Глава 4
Приближенное решение нелинейных уравнений м систем

§1. Метод итерации

Пусть дано уравнение

$$t = \varphi(t). \qquad (1)$$

Метод итерации для решения этого уравнения состоит в том, что, начиная с некоторого начального приближения t_0, строится последовательность

$$t_{s+1} = \varphi(t_s).$$

Очевидно, что если функция $\varphi(t)$ непрерывна и $t_s \to t^*$, то t^* — корень уравнения (1). Графически t^* это абсцисса точки пересечения графика функции $\varphi(t)$ с биссектрисой первого координатного угла. Из рассмотрения графика можно сделать следующие выводы.

1) Если $|\varphi'(t)| < 1$, то следует ожидать сходимость метода.

2) Если $|\varphi'(t)| > 1$, то следует ожидать расходимость метода.

3) Если $0 < \varphi'(t) < 1$, то приближения t_s лежат по одну сторону от t^*.

4) Если $-1 < \varphi'(t) < 0$, то имеет место альтернирующая сходимость, т.е. t^* лежит между последовательными приближениями t_s и t_{s+1}.

Однако делать строгие выводы мы будем сразу для системы нелинейных уравнений.

Пусть дана система нелинейных уравнений

$$\xi_k = \varphi_k(\xi_1, \ldots, \xi_n), \qquad k = 1, \ldots, n,$$

которую будем записывать в векторной форме

$$x = \Phi(x), \qquad (2)$$

где $x = (\xi, \ldots, \xi_n) \in \mathbb{R}^n$, $\Phi(x) = (\varphi_1(x), \ldots, \varphi_n(x))$. Метод итерации состоит в том, что, выбрав начальное приближение x_0, мы строим последовательные приближения по формуле $x_{s+1} = \Phi(x_s)$. Дальше мы считаем, что в $\mathbb{R}^n$ введена некоторая норма $\|.\|$, отображение Φ задано в некоторой области $\Omega \subset \mathbb{R}^n$. Кроме того будем пользоваться обозначением $S_r(y) = \{\, x \in \mathbb{R}^n \;\big|\; \|x-y\| \le r \,\}$ – шар (соответствующий введенной норме!) радиуса r с центром в точке y.

Теорема 1. Пусть $x_0 \in \Omega$ и выполнены следующие условия:

1^0. $\|\Phi(x_0) - x_0\| \le m$,

2^0. существует $q < 1$, такое что для любых точек $x', x'' \in \Omega$ выполняется неравенство $\|\Phi(x') - \Phi(x'')\| \le q\|x' - x''\|$,

3^0. $S_r(x_0) \subseteq \Omega$, где $r = m/(1-q)$.

Тогда

а) в области Ω существует и притом единственное решение x^* уравнения (1),

б) $x^* \in S_r(x_0)$,

в) $x_s \to x^*$,

г) выполняются оценки погрешности:

$$\|x_s - x^*\| \le \frac{mq^s}{1-q}, \qquad \|x_s - x^*\| \le \frac{q}{1-q}\|x_s - x_{s-1}\|.$$

Доказательство. Докажем сначала методом индукции, что при всех $s \ge 0$ α) $x_s \in S_r(x_0)$ и β) $\|x_{s+1} - x_s\| \le mq^s$. Действительно, при $s = 0$ α) очевидно и по условию 1^0 $\|x_1 - x_0\| \le m$, так что β) также выполнено. Докажем возможность индуктивного перехода от s к $s+1$:

α) $\|x_{s+1} - x_0\| \le \|x_{s+1} - x_s\| + \cdots + \|x_1 - x_0\| \le mq^s + mq^{s-1} + \cdots + m \le m/(1-q) = r$;

β) $\|x_{s+2} - x_{s+1}\| = \|\Phi(x_{s+1}) - \Phi(x_s)\| \le q\|x_{s+1} - x_s\| \le mq^{s+1}$.

