91-559

сообщения
объединенного
института
ядерных
исследований
дубна

P5-91-559

Т.Жанлав, И.В.Пузынин

СХОДИМОСТЬ ИТЕРАЦИОННОЙ НЬЮТОНОВСКОЙ СХЕМЫ

1991

Введение

Вопрос о построении эффективных алгоритмов решения нелинейных задач остается актуальным в вычислительной математике. Одним из методов решения таких задач является метод Ньютона. Главное его достоинство – квадратичная сходимость в близкой окрестности искомого решения. Однако при неудачно выбранном начальном приближении сходимость может быть медленной или отсутствовать. В настоящее время в расчетах задач теоретической физики успешно применяется непрерывный аналог метода Ньютона (НАМН), предложенный в /I/ и развитый в /2,3/.

В § I данной работы приведено обоснование алгоритма выбора итерационного параметра, предложенного в /4/, изучено свойство коэффициента перехода от итерации к итерации. Сходимость НАМН доказана в § 2.

§ I. Непрерывный аналог метода Ньютона и итерационная схема

Пусть φ –непрерывная нелинейная функция, переводящая банахово пространство X в банахово пространство Y. Рассмотрим уравнение

$$\varphi(z) = 0. \tag{I.I}$$

Пусть в некоторой окрестности решения z^* уравнения (I.I) функция φ дважды непрерывно дифференцируема, а оператор $(\varphi')^{-1}$ непрерывно обратим.

Как известно /I/, НАМН для уравнения (I.I) представляет эволюционное по дополнительному непрерывному параметру t $(0 \le t < \infty)$ уравнение

$$\frac{d}{dt}\varphi(z(t)) = -\varphi(z(t)), \qquad z(0) = z_0 . \tag{I.2}$$

При указанных выше ограничениях относительно функции φ и ее производных имеет место соотношение /I,2/

$$\lim_{t\to\infty} z(t) = z^*.$$

Дискретизация по параметру t уравнения (I.2) осуществляется методом Эйлера, который приводит к итерационному процессу:

$$\varphi'(z_n)v_n = -\varphi(z_n), \tag{I.3}$$

$$z_{n+1} = z_n + \tau_n v_n, \quad n=0.1\ \dots, \quad 0<\tau_n \le 1 .$$

При $\tau_n \equiv 1$ этот итерационный процесс совпадает с классическим методом Ньютона.

В работе[2] доказана сходимость итераций (I.3) к решению эволюционной задачи (I.2) на конечном отрезке $0 \le t \le T$ при $\tau_n \to 0$, если в окрестности решения z^* функция $\varphi(z)$ удовлетворяет некоторым дополнительным условиям.

В практических расчетах итерационный процесс (I.3) дополняется алгоритмами задания шага τ_n, от удачного выбора которого может зависеть как сходимость метода, так и скорость сходимости. В настоящее время существует несколько алгоритмов задания шага τ_n [4,5]. Мы рассмотрим итерационный процесс (I.3), в котором шаг τ_n задается по формуле[4]

$$\tau_n = \|\varphi(z_n)\|^{-1} \cdot \|\varphi(z_{n-1})\| \cdot \tau_{n-1}, \qquad 0<\tau_0 \le \tau_n \le 1. \tag{I.4}$$

Пусть в некоторой сфере D с центром в точке z_0 существуют производные $\varphi'(z)$, $\varphi''(z)$, причем линейный оператор $\varphi'(z)$ обратимый и имеют место неравенства

$$\|\varphi'(z)^{-1}\| \le B, \qquad \|\varphi''(z)\| \le M. \tag{I.5}$$

Предположим, что все приближения z_n, полученные с помощью итераций (I.3), принадлежат сфере D. Тогда легко показать[6], что

$$\|\varphi(z_{n+1})\| \le \psi_{n+1}(\tau_n)\ \|\varphi(z_n)\|, \quad n=0,1,\dots, \tag{I.6}$$

где коэффициент перехода $\psi_{n+1}(\tau)$ от итерации к итерации определяется по формуле

$$\psi_{n+1}(\tau) = 1-\tau + \frac{\tau^2}{2} MB^2\ \|\varphi(z_n)\|, \quad 0<\tau \le 1. \tag{I.7}$$

Выберем τ_n так, чтобы выполнялось условие

$$\psi_{n+1}(\tau_n) = \min_{\tau\in(0,1]} \psi_{n+1}(\tau).$$

Такой шаг τ_n мы назовем оптимальным и итерационный процесс (I.3) с оптимальным шагом τ_n назовем оптимальным по скорости сходимости.

