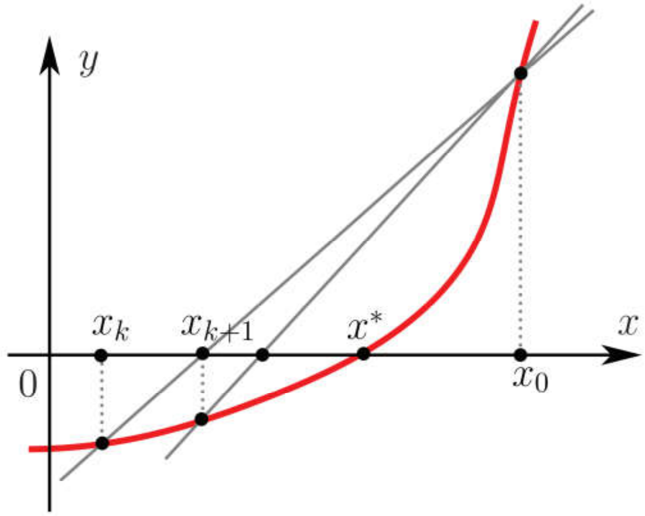

Рис. 4.5. Метод хорд

$\triangleright_5$ Обоснуйте это правило.

$\triangleright_6$ Определите порядок сходимости метода хорд.

Геометрический смысл метода хорд (4.11) изображен на рис. 4.5.

4.2. Решение систем нелинейных уравнений

4.2.1. Постановка задачи

Систему нелинейных уравнений

$$\begin{cases} f_1(x_1, \dots, x_n) = 0, \\ \dots\dots\dots\dots\dots\dots \\ f_n(x_1, \dots, x_n) = 0 \end{cases} \tag{4.12}$$

запишем в векторном виде

$$f(x) = 0, \tag{4.13}$$

где $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}^n$,

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}, \quad f_i : \mathbb{R}^n \to \mathbb{R}. \tag{4.14}$$

Как и ранее, точное решение уравнения (4.13) обозначим x^*. В дальнейшем будем предполагать, что f достаточно гладкая, т. е. существует необходимое число частных производных от f_i.

4.2.2. Метод Ньютона

В скалярном случае метод Ньютона имеет вид

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Для начала можно *формально* обобщить этот метод на случай системы (4.13):

$$x^{k+1} = x^k - f'(x^k)^{-1} f(x^k), \tag{4.15}$$

где $f' = \dfrac{\partial f}{\partial x}$ — матрица Якоби,

$$f' = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} \dfrac{\partial f_1}{\partial x_2} \cdots \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} \dfrac{\partial f_2}{\partial x_2} \cdots \dfrac{\partial f_2}{\partial x_n} \\ \dots\dots\dots\dots\dots \\ \dfrac{\partial f_n}{\partial x_1} \dfrac{\partial f_n}{\partial x_n} \cdots \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}. \tag{4.16}$$

Для обоснования формулы (4.15) рассмотрим $x \in \mathbb{R}^n$, приращение $\Delta x \in \mathbb{R}^n$ и воспользуемся формулой Тейлора для функции нескольких переменных:

$$f_i(x + \Delta x) = f_i(x) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x)\Delta x_j + O(\|\Delta x\|^2), \quad i = \overline{1, n}.$$

В векторной форме эти равенства можно записать как

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + O(\|\Delta x\|^2). \tag{4.17}$$

Пусть x^k — некоторое приближение к x^*. В идеале нам нужно найти такую поправку Δx^*, для которой $x^k + \Delta x^* = x^*$, или

$$f(x^k + \Delta x^*) = 0.$$

Отбрасывая остаточный член в формуле (4.17), имеем

$$0 = f(x^k + \Delta x^*) \approx f(x^k) + f'(x^k)\Delta x^* \Rightarrow \Delta x^* \approx -f'(x^k)^{-1} f(x^k) = \Delta x^k.$$

Вычисляя

$$x^* \approx x^{k+1} = x^k + \Delta x^k,$$

получим формулу (4.15).

Замечание 4.1. Важно понимать, что *формула* (4.15) ***в явном виде практически никогда не используется для расчетов***, так как для этого нужно иметь аналитический вид матрицы $f'(x)^{-1}$.

Вместо этого каждую итерацию метода Ньютона реализуют в два этапа:
1) сначала находят поправку Δx^k как решение СЛАУ

$$f'(x^k)\Delta x^k = -f(x^k); \tag{4.18а}$$

2) затем вычисляют

$$x^{k+1} = x^k + \Delta x^k. \tag{4.18б}$$

$\triangleright_1$ Оцените сложность реализации метода Ньютона по формулам (4.18).

