ОБ ОДНОМ ПРЯМОМ МЕТОДЕ МИНИМИЗАЦИИ НЕКОТОРЫХ ФУНКЦИОНАЛОВ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ

И. М. Глазман, Ю. Ф. Сенчук

1. *Настоящая статья посвящена минимизации функционалов вида*

$$\Phi[y] = \int_a^b [(y^{(n)})^2 + f(x, y)]\, dx \tag{1}$$

при краевых условиях *

$$y^{(s)}(a) = y^{(s)}(b) = 0 \quad (s = 0, 1, \ldots, n-1) \tag{2}$$

с помощью процесса градиентной релаксации

$$y_{k+1} = y_k - \gamma_k \nabla \Phi [y_k]. \tag{3}$$

Алгоритм определения множителей γ_k, описываемый в настоящей статье, будем обозначать через $\mathfrak{R}$.

Первое сообщение об алгоритме $\mathfrak{R}$ при $n = 1$ и условиях его сходимости было дано без доказательства в краткой заметке [1], посвященной, в основном, задаче минимизации конечномерных функционалов. Настоящая статья, помимо доказательства сходимости алгоритма $\mathfrak{R}$ для квази-квадратичного функционала (1) при любом n, содержит также некоторые дополнительные результаты, относящиеся к быстроте сходимости и к двусторонним оценкам последовательных приближений на каждом шаге. Заметим, что в случае, когда $f(x, y)$ есть полином от x и y, реализация алгоритма $\mathfrak{R}$ требует лишь рациональных операций и приводит к полиномиальной минимизирующей последовательности, если только начальное приближение представляет собой полином.

Введем гильбертово пространство H всех функций из $W_2^{(n)}$, удовлетворяющих краевым условиям (2), со скалярным произведением

$$(y, z) = \int_a^b y^{(n)} z^{(n)}\, dx.$$

Очевидно,

$$\max_{a \leqslant x \leqslant b} |y^{(s)}(x)| \leqslant A \|y\|, \quad (s = 0, 1, \ldots, n-1), \tag{4}$$

где $\|y\| = \sqrt{(y, y)}$ и A не зависит от y и s.

В дальнейшем предполагается, что функция $f(x, y)$ задана в полосе $a \leqslant x \leqslant b$, $-\infty < y < \infty$, непрерывна по x, дважды непрерывно дифференцируема по y и удовлетворяет неравенству

$$f(x, y) \geqslant \beta > -\infty, \tag{5}$$

где β не зависит от x и y.

* Случай неоднородных краевых условий не вносит дополнительных трудностей.

2. Для вычисления градиента $\nabla\Phi[y]$ функционала (1) в H положим при данном $y \in H$ и произвольном $\eta \in H$

$$\varphi(\alpha) = \Phi[y + \alpha\eta], \tag{6}$$

откуда

$$\varphi'(0) = \int_a^b [2y^{(n)}\eta^{(n)} + f'_y(x, y)\eta]\,dx. \tag{7}$$

Интегрируя второе слагаемое в (7) n раз по частям и учитывая (2), получаем

$$\int_a^b f'_y(x, y)\eta\,dx = \int_a^b [(-1)^n \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f'_y(t, y(t))\,dt + P_{n-1}(x)]\,\eta^{(n)}\,dx,$$

где $P_{n-1}(x)$ — произвольный многочлен степени $n-1$.

Теперь соотношение (7) можно представить в виде

$$\varphi'(0) = \int_a^b \zeta^{(n)}(x)\,\eta^{(n)}(x)\,dx = (\zeta, \eta), \tag{8}$$

где

$$\zeta(x) = 2y + (-1)^n \int_a^x \frac{(x-t)^{2n-1}}{(2n-1)!} f'_y(t, y(t)\,dt + P_{2n-1}(x). \tag{9}$$

Многочлен $P_{2n-1}(x)$ степени $2n-1$ мы выберем так, чтобы функция $\zeta(x)$ удовлетворяла условиям (2).

Из формулы (8) непосредственно вытекает равенство

$$\nabla\Phi[y] = \zeta(x). \tag{10}$$

Так как далее

$$f(x, y+u) = f(x, y) + f'_y(x, y)\,u + \frac{1}{2} f''_{yy}(x, \tilde{y})\,u^2,$$

где $\tilde{y} = y + \theta(x)\,u$ и $0 < \theta(x) < 1$, то при любых $y \in H$, $u \in H$ для приращения функционала (1) получаем выражение

$$\Phi[y+u] - \Phi[y] = \int_a^b [2y^{(n)}u^{(n)} + (u^{(n)})^2 + f'_y(x, y)\,u + \frac{1}{2} f''_{yy}(x, \tilde{y})\,u^2]\,dx,$$

или в силу (7), (8), (10)

$$\Phi[y+u] - \Phi[y] = (\nabla\Phi[y], u) + \int_a^b [(u^{(n)})^2 + \frac{1}{2} f''_{yy}(x, \tilde{y})\,u^2]\,dx. \tag{11}$$

3. Из (1), (4) и (5) легко вытекают следующие два свойства.

1°. *Часть Ω пространства H, в которой значения $\Phi[y]$ не превосходят заданной величины, ограничена. Обратно, в каждой ограниченной части пространства H функционал $\Phi[y]$ ограничен.*

2°. *Во всем пространстве H функционал $\Phi[y]$ полуограничен снизу.*

Из 2° следует существование минимизирующей последовательности для функционала $\Phi[y]$. С другой стороны, в силу неравенств (4), часть Ω пространства H компактна в метрике $C_{n-1}(a, b)$. Поэтому из минимизирующей последовательности можно выделить подпоследовательность $\{y_k\}$, сходящуюся в метрике $C_{n-1}(a, b)$. Далее, так же, как и в [2], легко установить абсолютную непрерывность $(n-1)$-ой производной предельной функции $y(x)$, а также слабую сходимость $y_k^{(n)}$ к $y^{(n)}$ в $L_2(a, b)$, т. е. предельное соотношение

$$\lim_{k\to\infty}\int_a^b [y^{(n)}(x) - y_k^{(n)}(x)]\, g(x)\, dx = 0$$

при любой $g(x) \in L_2(a, b)$. Все дальнейшие рассмотрения [2], относящиеся к доказательству теоремы Тонелли, также сохраняют силу для нашего случая, так что имеет место следующее предложение.

