

ГЛАВА 5. МЕТОДЫ РЕШЕНИЯ СИСТЕМ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

§ 1. ПОСТАНОВКА ЗАДАЧИ

Пусть требуется решить систему уравнений

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0, \\ f_2(x_1, x_2, \dots, x_n) = 0, \\ \dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) = 0, \end{cases} \quad (5.1)$$

где f_1, f_2, \dots, f_n – заданные, вообще говоря, нелинейные (среди них могут быть и линейные) вещественнозначные функции n вещественных переменных x_1, x_2, \dots, x_n .

Обозначив

$$\mathbf{x} = (x_1, x_2, \dots, x_n), \mathbf{F}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \dots \\ f_n(\mathbf{x}) \end{pmatrix}, \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix},$$

данную систему можно записать одним уравнением

$$\mathbf{F}(\mathbf{x}) = \mathbf{0} \quad (5.2)$$

относительно векторной функции \mathbf{F} векторного аргумента \mathbf{x} .

Таким образом, исходную задачу можно рассматривать как задачу о нулях нелинейного отображения $\mathbf{F}: R_n \rightarrow R_n$. В этой постановке она является прямым обобщением задачи главы 1 – задачи построения методов нахождения нулей одномерных нелинейных отображений. Фактически это та же задача, только в пространствах большей размерности. Поэтому можно как заново строить методы ее решения на основе разработанных выше подходов, так и осуществлять формальный перенос выведенных для скалярного случая расчетных формул. В любом случае следует позаботиться о правомочности тех или иных операций над векторными переменными и векторными функциями, а также о сходимости получаемых таким способом итерационных процессов. Не все результаты и методы можно перенести со случая $n = 1$ на случай $n \geq 2$, в то же время подобный переход вносит в задачу нахождения нулей нелинейного отображения свою специфику, что в совокупности приводит к новым методам и модификациям.

§ 2. МЕТОДЫ РЕШЕНИЯ СИСТЕМ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

2.1. Метод простых итераций

Метод простых итераций является наиболее простым методом решения системы нелинейных уравнений.

Преобразуем систему (5.1) к виду

$$\begin{cases} x_1 = \varphi_1(x_1, x_2, \dots, x_n), \\ x_2 = \varphi_2(x_1, x_2, \dots, x_n), \\ \dots\dots\dots \\ x_n = \varphi_n(x_1, x_2, \dots, x_n), \end{cases} \quad (5.3)$$

или иначе, в компактной записи

$$\mathbf{x} = \Phi(\mathbf{x}). \quad (5.4)$$

Для этой задачи о неподвижной точке нелинейного отображения $\Phi: R_n \rightarrow R_n$ запишем формально рекуррентное равенство

$$\mathbf{x}^{(k+1)} = \Phi(\mathbf{x}^{(k)}), k = 0, 1, 2, \dots, \quad (5.5)$$

которое определяет метод простых итераций (МПИ) (или метод последовательных приближений) для задачи (5.3).

Если начать процесс построения последовательности $(\mathbf{x}^{(k)})$ с некоторого вектора $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T$ (начальное приближение можно выбрать, например, графически) и продолжить по формуле (5.5), то при определенных условиях эта последовательность со скоростью геометрической прогрессии будет приближаться к вектору \mathbf{x}^* – неподвижной точке отображения $\Phi(\mathbf{x})$ (предполагается, что решение существует). Признаком того, что отображение $\Phi: R_n \rightarrow R_n$ является сжимающим, сформулированный в теореме 3.1 главы 3 для скалярной функции, т. е. при $n=1$, в общем случае формулируется несколько сложнее. Сформулируем первое и второе достаточные условия сходимости процесса итераций.

Определение 5.1. Область U пространства R_n называется выпуклой, если наряду с любыми двумя точками $\mathbf{a} = (a_1, a_2, \dots, a_n)$ и $\mathbf{b} = (b_1, b_2, \dots, b_n)$, принадлежащими области U , ей принадлежат и все точки отрезка \mathbf{ab} , содержащего эти точки.

Предполагается, что вектор-функции $\Phi(\mathbf{x})$ и $\Phi'(\mathbf{x})$ определены и непрерывны в выпуклой ограниченной замкнутой области $G \subset E_n$.

Будем пользоваться двумя нормами $\|\mathbf{x}\|_m = \max_i |x_i|$ и $\|\mathbf{x}\|_l = \sum_{i=1}^n |x_i|$.

Относительно области G введем две нормы $\|\Phi'(\mathbf{x})\|_l = \max_{\mathbf{x} \in G} \|\Phi'(\mathbf{x})\|_m$ и $\|\Phi'(\mathbf{x})\|_m = \max_{\mathbf{x} \in G} \|\Phi'(\mathbf{x})\|_l$, где $\|\Phi'(\mathbf{x})\|_m = \max_i \sum_{j=1}^n \left| \frac{\partial \varphi_i(\mathbf{x})}{\partial x_j} \right|$ и $\|\Phi'(\mathbf{x})\|_l = \max_j \sum_{i=1}^n \left| \frac{\partial \varphi_i(\mathbf{x})}{\partial x_j} \right|$ соответственно.