При больших n метод Ньютона становится слишком трудоемким. Кроме этого, он не применим, если нет возможности вычислить f', что довольно часто случается на практике. Для таких случаев разработан ряд модификаций метода Ньютона, направленных на снижение вычислительных затрат.

4.2.3. Метод Ньютона с постоянным якобианом

Для того чтобы снизить затраты на решение СЛАУ (4.18а), рассматривают метод Ньютона с «замороженной» матрицей Якоби:

$$x^{k+1} = x^k - J_0^{-1} f(x_k), \tag{4.19}$$

где $J_0 = f'(x_0)$.

Для реализации такого метода достаточно один раз построить LU-разложение матрицы J_0, а затем с помощью него решать СЛАУ вида

$$J_0 \Delta x^k = -f(x_k)$$

для нахождения поправок за $O(n^2)$ операций. Данная модификация на порядок снижает вычислительные затраты, но при этом, конечно, снижается скорость сходимости метода.

4.2.4. Обобщения метода секущих

Метод секущих (4.10) можно рассматривать с двух сторон.
1) С одной стороны, он может быть получен путем модификации метода Ньютона с помощью замены

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}},$$

т. е. метод секущих основан на приближенном вычислении f'.
2) С другой стороны, данный метод определяет уравнение секущей $y(x)$, которая пересекает график f в точках x_k и x_{k-1} (именно так мы этот метод строили в предыдущем пункте). С этой точки зрения мы строим приближение для f.

В одномерном случае два описанных подхода дают один и тот же метод, однако в многомерном случае мы получим два разных семейства методов.

Первый способ: приближенное вычисление якобиана. Рассмотрим набор приращений $\{h_1, \ldots, h_n\}$, $h_i \in \mathbb{R}$, и определим приближения для частных производных f по формуле

$$\frac{\partial f_i}{\partial x_j}(x) \approx \frac{f_i(x_1, \ldots, x_j + h_j, \ldots x_n) - f_i(x_1, \ldots, x_n)}{h_j} = \psi_{ij}(x), \quad (4.20)$$

откуда имеем $f'(x) \approx \Psi(x)$. Соответствующий метод имеет вид

$$x^{k+1} = x^k - \Psi(x^k)^{-1} f(x_k). \quad (4.21)$$

$\triangleright_2$ Укажите достоинства и недостатки этого метода.

Кроме (4.20) можно рассматривать и другие варианты приближенного вычисления $f'(x)$.

Второй способ приводит к методу Бройдена.

4.2.5. Метод Бройдена

Обобщением понятия прямой, задаваемой уравнением $y = ax + b$, для пространства $\mathbb{R}^n$ является множество пар $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, таких, что

$$y = Ax + b, \quad (4.22)$$

где x, y, b — векторы из $\mathbb{R}^n$; A — квадратная матрица.

Пусть нам известны два приближения к x^*: x^k и x^{k-1}. Тогда секущая вида (4.22) определяется соотношениями

$$\begin{cases} Ax^{k-1} + b = f(x^{k-1}), \\ Ax^k + b = f(x^k). \end{cases} \tag{4.23}$$

Очевидно, что эти условия не определяют A и b однозначно, но пока предположим, что мы каким-то образом нашли A. Тогда $f(x) \approx Ax + b$ и следующее приближение x^{k+1} строится из соотношения

$$Ax^{k+1} + b = 0 \quad \Rightarrow \quad x^{k+1} = -A^{-1}b,$$

откуда, выражая из (4.23) $b = f(x^k) - Ax^k$, получим знакомую формулу

$$x^{k+1} = x^k - A^{-1}f(x^k).$$

Общая схема итерации многомерного метода секущих

1) По двум предыдущим приближениям x^k и x^{k-1} вычисляется матрица $A = A_k$, определяющая уравнение секущей.

2) Находится новое приближение по формуле

$$x^{k+1} = x^k - A_k^{-1}f(x^k). \tag{4.24}$$

Рассмотрим теперь главный вопрос: как определять матрицу A из условий (4.23)? Это можно делать по-разному. Один из наиболее удачных способов был предложен Ч. Бройденом в 1965 г.

Обозначим $\Delta x = x^k - x^{k-1}$, $\Delta f = f(x^k) - f(x^{k-1})$. Из (4.23) имеем

$$A_k\Delta x = \Delta f. \tag{4.25}$$

Основная идея заключается в том, чтобы представить A_k в виде

$$A_k = A_{k-1} + u\Delta x^T, \tag{4.26}$$

где $u \in \mathbb{R}^n$ — вектор неизвестных параметров.