Точка минимума функции (I.7) есть

$$\tau_n^* = \frac{1}{MB^2\|\varphi(z_n)\|} = \frac{1}{\varepsilon_n}. \tag{I.8}$$

Пусть $\varepsilon_n \le 1$. Тогда $\tau_n^* \ge 1$, т.е. в этом случае $\tau_n = 1$ и поэтому $\psi_{n+1}(\tau_n) = \psi_{n+1}(1) \le 1/2$. Если же $\varepsilon_n > 1$, то $\tau_n^* < 1$, и, следовательно, $\psi^*_{n+1}(\tau_n^*) = 1 - 1/2\,\varepsilon_n > 1/2$. Однако, если в оценках (I.5) конкретные значения констант B и M неизвестны, то в этом случае мы не можем найти точку минимума τ_n^*. Вместо оптимального шага τ_n необходимо найти другой шаг, позволяющий уменьшать невязки от итерации к итерации. Заметим, что для функции (I.7) выполняется условие

$$\psi'_{n+1}(0) = \psi'_n(0) = -1. \tag{I.9}$$

Мы потребуем, чтобы дискретные аналоги производных $\psi'_{n+1}(0)$ и $\psi'_n(0)$ тоже удовлетворяли соотношению (I.9), то есть

$$\frac{\psi_{n+1}(\tau_n) - \psi_{n+1}(0)}{\tau_n} = \frac{\psi_n(\tau_{n-1}) - \psi_n(0)}{\tau_{n-1}}. \tag{I.10}$$

Подставляя выражение (I.7) в соотношение (I.10), получаем

$$\frac{\tau_n}{\tau_{n-1}} = \frac{\|\varphi(z_{n-1})\|}{\|\varphi(z_n)\|}, \quad n=1,2,\dots, \tag{I.11}$$

откуда следует формула (I.4). Экстремальные точки τ_n^* и τ_{n-1}^*, как видно из формулы (I.8), удовлетворяют соотношению (I.11), хотя каждая из них неизвестна в отдельности.

Выбор τ_n по формуле (I.11) позволяет уменьшать $\psi_{n+1}(\tau)$ при возрастании n. Используя соотношение (I.11), запишем $\psi_{n+1}(\tau_n)$ в виде

$$\psi_{n+1}(\tau_n) = 1-\tau_n(1-q_0), \quad n=0,1,\dots, \qquad q_0 = \frac{\tau_0}{2} MB^2\ \|\varphi(z_0)\|. \tag{I.12}$$

Отсюда следует, что

$$\psi_{n+1}(\tau_n) - \psi_n(\tau_{n-1}) = (1-q_0)(\tau_{n-1} - \tau_n), \quad n=1,2,\dots \tag{I.13}$$

Лемма. Пусть выбор шага τ_n осуществляется по формуле (I.4) и выполняется условие

$$q_o = \frac{\tau_o}{2} MB^2 \; \|\varphi(z_o)\| < 1 . \qquad (I.I4)$$

Тогда справедливо неравенство

$$\psi_{n+1}(\tau_n) < \psi_n(\tau_{n-1}) , \; n=1,2, \dots , \qquad (I.I5)$$

которое является необходимым и достаточным условием выполнения неравенства

$$\|\varphi(z_n)\| < \|\varphi(z_{n-1})\| , \; n=1,2,\dots . \qquad (I.I6)$$

Доказательство. Сначала докажем вторую часть леммы. Пусть выполнено неравенство (I.I6). Тогда из соотношения (I.II) следует, что $\tau_n > \tau_{n-1}$. Следовательно, в силу условия (I.I4), правая часть равенства (I.I3) становится отрицательной, и тем самым будет выполнено неравенство (I.I5).

