Л. В. Канторович, О методе Ньютона,
*Тр. МИАН СССР*, 1949, том 28, 104–144

https://www.mathnet.ru/tm1059

ТРУДЫ
МАТЕМАТИЧЕСКОГО ИНСТИТУТА ИМЕНИ В. А. СТЕКЛОВА
1949 г., том XXVIII

Л. В. КАНТОРОВИЧ

# О МЕТОДЕ НЬЮТОНА

Одним из наиболее эффективных методов нахождения корней алгебраических уравнений в случае, когда для корня известно приближенное значение, является метод Ньютона, иногда называемый также методом касательных. В этом методе последовательные приближения определяются формулами вида:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Сходимость этого метода была исследована Коши, а затем A. Ostrowski [1—3].

Этот метод был распространен на системы алгебраических уравнений. Условия сходимости его для этого случая были даны Willers [4], Стениным [5], Ostrowski [6].

Однако оказывается возможным применить этот метод и для случая любых нелинейных уравнений. В частности, по моему предложению, для случая нелинейных интегральных уравнений он был применен Д. М. Загадским [7,8].

Чтобы охватить одновременно все случаи, наиболее удобно развить его и исследовать в общем виде — для любых функциональных уравнений. С целью облегчить чтение работы читателю, незнакомому с теорией операций, в § 1—3 мы приводим те сведения, которые для нас существенны.

В § 4 излагается основная теорема о сходимости процесса Ньютона. В ней даются, в некотором смысле, окончательные условия сходимости метода и устанавливается быстрота сходимости. Эта теорема представляет одновременно некоторую теорему существования и единственности для нелинейных уравнений, а потому имеет и чисто теоретический интерес.

В § 5 изложен модифицированный метод, имеющий в некоторых случаях определенное преимущество.

В § 6 даны применения метода к алгебраическим уравнениям и системам, а в § 7 к нелинейным интегральным уравнениям; § 8 содержит изложение некоторого метода нахождения собственных чисел и векторов, получающегося в результате применения к этой задаче процесса Ньютона.

## § 1. Линейные функциональные операции*

Линейным нормированным пространством называется линейное или векториальное множество, т. е. множество $X = \{x\}$ элементов любой природы, для которых определены операции сложения $x+y$ и умножения элемента на вещественное число $\lambda x$, подчиненные обычным законам алгебры, а также определена норма элемента $\| x \|$ — число, обладающее обычными свойствами длины вектора.

Точнее говоря, для нормы должны быть удовлетворены следующие условия:

1) $\| x \| \geqslant 0$ и $\| x \| = 0$ тогда и только тогда, когда $x = 0$;

2) $\| x + y \| \leqslant \| x \| + \| y \|$;

3) $\| \lambda x \| = |\lambda| \, \| x \|$.

Наличие нормы позволяет определить понятие сходимости: говорят, что $x_n \to x$, или $\lim\limits_{n\to\infty} x_n = x$, если $\| x_n - x \| \to 0$ при $n \to \infty$. Пространство называется полным или типа $B$ (по имени С. Банаха), если для сходимости выполнен принцип Коши, т. е. из того, что

$$\lim_{n\to\infty} \| x_n - x_{n+p} \| = 0$$

равномерно относительно $p$ следует сходимость последовательности к некоторому элементу $x$, т. е. $x_n \to x$. Приведем примеры пространств, которыми нам придется пользоваться.

1. Множества вещественных или комплексных чисел представляют очевидным образом пространства типа $B$, если операции $x+y$ и $\lambda x$ определить как обычно для чисел и за норму принять модуль числа, т. е. $\| x \| = |x|$.

2. Пространство Эвклида $R^n$ или, что то же самое, пространство $n$-мерных векторов

$$x = (\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(n)})$$

представляет пространство типа $B$, если операции $x+y$ и $\lambda x$ производить, как обычно, для векторов (по компонентам), а в качестве нормы принять длину вектора

$$\| x \| = \sqrt{(\xi^{(1)})^2 + (\xi^{(2)})^2 + \ldots + (\xi^{(n)})^2}.$$

3. Из тех же элементов, что и в предыдущем примере, мы можем образовать другое пространство $m_n$, определив иным образом норму — взяв в качестве нее максимальную из абсолютных величин компонент

$$\| x \| = \max | \xi^{(i)} |.$$

---

* По поводу основных понятий функционального анализа см., например, статью Л. А. Люстерника [9].

Очевидно, и в этом случае все условия будут соблюдены. В частности, пространство будет полным и сходимость последовательности $x_m \to x$ будет означать здесь, так же как и в предыдущем случае, покоординатную сходимость: $\xi_m^{(i)} \to \xi^{(i)}$ при $m \to \infty$ $(i = 1, 2, \ldots, n)$.

4. Пространство $C$ непрерывных функций, определенных в некотором промежутке $(a, b)$, где принята норма

$$\| x \| = \max_{a \leqslant t \leqslant b} | x(t) | .$$

Это также полное пространство и сходимость в нем есть равномерная сходимость: $x_n \to x$ означает, что последовательность функций $x_n(t)$ равномерно сходится к $x(t)$.

5. Пространство $L^2$ интегрируемых с квадратом функций, определенных в промежутке $(a, b)$, где норма

$$\| x \| = \left[ \int_a^b x^2(t)\, dt \right]^{\frac{1}{2}},$$

а сходимость означает сходимость в среднем.

Операцией, переводящей одно такое пространство $X$ в другое $Y$, или оператором, называется функция $y = U(x)$, которая каждому элементу $x$ пространства $X$ относит элемент $y$ пространства $Y$. В частности, если $Y$ — пространство вещественных чисел, то операция называется функционалом. Операция $y = U(x)$ называется линейной, если она аддитивна:

$$U(x_1 + x_2) = U(x_1) + U(x_2)$$

и непрерывна, т. е. если $U(x_n) \to U(x)$ при $x_n \to x$. Линейная операция однородна. Это означает, что $U(\lambda x) = \lambda U(x)$.

Вместо непрерывности можно поставить равносильное условие — наличие такой постоянной $C$, что

$$\| U(x) \| \leqslant C \| x \| .$$

Наименьшая постоянная $C$, обеспечивающая выполнение этого неравенства, называется нормой данной линейной операции и обозначается $\| U \|$.

Отметим основное неравенство

$$\| U(x) \| \leqslant \| U \| \cdot \| x \| . \tag{1}$$

Множество линейных операций, переводящих $X$ в $Y$, будем обозначать $(X \to Y)$; оно само представляет также линейное множество и полное нормированное пространство, если норму $\| U \|$ определить, как указано выше.

Оператором, обратным данному $y = U(x)$, называется оператор $x = U^{-1}(y)$, отображающий $Y$ в $X$ такого рода, что

$$U(U^{-1}(y)) = y.$$

Такой обратный оператор существует не всегда. Отметим случай, когда его существование обеспечено, С. Банах [10].

Если $U$ — линейный оператор, отображающий $X$ в $X$ с нормой

$$\| U \| < 1$$

и $I$ — тождественный оператор в $X$, т. е. $Ix = x$, то оператор $I - U$ имеет обратный и при этом

$$\| (I - U)^{-1} \| \leqslant \frac{1}{1 - \| U \|}. \tag{2}$$

Для доказательства рассмотрим оператор

$$Hx = x + Ux + U^2 x + \ldots$$

Этот ряд сходится и определяет линейный оператор, так как мы имеем:

$$\| Hx \| \leqslant \| x \| + \| U \| \| x \| + \| U \|^2 \| x \| + \ldots \leqslant \frac{\| x \|}{1 - \| U \|}; \ \| H \| \leqslant \frac{1}{1 - \| U \|}.$$

Далее, очевидно,

$$(I - U) Hx = (x + Ux + U^2 x + \ldots) - (Ux + U^2 x + \ldots) = x,$$

а это показывает, что $H$ есть оператор, обратный $I - U$. Итак,

$$(I - U)^{-1} = H; \ \| (I - U)^{-1} \| = \| H \| \leqslant \frac{1}{1 - \| U \|}.$$

Приведем некоторые примеры линейных операций.

Пример 1. Рассмотрим линейные операции $y = U(x)$, переводящие пространство $X = m_n$ в $Y = m_\nu$. Введем элементы $x_1 = (1, 0, 0, \ldots, 0)$, $x_2 = (0, 1, 0, \ldots, 0), \ldots, x_n = (0, 0, 0, \ldots 1)$ и пусть

$$y_i = U(x_i) = (a_{i1}, a_{i2}, \ldots, a_{i\nu}) \quad (i = 1, 2, \ldots, n).$$

Тогда для любого

$$x = (\xi^{(1)}, \ldots, \xi^{(n)}) = \sum_{i=1}^{n} \xi^{(i)} x_i$$

имеем:

$$y = U(x) = (\eta^{(1)}, \ \eta^{(2)}, \ldots, \eta^{(\nu)}), \tag{3}$$

где

$$\eta^{(j)} = \sum_{i=1}^{n} a_{ij} \xi^{(i)} \quad (j = 1, 2, \ldots, \nu).$$

Таким образом, операция $U$ есть линейное преобразование, определяемое матрицей

$$A = \| a_{ij} \| \quad (i = 1, \ldots, n, \ j = 1, \ldots, \nu).$$

Покажем теперь, что норма этой операции определяется так:

$$\| U \| = \max_{j} \sum_{i} | a_{ij} |. \tag{4}$$

Действительно, имеем:

$$\| y \| = \| U(x) \| = \max_j |\eta^{(j)}| = \max_j \left| \sum_{i=1}^n \xi^{(i)} a_{ij} \right| \leqslant$$

$$\leqslant \max_i |\xi^{(i)}| \max_j \sum_{i=1}^n |a_{ij}| = \left( \max_j \sum_{i=1}^n |a_{ij}| \right) \| x \| .$$

Отсюда, собственно, следует, что $\| U \|$ не превосходит правой части (4), но легко проверить, что удовлетворяется и равенство.

Из (4) следует, очевидно, оценка

$$\| U \| \leqslant nL; \quad L = \max | a_{ij} | . \tag{4a}$$

Пример 2. Рассмотрим случай, когда $X = R^n$, $Y = R^\nu$ — эвклидовы пространства. В этом случае линейная операция дается в той же форме линейного преобразования (3) и лишь норма ее определяется иным образом. Можно показать, что в случае, когда матрица $A$ симметрична,

$$\| U \| = | \Lambda_n | , \tag{5}$$

где $\Lambda_n$ — наибольшее по модулю собственное значение этой матрицы. В случае, если матрица $A$ не симметрична, то

$$\| U \| = \sqrt{\Lambda_n}, \tag{6}$$

где $\Lambda_n$ — наибольшее собственное значение матрицы $AA^*$. Эти факты легко получаются на основании теорем об экстремальных свойствах собственных значений матриц, устанавливаемых в алгебре, либо на основании соображений, подобных приведенным ниже, при рассмотрении примера 5.

Кроме точного выражения нормы, может оказаться полезной и следующая оценка ее:

$$\| U \| \leqslant \left( \sum_{j=1}^{\nu} \sum_{i=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}. \tag{7}$$

Последняя устанавливается так:

$$\| y \|^2 = \sum_{j=1}^{\nu} \eta_j^2 = \sum_{j=1}^{\nu} \left( \sum_{i=1}^{n} a_{ij} \xi_i \right)^2 \leqslant \left( \sum_{j=1}^{\nu} \sum_{i=1}^{n} a_{ij}^2 \right) \sum_{i=1}^{n} \xi_i^2 =$$

$$= \left( \sum_{j=1}^{\nu} \sum_{i=1}^{n} a_{ij}^2 \right) \| x \|^2.$$

Пример 3. Рассмотрим линейные операции в пространстве комплексных чисел $X = \{x\}$; $x = x_1 + ix_2$. Такая операция, вообще говоря, определяется некоторой матрицей, но мы ограничимся операциями частного вида:

$$y = U(x) = wx, \tag{8}$$

где $w$ — комплексное число. Тогда очевидно, что

$$|y| = |w|\,|x|.$$

Так как в данном случае норма элемента равна модулю числа, то, очевидно,

$$\|U\| = |w|. \tag{9}$$

Пример 4. Рассмотрим операции, отображающие $C$ в $C$. Здесь мы ограничимся только операциями интегрального типа:

$$v = U(x);\ y(s) = \int_0^1 K(s, t)\, x(t)\, dt, \tag{10}$$

причем $K(s, t)$ считаем непрерывным. Эта операция, очевидно аддитивна. Оценим ее норму:

$$\|y\| = \max_s |y(s)| = \max_s \left| \int_0^1 K(s, t)\, x(t)\, dt \right| \leqslant$$

$$\leqslant \|x\| \max_{0 \leqslant s \leqslant 1} \int_0^1 |K(s, t)|\, dt.$$

Мы видим, что норма $U$ не превосходит второго множителя, но нетрудно показать, что она равна ему:

$$\|U\| = \max_{0 \leqslant s \leqslant 1} \int_0^1 |K(s, t)|\, dt. \tag{11}$$

В частности,

$$\|U\| \leqslant M, \text{ если } |K(s, t)| \leqslant M. \tag{11a}$$

Пример 5. Пусть теперь $X = Y = L^2$ есть пространство интегрируемых с квадратом функций. Ограничимся рассмотрением только операций частного вида:

$$y(s) = x(s) - \int_0^1 K(s, t)\, x(t)\, dt, \tag{12}$$

сначала для случая, когда ядро $K(s, t)$ симметрично.