Из β) следует, что при любом натуральном p

$$\|x_{s+p} - x_s\| \le \|x_{s+p} - x_{s+p-1}\| + \cdots + \|x_{s+1} - x_s\| \le \\ \le mq^{s+p-1} + \cdots + mq^s \le \frac{mq^s}{1-q} \to 0, \qquad (3)$$

так что последовательность $\{x_s\}$ сходится в себе, и потому сходится: $x_s \to x^*$. Из $x_s \in S_r(x_0)$ вытекает $x^* \in S_r(x_0)$, а из непрерывности

Φ — $x^* = \Phi(x^*)$. Этим существование решения и утверждения б) и в) теоремы доказаны. Первая из оценок г) получается предельным переходом при $p \to \infty$ из (3), а вторая также предельным переходом по p из следующего неравенства:

$$\|x_{s+p} - x_s\| \le \|x_{s+p} - x_{s+p-1}\| + \dots + \|x_{s+1} - x_s\| \le$$
$$\le (q^p + q^{p-1} + \dots + q)\|x_s - x_{s-1}\| \le \frac{q}{1-q}\|x_s - x_{s-1}\|,$$

которое следует из того, что $\|x_{j+1} - x_j\| = \|\Phi(x_j) - \Phi(x_{j-1})\| \le q\|x_j - x_{j-1}\|$. Единственность решения доказывается от противного. Пусть $x^*, x^{**} \in \Omega$ — два решения. Тогда

$$\|x^* - x^{**}\| = \|\Phi(x^*) - \Phi(x^{**})\| \le q\|x^* - x^{**}\|,$$

откуда следует, что $x^* = x^{**}$. ■

Замечание 1. Решение уравнения (2) обычно называют *неподвижной точкой* отображения Φ. Отображение Φ, удовлетворяющее условию 2^0 теоремы называют *сжатым* или *сжимающим*. В связи с этим теорему 1 (или некоторые ее модификации) обычно называют *принципом сжатых отображений*. Это один из *принципов неподвижной точки*.

Замечание 2. Первая из оценок г) является *априорной*, а вторая — *апостериорной*. Эти понятия уже были введены ранее в §5 главы 3.

Введем понятие интеграла от вектор-функции. Пусть вектор $x(t)$ непрерывно зависит от $t \in [a, b]$. Тогда

$$\int_a^b x(t)dt = y = (\eta_1, \dots, \eta_n), \text{ где } \eta_k = \int_a^b \xi_k(t)dt.$$

Лемма. Выполняется неравенство

$$\left\|\int_a^b x(t)dt\right\| \le \int_a^b \|x(t)\|dt.$$

Д о к а з а т е л ь с т в о. Подобное неравенство очевидно для римановых сумм, и остается совершить предельный переход. ■

Отображение Φ называется дифференцируемым в точке x_0, если в этой точке дифференцируемы все его составляющие φ_k. В этом случае $\Phi'(x_0)$ – матрица Якоби:

$$\Phi'(x_0) = \left\{ \frac{\partial \varphi_k}{\partial \xi_j}(x_0) \right\}.$$

Отображение Φ непрерывно дифференцируемо в Ω, если таковы все φ_k.

Мы по-прежнему считаем, что в $\mathbb{R}^n$ введена некоторая векторная норма, и норма матрицы — это всегда операторная норма, порожденная введенной векторной.

Теорема 2. Пусть область Ω выпукла, Φ непрерывно дифференцируемо в Ω и при всех $x \in \Omega \quad \|\Phi'(x)\| \le q$. Тогда для всех $x', x'' \in \Omega \quad \|\Phi(x') - \Phi(x'')\| \le q\|x' - x''\|$.