Пусть выполнено неравенство (I.I5). Тогда из соотношения (I.I3) следует, что $\tau_n > \tau_{n-1}$. Из соотношения (I.II) получаем неравенство (I.I6). Таким образом, условие (I.I5) необходимо и достаточно для выполнения неравенства (I.I6).

Теперь докажем справедливость неравенства (I.I5) при выполнении условия (I.I4). Из выражения (I.I2) следует, что $\psi_{n+1}(\tau_n) < 1$, $n = 0,1,\dots$. Поэтому из соотношения (I.6) получаем неравенство $\|\varphi(z_{n+1})\| < \|\varphi(z_n)\|$, $n = 0,1\dots$ Отсюда, согласно второй части леммы, следует неравенство (I.I5). Лемма доказана полностью.

§ 2. Сходимость метода

На вопросы о том, насколько шире область сходимости итераций НАМН, чем в обычном методе Ньютона, и какова их скорость сходимости отвечает следующая

Теорема I. Пусть в сфере $D_1 \; : \|z - z_o\| < \frac{B}{1 - q_1} \|\varphi(z_o)\|$ выполнены условия (I.5),(I.I4). Пусть существует целое положительное число k, такое, что

$$1 \geq \tau_n = \frac{\|\varphi(z_{n-1})\|}{\|\varphi(z_n)\|} \cdot \tau_{n-1} , \; n = 1,\dots,k-1 , \qquad (2.Ia)$$

$$\tau_{k+j} = 1, \; j = 0,1 \dots , \qquad (2.Iб)$$

$$\tau_o \geq q_1^{\,k} , \; q_1 = \psi_1(\tau_o). \qquad (2.2)$$

Тогда уравнение (I.I) имеет решение $z^* \in D_1$, к которому сходятся итерации (I.3)-(I.4), начиная с z_o. Скорость сходимости z_n к z^* определяется неравенством

$$\|z^* - z_n\| \leq B \cdot q_1^{2^{n-k}-1+k} \cdot \|\varphi(z_o)\| , \qquad n \geq k . \qquad (2.3)$$

Доказательство. Согласно способу задания τ_n (2.I) имеют место

$$\tau_o \|\varphi(z_o)\| = \tau_1 \|\varphi(z_1)\| = \dots = \tau_{k-1} \|\varphi(z_{k-1})\| .$$

Следовательно, $\psi_o(\tau_{n-1})$ можно записать в виде

$$\psi_n(\tau_{n-1}) = 1 - \tau_{n-1}(1-q_o) < 1 , \; n=1,\dots,k ,$$

причем, в силу леммы из § I, неравенство (I.I5) выполняется для всех $n=1,2,\dots,k-1$. Отсюда вытекает, что неравенство (I.I6) справедливо при $n=1,2,\dots,k$. Тогда из (I.II) следует, что

$$\tau_n > \tau_{n-1} \; , n=1,\dots,k-1 . \qquad (2.4)$$

Поскольку $z_o \in D_1$, то из (I.3) легко получаем

$$\|z_1 - z_o\| \leq \tau_o B \; \|\varphi(z_o)\| < B \; \|\varphi(z_o)\| , \qquad (2.5)$$

т.е. $z_1 \in D_1$. Далее, из (I.6) имеем

$$\|\varphi(z_1)\| \leq \psi_1(\tau_o) \; \|\varphi(z_o)\| = q_1 \|\varphi(z_o)\| . \qquad (2.6)$$

При этом, согласно (I.I4), справедливо $q_1 < 1$. Если учесть условия (I.5), (2.4), (2.Ia) и (2.6), то из уравнений (I.3) получаем

$$\|z_2 - z_1\| < B \cdot q_1 \; \|\varphi(z_o)\| .$$

Отсюда и с учетом оценки (2.5) получаем

$$\|z_2 - z_o\| \leq \|z_2 - z_1\| + \|z_1 - z_o\| < B(1+q_1) \|\varphi(z_o)\| .$$