Обозначим через $\lambda_i$ собственные числа ядра $K(s, t)$, и через $\varphi_i(t)$ — собственные функции его. Дополним систему функций $\{\varphi_i(t)\}$ функциями $\varphi_0(t), \varphi_{-1}(t), \dots$ так, чтобы в результате получилась полная система в $L^2$. При этом будем считать $\lambda_0 = \lambda_{-1} = \lambda_{-2} = \dots = \infty$. Далее обозначим через $\xi_i$ и $\eta_i$ коэффициенты Фурье функций $x(t)$ и $y(t)$ по системе $\varphi_i$. Тогда, подставляя билинейное разложение для $K$ и разложения для $x$ и $y$

$$K(s, t) = \sum_i \frac{\varphi_i(t)\, \varphi_i(s)}{\lambda_i};\ x(t) = \sum_i \xi_i \varphi_i(t);\ y(t) = \sum_i \eta_i \varphi_i(t)$$

в выражение оператора (12), найдем:

$$\eta_i = \xi_i - \frac{1}{\lambda_i}\xi_i.$$

Отсюда ясно, что, принимая $L$ равным точной верхней границе чисел

$$\left|1 - \frac{1}{\lambda_i}\right|,$$

имеем:

$$\| U(x) \|^2 = \| y \|^2 = \sum_i \eta_i^2 = \sum_i \left(1 - \frac{1}{\lambda_i}\right)^2 \xi_i^2 \leqslant L^2 \sum_i \xi_i^2 = L^2 \| x \|^2.$$

Следовательно, в данном случае

$$\| U \| \leqslant L = \sup \left|1 - \frac{1}{\lambda_i}\right|. \tag{13}$$

Если $| K(s, t) | \leqslant M$, то из $| \lambda_i | \geqslant \frac{1}{M}$ следует неравенство:

$$\| U \| \leqslant 1 + M. \tag{13a}$$

Также легко оценить отсюда и норму обратного оператора. Действительно, имеем:

$$\| x \|^2 = \sum_i \xi_i^2 = \sum \frac{1}{\left(1 - \frac{1}{\lambda_i}\right)} \eta_i^2 \leqslant L_1^2 \| y \|^2,$$

где $L_1 = \sup \frac{1}{\left|1 - \frac{1}{\lambda_i}\right|}$. Следовательно,

$$\| U^{-1} \| L_1 = \sup \frac{1}{\left|1 - \frac{1}{\lambda_i}\right|}. \tag{14}$$

В случае несимметричного ядра $K(s, t)$ легко убедиться, что

$$\int y^2(s)\,ds = \int x^2(s)\,ds - \iint \overline{K}(u, t)\,x(u)\,x(t)\,du\,dt,$$

где

$$\overline{K}(u, t) = K(u, t) + K(t, u) - \int K(s, t)\,K(s, u)\,ds. \tag{15}$$

Поэтому, если через $\Lambda_i$ обозначить собственные числа симметричного ядра $\overline{K}(u, t)$, то, рассуждая как и выше, найдем:

$$\| U \| \leqslant \sup_i \sqrt{\left|1 - \frac{1}{\Lambda_i}\right|}, \tag{16}$$

$$\| U^{-1} \| \leqslant \sup_i \frac{1}{\sqrt{\left|1 - \frac{1}{\Lambda_i}\right|}}. \tag{17}$$

## § 2. Билинейные операции

Рассмотрим линейную операцию

$$u = B(x),$$

переводящую пространство $X$ в пространство линейных операций, отображающих $X$ в $Y$, т. е. $u \in (X \to Y)$. Вычислим ее значение для произвольного элемента $x' \in X$. Тогда, полагая

$$B(x, x') = u(x') = B(x)(x'), \tag{1}$$

получаем, очевидно, операцию, определенную для пары элементов $x$ и $x'$, аддитивную по каждому аргументу и такую, что

$$\| B(x, x') \| \leqslant \| u \| \; \| x' \| \leqslant \| B \| \; \| x \| \; \| x' \| . \tag{2}$$

Операция, удовлетворяющая последним условиям, называется билинейной, а наименьший допустимый постоянный множитель в неравенстве типа (2) — ее нормой.

Обратно, пусть дана некоторая билинейная операция $B(x, x')$, аддитивная по обоим аргументам и удовлетворяющая условию

$$\| B(x, x') \| \leqslant C \| x \| \; \| x' \| .$$

Тогда ясно, что при $x$ постоянном $B(x, x')$ представляет некоторую линейную операцию $u(x')$, переводящую $X$ в $Y$. Полагая $B(x) = u$, имеем:

$$\| B(x) \| = \| u \| \leqslant C \| x \| .$$

При этом, так как $B(x)$ — аддитивная операция от $x$, то $u = B(x)$ есть линейная операция, отображающая $X$ в $(X \to Y)$ с нормой $\| B \| \leqslant C$. Таким образом, из сказанного ясно, что по существу эквивалентно рассматривать $B$ как операцию, отображающую $X$ в $(X \to Y)$ или как билинейную. Значение нормы в обоих случаях также одно и то же. Приведем примеры билинейных операций.

Пример 1. Рассмотрим билинейную операцию, переводящую пространство $X = m_n$ в $Y = m_\nu$. Легко усмотреть, что она имеет вид:

$$y = B(x, x') = \left\{ \sum_{i, j=1}^{n} a_{ijk} \xi_i \xi_j' \right\}_{k=1,2,\ldots,\nu}, \tag{3}$$

т. е. ее значение есть вектор $y$, компоненты которого — квадратичные формы. Ясно, что

$$\| y \| = \| B(x, x') \| = \max_k \left| \sum_{i, j=1}^{n} a_{ijk} \xi_i \xi_j' \right| \leqslant \max_k \left( \sum_{i, j=1}^{n} | a_{ijk} | \right) \| x \| \; \| x' \| ,$$

откуда следует, что

$$\| B \| \leqslant \max_k \sum_{i, j=k}^{n} | a_{ijk} | \leqslant n^2 M, \tag{4}$$

где $| a_{ijk} | \leqslant M$. Однако эти оценки не дают точного значения нормы.

Пример 2. Билинейная операция, отображающая $R^n$ в $R^\nu$, имеет тот же вид, что и в предыдущем случае. Однако норма ее определяется и оценивается иначе.

Именно:

$$\left|\sum_{i,j=1}^{n} a_{ijk}\xi_i\xi_j'\right|^2 \leqslant \left(\sum_{i,j=1}^{n} a_{ijk}^2\right)\left(\sum_{i,j=1}^{n} \xi_i^2\xi_j'^2\right) = \left(\sum_{i,j=1}^{n} a_{ijk}^2\right)\left(\sum_{i=1}^{n}\xi_i^2\right)\left(\sum_{j=1}^{n}\xi_j^2\right).$$

Отсюда

$$\| y \|^2 = \| B(x, x') \|^2 = \sum_{k=1}^{\nu}\left(\sum_{i,j=1}^{n} a_{ijk}\,\xi_i\,\xi_j'\right)^2 \leqslant \sum_{k=1}^{\nu}\sum_{i,j=1}^{n} a_{ijk}^2 \,\| x \|^2 \,\| x' \|^2$$

и, следовательно

$$\| B \| \leqslant \left(\sum_{k=1}^{\nu}\sum_{i,j=1}^{n} a_{ijk}^2\right)^{\frac{1}{2}} \leqslant n\sqrt{\nu}L, \tag{5}$$

если все $| a_{ijk} | \leqslant L$.

Пример 3. Если $X$ и $Y$ — пространства комплексных чисел, то примером билинейной операции будет операция вида:

$$y = B(x, x') = wxx', \tag{6}$$

где $w$ — комплексное число. Легко убедиться, что для такой операции

$$\| B \| = | w |. \tag{7}$$

Пример 4. Примером билинейной операции, отображающей $C$ в $C$, является интегральная операция вида:

$$y = B(x, x'); \quad y(s) = \int_0^1\int_0^1 K(s, t, u)\, x(t)\, x'(u)\, dt\, du. \tag{8}$$

Ее норма оценивается так:

$$\| B \| \leqslant \sup_s \int_0^1\int_0^1 | K(s, t, u) |\, dt\, du \leqslant M, \tag{9}$$

если $| K(s, t, u) | \leqslant M$.

Пример 5. Ту же операцию (8) можно рассматривать как операцию, отображающую $L^2$ в $L^2$. В этом случае ее норма может быть оценена так:

$$\| B \| \leqslant \left[\int_0^1\int_0^1\int_0^1 K^2(s, t, u)\, ds\, dt\, du\right]^{\frac{1}{2}}. \tag{10}$$

## § 3. Дифференцирование нелинейных операций

Пусть

$$y = P(x)$$

— нелинейная операция, переводящая пространство $X$ в $Y$. Говорят что она дифференцируема (в смысле Fréchet [11]) при данном значении $x$, если имеется такая линейная операция $U \in (X \to Y)$, что

$$\| [P(x + \Delta x) - P(x)] - U(\Delta x) \| \leqslant \| \Delta x \| \, \varepsilon(\| \Delta x \|); \qquad (1)$$

здесь $\varepsilon(\delta)$ — функция, стремящаяся к нулю, когда $\delta \to 0$. Эту операцию $U$ называют производной для операции $P(x)$ при данном $x$:

$$P'(x) = U. \qquad (2)$$

Как сказано, $P'(x)$ есть элемент пространства $(X \to Y)$.

В свою очередь, $U = P'(x)$ есть нелинейная операция, переводящая пространство $X$ в пространство $(X \to Y)$. Она также может оказаться дифференцируемой. Ее производная называется по отношению к нелинейной операции $P(x)$ второй производной:

$$V = [P'(x)]' = P''(x). \qquad (3)$$

Эта вторая производная представляет элемент пространства $[X \to (X \to Y)]$, т. е. пространства линейных операций, переводящих $X$ в $(X \to Y)$. Как мы видели (§ 2), рассмотрение такой операции эквивалентно рассмотрению билинейной операции, отображающей пространства $X$ в $Y$, так что $P''(x)$ можно рассматривать как такую билинейную операцию. В соответствии с этим под $\| P'(x) \|$ и $\| P''(x) \|$ следует понимать нормы, взятые соответственно в пространствах $(X \to Y)$ или $[X \to (X \to Y)]$. (См. Гавурин М. К. [17]).

Отметим некоторые предложения о производных, которыми нам придется пользоваться.

1. Если $y = \varphi(x)$, а $z = F(y) = F[\varphi(x)]$, причем функции $\varphi$ и $F$ дифференцируемы, то

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = F'[\varphi(x)]\,\varphi'(x). \qquad (4)$$

Здесь следует разуметь, что написанные рядом линейные операции $\frac{dz}{dy}$ и $\frac{dy}{dx}$ должны применяться последовательно (правило дифференцирования сложной функции).

Это правило устанавливается так же, как в случае обычных производных функций одного или нескольких переменных.

2. Если $y = P(x)$ — линейная операция, отображающая $X$ в $Y$, то, очевидно,

$$P'(x) = P; \quad P''(x) = 0, \qquad (5)$$

т. е. производная линейной операции совпадает с ней самой.

3. Если $U$ — линейная операция, отображающая $Y$ в $Z$, то

$$[U(P(x))]' = U(P'(x)), \tag{6}$$

т. е. постоянную операцию можно выносить за знак производной. Это вытекает сразу из правил 1 и 2. При этом

$$U(P'(x)) = V \in (X \to Z) \text{ и } V(x') = U[(P'(x))(x')].$$

4. Если $P(x)$ — дифференцируемая операция, то справедливо неравенство [12]:

$$\| P(\bar{x} + \Delta x) - P(x) \| \leqslant \sup \| P'(\bar{x}) \| \ \| \Delta x \| , \tag{7}$$

$$\bar{x} = x + \theta \Delta x; \ 0 \leqslant \theta \leqslant 1,$$

представляющее оценку приращения, подобную той, которая для обычной функции получается из формулы конечных приращений.

Для доказательства положим:

$$P(x + \Delta x) - P(x) = y,$$

и подберем, что всегда возможно [9], такой линейный функционал $T$ в пространстве $Y$, для которого

$$\| T \| = 1; \ T(y) = \| y \| .$$

Рассмотрим вещественную функцию вещественного переменного $t$:

$$f(t) = T[P(x + t\Delta x)].$$

Для ее производной, пользуясь при дифференцировании правилами 1 и 3, находим выражение:

$$f'(t) = TP'(x + t\Delta x)\Delta x.$$

Далее, пользуясь определением $f(t)$ и применяя обычную формулу конечных приращений, имеем:

$$T(y) = T[P(x + \Delta x) - P(x)] = f(1) - f(0) = f'(\theta) = TP'(x + \theta \Delta x)\Delta x.$$

Отсюда ясно, что

$$\| P(x + \Delta x) - P(x) \| = \| y \| = T(y) \leqslant$$

$$\leqslant \| T \| \ \| P'(x + \theta \Delta x) \| \ \| \Delta x \| \leqslant \sup \| P'(\bar{x}) \| \ \| \Delta x \| , \quad 0 \leqslant \theta \leqslant 1.$$

$$\bar{x} = x + \theta \Delta x.$$

5. Если $P(x)$ — дважды дифференцируемая функция, то справедливо следующее неравенство:

$$\| P(x + \Delta x) - P(x) - P'(x)\Delta x \| \leqslant \frac{1}{2} \sup \| P''(\bar{x}) \| \ \| \Delta x \|^2, \tag{8}$$

$$\bar{x} = x + \theta \Delta x,$$

связанное с формулой Тэйлора, подобно тому, как предыдущее связано с формулой конечных приращений.

Доказательство проводится аналогично предыдущему. Обозначая через $y$ элемент в левой части (8), норму которого надлежит оценить, вводим такой линейный функционал $T$, что $\| T \| = 1$ и $T(y) = \| y \|$. Далее строим вспомогательную функцию

$$f(t) = T[P(x + t\Delta x)].$$

Для нее

$$f'(t) = T[P'(x + t\Delta x)\Delta x],$$

$$f''(t) = T[P''(x + t\Delta x)\Delta x \Delta x],$$

где выражение в квадратных скобках означает, что билинейная операция $P''(x + t\Delta x)$ должна быть вычислена от пары аргументов, равных $\Delta x$. Теперь, применяя обычную формулу Тэйлора, находим, что

$$\| P(x + \Delta x) - P(x) - P'(x)\Delta x \| = \| y \| = T(y) =$$
$$= f(1) - f(0) - f'(0) = \frac{1}{2} f''(\theta) \leqslant \frac{1}{2} \| T \| \sup \| P''(x + t\Delta x) \| \| \Delta x \|^2;$$
$$0 \leqslant t \leqslant 1.$$

Приведем теперь некоторые примеры на дифференцирование операций.

