq U<\infty),$$

а также выполнены условия $h_2)$ - $h_5)$ теоремы I. Тогда, относительно R^S и уравнения (f) в силе утверждения теоремы I (при порядке притяжения $K=3$).

Доказательство.

Из (22) следует неравенство

$$\| x_{n+1} - x_n \| \leq g_1^S \rho_n \quad , \tag{23}$$

где $\quad g_1^S = e_n + M e_n^3 \rho_n \, , \quad e_n = \| (f'(x_n) + \tau_n)^{-1} \|$.

На основе тождества

$$f x_n - y + \frac{1}{2} f''(x_n)(\bar{\rho}_n, \bar{\rho}_n) + (f'(x_n) + \tau_n)\delta_n = 0 \; , \tag{24}$$

где $\quad \delta_n = x_{n+1} - x_n = -v_n^S \bar{\rho}_n \quad$ и

$\bar{\rho}_n = (f(x_n) + \tau_n)^{-1} (f x_n - y)$

и формулы Тейлора получается неравенство

$$\rho_{n+1} \leq M^2 e_n^4 (2 + M e_n^2 \rho_n)\rho_n^3 + U \| x_{n+1} - x_n \|^2 + \\ + \| \tau_n \| \| x_{n+1} - x_n \| \; . \tag{25}$$

Из неравенств (9), (23)-(25) следует оценка

$$\rho_{n+1} \leq g_2^S \rho_n \qquad (n = 0, 1, \ldots) \quad , \tag{26}$$

где

$$g_2^S = M^2 e_n^4 (2 + M e_n^2 \rho_n) + U(e_n + M e_n^3 \rho_n)^3 + N e_n (e_n + M e_n^3 \rho_n) \; .$$

Неравенства (23) и (26), вместе с (g_e), (g_1) и (g_2) дают основание заключить, что найденные функции g_1^S и g_2^S являются соответственно функциями сходимости и притяжения и, что порядок притяжения процесса R^S K =3. Теперь, применяя теорему I (для функции g_1^S и g_2^S) доказывается справедливость теоремы 4.

Очевидно, итерационная схема (22) содержит в себе рассматриваемые ранее R -процессы Ньютона-Канторовича[2] (теорема 4 обобщает теорему 3 из[2]).

R -процессы Гаусса Ньютона

Пусть X , Y - гильбертовы пространства. Пусть в D существует первая производная Фреше $f'(x)$, которая как оператор действует на всем X . Пусть $f'^*(x_n)$ обозначает оператор, сопряженный оператору $f'(x_n)$, действующий из пространства Y в X . Рассматривается следующий R -процесс типа $(1a^D)$ (процесс Гаусса-Ньютона)

$$R^{GN}: x_0 \, , \; x_{n+1} = x_n - v_n (f'^*(x_n) f'(x_n) + \tau_n)^{-1} f'^*(x_n)(f x_n - y)$$

$$x_n \in D \, , \; (\tau_n x, x) > 0 \text{ в } D , \quad (n = 0, 1, \ldots) . \tag{27}$$

Относительно R^{GN} имеет место следствие теоремы 3.

Теорема 5.

Пусть уравнение

$$f'^*(x)(f x - y) \qquad\qquad (f'^* f)$$

решается на основе процесса R^{GN} и пусть выполнено условие:

$h_1'')$ существуют постоянные L и $\bar{M} \in [0, \infty)$, такие, что для любого n =0,I,... в силе неравенства

$$\| f'^*(x_{n+1}) - f'^*(x_n) \| \leq L \| x_{n+1} - x_n \| \, , \tag{28}$$

$$\| f x_{n+1} - f x_n - f'(x_n) v_n^{-1} (x_{n+1} - x_n) \| \leq \bar{M} \| x_{n+1} - x_n \|^2 \tag{29}$$

$$(n = 0, I, \ldots),$$

а также условия $h_2) - h_5)$ теоремы I, в которых вместо соотношений (6) и (9) в силе соотношения (20) и (I9).

Тогда относительно процесса R^{GN} и уравнения $(f'^* f)$ имеют место утверждения теоремы I (при порядке притяжения K =I).

Доказательство.

Из (27), (g_ℓ) и (g_1) следует, что $g_1^{GN} = b_n$ является функцией сходимости процесса R_γ^{GN}.

