СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
Дубна

Р5-5790

Л. Александров

О РЕШЕНИИ НЕКОРРЕКТНО ПОСТАВЛЕННЫХ ЗАДАЧ ПОСТРОЕНИЕМ РЕГУЛЯРИЗОВАННЫХ И АВТОРЕГУЛЯРИЗОВАННЫХ ИТЕРАЦИОННЫХ ПРОЦЕССОВ

ЛАБОРАТОРИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ
И АВТОМАТИЗАЦИИ

1971

Р5-5790

Л. Александров

О РЕШЕНИИ НЕКОРРЕКТНО ПОСТАВЛЕННЫХ ЗАДАЧ ПОСТРОЕНИЕМ РЕГУЛЯРИЗОВАННЫХ И АВТОРЕГУЛЯРИЗОВАННЫХ ИТЕРАЦИОННЫХ ПРОЦЕССОВ

На основании метода регуляризации А.Н. Тихонова[1-3] для решения нелинейных операторных уравнений в[4-6] предлагаются "регуляризованные" и "авторегуляризованные" итерационные процессы типа Ньютона-Канторовича.

В настоящей работе рассматривается вопрос, являются ли регуляризованные итерационные процессы из[4-6] регуляризирующими алгоритмами.

Приводится следствие о сходимости процесса $R^{\ell}_{\epsilon_n}$ в линейном случае[5,6]. Даются численные примеры решения линейных, плохо обусловленных систем уравнений авторегуляризованным итерационным процессом $R^{\ell}_{\epsilon_n}$.

1. Пусть X - подмножество банахова пространства E и оператор f преобразует $x \in X$ в $y \in X$.

Пусть уравнение

$$fx - y = 0 \tag{1}$$

решается итерационным процессом

$$\mathcal{P}:\ x_0, x_{n+1} = x_n + R(x_n)(fx_n - y)\ (n = 0, 1, 2, \ldots,), \tag{2}$$

где оператор - функция $R(x)(x \in X)$ - как оператор действует из X в X.

Везде в этой работе для уравнения (1) будем предполагать, что если существует решение, то оно единственно.

Докажем предварительно следующее вспомогательное утверждение.

Лемма

Пусть в X существуют ограниченные первые производные Фреше операторов R и f . Пусть последовательность $\mathcal{P}$ сходится как для точного $\bar{y} \in X$, так и для любого приближения $\tilde{y} \in X$ ($\|\bar{y} - \tilde{y}\| \leq \delta$, $\delta > 0$) соответственно к $\bar{x}^*$, $\tilde{x}^* \in X$. Тогда нахождение предела $\tilde{x}^*$ является регуляризирующим алгоритмом.

Доказательство

Первое требование в определении Тихоновым регуляризирующего алгоритма/1/, очевидно, выполнено. Второе требование выполнено, если любому числу $\epsilon > 0$ можно сопоставить число $\delta(\epsilon) > 0$, такое, что если выполнено неравенство $\|\bar{y} - \tilde{y}\| \leq \delta$, то выполняется неравенство $\|\bar{x}^* - \tilde{x}^*\| \leq \epsilon$.

В нашем случае из сходимости последовательности $\mathcal{P}$ для $\bar{y}$ и $\tilde{y}$ следует существование убывающих при $n \to \infty$ функций $\bar{\eta}_n$ и $\tilde{\eta}_n$, таких, что

$$\|\bar{x}_n - \bar{x}^*\| \leq \bar{\eta}_n \,,$$

$$\|\tilde{x}_n - \tilde{x}^*\| \leq \tilde{\eta}_n \,, \ (n = 0,1,2,\dots) \,. \tag{3}$$

Используя тождество

$$(\bar{x}^* - \tilde{x}^*) + (\tilde{x}^* - \tilde{x}_{n+1}) - (\bar{x}^* - \bar{x}_{n+1}) = \bar{x}_{n+1} - \tilde{x}_{n+1}$$

и оценки (3), находим неравенство

$$\|\bar{x}^* - \tilde{x}^*\| \leq \bar{\eta}_{n+1} + \tilde{\eta}_{n+1} + \|\bar{x}_{n+1} - \tilde{x}_{n+1}\| \,. \tag{4}$$

Найдем оценку для нормы $\|\bar{x}_{n+1} - \tilde{x}_{n+1}\|$. Из (2) следует:

$$\bar{x}_{n+1} - \tilde{x}_{n+1} = (\bar{x}_n - \tilde{x}_n) + (\mathfrak{R}(\bar{x}_n) f \bar{x}_n - \mathfrak{R}(\tilde{x}_n) f \tilde{x}_n) + (\mathfrak{R}(\bar{x}_n)\bar{y} - \mathfrak{R}(\tilde{x}_n)\tilde{y}) \,. \tag{5}$$