Пример 1. Рассмотрим нелинейную операцию, переводящую $n$-мерное пространство в $\nu$-мерное. Она определяется совокупностью $\nu$ функций от $n$ переменных:

$$y = P(x), \quad \eta_k = f_k(\xi_1, \xi_2, \ldots, \xi_n), \quad (k = 1, 2, \ldots, \nu). \tag{9}$$

Будем предполагать, что функции $f_k$ имеют непрерывные частные производные 2-го порядка. Тогда для дифференциалов имеем:

$$d\eta_k = \sum_{i=1}^{n} \frac{\partial f_k}{\partial \xi_i} d\xi_i \quad (k = 1, 2, \ldots, \nu), \tag{10}$$

а приращение $\Delta y = \{\Delta \eta_k\}_{k=1,2,\ldots,\nu}$ выражается такой же системой форм от дифференциалов с точностью до бесконечно-малых высших порядков. Отсюда ясно, что в данном случае $P'(x)$ дается матрицей частных производных:

$$P'(x) = \left\| \frac{\partial f_k}{\partial \xi_i} \right\|_{\substack{i=1,2,\ldots,n \\ k=1,2,\ldots,\nu}}, \tag{11}$$

точнее говоря, $P'(x)$ есть линейное преобразование, соответствующее этой матрице [ср. § 1, (2)].

Аналогичным образом рассматривая приращение $P'(x)$ при приращении аргумента $\Delta x' = (\Delta \xi_1', \ldots, \Delta \xi_n')$, убедимся, что вторая производная определяется матрицей, зависящей от трех индексов:

$$P''(x) = \left\| \frac{\partial f_k}{\partial \xi_i \partial \xi_j} \right\|_{\substack{i,j=1,2,\ldots,n \\ k=1,2,\ldots,\nu}}. \tag{12}$$

Если рассматривать ее как билинейную операцию, то она будет определяться системой $\nu$ билинейных форм:

$$P''(x)\,xx' = \left\{ \sum_{i,\,j=1}^{n} \frac{\partial f_k}{\partial \xi_i \partial \xi_j} \xi_i\, \xi_j{}' \right\}_{k=1,2,\,\ldots,\,\nu} . \tag{13}$$

При выборе определенной нормы в $n$-мерных пространствах, соответствующих нормировке $R^n$ или $m_n$, можно на основании результатов §§ 1 и 2 указать оценки для $\| P'(x) \|$ и $\| P''(x) \|$.

Пример 2. Если в пространстве комплексных чисел рассмотреть аналитическую функцию

$$y = P(x),$$

то в данном случае

$$\Delta y = P'(x)\,\Delta x$$

с точностью до бесконечно малых высшего порядка, поэтому операция $P'(x)$ есть умножение на комплексное число $P'(x)$, а норма ее

$$\| P'(x) \| = | P'(x) | . \tag{14}$$

Таким же образом в данном случае и вторая производная совпадает с обычной второй производной, если последнюю рассматривать как билинейную операцию над парой комплексных чисел $P''(x)\cdot x\cdot x'$.

Пример 3. Рассмотрим нелинейную интегральную операцию:

$$y = P(x); \quad y(s) = \int_0^1 K(s, t, x(t))\, dt, \tag{15}$$

где $K(s, t, u)$ — дважды непрерывно-дифференцируемая функция своих аргументов. Тогда, с точностью до малых высших порядков, имеем

$$\Delta y(s) = \int_0^1 K_x{}'(s, t, x(t))\, \Delta x(t)\, dt,$$

откуда ясно, что $P'(x)$ в данном случае есть линейная интегральная операция с ядром $K(s, t) = K_x{}'(s, t, x(t))$:

$$P'(x)\,\Delta x = \int_0^1 K(s, t)\,\Delta x(t)\, dt = \int_0^1 K_x{}'(s, t;\ x(t))\,\Delta x(t)\, dt. \tag{16}$$

Придавая теперь $x(t)$ приращение $\Delta' x(t)$, убеждаемся в том, что с точностью до малых высших порядков

$$[P'(x + \Delta' x) - P'(x)]\,\Delta x = [\Delta P'(x)]\,\Delta x = \int_0^1 K''_{x^2}(s, t, x(t))\,\Delta' x(t)\,\Delta x(t)\, dt,$$

откуда ясно, что вторая производная в данном случае есть билинейная интегральная операция специального вида:

$$P''(x)\,\Delta x\,\Delta x' = \int_0^1 K_2(s, t)\,\Delta x(t)\,\Delta' x(t)\, dt,$$

где

$$K_2(s, t) = K''_{x^2}(s, t, x(t)). \tag{17}$$

## § 4. Сходимость процесса Ньютона

Рассмотрим применение процесса Ньютона к нелинейному функциональному уравнению

$$P(x)=0, \tag{1}$$

где $P(x)$ — операция, переводящая пространство $X$ в $Y$, которая предполагается дважды дифференцируемой. Формулы, связывающие последовательные приближения, строятся на основании соображений, аналогичных тому, как это делается в случае вещественных уравнений.

Пусть $x_0$ — начальное приближение к решению. Заменяя приращения $P(x)-P(x_0)$ на дифференциал в точке $x_0$, заменим данное уравнение приближенно на линейное:

$$P(x)\approx P(x_0)+P'(x_0)(x-x_0)=0. \tag{2}$$

Решение $x_1$ этого уравнения и дает новое приближенное значение корня. Если оператор $P'(x_0)$ имеет обратный $[P'(x_0)]^{-1}\in(Y\to X)$, то, пользуясь им, выражение $x_1$ можно получить в явном виде. Действительно, применяя указанный оператор $[P'(x_0)]^{-1}$ к обеим частям равенства (2), найдем:

$$x_1=x_0-[P'(x_0)]^{-1}P(x_0). \tag{3}$$

Аналогичным образом выражаются последовательно одно через другое и дальнейшие приближения:

$$x_{n+1}=x_n-[P'(x_n)]^{-1}P(x_n). \tag{4}$$

Условия сходимости последовательности $x_n$ к точному решению уравнения (1) и одновременно достаточные условия для существования этого решения даются следующей теоремой.

**Теорема 1.** О существовании решения и о сходимости процесса Ньютона. *Пусть выполнены следующие условия:*

1) *для элемента $x_0$ — начального приближения, $P'(x_0)$ — оператор, отображающий $X$ в $Y$, имеет обратный $\Gamma_0=[P'(x_0)]^{-1}$, и известна оценка для его нормы:*

$$\|\Gamma_0\|\leqslant B_0; \tag{5}$$

2) *элемент $x_0$ удовлетворяет приближенно уравнению (1), причем известна оценка выражения $\Gamma_0 P(x_0)$:*

$$\|\Gamma_0 P(x_0)\|\leqslant\eta_0; \tag{6}$$

3) *вторая производная $P''(x)$ ограничена в интересующей нас области, определяемой неравенством (9), именно:*

$$\|P''(x)\|\leqslant K; \tag{7}$$

4) *постоянные* $B_0, \eta_0, K$ *удовлетворяют соотношению*

$$h_0 = B_0 \eta_0 K \leqslant \frac{1}{2}. \tag{8}$$

***Тогда уравнение*** (1) *имеет решение* $x^*$, *которое находится в области вблизи* $x_0$, *определяемой неравенством*

$$\| x^* - x_0 \| \leqslant N(h_0)\eta_0 = \frac{1 - \sqrt{1 - 2h_0}}{h_0}\eta_0, \tag{9}$$

*и последовательные приближения процесса Ньютона* $x_n$ *сходятся к нему, причем быстрота сходимости характеризуется оценкой:*

$$\| x_n - x^* \| \leqslant \frac{1}{2^{n-1}} (2h_0)^{2^n - 1} \eta_0. \tag{10}$$

Доказательство. Введем обозначение:

$$F_0(x) = x - \Gamma_0 P(x). \tag{11}$$

Пользуясь им, соотношение (3), связывающее $x_1$ и $x_0$, можем записать так:

$$x_1 = x_0 - [P'(x_0)]^{-1} P(x_0) = F_0(x_0). \tag{12}$$

Покажем, что когда мы переходим от значения $x_0$ к $x_1$, все условия 1—4 будут выполняться попрежнему. Прежде всего имеем:

$$\| x_1 - x_0 \| = \| \Gamma_0 P(x_0) \| \leqslant \eta_0. \tag{13}$$

Далее, применяя аналог формулы Лагранжа (§ 3 (7)) к $P'(x)$, получаем оценку для нормы следующего оператора, переводящего $X$ в $X$:

$$\| \Gamma_0 [P'(x_0) - P'(x_1)] \| \leqslant B_0 \sup \| P''(\bar{x}) \| \; \| x_1 - x_0 \| \leqslant B_0 K \eta_0 = h_0 < 1,$$
$$\bar{x} = x_1 + \theta(x_0 - x_1). \tag{14}$$

Отсюда на основании предложения С. Банаха [§ 1, (2)] следует, что существует обратный оператор

$$H = [I - \Gamma_0 (P'(x_0) - P'(x_1))]^{-1},$$

где $I$ — единичный оператор, отображающий $X$ в $X$, и что для его нормы справедлива оценка

$$\| H \| \leqslant \frac{1}{1 - h_0}. \tag{15}$$

Отсюда, полагая $\Gamma_1 = H\Gamma_0$ и пользуясь для операторов правилом $[AB]^{-1} = B^{-1} A^{-1}$, находим:

$$\Gamma_1 = H\Gamma_0 = \{I - \Gamma_0 [P'(x_0) - P'(x_1)]\}^{-1} [P'(x_0)]^{-1} =$$
$$= \{P'(x_0)(I - \Gamma_0 [P'(x_0) - P'(x_1)])\}^{-1} = [P'(x_1)]^{-1},$$

что и доказывает существование этого обратного оператора. При этом, на основании неравенства (15), получаем оценку для него:

$$\| \Gamma_1 \| = \| [P'(x_1)]^{-1} \| \leqslant \frac{B_0}{1-h_0} = B_1. \tag{16}$$

Условие 1 проверено.

Теперь, пользуясь тем, что

$$F_0'(x_0) = I - \Gamma_0 P'(x_0) = 0 \tag{17}$$

(на основании правил 2 и 3 § 3), а также используя (11) для $x = x_1$, имеем:

$$-\Gamma_0 P(x_1) = F_0(x_1) - F_0(x_0) - F_0'(x_0)(x_1 - x_0).$$

Тогда, применяя для $P = F_0$ и $\Delta x = x_1 - x_0$ аналог формулы Тэйлора (8) § 3, находим:

$$\| \Gamma_0 P(x_1) \| \leqslant \frac{1}{2} \sup \| F_0''(\bar{x}) \| \; \| x_1 - x_0 \|^2 =$$

$$= \frac{1}{2} \sup \| \Gamma_0 P''(\bar{x}) \| \; \| x_1 - x_0 \|^2 \leqslant \frac{1}{2} B_0 K \eta_0^2 = \frac{1}{2} h_0 \eta_0, \tag{18}$$

$$(\bar{x} = x_0 + \theta(x_1 - x_0)).$$

Наконец, пользуясь (15) и (18), имеем:

$$\| \Gamma_1 P(x_1) \| = \| H \Gamma_0 P(x_1) \| \leqslant \| H \| \; \| \Gamma_0 P(x_1) \| \leqslant \frac{1}{2} \frac{h_0 \eta_0}{1-h_0} = \eta_1. \tag{19}$$

Условие 3 будет выполнено для точки $x_1$, так как соответствующая ей сфера, как мы убедимся ниже, не выходит за пределы сферы, определяемой неравенством (9).

Наконец, непосредственно проверяем и условие 4, пользуясь (16) и (19). Действительно,

$$h_1 = B_1 \eta_1 K = \frac{B_0}{1-h_0} \frac{1}{2} \frac{h_0 \eta_0}{1-h_0} K = \frac{h_0^2}{2(1-h_0)^2} \leqslant 2h_0^2 \leqslant \frac{1}{2}, \tag{20}$$

так как $h_0 \leqslant \frac{1}{2}$.

Итак, для $x = x_1$ выполнены условия вида условий 1—4 с заменой чисел $B_0, \eta_0$ и $h_0$ на $B_1$, $\eta_1$ и $h_1$. Это дает возможность продолжать последовательное определение элементов $x_n$ и связанных с ними чисел $B_n$, $\eta_n$ и $h_n$, которые будут связаны друг с другом формулами, аналогичными (13), (16), (19) и (20).

$$\| x_n - x_{n+1} \| \leqslant \eta_n; \tag{13a}$$

$$B_n = \frac{B_{n-1}}{1-h_{n-1}}; \tag{16a}$$

$$\eta_n = \frac{1}{2} \frac{h_{n-1} \eta_{n-1}}{1-h_{n-1}}; \tag{19a}$$

$$h_n = \frac{1}{2} \frac{h_{n-1}^2}{(1-h_{n-1})^2}. \tag{20a}$$

Далее для них будем иметь следующие оценки:

$$h_2 \leqslant 2h_1^2 \leqslant 8h_0^4; \ \ldots; \quad h_n \leqslant \frac{1}{2}(2h_0)^{2^n}; \tag{21}$$

$$\eta_n = \frac{1}{2}\frac{h_{n-1}}{1-h_{n-1}}\eta_{n-1} \leqslant h_{n-1}\eta_{n-1} \leqslant \ldots \leqslant h_{n-1}h_{n-2}\ldots h_0\eta_0 \leqslant$$
$$\leqslant \frac{1}{2^n}(2h_0)^{2^{n-1}}(2h_0)^{2^{n-2}}\ldots(2h_0)\eta_0 \leqslant \frac{1}{2^n}(2h_0)^{2^n-1}\eta_0. \tag{22}$$

Наконец, отметим следующее тождество

$$\eta_n N(h_n) - \eta_{n+1}N(h_{n+1}) = \eta_n, \tag{23}$$

которое проверяется непосредственным вычислением:

$$\eta_{n+1}N(h_{n+1}) = \eta_{n+1}\frac{1-\sqrt{1-2h_{n+1}}}{h_{n+1}} = \frac{1}{2}\frac{h_n\eta_n}{1-h_n}\frac{1-\sqrt{1-\frac{h_n^2}{(1-h_n)^2}}}{\frac{1}{2}\frac{h_n^2}{(1-h_n)^2}} =$$
$$= \eta_n\frac{1-h_n-\sqrt{1-2h_n}}{h_n} = \eta_n N(h_n) - \eta_n.$$

Используя (13a), (23) и (22), находим:

$$\| x_{n+p} - x_n \| \leqslant \| x_{n+1} - x_n \| + \| x_{n+2} - x_{n+1} \| + \ldots + \| x_{n+p} -$$
$$- x_{n+p-1} \| \leqslant \eta_n + \eta_{n+1} + \ldots + \eta_{n+p-1} = N(h_n)\eta_n -$$
$$- N(h_{n+p})\eta_{n+p} \leqslant N(h_n)\eta_n \leqslant 2\eta_n \leqslant \frac{1}{2^{n-1}}(2h_0)^{2^n-1}\eta_0. \tag{24}$$

Это доказывает существование предела

$$\lim_{n\to\infty} x_n = x^*. \tag{25}$$

Отсюда также, беря $p\to\infty$, устанавливаем справедливость неравенства (10) и, беря затем $n=0$, — и справедливость (9).