Подставляя (4.26) в (4.25), получим

$$(A_{k-1} + u\Delta x^T)\Delta x = \Delta f,$$

откуда

$$u = \frac{1}{\|\Delta x\|_2^2}\Big(\Delta f - A_{k-1}\Delta x\Big). \tag{4.27}$$

Здесь уже можно остановиться: по u находим A_k, после чего найдем x^{k+1} по формуле (4.24). При этом необходимо решить СЛАУ вида

$$A_k(x^{k+1} - x^k) = -f(x^k). \tag{4.28}$$

При определенной схеме организации вычислений решение таких систем можно получить за $O(n^2)$ операций, но описание такого алгоритма не входит в рамки нашего курса [18].

В самом начале процесса для получения приближения x^1 можно использовать метод Ньютона или метод (4.21). Тогда матрицей A_1 будет матрица $f'(x^0)$ или $\Psi(x^0)$ соответственно.

4.2.6. Другие методы

В заключение рассмотрим *нелинейный метод Гаусса – Зейделя*. По аналогии со случаем СЛАУ здесь в цикле по i от 1 до n по очереди вычисляются компоненты x_i^{k+1} путем решения скалярного уравнения вида

$$f_i(x_1^{k+1}, \ldots, x_{i-1}^{k+1}, x_i, x_{i+1}^k, \ldots, x_n^k) = 0, \tag{4.29}$$

относительно x_i. После этого полагается $x_i^{k+1} = x_i$. Для решения скалярных уравнений (4.29) могут быть использованы соответствующие методы, рассмотренные ранее.

$\triangleright_3$ Постройте нелинейные аналоги методов Якоби и релаксации.

4.3. Анализ сходимости итерационных процессов

Как уже не раз упоминалось, многие из рассмотренных нами методов решения нелинейных уравнений и систем представимы в виде $x^{k+1} = \varphi(x^k)$. Рассмотрим классический способ доказательства сходимости таких итерационных процессов.

4.3.1. Принцип сжимающих отображений

Расстояние между двумя точками $x, y \in \mathbb{R}^n$ условимся обозначать

$$\rho(x, y) = \|x - y\|. \tag{4.30}$$

Определение 4.4. Если для функции $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ существует константа $L < \infty$, такая, что

$$\rho(\varphi(x), \varphi(y)) \leqslant L\rho(x, y), \quad \forall x, y \in D \subset \mathbb{R}^n,$$

то говорят, что φ удовлетворяет *условию Липшица* (или просто *липшицева*) на множестве D. Число L называют *константой Липшица*.

Определение 4.5. *Сжимающим отображением* называют функцию φ, удовлетворяющую условию Липшица на $\mathbb{R}^n$ с константой $L < 1$. Если же это условие выполнено лишь на некоторой области $D \subset \mathbb{R}^n$, то такое отображение будем называть *локально-сжимающим*.

Схема действия сжимающего отображения $\mathbb{R}^2 \to \mathbb{R}^2$ изображена на рис. 4.6.

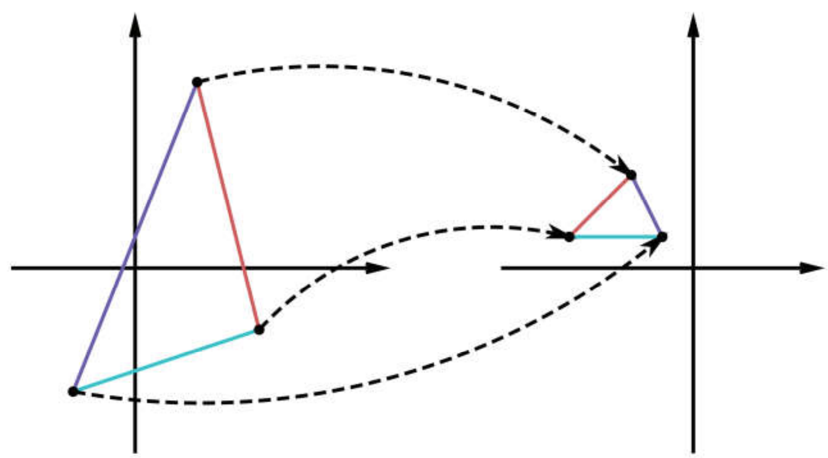

Рис. 4.6

Определение 4.6. Последовательность точек (x^k), $x^k \in \mathbb{R}^n$, называется *фундаментальной* или *последовательностью Коши*, если

$$\rho(x^k, x^m) \to 0 \quad \text{при} \quad k, m \to \infty.$$

Согласно критерию Коши, последовательность (x^k) сходится тогда и только тогда, когда она фундаментальна.