3°. *Существует минимизирующая последовательность функций $y_k(x) \in H$, сходящаяся в метрике $C_{n-1}(a, b)$ к некоторой функции $\breve{y}(x) \in H$, на которой функционал $\Phi[y]$ достигает своего абсолютного минимума*.*

Далее, полагая в (11) $y = \breve{y}$, легко устанавливаем следующее свойство.

4°. *Точка $y = \breve{y}$ абсолютного минимума функционала $\Phi[y]$ является его стационарной точкой, т. е. $\nabla\Phi[\breve{y}] = 0$.*

Полагая в (9) $y = \breve{y}$, устанавливаем свойство.

5°. *Функция $\breve{y}(x)$ $2n$-кратно непрерывно дифференцируема и удовлетворяет уравнению*

$$(-1)^n y^{(2n)} + \frac{1}{2} f'_y(x, y) = 0. \tag{12}$$

Начиная со следующего $n°$, на протяжении всей статьи считается выполненным такое свойство.

6°. *Точка $\breve{y}$ абсолютного минимума функционала $\Phi[y]$ является его единственной стационарной точкой.*

В силу (11) это свойство имеет место, в частности, в том случае, когда всюду в полосе $a \leqslant x \leqslant b$, $-\infty < y < \infty$ выполняется условие

$$f''_{yy}(x, y) \geqslant 0. \tag{13}$$

Отметим, наконец, следующее свойство.

7°. *Если в точке $y = \breve{y}$ абсолютного минимума выполнено условие (13), то минимизирующая последовательность, существующая в силу 3°, сходится к $\breve{y}$ в метрике H.*

Доказательство непосредственно следует из соотношения (11).

4. Установим теперь некоторые свойства градиента $\nabla\Phi[y]$.

1°. *В каждой ограниченной части пространства H градиент $\nabla\Phi[y]$ есть ограниченный оператор относительно y.*

Доказательство легко может быть получено из выражения (9).

2°. *В части F пространства H, определяемой условием $F = F\{y : \Phi[\breve{y}] + \delta \leqslant \Phi[y] \leqslant N^2$, где $\delta > 0$ и $N < \infty$, норма градиента $\nabla\Phi[y]$ в H имеет положительную нижнюю грань.*

Для доказательства этого свойства предположим противное. Тогда для некоторой последовательности $y_k \in F$ будет $\|\nabla\Phi[y_k]\| \to 0$. В силу рассмотрений из [2], уже использованных в предыдущем $n°$, можно считать последовательность $\{y_k\}$ сходящейся в метрике $C_{n-1}(a, b)$ к некоторой функции $y^*(x)$ с абсолютно непрерывной производной $(n-1)$-го порядка. При этом последовательность $\{y_k^{(n)}\}$ в метрике $L_2(a, b)$ слабо сходится к $y^{*(n)}(x)$. Полагая в (9) $y = y_k$ и переходя к пределу в метрике $C(a, b)$, получаем

$$0 = 2y^* + (-1)^n \int_a^x \frac{(x-t)^{2n-1}}{(2n-1)!} f'_y(t, y^*(t))\, dt + P^*_{2n-1}(x). \tag{14}$$

*) Из одной теоремы М. А. Красносельского (см. [4], стр. 303) следует сходимость $y_k(x) \to \breve{y}(x)$ в метрике H.

Так как правая часть этого равенства есть $\nabla \Phi[y^*]$, то $\nabla \Phi[y^*] = 0$, и по условию 6° n°3 должно быть $y^* = \breve{y}$.

Теперь снова вернемся к предельному переходу от (9) при $y = y_k$ к (14), но на этот раз уже в метрике H. Так как при этом левая часть (9) и последние два слагаемых в правой части (9) сходятся в метрике H, то и первое слагаемое в правой части сходится в этой метрике, т. е. $\lim\limits_{k\to\infty} \| \breve{y} - y_k \| = 0$. Но тогда в силу очевидной непрерывности функционала $\Phi[y]$ должно быть $\Phi[y_k] \to \Phi[\breve{y}]$, что противоречит положительности числа δ.

5. Вернемся теперь к основной задаче статьи — алгоритмизации процесса минимизации функционала (1) с помощью градиентной релаксации, что сводится к определению последовательности множителей γ_k в (3). Очевидно, эта последовательность не должна стремиться к нулю слишком быстро, ибо в противном случае процесс (3) может привести к точке y, в которой еще $\Phi[y] > \Phi[\breve{y}]$. С другой стороны, если последовательность γ_k не стремится к нулю или стремится к нулю недостаточно быстро, то не исключено зацикливание, при котором y_k не имеют предела, а $\lim\limits_{k\to\infty} \Phi[y_k] > \Phi[\breve{y}]$. Для преодоления указанных опасностей последовательность γ_k должна не задаваться a priori, а вырабатываться в ходе процесса (3). Предлагаемый для этого алгоритм $\mathfrak{R}$ состоит в следующем (см. [1]).

Для управления процессом (3) произвольно задаем возрастающую последовательность положительных чисел $r_m \to \infty$. На каждом шаге процесса после конечного числа проб будет определяться очередной множитель γ_k, следующее приближение y_{k+1}, текущая длина релаксационного пути $l_k = \sum\limits_{j=0}^{k-1} \| y_{j+1} - y_j \|$, а также значение некоторой целочисленной функции $i(k)$, которая будет осуществлять обратную связь с управляющей последовательностью $\{r_m\}$.