Далее, используя лемму (неравенство (I.I5) и неравенство (2.6), из соотношения (I.6) легко получаем

$$\|\varphi(z_2)\| < q_1^2 \; \|\varphi(z_o)\| .$$

Продолжая эти рассуждения ($k - 2$) раз, мы приходим к неравенствам

$$\|z_n - z_o\| < B(1+q_1+q_1^2 + \dots + q_1^{n-1}) \; \|\varphi(z_o)\| , \qquad (2.7а)$$

$$\|\varphi(z_n)\| < q_1^n \; \|\varphi(z_o)\| , \; n = 1, \dots , k . \qquad (2.7б)$$

В силу неравенств (2.7), (2.2) и (I.14) справедливо

$$\bar q_k = \frac{1}{2} MB^2 \cdot \|\varphi(z_k)\| < q_o < 1 \quad . \tag{2.8}$$

Тогда при $\tau_{k+j} = 1$, $j = 0,1...$ итерационный процесс (I.3) фактически превращается в метод Ньютона, начатый с $\bar z_o = z_k$. В этом случае из соотношения (I.6) и формул (I.3) следует, что

$$\|\varphi(z_{k+1})\| \leq \Psi_{k+1}(1) \; \|\varphi(z_k)\| = \bar q_k \; \|\varphi(z_k)\| \quad ,$$

$$\|z_{k+1} - z_k\| \leq \tau_k \cdot B \; \|\varphi(z_k)\| = B \cdot \|\varphi(z_k)\| \quad .$$

Учитывая неравенства (2.7), (2.8) и $q_o < q_1$, последние неравенства можно записать в виде

$$\|\varphi(z_{k+1})\| < q_1 \; \|\varphi(z_k)\| \quad , \quad \|z_{k+1} - z_k\| < B \cdot q_1^k \cdot \|\varphi(z_o)\| \quad .$$

Применение неравенства треугольника, с учетом (2.7а), дает

$$\|z_{k+1} - z_o\| < B(1+q_1+q_1^2+...+q_1^k) \; \|\varphi(z_o)\| \quad .$$

Докажем по индукции неравенства

$$\|z_n - z_o\| < B(1+q_1+q_1^2 + ... + q_1^{n-1}) \; \|\varphi(z_o)\| \quad , \tag{2.9а}$$

$$\|\varphi(z_n)\| \leq \bar q_k^{\,2^{n-k}-1} \cdot \|\varphi(z_k)\| \quad , \; n \geq k \quad . \tag{2.9б}$$

Действительно, мы доказали неравенства (2.9) при $n = k+1$, пусть они верны при $n = k + \bar m$ и мы докажем соотношения (2.9) на следующем шаге $n + 1 = k + \bar m + 1$. Так как $z_n \in D_1$, то из (I.3) с учетом оценок (I.5), (2.9б) и (2.7б) получаем

$$\|z_{n+1} - z_n\| \leq \tau_n \cdot B \; \|\varphi(z_n)\| < B \; q_1^{2^{n-k}-1+k} \|\varphi(z_o)\| .$$

Отсюда, используя неравенство (2.9а) и элементарное неравенство $2^n - 1 \geq n$, перейдем к оценке

$$\|z_{n+1} - z_o\| \leq B(1+q_1+q_1^2 + ... + q_1^n) \; \|\varphi(z_o)\| .$$

Далее, из неравенства (I.6), с учетом соотношений (2.8), (2.9б), получаем

$$\|\varphi(z_{n+1})\| \leq \Psi_{n+1}(1) \cdot \|\varphi(z_n)\| \leq \frac{1}{2} MB^2 \; \bar q_k^{\,2(2^{n-k}-1)} \; \|\varphi(z_k)\|^2 \leq$$

$$\leq \bar q_k^{\,2^{n+1-k}-1} \cdot \|\varphi(z_k)\| \quad ,$$

т.е. неравенства (2.9) доказаны на (n + 1)- м шаге. Из (2.9) видно, что для всех n приближения z_n принадлежат сфере D_1. Покажем фундаментальность последовательности $\{z_n\}$. Из уравнений (I.3) с учетом оценок (2.7), (2.9) легко получаем

$$\|z_{n+1} - z_n\| < B \; q_1^n \; \|\varphi(z_o)\| \quad , \quad n > k \quad ,$$

$$\|z_{n+2} - z_{n+1}\| < B \cdot q_1^{n+1} \; \|\varphi(z_o)\| \quad ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\|z_{n+p} - z_{n+p-1}\| < B \cdot q_1^{n+p-1} \; \|\varphi(z_o)\| \quad ,$$

