

Лекция 2

2. Численное нахождение корней уравнений и решений систем уравнений

Мы начнем изложение методов исследования нелинейных задач с простейшего случая, когда объектом изучения является простое (не дифференциальное) нелинейное уравнение или система таких уравнений вида

$$F_1(x_1^{(n)}, x_2^{(n)}, \dots, x_N^{(n)}) = 0$$

$$F_2(x_1^{(n)}, x_2^{(n)}, \dots, x_N^{(n)}) = 0$$

...

$$F_N(x_1^{(n)}, x_2^{(n)}, \dots, x_N^{(n)}) = 0$$

которую удобно записать в векторном виде

$$\mathbf{F}(\mathbf{x}) = 0 \tag{1}$$

Для дальнейшего оказывается полезным переписать уравнение (1) в виде

$$\mathbf{x} = \mathbf{f}(\mathbf{x}) \tag{2}$$

и находить решение уравнения (1), путем построения последовательных приближений.

2.1. Необходимые сведения из линейной алгебры

Напомним следующее определение.

Пусть M - линейное пространство. Функция $\|\cdot\| : M \rightarrow \mathbb{R}$ является *векторной нормой*, если для всех $\mathbf{x}, \mathbf{y} \in M$ выполняются следующие условия

- $\|\mathbf{x}\| \geq 0$, причем $\|\mathbf{x}\| = 0$ тогда и только тогда, когда $\mathbf{x} = 0$;
- $\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$;
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

В пространстве \mathbb{R}^N можно ввести векторную норму различными способами. В частности, используются следующие нормы

$$\|\mathbf{x}\|_\infty = \max_i |x_i|,$$

$$\|\mathbf{x}\|_1 = |x_1| + \dots + |x_N|,$$

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_N^2}$$

В традиционном курсе линейной алгебры наиболее часто используется последняя, *евклидова* норма элемента. Вместе с тем, все три приведенных векторных нормы эквивалентны.

Нормы $\|\cdot\|_a$ и $\|\cdot\|_b$ называют *эквивалентными*, если существуют такие константы C_1 и C_2 , что для любого элемента $\mathbf{x} \in \mathbb{R}$ выполняются соотношения

$$C_1 \|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq C_2 \|\mathbf{x}\|_b$$

Приведем еще одно определение.

Пусть \mathbf{A} - линейный оператор, действующий в M - линейном пространстве, снабженном векторной нормой $\|\cdot\|$. Нормой оператора \mathbf{A} , *подчиненной* этой векторной норме называется число

$$\|\mathbf{A}\| \equiv \sup_{\mathbf{x} \in M} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x} \in \{\mathbf{x}: \|\mathbf{x}\|=1\}} \|\mathbf{Ax}\|$$

В пространстве \mathbb{R}^N в качестве $\|\mathbf{x}\|$ можно использовать любую из норм, определенных выше. Норму \mathbf{A} можно определить при этом как

$$\|\mathbf{A}\| = \max_{\mathbf{x} \in \mathbb{S}^{N-1}} \|\mathbf{Ax}\|$$

где \mathbb{S}^{N-1} - сфера в \mathbb{R}^N . Замена \sup на \max оправдана, так как \mathbb{S}^{N-1} - компактное множество.

Укажем явные формулы для матричных норм, подчиненных представленным выше векторным нормам:

$$\|\mathbf{A}\|_{\infty} = \max_i \left(\sum_{j=1}^N |a_{ij}| \right),$$

$$\|\mathbf{A}\|_1 = \max_j \left(\sum_{i=1}^N |a_{ij}| \right),$$

$$\|\mathbf{A}\|_2 = \sqrt{\max_i \lambda(A^T A)}$$

В последнем равенстве $\lambda(A^T A)$ - набор собственных значений симметричной матрицы $A^T A$.

2.2. Итерационные процессы и их скорость сходимости

Пусть $\mathbf{f}(\mathbf{x})$ - отображение из \mathbb{R}^N в \mathbb{R}^N . Назовем *итерационным процессом* построение последовательности $\{\mathbf{x}^{(n)}\}$ по следующему рекуррентному правилу

$$\mathbf{x}^{(n+1)} = \mathbf{f}(\mathbf{x}^{(n)}) \quad (3)$$

Записанное в развернутом виде, (3) представляется формулами

$$x_1^{(n+1)} = f_1(x_1^{(n)}, x_2^{(n)}, \dots, x_N^{(n)})$$

$$x_2^{(n+1)} = f_2(x_1^{(n)}, x_2^{(n)}, \dots, x_N^{(n)})$$

...

$$x_N^{(n+1)} = f_N(x_1^{(n)}, x_2^{(n)}, \dots, x_N^{(n)})$$

Существует общепринятая терминология для описания скорости сходимости итерационного процесса.

