С.Ф.Гилязов, В.В.Черный

РЕГУЛЯРИЗИРУЮЩИЙ МЕТОД СОПРЯЖЕННЫХ ГРАДИЕНТОВ ДЛЯ РЕШЕНИЯ ЛИНЕЙНЫХ ОПЕРАТОРНЫХ УРАВНЕНИЙ С САМОСОПРЯЖЕННЫМ ОПЕРАТОРОМ[1]

Аннотация

Обосновывается регуляризирующий алгоритм на основе итерационного метода сопряженных градиентов для решения линейных некоррректных операторных уравнений первого рода с самосопряженным незнакоопределенным оператором. Устойчивость по отношению к возмущениям исходных данным достигается за счет согласования номера итераций (момента останова итерационного процесса) и уровня погрешности исходных данных.

1. Введение

Итерационный метод сопряженных градиентов широко используется для решения некорректных линейных операторных уравнений первого рода в гильбертовых пространствах [1–5]. Регуляризирующий эффект достигается за счет согласования номера итераций (момента останова) и уровня погрешности исходных данных. В качестве критерия согласования может быть использован, например, критерий невязки.

Для произвольного линейного ограниченного оператора задача решения уравнения первого рода сводится к задаче (некорректной) минимизации нормы невязки уравнения. В специальном случае самосопряженного неотрицательно определенного оператора решается задача минимизации соответствующей билинейной формы.

В случае самосопряженного незнакоопределенного, т.е. имеющего незнакоопределенную билинейную форму, оператора метод сопряженных градиентов может быть реализован по тем же формулам, что и в случае самосопряженного знакоопределенного оператора. Сходимость такого итерационного метода для решения систем линейных алгебраических уравнений изучалась в работах [6, 7].

Настоящая статья посвящена исследованию регуляризирующего алгоритма на основе итерационного метода сопряженных градиентов для

[1]Работа выполнена при финансовой поддержке РФФИ (грант 98–01–00345а) и ФНП "Университеты России–фундаментальные исследования" (грант 8854)

решения линейных некоррректных операторных уравнений первого рода с самосопряженным незнакоопределенным оператором.

2. Постановка задачи

1. Рассмотрим задачу решения линейного операторного уравнения

$$Au = f, \tag{1}$$

где $A : U \to U$ — линейный ограниченный самосопряженный оператор, действующий в гильбертовом пространстве U, $f \in U$ — заданный, а $u \in U$ — искомый элемент. Пространство U предполагается вещественным. Всюду в дальнейшем мы будем считать, что уравнение (1) имеет единственное решение $\hat{u} \in U$.

Если дополнительно известно, что оператор A положительный, то есть $(Au, u) > 0$ для всех $u \in U, \|u\| \neq 0$, то задача решения уравнения (1) эквивалентна задаче минимизации квадратичного функционала $J(u) = (Au, u) - 2(u, f)$ [8, стр. 80]. Аналогично, если $(Au, u) < 0$, то минимизируется функционал $-J(u)$.

Если квадратичная форма (Au, u) незнакоопределена, то минимум функционала $J(u)$ на всем пространстве U может не существовать даже в случае разрешимости уравнения (1).

Пример 1. *Пусть A — линейный самосопряженный вполне непрерывный оператор, имеющий собственные значения разных знаков. Выберем собственный элемент y_0, соответствующий собственному числу $\lambda_0 < 0$, и рассмотрим последовательность элементов пространства U вида $y_k = ky_0 + \hat{u}$, $k = 1, 2, \ldots$. Тогда значение функционала на последовательности y_k будет выражаться формулой*

$$\begin{aligned} J(y_k) &= (Ay_k, y_k) - 2\,(y_k, f) = \\ &= \lambda_0 k^2 \|y_0\|^2 - (\hat{u}, f) \end{aligned}$$

и $\lim\limits_{k\to\infty} J(y_k) = -\infty$.

Аналогично строится последовательность $z_k : \lim\limits_{k\to\infty} J(z_k) = \infty$. Таким образом, функционалы $J(u)$ и $-J(u)$ не имеют минимума на всем пространстве U.

2. Для вычисления решения уравнения (1) с самосопряженным оператором A по заданному начальному приближению $u_0 \in U$ построим последовательность $\{u_k\}$ вида

$$u_{k+1} = u_k - \tau_k p_k, \quad k = 0, 1, \ldots, \tag{2}$$

$$p_{k+1} = g_{k+1} + \beta_k p_k, \;\; g_k = Au_k - f, \; p_0 = g_0, \tag{3}$$

$$\tau_k = \frac{\|g_k\|^2}{(Ap_k, p_k)}, \qquad \beta_k = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}. \tag{4}$$

Свойства последовательности (2)–(4) описывает

Лемма 1. *[9, 10] Пусть оператор A самосопряженный, тогда для всех номеров k таких, что $(Ap_k, p_k) \neq 0$, $\|g_k\| \neq 0$ выполнены равенства*

$$(g_i, g_j) = 0, \;\; (Ap_i, p_j) = 0, \;\; (g_j, u_k - \hat{u}) = (g_k, u_k - \hat{u}), \;\; 0 \leq i \neq j \leq k. \tag{5}$$

Если оператор A положительный, то соотношения (2)–(4) определяют метод сопряженных градиентов для минимизации функционала $(Au, u) - 2(u, f)$.

Справедлива

Теорема 1. *[4, стр. 105] Если $A > 0$, то для любого начального приближения $u_0 \in U$ последовательность метода (2)–(4) минимизирует функционал $(Au, u) - 2(u, f)$:*

$$\lim_{k\to\infty} (Au_k, u_k) - 2(u_k, f) = -(\hat{u}, f)$$

и сильно сходится к решению уравнения (1).

Если квадратичная форма (Au, u) незнакоопределена, то метод (2)–(4) в общем случае нельзя рассматривать как метод минимизации функционала $(Au, u) - 2(u, f)$. В работах [6, 7, 10] для обоснования сходимости метода в конечномерном пространстве использовано свойство ортогональности невязок g_k. Из результатов работ [6, 7] вытекает

Теорема 2. *Пусть A — симметричная матрица порядка $n \times n$. Если в соотношениях (2)–(4) $(Ap_k, p_k) \neq 0$, $0 \leq k \leq n-1$, то $\|g_n\| = 0$ и $u_n = \hat{u}$.*

Для самосопряженного оператора в бесконечномерном пространстве достаточные условия сходимости метода (2)–(4) не исследованы.