По теореме о среднем значении имеем оценку

$$\| \mathcal{R}(\bar{x})\, f\bar{x} - \mathcal{R}(\tilde{x})\, f\tilde{x} \| \le PQ \| \bar{x} - \tilde{x} \| . \tag{6}$$

(Здесь используются обозначения: $\| \mathcal{R}'(x) \| \le P$ и $\| f'(x) \| \le Q$ для всех $x \in X$).

По формуле Тейлора имеем представление

$$\mathcal{R}(\tilde{x}) = \mathcal{R}(\bar{x}) + \sigma(\bar{x}, \tilde{x}) \| \bar{x} - \tilde{x} \| , \tag{7}$$

где $\lim\limits_{\tilde{x} \to \bar{x}} \sigma(\bar{x}, \tilde{x}) = 0$.

На основе (5), (6) и (7) запишем неравенства

$$\| \bar{x}_{n+1} - \tilde{x}_{n+1} \| \le \| \bar{x}_n - \tilde{x}_n \| + PQ \| \bar{x}_n - \tilde{x}_n \| +$$

$$+ \| \sigma \| \, \| \tilde{y} \| \, \| \bar{x}_n - \tilde{x}_n \| + \| \mathcal{R}(\bar{x}_n)(\bar{y} - \tilde{y}) \| \le$$

$$\le (1 + PQ + \| \sigma \| \, \| \tilde{y} \|) \| \bar{x}_n - \tilde{x}_n \| + \mathcal{R} \| \bar{y} - \tilde{y} \| .$$

(Здесь – $\| \mathcal{R}(x) \| \le \mathcal{R} < \infty$ для всех $x \in X$).

Введем числа S и δ_0 , такие, что: $1 + PQ + \| \sigma \| \, \| \tilde{y} \| = S$ и $\| \bar{y} - \tilde{y} \| \le \delta_0$. Из последних неравенств имеем

$$\| \bar{x}_{n+1} - \tilde{x}_{n+1} \| \le S \| \bar{x}_n - \tilde{x}_n \| + R \delta_0 \quad (n = 0, 1, 2, \dots) .$$

Из этого рекуррентного соотношения находим оценку

$$\| \bar{x}_{n+1} - \tilde{x}_{n+1} \| \le R \frac{S^{n+1} - 1}{S - 1} \delta_0 \equiv t_n \delta_0 . \tag{8}$$

Подставляя (8) в (4), получаем окончательную оценку

$$\| \bar{x}^* - \tilde{x}^* \| \le \bar{\eta}_n + \tilde{\eta}_n + t_n \delta_0 . \tag{9}$$

Для произвольного фиксированного $\epsilon > 0$ всегда существует достаточно большой номер n^* и достаточно маленькое число $\delta_0^* > 0$, такие, что имеет место неравенство

$$\epsilon \ge \bar{\eta}_{n^*} + \tilde{\eta}_{n^*} + t_{n^*} \delta_0^* . \tag{10}$$

На основе (9) и (10) заключаем, что для любого положительного числа $\delta \le \delta_0^* \le \frac{1}{t} (\epsilon - \bar{\eta}_{n^*} - \tilde{\eta}_{n^*})$ в силе импликация $\| \bar{y} - \tilde{y} \| \le \delta \longrightarrow$ $\longrightarrow \| \bar{x}^* - \tilde{x}^* \| \le \epsilon$, что и требовалось доказать.

2. Пусть E–гильбертово пространство и пусть уравнение (1) решается авторегуляризованным итерационным процессом типа Ньютона-Канторовича $\mathfrak{R}^{\ell}_{\epsilon_n}$ /6/. В этом случае для оператора $\mathfrak{R}$ имеем

$$\mathfrak{R}(x_n) = -(f'^*(x_n)f'(x_n) + \epsilon_n E)^{-1} f'^*(x_n),$$

где числа ϵ_n $(n = 0,1,2,\dots)$ задаются выражением

$$\epsilon_n = a_2 \left\{ \sqrt{\frac{\|f'(x_n)\|^4}{4} + \frac{a_1}{\|f'^*(x_0)(fx_0 - y)\|}(\|f'(x_0)\|^2\epsilon_0 + \epsilon_0^2)\|f'^*(x_n)(fx_n - y)\|} - \frac{1}{2}\|f'(x_n)\|^2 \right\},$$

$a_1 \in (0,1]$, $a_2 \in [0,1]$ – произвольные фиксированные числа и $\epsilon_0 > 0$.