То обстоятельство, что $x^*$ — корень уравнения (1), получаем, переходя к пределу в соотношении,

$$P'(x_n)(x_{n+1} - x_n) + P(x_n) = 0.$$

Действительно, исходя из того, что $\| x_{n+1} - x_n \| \to 0$, а $\| P'(x_n) \|$ ограничена, так как

$$\| P'(x_n) \| \leqslant \| P'(x_0) \| + \| P'(x_n) - P'(x_0) \| \leqslant \| P'(x_0) \| + K \| x_n - x_0 \|$$
$$\leqslant \| P'(x_0) \| + KN(h_0)\eta_0,$$

получаем, что $\| P(x_n) \| \to 0$, и так как $x_n \to x^*$, то по непрерывности $P(x)$

$$P(x^*) = 0.$$

Отметим еще, что использованное в ходе доказательства утверждение, что сфера

$$\| x - x_n \| \leqslant N(h_n)\eta_n \tag{26}$$

не выходит за пределы сферы (9), теперь совершенно очевидно. Действительно, если $x$ находится в сфере (26), то

$$\| x - x_0 \| \leqslant \| x_n - x_0 \| + \| x - x_n \| \leqslant [N(h_0)\eta_0 - N(h_n)\eta_n] + \\ + N(h_n)\eta_n = N(h_0)\eta_0,$$

т. е. $x$ попадает и в сферу (9). Итак, теорема полностью доказана.

**Теорема 2.** Об единственности решения и о сходимости процесса при других начальных значениях. *Пусть выполнены условия* 1—4 *предыдущей теоремы, с той лишь разницей, что неравенство* (7)

$$\| P''(x) \| \leqslant K$$

*будем теперь предполагать выполненным в сфере, определяемой неравенством* (27). *Тогда решение уравнения* (1) *в области*

$$\| x - x_0 \| < L(h_0)\eta_0 = \frac{1+\sqrt{1-2h_0}}{h_0}\eta_0 \tag{27}$$

*единственно* $\left(\textit{в случае, когда } h_0 = \frac{1}{2}, \textit{ знак } < \textit{ в этом неравенстве должен быть заменен знаком } \leqslant\right)$. *К этому решению сходится процесс последовательных приближений, если его начать с любого значения* $x$, *удовлетворяющего неравенству*

$$\| x - x_0 \| \leqslant \Delta = \frac{1-2h_0}{4h_0}\eta_0. \tag{28}$$

Доказательство. Для установления единственности предположим, что имеется некоторое решение $\breve{x}$ уравнения $P(x) = 0$, удовлетворяющее условию

$$\| \breve{x} - x_0 \| \leqslant \theta L(h_0)\eta_0 \qquad (0 \leqslant \theta < 1) \tag{29}$$

$\left(\text{мы ограничиваемся сначала случаем, когда } h_0 < \frac{1}{2}\right)$. Так как для данного значения $P(\breve{x}) = 0$, то [см. (11)]:

$$F_0(\breve{x}) = \breve{x}.$$

Имеем, как и выше, при доказательстве теоремы 1 [ср. (17) и (18)]:

$$\| \breve{x} - x_1 \| = \| F_0(\breve{x}) - F_0(x_0) \| = \| F_0(\breve{x}) - F_0(x_0) - F_0'(x_0)(\breve{x} - x_0) \| \leqslant \\ \leqslant \frac{1}{2} B_0 K \| \breve{x} - x_0 \|^2 \leqslant \frac{1}{2} B_0 K \theta^2 L^2(h_0)\eta_0^2 = \theta^2 L(h_1)\eta_1. \tag{30}$$

Справедливость последнего равенства проверяем непосредственным вычислением:

$$L(h_1)\eta_1 = \frac{1+\sqrt{1-2h_1}}{h_1}\eta_1 = \frac{1+\sqrt{1-\frac{h_0^2}{(1-h_0)^2}}}{\frac{1}{2}\frac{h_0^2}{(1-h_0)^2}}\frac{1}{2}\frac{h_0}{1-h_0}\eta_0 =$$

$$= \frac{1-h_0+\sqrt{1-2h_0}}{h_0}\eta_0 = (L(h_0)-1)\eta_0;$$

$$\frac{1}{2}B_0KL^2(h_0)\eta_0^2 = \frac{1}{2}h_0\left[\frac{2-2h_0+2\sqrt{1-2h_0}}{h_0^2}\right]\eta_0 = (L(h_0)-1)\eta_0.$$

Установленное неравенство (30) показывает, что для элементов $\bar{x}$ и $x_1$ выполнено условие, аналогичное (29), лишь с заменой $\theta$ на $\theta^2$. Поэтому можем продолжить оценки, беря вместо $x_0$ элемент $x_1$, затем $x_2$ и т. д. В результате найдем, что

$$\| \bar{x} - x_n \| \leqslant \theta^{2^n} L(h_n)\eta_n \leqslant \theta^{2^n}\frac{2}{B_nK}, \tag{31}$$

последнее неравенство следует из того, что $L(h_n) < \frac{2}{h_n}$.

Полученная оценка (31) показывает, так как $B_n \geqslant B_0$ и $\theta < 1$, что $\| \bar{x} - x_n \| \to 0$. Поэтому

$$\bar{x} = \lim_{n\to\infty} x_n = x^*, \tag{32}$$

что и доказывает единственность.

Случай, когда $h_0 = \frac{1}{2}$, должен быть рассмотрен особо. Здесь, благодаря тому, что в условии (27) допускается знак $=$, в соотношениях (29) и (31) приходится принять $\theta = 1$, но в случае $h_0 = \frac{1}{2}$, как это ясно из (16), имеем: $B_1 = 2B_0$ и вообще $B_n = 2^n B_0$, а потому попрежнему, на основании (31), заключаем, что $\| x_n - \bar{x} \| \to 0$, откуда следует единственность.

Предположим теперь, что в качестве начального взято некоторое значение $x'$, лежащее в сфере (28):

$$\| x' - x_0 \| \leqslant \Delta = \frac{1-2h_0}{4h_0}\eta_0, \tag{33}$$

и докажем, что ряд последовательных приближений, начинающихся с этого значения, также сходится к $x^*$.

Покажем прежде всего, что для $x = x'$ будут соблюдены условия, аналогичные условиям 1—4 теоремы 1 для $x = x_0$. Имеем, повторяя частично рассуждения теоремы 1, что

$$\| \Gamma_0 [P'(x_0) - P'(x')] \| \leqslant B_0 K \Delta < 1. \tag{34}$$

Отсюда [ср. (15)]

$$\|H'\| = \|[I - \Gamma_0[P'(x_0) - P'(x')]]^{-1}\| \leqslant \frac{1}{1 - B_0K\Delta}, \tag{35}$$

$$\|[P'(x')]^{-1}\| = \|\Gamma'\| = \|H'\Gamma_0\| \leqslant \frac{B_0}{1 - B_0K\Delta} = B'. \tag{36}$$

Таким образом, условие 1 проверено.

Далее имеем [ср. (6) и (34)]:

$$\|\Gamma_0 P(x')\| \leqslant \|\Gamma_0[P(x') - P(x_0) - P'(x_0)(x' - x_0)]\| + \|\Gamma_0[P(x_0) + P'(x_0)(x' - x_0)]\| \leqslant \frac{1}{2}B_0K\Delta^2 + \eta_0 + \Delta. \tag{37}$$

Тогда

$$\|\Gamma' P(x')\| = \|H'\Gamma_0 P(x')\| \leqslant \frac{1}{1 - B_0K\Delta}\left(\frac{1}{2}B_0K\Delta^2 + \eta_0 + \Delta\right) = \eta', \tag{38}$$

т. е. проверено и условие 2. В выполнении условия 3 мы убедимся несколько дальше. Проверим условие 4. Имеем

$$h' = B'K\eta' = \frac{\frac{1}{2}B_0K\Delta^2 + \eta_0 + \Delta}{(1 - B_0K\Delta)^2}B_0K = \frac{h_0 + B_0K\Delta + \frac{1}{2}B_0^2K^2\Delta^2}{(1 - B_0K\Delta)^2} =$$

$$= \frac{1}{2} + \frac{h_0 - \frac{1}{2} + 2B_0K\Delta}{(1 - B_0K\Delta)^2} = \frac{1}{2} + \frac{h_0 - \frac{1}{2} + \left(\frac{1}{2} - h_0\right)}{(1 - B_0K\Delta)^2} = \frac{1}{2}.$$

Таким образом, условие 4 также выполнено. Это позволяет заключить, что ряд последовательных приближений $x_n'$, начинающихся с $x_0' = x'$, сходится к некоторому решению $x^{*\prime}$, лежащему в сфере

$$\|x - x'\| \leqslant N(h')\eta' \leqslant \frac{\eta'}{h'} = \frac{1}{B'K} = \frac{1 - B_0K\Delta}{B_0K}. \tag{39}$$

Но эта сфера лежит внутри сферы (27). Действительно,

$$\|x - x_0\| \leqslant \|x' - x_0\| + \|x' - x\| \leqslant \Delta + \frac{1 - B_0K\Delta}{B_0K} =$$

$$= \frac{1}{B_0K} = \frac{\eta_0}{h_0} < L(h_0)\eta_0^{*}.$$

Следовательно, решение $x^{*\prime}$ также лежит в сфере (27) и по доказанной уже единственности решение должно совпадать с $x^*$. Итак,

$$\lim_{n \to \infty} x_n' = x^*.$$

Попутно из того факта, что сфера (39) лежит в (27), в которой неравенство $\|P''(x)\| \leqslant K$ выполнено, ясно выполнение условия 3 для точки $x = x'$, проверка которого не была произведена выше.

---

* Если $h_0 = \frac{1}{2}$, то здесь будет знак $=$, но тогда и в (27) допускается знак равенства.

Итак, доказана и вторая часть теоремы.

Сделаем несколько замечаний, относящихся к обеим доказанным теоремам.

Замечание 1. Отметим, что условие 2 могло бы быть заменено на более просто формулируемое условие 2', состоящее в выполнении неравенства:

$$\| P(x_0) \| \leqslant \eta_0', \tag{40}$$

которое дает непосредственно характеристику того, в какой степени удовлетворено уравнение (1) начальным значением. Выполнение условия 2' влечет за собой выполнение условия 2 для $\eta_0 = B_0 \eta_0'$. Действительно,

$$\| \Gamma_0 P(x_0) \| \leqslant \| \Gamma_0 \| \ \| P(x_0) \| \leqslant B_0 \eta_0' = \eta_0.$$

В соответствии с этим, если вместо условия 2 выполнено условие 2', то число $h_0$ и неравенство (8) запишутся так:

$$h_0 = B_0 K \eta_0 = B_0^2 K \eta_0' \leqslant \frac{1}{2}. \tag{41}$$

Попутно отметим также, что и само условие 2 допускает иную запись, иногда более удобную в применениях. Именно, пользуясь (13), ему можно придать вид:

$$\| \Gamma_0 P(x_0) \| = \| x_1 - x_0 \| \leqslant \eta_0 \tag{42}$$

и таким образом оно с удобством проверяется после нахождения первого приближения.

Замечание 2. Область расположения решения $x^*$ по данным в условиях теоремы дается неравенством (9). Однако сравнительно легко получить более точное неравенство для $x^*$. Именно, зная $\Gamma_0$, можем определить $x_1 = x_0 - \Gamma_0 P(x_0)$ и $P(x_1)$, а тогда, применяя теорему 1 к $x_1$ вместо $x_0$ [ср. (9) и (6)] и пользуясь затем оценкой для $\| \Gamma_1 \|$ (16), имеем:

$$\| x_1 - x^* \| \leqslant N(h_1) \| \Gamma_1 P(x_1) \| \leqslant N(h_1) \| \Gamma_1 \| \ \| P(x_1) \| \leqslant$$

$$\leqslant N(h_1) \frac{B_0}{1-h_0} \| P(x_1) \| = N(h_1) \frac{1}{2} \frac{h_0 \eta_0}{1-h_0} \frac{2B_0}{h_0 \eta_0} \| P(x_1) \| =$$

$$= N(h_1) \eta_1 \frac{2B_0}{h_0 \eta_0} \| P(x_1) \| = [\eta_0 N(h_0) - \eta_0] \frac{2B_0}{h_0 \eta_0} \| P(x_1) \| =$$

$$= 2 \frac{1 - h_0 - \sqrt{1-2h_0}}{h_0^2} B_0 \| P(x_1) \| = \frac{2}{1 - h_0 + \sqrt{1-2h_0}} B_0 \| P(x_1) \| .$$

Полученной оценкой

$$\| x_1 - x^* \| \leqslant \frac{2}{1 - h_0 + \sqrt{1-2h_0}} B_0 \| P(x_1) \| \tag{43}$$

мы воспользуемся в дальнейшем. Отметим, что при $h_0$ малом она имеет порядок $B_0 \| P(x_1) \|$.