3. Будем предполагать, что исходные данные для уравнения (1), то есть оператор A и правая часть f, заданы с погрешностями. А именно, определены линейный ограниченный самосопряженный оператор $\tilde{A} : U \to F$, элемент $\tilde{f} \in U$, а также числа $\eta \geq 0$, $\delta \geq 0$, $\eta^2 + \delta^2 > 0$ такие, что

$$\|\tilde{A} - A\| \leq \eta, \;\; \|\tilde{f} - f\| \leq \delta. \tag{6}$$

Решение уравнения (1) не предполагается непрерывно зависящим от возмущений исходных данных, то есть задача является некорректно поставленной [11]. Последовательность $\{\tilde{u}_k\}$, задаваемая соотношениями

(2)–(4) с заменой A, f на $\tilde{A}$, $\tilde{f}$, даже если и сходится при $k \to \infty$, то ее предел может быть сколь угодно сильно удален от решения $\hat{u}$ уравнения (1). Однако можно указать условия согласования номера итерации k с уровнем погрешности исходных данных $\sigma = (\eta, \delta)$ такие, что семейство $\{u_{k(\sigma)}\}$ сходится к $\hat{u}$, причем как в смысле слабой сходимости, так и сходимости по норме пространства U. Для метода (2)–(4) с положительным оператором A такие условия получены в [2-5].

3. Доказательство сходимости регуляризирующего метода сопряженных градиентов

1. Рассмотрим итерационный процесс (2)–(4) в предположении незнакоопределенности квадратичной формы (Au, u).

Для последовательности (2)–(4) введем обозначение

$$\theta_k^{-1} = \sum_{i=0}^{k-1} \frac{\|p_i\|^2}{|(Ap_i, p_i)|}. \tag{7}$$

Лемма 2. *Справедлива формула*

$$\theta_k^{-1} = \sum_{j=0}^{k-1} \frac{1}{\|g_j\|^2} \left(\sum_{i=j}^{k-1} |\tau_i| \, \|g_i\|^2 \right). \tag{8}$$

Доказательство. Так как формулы, задающие метод сопряженных градиентов в точном и приближенном случаях, совпадают с точностью до обозначений величин, мы будем проводить рассуждения для точного случая, подразумевая при этом, что все сказанное верно и для приближенных данных.

Используя определение поправки p_i, имеем $p_i = \sum_{j=0}^{i} \frac{\|g_i\|^2}{\|g_j\|^2} g_j$, откуда, используя ортогональность $\{g_j\}$, получаем $\|p_i\|^2 = \sum_{j=0}^{i} \frac{\|g_i\|^4}{\|g_j\|^2}$. Отсюда следует равенство

$$\frac{\|p_i\|^2}{\|g_i\|^2} = \sum_{j=0}^{i} \frac{\|g_i\|^2}{\|g_j\|^2}. \tag{9}$$

Меняя порядок суммирования в (8) и используя формулу (4) для τ_k, с учетом равенства (9) имеем

$$\sum_{j=0}^{k-1} \frac{1}{\|g_j\|^2} \left(\sum_{i=j}^{k-1} |\tau_i| \, \|g_i\|^2 \right) = \sum_{i=0}^{k-1} |\tau_i| \sum_{j=0}^{i} \frac{\|g_i\|^2}{\|g_j\|^2} = \sum_{i=0}^{k-1} \frac{\|p_i\|^2}{|(Ap_i, p_i)|}.$$

Лемма доказана.

В дальнейшем нам потребуется

Лемма 3. *Пусть семейство $\{v_\nu\} \subset V$ ограничено, V — гильбертово пространство, $0 < \nu < +\infty$ и пусть $J(v)$ — неотрицательный слабо полунепрерывный снизу функционал, определенный на замкнутом выпуклом множестве $M \subset V$, достигающий минимума в единственной точке $\hat{v} \in M$, то есть $J(\hat{v}) = \min\limits_{v \in M} J(v) = 0$.*

Тогда если $\lim\limits_{\nu \to 0} J(v_\nu) = 0$, то v_ν сходится слабо к $\hat{v}$ при $\nu \to 0$.

Доказательство полностью аналогично доказательству теоремы 2 из [10, стр. 49].

2. Справедлива следующая

Теорема 3. *Пусть для последовательности $\{u_k\}$, определяемой соотношениями (2)–(4), выполнены условия*

$$(Ap_k, p_k) \neq 0 \ \forall k \geq 0, \ \ \sup_{k \geq 0} \theta_k^{-1} \|g_k\| < +\infty.$$

Тогда $\{u_k\}$ слабо сходится к $\hat{u}$ при $k \to \infty$.

Если дополнительно $\lim\limits_{k \to \infty} \theta_k^{-1} \|g_k\| = 0$, то $\lim\limits_{k \to \infty} \|u_k - \hat{u}\| = 0$.

Доказательство. Просуммируем равенства $u_{i+1} = u_i - \tau_i p_i$ для $i = 0, \ldots, k-1$ и вычтем $\hat{u}$:

$$u_k - \hat{u} = u_0 - \hat{u} - \sum_{i=0}^{k-1} \tau_i p_i. \tag{10}$$

Подставим в (10) полученное в лемме 2 выражение для поправки p_i, тогда

$$u_k - \hat{u} \ = u_0 - \hat{u} - \sum_{i=0}^{k-1} \tau_i p_i = u_0 - \hat{u} - \sum_{i=0}^{k-1} \tau_i \sum_{j=0}^{i} \frac{\|g_i\|^2}{\|g_j\|^2} g_j. \tag{11}$$

Для удобства обозначений положим $z_k = u_k - \hat{u}$. Умножим (11) скалярно на z_k и поменяем порядок суммирования

$$\|z_k\|^2 = (z_0, z_k) - \sum_{j=0}^{k-1} \frac{1}{\|g_j\|^2} \left(\sum_{i=j}^{k-1} \tau_i \|g_i\|^2 \right) (g_j, z_k). \tag{12}$$