Первый шаг алгоритма $\mathfrak{R}$ начинается с выбора произвольного нулевого приближения $y_0 \in H$ и произвольного пробного множителя $\gamma_0' > 0$, который испытывается на релаксационность с помощью неравенства

$$\Phi[y_0 - \gamma_0' \nabla \Phi[y_0]] < \Phi[y_0]. \tag{15}$$

Если (15) верно, то последовательно удваиваем γ_0' пока (15) не нарушится, и последнее в этом ряду значение пробного множителя принимаем за окончательное значение γ_0. Если же (15) неверно, то делим последовательно γ_0' пополам пока не будет выполнено (15), и последнее значение пробного множителя принимаем за γ_0.

Теперь, определив по формуле (3) y_1, вычисляем l_1 и полагаем $i(1) = 1$, на чем заканчивается построение первого шага алгоритма $\mathfrak{R}$.

Пусть выполнен k-ый шаг. Осуществление следующего шага начинаем с проверки соотношения

$$l_k < l_1 + r_{i(k)}. \tag{16}$$

Если (16) верно, то определяем начальное пробное значение γ_k' очередного релаксационного множителя по формуле $\gamma_k' = \gamma_{k-1}$, полагаем $i(k+1) = i(k)$ и проверяем множитель γ_k' на релаксационность с помощью неравенства

$$\Phi[y_k - \gamma_k' \nabla \Phi[y_k]] < \Phi[y_k]. \tag{17}$$

Если же (16) неверно, то определяем γ_k' по формуле $\gamma_k' = \frac{1}{2}\gamma_{k-1}$, полагаем $i(k+1) = i(k) + 1$ и проверяем γ_k' на релаксационность с помощью (17).

Если для γ'_k выполнено (17), то полагаем $\gamma'_k = \gamma_k$. Если же для γ'_k (17) неверно, то делим последовательно γ'_k пополам, пока не удовлетворим (17), и последнее значение получаемого таким путем пробного множителя принимаем за γ_k.

Теперь вычисляем по формуле (3) очередное приближение y_{k+1} и текущую длину релаксационного пути $l_{k+1} = l_k + \| y_{k+1} - y_k \|$, а значение $i(k+1)$ уже было вычислено выше.

Таким образом, $(k+1)$-й шаг выполнен и тем самым описание алгоритма $\mathfrak{R}$ завершено.

6. Для доказательства сходимости алгоритма $\mathfrak{R}$ предварительно установим два вспомогательных предложения (см. [1]).

Лемма 1. *В области* $\Omega\{y : \Phi[y] < N^2\}$ *множитель полной релаксации* $\alpha = \alpha[y]$ *ограничен снизу положительной константой.*

Доказательство. Введем в рассмотрение функцию $\psi(t) = \Phi[y - t\nabla\Phi[y]]$. Поскольку она совпадает с функцией (6) при $\alpha = -t$ и $\eta = \nabla\Phi[y]$, то, согласно (8) и (10), имеем

$$\psi'(0) = -\|\zeta\|^2. \tag{18}$$

С другой стороны, легко получить равенство

$$\psi''(t) = 2\|\zeta\|^2 + \int_a^b f''_{yy}(x,\ y - t\zeta)\,\zeta^2\,dx. \tag{19}$$

Будем считать, что $y \in \Omega$ и множитель t — релаксационный, так что и $z = y - t\zeta \in \Omega$. В то же время из неравенства $\Phi[y] < N^2$ в силу (1) и (5) следует, что

$$\int_a^b (y^{(n)})^2\,dx < N^2 - \beta(b-a),$$

а отсюда и из (4) при $s = 0$ вытекает неравенство $|z(x)| < B$, где константа B не зависит ни от x, ни от выбора $y \in \Omega$, ни от значения релаксационного множителя t. Поэтому при некотором $\mu < \infty$ будет $|f''_{yy}(x,\ y - t\zeta)| \leqslant \mu$ для всех $y \in \Omega$, $x \in [a, b]$ и релаксационных t, а тогда, учитывая неравенство (4) при $s = 0$, получаем

$$\left|\int_a^b f''_{yy}(x,\ y - t\zeta)\,\zeta^2\,dx\right| \leqslant \mu\int_a^b \zeta^2\,dx \leqslant \mu_1\|\zeta\|^2.$$

Из полученного неравенства и (19) следует соотношение

$$|\psi''(t)| \leqslant \mu_2\|\zeta\|^2. \tag{20}$$

Далее, из равенства

$$\psi'(t) = \int_0^t \psi''(t)\,dt + \psi'(0)$$

и (18) следует, что

$$\psi'(t) = \int_0^t \psi''(t)\,dt - \|\zeta\|^2,$$

откуда при $t = \alpha$ получаем

$$\int_0^\alpha \psi''(t)\,dt = \|\zeta\|^2. \tag{21}$$

Из (20) и (21) следует неравенство

$$\mu_2 \| \zeta \|^2 \int_0^\alpha dt \geqslant \| \zeta \|^2,$$

или

$$\alpha \geqslant \frac{1}{\mu_2},$$

что и требовалось установить.

Лемма 2. *Если в* (3) $\lim_{k\to\infty} \gamma_k = 0$, *но* $\lim_{k\to\infty} \Phi[y_k] > \Phi[\breve{y}]$, *то длина релаксационного пути конечна, т. е.*

$$\sum_{k=0}^{\infty} \| y_{k+1} - y_k \| < \infty. \tag{22}$$

Доказательство. Из (3) и свойства 1° n° 4 следует, что $\| y_{k+1} - y_k \| \to 0$. Далее, из формул (11) и (3) получаем соотношение

$$\Phi[y_k] - \Phi[y_{k+1}] = \| \nabla \Phi[y_k] \| \cdot \| y_k - y_{k+1} \| + O(\| y_k - y_{k+1} \|) = \\ = \| y_k - y_{k+1} \| \{ \| \nabla \Phi[y_k] \| + O(\| y_k - y_{k+1} \|) \}.$$

В силу свойства 2° n° 4 из последнего соотношения следует (22).