Введем обозначения:

$$\rho_n = \| f'^*(x_n)(fx_n - y)\|, \quad b_n = \|(f'^*(x_n)f'(x_n) + \epsilon_n E)^{-1}\|,$$

$$N_0 = \frac{a_1}{\rho_0}(2\tau_0 \epsilon_0 + \epsilon_0^2), \quad \tau_n = \frac{1}{2}\|f'(x_n)\|^2, \quad \|f'(x)\| \le T < \infty \text{ в } X.$$

Сформулируем следующие условия /6/, связанные с процессом $\mathfrak{R}^{\ell}_{\epsilon_n}$:

а) в X существует ограниченная вторая производная Фреше $f''(x)$:

$$\frac{1}{2}\|f''(x)\| \le M < \infty;$$

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

$$\|f'^*(x) - f'^*(y))f\; z\| \le P\,\|x - y\|^2; \tag{б.1}$$

в) для начальных x_0 и ϵ_0 выполняются неравенства (начальные условия)

$$0 \le \phi_0 \le 1-\beta < 1, \quad \beta < 1, \tag{в.1}$$

$$\beta + \frac{\phi_0}{1-\phi_0} \frac{\gamma^2}{(\gamma - b_0) b_0} \le 1, \tag{в.2}$$

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

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

г) для чисел ϵ_n выполняются неравенства

$$|\epsilon_n - \epsilon_{n+1}| \le \epsilon_n \quad (n = 0,1,2,\ldots) \ ;$$

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

На основе теоремы из работы /6/ и леммы настоящей работы имеет место следующая

Теорема 1

Пусть уравнение (1) решается с помощью процесса $R^{\ell}_{\epsilon_n}$ и условия а)–д) выполнены как для точного $\bar{y} \in x$, так и для любого приближения $\tilde{y} \in X$ ($\| \bar{y} - \tilde{y} \| \le \delta$, $\delta > 0$). Тогда нахождение предела $\tilde{x}^*$ процесса $\mathfrak{R}^{\ell}_{\epsilon_n}$ (относительно уравнения $fx - \tilde{y} = 0$) является регуляризующим алгоритмом.

Утверждения, аналогичные теореме 1, получаются и для регуляризованных итерационных процессов типа Ньютона–Канторовича $\mathfrak{R}_{\omega_n}$ /4/ и $\mathfrak{R}^{\ell}_{\omega_n}$ /5/.

3. Исследование возможности авторегуляризованных итерационных процессов для решения линейных некорректно поставленных задач представляет определенный интерес.

Приведем относительно процесса $\mathfrak{R}^{\ell}_{\epsilon_n}$ одну из возможных теорем о сходимости в линейном случае. Теперь утверждения типа теоремы 3/4/

и теоремы из /6/ упрощаются $(M=P=0)$ и трудно проверяемые условия а) и б) отсутствуют.

В гильбертовом пространстве E рассмотрим линейное уравнение

$$f^*(fx-y)=0, \tag{11}$$

для которого оператор f^*f – обратим.

Теорема 2

Пусть уравнение (11) решается итерационным процессом $\mathfrak{R}^{\ell}_{\epsilon_n}$ и пусть для произвольного фиксированного числа $\epsilon_0>0$ константы γ, α_1 и числа ϵ_n $(n=1,2,\dots)$ выбраны так, чтобы удовлетворялись соотношения

$$\gamma = 3b_0, \tag{12}$$

$$\alpha_1 \leq \frac{1}{4(\epsilon_0^2+2\epsilon_0 T)\gamma^2} \tag{13}$$

(в линейном случае положено $r_n = T$),

$$|\epsilon_n-\epsilon_{n+1}|\leq\epsilon_n. \tag{14}$$

Пусть выполнено условие д) теоремы 1. Тогда для любого начального приближения $x_0\in X$ процесс $\mathfrak{R}^{\ell}_{\epsilon_n}$ сходится и $\lim\limits_{n\to\infty} x_n=x_\epsilon\in W$. Точка x_ϵ является решением уравнения (11) в области X, и справедливы оценки

$$\|x_n-x_\epsilon\|\leq\frac{3b_0\rho_0\beta^{2^{n-1}}}{1-\beta^{2^n}}\leq 12\frac{b_0\rho_0}{4^{2^n}-1} \qquad (n=1,2,\dots).$$