Теорема 5.1. Пусть функции $\Phi(\mathbf{x})$ и $\Phi'(\mathbf{x})$ непрерывны в области G , причем в G выполнено одно из неравенств $\|\Phi'(\mathbf{x})\|_l \leq q < 1$ или $\|\Phi'(\mathbf{x})\|_m \leq q < 1$, где q – некоторая постоянная. Если $\mathbf{x}^{(k)} \in G$ ($k = 0, 1, 2, \dots$), то итерационный процесс (5.5) сходится и предельный вектор $\mathbf{x}^* = \lim_{k \rightarrow \infty} \mathbf{x}^{(k)}$ является в области G единственным решением системы (5.4).

Следствие 5.1. Процесс итерации (5.5) сходится, если при $\mathbf{x} \in G$ справедливы неравенства

$$\sum_{j=1}^n \left| \frac{\partial \varphi_i(\mathbf{x})}{\partial x_j} \right| \leq q_i < 1 \text{ или } \sum_{j=1}^n \left| \frac{\partial \varphi_j(\mathbf{x})}{\partial x_i} \right| \leq q_i < 1 \quad (i = 1, 2, \dots, n).$$

Таким образом, для вычисления решения \mathbf{x} уравнения $\mathbf{F}(\mathbf{x}) = 0$ методом простых итераций надо стараться выбрать равносильное уравнение $\mathbf{x} = \Phi(\mathbf{x})$ так, чтобы отображение в области G , содержащей решение \mathbf{x} , было сжимающим с коэффициентом сжатия q , как можно более близким к нулю. Существуют различные способы приведения, рассмотрим один из них. Умножим, например, (5.2) на некоторую неособенную $n \times n$ -матрицу \mathbf{A} и прибавим к обеим частям уравнения $\mathbf{AF}(\mathbf{x}) = 0$ вектор неизвестных \mathbf{x} . Полученная система $\mathbf{x} = \mathbf{x} + \mathbf{AF}(\mathbf{x})$

эквивалентна исходной системе и имеет вид задачи о неподвижной точке (5.4).

Проблема теперь состоит лишь в подборе матричного параметра \mathbf{A} такого, при котором вектор-функция $\Phi(\mathbf{x}) = \mathbf{x} + \mathbf{AF}(\mathbf{x})$ обладала бы нужными свойствами.

Из теоремы 5.1 следует, что процесс итерации для уравнения (5.4) быстро сходится, если $\Phi'(\mathbf{x})$ мала по норме. Учитывая это обстоятельство, выбираем матрицу \mathbf{A} так, чтобы $\Phi'(\mathbf{x}^{(0)}) = \mathbf{E} + \mathbf{A}\mathbf{F}'(\mathbf{x}^{(0)}) = \mathbf{0}$. Отсюда, если $\mathbf{F}'(\mathbf{x}^{(0)})$ – неособенная, имеем $\mathbf{A} = -[\mathbf{F}'(\mathbf{x}^{(0)})]^{-1}$. В случае, если $\det \mathbf{F}'(\mathbf{x}^{(0)}) = 0$, то следует выбрать другое начальное приближение $\mathbf{x}^{(0)}$.

Употребляются также и иные способы замены системы (5.2) эквивалентной ей системой (5.4).

Систему (5.3) можно решать и методом Зейделя, напоминающим метод Гаусса-Зейделя решения систем линейных уравнений (см. гл. 4). Значение $x_i^{(k)}$ находится из i -го уравнения системы (5.3) с использованием уже вычисленных на текущей итерации значений неизвестных. Таким образом, значения неизвестных на k -й итерации будут находиться с помощью соотношения

$$x_i^{(k)} = \varphi_i(x_1^{(k)}, \dots, x_{i-1}^{(k)}, x_i^{(k-1)}, \dots, x_n^{(k-1)}), i = 1, 2, \dots, n.$$

Итерационный процесс в обоих методах продолжается до тех пор, пока изменения всех неизвестных в двух последовательных итерациях не станут малыми.

При использовании метода простой итерации и метода Зейделя успех во многом определяется удачным выбором начальных приближений неизвестных: они должны быть достаточно близкими к истинному решению. В противном случае итерационный процесс может не сойтись.

Пример 5.1. Методом итерации приближенно решить систему

$$\begin{cases} x_1^2 + x_2^2 = 1, \\ x_1^3 - x_2 = 0. \end{cases}$$

Решение. Из графического построения видно (рис. 5.1), что данная система имеет два решения, отличающихся только знаком. Ограничимся нахождением положительного решения.

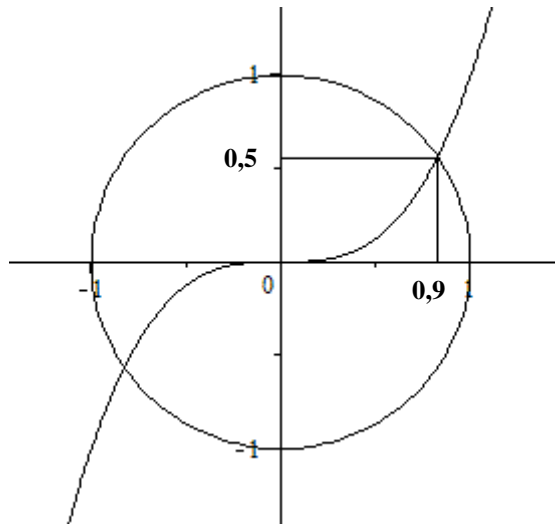


Рис. 5.1. Графическая локализация корней системы $\begin{cases} x_1^2 + x_2^2 = 1, \\ x_1^3 - x_2 = 0. \end{cases}$