Теорема 4.1 (принцип сжимающих отображений). *Пусть* $\varphi : \mathbb{R}^n \to \to \mathbb{R}^n$ *— сжимающее отображение. Тогда уравнение*

$$\varphi(x) = x \tag{4.31}$$

имеет единственное решение $x = x^*$ *и для любого* $x_0 \in \mathbb{R}^n$ *последовательность*

$$x^{k+1} = \varphi(x^k), \quad k = 0, 1, 2, \ldots \tag{4.32}$$

сходится к x^*:

$$\rho(x^k, x^*) \xrightarrow[k\to\infty]{} 0.$$

Доказательство. Рассмотрим последовательность

$$x^1 = \varphi(x^0),\ x^2 = \varphi(x^1),\ \ldots,\ x^{k+1} = \varphi(x^k),\ \ldots$$

и докажем, что она является фундаментальной:

$$\rho(x^k, x^{k+1}) = \rho\left(\varphi(x^{k-1}), \varphi(x^k)\right) \leqslant L\rho(x^{k-1}, x^k) \leqslant \ldots \leqslant L^k\rho(x^0, x^1).$$

Пусть $m > k$. Тогда по неравенству треугольника

$$\rho(x^k, x^m) \leqslant \rho(x^k, x^{k+1}) + \rho(x^{k+1}, x^{k+2}) + \ldots + \rho(x^{m-1}, x^m) \leqslant$$
$$\leqslant (L^k + L^{k+1} + \ldots + L^{m-1})\rho(x^0, x^1),$$

откуда получим

$$\rho(x^k, x^m) \leqslant \frac{L^k}{1-L}\rho(x^0, x^1), \tag{4.33}$$

т. е. (x^k) фундаментальна ($L < 1$ по условию), и, следовательно, сходится к какому-то вектору $x^* \in \mathbb{R}^n$. Переходя к пределу в (4.32), получим, что x^* удовлетворяет уравнению (4.31).

Докажем единственность. Пусть $\exists\, x^{**} \neq x^*$ такой, что $\varphi(x^{**}) = x^{**}$. Это приводит к противоречию:

$$\rho(x^*, x^{**}) = \rho\left(\varphi(x^*), \varphi(x^{**})\right) \leqslant L\rho(x^*, x^{**}) < \rho(x^*, x^{**}). \qquad \square$$

Следствие 4.1. *Устремляя $m \to \infty$ в формуле* (4.33)*, получим априорную оценку погрешности:*

$$\rho(x^k, x^*) \leqslant \frac{L^k}{1-L}\rho(x^0, x^1). \tag{4.34}$$

$\triangleright_1$ Получите из формулы (4.34) оценку для количества итераций N, достаточного для получения решения с погрешностью, не превышающей ϵ.

Кроме оценки (4.34) можно получить и *апостериорную* оценку погрешности, которая, как правило, более точна, но для получения которой необходимо проделать k итераций. Положив в (4.34) $x^0 = x^{k-1}$, имеем

$$\rho(x^k, x^*) \leqslant \frac{L}{1-L}\rho(x^{k-1}, x^k). \tag{4.35}$$

Замечание 4.2. Функция ρ (метрика), которая измеряет расстояние между двумя точками, не обязательно должна иметь вид (4.30). Достаточно лишь, чтобы она удовлетворяла аксиомам метрики:
1) $\rho(x, y) = 0 \Leftrightarrow x = y$;
2) $\rho(x, y) = \rho(y, x)$;
3) $\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$.

Замечание 4.3. Принцип сжимающих отображений играет важную роль в функциональном анализе и других разделах математики. Он может быть сформулирован не только для $\mathbb{R}^n$, но и для произвольных метрических пространств, обладающих свойством *полноты*. В частности, с его помощью легко доказывается существование решений для обыкновенных дифференциальных уравнений, интегральных уравнений и т. д.

4.3.2. Локальный принцип сжимающих отображений

Рассмотрим теперь случай, когда отображение φ является лишь локально-сжимающим.