В силу равенства $(g_j, z_k) = (g_k, z_k)$ (см. Лемму 1) из (12) получаем

$$\|z_k\|^2 = (z_0, z_k) - \sum_{j=0}^{k-1} \frac{1}{\|g_j\|^2} \left(\sum_{i=j}^{k-1} \tau_i \|g_i\|^2 \right) (g_k, z_k). \tag{13}$$

Применяя к правой части (13) неравенство Коши–Буняковского $-(g_k, z_k) \leq \|g_k\|\|z_k\|$, с учетом леммы 2 получаем

$$\begin{aligned} \|z_k\|^2 &\leq \|z_0\|\,\|z_k\| + \left\| \sum_{j=0}^{k-1} \tfrac{1}{\|g_j\|^2} \left(\sum_{i=j}^{k-1} \tau_i \,\|g_i\|^2 \right) g_k \right\| \|z_k\| \leq \\ &\leq \left(\|z_0\| + \theta_k^{-1} \|g_k\| \right) \|z_k\| . \end{aligned} \tag{14}$$

Из условия теоремы и (14) следует, что последовательность $\{u_k\}$ ограничена.

Далее, заметим, что $\theta_k^{-1} \to \infty$ при $k \to \infty$. Действительно, в силу неравенства Коши–Буняковского $|(Ap_i, p_i)| \leq \|A\|\,\|p_i\|^2$, поэтому $\theta_k^{-1} \geq \frac{k}{\|A\|} \to \infty$ при $k \to \infty$. Из условия теоремы получаем, что

$$\|g_k\| = \left(\theta_k^{-1}\|g_k\|\right)\theta_k \leq \theta_k \sup_k \theta_k^{-1}\|g_k\| \to 0 \text{ при } k \to \infty.$$

Таким образом, для последовательности $\{u_k\}$ выполнены условия леммы 3 и, следовательно, $u_k \rightharpoonup \hat{u}$.

Пусть выполнено дополнительное условие теоремы: $\lim\limits_{k\to\infty} \theta_k^{-1}\|g_k\| = 0$. В этом случае в силу первой части теоремы $(z_0, z_k) \to 0$ при $k \to \infty$. Кроме того, из (13) следует оценка

$$\|z_k\|^2 \leq |(z_0, z_k)| + \theta_k^{-1}\|g_k\|\|z_k\|.$$

Решая квадратное неравенство относительно $\|z_k\|$, получаем

$$\|z_k\| \leq \sqrt{|(z_0, z_k)|} + \theta_k^{-1}\|g_k\| \to \infty \text{ при } k \to \infty.$$

Теорема доказана.

3. Пусть в методе (2)–(4) $k = k_0 \geq 0$ — первый номер итерации такой, что $(Ap_{k_0}, p_{k_0}) = 0$, тогда итерации для $k > k_0$ построить невозможно. Если при этом $p_{k_0} = 0$, то $g_{k_0} = 0$ и решение уравнения (1) вычислено. Действительно, для $k_0 > 0$ в силу (9) имеем

$$0 = \|p_{k_0}\|^2 = \|g_{k_0}\|^2 + \sum_{j=0}^{k_0-1} \frac{\|g_{k_0}\|^4}{\|g_j\|^2} \geq \|g_{k_0}\|^2 .$$

Если $k_0 = 0$, то $p_0 = g_0$ по определению.

Для незнакоопределенной квадратичной формы (Ap_{k_0}, p_{k_0}) из равенства $(Ap_{k_0}, p_{k_0}) = 0$, вообще говоря, не следует равенство $p_{k_0} = 0$. Пусть

$(Ap_{k_0}, p_{k_0}) = 0$, $k_0 \geq 0$, но $g_{k_0} \neq 0$. Если $k_0 = 0$, то следует выбрать другое начальное приближение. Пусть $k_0 > 0$, тогда, используя (3), имеем

$$\begin{aligned} (Ap_{k_0}, p_{k_0}) &= (A(g_{k_0} + \beta_{k_0-1}p_{k_0-1}), g_{k_0} + \beta_{k_0-1}p_{k_0-1}) = \\ &= (Ag_{k_0}, g_{k_0}) + 2\beta_{k_0-1}(Ap_{k_0-1}, g_{k_0}) + \\ &\quad + \beta_{k_0-1}^2 (Ap_{k_0-1}, p_{k_0-1}). \end{aligned} \tag{15}$$

Применим к обеим частям (2) оператор A и вычтем f, тогда

$$Ap_{k_0-1} = \frac{1}{\tau_{k_0-1}}(g_{k_0-1} - g_{k_0}).$$

Используя лемму 1, имеем

$$\begin{aligned} 2\beta_{k_0-1}(Ap_{k_0-1}, g_{k_0}) &= 2\frac{\|g_{k_0}\|^2}{\|g_{k_0-1}\|^2}\frac{(Ap_{k_0-1}, p_{k_0-1})}{\|g_{k_0-1}\|^2}(g_{k_0-1} - g_{k_0}, g_{k_0}) = \\ &= -2\beta_{k_0-1}^2 (Ap_{k_0-1}, p_{k_0-1}). \end{aligned} \tag{16}$$

Складывая (15) и (16), получаем

$$(Ap_{k_0}, p_{k_0}) = (Ag_{k_0}, g_{k_0}) - \beta_{k_0-1}^2 (Ap_{k_0-1}, p_{k_0-1}). \tag{17}$$

Из (17) следует, что $(Ag_{k_0}, g_{k_0}) \neq 0$. Действительно, если это не так, то

$$\beta_{k_0-1}^2 (Ap_{k_0-1}, p_{k_0-1}) = (Ag_{k_0}, g_{k_0}) - (Ap_{k_0}, p_{k_0}) = 0,$$

что противоречит предположению.

Используя приближение u_{k_0}, вычислим следующее приближение по формуле

$$u_{k_0+1} = u_{k_0} - \frac{\|g_{k_0}\|^2}{(Ag_{k_0}, g_{k_0})} g_{k_0}. \tag{18}$$

Для $A > 0$ формула (18) совпадает с формулой метода наискорейшего спуска для минимизации функционала $(Au, u) - 2(u, f)$. Заметим, что первая итерация метода (2)–(4) также выполняется по формуле (18).