Д о к а з а т е л ь с т в о. При $t \in [0, 1]$ рассмотрим

$$y(t) = \big(\eta_1(t), \dots, \eta_n(t)\big) = \Phi\big((1-t)x' + tx''\big) = \Phi\big(x' + t(x'' - x')\big)$$

(ввиду выпуклости $\Omega \;\; (1-t)x' + tx'' \in \Omega$). Очевидно

$$y(0) = \Phi(x'), \quad y(1) = \Phi(x''), \quad \eta_k(t) = \varphi_k\big(x' + t(x'' - x')\big).$$

Теперь имеем:

$$\eta_k(1) - \eta_k(0) = \int_0^1 \eta_k'(t)dt,$$

$$\eta_k'(t) = \sum_{j=1}^n \frac{\partial \varphi_k}{\partial \xi_j}\big(x' + t(x'' - x')\big) \cdot (\xi_j'' - \xi_j'),$$

$$y(1) - y(0) = \int_0^1 \Phi'\big(x' + t(x'' - x')\big) \cdot (x'' - x')dt,$$

$$\|\Phi(x'') - \Phi(x')\| = \|y(1) - y(0)\| \le \int_0^1 \|\Phi'\big(x' + t(x'' - x')\big)\| \cdot \|x'' - x'\| dt \le q\|x'' - x'\|. \quad \blacksquare$$

Заметим, что при применении этой теоремы удобно использовать нормы $\|\cdot\|_\infty$ и $\|\cdot\|_1$, поскольку именно для этих векторных норм порожденные ими матричные вычисляются по простым формулам.

В принципе сжатых отображений (теорема 1) в случае выпуклой области Ω условие 2^0 можно заменить на: $\|\Phi'(x)\| \le q < 1$ $\forall x \in \Omega$.

Задача 1. Сформулировать принцип сжатых отображений в той форме, как это предлагается в последнем абзаце. Обратить внимание, что предположения j том, что область Ω выпукла, не требуется (за исключением утверждения о единственности) — достаточно воспользоваться выпуклостью $S_r(x_0)$.

Задача 2. Показать, что правая часть во второй из оценок г) в теореме 1 всегда не больше правой части в первой.

Задача 3. Показать, что теорема 2 (§5 главы 3) может быть получена как следствие теоремы 1 этого параграфа.

§2. Метод итерации (продолжение)

В этом параграфе сосредоточены результаты, имеющие локальный характер.

Напомним из курса анализа: функция n переменных $\varphi(x)$ ($x \in \mathbb{R}^n$) называется дифференцируемой в точке x_0, если по всякому $\varepsilon > 0$ найдется такое $\delta > 0$, что при $\|x - x_0\|_2 < \delta$

$$\left|\varphi(x) - \varphi(x_0) - \sum_{j=1}^{n} \frac{\partial \varphi}{\partial \xi_j}(\xi_j - \xi_j^0)\right| < \varepsilon\|x - x_0\|_2.$$

Ввиду эквивалентности всех заданных в $\mathbb{R}^n$ норм норму $\|\cdot\|_2$ можно заменить на любую другую – получится эквивалентное определение.

Отсюда легко следует, что данному в предыдущем параграфе определению дифференцируемости отображения $\Phi(x)$ в точке x_0 эквивалентно следующее: по всякому $\varepsilon > 0$ найдется такое $\delta > 0$, что при $\|x - x_0\| < \delta$

$$\|\Phi(x) - \Phi(x_0) - \Phi'(x_0)(x - x_0)\| < \varepsilon \|x - x_0\|.$$

Последнее определение инвариантно относительно выбранной в $\mathbb{R}^n$ нормы.

Ниже, как и в предыдущем параграфе, норма матрицы — это всегда операторная норма, порожденная введенной в $\mathbb{R}^n$ векторной.

Пусть $\Phi: \ \Omega \mapsto \mathbb{R}^n$ $(\Omega \subset \mathbb{R}^n)$ – отображение, имеющее неподвижную точку $x^* \in \Omega$: $x^* = \Phi(x^*)$. Пусть x^* принадлежит Ω вместе с некоторой окрестностью. Рассматривается итеративная последовательность: $x_{s+1} = \Phi(x_s)$ при некотором начальном приближении x_0.