Доказательство

Пусть $\epsilon_0>0$ – фиксированное число. Согласно (12) и (13) выбираем константы γ и α_1. Тогда справедливо неравенство

$$\alpha_1(\epsilon_0^2+2\epsilon_0 T)\gamma^2\leq\frac{1}{4}. \tag{15}$$

Используя обозначения $N_0=\frac{\alpha_1}{\rho_0}(\epsilon_0^2+2\epsilon_0 T)$, $\beta=N_0\gamma^2\rho_0$, неравенство (15) записываем как

$$\beta\leq\frac{1}{4}. \tag{16}$$

В линейном случае $\phi_0=N_0\gamma\rho_0 b_0$ и на основе (16) и (12) находим неравенства

$$\phi_0\leq\frac{1}{12}, \quad \phi_0<1-\beta. \tag{17}$$

Таким образом, выполнено неравенство (B.1) в условии B) теоремы из/6/. Из первого неравенства (17) следует, что

$$3\phi_0+\frac{9}{2}\,\frac{\phi_0}{1-\phi_0}\leq 1. \tag{18}$$

Теперь на основе (12) и равенства $\beta=3\phi_0$ из (18) заключаем, что выполнено неравенство (B.2) условия B) теоремы из/6/. Имея в виду, что выполнено и условие Д) этой же теоремы, делаем вывод о справедливости доказываемой теоремы.

4. Пусть при решении уравнения (11) оператор f^*f необратим. В этом случае итерационный процесс $\mathfrak{R}^{\ell}_{\epsilon_n}$ может быть использован для решения некоторого "аппроксимирующего" уравнения типа

$$f^*fx+\alpha x=f^*y, \tag{19}$$

где $\alpha>0$.

Прерывание процесса $\mathcal{R}^{\ell}_{\epsilon_n}$ для некоторого $n=N^*$ можно связать с параметром α. При решении уравнения (19) с помощью процесса $R^{\ell}_{\epsilon_n}$ имеет место равенство

$$\| f^* f x_{n+1} + \alpha x_{n+1} - f^* y \| = | \epsilon_n - \alpha | \, \| x_n - x_{n+1} \|, \quad (n = 1,2,\dots).$$

Отсюда заключаем, что для $n=N^*$, при котором $\epsilon_{N^*}=\alpha$, точка x_{N^*+1} является решением уравнения (19).

В ряде случаев параметр α может быть неизвестным (т.е. можно поставить задачу на аппроксимацию уравнения (11) посредством (19) с параметром α, являющимся в некотором конкретном смысле "наилучшим").

В качестве примера опишем одну задачу этого класса. Пусть решается линейная система уравнений типа (11) и пусть матрица f^*f близка к вырождению настолько, что данная стандартная программа (на определенной ЭВМ) не может выполнить операцию обращения. Пусть ищется верхняя оценка минимального значения параметра α, для которого задача (19) решается стандартной программой.

Приближенное решение этой задачи часто ищется методами, восходящими к эвристическому программированию (и прежде всего методом "проб-ошибок").

Поставленную выше задачу можно решать и на основе авторегуляризованного процесса $\mathcal{R}^{\ell}_{\epsilon_n}$. В нашем случае имеет место равенство

$$f^* f x_{n+1} + \epsilon_n x_{n+1} - f^* y = \epsilon_n x_n \quad (n=1,2,\dots).$$

Процесс $\mathcal{R}^{\ell}_{\epsilon_n}$ сам остановится на $n=N^*$, для которого обращение матрицы $f^*f+\epsilon_{N^*}E$ уже не выполняется программой. На основе последнего равенства для искомой оценки минимального значения α можно взять число ϵ_{N^*-1}. Надо иметь в виду, что при фиксированном начальном приближении x_0 большим ϵ_0 соответствуют более медленные процессы, а искомая оценка параметра α находится более точно.

Очевидно, решение такой задачи построением авторегуляризованного процесса $\mathcal{R}^{\ell}_{\epsilon_n}$ имеет преимущество перед эвристическими методами.

5. Рассмотрим два примера решения плохо обусловленных систем уравнений авторегуляризованным процессом $\mathcal{R}^{\ell}_{\epsilon_n}$.