В заключении этого пункта отметим, что любой постоянный множитель релаксации $\gamma < \alpha[y]$, где $y \in \Omega_0 \{ y : \Phi[y], < \Phi[y_0] \}$, обеспечил бы сходимость процесса (3) к $\breve{y}$. Однако, как это видно из доказательства леммы 1, оценка $\alpha[y]$ снизу требует оценки сверху функции $|f''_{yy}(x, y)|$ в ограниченной области изменения переменных. Очевидно, невозможно построить универсальный алгоритм для решения этой последней задачи в классе всех функций $f(x, y)$, удовлетворяющих лишь условиям n° 1.

7. Перейдем к доказательству сходимости алгоритма $\mathfrak{R}$ (см. [1]).

Теорема 1. *При любом начальном приближении* $y_0 \in H$ *алгоритм* $\mathfrak{R}$ *обеспечивает сходимость в метрике* $C_{n-1}(a, b)$ *последовательности* y_k, *построенной по формуле* (3), *к функции* $\breve{y} \in H$, *дающей абсолютный минимум функционалу* $\Phi[y]$.

Доказательство. При реализации алгоритма $\mathfrak{R}$ последовательность $i(k)$ либо остается ограниченной, либо стремится к бесконечности.

В первом случае из ограниченности $i(k)$ следует конечность длины релаксационного пути, т. е. неравенство (22), откуда вытекает существование $\lim_{k\to\infty} y_k = z$ в H, а тем более и в $C_{n-1}(a, b)$. Далее, поскольку в силу леммы 1 любой множитель, удовлетворяющий неравенству $\gamma_k < \frac{1}{\mu_2}$, является релаксационным, то из ограниченности $i(k)$ вытекает также стационарность процесса (3). Переходя к пределу в (3), получаем $\nabla \Phi(z] = 0$, откуда по условию 6° n° 3 $z = \breve{y}$.

Во втором случае из неограниченности $i(k)$ следует, что $\gamma_k \to 0$. С другой стороны, из релаксационности процесса (3) вытекает существование предела $\lim_{k\to\infty} \Phi[y_k] = \nu$. При этом должно быть $\nu = \Phi[\breve{y}]$, ибо в противном случае по лемме 2 длина релаксационного пути была бы конечной, в то время как неограниченность $i(k)$ свидетельствует о бесконечности релаксационного пути. Таким образом, $\nu = \Phi[\breve{y}]$, т. е. $\lim_{k\to\infty} \Phi[y_k] = \Phi[\breve{y}]$, и остается установить предельное соотношение

$$\lim_{k\to\infty} y_k = \breve{y} \tag{23}$$

в метрике $C_{n-1}(a, b)$. Это соотношение вытекает из доказательства свойства 3° и условия 6° n° 3.

Приведенное выше доказательство сходимости алгоритма $\mathfrak{R}$ существенно опирается на лемму 2.

Из этой же леммы непосредственно вытекает сходимость к $\check{y}$ любой релаксационной последовательности вида (3) при единственном условии расходимости ряда $\sum_{k=0}^{\infty} \gamma_k$.

Однако уже в конечномерном случае легко обнаруживается медленная сходимость таких релаксационных последовательностей по сравнению со сходимостью алгоритма $\mathfrak{R}$.

Последний, как будет показано ниже в n° 12, обладает экспоненциальной скоростью сходимости, в то время как последовательность с расходящимся рядом $\sum_{k=0}^{\infty} \gamma_k$ обладает лишь степенным характером сходимости.

8. Переходя к исследованию скорости сходимости алгоритма $\mathfrak{R}$, будем предполагать выполненным условие (13) при $y = \check{y}(x)$ и считать функцию $f(x, y)$ трижды непрерывно дифференцируемой по y.

Тогда имеем

$$\Phi[y] = \int_a^b [(y^{(n)})^2 + f(x, \check{y}) + f'_y(x, \check{y})(y - \check{y}) + \frac{1}{2} f''_{yy}(x, \check{y})(y - \check{y})^2 + \\ + \frac{1}{6} f'''_{y^3}(x, \hat{y})(y - \check{y})^3]\, dx. \qquad (24)$$

Полагая здесь $y - \check{y} = z$, легко получим для квадратичной части функционала $\Phi[y]$ выражение

$$\int_a^b \left\{(z^{(n)})^2 + \frac{1}{2} f''_{yy}(x, \check{y}) z^2 + [f'_y(x, \check{y}) + 2(-1)^n \check{y}^{(2n)}]\, z + (\check{y}^{(n)})^2 + f(x, \check{y})\right\} dx. \quad (25)$$

Учитывая уравнение Эйлера — Лагранжа (12), заключаем, что (25) лишь на константу отличается от квадратичного функционала

$$F[z] = \int_a^b \left[(z^{(n)})^2 + \frac{1}{2} f''_{yy}(x, \check{y}) z^2\right] dx.$$

Поэтому, не ограничивая общности, будем считать, что функционал (24) имеет вид

$$\Phi[y] = \int_a^b [(y^{(n)})^2 + q(x) y^2]\, dx + \int_a^b r(x, \hat{y})\, y^3\, dx, \qquad (26)$$