Укажем условия, при выполнении которых $\|u_{k_0+1} - \hat{u}\| \leq \|u_{k_0} - \hat{u}\|$. Вычтем из обеих частей (18) решение $\hat{u}$ уравнения (1):

$$z_{k_0+1} = z_{k_0} - \frac{\|g_{k_0}\|^2}{(Ag_{k_0}, g_{k_0})} g_{k_0}. \tag{19}$$

Используя формулу (19), перейдем к равенству норм

$$\|z_{k_0+1}\|^2 = \|z_{k_0}\|^2 - 2\frac{\|g_{k_0}\|^2}{(Ag_{k_0},g_{k_0})}(g_{k_0},z_{k_0}) + \frac{\|g_{k_0}\|^4}{(Ag_{k_0},g_{k_0})^2}\|g_{k_0}\|^2. \tag{20}$$

Используя свойства скалярного произведения и формулу (19), получаем

$$\begin{aligned}(g_{k_0+1},z_{k_0+1}) &= (Az_{k_0+1},z_{k_0+1}) = \\ &= (g_{k_0},z_{k_0}) - 2\frac{\|g_{k_0}\|^2}{(Ag_{k_0},g_{k_0})}\|g_{k_0}\|^2 + \frac{\|g_{k_0}\|^4}{(Ag_{k_0},g_{k_0})^2}(Ag_{k_0},g_{k_0}) = \\ &= (g_{k_0},z_{k_0}) - \frac{\|g_{k_0}\|^4}{(Ag_{k_0},g_{k_0})}.\end{aligned} \tag{21}$$

Преобразуем равенство (20) с учетом (21)

$$\begin{aligned}\|z_{k_0}\|^2 - \|z_{k_0+1}\|^2 &= \frac{\|g_{k_0}\|^2}{(Ag_{k_0},g_{k_0})}\left[(g_{k_0},z_{k_0}) + (g_{k_0},z_{k_0}) - \frac{\|g_{k_0}\|^4}{(Ag_{k_0},g_{k_0})^2}\right] = \\ &= \frac{\|g_{k_0}\|^2}{(Ag_{k_0},g_{k_0})}\left[(g_{k_0},z_{k_0}) + (g_{k_0+1},z_{k_0+1})\right] = \\ &= \frac{1}{\|g_{k_0}\|^2}\left[(g_{k_0},z_{k_0})^2 - (g_{k_0+1},z_{k_0+1})^2\right].\end{aligned}$$

Отсюда следует, что если $(g_{k_0},z_{k_0})^2 \geq (g_{k_0+1},z_{k_0+1})^2$, то $\|z_{k_0+1}\|^2 \leq \|z_{k_0}\|^2$.

4. Пусть исходные данные уравнения (1) заданы с погрешностью и выполнены условия аппроксимации (6). Рассмотрим последовательность $\{u_k\}$, определяемую соотношениями

$$\tilde{u}_{k+1} = \tilde{u}_k - \tilde{\tau}_k\tilde{p}_k, \quad k = 0,1,\ldots, \tag{22}$$

$$\tilde{p}_{k+1} = \tilde{g}_{k+1} + \tilde{\beta}_k\tilde{p}_k, \ \ \tilde{g}_k = \tilde{A}\tilde{u}_k - \tilde{f}, \ \tilde{p}_0 = \tilde{g}_0, \tag{23}$$

$$\tilde{\tau}_k = \frac{\|\tilde{g}_k\|^2}{(\tilde{A}\tilde{p}_k,\tilde{p}_k)}, \qquad \tilde{\beta}_k = \frac{\|\tilde{g}_{k+1}\|^2}{\|\tilde{g}_k\|^2}. \tag{24}$$

Определим величины

$$\tilde{\theta}_k^{-1} = \sum_{i=0}^{k-1}\frac{\|\tilde{p}_i\|^2}{|(\tilde{A}\tilde{p}_i,\tilde{p}_i)|},$$

$$\Delta = \eta\|\hat{u}\| + \delta,$$

которые будут использованы в дальнейшем.

Теорема 4. *Пусть для последовательности $\{\tilde{u}_k\}$, определяемой соотношениями (22)–(24), выполнены условия*

$$\lim_{\sigma\to 0}\sup_{k\geq 0}\tilde{\theta}_k^{-1}\|\tilde{g}_k\| < +\infty, \quad (\tilde{A}\tilde{p}_k, \tilde{p}_k) \neq 0 \ \forall k \geq 0,$$

и пусть номер $k(\sigma)$ выбран так, что

$$\lim_{\sigma\to 0}\tilde{\theta}_{k(\sigma)} = 0, \ \lim_{\sigma\to 0}\tilde{\theta}_{k(\sigma)}^{-1}\sqrt{\eta+\delta} < +\infty.$$

Тогда $\tilde{u}_{k(\sigma)}$ слабо сходится к $\hat{u}$ при $\sigma \to 0$. Если дополнительно выполнены условия

$$\lim_{\sigma\to 0}\tilde{\theta}_{k(\sigma)}^{-1}\left\|\tilde{g}_{k(\sigma)}\right\| = 0, \ \lim_{\sigma\to 0}\tilde{\theta}_{k(\sigma)}^{-1}\sqrt{\eta+\delta} = 0,$$

то

$$\lim_{\sigma\to 0}\|\tilde{u}_{k(\sigma)} - \hat{u}\| = 0.$$

Доказательство. При доказательстве данной теоремы мы будем опираться на доказательство теоремы 3. Аналогично (12) выводится соотношение

$$\|\tilde{z}_k\|^2 = (z_0, \tilde{z}_k) - \sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left(\sum_{i=j}^{k-1}\tilde{\tau}_i\|\tilde{g}_i\|^2\right)(\tilde{g}_j, \tilde{z}_k), \tag{25}$$

где $\tilde{z}_k = \tilde{u}_k - \hat{u}$.