На основе тождества

$$f'^*(x_n)(fx_n - y) + (f'^*(x_n)f'(x_n) + \tau_n)V_n^{-1}\delta_n = 0$$

(где $\delta_n = x_{n+1} - x_n$) получаются соотношения

$$\rho_{n+1} = \| f'^*(x_{n+1})(fx_{n+1} - y) - f'^*(x_n)(fx_n - y) - (f'^*(x_n)f'(x_n) + \tau_n)V_n^{-1}\delta_n \| \le$$

$$\le \| (f'^*(x_{n+1}) - f'^*(x_n))(fx_{n+1} - y)\| + \| f'^*(x_n)\| \| fx_{n+1} - fx_n - f'(x_n)V_n^{-1}\delta_n \| + \| \tau_n V_n^{-1}\delta_n \| .$$

Отсюда после применения неравенств (28)-(29), а также ограничений (I9) и (20), получается оценка

$$\rho_{n+1} \le (LR + (\bar{M}T + N)b_n\rho_n)b_n\rho_n . \tag{29}$$

Здесь R – постоянная со свойством $\|fx - y\| \le R < \infty$ для всех $x \in D$.

Из соотношений (g_ℓ), (g_2) и (29) следует, что

$$g_2^{GN} = (LR + (\bar{M}T + N)b_n\rho_n)b_n$$

является для процесса R^{GN} функцией притяжения и что порядок притяжения $k = I$.

Далее, теорема 5 следует непосредственно из теоремы 3.

Замечание 4.

Гипотеза h_1'') является достаточно естественной, так как для частного случая $V_n = E^D$ неравенство (29) следует из условия Липшица.

О процессе R^{GN}, при $V_n = E^D$, можно доказать утверждение, аналогичное теореме 5, в котором порядок притяжения $k = 2$. Для этой цели надо предполагать вместо выполнения (28) выполнение соотношения/2,5/

$$\| (f'^*(x) - f'^*(\bar{x}))(fx - y)\| \le Q \| f'^*(x)(fx - y)\| ,$$

где $Q \in [0,1)$, для всех x и $\bar{x}$ из W.

Однако это соотношение уменьшает сферу сходимости рассматриваемого процесса R_γ^{GN} в сравнении со сферой сходимости, обоснованной теоремой 5.

Сопоставляя теорему 5 с соответствующими утверждениями из/2,5/, можно заключить следующее.

Линейный порядок притяжения ($k = I$) является типичным для большей части итерационного процесса R_γ^{GN} ($V_n = E^D$), но для достаточно большого номера итерации n квадратический закон притяжения ($k = 2$) тоже возможен.

R-процессы Гаусса-Ньютона с компенсированной регулиризацией.

При условиях, в которых рассматривается процесс R^{GN}, можно рассматривать также итерационный процесс

$$R_C^{GN}: x_0, x_{n+1} = x_n - (f'^*(x_n)f'(x_n) + \tau_n)^{-1}\varphi_n^C(fx_n - y)$$

$$x_n \in D \qquad (n = 0, I, \ldots),$$

где линейные операторы φ_n^C имеют вид

$$\varphi_n^C = (E^D - \tau_n(f'^*(x)f'(x_n) + \tau_n)^{-1})f'^*(x_n).$$

Процесс R_C^{GN} получается аналогично процессу R_γ^{GN}, с той только разницей, что при решении линеаризованного уравнения прибавление регуляризующего вектора $\tau_n(x_{n+1} - x_n)$ компенсируется добавлением в правой части/I8/ нового вектора

$$z_n \, (f'^*(x_n) f'(x_n) + z_n)^{-1} f'^*(x_n)(f x_n - y) \; .$$

Для получения таким образом процесса R_C^{GN} в силе утверждение о сходимости и существовании решения, аналогичное теореме 5.

Процесс R_C^{GN} сходится к решению уравнения

$$\varphi^C(x)(fx-y) = \theta \; .$$

§ 5. О R-процессах ньютоновского типа, регуляризованных во время выполнения

Пусть для процесса ($1\bar{a}$) задан некоторый уровень регуляризованности $\bar{\gamma} < \infty$ и пусть операторы v_n и z_n подбираются так, что ограничения

$$\| v_n \, (w_n u_n + z_n)^{-1} \| \le \bar{\gamma} \tag{30}$$

выполнены для всех $n = 0, 1, \ldots, n^*$, где n^* – достаточно большое число. Тогда процесс ($1\bar{a}$) понимается как регуляризованный во время выполнения. R-процессы Гаусса-Ньютона существенно отличаются от других процессов ньютоновского типа по возможности алгоритмически простым обеспечением неравенств (30) – на основе только a-регуляризации. В этом случае используется, что операторы $f'^*(x_n) f'(x_n)$ являются неотрицательно определенными в X. О сходимости R-процессов, регуляризованных во время выполнения, имеет место следующее общее утверждение.