Замечание 3. Отметим, что полученные в теоремах 1 и 2 оценки (8), (9) и (27) не могут быть улучшены даже для случая вещественного уравнения второй степени, как показывает следующий пример:

$$P(x)=\frac{1}{2}x^2-x+h=0 \quad (h>0), \quad x_0=0.$$

Действительно, в данном случае

$$P'(x_0)=-1, \quad \|\Gamma_0\|=\frac{1}{|P'(x_0)|}=1; \quad B_0=1;$$

$$\|\Gamma_0 P(x_0)\|=1\cdot h=h; \quad \eta_0=h; \quad \|P''(x)\|=1=K;$$

$$h_0=B_0K\eta_0=h.$$

При этом корни уравнения будут:

$$x_{1,2}^*=1\pm\sqrt{1-2h}=N(h_0)\eta_0 \text{ или } L(h_0)\eta_0$$

и, следовательно, они будут вещественны только при условии, когда $h_0=h\leqslant\frac{1}{2}$. При этом первый корень лежит на расстоянии $N(h_0)\eta_0$, а второй — на расстоянии $L(h_0)\eta_0$. Это показывает, что и границы области расположения корня и области единственности определены точно.

Замечание 4. Отметим, что для случая вещественных и комплексных уравнений, теоремы, аналогичные теореме 1, были получены A. Ostrowski [1, 2]. Эти теоремы следуют из доказанной теоремы 1 для случая вещественного уравнения совершенно непосредственно, а для случая комплексных уравнений, если учесть смысл операции и значение нормы для этого случая, в соответствии с рассмотренными выше примерами (пример 3 § 1 и пример 2 § 3), а именно:

$$\|P(x_0)\|=|P(x_0)|; \quad \|\Gamma_0\|=\|[P'(x_0)]^{-1}\|=\frac{1}{|P'(x_0)|};$$

$$K=\max\|P''(x)\|=\max|P''(x)|;$$

$$\eta_0'=|P(x_0)|; \quad \eta_0=\frac{|P(x_0)|}{|P'(x_0)|}.$$

В соответствии с этим, теоремы 1 и 2 для случая вещественных и комплексных уравнений формулируются так:

Если выполнено условие

$$h_0=\frac{|P(x_0)|K}{|P'(x_0)|^2}\leqslant\frac{1}{2}, \quad K=\max|P''(x)|,$$

то процесс Ньютона сходится к решению уравнения $P(x)=0$, которое существует и лежит в области

$$|x-x_0|\leqslant N(h_0)\eta_0=\frac{(1-\sqrt{1-2h_0})|P'(x_0)|}{K}$$

и единственно в области

$$|x - x_0| < L(h_0)\eta_0 = \frac{(1+\sqrt{1-2h_0})\,|P'(x_0)|}{K}.$$

Отметим, что рассматривалась также сходимость метода Ньютона при других условиях [13,3]*, но в этих теоремах ставилась условием ограниченность величины $\frac{1}{|P'(x)|}$ в целом промежутке, а не только в точке $x_0$, что менее удобно для проверки.

Замечание 5. Отметим, что, в частности, если значение $x_0 = x^*$ есть корень уравнения, то в этом случае $h_0 = 0$; $\eta_0 = 0$, но $L(h_0)\eta_0$ следует заменить на $\frac{2}{B_0K}$. В соответствии с этим можно заключить, что $x_0$ — единственное решение уравнения в области

$$\|x - x^*\| < \frac{2}{B_0K} \tag{44}$$

и что к нему сходится метод последовательных приближений, если начальное значение взято из области

$$\|x - x^*\| \leqslant \Delta = \frac{1}{4B_0K}. \tag{45}$$

Замечание 6. В случае, если оператор $P$ вполне непрерывен, то существование решения в сфере (9) при условиях 1—4 могло бы быть получено также на основании теоремы Шаудера о фикспунктах [14]. Именно, уравнение $P(x) = 0$ заменяется эквивалентным ему $x = F_0(x)$, а, как можно проверить, оператор $F_0(x)$ переводит сферу (9) в себя. Необходимо сказать, однако, что эта проверка потребовала бы проведения значительной части рассуждений, использованных в теореме, и в то же время этот путь не дает ряда других важных результатов — сходимость процесса Ньютона и теорему единственности. Принцип Caccioppoli-Banach [14] позволил бы получить существование решения лишь при гораздо более грубых условиях, примерно $h_0 = 0.1$.

Замечание 7. Отметим, наконец, некоторое принципиальное значение доказанных теорем. Именно, они дают не только установление сходимости определенного алгорифма, но и представляют теоремы о существовании, единственности и области расположения решения. При этом существенным условием для возможности применения этих теорем является наличие в нашем распоряжении начального значения $x_0$, представляющего грубое приближение к решению. При этом такое начальное значение $x_0$, для которого выполнены условия теоремы 1 и 2, всегда должно быть, если решение существует и оно простое, т. е. для него $\|[P'(x^*)]^{-1}\| < +\infty$. Действительно, в этом

* См. также работу И. П. Мысовских в данном сборнике, где эта теорема перенесена на общие функциональные уравнения.

случае всем условиям будет удовлетворять каждая точка $x_0$, достаточно близкая к $x^*$.

Такое начальное значение $x_0$ фактически может быть получено в результате грубого численного или приближенного решения задач. В частности, в механике и в других областях прикладной математики такое приближенное решение часто получается в результате рассмотрения проблемы в упрощенных условиях. После же того как приближенное решение найдено и окажется, что для него выполнены условия 1—4, на основании данных теорем можно заключить о наличии точного решения, единственности его и области расположения, т. е. произвести довольно полное теоретическое исследование проблемы.

Таким образом, данные теоремы показывают, что приближенное решение задачи полезно не только для получения некоторых численных результатов, но и для теоретического анализа вопроса.

## § 5. Некоторые модификации процесса Ньютона

Применяя процесс Ньютона, приходится при каждом последовательном приближении находить обратный оператор $[P'(x_n)]^{-1}$ или, во всяком случае, решать уравнение:

$$P'(x_n)(x_n - x_{n+1}) = P(x_n), \tag{1}$$

что может оказаться затруднительным. Поэтому при практическом применении процесса Ньютона рекомендуется его видоизменить, заменяя каждый раз $[P'(x_n)]^{-1}$ на $[P'(x_0)]^{-1} = \Gamma_0$ или, что то же самое, определяя последовательные приближения из уравнения

$$P'(x_0)(x_n' - x'_{n+1}) = P(x_n'), \tag{2}$$

откуда

$$x'_{n+1} = x_n' - [P'(x_0)]^{-1} P(x_n'). \tag{3}$$

Очевидно, что первый шаг в обоих процессах совпадает, т. е. $x_1' = x_1$. В следующей теореме мы устанавливаем сходимость такого модифицированного процесса.

**Теорема 3.** *Если выполнены условия теоремы 1 и при этом $h_0 < \frac{1}{2}$, то модифицированный процесс Ньютона, при котором последовательные приближения определяются по* (3), *оказывается сходящимся к решению*

$$\lim_{n \to \infty} x_n' = x^*. \tag{4}$$

*При этом сходимость имеет порядок геометрической прогрессии, точнее*

$$\| x_n' - x^* \| < q^{n-1} \| x_1 - x^* \|; \quad q = 1 - \sqrt{1 - 2h_0} < 1. \tag{5}$$

Доказательство. Для доказательства этой теоремы установим предварительное предложение: если для элемента $x$ выполнены условия

$$\| x - x^* \| \leqslant \| x_1 - x^* \| , \tag{6}$$

$$\| x - x_0 \| \leqslant N(h_0)\,\eta_0, \tag{7}$$

то для элемента $x' = F_0(x)$ будет:

$$\| x' - x^* \| \leqslant q \, \| x - x^* \| , \tag{6a}$$

$$\| x' - x_0 \| \leqslant N(h_0)\,\eta_0. \tag{7a}$$

В самом деле, имеем:

$$\| x' - x^* \| = \| F_0(x) - F_0(x^*) \| \leqslant \sup \| F_0'(\bar{x}) \| \; \| x - x^* \| , \tag{8}$$

$$\bar{x} = x + \theta\,(x^* - x).$$

Далее [ср. § 4 (13), (7), (9)]:

$$\| F_0'(\bar{x}) \| = \| F_0'(\bar{x}) - F_0'(x_0) \| \leqslant \sup \| F_0''(\tilde{x}) \| \; \| \bar{x} - x_0 \| \leqslant$$

$$\leqslant B_0 K \max ( \| x - x_0 \| , \| x_0 - x^* \| ) \leqslant B_0 K N(h_0)\,\eta_0 =$$

$$= 1 - \sqrt{1 - 2h_0} = q, \quad (\tilde{x} = x_0 + \theta_1 (\bar{x} - x_0)).$$

Это неравенство вместе с предыдущим (8) дает (6a). Остается установить (7a). В самом деле,

$$\| x' - x_0 \| = \| F_0(x) - x_0 \| = \| x - \Gamma_0 P(x) - x_0 \| =$$

$$= \| \Gamma_0 [P(x) - P'(x_0)(x - x_0)] \| \leqslant \| \Gamma_0 P(x_0) \| +$$

$$+ \| \Gamma_0 [P(x) - P(x_0) - P'(x_0)(x - x_0)] \| \leqslant \eta_0 +$$

$$+ \frac{1}{2} B_0 K \, \| x - x_0 \|^2 \leqslant \eta_0 + \frac{1}{2} B_0 K \, [N(h_0)\,\eta_0]^2 =$$

$$= \eta_0 \left[ 1 + \frac{1}{2} B_0 K \left( \frac{1 - \sqrt{1 - 2h_0}}{h_0} \right)^2 \eta_0 \right] = \eta_0 N(h_0). \tag{9}$$

Итак, предварительное предложение доказано. Определяя последовательные приближения, начиная с $x_0$, найдем: $x_1' = F_0(x_0) = x_1$. Для этого приближения условия (6) и (7), очевидно, выполнены. Тогда для следующего приближения $x_2' = F_0(x_1')$, согласно предварительному предложению, применяя его для $x = x_1'$, найдем:

$$\| x_2' - x^* \| \leqslant q \, \| x_1 - x^* \| .$$

Продолжая таким же образом, придем к (5), откуда следует также и (4).

Д. А. Граве [15] предложил и исследовал для случая уравнений с вещественным переменным несколько иную модификацию процесса Ньютона, именно когда последовательные приближения при решении уравнения $f(x) = 0$ определяются по формуле:

$$x_{n+1} = x_n - Q \frac{f(x_n)}{f'(x_n)} , \tag{10}$$

где $Q$ — некоторое число $0 < Q < 1$.

Для абстрактных уравнений аналогичная формула, связывающая последовательные приближения, принимает вид:

$$x_{n+1} = x_n - Q\,[P'(x_n)]^{-1}\,P(x_n)^*. \tag{11}$$

Сходимость этого процесса может быть доказана при том же условии, что и в теореме 1, именно при $h_0 \leqslant \frac{1}{2}$.

**Теорема 4.** *При условиях теоремы 1 обобщенный процесс Д. А. Граве* [см. (11)] *сводится к решению уравнения.*

Доказательство. Введем функцию

$$F_Q(x) = x - Q\,\Gamma_0\,P(x).$$

Тогда приближение $x_1$ через $x_0$ можно выразить так:

$$x_1 = x_0 - Q\Gamma_0\,P(x_0) = F_Q(x_0).$$

Прежде всего имеем:

$$\| x_1 - x_0 \| \leqslant Q\,\| \Gamma_0 P(x_0) \| \leqslant Q\eta_0. \tag{12}$$

Поступая так же, как при доказательстве теоремы 1, находим:

$$\| F_Q(x_1) - F_Q(x_0) \| \leqslant \| F'_Q(x_0)(x_1 - x_0) \| +$$

$$+ \frac{1}{2} \sup \| F''_Q(\bar{x}) \| \; \| x_1 - x_0 \|^2 \leqslant \| I - QI \| \; \| x_1 - x_0 \| +$$

$$+ \frac{1}{2} QB_0K\,\| x_1 - x_0 \|^2 = (1 - Q)\,Q\eta_0 + \frac{1}{2}\,QB_0K\,(Q\eta_0)^2,$$

$$(\bar{x} = x_0 + \theta\,(x - x_0));$$

$$\| \Gamma_0 P(x_1) \| = \frac{1}{Q}\,\| F_Q(x_1) - x_1 \| = \frac{1}{Q}\,\| F_Q(x_1) - F_Q(x_0) \| \leqslant$$

$$\leqslant (1 - Q)\,\eta_0 + \frac{1}{2}\,Q^2 h_0 \eta_0.$$

Далее,

$$\| \Gamma_0 [P'(x_0) - P'(x_1)] \| < K\,(Q\eta_0)\,B_0 = Qh_0 < 1;$$

$$\Gamma_1 = [P'(x_1)]^{-1} = [I - \Gamma_0\,(P'(x_0) - P'(x_1))]^{-1}\,\Gamma_0 = H\Gamma_0;$$

$$\| H \| \leqslant \frac{1}{1 - Qh_0}\,; \qquad \| \Gamma_1 \| \leqslant \frac{B_0}{1 - Qh_0} = B_1.$$

---

* Здесь $x_n$ и в дальнейшем $B_1$, $\Gamma_1$ и т. д. означают не то же самое, что в теореме 1.