Из чертежа видно, что за начальное приближение положительного решения системы можно принять:

$$\mathbf{x}^{(0)} = \begin{bmatrix} 0,9 \\ 0,5 \end{bmatrix}.$$

Полагая $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} x_1^2 + x_2^2 - 1 \\ x_1^3 - x_2 \end{bmatrix}$, получим $\mathbf{F}'(\mathbf{x}) = \begin{bmatrix} 2x_1 & 2x_2 \\ 3x_1 & -1 \end{bmatrix}$.

Отсюда $\mathbf{F}'(\mathbf{x}^{(0)}) = \begin{bmatrix} 1,8 & 1 \\ 2,43 & -1 \end{bmatrix}$ и $\det \mathbf{F}'(\mathbf{x}^{(0)}) = -4,23$. Так как матрица $\mathbf{F}'(\mathbf{x}^{(0)})$ –

неособенная, то существует обратная матрица $[\mathbf{F}'(\mathbf{x}^{(0)})]^{-1} = -\frac{1}{4,23} \begin{bmatrix} -1 & -1 \\ -2,43 & 1,8 \end{bmatrix}$. Таким

образом, $\mathbf{A} = -[\mathbf{F}'(\mathbf{x}^{(0)})]^{-1} = \frac{1}{4,23} \begin{bmatrix} -1 & -1 \\ -2,43 & 1,8 \end{bmatrix}$.

Полагая $\boldsymbol{\varphi}(\mathbf{x}) = \mathbf{x} + \mathbf{A}\mathbf{F}(\mathbf{x}) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{4,23} \begin{bmatrix} 1 & 1 \\ 2,43 & -1,8 \end{bmatrix} \cdot \begin{bmatrix} x_1^2 + x_2^2 - 1 \\ x_1^3 - x_2 \end{bmatrix}$, перейдем к

эквивалентной системе $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{x})$.

Пользуясь формулой (5.5), находим для решения системы последовательные приближения

$$\mathbf{x}^{(1)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix} - \frac{1}{4,23} \begin{bmatrix} 1 & 1 \\ 2,43 & -1,8 \end{bmatrix} \cdot \begin{bmatrix} x_1^{(0)2} + x_2^{(0)2} - 1 \\ x_1^{(0)3} - x_2^{(0)} \end{bmatrix} = \begin{bmatrix} 0,8317 \\ 0,5630 \end{bmatrix};$$

$$\mathbf{x}^{(2)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} - \frac{1}{4,23} \begin{bmatrix} 1 & 1 \\ 2,43 & -1,8 \end{bmatrix} \cdot \begin{bmatrix} x_1^{(1)2} + x_2^{(1)2} - 1 \\ x_1^{(1)3} - x_2^{(1)} \end{bmatrix} = \begin{bmatrix} 0,8268 \\ 0,5633 \end{bmatrix};$$

$$\mathbf{x}^{(3)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \end{bmatrix} - \frac{1}{4,23} \begin{bmatrix} 1 & 1 \\ 2,43 & -1,8 \end{bmatrix} \cdot \begin{bmatrix} x_1^{(2)2} + x_2^{(2)2} - 1 \\ x_1^{(2)3} - x_2^{(2)} \end{bmatrix} = \begin{bmatrix} 0,8261 \\ 0,5631 \end{bmatrix};$$

$$\mathbf{x}^{(4)} = \begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \end{bmatrix} - \frac{1}{4,23} \begin{bmatrix} 1 & 1 \\ 2,43 & -1,8 \end{bmatrix} \cdot \begin{bmatrix} x_1^{(3)2} + x_2^{(3)2} - 1 \\ x_1^{(3)3} - x_2^{(3)} \end{bmatrix} = \begin{bmatrix} 0,8261 \\ 0,5636 \end{bmatrix}$$

и т.д.

Ограничиваясь четвертым приближением, имеем корни $x_1 = 0,8261$ и $x_2 = 0,5636$ с точностью до 10^{-4} .

2.2. Метод Ньютона

Рассмотрим систему нелинейных уравнений (5.2) $\mathbf{F}(\mathbf{x}) = 0$. Для решения данной системы будем пользоваться методом последовательных приближений. Предположим, что найдено k -е приближение

$$\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$$

одного из изолированных корней $\mathbf{x} = (x_1, x_2, \dots, x_n)$ векторного уравнения (5.2). Тогда точный корень уравнения (5.2) можно представить в виде

$$\mathbf{x} = \mathbf{x}^{(k)} + \boldsymbol{\varepsilon}^{(k)}, \quad (5.6)$$

где $\boldsymbol{\varepsilon}^{(k)} = (\varepsilon_1^{(k)}, \varepsilon_2^{(k)}, \dots, \varepsilon_n^{(k)})$ – поправка (погрешность корня).