Преобразуем выражение $(\tilde{g}_j, \tilde{z}_k)$. Выражая $\tilde{u}_j$ через $\tilde{u}_k$ для $j \leq k$, применяя оператор $\tilde{A}$, вычитая из обеих частей равенства элемент $\tilde{f}$ и умножая полученное равенство скалярно на $\tilde{z}_k$, получаем

$$\begin{aligned}(\tilde{g}_j, \tilde{z}_k) &= (\tilde{g}_k, \tilde{z}_k) + \sum_{s=j}^{k-1}\tilde{\tau}_s(\tilde{A}\tilde{p}_s, \tilde{z}_k) = \\ &= (\tilde{g}_k, \tilde{z}_k) + \sum_{s=j}^{k-1}\tilde{\tau}_s(\tilde{p}_s, \tilde{A}\tilde{u}_k - \tilde{f} + \tilde{f} - \tilde{A}\hat{u}) = \\ &= (\tilde{g}_k, \tilde{z}_k) + \sum_{s=j}^{k-1}\tilde{\tau}_s(\tilde{p}_s, \tilde{g}_k) + \sum_{s=j}^{k-1}\tilde{\tau}_s(\tilde{p}_s, \tilde{f} - \tilde{A}\hat{u}).\end{aligned} \tag{26}$$

Из леммы 1 об ортогональности $\{\tilde{g}_k\}$ для всех $s < k$ следуют равенства $(\tilde{p}_s, \tilde{g}_k) = \sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2}(\tilde{g}_l, \tilde{g}_k) = 0$. С учетом этих равенств выражение (25) принимает вид

$$\begin{aligned}\|\tilde{z}_k\|^2 &= (z_0, \tilde{z}_k) - \sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left(\sum_{i=j}^{k-1}\tilde{\tau}_i\|\tilde{g}_i\|^2\right)(\tilde{g}_k, \tilde{z}_k) + \\ &+ \sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left(\sum_{i=j}^{k-1}\tilde{\tau}_i\|\tilde{g}_i\|^2\right)\sum_{s=j}^{k-1}\tilde{\tau}_s\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2}(\tilde{g}_l, \tilde{A}\hat{u} - \tilde{f}).\end{aligned} \tag{27}$$

Используя неравенство Коши–Буняковского, оценки $\|\tilde{g}_l\| \le \sup\limits_{0\le i\le k} \|\tilde{g}_i\|$ и $\left\|\tilde{A}\hat{u}-\tilde{f}\right\| \le \Delta$, имеем

$$\begin{aligned}
\|\tilde{z}_k\|^2 &< \|z_0\|\,\|\tilde{z}_k\| + \left\|\sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left(\sum_{i=j}^{k-1}\tilde{\tau}_i\,\|\tilde{g}_i\|^2\right)\tilde{g}_k\right\|\|\tilde{z}_k\| + \\
&\quad + \sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left|\sum_{i=j}^{k-1}\tilde{\tau}_i\,\|\tilde{g}_i\|^2\right|\sum_{s=j}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2}\|\tilde{g}_l\|\left\|\tilde{A}\hat{u}-\tilde{f}\right\| < \\
&< \|z_0\|\,\|\tilde{z}_k\| + \tilde{\theta}_k^{-1}\|\tilde{g}_k\|\,\|\tilde{z}_k\| + \\
&\quad + \Delta\sup_{0\le i\le k}\|\tilde{g}_i\|\sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left|\sum_{i=j}^{k-1}\tilde{\tau}_i\,\|\tilde{g}_i\|^2\right|\sum_{s=j}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2}.
\end{aligned} \tag{28}$$

Оценим величину

$$\sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left|\sum_{i=j}^{k-1}\tilde{\tau}_i\,\|\tilde{g}_i\|^2\right|\sum_{s=j}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2}.$$

Изменяя пределы суммирования с $s=j$ на $s=0$ и затем внося знак модуля под знак суммы, получаем

$$\begin{aligned}
\sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left|\sum_{i=j}^{k-1}\tilde{\tau}_i\,\|\tilde{g}_i\|^2\right|\sum_{s=j}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2} &< \sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\left|\sum_{i=j}^{k-1}\tilde{\tau}_i\,\|\tilde{g}_i\|^2\right|\sum_{s=0}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2} \le \\
&\le \sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\sum_{i=j}^{k-1}|\tilde{\tau}_i|\,\|\tilde{g}_i\|^2\sum_{s=0}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2}.
\end{aligned}$$

Преобразуем правую часть полученного неравенства. Меняя порядок суммирования по индексам i, j и используя лемму 2, получаем

$$\sum_{j=0}^{k-1}\frac{1}{\|\tilde{g}_j\|^2}\sum_{i=j}^{k-1}|\tilde{\tau}_i|\,\|\tilde{g}_i\|^2\sum_{s=0}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2} = \left(\sum_{s=0}^{k-1}|\tilde{\tau}_s|\sum_{l=0}^{s}\frac{\|\tilde{g}_s\|^2}{\|\tilde{g}_l\|^2}\right)^2 = \tilde{\theta}_k^{-2}.$$

Подставляя полученную оценку в (28), имеем

$$\|\tilde{z}_k\|^2 < \left(\|z_0\| + \tilde{\theta}_k^{-1}\|\tilde{g}_k\|\right)\|\tilde{z}_k\| + \sup_{0\le i\le k}\|\tilde{g}_i\|\,\tilde{\theta}_k^{-2}\Delta.$$

Выразим отсюда оценку для $\|\tilde{z}_k\|$:

$$\|\tilde{z}_k\| \le \|z_0\| + \tilde{\theta}_k^{-1}\|\tilde{g}_k\| + \sqrt{\sup_{0\le i\le k}\|\tilde{g}_i\|\,\tilde{\theta}_k^{-2}\Delta}.$$

Заметим, что $\lim\limits_{\sigma\to 0}\sup\limits_{0\le i\le k}\|\tilde g_i\| < +\infty$. Действительно, из условия теоремы следует неравенство

$$\|\tilde g_k\| = \tilde\theta_k(\tilde\theta_k^{-1}\|\tilde g_k\|) \le \tilde\theta_k \sup_{0\le i\le k}(\tilde\theta_k^{-1}\|\tilde g_k\|) \le const\frac{|(\tilde A\tilde p_0, \tilde p_0)|}{\|\tilde p_0\|^2},$$