Наконец,

$$\left.\begin{aligned} &\| \Gamma_1 P(x_1) \| = \| H\Gamma_0 P(x_1) \| \leqslant \frac{1}{1-Qh_0}\left[(1-Q)+\frac{1}{2}Q^2h_0\right]\eta_0 = \eta_1,\\ &\quad h_1 = B_1K\eta_1 = \frac{B_0K}{(1-Qh_0)^2}\left[(1-Q)+\frac{1}{2}Q^2h_0\right]\eta_0 =\\ &\quad = h_0\frac{1-Q+\frac{1}{2}Q^2h_0}{(1-Qh_0)^2} = \frac{1}{2}-\frac{\frac{1}{2}-h_0}{(1-Qh_0)^2}. \end{aligned}\right\} \quad (13)$$

Отсюда видно, что, так как $h_0 \leqslant \frac{1}{2}$, то и $h_1 \leqslant \frac{1}{2}$, т. е. процесс последовательных приближений может продолжаться.

Для установления сходимости и быстроты ее мы ограничимся рассмотрением наименее выгодного случая, когда $h_0 = \frac{1}{2}$. Полагая в (13) $h_0 = \frac{1}{2}$, имеем:

$$\eta_1 = \frac{1}{1-\frac{1}{2}Q}\left[1-Q+\frac{1}{4}Q^2\right]\eta_0 = \left(1-\frac{Q}{2}\right)\eta_0.$$

Следовательно, продолжая процесс, найдем, что

$$\eta_n \leqslant \left(1-\frac{Q}{2}\right)^n \eta_0,$$

и так как [ср. (12)]

$$\| x_{n+1} - x_n \| \leqslant Q\eta_n,$$

то ясно, что существует

$$\lim_{n\to\infty} x_n = x^*$$

и что

$$\| x_n - x^* \| \leqslant Q(\eta_n + \eta_{n+1} + \cdots) \leqslant 2\left(1-\frac{Q}{2}\right)^n \eta_0.$$

Таким образом, даже и в этом, наихудшем случае, быстрота сходимости имеет порядок геометрической прогрессии.

## § 6. Метод Ньютона для систем алгебраических уравнений

Метод Ньютона для решения систем $m$ алгебраических уравнений с $m$ неизвестными

$$\eta^{(i)} = f_i(\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(m)}) = 0 \quad (i = 1, 2, \ldots, m) \tag{1}$$

является естественным обобщением того же метода для одного уравнения. Последовательные приближения для корня — первое приближение $(\xi_1^{(1)}, \xi_1^{(2)}, \ldots, \xi_1^{(m)})$ по нулевому $(\xi_0^{(1)}, \xi_0^{(2)}, \ldots, \xi_0^{(m)})$ — определяются из системы уравнений

$$\left(\frac{\partial f_i}{\partial \xi^{(1)}}\right)_0 (\xi_1^{(1)} - \xi_0^{(1)}) + \ldots + \left(\frac{\partial f_i}{\partial \xi^{(m)}}\right)_0 (\xi_1^{(m)} - \xi_0^{(m)}) + \\ + f_i(\xi_0^{(1)}, \xi_0^{(2)}, \ldots, \xi_0^{(m)}) = 0 \\ (i = 1, 2, \ldots, m).$$

В таком виде этот метод указан у Рунге. Некоторая попытка дать условия сходимости процесса дана Willers [4] для двух уравнений с использованием производных третьего порядка. Для общего случая $m$ уравнений некоторые достаточные условия, использующие только производные первого и второго порядков, даны Н. П. Стениным [5], см. также [19]. Для случая двух уравнений с двумя неизвестными наиболее точные условия были даны A. Ostrowski [6]. Эта теорема рассмотрена в следствии 1 к теореме 5.

Здесь мы рассмотрим применение к системе алгебраических уравнений общих теорем о сходимости процесса Ньютона.

Данную систему можно рассматривать как одно уравнение

$$y = P(x) = 0, \tag{2}$$

где $P$ — оператор, переводящий $m$-мерное пространство в $m$-мерное. В частности, принимая метрику примера 1 § 1, получаем следующую теорему.

**Теорема 5.** *Если выполнены условия*:

1) $$|f_i(\xi_0^{(1)}, \xi_0^{(2)}, \ldots, \xi_0^{(m)})| \leqslant \bar{\eta} \quad (i = 1, 2, \ldots, m); \tag{3}$$

2) *матрица* $\left\| \left( \frac{\partial f_i}{\partial \xi^{(k)}} \right)_0 \right\|$ *имеет определитель* $\Delta$, *отличный от нуля, и если через* $A_{ik}$ *обозначить алгебраические дополнения его элементов, то выполнено условие:*

$$\max_i \frac{1}{|\Delta|} \sum_{k=1}^{m} |A_{ik}| \leqslant B; \tag{4}$$

3) $$\left| \frac{\partial^2 f_i}{\partial \xi^{(j)} \partial \xi^{(k)}} \right| \leqslant L \tag{5}$$

*в интересующей нас области*;

4) $$h = B^2 \bar{\eta} L m^2 \leqslant \frac{1}{2}, \tag{6}$$

*то данная алгебраическая система имеет решение, которое может быть получено процессом Ньютона.*

Доказательство. Действительно, проверим, что поставленные здесь условия обеспечивают выполнение условий 1—4 теоремы 1 (по поводу условия 1 см. замечание 1).

Для данного случая оператор $P'(x_0)$ определяется матрицей $\left\| \left( \frac{\partial f_i}{\partial \xi^{(k)}} \right)_0 \right\|$ или преобразованием [ср. § 3 (10)]:

$$\Delta\eta^{(1)} = \left( \frac{\partial f_1}{\partial \xi^{(1)}} \right)_0 \Delta\xi^{(1)} + \ldots + \left( \frac{\partial f_1}{\partial \xi^{(m)}} \right)_0 \Delta\xi^{(m)},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\Delta\eta^{(m)} = \left( \frac{\partial f_m}{\partial \xi^{(1)}} \right)_0 \Delta\xi^{(1)} + \ldots + \left( \frac{\partial f_m}{\partial \xi^{(m)}} \right)_0 \Delta\xi^{(m)}.$$

Обратная операция $\Gamma_0 = [P'(x)]^{-1}$ определяется обратной матрицей $\left\| \frac{A_{ik}}{\Delta} \right\|_{i, k=1,2,\ldots, m}$, а норма ее, согласно § 1 (4), может быть оценена так:

$$\| \Gamma_0 \| \leqslant \max_i \sum_k \frac{|A_{ik}|}{|\Delta|} \leqslant B,$$

т. е. выполнено условие 1.

Далее имеем:

$$\| P(x_0) \| = \max_i | f_i(\xi_0^{(1)}, \xi_0^{(2)}, \ldots, \xi_0^{(m)}) | \leqslant \bar{\eta},$$

так что выполнено условие 2′ [§ 4, (40)].

Наконец, для $\| P''(x) \|$, согласно § 3 (12), § 2 (4) имеем оценку

$$\| P''(x) \| \leqslant \max_k \sum_{i, j=1}^{m} \left| \frac{\partial f_k}{\partial \xi^{(i)} \partial \xi^{(j)}} \right| \leqslant m^2 L (= K),$$

т. е. выполнено условие 3.

Наконец, сопоставляя полученные оценки, мы убедимся, что выполнено и условие 4. Поэтому на основании теоремы 1 заключаем о сходимости процесса Ньютона. Также могут быть сформулированы для данного случая остальные заключения теорем 1—4.

Следствие 1. Рассмотрим случай системы двух уравнений с двумя неизвестными. Здесь формулировка условий может быть несколько упрощена, именно: если через $l$ обозначить максимум модулей $\left| \frac{\partial f_i}{\partial f^{(k)}} \right| \leqslant l$, то, учитывая, что в данном случае определитель $\Delta$ второго порядка и миноры его это его элементы, имеем: $|A_{ik}| \leqslant l$ и, следовательно, можно принять:

$$\max_i \sum_{k=1}^{2} \frac{|A_{ik}|}{|\Delta|} = \max_i \sum_{k=1}^{2} \frac{\left| \left( \frac{\partial f_i}{\partial \xi^{(k)}} \right)_0 \right|}{|\Delta|} \leqslant \frac{2l}{|\Delta|} = B^*.$$

В соответствии с этим условие (6) примет вид:

$$h_0 = \frac{4l^2}{|\Delta|^2} \bar{\eta} L \cdot 2^2 \leqslant \frac{1}{2}$$

или

$$32 l^2 L \bar{\eta} \leqslant |\Delta|^2.$$

В таком виде эта теорема была получена A. Ostrowski [6]. Любопытно, что примененное им для данного частного случая рассуждение, пожалуй, сложнее доказательства общей теоремы 1 и, повиди-

---

* За более точное значение для $B$ можно принять:

$$B = \max \left( \left| \left( \frac{\partial f_1}{\partial \xi^{(1)}} \right)_0 \right| + \left| \left( \frac{\partial f_1}{\partial \xi^{(2)}} \right)_0 \right| ; \ \left| \left( \frac{\partial f_2}{\partial \xi^{(1)}} \right)_0 \right| + \left| \left( \frac{\partial f_2}{\partial \xi^{(2)}} \right)_0 \right| \right).$$

мому, эта сложность доказательства заставила его ограничиться случаем двух уравнений. У него отсутствует также заключение об области единственности.

Замечание. Отметим, что мы могли бы применить и условие 2. Оно имеет вид:

$$\| x_1 - x_0 \| = \| \Gamma_0 P(x_0) \| \leqslant \eta.$$

В данном случае это будет оценка первых поправок

$$\| x_1 - x_0 \| = \max_i | \xi_1^{(i)} - \xi_0^{(i)} | \leqslant \eta.$$

При введении такой величины условие 4 нужно было бы заменить на

$$h_0 = BLm^2\eta \leqslant \frac{1}{2}.$$

Другую теорему о системах алгебраических уравнений мы получим, применив метрику пространства $R^m$.

**Теорема 5а.** *Если выполнены условия:*

1) *матрица* $\left\| \left( \frac{\partial f_i}{\partial \xi^{(k)}} \right)_0 \right\|$ *имеет обратную* $\left\| \frac{A_{ik}}{\Delta} \right\|$, *причем*

$$\frac{1}{|\Delta|} \left( \sum_{i=1}^{m} \sum_{k=1}^{m} A_{ik}^2 \right)^{\frac{1}{2}} \leqslant B$$

(*за $B$ может быть принята и иная оценка нормы обратной матрицы.*)

2) $\sum_{i=1}^{m} | \xi_1^{(i)} - \xi_0^{(i)} |^2 \leqslant \eta^2$,

3) $\sum_{i,j,k=1}^{m} \left( \frac{\partial^2 f_i}{\partial \xi^{(k)} \partial \xi^{(j)}} \right)^2 \leqslant K^2$

(*можно принять $K = m\sqrt{m}L$, где* $\left| \frac{\partial^2 f_i}{\partial \xi^{(k)} \partial \xi^{(j)}} \right| \leqslant L$),

4) $\eta_0 = BK\eta \leqslant \frac{1}{2}$,

*то процесс Ньютона сходится. При этом решение системы $x^* = (\xi^{(1)*}, \xi^{(2)*}, \ldots, \xi^{(m)*})$ лежит в области*

$$\| x^* - x_0 \| = \left[ \sum_{i=1}^{m} | \xi^{(i)*} - \xi_0^{(i)} |^2 \right]^{\frac{1}{2}} \leqslant \frac{1 - \sqrt{1 - 2h_0}}{h_0} \eta.$$

Доказательство. Доказательство проводится, так же как и доказательство теоремы 5, сведением к теореме 1. Только в данном случае нужно воспользоваться выражениями для норм операций, отображающих $R^m$ в $R^m$ [§ 1 (7), § 2 (5)].

В качестве примера рассмотрим следующую систему [16, 6]

$$f \equiv 2x^3 - y^2 - 1 = 0,$$
$$g \equiv xy^3 - y - 4 = 0.$$

За первое приближение возьмем точку $T_0$ с $x_0 = 1.2$; $y_0 = 1.7$. Тогда

$$f(T_0) = -0.434;\quad g(T_0) = 0.1956,$$

система для определения первых поправок $\Delta x_1 = x_1 - x_0$ и $\Delta y_1 = y_1 - y_0$ имеет вид:

$$f_x'(T_0)\Delta x_1 + f_y'(T_0)\Delta y_1 + f(T_0) = 0,\quad 8.64\Delta x_1 - 3.4\Delta y_1 - 0.434 = 0,$$
$$g_x'(T_0)\Delta x_1 + g_y'(T_0)\Delta y_1 + g(T_0) = 0,\quad 4.913\Delta x_1 + 9.404\Delta y_1 + 0.1956 = 0,$$

откуда их значения будут:

$$\Delta x_1 = x_1 - x_0 = 0.0349;\quad \Delta y_1 = y_1 - y_0 = -0.0390.$$

При этом определитель системы $\Delta = 97.95$.

Сначала оценим $h_0$, следуя Островскому. Примем

$$l = \max_{i,k}\left|\left(\frac{\partial f_i}{\partial \xi^{(k)}}\right)_0\right| = 9.404;\ \bar{\eta} = \max(\,|f(T_0)|\,,\,|g(T_0)|\,) = 0.434.$$

Для нахождения $L$ выпишем вторые производные:

$$f''_{xx} = 12x;\quad f''_{xy} = 0;\quad f''_{yy} = -2;\quad g''_{xx} = 0;\quad g''_{xy} = 3y^2;\quad g''_{yy} = 6xy;$$

оценивая их в прямоугольнике $0 \leqslant x \leqslant 1.3$, $0 \leqslant y \leqslant 1.8$, откуда не должны выйти последовательные приближения, получим: $L = 12\cdot 1.3 = 15.6$. Следовательно, значение $h_0$ будет:

$$h_0 = \frac{16l^2}{\Delta^2}\eta L = \frac{16\cdot 9.404^2\cdot 0.434\cdot 15.6}{97.95^2} = 0.998 > 0.5.$$

Таким образом, на основании теоремы Островского нельзя сделать заключения о сходимости процесса.

Попробуем применить теорему 5а.