Подставив выражение (5.6) в уравнение (5.2), получим

$$\mathbf{F}(\mathbf{x}^{(k)} + \boldsymbol{\varepsilon}^{(k)}) = 0. \quad (5.7)$$

Предполагая, что функция $\mathbf{F}(\mathbf{x})$ непрерывно дифференцируема в некоторой выпуклой области, содержащей \mathbf{x} и $\mathbf{x}^{(k)}$, разложим левую часть уравнения (5.7) по степеням малого вектора $\boldsymbol{\varepsilon}^{(k)}$, ограничиваясь линейными членами,

$$\mathbf{F}(\mathbf{x}^{(k)} + \boldsymbol{\varepsilon}^{(k)}) = \mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{F}'(\mathbf{x}^{(k)})\boldsymbol{\varepsilon}^{(k)} = 0. \quad (5.8)$$

Система (5.8) представляет собой линейную систему относительно поправок $\varepsilon_i^{(k)}$ ($i = 1, 2, \dots, n$). Под производной $\mathbf{F}(\mathbf{x})$ следует понимать матрицу Якоби системы функций $\{f_i\}_1^n$ относительно переменных x_1, x_2, \dots, x_n , т. е.

$$\mathbf{F}'(\mathbf{x}) = \mathbf{W}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} = \left[\frac{\partial f_i}{\partial x_j} \right] \quad (i, j = 1, 2, \dots, n).$$

Перепишем систему (5.8) в виде $\mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{W}(\mathbf{x}^{(k)})\boldsymbol{\varepsilon}^{(k)} = 0$. Отсюда, предполагая, что матрица $\mathbf{W}(\mathbf{x})$ – неособенная, получим $\boldsymbol{\varepsilon}^{(k)} = -[\mathbf{W}(\mathbf{x}^{(k)})]^{-1}\mathbf{F}(\mathbf{x}^{(k)})$. Следовательно,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - [\mathbf{W}(\mathbf{x}^{(k)})]^{-1}\mathbf{F}(\mathbf{x}^{(k)}) \quad (k = 0, 1, 2, \dots). \quad (5.9)$$

Итерационный процесс (5.9) называется методом Ньютона.

Отметим, что в достаточно малой окрестности итерации сходятся, если $\det \mathbf{W}(\mathbf{x}) \neq 0$, причем сходимость квадратичная. Следовательно, если нулевое приближение выбрано удачно, то метод Ньютона сходится, причем очень быстро. Поэтому на практике этот метод используют чаще всего.

Отрицательным фактором, осложняющим применение метода Ньютона к n -мерному случаю, является необходимость решения n -мерных линейных задач на каждом шаге итерации (обращение матрицы Якоби), требующего в зависимости от n значительных вычислительных затрат. Уменьшение таких затрат – одно из направлений модификации метода Ньютона.

Если матрицу Якоби $\mathbf{W}(\mathbf{x})$ вычислить и обратить лишь один раз – в начальной точке $\mathbf{x}^{(0)}$, то от метода Ньютона (5.9) придем к модифицированному методу Ньютона

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - [\mathbf{W}(\mathbf{x}^{(0)})]^{-1}\mathbf{F}(\mathbf{x}^{(k)}) \quad (k = 0, 1, 2, \dots). \quad (5.10)$$

Этот метод требует гораздо меньших вычислительных затрат на один итерационный шаг по сравнению с основным методом Ньютона, однако, являясь частным случаем МПИ, имеет лишь скорость сходимости геометрической прогрессии.

Существует еще множество различных модификаций метода Ньютона.

Пример 5.2. Рассмотрим применение метода Ньютона на примере системы двух нелинейных уравнений

$$\begin{cases} f_1(x_1, x_2) = x_1 + 3 \lg x_1 - x_2^2 = 0, \\ f_2(x_1, x_2) = 2x_1^2 - x_1x_2 - 5x_1 + 1 = 0. \end{cases}$$

Решение. Графическое решение этой системы дает две точки пересечения $M_1(1,4; -1,5)$ и $M_2(3,4; 2,2)$. Зададим начальное приближение $\mathbf{x}^{(0)} = \begin{bmatrix} 3,4 \\ 2,2 \end{bmatrix}$. Составим матрицу Якоби

$$\mathbf{W}(\mathbf{x}) = \begin{bmatrix} 1 + \frac{3}{x_1 \ln 10} & -2x_2 \\ 4x_1 - x_2 - 5 & -x_1 \end{bmatrix}.$$

Построим итерационный процесс по формуле (5.9)

$$\mathbf{W}(\mathbf{x}^{(0)}) = \begin{bmatrix} 1,3832 & -4,4 \\ 6,4 & -3,4 \end{bmatrix}; \det \mathbf{W}(\mathbf{x}^{(0)}) = 23,457; [\mathbf{W}(\mathbf{x}^{(0)})]^{-1} = \frac{1}{23,457} \begin{bmatrix} -3,4 & 4,4 \\ 6,4 & 1,3832 \end{bmatrix};$$

$$\mathbf{x}^{(1)} = \begin{bmatrix} 3,4 \\ 2,2 \end{bmatrix} - \frac{1}{23,457} \begin{bmatrix} -3,4 & 4,4 \\ 6,4 & 1,3832 \end{bmatrix} \begin{bmatrix} 0,1544 \\ -0,36 \end{bmatrix} = \begin{bmatrix} 3,4899 \\ 2,2633 \end{bmatrix}.$$

Аналогично получим:

$$\mathbf{x}^{(2)} = \begin{bmatrix} 3,4891 \\ 2,2621 \end{bmatrix}; \mathbf{x}^{(3)} = \begin{bmatrix} 3,4875 \\ 2,2616 \end{bmatrix}.$$

Остановливаясь на 3-м приближении, имеем $x_1 = 3,4875$ и $x_2 = 2,2616$, причем вектор невязки $\mathbf{F}(\mathbf{x}^{(3)}) = \begin{bmatrix} 0,0002 \\ 0,0000 \end{bmatrix}$.