Теорема 6.

Пусть (f) решается на основе регуляризованного во время

выполнения процесса $R_{\bar{\gamma}}$ типа ($1\bar{a}$) и пусть выполнены условия $h_1)$ и $h_6)$ теоремы 1, а также выполнено условие:

$h_3')$ для функции притяжения g_2 и заданного уровня регуляризованности $\bar{\gamma}$ в силе неравенство

$$\bar{\beta} = g_2(\bar{\gamma}, \rho_0)\rho_0^{k-1} < 1 \; . \tag{31}$$

Тогда относительно процесса $R_{\bar{\gamma}}$ имеют место утверждения теоремы 1.

При доказательстве этой теоремы сначала устанавливается справедливость утверждений а) и д) (см. доказательство теоремы 1). При этом имеется разница только в установлении а). Для процесса $R_{\bar{\gamma}}$ приближение x_{n+1} всегда существует и согласно д) принадлежит D. Из $h_1)$ и утверждения а) следуют неравенства

$$\rho_{n+1} \le g_2(b_n, \rho_n)\rho_n^k \le g_2(\bar{\gamma}, \rho_0)\rho_n^k \le$$

$$\le g_2(\bar{\gamma}, \rho_0)(\rho_0 B_n)^k = \beta \rho_0 B_n^k = \rho_0 B_{n+1} \; . \tag{32}$$

Условие $h_3')$ дает, что $B_{n+1} < 1$.

На основе (32) заключается, что в силе

$$\rho_{n+1} \le \rho_0 B_{n+1} \le \rho_0 \; .$$

Окончание доказательства теоремы 6 совпадает с частью 2 доказательства теоремы 1.

В сравнении с теоремой 1, теорема 6 имеет меньше условий (отсутствуют аналоги условий $h_2)$ и $h_4)$), а условие $h_3')$ проще, чем условие $h_3)$.

В частном случае R-процессов Ньютона-Канторовича теорема 6 обобщает известную теорему /19/.

Для R-процессов Гаусса-Ньютона, регуляризованных во времени выполнения, здесь приводится только утверждение в случае α-регуляризации ($\mathcal{V}_n = E^{D}$).

Предварителтно формулируются следующие условия:

h_2^*) в D существует ограниченная вторая производная Фреше $f''(x)\,(\frac{1}{2}\|f''(x)\| \le M < \infty)$;

h_5^*) операторы регуляризации подчинены ограничениям

$$\|r_n\| \le \frac{1}{2}\left(\sqrt{\|f'^*(x_n)f'(x_n)\|^2 + 4\bar{N}\|f'^*(x_n)(fx_n - y)\|} - \|f'^*(x_n)f'(x_n)\|\right)$$

$$(n = 0, 1 \ldots), \qquad (33)$$

где $\bar{N}$ – постоянная со свойством

$$0 \le \bar{N} < \frac{1 - MT\rho_0\gamma^2 - 2MR\bar{\gamma}}{\bar{\gamma}^2\rho_0}; \qquad (34)$$

h_6^*) сфера

$$W = \left\{ x \in X \;\middle|\; \|x_1 - x\| \le \frac{\bar{\gamma}\bar{\beta}\rho_0}{1 - \bar{\beta}} \right\},$$

где $\bar{\beta} = (2M\bar{\gamma} + (MT + \bar{N})\bar{\gamma}\rho_0)\bar{\gamma}$,

принадлежит области D.

Теорема 7.

Пусть в силе условие h_2^*) и пусть уравнение (f'^*f) решается на основе α-регуляризованного (во время выполнения) процесса $R_{\bar{\gamma}}^{GN}$, уровень регуляризации которого (в случае $T \ne 0$ и $M \ne 0$) подчинен условию

$$\bar{\gamma} \le \sqrt{\left(\frac{R}{T\rho_0}\right)^2 + \frac{1}{MT\rho_0}} - \frac{R}{T\rho_0}. \qquad (35)$$

Пусть также выполнены условия h_5^*) и h_6^*). Тогда процесс $R_{\bar{\gamma}}^{GN}$ сходится. Предельная точка $x_r = \lim_{n\to\infty} x_n$ лежит в сфере W и является решением уравнения (f'^*f). Имеет место оценка

$$\|x_n - x_r\| \le \frac{\bar{\gamma}\bar{\beta}^n}{1 - \bar{\beta}} \qquad (n = 0, 1, \ldots)$$