Для нахождения $B$ рассмотрим матрицу

$$\Gamma_0 = \begin{Vmatrix}(f_x')_0, & (f_y')_0\\ (g_x')_0, & (g_y')_0\end{Vmatrix}^{-1} = \begin{Vmatrix}8.64 & -3.4\\ 4.913 & 9.404\end{Vmatrix}^{-1}. \tag{7}$$

Норма линейного преобразования в $R^2$, осуществляемого матрицей (7), есть, согласно примеру 2, § 1,

$$\|\Gamma_0\| = \sqrt{\Lambda_{\max}},$$

где $\Lambda_{\max}$ — наибольшее собственное значение матрицы $\Gamma_0\Gamma_0^*$, т. е. наибольший корень уравнения

$$\Lambda^2 - \frac{8.64^2 + 3.4^2 + 4.913^2 + 9.404^2}{\Delta^2}\Lambda + \frac{1}{\Delta^2} = 0,$$

т. е. уравнения

$$\Lambda^2 - 0.02072\Lambda + 0.00010422 = 0,$$

откуда

$$\Lambda_{max} = 0.01036 + \sqrt{0.00010733-0.00010422} = 0.0121$$

и

$$B = \| \Gamma_0 \| = \sqrt{\Lambda_{max}} = 0.11.$$

За $\eta$ берем величину

$$\eta = \| T_1 - T_0 \| = \sqrt{\Delta x_1^2 + \Delta y_1^2} = 0.0524.$$

И, наконец, $K$ оцениваем (в том же прямоугольнике, что и выше), согласно примеру 2 § 2:

$$K \leqslant (15.6^2 + 2^2 + 2\cdot 9.72^2 + 14.04^2)^{\frac{1}{2}} = \sqrt{631.72} = 25.14.$$

Теперь находим $h_0$:

$$h_0 = BK\eta = 0.11\cdot 25.14\cdot 0.0524 = 0.15 < 0.5.$$

Это показывает быструю сходимость процесса.

## § 7. Метод Ньютона для нелинейных интегральных уравнений

Будем рас сматривать нелинейное интегральное уравнение вид

$$x(s) = \int_0^1 K(s, t, x(t))\, dt, \tag{1}$$

где $K$ — непр ерывная функция своих аргументов. Если ввести оператор

$$y = P(x), \quad y(s) = x(s) - \int_0^1 K(s, t, x(t))\, dt, \tag{2}$$

то уравнение примет вид, рассмотренный в § 4.

Процесс Ньютона для него строится следующим образом. Задаемся начальным приближением $x_0(s)$. Тогда следующее приближение — функция $x_1(t)$ — должна определяться из линейного интегрального уравнения

$$x_1(s) - x_0(s) - \int_0^1 K_x'(s, t, x_0(t))(x_1(t) - x_0(t))\, dt = \varepsilon_0(s), \tag{3}$$

где

$$\varepsilon_0(s) = \int_0^1 K(s, t, x_0(t))\, dt - x_0(s).$$

Это уравнение может быть получено непосредственно или из общей формулы, связывающей $x_1$ и $x_0$ [§ 4, (2)]:

$$P'(x_0)(x_1 - x_0) = -P(x_0),$$

если учесть также смысл $P'(x)$ для данного случая [ср. § 3 (16)].

Таким образом, для нахождения последовательных приближений требуется при каждом шаге решать линейное интегральное уравнение.

Сходимость этого процесса для данного случая была, по моему предложению, исследована непосредственно в диссертации Д. М. Загадского [7, 8]. Однако полученные им условия сходимости $\left(h_0 \leqslant \frac{1}{10}\right)$ более жестки по сравнению с теми, которые получаются на основании общих теорем § 4. Из них получается теорема 6, если рассматривать операцию (2) как операцию в пространстве $C$.

**Теорема 6.** *Если выполнены следующие условия:*

1) *для начального значения* $x_0(s)$ *ядро*

$$K_x'(s, t, x_0(t)) = K(s, t) \tag{4}$$

*имеет резольвенту* $G(s,t)$, *причем*

$$\int_0^1 |G(s,t)|\, dt \leqslant B, \quad 0 \leqslant s \leqslant 1; \tag{5}$$

2) $$|\varepsilon_0(s)| = |x_0(s) - \int_0^1 K(s, t, x_0(t))\, dt| \leqslant \overline{\eta}; \tag{6}$$

3) $$|K_{u^2}''(s, t, u)| \leqslant K \tag{7}$$

*в области, определяемой неравенством* (9) [*или* (10)];

4) $$h = (B+1)^2 \eta K \leqslant \frac{1}{2}, \tag{8}$$

*то процесс Ньютона для интегрального уравнения* (*1*) *с начальным значением* $x_0(s)$ *сходится к решению этого уравнения, которое существует и лежит в области*

$$|x^*(s) - x_0(s)| \leqslant N(h)(B+1)\overline{\eta} \tag{9}$$

*и единственно в области*

$$|x(s) - x_0(s)| \leqslant L(h)(B+1)\overline{\eta}. \tag{10}$$

Доказательство. Для доказательства достаточно применить теоремы 1 и 2 (с условием 2' вместо 2), принимая в качестве пространств $X$ и $Y$ пространство $C$. При этом для оценки оператора $\Gamma_0 = [P'(x_0)]^{-1}$ нужно учесть, что в данном случае оператор $P'(x_0)$ есть оператор левой части уравнения (3), а обратный ему оператор $\Gamma_0$ дает выражение решения через свободный член и записывается с помощью резольвенты

$$x = \Gamma_0(y), \quad x(s) = y(s) - \int_0^1 G(s,t)\, y(t)\, dt.$$

Отсюда $\|\Gamma_0\|$ легко оценивается по § 1 (11) или непосредственно

$$|x(s)| \leqslant |y(s)| + \int_0^1 |G(s, t) y(t)|\, dt \leqslant (1 + \max_s \int_0^1 |G(s, t)|\, dt)\, \|y\|.$$

Поэтому

$$\| \Gamma_0 \| \leqslant 1 + B$$

и, следовательно, условие 1 выполнено с $B_0 = 1 + B$.

Условие 2' очевидным образом также выполнено, благодаря условию 2 данной теоремы, так как $\varepsilon_0(s)$ как раз и есть $P(x_0)$. Наконец, если учесть, что вторая производная тождественного оператора равна нулю, то $P''(x)$, согласно примеру 3 § 3, представляет билинейную операцию с ядром $K_{x^2}''(s, t, x_0(t))$, а потому, согласно § 2 (9) и условию 3 данной теоремы, будет

$$\| P''(x) \| \leqslant K.$$

Наконец, условие 4 эквивалентно форме условия 4, указанной в замечании 1 § 4. Итак, все условия теоремы 1 и 2 соблюдены: следовательно, в результате их применения приходим к утверждениям теоремы.

Как уже упоминалось, теорема, подобная теореме 6, была получена Д. М. Загадским. Она отличается от доказанной лишь заменой числа $\frac{1}{2}$ на $\frac{1}{10}$ в условии 4.

Выбирая в качестве пространств $X$ и $Y$ пространство $L^2$, получим следующую теорему.

**Теорема 7.** *Пусть выполнены условия*:

1) $$\int_0^1 \left[ x_0(s) - \int_0^1 K(s, t, x_0(t))\, dt \right]^2 ds \leqslant \eta^2.$$

2) *Соблюдено неравенство*

$$\frac{|\lambda_n|}{|1-\lambda_n|} \leqslant B,$$

*где* $\lambda_n$ — *собственные числа ядра*

$$K_x'(s, t, x_0(t)) = K(s, t),$$

*если последнее симметрично*; *если ядро не симметрично, то следует потребовать выполнения неравенства*

$$\frac{|\Lambda_n|}{|1-\Lambda_n|} \leqslant B^2,$$

*где* $\Lambda_n$ — *собственное значение симметричного ядра, построенного по данному по формуле* (15) § 1;

3) $|K_{u^2}''(s, t, u)| \leqslant K$

*для всех конечных значений* $u$;

4) $B^2 K \eta \leqslant \frac{1}{2}$.

*Тогда уравнение* (1) *имеет решение, которое может быть найдено процессом Ньютона.*

Доказательство теоремы 7 получается непосредственно на основании теоремы 1, так как можно без труда проверить, что все условия ее выполнены.

Пример. В качестве примера применения теорем 6 и 7 рассмотрим интегральное уравнение (ср. Загадский [7])

$$x(s) = 1 - 0.4854\,s + s^2 + \int_0^1 st \operatorname{arc\,tg} x(t)\,dt;$$

точное его решение есть $x^*(s) = 1 + s^2$.

Возьмем начальное приближение $x_0(s) \equiv \frac{3}{2}$ и применим теорему 6. Ядро

$$K(s,t) = K_u'(s, t, x_0(t)) = \frac{4}{13} st,$$

поэтому его резольвента $G(s, t)$ должна иметь вид:

$$G(s,t) = Cst.$$

Постоянную $C$ определяем из интегрального уравнения резольвенты

$$G(s,t) = K(s,t) + \int_0^1 K(s,u)\,G(u,t)\,du,$$

что дает:

$$C = \frac{4}{13} + \frac{4}{39} C; \qquad C = \frac{12}{35}.$$

Находим теперь $B$:

$$B = \max_s \int_0^1 |G(s,t)|\,dt = \frac{6}{35}.$$

Определяем $\overline{\eta}$. Так как

$$\varepsilon_0(s) = 1 - 0.4854\,s + s^2 + s\int_0^1 t\left(\operatorname{arc\,tg}\frac{3}{2}\right)dt - \frac{3}{2} = s^2 + 0.006\,s - 0.5,$$

то

$$\overline{\eta} = \max_s |\varepsilon_0(s)| = \varepsilon_0(1) = 0.506.$$

Наконец,

$$K = \max_{\substack{0\leqslant s\leqslant 1\\ 0\leqslant t\leqslant 1}} |K_{u^2}''(s,t,u)| = \max\left|\frac{2stu}{(1+u^2)^2}\right| = \frac{2\sqrt{\frac{1}{3}}}{\left(1+\frac{1}{3}\right)^2} = \frac{3\sqrt{3}}{8}.$$

Подсчитываем $h$:

$$h = (B+1)^2 K\overline{\eta} = \left(\frac{41}{35}\right)^2 \cdot 0.506 \cdot \frac{3\sqrt{3}}{8} = 0.451 < 0.5.$$

Таким образом, сходимость процесса обеспечена*.

---

* Применение теоремы 7 позволило бы получить значительно меньшее значение для $h$, а именно: $h = 0.141$.

Находим следующее приближение. Согласно общей теории, поправку $\Delta x = x_1 - x_0$ определяем из уравнения

$$\Delta x(s) = \int_0^1 \frac{st}{1 + [x_0(t)]^2} \Delta x(t)\, dt + \varepsilon_0(s),$$

т. е.

$$\Delta x(s) = \frac{4}{13} s \int_0^1 t \Delta x(t)\, dt + s^2 + 0.006 s - 0.5.$$

Следовательно,

$$\Delta x(s) = s^2 + 0.006 s - 0.5 + \int_0^1 G(s, t)(t^2 + 0.006 t - 0.5)\, dt = s^2 + 0.006 s -$$

$$- 0.5 + \int_0^1 \frac{12}{35} st (t^2 + 0.006 t - 0.5)\, dt = s^2 + 0.0067 s - 0.5.$$

Таким образом, первое приближение равно

$$x_1(s) = x_0(s) + \Delta x(s) = s^2 + 0.0067 s + 1,$$

что отличается от точного решения меньше, чем на 0.01.

Отметим, что следующие приближения в подобных случаях, когда ядро $K_x'(s, t, x_0(t))$ особенно просто, удобнее находить, пользуясь модифицированным процессом. Впрочем, для данного примера это не существенно, так как он носит чисто иллюстративный характер — точное, решение его легко приводится к алгебраическому уравнению.

К решению многих задач, относящихся к нелинейным дифференциальным уравнениям, также может быть применен метод Ньютона. При этом исследование его может быть проведено непосредственно, применением теоремы 1, или после предварительного сведения проблемы к нелинейному интегральному или интегро-дифференциальному уравнению.

## § 8. Метод Ньютона для задачи о собственных значениях

Мы проведем рассмотрение задачи о собственных значениях для случая, когда $X = H$ есть пространство Гильберта (например, $L^2$, $R^n$), а оператор $A$ самосопряженный и вполне непрерывный (с изолированными собственными значениями). Задача состоит в отыскании пар $(x, \lambda)$, удовлетворяющих уравнению

$$Ax - \lambda x = 0,$$

причем $x \neq 0$. Такого рода значение $x$ называется собственным вектором, а значение $\lambda$ — собственным числом. Если возможно найти решение $x \neq 0$, то можно найти и такое решение $x$, что $\| x \|^2 = (x, x) = 1$*. При этом дополнительном условии, как правило, задача перестает быть неопределенной и в такой форме мы и будем ставить ее:

$$\left.\begin{aligned} &Ax - \lambda x = 0, \\ &\frac{1}{2}[(x, x) - 1] = 0. \end{aligned}\right\} \tag{1}$$

Если ввести оператор

$$P\begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} Ax - \lambda x \\ \frac{1}{2}[(x, x) - 1] \end{pmatrix}, \tag{2}$$

переводящий пространство $X'$ пар $(x, \lambda)$, где $x \in H$, а $\lambda$ — вещественное число, в то же самое пространство, то систему уравнений (1) можно заменить одним уравнением:

$$P\begin{pmatrix} x \\ \lambda \end{pmatrix} = 0. \tag{3}$$

Рассмотрим применение метода Ньютона именно к этому уравнению. Предварительно найдем производные оператора $P\begin{pmatrix} x \\ \lambda \end{pmatrix}$. Имеем:

$$P\begin{pmatrix} x + \Delta x \\ \lambda + \Delta\lambda \end{pmatrix} = \begin{pmatrix} Ax + A\Delta x - x\Delta\lambda - \lambda\Delta x - \lambda x - \Delta\lambda\Delta x \\ \frac{1}{2}(x, x) + (x, \Delta x) + \frac{1}{2}(\Delta x, \Delta x) - \frac{1}{2} \end{pmatrix},$$