2.3. Метод Брауна

В отличие от пошаговой линеаризации векторной функции $\mathbf{F}(\mathbf{x})$, приведшей к методу Ньютона (5.9), Брауном (1966) предложено проводить поочередную линеаризацию компонент вектор-функции $\mathbf{F}(\mathbf{x})$, линеаризовать в системе (5.1) сначала f_1 , затем f_2 и т.д., и последовательно решать получаемые таким образом уравнения. Рассмотрим данный метод в двумерном случае.

Пусть требуется найти решение системы

$$\begin{cases} f(x, y) = 0, \\ g(x, y) = 0, \end{cases} \quad (5.11)$$

и пусть уже получены приближения x_k и y_k .

Заменим первое уравнение системы (5.11) линейным, полученным по формуле Тейлора для функции двух переменных:

$$f(x, y) \approx f(x_k, y_k) + f'_x(x_k, y_k)(x - x_k) + f'_y(x_k, y_k)(y - y_k) = 0.$$

Отсюда выражаем x (обозначим этот результат через \tilde{x}):

$$\tilde{x} = x_k - \frac{1}{f'_x(x_k, y_k)} \left[f(x_k, y_k) + f'_y(x_k, y_k)(y - y_k) \right]. \quad (5.12)$$

При $y = y_k$ находим значение \tilde{x}_k переменной \tilde{x} :

$$\tilde{x} = x_k - \frac{f(x_k, y_k)}{f'_x(x_k, y_k)}.$$

Подставив в $g(x, y)$ вместо x переменную $\tilde{x} = \tilde{x}(y)$ придем к некоторой функции $G(y) = g(\tilde{x}, y)$ одной переменной y . Это позволяет линеаризовать второе уравнение системы (5.11) с помощью формулы Тейлора для функции одной переменной:

$$g(\tilde{x}, y) \approx G(y_k) + G'(y_k)(y - y_k) = 0. \quad (5.13)$$

При нахождении производной $G'(y)$ необходимо учесть, что $G(y)$ – сложная функция

($G(y) = g(\tilde{x}(y), y)$):

$$G'(y) = g'_x(\tilde{x}, y) \cdot \tilde{x}' + g'_y(\tilde{x}, y). \quad (5.14)$$

Дифференцируя по y равенство (5.12), получим

$$\tilde{x}'_y = -\frac{f'_y(x_k, y_k)}{f'_x(x_k, y_k)}.$$

Подстановка последнего в (5.14) при $y = y_k$ и $x = x_k$ даст

$$G'(y_k) = -g'_x(\tilde{x}_k, y_k) \cdot \frac{f'_y(x_k, y_k)}{f'_x(x_k, y_k)} + g'_y(\tilde{x}_k, y_k).$$

При известных значениях $G(y_k) = g(\tilde{x}_k, y_k)$ и $G'(y_k)$ можно разрешить линейное уравнение (5.13) относительно y :

$$y_{k+1} = y_k - \frac{G(y_k)}{G'(y_k)} = y_k - \frac{g(\tilde{x}_k, y_k) f'_x(x_k, y_k)}{f'_x(x_k, y_k) g'_y(\tilde{x}_k, y_k) - f'_y(x_k, y_k) g'_x(\tilde{x}_k, y_k)}.$$

Заменяя в (5.12) y полученным значением y_{k+1} , придем к значению x_{k+1} :

$$x_{k+1} = \tilde{x}(y_{k+1}) = x_k - \frac{1}{f'_x(x_k, y_k)} \left[f(x_k, y_k) + f'_y(x_k, y_k)(y_{k+1} - y_k) \right].$$

Таким образом, реализация метода Брауна решения двумерных нелинейных систем вида (5.11) сводится к следующему.

При выбранных начальных значениях x_0, y_0 каждое последующее приближение по методу Брауна находится с помощью совокупности формул

$$\begin{aligned} \tilde{x} &= x_k - \frac{f(x_k, y_k)}{f'_x(x_k, y_k)}, \quad q_k = \frac{g(\tilde{x}_k, y_k) f'_x(x_k, y_k)}{f'_x(x_k, y_k) g'_y(\tilde{x}_k, y_k) - f'_y(x_k, y_k) g'_x(\tilde{x}_k, y_k)}, \\ p_k &= \frac{f(x_k, y_k) - q_k f'_y(x_k, y_k)}{f'_x(x_k, y_k)}, \quad x_{k+1} = x_k - p_k, \quad y_{k+1} = y_k - q_k. \end{aligned}$$

Вычисления в методе Брауна естественно заканчивать, когда выполнится неравенство $\max\{|p_k|, |q_k|\} < \varepsilon$ с результатом $(x^*, y^*) = (x_k, y_k)$. В ходе вычислений следует контролировать величину знаменателей расчетных формул. Заметим, что функции f и g в этом методе неравноправны, и перемена их ролями может изменить ситуацию со сходимостью.