Пусть $\{\mathbf{x}^{(n)}\} \subset \mathbb{R}^N$ и $\mathbf{x}^* \in \mathbb{R}^N$. Будем говорить, что

- $\{\mathbf{x}^{(n)}\}$ сходится к \mathbf{x}^* *линейно*, если $\{\mathbf{x}^{(n)}\} \rightarrow \mathbf{x}^*$ и существует такое $q > 0$, что

$$\|\mathbf{x}^{(n+1)} - \mathbf{x}^*\| \leq q \|\mathbf{x}^{(n)} - \mathbf{x}^*\|$$

- $\{\mathbf{x}^{(n)}\}$ сходится к \mathbf{x}^* *квадратично*, если $\{\mathbf{x}^{(n)}\} \rightarrow \mathbf{x}^*$ и существует такое $K > 0$, что

$$\|\mathbf{x}^{(n+1)} - \mathbf{x}^*\| \leq K \|\mathbf{x}^{(n)} - \mathbf{x}^*\|^2$$

Используя это определение можно ввести понятия *линейно* сходящегося итерационного процесса и *квадратично* сходящегося процесса:

Скажем, что итерационный процесс (3) сходится линейно (квадратично) к \mathbf{x}^* - решению уравнения (2), если для любого начального приближения $\mathbf{x}^{(0)}$, расположенного в некоторой окрестности \mathbf{x}^* , последовательность $\{\mathbf{x}^{(n)}\}$ сходится к \mathbf{x}^* линейно (квадратично).

Очевидно, квадратично сходящийся итерационный процесс сходится быстрее, чем процесс, сходящийся только линейно. Однако для того, чтобы итерационный процесс вообще сходил, необходимо, чтобы правая часть в (3) обладала достаточно “хорошими” свойствами. О них речь пойдет в разделе 2.5.

2.3. Уравнения вида $F(x) = 0$

Численному исследованию уравнений вида

$$F(x) = 0, \quad x \in \mathbb{R} \tag{4}$$

посвящен значительный раздел курса численных методов. Представить себе расположение корней этого уравнения *в целом* можно построив график этого уравнения и исследовав свойства функции $y =$

$F(x)$. Для более точного нахождения корней можно вспомнить следующие методы, пригодные для численного решения уравнения (4).

1. Метод *дихотомии*. Описание этого простого и надежного метода можно найти практически в любом учебнике численного анализа. К сожалению, этот метод, вообще говоря, неприменим для систем с более чем одной неизвестной, поэтому мы не будем на нем подробно останавливаться.

2. Метод *простой итерации*. Он заключается в переписывании уравнения (4) тем или иным способом в виде (2), то есть

$$x = f(x) \quad (5)$$

с последующим решением этого уравнения при помощи итерационного процесса

$$x^{(n+1)} = f(x^{(n)}), \quad n = 0, 1, \dots \quad (6)$$

Метод простой итерации удобно иллюстрировать графически при помощи так называемых *диаграмм Ламерея*. Построим графики функций $y = f(x)$ и $y = x$. Корню $x = x^*$ уравнения (5) соответствует точка пересечения этих графиков. Рисунок 1 показывает как могут выглядеть последовательные приближения к этому корню.

Наблюдение за диаграммами Ламерея показывает, что случаи $|f'(x^*)| > 1$ и $|f'(x^*)| < 1$ существенно различаются. В первом из этих случаев диаграмма Ламерея показывает, что сходимости к корню x^* нет, по крайней мере для начальных значений, расположенных близко к x^* . Во втором случае такая сходимость имеется. Действительно, в традиционном курсе численных методов доказывается следующее утверждение:

Теорема 0.1 (о сходимости метода простой итерации) Пусть уравнение (5) имеет решение x^* , функция $f(x)$ непрерывно дифференцируема на интервале $U_\delta = (x^* - \delta; x^* + \delta)$, причем на этом интервале $|f'(x)| \leq q < 1$. Тогда x^* является единственным корнем на этом промежутке, итерационный процесс (6) сходится к этому корню, причем

$$|x^{(n)} - x^*| \leq \frac{q^n |x^{(1)} - x^{(0)}|}{1 - q}$$

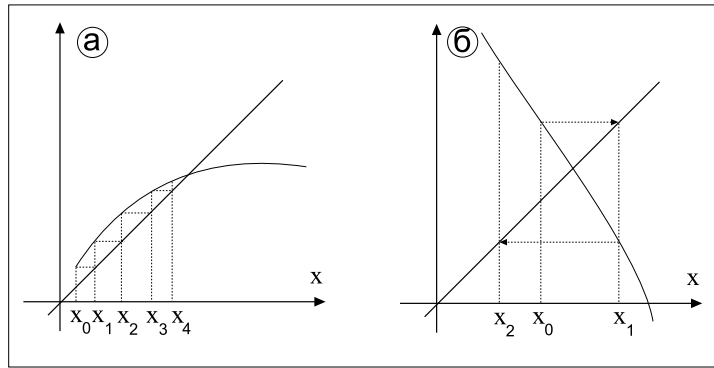


Рис. 1.