Пусть относительно неизвестных x_i, $i=1,\dots,5$ заданы следующие системы уравнений:

$$\left.\begin{aligned}
&1{,}00000001\,x_1 + x_2 + x_3 + x_4 + x_5 = 5{,}00000001,\\
&x_1 + 1{,}00000002\,x_2 + x_3 + x_4 + x_5 = 5{,}00000002,\\
&x_1 + x_2 + 1{,}00000003\,x_3 + x_4 + x_5 = 5{,}00000003,\\
&x_1 + x_2 + x_3 + 1{,}00000004\,x_4 + x_5 = 5{,}00000004,\\
&x_1 + x_2 + x_3 + x_4 + 1{,}00000005\,x_5 = 5{,}00000005.
\end{aligned}\right\} \tag{20}$$

$$\left.\begin{aligned}
&1{,}00000001\,x_1 + x_2 + x_4 + x_5 = 5{,}00000001,\\
&x_1 + 1{,}00000002\,x_2 + x_3 + x_4 + x_5 = 5{,}00000002,\\
&x_1 + x_2 + 1{,}00000003\,x_3 + x_4 + x_5 = 5{,}00000003,\\
&x_1 + x_2 + x_3 + 46{,}0000004\,x_4 + x_5 = 50{,}0000004,\\
&x_1 + x_2 + x_3 + x_4 + 46{,}0000005\,x_5 = 50{,}0000005.
\end{aligned}\right\} \tag{21}$$

Эти системы уравнений плохо обусловлены, так как числа обусловленности оператора f^*f (см., например,/7/) являются большими чем 10^9 (использована норма Чебышева).

В таблице показаны итерационные процессы $\mathcal{R}^{\ell}_{\epsilon_n}$ (реализованные программой AUTORE /6/) соответственно для систем (20) и (21). Уже на шестой итерации в первом случае и на четвертой – во втором найдено решение с точностью до 10^{-9}.

В нашем случае при решении систем (20) и (21) соблюдение условий (12), (13) и (14) теоремы 2 было не обязательно.

На рис. 1 показана зависимость номера N^* окончания итерационного процесса (итерация N^*+1 неосуществима по причине практической необратимости оператора $f^*f + \epsilon_{N^*+1}E$) от начального числа ϵ_0. Отметим еще, что при возмущении правой части системы (21) вектором

$$\Delta y = \begin{bmatrix} 7.5 \; 10^{-8} \\ -1.9 \; 10^{-8} \\ 2.0 \; 10^{-8} \\ -3.0 \; 10^{-} \\ 3.0 \; 10^{-7} \end{bmatrix}$$

изменения процессов $\mathfrak{R}^{\ell}_{\epsilon_n}$ с точностью до 10^{-8} не замечаются.

Пусть решение систем (20) и (21) рассматривается как задача, поставленная в п.4. Из таблицы следует, что для верхних оценок минимальных значений параметров α (для которых соответствующие уравнения типа (19) решаются) можно взять соответственно числа $\epsilon_5 = 0{,}00000105$ и $\epsilon_3 = 0{,}00000153$.

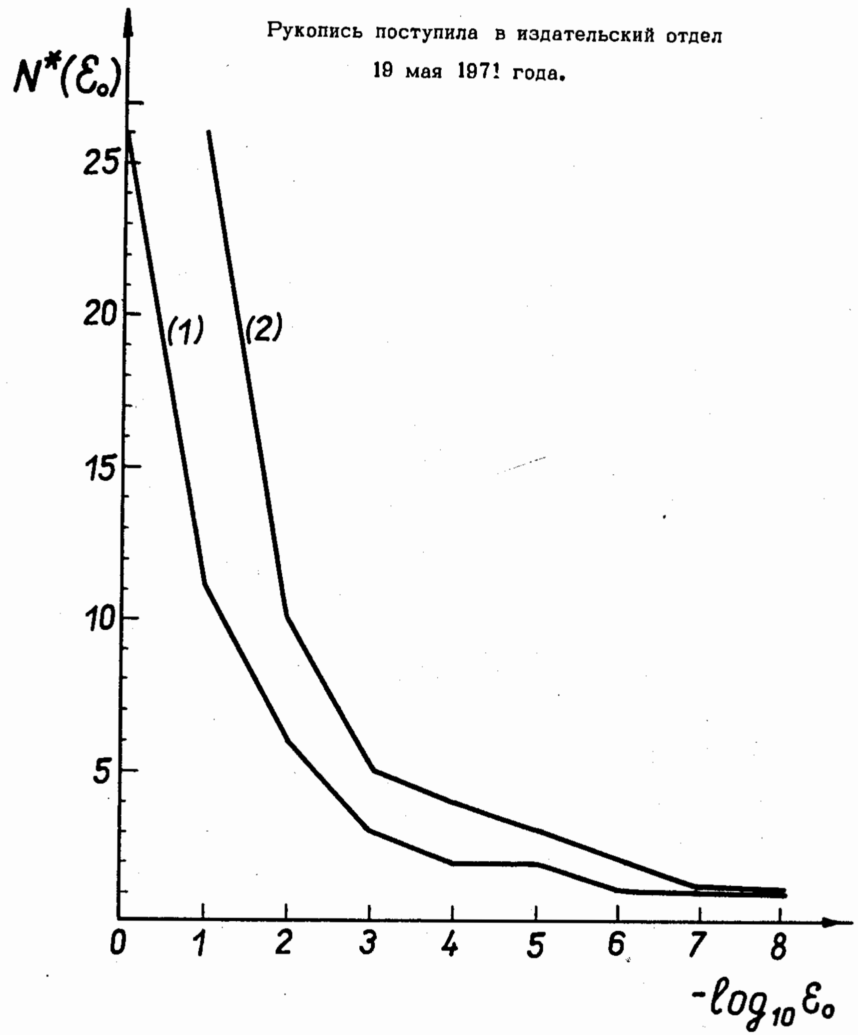
Системы (10) и (11) решались на ЭВМ БЭСМ-6 Объединенного института ядерных исследований. В нашем случае употреблялась стандартная программа **SYMINV** и язык **DUBNA-FORTRAN** /8/.