Метод Брауна обладает квадратичной сходимостью. Ждать большей по сравнению с методом Ньютона эффективности можно лишь, заменяя фигурирующие в формулах частные производные разностными соотношениями.

2.4. Метод скорейшего спуска (метод градиента)

Общий недостаток всех рассмотренных выше методов решения систем нелинейных уравнений состоит в сугубо локальном характере сходимости, затрудняющем их применение в случаях, когда имеются проблемы с выбором хороших начальных приближений. Помощь здесь может прийти со стороны численных методов оптимизации. Для этого нужно поставить задачу нахождения решений данной нелинейной системы как оптимизационную или, иначе, экстремальную задачу.

Рассмотрим систему нелинейных уравнений (5.2). Система нелинейных уравнений вида (5.2) эквивалентна системе

$$\Phi(\mathbf{x}) = 0, \tag{5.15}$$

где $\Phi(\mathbf{x}) = \sum_{i=1}^n f_i^2(\mathbf{x}) = (f(\mathbf{x}), f(\mathbf{x}))$. Очевидно, решениями уравнения (5.15) являются точки нулевых минимумов функции $\Phi(\mathbf{x}) = \Phi(x_1, x_2, \dots, x_n)$.

Будем предполагать, что система (5.2) имеет лишь изолированное решение, которое представляет собой точку строгого минимума функции $\Phi(\mathbf{x})$. Пусть \mathbf{x}^* – вектор-корень системы (5.2) и $\mathbf{x}^{(0)}$ – его нулевое приближение. Через точку $\mathbf{x}^{(0)}$ проведем поверхность уровня функции $\Phi(\mathbf{x})$. Из точки $\mathbf{x}^{(0)}$ двигаемся по нормали к поверхности $\Phi(\mathbf{x}) = \Phi(\mathbf{x}^{(0)})$ до тех пор, пока эта нормаль не коснется в некоторой точке $\mathbf{x}^{(1)}$ какой-то другой поверхности уровня (рис. 5.2). Затем, отправляясь от точки $\mathbf{x}^{(1)}$, снова двигаемся по нормали к поверхности уровня $\Phi(\mathbf{x}) = \Phi(\mathbf{x}^{(1)})$ до тех пор, пока эта нормаль не коснется в некоторой точке $\mathbf{x}^{(2)}$ новой поверхности уровня $\Phi(\mathbf{x}) = \Phi(\mathbf{x}^{(2)})$ и т. д. Так как $\Phi(\mathbf{x}^{(0)}) > \Phi(\mathbf{x}^{(1)}) > \Phi(\mathbf{x}^{(2)}) > \dots$, то, двигаясь по такому пути, мы приближаемся к точке с наименьшим значением $\Phi(\mathbf{x})$ (дно «ямы»), которая соответствует искомому корню \mathbf{x}^* системы.

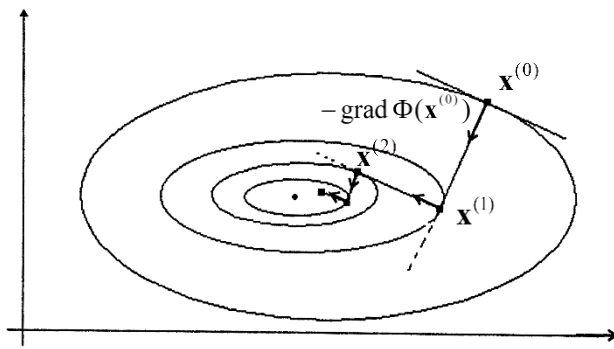


Рис. 5.2. Метод градиента.

Допустим, что $\Phi(\mathbf{x})$ дважды непрерывно дифференцируема в области, содержащей изолированное решение \mathbf{x}^* , в окрестности которого поверхности уровня функции Φ имеют вид, изображенный на рис. 5.2. Обозначим через

$$\text{grad } \Phi(\mathbf{x}) = \begin{bmatrix} \frac{\partial \Phi}{\partial x_1} \\ \vdots \\ \frac{\partial \Phi}{\partial x_n} \end{bmatrix}$$

градиент функции $\Phi(\mathbf{x})$.

Задавшись начальным приближением $\mathbf{x}^{(0)}$, строим итерационный процесс по формулам

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \lambda_k \text{grad } \Phi(\mathbf{x}^{(k)}), k = 0, 1, \dots,$$

где λ_k – минимальный неотрицательный корень уравнения $\frac{d}{d\lambda} \Phi(\mathbf{x}^{(k)} - \lambda \text{grad } \Phi(\mathbf{x}^{(k)})) = 0$.