Метод простой итерации, вообще говоря, сходится *линейно*. Сформулированная теорема позволяет оценить погрешность метода. Значение q можно оценить, предполагая, что закон, описывающий сходимость в непосредственной близости от x^* , соответствует геометрической прогрессии. Тогда можно считать, что

$$q = \frac{x^{(n)} - x^{(n-1)}}{x^{(n-1)} - x^{(n-2)}}$$

и процесс можно прерывать, когда величина

$$\Delta = \left| \frac{q(x^{(n)} - x^{(n-1)})}{1 - q} \right| = \frac{(x^{(n)} - x^{(n-1)})^2}{|x^{(n)} + x^{(n-2)} - 2x^{(n-1)}|}$$

окажется меньше заданной погрешности ε . Более детальное обсуждение сходимости метода простой итерации для более общего случая см. в параграфе 2.5.

3. Метод *Ньютона*. В методе Ньютона нахождение корня уравнения производится при помощи следующего итерационного алгоритма

$$x^{(n+1)} = x^{(n)} - \frac{F(x^{(n)})}{F'(x^{(n)})}$$

Такой алгоритм очень быстро (квадратично) сходится, если начальное приближение оказывается близко к искомому корню. О методе

Ньютона как методе нахождения корней уравнения $F(x) = 0$ достаточно подробно говорилось в курсе численных методов. Здесь подчеркнем два важных момента.

Во-первых, метод Ньютона является специальным частным случаем метода простой итерации. При этом правая часть $f(x) = x - F(x)/F'(x)$ имеет производную

$$f'(x) = \frac{F(x)F''(x)}{(F'(x))^2}$$

которая зануляется в точке x^* , если $F(x^*) = 0$. Таким образом если $F'(x^*) \neq 0$, функция $f(x)$ имеет в точке x^* экстремум.

Во-вторых, стоит напомнить о полезной интерпретации метода Ньютона при помощи касательных к графику функции $y = F(x)$, см. любой курс численного анализа.

В-третьих, в силу быстрой сходимости, выход из итерационного процесса можно производить как только разность между значениями на соседних итерациях станут меньше заданной погрешности

$$|x^{(n)} - x^{(n-1)}| < \varepsilon |x^{(n)}|.$$

На практике одним из недостатков метода Ньютона является достаточно узкая область сходимости. Для начального значения вне этой области метод Ньютона очень быстро расходится. Некоторые строгие утверждения о сходимости метода Ньютона приведены ниже, в 2.6.

2.4. Системы уравнений вида $F_1(x, y) = 0$, $F_2(x, y) = 0$

Исследование системы уравнений вида

$$F_1(x, y) = 0 \tag{7}$$

$$F_2(x, y) = 0 \tag{8}$$

также естественно начать с грубого анализа числа решений. Эта задача, как правило, оказывается сложнее, чем исследование уравнения (4). Вместе с тем, используя если необходимо компьютерные программы визуализации, можно попытаться построить тем или иным

способом геометрическое место точек на плоскости (x, y) , удовлетворяющих соотношению $F_1(x, y) = 0$, затем - точек, удовлетворяющих соотношению $F_2(x, y) = 0$ и выяснить примерное расположение точек пересечения этих множеств и их число.

Пример 1. Рассмотрим систему уравнений

$$y = x^3 - 4x; \quad x = y^3 - 4y$$

Первому из этих уравнений соответствует график кубической параболы, второму - тот же график, отраженный относительно прямой $y = x$. Окончательно получаем, что имеется девять точек пересечения кривых и, следовательно, девять решений рассматриваемой системы уравнений.

Задача. Решите эту систему уравнений в явном виде.

Пример 2. Рассмотрим комплексное уравнение

$$\sin z = z$$

Один из его корней $z = 0$ очевиден. Для исследования остальных его корней $z = x + iy$ выделим действительную и мнимую части этого уравнения

$$\sin x \operatorname{ch} y = x; \quad \cos x \operatorname{sh} y = y$$

Далее, на плоскости (x, y) находим точки, удовлетворяющие соотношению $\operatorname{ch} y = x / \sin x$ и точки, отвечающие равенству $\cos x = y / \operatorname{sh} y$. Окончательно, получаем, что на комплексной плоскости имеется бесконечное число корней нашего уравнения, расположенных комплексно сопряженными парами в полосах $[2\pi; 3\pi]$, $[4\pi; 5\pi]$, ..., а также $[-3\pi; -2\pi]$, $[-5\pi; -4\pi]$ и т.д. В каждой из этих полос имеется единственная пара комплексно сопряженных точек, соответствующих решениям нашей системы. Кроме того, в полосе $[-\pi; \pi]$ имеется единственная точка $z = 0$.