откуда, сохраняя малые первого порядка, найдем:

$$dP\begin{pmatrix} x \\ \lambda \end{pmatrix} = P'\begin{pmatrix} x \\ \lambda \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = \begin{pmatrix} (A - \lambda I)\Delta x - x\Delta\lambda \\ (x, \Delta x) \end{pmatrix}. \tag{4}$$

Аналогично для второй производной, записанной как билинейная операция, получаем:

$$P''\begin{pmatrix} x \\ \lambda \end{pmatrix}\left[\begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix}, \begin{pmatrix} \Delta' x \\ \Delta'\lambda \end{pmatrix}\right] = \left[P'\begin{pmatrix} x + \Delta' x \\ \lambda + \Delta'\lambda \end{pmatrix} - P'\begin{pmatrix} x \\ \lambda \end{pmatrix}\right]\begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} =$$
$$= \begin{pmatrix} -\Delta'\lambda\Delta x - \Delta' x\Delta\lambda \\ (\Delta' x, \Delta x) \end{pmatrix}. \tag{5}$$

Зададимся в качестве начальных некоторыми значениями $\lambda$ и $x$. При этом будем считать, что $x$ нормировано, или, во всяком случае, $(x, x) \approx 1$. Тогда при подстановке в уравнение (3) будем иметь:

$$P\begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} -y \\ -t \end{pmatrix}, \quad \begin{aligned} y &= -(Ax - \lambda x), \\ t &= -\frac{1}{2}[(x, x) - 1]. \end{aligned} \tag{6}$$

* Вообще $(x, y)$ обозначает скалярное произведение векторов $x$ и $y$. В случае, если $X = L^2$, то

$$(x, y) = \int_0^1 x(t)\, y(t)\, dt$$

Уравнение для определения поправки $\Delta x = x_1 - x_0$ по методу Ньютона в общем случае имеет вид (§ 4 (2)):

$$P'(x_0)(x_1 - x_0) + P(x_0) = 0,$$

а потому в рассматриваемом случае оно запишется так:

$$P'\begin{pmatrix} x \\ \lambda \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -P\begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} y \\ t \end{pmatrix},$$

или же, принимая во внимание выражение $P'\begin{pmatrix} x \\ \lambda \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix}$,

$$\left.\begin{aligned} &(A - \lambda I)\Delta x - x\Delta\lambda = y, \\ &(x, \Delta x) = t. \end{aligned}\right\} \tag{7}$$

Из этих линейных уравнений и определяются последовательные поправки. При этом, в случае метода Ньютона, в чистом виде значения $x$ и $\lambda$ в левой части заменяются последовательно их исправленными значениями; в случае же модифицированного процесса (§ 5), изменение вносится только в свободные члены системы (7).

Для исследования сходимости процесса Ньютона необходимо прежде всего составить обратный оператор

$$\Gamma = \left[P'\begin{pmatrix} x \\ \lambda \end{pmatrix}\right]^{-1}, \tag{8}$$

который находится, если систему (7) решить в общем виде, а именно: при произвольных $y$ и $t$ оказывается

$$\begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = \Gamma\begin{pmatrix} y \\ t \end{pmatrix}. \tag{9}$$

Если такое выражение найдено, то имеем возможность оценить непосредственно

$$\left\|\left[P'\begin{pmatrix} x \\ \lambda \end{pmatrix}\right]^{-1}\right\| = \|\Gamma\|.$$

Тогда $\|\Gamma\|$ оценивается как норма оператора, переводящего пространство пар $\begin{pmatrix} x \\ \lambda \end{pmatrix}$ в себя, причем норму элемента в этом пространстве принимаем равной

$$\left\|\begin{pmatrix} x \\ \lambda \end{pmatrix}\right\| = \sqrt{\|x\|^2 + |\lambda|^2} = \sqrt{(x, x) + \lambda^2}. \tag{10}$$

Далее для применения теоремы нам необходимо оценить еще $\left\|P''\begin{pmatrix} x \\ \lambda \end{pmatrix}\right\|$. Это может быть сделано в общем виде на основании выражения (5) для $P''$.

Имеем:

$$\left\| P''\begin{pmatrix}x\\ \lambda\end{pmatrix}\begin{pmatrix}\Delta x\\ \Delta\lambda\end{pmatrix}\begin{pmatrix}\Delta' x\\ \Delta'\lambda\end{pmatrix}\right\|^2=\left\|\begin{pmatrix}-\Delta'\lambda\Delta x-\Delta' x\Delta\lambda\\ (\Delta' x,\ \Delta x)\end{pmatrix}\right\|^2=(\Delta' x,\ \Delta x)^2+$$

$$+(\Delta'\lambda\Delta x+\Delta' x\Delta\lambda,\ \Delta'\lambda\Delta x+\Delta' x\Delta\lambda)\leqslant\|\Delta' x\|^2\|\Delta x\|^2+(\Delta'\lambda)^2\|\Delta x\|^2+$$

$$+(\Delta\lambda)^2\|\Delta' x\|^2+2|\Delta'\lambda||\Delta\lambda|\|\Delta x\|\ \|\Delta' x\|\leqslant\|\Delta' x\|^2\|\Delta x\|^2+$$

$$+(\Delta'\lambda)^2\|\Delta x\|^2+(\Delta\lambda)^2\|\Delta' x\|^2+2\ \frac{(\Delta'\lambda)^2+(\Delta\lambda)^2}{2}\frac{\|\Delta' x\|^2+\|\Delta x\|^2}{2}\leqslant$$

$$\leqslant\frac{3}{2}\left(\|\Delta x\|^2+|\Delta\lambda|^2\right)\left(\|\Delta' x\|^2+|\Delta'\lambda|^2\right)=\frac{3}{2}\left\|\begin{pmatrix}\Delta x\\ \Delta\lambda\end{pmatrix}\right\|^2\left\|\begin{pmatrix}\Delta' x\\ \Delta'\lambda\end{pmatrix}\right\|^2.$$

Отсюда следует, что

$$\left\|P''\begin{pmatrix}x\\ \lambda\end{pmatrix}\right\|\leqslant\sqrt{\frac{3}{2}}. \tag{11}$$

Таким образом, в соответствии с теоремой 3 условие сходимости процесса Ньютона и модифицированного процесса может быть записано в виде:

$$h_0=\sqrt{\frac{3}{2}}\,\|\Gamma\|\left\|\begin{pmatrix}\Delta x\\ \Delta\lambda\end{pmatrix}\right\|<\frac{1}{2}, \tag{12}$$

где, как сказано, $\|\Gamma\|$ должна быть оценена непосредственно на основании ее выражения (9), $\Delta x$ и $\Delta\lambda$ — значения поправок при переходе от начальных значений к первому приближению, т. е. решения системы (7), а норма определяется по (10).

Необходимо отметить, что при выполнении условия (12) мы можем утверждать наличие собственной пары $\begin{pmatrix}x\\ \lambda\end{pmatrix}$ и сходимость к ней процесса Ньютона, однако, нельзя гарантировать, что это есть пара, удовлетворяющая каким-либо определенным условиям, например, что это пара, соответствующая наименьшему собственному значению.

Данный процесс, в частности, может применяться для нахождения собственных чисел матричного оператора, что, как известно, эквивалентно решению так называемого «векового» уравнения.

В этом случае применение процесса требует на каждом шагу решения системы $n+1$ линейных алгебраических уравнений, так как именно такой вид будет иметь в данном случае система (7). При этом в случае применения модифицированного процесса коэффициенты этой системы будут для всех шагов процесса одни и те же. Система может быть еще несколько упрощена, так как второе из уравнений (1) можно заменить более простым $(x, U)=1$, где $U$ — постоянный вектор, например, $U=(1, 0, \ldots, 0)$.

Докажем, что условие сходимости процесса Ньютона (12) будет обязательно выполнено, если оператор $A$ имеет изолированное простое собственное значение и начальное приближение достаточно близко к нему. Впрочем, для некоторого упрощения записи проведем рассуждения для случая, когда спектр оператора $A$ дискретный.

Пусть $\lambda_0$ и $x_0$ — данное собственное значение и элемент. Прочие собственные значения и элементы обозначим $\lambda_n$, $x_n$ $(n=1, 2, \ldots)$; среди этих $\lambda_n$ могут быть и равные, но элементы $x_n$ все ортогональны один к другому. Пусть $\lambda_1$ — значение, ближайшее к $\lambda_0$. Далее, добавим еще элементы $x_{-1}, x_{-2}, \ldots$ так, чтобы вместе с элементами $x_0, x_1, x_2, \ldots$ они образовали полную ортогональную систему. Будем считать $\lambda_n=0$ $(n=-1, -2, \cdots)$. Теперь постараемся оценить норму оператора $\Gamma=\left[P'\begin{pmatrix}x\\ \lambda\end{pmatrix}\right]^{-1}$. Этот оператор выражает решение системы (7) [ср. (9)], которая в данном случае имеет вид:

$$\left.\begin{aligned}(A-\lambda_0 I)\,\Delta x-x_0\Delta\lambda&=y,\\ (x_0, \Delta x)&=t.\end{aligned}\right\} \tag{7a}$$

Пусть

$$y=\sum_{-\infty}^{+\infty}\alpha_n x_n.$$

Будем искать решение в виде $\Delta x=\Sigma\beta_n x_n$. Тогда система (7a), если учесть, что $Ax_n=\lambda_n x_n$, примет вид:

$$\sum_{-\infty}^{+\infty}\beta_n(\lambda_n-\lambda_0)\,x_n-x_0\,\Delta\lambda=\sum\alpha_n x_n,$$
$$\beta_0=t.$$

Отсюда, приравнивая коэффициенты, получим:

$$\beta_0=t,\quad -\Delta\lambda=\alpha_0,\quad \beta_n=\frac{1}{\lambda_n-\lambda_0}\alpha_n\quad (n=1, 2, \ldots);$$
$$\beta_n=\frac{-1}{\lambda_0}\alpha_n\quad (n=-1, -2, \ldots)$$

и, следовательно,

$$\left\|\Gamma\begin{pmatrix}y\\ t\end{pmatrix}\right\|^2=\left\|\begin{matrix}\Delta x\\ \Delta\lambda\end{matrix}\right\|^2=\|x\|^2+|\Delta\lambda|^2=\alpha_0^2+\sum\beta_n^2=\alpha_0^2+t^2+$$
$$+\sum_{n=1}^{\infty}\frac{1}{(\lambda_n-\lambda_0)^2}\alpha_n^2+\frac{1}{\lambda_0^2}\sum_{n=-1}^{-\infty}\alpha_n^2\leqslant t^2+\max\left(1, \frac{1}{|\lambda_0|^2}, \frac{1}{|\lambda_1-\lambda_0|^2}\right)\sum_{-\infty}^{+\infty}\alpha_k^2\leqslant$$
$$\leqslant\max\left(1, \frac{1}{|\lambda|_0^2}, \frac{1}{|\lambda_1-\lambda_0|^2}\right)\left\|\begin{matrix}y\\ t\end{matrix}\right\|^2.$$

Отсюда видно, что

$$\|\Gamma\|=\left\|\left[P'\begin{pmatrix}x_0\\ \lambda_0\end{pmatrix}\right]^{-1}\right\|\leqslant\max\left(1, \frac{1}{|\lambda_0|}, \frac{1}{|\lambda_1-\lambda_0|}\right).$$

Если начальные значения $x$, $\lambda$ близки к $x_0$ и $\lambda_0$, то $\left\|\left[P'\begin{pmatrix}x\\ \lambda\end{pmatrix}\right]^{-1}\right\|$ будет ограничена, а $\eta$ будет мало, так как $\Delta x$ и $\Delta\lambda$ будут малы, а потому условие (12) будет обязательно соблюдено.

Иным образом границу сходимости можно определить, используя замечание 5 § 4. Здесь соответствующее неравенство [§ 4 (45)] примет вид:

$$\left\|\begin{pmatrix}x\\ \lambda\end{pmatrix}-\begin{pmatrix}x_0\\ \lambda_0\end{pmatrix}\right\|=\sqrt{\|x-x_0\|^2+(\lambda-\lambda_0)^2}\leqslant\frac{1}{4\,\|\Gamma\|\sqrt{\frac{3}{2}}},$$

т. е. удовлетворяющая этому неравенству пара значений $\begin{pmatrix}x\\ \lambda\end{pmatrix}$, взятая в качестве начальных значений, обеспечивает сходимость к собственной паре $\begin{pmatrix}x_0\\ \lambda_0\end{pmatrix}$.

## ЛИТЕРАТУРА


1. Ostrowski A. Сборник работ памяти Д. А. Граве, стр. 213, 1940.
2. Ostrowski A. Математический сборник 2, стр. 1073, 1937.
3. Ostrowski A. Математический сборник 3 (45), 1938.
4. Willers F. A. Methoden der praktischen Analysis, Berlin, 1928.
5. Стенин Н. П. Сборник работ. Конформное отображение. М. — Л., 1937.
6. Ostrowski A. Comment. Mathem. Helv., 9, p. 79, 1937.
7. Загадский Д. М. Приближенное решение нелинейных интегральных уравнений. Диссертация. Пед. ин-т им. Герцена, 1946.
8. Загадский Д. М. ДАН СССР 59, № 6, 1948.
9. Люстерник Л. А. Основные идеи функционального анализа. Успехи матем. наук, вып. 1.
10. Banach S. Théorie des opérations linéaires.
11. Fréchet M. Ann. Ec. Norm., p. 213, 1925.
12. Kerner M. Studia Mathem., 3.
13. Cauchy A. Oeuvres complètes (II), 4, p. 273.
14. Немыцкий В. Метод неподвижных точек в анализе. Успехи матем. наук, вып. 1.
15. Граве Д. Журн. Матем. ин-та АН УССР 2, 3, 1936.
16. Runge C. und Konig H. Vorlesungen über numerisches rechnen. Berlin, 1924.
17. Гавурин М. К. ДАН СССР 22 (1939), 547—556.
18. Л. В. Канторович. ДАН СССР 56, 117, 1948.
19. Л. В. Канторович. Учен. Зап. Л. Г. У., 1937, № 17.