__Доказательство.__

Сначала отметим, что неравенство (34) хорошо определяет постоянную $\bar{N}$ (при $T \ne 0$ и $M \ne 0$), так как из (36) следует неравенство

$$\left(\bar{\gamma} - \sqrt{\left(\frac{R}{T\rho_0}\right)^2 + \frac{1}{MT\rho_0}} + \frac{R}{T\rho_0}\right)\left(\bar{\gamma} + \sqrt{\left(\frac{R}{T\rho_0}\right)^2 + \frac{1}{MT\rho_0}} + \frac{R}{T\rho_0}\right) < 0$$

и неравенство

$$MT\rho_0\bar{\gamma}^2 + 2MR\bar{\gamma} < 1. \qquad (36)$$

Из условия h_2^*) следует, что для всех x, $\bar{x} \in D$ в силе соотношения

$$\|f'^*(x) - f'^*(\bar{x})\| \le 2M\|x - \bar{x}\|, \qquad (37)$$

$$\|fx - f\bar{x} - f'(\bar{x})(x - \bar{x})\| \le M\|x - \bar{x}\|. \qquad (38)$$

Неравенства (33) вместе с теоремой 2 обеспечивают выполнение ограничений

$$\|r_n\| \le \bar{N} b_n \rho_n \qquad (n = 0, 1, \ldots) \qquad (39)$$

(Здесь $b_n = \|(f'^*(x_n)f'(x_n) + r_n)^{-1}\|$ и $\rho_n = \|f'^*(x_n)(fx_n - y)\|$).

Из неравенств (37)-(39) и свойств (g_l), (g_1), (g_2) следует, что функции

$$g_1^{GN} = b_n,$$
$$g_2^{GN} = (2MR + (MT + \bar{N})b_n\rho_n)b_n$$

являются функциями сходимости и притяжения для процесса $R_{\bar{\gamma}}^{GN}$ (порядок сходимости $k = 1$).

Определение $\bar{\beta} = g_2^{GN}(\bar{\gamma}, \rho_0)$ и второе неравенство (34) дают неравенство $\bar{\beta} < 1$. Для окончания доказательства применяется теорема 6.

ЛИТЕРАТУРА

I. Л.Александров. Сообщение ОИЯИ, Р5-5I37, Дубна, I970.

2. Л.Александров. ЖВМ и МФ, т.II, № I, I97I, стр. 36-43.

3. Л.Александров. Сообщение ОИЯИ, Р5-52I5, Дубна, I970.

4. В.Е.Шаманский. Укр. матем. журнал. т. 22, №4, I970.

5. Л.Александров. Сообщение ОИЯИ, Р5-55I5, Дубна, I970.

6. В.Н.Кублановская. Зап. науч. сем. мат. и-та В.А.Стеклова, АН СССР, т. 23, Ленинград, I97I.

7. А.Фиакко, Г.Мак-Кормик. Нелинейное программирование. М., I972.

8. Y. Bard. SiAM J. Numer. Anal., v. 7, No. 1, 1970, 157-186.

9. Ч.Стоянов, Л.Александров, В.Гаджоков. J. of Radio . Chem., v. 10, 1972, pp 75-81.

I0. Л.Александров, В.Гаджоков. J. of Radio. Chem., v. 9, 1971, pp 279-292.

II. Л.Александров, Д.Караджов, И.Н.Михайлов, Г.Ходжаев. Изв. АН СССР, т. 36, М. (I972).

I2. Л.Александров. Препринт ОИЯИ, 5-682I, Дубна, I972.

I3. М.К.Гавурин. Изв.ВУЗ, сер. математ., I958, №5, стр. I063-I066.

I4. K.Levenberg. Quart. Appl. Math. 2 (1944), pp 164-168.

I5. D.W.Marquardt. SiAM J. Appl. Math. v. 11, No.2, 1963, 431-441.

I6. С.Н.Соколов, И.Н.Силин. Препринт ОИЯИ, Д-8I0, Дубна, I96I.

I7. Л.Коллатц. Функциональный анализ и вычислительная математика, М., I969.

I8. В.А.Морозов. Вычислительные методы и программирование. (Сборник) МГУ, I967, стр. 63-95.

I9. И.П.Мысовских. Т. матем. ин-та АН СССР, т. 29, I949.

Рукопись поступила в издательский отдел
26 июня I973 года.