где $q(x) = \frac{1}{2} f''_{yy}(x, \check{y})$ и $r(x, \hat{y}) = \frac{1}{6} f'''_{y^3}(x, \hat{y})$. Функционал (26) на функции $y(x) \equiv 0$ принимает свое минимальное значение, равное нулю. Из соотношения (26) и неравенства (4) при $s = 0$ вытекает равенство

$$\Phi[y] = F[y] + O(\|y\|^3). \qquad (27)$$

Представление (27) функционала $\Phi[y]$ в виде суммы квадратичного функционала и остаточного члена вместе с аналогичным представлением для градиента

$$\nabla \Phi[y] = \nabla F[y] + O(\|y\|^2) \qquad (28)$$

является исходным пунктом для исследования скорости сходимости алгоритма $\mathfrak{R}$. Соотношение (28) легко выводится из (9) и формулы Тейлора

$$f'_y(x, y) = f'_y(x, 0) + f''_{yy}(x, 0)\, y + \frac{1}{2} f'''_{y^3}(x, y^*)\, y^2.$$

Для дальнейшего существенно, что функционал $\Phi[y]$ не вырождается в точке минимума, т. е. первое слагаемое в (28) удовлетворяет неравенству

$$\|\nabla F[y]\| \geqslant r\|y\| \tag{29}$$

при некотором $r > 0$. Справедливость этого неравенства (при $r = 2$) легко вытекает из вида градиента $\nabla F[y]$. Действительно, заменяя в (9) Φ на F и дифференцируя обе части n раз, получаем

$$\zeta^{(n)} = 2y^{(n)} + 2(-1)^n \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} q(t)\, y(t)\, dt + P_{n-1}(x),$$

или

$$\frac{1}{2}\zeta^{(n)}y^{(n)} = (y^{(n)})^2 + (-1)^n y^{(n)} \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} q(t)\, y(t)\, dt + \frac{1}{2} y^{(n)} P_{n-1}(x).$$

Отсюда с помощью интегрирований по частям при краевых условиях (2) получаем соотношение

$$\frac{1}{2}\int_a^b \zeta^{(n)}y^{(n)}\, dx = \int_a^b [(y^{(n)})^2 + qy^2]\, dx,$$

откуда

$$\|y\|^2 \leqslant \frac{1}{2}(\zeta, y) \leqslant \frac{1}{2}\|\zeta\| \cdot \|y\|,$$

т. е. $\|y\| \leqslant \frac{1}{2}\|\zeta\|$, что совпадает с (29) при $r = 2$.

9. Рассмотрим некоторые свойства процесса (3) для $\Phi[y]$ по отношению к $F[y]$. При этом мы не предполагаем, что процесс (3) построен с помощью алгоритма $\mathfrak{R}$.

Лемма 1. Если релаксационный процесс (3) *сходится в H к точке минимума функционала $\Phi[y]$, то этот процесс является асимптотически градиентным для функционала $F[y]$.*

Доказательство. Положим $\nabla\Phi[y_k] - \nabla F[y_k] = w_k$. Тогда для «косинуса угла» между градиентами $\nabla\Phi[y_k]$ и $\nabla F[y_k]$ в H на основании (28) и (29) получаем при $k \to \infty$

$$\frac{(\nabla\Phi[y_k],\ \nabla F[y_k])}{\|\nabla\Phi[y_k]\| \cdot \|\nabla F[y_k]\|} = \frac{(\nabla F[y_k] + w_k,\ \nabla F[y_k])}{\|\nabla F[y_k] + w_k\| \cdot \|\nabla F[y_k]\|} \to 1.$$

Лемма 2. Если релаксационный процесс (3) *сходится в H к точке минимума функционала $\Phi[y]$ и при этом $\gamma_k \to 0$, то, начиная с некоторого места, процесс* (3) *будет релаксационным и для функционала $F[y]$.*

Доказательство. Заменив в (11) Φ на F и положив $y = y_{k+1}$, $y + u = y_k$, получаем

$$\begin{aligned} F[y_k] - F[y_{k+1}] &= \gamma_k(\nabla F[y_k],\ \nabla\Phi[y_k]) + \gamma_k^2 O(\|\nabla\Phi[y_k]\|^2) = \\ &= \gamma_k\|\nabla F[y_k]\|^2 + \gamma_k(\nabla F[y_k],\ \nabla\Phi[y_k] - \nabla F[y_k]) + \\ &\quad + \gamma_k^2 O(\|\nabla\Phi[y_k]\|^2). \end{aligned} \tag{30}$$

С другой стороны, из (9) непосредственно вытекает, что $\|\nabla\Phi[y_k]\| = O(\|y_k\|)$ и $\|\nabla F[y_k]\| = O(\|y_k\|)$. Поэтому в силу (28) равенство (30) можно представить в виде

$$F[y_k] - F[y_{k+1}] = \gamma_k\|\nabla F[y_k]\|^2 + O(\|y_k\|^3) + \gamma_k O(\|y_k\|^2).$$

Отсюда и из (29) следует требуемое неравенство $F[y_{k+1}] \leqslant F[y_k]$.

10. Для дальнейшего удобно ввести в рассмотрение линейный оператор T, порождаемый квадратичным функционалом $F[y]$. На плотном в H линейном многообразии достаточно гладких функций $y(x)$ мы определим его формулой $Ty = B^{-1}Ay$, где

$$Ay = (-1)^n y^{(2n)} + q(x) y$$

и B^{-1} есть обращение оператора $(-1)^n y^{(2n)}$ при краевых условиях (2). Очевидно, $(Ty, z) = (y, Tz)$ и

$$(y, y) \leqslant (Ty, y) \leqslant M(y, y). \tag{31}$$

В частности, из последнего неравенства вытекает ограниченность оператора T, и поэтому можно продолжить его по непрерывности на все H. Через T мы впредь обозначаем это продолжение. Очевидно, $F[y] = (Ty, y)$ и

$$\nabla F[y] = 2Ty. \tag{32}$$

Далее, имеем

$$(T - I) y = B^{-1}Ay - y = B^{-1}(By + qy) - y = B^{-1}(qy)$$

и, следовательно,

$$(T - I) y = \int_a^b K(x, s) q(s) y(s) ds, \tag{33}$$

где $K(x, s)$ — функция Грина оператора B. Отсюда, дифференцируя обе части (33) n раз по x, легко устанавливаем полную непрерывность оператора $T - I$ в H. Привлекая еще соотношение (31), заключаем, что спектр самосопряженного оператора T исчерпывается изолированными собственными значениями конечной кратности, расположенными правее единицы, с единственной предельной точкой $\lambda = 1$.