откуда приходим к неравенству

$$\|z_{n+p} - z_n\| < B \left(\sum_{i=0}^{p-1} q_1^{n+i}\right) \cdot \|\varphi(z_o)\| \quad .$$

Отсюда ясно, что $\|z_{n+p} - z_n\| \to 0$ при $n \to \infty$ равномерно по p. Итак $\{z_n\}$ – фундаментальная, а вследствие полноты пространства X –сходящаяся последовательность. Обозначим через z^* ее предел. Переходя к пределу в первом неравенстве (2.9), получаем

$$\|z^* - z_o\| < B \left(\sum_{i=0}^{\infty} q_1^i\right) \; \|\varphi(z_o)\| = \frac{B}{1-q_1} \; \|\varphi(z_o)\| \; ,$$

т.е. $z^* \in D_1$. Предельный переход во втором из неравенств (2.9) дает

$$\|\varphi(z^*)\| = 0 \; .$$

Неравенство (2.3) следует из равенства

$$z^* - z_n = \varphi'(\bar z_n)^{-1}(\varphi(z^*) - \varphi(z_n)) \; , \; \bar z_n = (1-\alpha_n)z_n + \alpha_n z^* \; ,$$

$$\alpha_n \in (0,1) \; ,$$

с учетом неравенств (2.7) и (2.9). Теорема доказана полностью.

<u>Следствие I</u>. Пусть начальное приближение z_o таково, что

$$\bar q_o = \frac{1}{2} MB^2 \| \varphi(z_o) \| < 1 \; . \tag{2.I'}$$

Тогда классический метод Ньютона сходится и справедлива оценка

$$\|z^* - z_n\| \leq B \cdot \bar q_o^{\,2^n-1} \|\varphi(z_o)\| . \tag{2.3'}$$

Действительно, условие (2.8) выполняется здесь при $k = 0$. При $\tau_0 = 1$ имеем $q_o = \bar q_o$, а неравенство (2.2) автоматически выполняется, поскольку здесь $k = 0$. Неравенство (2.3) превращается в (2.3'). Отметим, что оценка (2.3') лучше, чем известная оценка для метода Ньютона (см. /7/ стр. 403).

Следствие 2. Пусть $\bar{q}_0 \le 0{,}5$. Тогда классический метод Ньютона является оптимальным по скорости сходимости в смысле определения, данного в § I. Действительно, легко проверить, что все экстремальные точки τ_n^* функции (I.7) удовлетворяют условию $\tau_n^* \ge 1$, $n = 0,1\ldots$. Следовательно,

$$\psi_{n+1}(\tau_n) = \min \psi_{n+1}(\tau) = \psi_{n+1}(1).$$

Отсюда следует рассматриваемое утверждение.

Замечание I. Сравнение условий (I.I4) и (2.I') показывает, насколько шире область сходимости итераций НАМН, чем у метода Ньютона[8].

Литература

I. Гавурин М.К. Изв. вузов, математика, 1958.т.5/6.с.I8.
2. Жидков Е.П., Пузынин И.В. ДАН СССР, I968, т.I80, №I, с.I8.
3. Гареев Ф.А. и др. ЖВМиМФ, I077, т.I7, №2, с.407.
4. Пузынин И.В. Автореферат дис.канд.физ.-мат. наук, ОИЯИ, II-4735, Дубна, I969.
5. Ермаков В.В., Калиткин Н.Н. ЖВМ и МФ, I98I, т.2I, № 2, с.49I.
6. Жанлав Т., Пузынин И.В. ОИЯИ, PII-9I-I00, Дубна, I99I.
7. Треногин В.А. Функциональный анализ - М.: Наука, I980.
8. Канторович Л.В., Акилов Г.П. Функциональный анализ.—М.:Наука,I984.

Рукопись поступила в издательский отдел
I9 декабря I99I года.