и тогда

$$\lim_{\sigma\to 0}\sup_{0\le i\le k}\|\tilde g_k\| \le const\|A\| < +\infty.$$

Пусть номер $k(\sigma)$ выбран в соответствии с условиями теоремы, тогда семейство $\left\{\tilde u_{k(\sigma)}\right\}$ ограничено. С другой стороны, из условия теоремы следует, что семейство $\left\{\tilde u_{k(\sigma)}\right\}$ минимизирует функционал $\|Au - f\|$, так как

$$\begin{aligned}
\left\|A\tilde u_{k(\sigma)} - f\right\| &\le \eta\left\|\tilde u_{k(\sigma)}\right\| + \left\|\tilde g_{k(\sigma)}\right\| + \delta \le \\
&\le \eta\left\|\tilde u_{k(\sigma)}\right\| + \tilde\theta_{k(\sigma)}^{-1}\left\|\tilde g_{k(\sigma)}\right\|\tilde\theta_{k(\sigma)} + \delta < \\
&< const\cdot\left(\eta + \delta + \tilde\theta_{k(\sigma)}\right) \to 0
\end{aligned}$$

при $\sigma \to 0$. Таким образом, для семейства $\left\{\tilde u_{k(\sigma)}\right\}$ выполнены условия леммы 3 и, следовательно, $\tilde u_{k(\sigma)} \rightharpoonup \hat u$ при $\sigma \to 0$.

Пусть выполнены дополнительные условия теоремы. Повторяя рассуждения в первой части доказательства, запишем оценку

$$\|\tilde z_k\|^2 \le (z_0, \tilde z_k) + \tilde\theta_k^{-1}\|\tilde g_k\|\|\tilde z_k\| + \sup_{0\le i\le k}\|\tilde g_i\|\tilde\theta_k^{-2}\Delta, \tag{29}$$

откуда в силу слабой сходимости $\tilde z_{k(\sigma)} \rightharpoonup 0$ получаем $\tilde u_{k(\sigma)} \to \hat u$ при $\sigma \to 0$.

Теорема доказана.

4. Численные результаты

Возьмем в качестве модельной задачу решения интегрального уравнения, оператор которого имеет собственные числа разных знаков [12]. Пусть решается интегральное уравнение

$$\int\limits_{-\pi}^{\pi}\mathcal K(x-s)u(s)ds = f(x), \quad x\in[-\pi,\pi], \tag{30}$$

$$\mathcal K(x-s) = \frac{1}{2\pi}\frac{1-h^2}{1-2h\cos(x-s)+h^2}, \quad h = -0.8.$$

Известно [12, стр. 105], что интегральный оператор с таким ядром имеет при $|h| < 1$ собственные функции 1, $\sin kx$, $\cos kx$, соответствующие собственным числам 1, h^k, h^k. Возьмем в качестве точного решения функцию $u(s) = \pi^2 - s^2$ и правую часть уравнения будем вычислять как значение оператора на точном решении.

Для дискретизации уравнения (30) введем равномерную сетку узлов по переменным x и s: $x_i = a+(i-1)h$, $s_j = a+(j-1)h$, где $h = (b-a)/(N-1)$ — шаг, а N — число узлов сетки. Используя квадратурную формулу трапеций, аппроксимируем интегральное уравнение системой линейных алгебраических уравнений

$$A_{\text{Д}} u = f_{\text{Д}}, \tag{31}$$

где $A_{\text{Д}}$ – квадратная матрица с элементами вида $a_{ij} = \gamma_j K(x_i, s_j)$, $i, j = 1 \dots N$,

$$\gamma_j = \begin{cases} \frac{h}{2}, & j = 1, \\ h, & j = 2 \dots N-1, \\ \frac{h}{2}, & j = N, \end{cases} \tag{32}$$

— квадратурные коэффициенты формулы трапеций.

Будем моделировать погрешности в правой части следующим образом. Примем в качестве вектора правой части вектор $\tilde{f}$ с координатами

$$\tilde{f}_i = (1 + \delta\xi)(f_{\text{Д}})_i,$$

где ξ — равномерно распределенная случайная величина на отрезке $[-1, 1]$, $\delta \in [0, 1]$ — уровень возмущения, $f_{\text{Д}}$ — вектор, вычисленный умножением матрицы оператора на дискретизованное точное решение.

Таким образом, мы получаем систему линейных алгебраических уравнений

$$A_{\text{Д}} u = \tilde{f}, \tag{33}$$

для решения которой будем применять метод (22)–(24).

В качестве правила останова итераций метода мы будем рассматривать два критерия:

1) Критерий невязки. Процесс останавливается, как только для очередного номера k выполняется неравенство

$$\|A_{\text{Д}}\tilde{u}_k - \tilde{f}\|_2 \leq \delta \|A_{\text{Д}} u_0 - \tilde{f}\|_2. \tag{34}$$

2) θ–критерий. Процесс останавливается, как только для очередного номера k выполняется неравенство

$$\tilde{\theta}_k^{-1} \geq \frac{1}{\sqrt{\delta}}. \tag{35}$$

Здесь
$$\tilde{\theta}_k^{-1} = \sum_{m=0}^{k-1} \frac{\|\tilde{p}_m\|_2^2}{|(A_{\text{Д}}\tilde{p}_m, \tilde{p}_m)|},$$
$$\|x\|_2 = \sqrt{(x,x)},\ (x,y) = \frac{1}{2}x_1 y_1 + \sum_{l=2}^{N-1} x_l y_l + \frac{1}{2}x_N y_N.$$

Отметим, что θ–критерий выражает условие ограниченности величины $\tilde{\theta}_{k(\sigma)}^{-1}\sqrt{\eta+\zeta+\delta}$ (см. условие теоремы 4).

Запишем алгоритм решения уравнения (33) на основе метода (22)–(24).

Шаг 1. Выбрать начальное приближение $u_0 \in U : (A_{\text{Д}}\tilde{g}_0, \tilde{g}_0) \neq 0$, задать $k = 0$.

Шаг 2. Если $\|A_{\text{Д}}\tilde{u}_k - \tilde{f}\|_2 = 0$, то задача решена и точное решение найдено.