11. Если процесс (3) удовлетворяет условиям лемм 1 и 2 $n°$ 9, то он по отношению к квадратичному функционалу $F[y]$ является сходящимся релаксационным и асимптотически градиентным. Мы утверждаем, что для такого процесса релаксационный путь конечен. Это утверждение было усилено и доказано Ю. И. Любичем. В настоящем $n°$ мы воспроизводим его формулировку и доказательство, сообщенные нам устно. При этом H означает абстрактное гильбертово пространство, а T — произвольный ограниченный и положительно определенный самосопряженный оператор в H. Любой релаксационный процесс для функционала $F[y] = (Ty, y)$ можно (см. [3]) представить в виде

$$y_{k+1} = y_k - \beta_k z_k, \tag{34}$$

где

$$\beta_k = \| y_{k+1} - y_k \| > 0, \quad \| z_k \| = 1.$$

Так как

$$F[y_{k+1}] - F[y_k] = -2\beta_k (Ty_k, z_k) + \beta_k^2 (Tz_k, z_k), \tag{35}$$

то релаксационность процесса (34) равносильна представимости β_k в виде

$$\beta_k = \gamma_k \frac{(Ty_k, z_k)}{(Tz_k, z_k)}, \tag{36}$$

где $0 \leqslant \gamma_k \leqslant 2$. Коэффициент γ_k в (36) называется множителем релаксации в направлении z_k, а угол $\theta_k \left(0 \leqslant \theta_k \leqslant \frac{\pi}{2}\right)$, определяемый равенством

$$(Ty_k, z_k) = \| Ty_k \| \cos \theta_k, \tag{37}$$

углом релаксации.

Лемма. *Если* $\sup\gamma_k < 2$ *и* $\inf\cos\theta_k > 0$, *то*

$$\sum_{k=0}^{\infty} \| y_{k+1} - y_k \| < \infty. \tag{38}$$

Доказательство. Так как в силу (34), (36), (37)

$$y_{k+1} - y_k = -\frac{\gamma_k \| Ty_k \| \cos\theta_k}{(Tz_k, z_k)} z_k,$$

то достаточно доказать, что

$$\sum_{k=0}^{\infty} \gamma_k \| Ty_k \| < \infty. \tag{39}$$

Положив $\Delta_k = (Ty_k, y_k)$, получаем из (35), (36), (37)

$$\Delta_k - \Delta_{k+1} = \gamma_k (2 - \gamma_k) \cos^2\theta_k \Delta_k h_k, \tag{40}$$

где

$$h_k = \frac{\| Ty_k \|^2}{(Tz_k, (z_k)(Ty_k, y_k)}.$$

Полагая далее $\delta = \inf\{(2-\gamma_k)\cos^2\theta_k h_k\}$ и учитывая, что, в силу условий леммы, $\delta > 0$, получаем из (40) соотношение

$$\Delta_{k+1} = [1 - \gamma_k (2 - \gamma_k)\cos^2\theta_k h_k] \Delta_k \leqslant (1 - \delta\gamma_k)\Delta_k \leqslant e^{-\delta\gamma_k} \Delta_k.$$

Отсюда путем перемножения по k выводим неравенство

$$\Delta_m \leqslant \Delta_0 e^{-\delta \sum\limits_{k=0}^{m-1} \gamma_k},$$

а так как $\Delta_m = (Ty_m, y_m) \geqslant \frac{1}{\| T \|} \| Ty_m \|^2$, то

$$\| Ty_m \| \leqslant \sqrt{\| T \| \Delta_0} e^{-\frac{\delta}{2} \sum\limits_{k=0}^{m-1} \gamma_k}.$$

Теперь вместо (39) достаточно доказать, что

$$\sum_{m=0}^{\infty} \gamma_m e^{-\frac{\delta}{2} \sum\limits_{k=0}^{m-1} \gamma_k} < \infty,$$

или

$$\sum_{m=0}^{\infty} (S_{m+1} - S_m) e^{-S_m} < \infty, \tag{41}$$

где

$$S_m = \frac{\delta}{2} \sum_{k=0}^{m-1} \gamma_k.$$

Соотношение (41) проверяется элементарно. А именно, поскольку $0 \leqslant S_{m+1} - S_m = \frac{\delta}{2}\gamma_m < \delta$, то можно при некотором Q считать

$$S_{m+1} - S_m < Q(1 - e^{-(S_{m+1} - S_m)}),$$

а тогда

$$\sum_{m=0}^{\infty} (S_{m+1} - S_m) e^{-S_m} < Q \sum_{m=}^{\infty} (e^{-S_m} - e^{-S_{m+1}}) = Q(e^{-S_0} - e^{-S_\infty}) < \infty.$$

12. Теперь можно приступать к исследованию скорости сходимости алгоритма $\mathfrak{R}$. В основе этого исследования лежит следующее предложение.