Пример 3. Рассмотрим систему уравнений

$$(x^2 + y^2)^3 = 4x^2y^2, \quad 4y^2 = \cos x$$

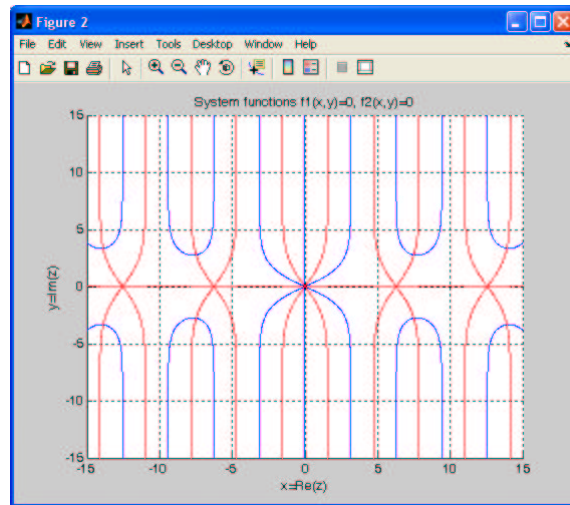


Рис. 2.

Первому из уравнений на плоскости (x, y) соответствует так называемая *четырёхлепестковая роза* (легко строится в полярных координатах). Множество, соответствующее второму уравнению также легко строится. Получаем, что система имеет 8 решений.

Уравнения в системе (7)-(8) могут быть заданы неявно, а для нахождения значений функций $F_1(x, y)$, $F_2(x, y)$ может потребоваться решение дифференциальных уравнений, вычисление интегралов и т.д. Это не меняет структуры самой системы и методов ее исследования.

Пример 4. Рассмотрим задачу о нахождении решения $u(t)$ следующего интегродифференциального уравнения

$$u_t + \left(\int_0^1 u^2(\xi) d\xi \right) s(t, u) = 0$$

с граничными условиями

$$u(0) = u(1),$$

где $s(t, u)$ - некоторая заданная функция. Перевести эту задачу на язык систем (7)-(8) позволит следующая вспомогательная задача. Рассмотрим уравнение

$$u_t + ys(\xi, u) = 0 \quad (9)$$

Если заданы значения y и $x = u(0)$, задача Коши для уравнения (9) однозначно разрешима. Ее решением является функция $u(t)$. Соответственно, однозначно определено значение $u(1) \equiv f_1(x, y)$. Также однозначно определен интеграл $\int_0^1 u^2(\xi) d\xi \equiv f_2(x, y)$. Учитывая граничное условие, получаем, что решению исходной задачи соответствует решение системы

$$x = f_1(x, y), \quad y = f_2(x, y)$$

Несмотря на сложный вид функций f_1 и f_2 , не представляет труда написать программу, которая бы рисовала линии, на которых $F_1(x, y) = x - f_1(x, y)$ и $F_2(x, y) = x - f_2(x, y)$ были бы равны нулю и локализовать точки их пересечения. Далее, численно решив данную систему, необходимо найти $x = u(0)$ и восстановить $u(t)$ путем численного решения задачи Коши для (9).

После того, как решения системы (7)-(8) локализованы, для их уточнения естественно использовать итерационные методы. Однако некоторые методы, применимые для уравнения (4) оказываются неприменимы в случае системы уравнений (мы не рассматриваем случай, когда одну из переменных в одном из уравнений можно исключить и свести задачу к уравнению вида $F(x) = 0$.) Укажем методы применимые к системе (7)-(8).

1. *Различные модификации метода простых итераций.* Аналогично случаю уравнения (4), можно попытаться переписать систему (7)-(8) в виде

$$x = f_1(x, y) \quad (10)$$

$$y = f_2(x, y) \quad (11)$$

и использовать итерационный процесс

$$x^{(n+1)} = f_1(x^{(n)}, y^{(n)})$$

$$y^{(n+1)} = f_2(x^{(n)}, y^{(n)})$$

Возможны и другие итерационные процессы для решения системы (7)-(8). Например, можно попробовать использовать следующий двух-шаговый итерационный процесс. На первом шаге по заданным значениям $(x^{(n)}, y^{(n)})$ найти значение $y^{(n+1)}$, используя какой-либо численный метод для решения уравнения

$$F_1(x^{(n)}, y^{(n+1)}) = 0$$

На втором шаге из уравнения

$$F_2(x^{(n+1)}, y^{(n+1)}) = 0$$

по только что найденному значению $y^{(n+1)}$, найти и значение $x^{(n+1)}$, также используя какой-либо метод решения уравнения с одной переменной.

Задача: Предложите геометрическую интерпретацию такого алгоритма.

2. Метод Ньютона. Изложение метода Ньютона для случая систем нелинейных уравнений можно найти практически в любом учебнике численных методов (см. напр. учебник Н.С.Бахвалова “Численные методы”, гл. VII, параграф 2). Итерационный процесс в этом случае описывается формулой

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} F_{1,x}(x_n, y_n) & F_{1,y}(x_n, y_n) \\ F_{2,x}(x_n, y_n) & F_{2,y}(x_n, y_n) \end{pmatrix}^{-1} \begin{pmatrix} F_1(x_n, y_n) \\ F_2(x_n, y_n) \end{pmatrix}$$

Здесь x_n, y_n - приближение к решению на n -й итерации. Матрица в правой части представлена частными производными функций $F_1(x, y)$ и $F_2(x, y)$. Так же как и в случае одного уравнения, скорость сходимости метода Ньютона *квадратичная*, то есть этот метод сходится очень быстро, если начальное приближение находится достаточно близко от решения. На практике сходимость метода Ньютона во многом определяется именно удачным начальным приближением: если это приближение задано неудачно, быстрая сходимость, как правило, пропадает, и вообще итерационный процесс может разойтись.