Последнее уравнение, вообще говоря, нужно решать численно. Поэтому укажем метод приближенного нахождения чисел λ_k :

$$\lambda_k = \frac{(\mathbf{F}(\mathbf{x}^{(k)}), \mathbf{W}(\mathbf{x}^{(k)}) \text{grad } \Phi(\mathbf{x}^{(k)}))}{(\mathbf{W}(\mathbf{x}^{(k)}) \text{grad } \Phi(\mathbf{x}^{(k)}), \mathbf{W}(\mathbf{x}^{(k)}) \text{grad } \Phi(\mathbf{x}^{(k)}))},$$

где $\mathbf{W}(\mathbf{x}^{(k)})$ – матрица Якоби $\mathbf{W}(\mathbf{x})$ для функции $\mathbf{F}(\mathbf{x})$ со значениями $\mathbf{x}^{(k)}$. Поскольку $\text{grad } \Phi(\mathbf{x}) = 2\mathbf{W}^T(\mathbf{x})\mathbf{F}(\mathbf{x})$, окончательно получим

$$\mu_k = 2\lambda_k = \frac{(\mathbf{F}^{(k)}, \mathbf{W}_k \mathbf{W}_k^T \mathbf{F}^{(k)})}{(\mathbf{W}_k \mathbf{W}_k^T \mathbf{F}^{(k)}, \mathbf{W}_k \mathbf{W}_k^T \mathbf{F}^{(k)})},$$

где для краткости положено $\mathbf{F}^{(k)} = \mathbf{F}(\mathbf{x}^{(k)})$, $\mathbf{W}_k = \mathbf{W}(\mathbf{x}^{(k)})$. Решение будем искать в виде

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mu_k \mathbf{W}_k^T \mathbf{F}^{(k)}, k = 0, 1, \dots$$

Итерационный процесс продолжается до тех пор, пока вектор невязки $\mathbf{F}(\mathbf{x}^{(k+1)})$ не станет близок к нулю.

Главное достоинство метода спуска решения нелинейных систем – глобальная сходимость. Процесс наискорейшего (градиентного) спуска приведет к какой-либо точке минимума функции из любой начальной точки. При определенных условиях найденная точка минимума будет искомым решением исходной нелинейной системы, однако существует вероятность попасть в точку относительного минимума. Кроме того следующий недостаток – медленная сходимость. Доказано, что сходимость данных методов линейная.

Пример 5.3. Методом скорейшего спуска приближенно вычислить корни системы

$$\begin{cases} x + x^2 - 2yz = 0,1, \\ y - y^2 + 3xz = -0,2, \\ z + z^2 + 2xy = 0,3, \end{cases}$$

расположенные в окрестности начала координат.

Решение. Имеем:

$$\mathbf{F} = \begin{bmatrix} x + x^2 - 2yz - 0,1 \\ y - y^2 + 3xz + 0,2 \\ z + z^2 + 2xy - 0,3 \end{bmatrix}; \mathbf{W} = \begin{bmatrix} 1+2x & -2z & -2y \\ 3z & 1-2y & 3x \\ 2y & 2x & 1+2z \end{bmatrix}.$$

Выберем начальное приближение:

$$\mathbf{x}^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \mathbf{F}^{(0)} = \begin{bmatrix} -0,1 \\ 0,2 \\ -0,3 \end{bmatrix}; \mathbf{W}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{E}.$$

По вышеприведенным формулам найдем первое приближение:

$$\mu_0 = \frac{(\mathbf{F}^{(0)}, \mathbf{F}^{(0)})}{(\mathbf{F}^{(0)}, \mathbf{F}^{(0)})} = 1; \mathbf{x}^{(1)} = \mathbf{x}^{(0)} - 1 \cdot \mathbf{E} \cdot \mathbf{F}^{(0)} = \begin{bmatrix} 0,1 \\ -0,2 \\ 0,3 \end{bmatrix}.$$

Аналогичным образом находим следующее приближение:

$$\mathbf{F}^{(1)} = \begin{bmatrix} 0,13 \\ 0,05 \\ 0,05 \end{bmatrix}; \mathbf{W}_1 = \begin{bmatrix} 1,2 & -0,6 & 0,4 \\ 0,9 & 1,4 & 0,3 \\ -0,4 & 0,2 & 1,6 \end{bmatrix}; \mathbf{W}_1^T \mathbf{F}^{(1)} = \begin{bmatrix} 0,181 \\ 0,002 \\ 0,147 \end{bmatrix}; \mathbf{W}_1 \mathbf{W}_1^T \mathbf{F}^{(1)} = \begin{bmatrix} 0,2748 \\ 0,2098 \\ 0,1632 \end{bmatrix};$$

$$\mu_1 = \frac{0,054374}{0,146198} = 0,3719; \mathbf{x}^{(2)} = \begin{bmatrix} 0,1 \\ -0,2 \\ 0,3 \end{bmatrix} - 0,3719 \cdot \begin{bmatrix} 0,181 \\ 0,002 \\ 0,147 \end{bmatrix} = \begin{bmatrix} 0,0327 \\ -0,2007 \\ 0,2453 \end{bmatrix}.$$

$$\text{Ограничимся двумя итерациями и оценим невязку: } \mathbf{F}^{(2)} = \begin{bmatrix} 0,032 \\ -0,017 \\ -0,007 \end{bmatrix}.$$

ЛАБОРАТОРНАЯ РАБОТА № 5

Цель работы

Исследовать рассмотренные выше итерационные методы к решению систем нелинейных уравнений; выявить их преимущества и недостатки друг перед другом.