Лемма. *При реализации алгоритма $\mathfrak{R}$ релаксационный множитель γ_k в конце концов стабилизируется (т. е., начиная с некоторого шага, перестает зависеть от k), а значит (по построению алгоритма $\mathfrak{R}$) релаксационный путь конечен.*

Доказательство. Предположение противного влечет, по построению алгоритма $\mathfrak{R}$, соотношение $\gamma_k \to 0$. Но тогда в силу лемм n° 9 процесс (3) является релаксационным и притом асимптотически градиентным для квадратичного функционала $F[y]$. Из свойства асимптотической градиентности и условия $\gamma_k \to 0$ заключаем, на основании леммы n° 11, о конечности релаксационного пути. Однако это заключение противоречит построению алгоритма $\mathfrak{R}$, в силу которого при $\gamma_k \to 0$ релаксационный путь должен быть бесконечным. Лемма доказана.

Стабилизация релаксационного множителя γ_k делает естественным предположение о том, что алгоритм $\mathfrak{R}$ обеспечивает сходимость со скоростью геометрической прогрессии. Если бы установившееся значение $\gamma > 0$ множителя γ_k удовлетворяло неравенству $\gamma < \frac{1}{\| T \|}$, то это предположение можно было бы установить без труда. Однако в действительности алгоритм $\mathfrak{R}$ не обеспечивает стабилизации γ_k на значении γ, меньшем $\frac{1}{\| T \|}$. Это обстоятельство существенно усложняет ситуацию. Основные трудности доказательства следующей теоремы были преодолены Ю. И. Любичем. При доказательстве этой теоремы мы следуем в основном устному сообщению Ю. И. Любича.

Теорема 2. *Скорость сходимости в H последовательных приближений* (3), *построенных по алгоритму $\mathfrak{R}$, к точке минимума $\check{y}$ не менее скорости сходимости некоторой геометрической прогрессии.*

Доказательство. Напомним, что после выкладок n° 8 мы, не ограничивая общности, считаем $\check{y} \equiv 0$.

Положим для краткости

$$y - \gamma \nabla \Phi[y] = \bar{y},$$

где γ есть установившееся при реализации алгоритма $\mathfrak{R}$ постоянное значение релаксационного множителя. В силу (28) и (32) имеем

$$\bar{y} = y - \gamma [2Ty + O(\| y \|^2)] = y - 2\gamma Ty + O(\| y \|^2),$$

или, поскольку, очевидно, $\Phi[y] \geqslant \| y \|^2$,

$$\bar{y} = (I - 2\gamma T) y + O(\Phi[y]). \tag{42}$$

Отсюда в силу (27) получаем

$$\Phi[\bar{y}] = (T\bar{y}, \bar{y}) + O(\| y \|^3) = (T[I - 2\gamma T] y, [I - 2\gamma T] y) + O(\| y \|^3) = (T[I - 2\gamma T]^2 y, y) + O(\| y \|^3),$$

а значит

$$\Phi[y] - \Phi[\bar{y}] = (Ty, y) - (T[I - 2\gamma T]^2 y, y) + O(\| y \|^3) = (Sy, y) + O(\| y \|^3), \tag{43}$$

где $S = 4\gamma T^2 (I - \gamma T)$.

Обозначим через λ_i собственные значения оператора T, а через μ_i — соответствующие собственные значения оператора S, так что

$$\mu_i = 4\gamma\lambda_i^2 (1 - \gamma\lambda_i). \tag{44}$$

Заметим, что среди чисел μ_j могут быть нулевые и даже отрицательные, поскольку, как уже отмечалось выше, не обязательно $\gamma \leqslant \frac{1}{\|T\|}$.

Из (42) имеем

$$y_{k+1} = (I - 2\gamma T)\, y_k + v_k, \tag{45}$$

где $v_k = O(\Phi[y_k])$. При этом легко видеть, что оценка для v_k равномерна по k, т. е. существует константа σ такая, что при всех k

$$\|v_k\| < \sigma\Phi[y_k]. \tag{46}$$

Далее, согласно (45),

$$y_k = (I - 2\gamma T)^{-1} y_{k+1} - (I - 2\gamma T)^{-1} v_k,$$

откуда получаем при любом $m > k$

$$y_k = (I - 2\gamma T)^{-(m-k)} y_m - (I - 2\gamma T)^{-(m-k)} v_{m-1} - \cdots - (I - 2\gamma T)^{-1} v_k. \tag{47}$$

Обозначим через G спектральное подпространство, соответствующее неположительным собственным значениям μ_j, а через φ — проекцию y на G. Тогда, вводя оператор проектирования P на G, получаем из (47)

$$\varphi_k = P(I - 2\gamma T)^{-(m-k)} - P(I - 2\gamma T)^{-(m-k)} v_{m-1} - \cdots - P(I - 2\gamma T)^{-1} v_k.$$

Так как при $\mu_j \leqslant 0$ в силу (44) должно быть $\lambda_j \leqslant \frac{1}{\gamma}$, а значит $1 - 2\gamma\lambda_j \leqslant -1$, то норма оператора $P(I - 2\gamma T)^{-1}$ не превосходит единицы. Поэтому, учитывая (46) и установленную в $n^\circ 7$ сходимость $y_m \to 0$, получаем

$$\|\varphi_k\| \leqslant \sum_{j=k}^{\infty} \|v_j\| \leqslant \sigma \sum_{j=k}^{\infty} \Phi[y_j] = O\Big(\sum_{j=k}^{\infty} \Phi[y_j]\Big). \tag{48}$$

С другой стороны, в силу (3) и установленной в лемме этого n° конечности релаксационного пути сходится ряд $\sum_{k=0}^{\infty} \|\nabla\Phi[y_k]\|$, а, следовательно, в силу (27) и ряд $\sum_{k=0}^{\infty} \|y_k\|$. Но тогда, очевидно, сходится также ряд $\sum_{k=0}^{\infty} \sqrt{\Phi[y_k]}$, т. е.