Проиллюстрируем вывод итерационной формулы, приведенной выше, при помощи следующих простых нестрогих рассуждений. Пусть

$x = x^{(n)}$, $y = y^{(n)}$ является приближением к искомому решению $x = x^*$, $y = y^*$. Введем поправки $\Delta x^{(n)}$ и $\Delta y^{(n)}$, потребовав, чтобы $F_1(x^{(n)} + \Delta x^{(n)}, y^{(n)} + \Delta y^{(n)}) = 0$, $F_2(x^{(n)} + \Delta x^{(n)}, y^{(n)} + \Delta y^{(n)}) = 0$

Найти эти поправки в явном виде из этих уравнений мы, конечно, не сможем. Однако, считая, что $x^{(n)}$ и $y^{(n)}$ уже достаточно близки к искомому решению, мы можем заменить приближенно значения функций F_1 и F_2 , используя первый дифференциал

$$\begin{aligned} F_1(x^{(n)} + \Delta x^{(n)}, y^{(n)} + \Delta y^{(n)}) &\approx \\ &\approx F_1(x^{(n)}, y^{(n)}) + F_{1,x}(x^{(n)}, y^{(n)})\Delta x^{(n)} + F_{1,y}(x^{(n)}, y^{(n)})\Delta y^{(n)} = 0, \\ F_2(x^{(n)} + \Delta x^{(n)}, y^{(n)} + \Delta y^{(n)}) &\approx \\ &\approx F_2(x^{(n)}, y^{(n)}) + F_{2,x}(x^{(n)}, y^{(n)})\Delta x^{(n)} + F_{2,y}(x^{(n)}, y^{(n)})\Delta y^{(n)} = 0. \end{aligned}$$

Перепишав эту систему уравнений в матричном виде, имеем

$$\begin{pmatrix} \Delta x^{(n)} \\ \Delta y^{(n)} \end{pmatrix} = - \begin{pmatrix} F_{1,x}(x^{(n)}, y^{(n)}) & F_{1,y}(x^{(n)}, y^{(n)}) \\ F_{2,x}(x^{(n)}, y^{(n)}) & F_{2,y}(x^{(n)}, y^{(n)}) \end{pmatrix}^{-1} \begin{pmatrix} F_1(x^{(n)}, y^{(n)}) \\ F_2(x^{(n)}, y^{(n)}) \end{pmatrix}$$

Полагая $\Delta x^{(n)} = x^{(n+1)} - x^{(n)}$, $\Delta y^{(n)} = y^{(n+1)} - y^{(n)}$, получаем основную формулу ньютоновского итерационного процесса.

2.5. Системы нелинейных уравнений: общий случай. Сжимающие отображения

Пусть $\mathbf{x} \in \mathbb{R}^N$ и $\mathbf{F}(\mathbf{x})$ - отображение из \mathbb{R}^N в \mathbb{R}^N . Задача о нахождении решений уравнения $\mathbf{F}(\mathbf{x}) = 0$ при $N \geq 3$ уже не допускает достаточно простой визуализации. В этом случае для локализации решений приходится применять подчас довольно сложные методы. Вместе с тем, для численного нахождения решений можно снова использовать итерационные методы, переписав исходное уравнение в виде $\mathbf{x} = \mathbf{f}(\mathbf{x})$, где

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, \quad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(x_1, x_2, \dots, x_N) \\ f_2(x_1, x_2, \dots, x_N) \\ \vdots \\ f_N(x_1, x_2, \dots, x_N) \end{pmatrix}$$

Если все компоненты $f_k(x_1, x_2, \dots, x_N)$, $k = 1, \dots, N$ являются непрерывно дифференцируемыми функциями в некоторой области пространства $\Omega \subset \mathbb{R}^N$, то мы будем говорить, что отображение \mathbf{f} является непрерывно дифференцируемым в Ω . В этом случае для $\mathbf{x} \in \Omega$ можно определить матрицу Якоби

$$\mathbf{f}'(\mathbf{x}) = \begin{pmatrix} f_{1,x_1}(x_1, \dots, x_N) & f_{1,x_2}(x_1, \dots, x_N) & \cdots & f_{1,x_N}(x_1, \dots, x_N) \\ f_{2,x_1}(x_1, \dots, x_N) & f_{2,x_2}(x_1, \dots, x_N) & \cdots & f_{2,x_N}(x_1, \dots, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ f_{N,x_1}(x_1, \dots, x_N) & f_{N,x_2}(x_1, \dots, x_N) & \cdots & f_{N,x_N}(x_1, \dots, x_N) \end{pmatrix}$$

Здесь f_{i,x_j} - частная производная i -ой компоненты f_i по переменной x_j .