Шаг 3. Если $(A_{\text{Д}}\tilde{p}_k, \tilde{p}_k) = 0$, принять направление спуска равным градиенту.

Шаг 4. Сделать одну итерацию по схеме метода (22)–(24). Для $\tilde{u}_{k+1}$ проверить выполнение выбранного критерия останова.

Шаг 5. Если критерий останова выполнен, принять $\tilde{u}_{k(\sigma)} = \tilde{u}_{k+1}$, если нет, принять $k = k+1$ и переход к шагу 2.

На рисунках 1, 2, 3 представлены результаты решения модельной задачи при $N = 101$ и нулевом начальном приближении.

Основное отличие метода (22)–(24) для произвольного самосопряженного оператора от случая неотрицательного самосопряженного оператора в том, что в первом случае теоретически не гарантируется (хотя практически наблюдается) монотонность невязки. Сходимость метода сопряженных градиентов в данном случае происходит за счет ортогонализации векторов Крылова, а не за счет минимизации какого–либо функционала. Теорема 4 дает альтернативный критерий останова (θ–критерий).

На рис. 2, 3 представлены результаты решения модельной задачи при $\delta = 0.005$ и $\delta = 0.01$. В качестве критериев останова используются критерий невязки и θ–критерий. Оба критерия выполняются на одних и тех же итерациях.

Отметим, что при проведении расчетов ни разу не была достигнута ситуация, когда $(A_{\text{Д}}\tilde{p}_k, \tilde{p}_k) = 0$.

Авторы выражают благодарность В. А. Морозову, А. И. Гребенникову за обсуждение полученных результатов и ценные замечания.

Литература

[1] **Старостенко В. И.** Устойчивые численные методы в задачах гравиметрии. Киев, Наукова думка, 1978.

[2] **Алифанов О. М., Артюхин Е. А., Румянцев С. В.** Экстремальные методы решения некорректных задач. М.: Наука, 1988.

[3] **Немировский А. С.** О регуляризующих свойствах метода сопряженных градиентов на некорректных задачах// ЖВМ и МФ, 1986. Т.26. N3. С.332–347.

[4] **Гилязов С. Ф.** Приближенное решение некорректных задач. Теория и алгоритмы. М.: Изд–во Моск.ун–та, 1995.

[5] **Hanke M.** Conjugate gradient type methods for ill–posed problems. Longman Scientific & technical, 1995.

[6] **Воеводин В. В.** О методах сопряженных направлений // ЖВМ и МФ, 1979, Т.19. N5. С.1313–1317.

[7] **Воеводин В. В., Тыртышников Е. Е.** Об обобщении методов сопряженных направлений // Численные методы алгебры. М.: Изд–во Моск.ун–та, 1981. С.3-9.

[8] **Михлин С. Т.** Вариационные методы в математической физики. Изд. 2-е. М.: Наука, 1970.

[9] **Фаддеев Д. К., Фаддеева В. Н.** Вычислительные методы линейной алгебры. – М. – Л.: Физматгиз, 1963.

[10] **Васильев Ф. П.** Методы решения экстремальных задач. М.: Наука, 1981.

[11] **Тихонов А. Н., Арсенин В. Я.** Методы решения некорректных задач. М.: Наука, 1986.

[12] **Краснов М. Л., Киселев А. И., Макаренко Г. И.** Интегральные уравнения. М.: Наука, 1968.

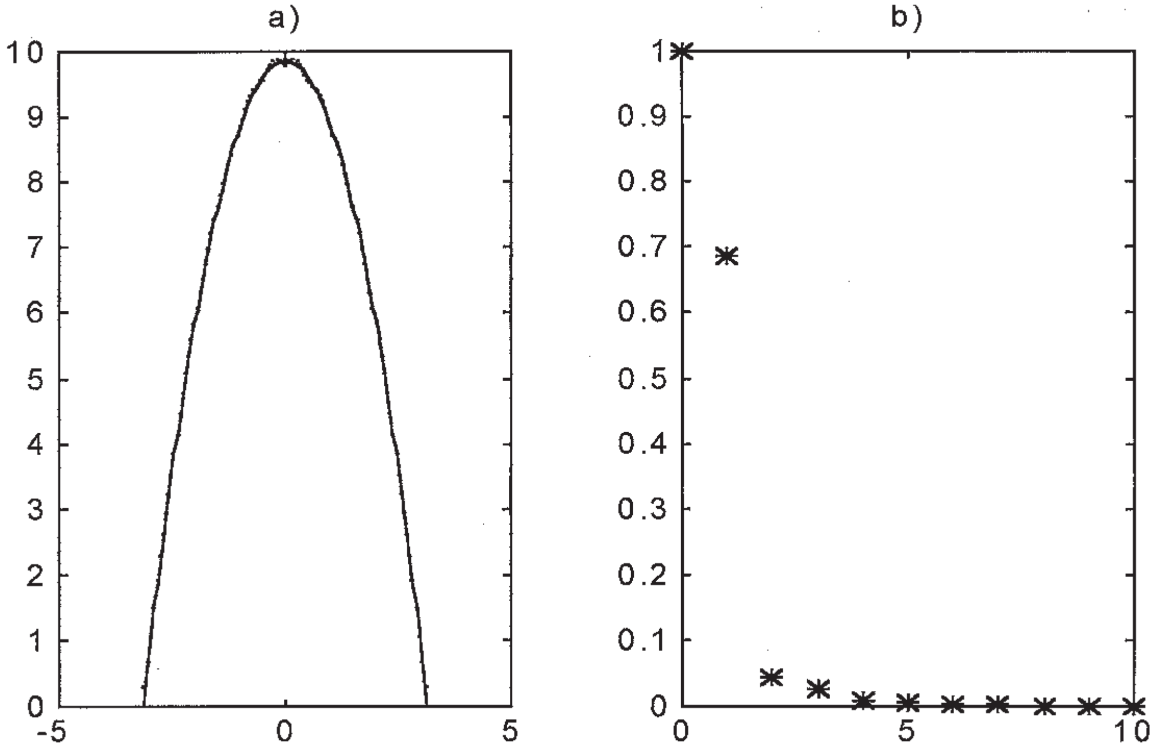

Рис. 1: Решение модельной задачи (a) при $\delta = 0$ и $k = 10$: (-) – точное решение, (.) – приближенное решение; график невязки (b): $\|g_k\|_2 \,/\, \|g_0\|_2$.