Определение 4.7. Гипершар радиусом r с центром в точке $x \in \mathbb{R}^n$ будем обозначать

$$B(x, r) = \{y \in \mathbb{R}^n \mid \rho(x, y) \leqslant r\}.$$

Лемма 4.2. *Пусть отображение φ имеет неподвижную точку $x^* = \varphi(x^*)$ и является локально-сжимающим в некоторой ее окрестности $B^* = B(x^*, r)$. Тогда для всех $x_0 \in B^*$ итерационный процесс* (4.32) *сходится к x^*.*

$\triangleright_2$ Докажите лемму.

Для применения леммы 4.2 необходимо обладать достаточно точной информацией о местоположении корня x^*. Следующий признак сходимости опирается лишь на информацию о φ в окрестности начального приближения x^0.

Теорема 4.2. *Пусть отображение φ является локально-сжимающим на множестве $B = B(x, r)$ с константой $L < 1$. Если имеет место неравенство*

$$\rho(x, \varphi(x)) \leqslant (1 - L)r, \tag{4.36}$$

то x^* — искомый корень уравнения $\varphi(x) = x$ — лежит в B и итерационный процесс (4.32) *сходится к x^* для всех $x_0 \in B$.*

Доказательство. Если мы докажем, что $\varphi(y) \in B$ при любых $y \in B$, то по общему принципу сжимающих отображений из этого автоматически будет следовать утверждение теоремы. Итак, пусть $y \in B$, тогда

$$\begin{aligned}\rho(\varphi(y), x) &\leqslant \rho(\varphi(y), \varphi(x)) + \rho(x, \varphi(x)) \leqslant \\ &\leqslant L\rho(y, x) + \rho(x, \varphi(x)) \leqslant Lr + (1 - L)r = r. \end{aligned} \qquad \square$$

4.3.3. Применение принципа сжимающих отображений

Рассмотрим случай функции одной переменной (одного нелинейного уравнения).

Теорема 4.3 (Лагранжа о среднем значении). *Если функция $\varphi : [a, b] \to \mathbb{R}$ непрерывна на $[a, b]$ и дифференцируема на (a, b), то $\exists\, \xi \in (a, b)$ такое, что*

$$\varphi(b) - \varphi(a) = \varphi'(\xi)(b - a).$$

В случае $\varphi : \mathbb{R} \to \mathbb{R}$ существует простое достаточное условие липшицевости: если φ достаточно гладкая, то согласно теореме Лагранжа имеем

$$|\varphi(x) - \varphi(y)| = |\varphi'(\xi)| \cdot |x - y|, \quad \xi \in (x, y),$$

поэтому если

$$|\varphi'(x)| \leqslant M \quad \forall\, x \in [a, b],$$

то φ — липшицева на $[a, b]$ с константой $L = M$.

Таким образом, получаем **признак сходимости 0**: если

$$|\varphi'(x)| \leqslant M < 1 \quad \forall\, x \in \mathbb{R}, \tag{4.37}$$

то итерационный процесс (4.32) сходится.

Очевидно, что ограничение вида (4.37) является слишком сильным. Как правило, мы можем рассчитывать на выполнение условий такого рода лишь на некотором подмножестве $\mathbb{R}$. Лемма 4.2 дает **признак сходимости 1**: если

$$|\varphi'(x)| \leqslant M < 1 \quad \forall\, x \in B^* = [x^* - r, x^* + r], \tag{4.38}$$

то итерационный процесс (4.32) сходится к x^* при всех $x_0 \in B^*$.

Из теоремы 4.2 получаем также **признак сходимости 2**: если

$$|\varphi'(x)| \leqslant M < 1 \quad \forall\, x \in B = [\xi - r, \xi + r]$$

и при этом

$$|\xi - \varphi(\xi)| \leqslant (1 - M)r,$$

то итерационный процесс (4.32) сходится к $x^* = \varphi(x^*)$ при всех $x_0 \in B$.

Важно понимать, что все рассмотренные выше признаки являются лишь *достаточными* условиями сходимости.

Применение к методу Ньютона

Напомним, что метод Ньютона для решения уравнения $f(x) = 0$ представляет собой метод типа (4.32), где

$$\varphi(x) = x - \frac{f(x)}{f'(x)}$$

и

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

Пусть $x^* \in [a, b]$, производные $f'(x)$ и $f''(x)$ непрерывны и не обращаются в нуль при всех $x \in [a, b]$. Тогда $\varphi'(x)$ непрерывна и из тождества $\varphi'(x^*) = 0$ следует, что существует целая окрестность $B^* = B(x^*, r)$ такая, что

$$|\varphi'(x)| \leqslant M < 1 \quad \forall\, x \in B^*.$$

Следовательно, по признаку сходимости 1 метод Ньютона будет сходиться при всех x_0 из B^*.