Введем следующее определение:

Будем говорить, что отображение $\mathbf{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ является *сжимающим* в области $\Omega \subset \mathbb{R}^N$, если при некотором $q < 1$ для любых $\mathbf{u}_1 \in \Omega$ и $\mathbf{u}_2 \in \Omega$ выполняется соотношение

$$\|\mathbf{f}(\mathbf{u}_2) - \mathbf{f}(\mathbf{u}_1)\| \leq q \|\mathbf{u}_2 - \mathbf{u}_1\|.$$

Докажем следующее утверждение:

Теорема 0.2. Пусть $\mathbf{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ является сжимающим в области Ω с некоторой константой q , причем $\mathbf{f}(\Omega) \subset \Omega$. Тогда

- (a) уравнение (2) имеет в Ω единственное решение \mathbf{x}^* ;
- (b) для итерационного процесса (3) с начальным значением $\mathbf{x}^{(0)} \in \Omega$ справедлива оценка

$$\|\mathbf{x}^* - \mathbf{x}^{(n)}\| \leq \frac{q^n}{1 - q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|; \quad (12)$$

- (c) итерационный процесс (3) с начальным значением $\mathbf{x}^{(0)} \in \Omega$ сходится к \mathbf{x}^* как минимум линейно.

Доказательство: Имеем

$$\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| = \|\mathbf{f}(\mathbf{x}^{(n)}) - \mathbf{f}(\mathbf{x}^{(n-1)})\| \leq q\|\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}\|$$

поэтому

$$\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| \leq q^n \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|$$

Пусть $p > n$, тогда

$$\begin{aligned} \|\mathbf{x}^{(p)} - \mathbf{x}^{(n)}\| &\leq \|\mathbf{x}^{(p)} - \mathbf{x}^{(p-1)}\| + \dots + \|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| \leq \\ &\leq q^{p-1} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| + \dots + q^n \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| \leq q^n \sum_{k=0}^{\infty} q^k \cdot \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| = \\ &= \frac{q^n}{1-q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| \end{aligned}$$

По критерию Коши последовательность $\mathbf{x}^{(n)}$ имеет предел \mathbf{x}^* . Перейдем к пределу при $p \rightarrow \infty$. Получим

$$\|\mathbf{x}^* - \mathbf{x}^{(n)}\| \leq \frac{q^n}{1-q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|$$

Последняя формула означает справедливость утверждения (б), при условии, что мы докажем, что \mathbf{x}^* является решением (2), и притом единственным. Запишем

$$\begin{aligned} \|\mathbf{x}^* - \mathbf{f}(\mathbf{x}^*)\| &\leq \|\mathbf{x}^* - \mathbf{x}^{(n+1)}\| + \|\mathbf{x}^{(n+1)} - \mathbf{f}(\mathbf{x}^*)\| = \\ &= \|\mathbf{x}^* - \mathbf{x}^{(n+1)}\| + \|\mathbf{f}(\mathbf{x}^{(n)}) - \mathbf{f}(\mathbf{x}^*)\| \leq \\ &\leq \|\mathbf{x}^* - \mathbf{x}^{(n+1)}\| + q\|\mathbf{x}^{(n)} - \mathbf{x}^*\| \leq \frac{2q^{n+1}}{1-q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|. \end{aligned}$$

Так как n - произвольно, то $\|\mathbf{x}^* - \mathbf{f}(\mathbf{x}^*)\| = 0$, то есть \mathbf{x}^* удовлетворяет (2). Далее, если имеется еще одно решение (2), $\mathbf{x}^{**} \in \Omega$, то

$$\|\mathbf{x}^* - \mathbf{x}^{**}\| = \|\mathbf{f}(\mathbf{x}^*) - \mathbf{f}(\mathbf{x}^{**})\| \leq q\|\mathbf{x}^* - \mathbf{x}^{**}\|$$

что невозможно. Таким образом, \mathbf{x}^* является единственным решением (2) в Ω , тем самым доказано (а) и, следовательно, (б). Для доказательства (с) заметим, что

$$\|\mathbf{x}^{(n+1)} - \mathbf{x}^*\| = \|\mathbf{f}(\mathbf{x}^{(n)}) - \mathbf{f}(\mathbf{x}^*)\| \leq q\|\mathbf{x}^{(n)} - \mathbf{x}^*\|$$

Теорема доказана. ■

Таким образом, если итерационный процесс порожден сжимающим отображением, то он сходится (обратное, вообще говоря, неверно, см. задачи после этой главы). Приведем достаточное условие того, что отображение является сжимающим.

Теорема 0.3. Пусть $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ непрерывно дифференцируемое отображение, уравнение $x = f(x)$ имеет решение $x = x^*$, причем $\|f'(x^*)\| = q < 1$. Тогда существует некоторая окрестность точки x^* , в которой f является сжимающим.