Литература

1. А.Н. Тихонов. ДАН, 153, №1 (1963).
2. А.Н. Тихонов. ДАН, 161, №5 (1965).
3. А.Н. Тихонов. Статья в сборнике "Вычислительные методы и программирование", изд. МГУ, 1967.
4. Л. Александров. ЖВМ и МФ, т. 11, №1 (1971).
5. Л. Александров. Сообщение ОИЯИ, Р5-5134, Дубна, 1970.
6. Л. Александров. Сообщение ОИЯИ, Р5-5515, Дубна, 1970.

7. Л. Колатц. Функциональный анализ и вычислительная математика, М., 1969 г., стр. 105.
8. Язык ФОРТРАН, ред. В.П. Шириков. Препринт ОИЯИ, 11-4818, Дубна, 1969.

Рукопись поступила в издательский отдел
19 мая 1971 года.



Зависимость номера N^* окончания итерационного процесса $\mathfrak{R}^{\ell}_{\epsilon_n}$ от начального числа ϵ_0. Кривая (1) относится к системе (20), а кривая (2) - к системе (21).

ТАБЛИЦА

Решение системы (20)

m	0	I	2	3	4	5	6
ε_n	I7.I4875I8	8.9524077	2.9747048	0.35699206	0.005I4368	0.00000I05	
C_n	3.33253I3	5.46807I8	I4.4467I27	II3.0473I0I	7822.I037I	38226274.3	
ρ_n	4975.000002	2024.I4I56	533.7I59I6	56.752959	0.79900470	0.000I6344	
τ_n	995.000002	404.8283I2	I06.743I83	II.350592	0.I5980096	0.000032707	
x^n	200	8I.9656622	22.3486366	3.270II836	I.03I960I9	I.00000654	I.0000000
x_2^n	200	8I.9656622	22.3486366	3.270II836	I.03I960I9	I.00000654	I.0000000
x_3^n	200	8I.9656622	22.3486366	3.270II836	I.03I960I9	I.00000654	I.0000000
x_4^n	200	8I.9656622	22.3486366	3.270II836	I.03I960I9	I.00000654	I.0000000
x_5^n	200	8I.9656622	22.3486366	3.270II836	I.03I960I9	I.00000654	I.0000000

Решение системы (2I)

m	0	I	2	3	4	5	6
ε_n	29.6558074	0.39I45855	0.0035508	0.00000I53			
C_n	89.4532634	7977.60082	887975.749I	205439369.2			
ρ_n	4470635.0I	6II7.805I4	55.48I707	0.0239855			
τ_n	9950.0000I	430.87805	I9.35549I	0.00837458			
x_1^n	200.	I49.488458	7.7577248	I.00292402	I.0000000		
x_2^n	200	I49.488458	7.7577248	I.00292402	I.0000000		
x_3^n	200	I49.488458	7.7577248	I.00292402	I.0000000		
x_4^n	200	- 6.2936629	0.54II5834	0.99980I25	0.9999999		
x_5^n	200	- 6.2936629	0.54II5834	0.99980I25	0.9999999		