Теорема 1. Если отображение Φ дифференцируемо в точке x^* и выполняется неравенство $\|\Phi'(x^*)\| < 1$, то найдется такое $\delta > 0$, что при любом начальном приближении, удовлетворяющем неравенству $\|x_0 - x^*\| < \delta$, последовательность $\{x_s\}$ сходится к x^*.

Доказательство. Выберем число q так, что $\|\Phi'(x^*)\| < q < 1$ и положим $\varepsilon = q - \|\Phi'(x^*)\|$. По этому ε найдется такое $\delta > 0$, что из неравенства $\|x - x^*\| < \delta$ следует

$$\|\Phi(x) - \Phi(x^*) - \Phi'(x^*)(x - x^*)\| < \varepsilon \|x - x^*\|.$$

Если $\|x_0 - x^*\| < \delta$, то

$$\begin{aligned} \|x_1 - x^*\| &= \|\Phi(x_0) - \Phi(x^*)\| \le \\ &\le \|\Phi'(x^*)\| \cdot \|x_0 - x^*\| + \varepsilon \|x_0 - x^*\| = q \|x_0 - x^*\|. \end{aligned}$$

Методом индукции отсюда легко можно получить, что при всех $s = 1, 2, \ldots$ $\|x_s - x^*\| < \delta$ и $\|x_s - x^*\| \le q^s \|x_0 - x^*\|$, откуда и следует сходимость. ■

При применении этой теоремы в распоряжении исследователя находится параметр q, лежащий в указанных пределах. Увеличение этого параметра ослабляет напрашивающееся утверждение о скорости сходимости, но увеличивает возможную область выбора начального приближения. Кроме того, формулировка теоремы *не* инвариантна относительно введенной в $\mathbb{R}^n$ нормы, так что эта норма

– еще один инструмент, находящийся в руках исследователя. В действительности верно несколько более сильное утверждение: условие $\|\Phi'(x^*)\| < 1$ может быть заменено на $\rho\big(\Phi'(x^*)\big) < 1$. Здесь $\rho(A)$ – спектральный радиус матрицы A.

Определение. Пусть $\alpha > 1$ и дана последовательность векторов $\{a_s\}$. Говорят, что эта последовательность сходится к вектору a^* *с порядком* α, если $a_s \to a^*$ и существует такая постоянная c, что при всех s $\quad \|a_{s+1} - a^*\| \le c\|a_s - a^*\|^\alpha$.

Пусть имеется некоторый метод, который исходя из произвольного вектора x_0 строит последовательность приближений x_s к вектору x^*. Говорят, что этот метод сходится *с порядком* α, если существует такое $\delta > 0$, что при выполнении условия $\|x_0 - x^*\| < \delta$ последовательность $\{x_s\}$ сходится к x^* с порядком α.

Замечание 1. Сходимость с порядком $\alpha = 2$ называется квадратичной. Существуют еще и такие термины. Последовательность $\{a_s\}$ сходится к a^* *линейно*, если существует такое $q < 1$, что $\|a_{s+1} - a^*\| \le q\|a_s - a^*\|$, и *сверхлинейно*, если $\|a_{s+1} - a^*\|/\|a_s - a^*\| \to 0$. Понятие линейной и сверхлинейной сходимости переносится и на методы построения последовательностей.

Замечание 2. Легко видеть, что понятие сходимости с порядком α и сверхлинейной сходимости инвариантно относительно выбранной в $\mathbb{R}^n$ нормы. К линейной сходимости это не относится.

Лемма. Пусть x^* – неподвижная точка отображения Φ: $x^* = \Phi(x^*)$, и пусть нашлись такие $r > 0$, $c > 0$ и $\alpha > 1$, что $S_r(x^*) \subset \Omega$ и для любого $x \in S_r(x^*)$ выполняется неравенство $\|\Phi(x) - \Phi(x^*)\| \le c\|x - x^*\|^\alpha$. Тогда метод итерации для нахождения x^* сходится с порядком α.