Доказательство: По теореме об оценке конечных приращений (см. напр. Г.Е.Шилов, Математический анализ, функции нескольких вещественных переменных ч.1-2, Москва, Наука, 1972, формула параграф 1.4 формула (3)), в любой выпуклой окрестности Ω точки $x = x^*$ для $u, v \in \Omega$ имеем

$$\|f(u) - f(v)\| = \sup_{w \in \Omega} \|f'(w)\| \cdot \|u - v\|$$

Так как $\|f'(x^*)\| = q < 1$, то существует область Ω_0 , такая, что где $\|f'(w^*)\| < \tilde{q} < 1$, $w \in \Omega_0$, $\tilde{q} < 1$. Теорема доказана. ■

Можно доказать и утверждение, “дополнительное” к приведенной теореме. Оказывается, что если норма матрицы Якоби f' в точке $x = x^*$, соответствующей решению, больше единицы, ни в какой окрестности этой точки отображение *не будет сжимающим*. Это простое утверждение, его предлагается доказать самостоятельно. Стоит отметить, что если изображение не является сжимающим, это еще не означает, что итерационный процесс расходится (см задачу 6 ниже). Вместе с тем, сжимающие отображения, конечно, предпочтительнее.

Суммируем сказанное. Для нахождения решения x^* уравнения (2) можно использовать итерационный процесс вида (3). Если $\|f'(x^*)\| < 1$, то в некоторой окрестности точки x^* он будет сходиться к решению. Оценка погрешности найденного решения x^* дается формулой (12). Если же удастся выяснить, что $\|f'(x^*)\| > 1$, то выбранное представление (3) может оказаться неудачным, так как отображение f не является сжимающим.

2.6. Системы нелинейных уравнений: общий случай. Метод Ньютона

Метод Ньютона для общего случая системы N уравнений (1) имеет вид итерационного процесса

$$\mathbf{x}_{n+1} = \mathbf{x}_n - [\mathbf{F}'(\mathbf{x}_n)]^{-1} \mathbf{F}(\mathbf{x}_n), \quad \mathbf{x}_n = \begin{pmatrix} x_n^{(1)} \\ \vdots \\ x_n^{(m)} \end{pmatrix}$$

Здесь $\mathbf{F}'(\mathbf{x}_n)$ - матрица Якоби, имеющая вид

$$\mathbf{F}'(\mathbf{x}_n) = \begin{pmatrix} F_{1,x^{(1)}}(x^{(1)}, \dots, x^{(m)}) & \dots & F_{1,x^{(m)}}(x^{(1)}, \dots, x^{(m)}) \\ \vdots & \ddots & \vdots \\ F_{m,x^{(1)}}(x^{(1)}, \dots, x^{(m)}) & \dots & F_{m,x^{(m)}}(x^{(1)}, \dots, x^{(m)}) \end{pmatrix}$$

Для нестрогого обоснования метода Ньютона можно использовать рассуждения, аналогичные приведенным в разделе 2.4. Приведем строгий результат о сходимости метода Ньютона в общем случае.

Пусть $\Omega_a = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^*\| \leq a\}$. Пусть при некоторых $a > 0$, $0 \leq a_1, a_2 < \infty$ выполняются условия

$$1. \quad \|(\mathbf{F}'(\mathbf{x}))^{-1}\| \leq a_1, \quad \text{при } \mathbf{x} \in \Omega_a \quad (13)$$

$$2. \quad \|\mathbf{F}(\mathbf{u}_1) - \mathbf{F}(\mathbf{u}_2) - \mathbf{F}'(\mathbf{u}_2)(\mathbf{u}_1 - \mathbf{u}_2)\| \leq a_2 \|\mathbf{u}_1 - \mathbf{u}_2\|^2, \quad (14)$$

$$\mathbf{u}_1, \mathbf{u}_2 \in \Omega_a$$

Справедлива следующая теорема:

Теорема 0.4. Пусть выполняются условия (13)-(14) и $c = a_1 a_2$. Тогда при $\mathbf{x}^{(0)} \in \Omega_b$ где $b = \min(a, 1/c)$ итерационный процесс Ньютона сходится с оценкой погрешности

$$\|\mathbf{x}^{(n)} - \mathbf{x}^*\| \leq c^{-1} \left(c \|\mathbf{x}^{(0)} - \mathbf{x}^*\| \right)^{2^n}$$

Доказательство этой теоремы можно найти, например, в книге Н.С.Бахвалова “Численные методы”, Москва, Наука, глава VII, параграф 2.

Задачи

1. Проиллюстрировать графически, как будет сходиться алгоритм простых итераций (с наиболее очевидным выбором функции $f(x)$) в случае уравнений

$$(a) \ x = \operatorname{arctg} 2x, \quad (b) \ x = x^2, \quad (c) \ x = 4/x$$

В каждом из этих уравнений имеется корень, к которому алгоритм простых итераций не сходится. Как можно модифицировать этот алгоритм, чтобы сходимость к этому корню появилась?