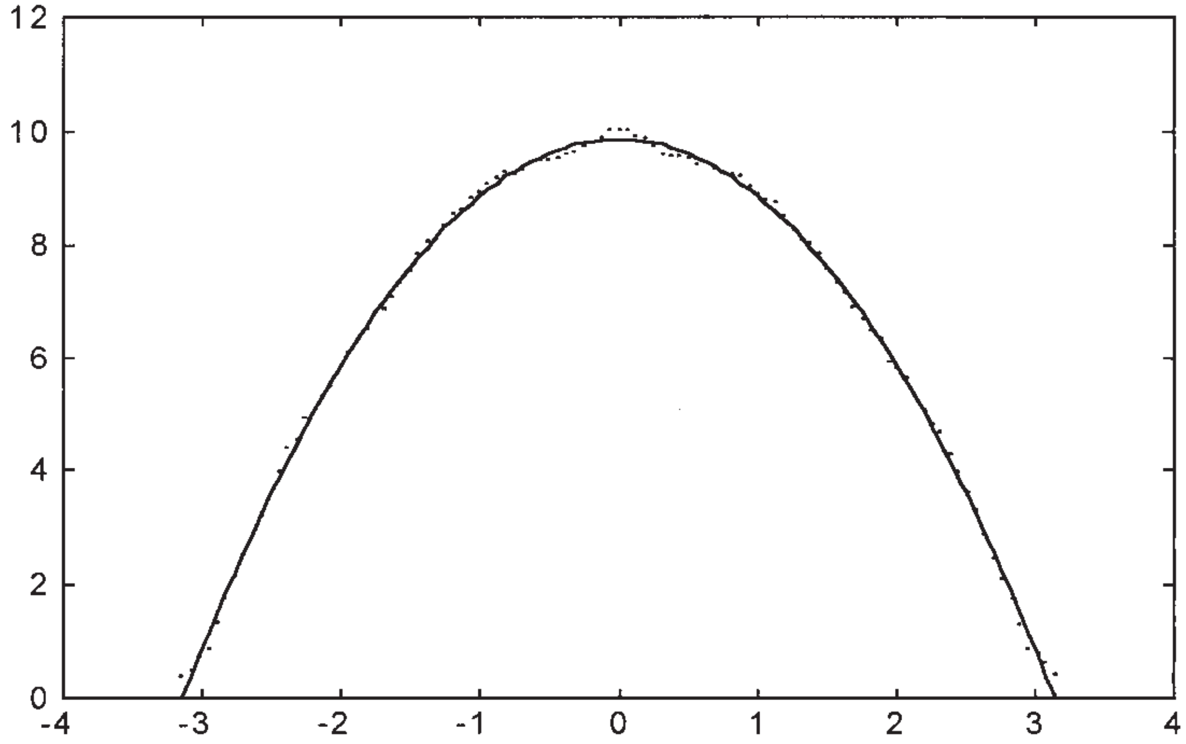

Рис. 2: Решение модельной задачи при $\delta = 0.005$, $k(\delta) = 5$.

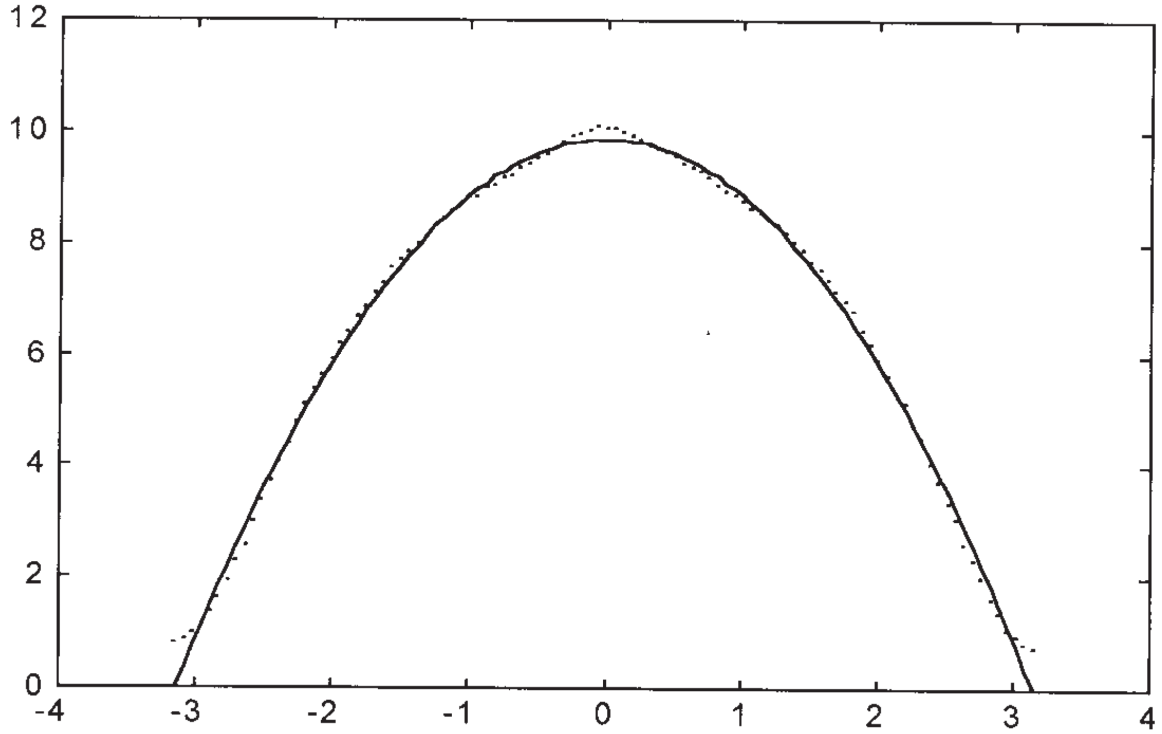

Рис. 3: Решение модельной задачи при $\delta = 0.01$, $k(\delta) = 3$.