Задание к лабораторной работе

Найти с точностью $\varepsilon = 10^{-6}$ корень системы нелинейных уравнений

$$\begin{cases} f_1(x_1, x_2) = 0, \\ f_2(x_1, x_2) = 0, \end{cases}$$

используя итерационные методы

- 1) Метод простых итераций;
- 2) Метод Ньютона;
- 3) Модифицированный метод Ньютона;
- 4) Метод Брауна;
- 5) Метод наискорейшего спуска;

сравнить скорости сходимости данных методов; сделать выводы.

N	Система уравнений	N	Система уравнений
1	$\begin{cases} \sin(x_1 + x_2) - x_2 - 1.2 = 0 \\ 2x_1 + \cos x_2 - 2 = 0 \end{cases}$	2	$\begin{cases} \sin(x_1 + x_2) - x_1 + 0.1 = 0 \\ x_2 - \cos(3x_1) + 0.1 = 0 \end{cases}$
3	$\begin{cases} \cos(x_1 - 1) + x_2 - 0.5 = 0 \\ \sin x_1 + 2x_2 - 2 = 0 \end{cases}$	4	$\begin{cases} \cos(x_1 - 1) + x_2 - 1 = 0 \\ \sin x_2 + 2x_1 - 1.6 = 0 \end{cases}$
5	$\begin{cases} \sin(0.5x_1 + x_2) - 1.2x_1 - 1 = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$	6	$\begin{cases} \sin(x_1 + 1) - x_2 - 1.2 = 0 \\ 2x_1 + \cos(x_2) - 2 = 0 \end{cases}$
7	$\begin{cases} \operatorname{tg}(x_1 x_2 + 0.3) - x_1^2 = 0 \\ 0.9x_1^2 + 2x_2^2 - 1 = 0 \end{cases}$	8	$\begin{cases} \operatorname{tg}(x_1 \cdot x_2 + 0.1) - x_1^2 = 0 \\ x_1^2 + 2x_2^2 - 1 = 0 \end{cases}$
9	$\begin{cases} \sin(x_2 + 1) - x_1 - 1.2 = 0 \\ 2x_1^2 + x_2 - 2 = 0 \end{cases}$	10	$\begin{cases} \sin(x_1 + x_2) - 1.5x_1 = 0 \\ x_1^2 + x_2^2 = 1 \end{cases}$
11	$\begin{cases} \cos(x_2 - 1) + x_1 - 0.8 = 0 \\ x_2 - \cos x_1 - 2 = 0 \end{cases}$	12	$\begin{cases} \operatorname{tg}(x_1 \cdot x_2) = x_1^2 \\ 0.7x_1^2 + 2x_2^2 = 1 \end{cases}$
13	$\begin{cases} \sin(x_1 + x_2) - 1.6x_1 - 1 = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$	14	$\begin{cases} \cos(x_1 - 1) + x_2 - 0.8 = 0 \\ 2x_1 - \cos x_2 + 2 = 0 \end{cases}$
15	$\begin{cases} \operatorname{tg}(x_1 x_2 + 0.2) - x_1^2 = 0 \\ 0.6x_1^2 + 2x_2^2 - 1 = 0 \end{cases}$		

Требования к оформлению отчета по работе

Отчет по работе должен содержать:

- 1) цель работы;
- 2) постановку задачи;
- 3) листинги программ с необходимыми комментариями;

- 4) результаты работы программ (оформлены в виде таблиц, увязывающих входные данные – точность вычислений и результаты – количество итераций);
- 5) анализ полученных результатов (оценка и сравнение числа итераций n для рассматриваемых методов при заданной точности);
- 6) выводы по проделанной работе.

Контрольные вопросы

1. Какую систему уравнений называют нелинейной?
2. Какова основная идея метода простых итераций решения систем нелинейных уравнений?
3. По каким формулам строится итерационный процесс метода простых итераций?
4. Что называют сжимающим отображением?
5. Какую область пространства R_n называют выпуклой?
6. Сформулируйте первое и второе достаточные условия сходимости метода простых итераций.
7. Как привести систему нелинейных уравнений к виду, удобному для итераций?
8. В чем состоит отличие метода Зейделя решения системы нелинейных уравнений от метода простых итераций?
9. По каким формулам строится итерационный процесс метода Зейделя?
10. Какой принцип лежит в основе реализации метода Ньютона решения систем нелинейных уравнений?
11. По каким формулам строится итерационный процесс согласно методу Ньютона?
12. На что направлены модификации метода Ньютона решения систем нелинейных уравнений?
13. По каким формулам строится итерационный процесс согласно модифицированному методу Ньютона?
14. Какой принцип лежит в основе реализации метода Брауна решения систем нелинейных уравнений?
15. По каким формулам строится итерационный процесс согласно методу Брауна?
16. В чем состоит общий недостаток методов простых итераций, Зейделя, Ньютона и Брауна?
17. Какой принцип лежит в основе реализации метода наискорейшего спуска решения систем нелинейных уравнений?
18. Что характеризует градиент функции?
19. По каким формулам строится итерационный процесс согласно методу наискорейшего спуска?
20. В чем состоит главное достоинство метода наискорейшего спуска?
21. В чем состоит недостаток метода наискорейшего спуска?

Задачи