2. Доказать сходимость метода простых итераций для указанного диапазона $x^{(0)}$ и найти $\lim_{n \rightarrow \infty} x^{(n)}$.

$$(a) \ x^{(n+1)} = x^{(n)} + \sin x^{(n)}$$

$$(b) \ x^{(n+1)} = \frac{x^{(n)}((x^{(n)})^2 + 3)}{3(x^{(n)})^2 + 1}; \quad x^{(0)} \in \mathbb{R}$$

$$(c) \ x^{(n+1)} = \frac{1}{2} \left(x^{(n)} + \frac{1}{x^{(n)}} \right); \quad x^{(0)} > 0$$

$$(d) \ x^{(n+1)} = \frac{2}{\pi} \arcsin \frac{2x^{(n)}}{(x^{(n)})^2 + 1}; \quad x^{(0)} > 0$$

3. Выяснить, для каких начальных значений $x^{(0)}$ метод простых итераций сходится, и найти $\lim_{n \rightarrow \infty} x^{(n)}$.

$$(a) \ x^{(n+1)} = x^{(n)} + \sin x^{(n)}$$

$$(b) \ x^{(n+1)} = x^{(n)}(x^{(n)} - 1)$$

$$(c) \ x^{(n+1)} = (x^{(n)})^2 + c, \quad c \leq 1/4$$

4. Для итерационного процесса

$$x^{(n+1)} = \sin x^{(n)}$$

выясните (i) является ли он сходящимся; (ii) является ли соответствующее отображение сжимающим.

5. Нарисовать на плоскости множества $\|\mathbf{u}\|_\infty = 1$, $\|\mathbf{u}\|_1 = 1$, $\|\mathbf{u}\|_2 = 1$ (единичные сферы \mathbb{S}^1 для трех различных норм).

6. Рассмотрим в \mathbb{R}^2 линейное преобразование \mathbf{A} , заданное матрицей.

$$\mathbf{A} = \begin{pmatrix} 0.9 \cos \frac{\pi}{4} & -0.9 \sin \frac{\pi}{4} \\ 0.9 \sin \frac{\pi}{4} & 0.9 \cos \frac{\pi}{4} \end{pmatrix}$$

(a) Выясните геометрический смысл этого преобразования.

(б) Выясните, чему равна норма оператора \mathbf{A} , если \mathbb{R}^2 снабжено каждой из норм $\|\cdot\|_\infty$, $\|\cdot\|_1$, $\|\cdot\|_2$.

(в) Выясните, является ли \mathbf{A} сжимающим в каждой из этих норм.

7. Выяснить число комплексных корней и локализовать их для уравнений

$$(a) \cos z = z, \quad (b) e^z = z$$

8. Выяснить число корней и локализовать их для систем уравнений

$$(a) x^2 - 3xy + y^2 = 0, \quad e^x + e^y = 3$$

$$(b) y^4 + x^4 - 4xy = 1, \quad x^2 + y^2 = 1$$

$$(c) x^2 + xy + x^2 = \frac{3}{2}, \quad e^{x+y} = 1 + x + y$$

9. Найти достаточное условие на функции $f(x)$ и $g(x)$, при которых метод простых итераций

$$x^{(n+1)} = x^{(n)} f(y^{(n)})$$

$$y^{(n+1)} = y^{(n)} g(x^{(n)})$$

будет сходиться к решению $x = 0$, $y = 0$.

10. Найти достаточное условие на функции $f(x)$ и $g(x)$, при которых метод простых итераций

$$x^{(n+1)} = y^{(n)} f(x^{(n)})$$

$$y^{(n+1)} = x^{(n)} g(y^{(n)})$$

будет сходиться к решению $x = 0, y = 0$.

11. Аналогично примеру 4, свести к системе $F_1(x, y) = 0, F_2(x, y) = 0$ следующие задачи:

(а) $u_{tt} + f(t, u) = 0, u(0) = u(2\pi), u_t(0) = u_t(2\pi), f(t, u)$ - функция, периодическая по t с периодом 2π .

(б) $u_{tt} + u_t + g\left(\frac{1}{2\pi} \int_0^{2\pi} u(\xi) d\xi\right) = 0, u(0) = 0, u(2\pi) = 1, g(\cdot)$ - некоторая функция.

12. Докажите, что отображение комплексной плоскости $w = e^z$ не является сжимающим ни в какой окрестности ни одного из комплексных корней уравнения $z = e^z$.

13. Пусть отображение

$$x^{(n+1)} = f_1(x^{(n)}, y^{(n)})$$

$$y^{(n+1)} = f_2(x^{(n)}, y^{(n)})$$

является сжимающим в некоторой области $\Omega \subset \mathbb{R}^2$. Докажите, что существуют промежутки $(a; b)$ и $(c; d)$, такие, что для любого $\tilde{y} \in (c; d)$ отображение

$$x^{(n+1)} = f_1(x^{(n)}, \tilde{y})$$

является сжимающим на промежутке $x \in (a; b)$.