$$\lim_{k\to\infty} \sum_{j=k}^{\infty} \sqrt{\Phi[y_j]} = 0. \tag{49}$$

В то же время

$$\sum_{j=k}^{\infty} \Phi[y_j] = \sum_{j=k}^{\infty} \sqrt{\Phi[y_j]}\sqrt{\Phi[y_j]} \leqslant \sqrt{\Phi[y_k]} \sum_{j=k}^{\infty} \sqrt{\Phi[y_j]},$$

откуда на основании (49) получаем соотношение

$$\sum_{j=k}^{\infty} \Phi[y_j] = o\big(\sqrt{\Phi[y_k]}\big) = o(\|y_k\|).$$

Отсюда и из (48) получаем

$$\|\varphi_k\| = o(\|y_k\|). \tag{50}$$

Это последнее соотношение показывает, что проекция последовательных приближений (3) на неблагоприятное спектральное подпространство G (в котором стационарный градиентный релаксационный процесс для квадратичной части $F[y]$ функционала $\Phi[y]$ с данным множителем был бы расходящимся) составляет несущественную часть этих последовательных приближений. Теперь уже нетрудно завершить доказательство.

Положим $y_k = \psi_k + \varphi_k$, где $\varphi_k \in G$, $\psi_k \perp G$. Тогда в силу (50)

$$(Sy_j, y_j) = (S\psi_j, \psi_j) + (S\varphi_j, \varphi_j) \geqslant m_1 \| \psi_j \|^2 + O(\| \varphi_j \|^2) > m_2 \| y_j \|^2, \quad (51)$$

где $m_1 > 0$ есть нижняя грань оператора S в $H \ominus G$, а m_2 — некоторое положительное число. Так как согласно (43)

$$\Phi[y_k] - \Phi[y_{k+1}] = (Sy_k, y_k) + O(\| y_k \|^3),$$

то

$$\Phi[y_k] = \sum_{j=k}^{\infty} [(Sy_j, y_j) + O(\| y_j \|^3)],$$

откуда на основании (51) при некотором $m_3 > 0$ будет

$$\Phi[y_k] > m_3 \sum_{j=k}^{\infty} \| y_j \|^2,$$

т. е. при некотором $M_1 < \infty$

$$\sum_{j=k}^{\infty} \| y_j \|^2 < M_1 \| y_k \|^2.$$

Обозначая левую часть последнего неравенства через R_k, имеем

$$R_k < M_1 (R_k - R_{k+1}),$$

или

$$R_{k+1} < qR_k, \quad (0 < q < 1),$$

а значит при некотором c

$$R_k < cq^k.$$

Поэтому

$$\| y_k \|^2 = R_k - R_{k+1} < R_k + R_{k+1} < c(1+q)q^k,$$

и теорема доказана.

13. Укажем теперь способ оценки данного приближения независимо от алгоритма его построения. В основе этого способа лежит свойство выпуклости функционала $\Phi[y]$. Это свойство вытекает из условия (13), которое в этом $n°$ предполагается выполненным во всей полосе $a \leqslant x \leqslant b$, $-\infty < y < \infty$. Ниже мы уже не считаем $\breve{y} \equiv 0$.

Подставляя в (1) какую-нибудь функцию $y_0(x)$, получаем

$$\Phi[\breve{y}] = \int_a^b [(\breve{y}^{(n)})^2 + f(x, \breve{y})]\, dx \leqslant \Phi[y_0],$$

или согласно (5)

$$\| \breve{y} \|^2 \leqslant \Phi[y_0] - \beta(b-a).$$

Таким образом, искомая точка $\breve{y}$ принадлежит сфере σ_N пространства H с центром в точке $y \equiv 0$ радиуса

$$N = \sqrt{\Phi[y_0] - \beta(b-a)}.$$

Из (11) легко следует, что при любом $y \in H$

$$\Phi[y] \geqslant L[y], \quad (52)$$

где $L[y] = \Phi[y_k] + (\nabla\Phi[y_k], y - y_k)$.

Для минимума неоднородного линейного функционала $L[y]$ в сфере σ_N имеем

$$\min_{\| y \| \leqslant N} L[y] = \Phi[y_k] - (\nabla\Phi[y_k], y_k) - N \| \nabla \Phi[y_k] \|.$$

Так как в силу (52) $\Phi[\check{y}] \geqslant L[\check{y}] \geqslant \min\limits_{\|y\| < N} L[y]$, то получаем следующую двустороннюю оценку:

$$\Phi[y_k] - N\,\|\nabla\Phi[y_k]\| - (\nabla\Phi[y_k],\ y_k) \leqslant \Phi[\check{y}] \leqslant \Phi[y_k]. \tag{53}$$

Если $y_k \to \check{y}$ в метрике H, то левая и правая части неравенства (53) стремятся к средней его части $\Phi[\check{y}]$. Далее, из (11) и (13) имеем

$$\|y_k - \check{y}\|^2 \leqslant \Phi[y_k] - \Phi[\check{y}].$$

Отсюда на основании (53) имеем для погрешности $\|y_k - \check{y}\|$ оценку

$$\|y_k - \check{y}\| \leqslant \sqrt{N\,\|\nabla\Phi[y_k]\| + (\nabla\Phi[y_k],\ y_k)}. \tag{54}$$

Заметим, что оценки (53) и (54) можно несколько улучшить, беря на каждом шаге вместо N число $N_k = \sqrt{\Phi[y_k] - \beta(b-a)}$.

ЛИТЕРАТУРА

1. И. М. Глазман. О градиентной релаксации для неквадратичных функционалов. ДАН СССР, **154**, №5, 1011—1014, 1964.

2. Н. И. Ахиезер. Лекции по вариационному исчислению. Гостехиздат, М., 1955.

3. Ю. И. Любич. Общие теоремы о квадратичной релаксации. ДАН СССР, **161**, № 6, 1274—1277, 1965.

4. М. А. Красносельский. Топологические методы в теории нелинейных интегральных уравнений, Гостехиздат, М., 1956.