Доказательство. Возьмем произвольное $q < 1$ и найдем $\delta > 0$ из условий $\delta < r$ и $c\delta^{\alpha-1} \le q$. Тогда если $\|x_0 - x^*\| < \delta$, то

$$\|x_1 - x^*\| = \|\Phi(x_0) - \Phi(x^*)\| \le c\|x_0 - x^*\|^\alpha \le q\|x_0 - x^*\| \le \delta.$$

Отсюда методом индукции легко заключить, что при всех s

$$\|x_s - x^*\| \le \|x_0 - x^*\|q^s \to 0,$$

причем $\|x_{s+1} - x^*\| = \|\Phi(x_s) - \Phi(x^*)\| \le c\|x_s - x^*\|^\alpha$. ■

Теорема 2. Пусть отображение Φ дважды непрерывно дифференцируемо в некоторой окрестности неподвижной точки x^* и $\Phi'(x^*) = \mathbb{O}$. Тогда метод итерации для нахождения x^* сходится квадратически.

Доказательство. Используя инвариантность понятия квадратичной сходимости относительно выбранной в $\mathbb{R}^n$ нормы, проверим выполнение условий леммы при $\alpha = 2$ для нормы $\|\cdot\|_\infty$. Пусть отображение Φ дважды непрерывно дифференцируемо в шаре $S_r(x^*)$ ($\|x - x^*\| \le r$). Положим

$$M = \max\left\{ \left|\frac{\partial^2\varphi_k}{\partial\xi_j\partial\xi_l}\right| \;\middle|\; 1 \le k,j,l \le n,\ x \in S_r(x^*) \right\}.$$

Пусть $x \in S_r(x^*)$. Определим функции

$$u_k(t) = \varphi_k\big((1-t)x^* + tx\big) = \varphi_k\big(x^* + t(x - x^*)\big).$$

Тогда имеем

$$u_k'(t) = \sum_{j=1}^n \frac{\partial\varphi_k}{\partial\xi_j}\cdot(\xi_j - \xi_j^*),$$
$$u_k''(t) = \sum_{j,l=1}^n \frac{\partial^2\varphi_k}{\partial\xi_j\partial\xi_l}\cdot(\xi_j - \xi_j^*)(\xi_l - \xi_l^*),$$

где все производные вычисляются в точке $x^* + t(x - x^*)$. Отсюда

$$u_k'(0) = 0;$$
$$|u_k''(t)| \le Mn^2\|x - x^*\|_\infty^2 \quad \forall t \in [0,1];$$
$$\varphi_k(x) - \varphi_k(x^*) = u_k(1) - u_k(0) = u_k'(0) + \frac{1}{2}u_k''(\tau_k) = \frac{1}{2}u_k''(\tau_k);$$
$$|\varphi_k(x) - \varphi_k(x^*)| \le \frac{1}{2}Mn^2\|x - x^*\|_\infty^2;$$
$$\|\Phi(x) - \Phi(x^*)\|_\infty \le c\|x - x^*\|_\infty^2$$

при $c = \frac{1}{2}Mn^2$. ■

Задача. Показать, что в случае одного уравнения $t = \varphi(t)$ с трижды непрерывно дифференцируемой функцией φ условие $\varphi(t^*) = \varphi'(t^*) = \varphi''(t^*) = 0$ гарантирует сходимость третьего порядка метода итерации для корня t^*.

§3. Метод Ньютона

Метод Ньютона сначала рассмотрим для одного уравнения $f(t) = 0$. Пусть функция f дважды непрерывно дифференцируема в окрестности корня t^* этого уравнения и пусть нам известно достаточно близкое приближение t_0 к этому корню. Тогда

$$0 = f(t^*) = f(t_0) + (t^* - t_0)f'(t_0) + \frac{1}{2}(t^* - t_0)^2 f''(\tau).$$

Последнее слагаемое в правой части мало, и им можно пренебречь, так что t^* с хорошей точностью удовлетворяет уравнению $f(t_0) + (t - t_0)f'(t_0) = 0$. Решение этого уравнения $t_1 = t_0 - f(t_0)/f'(t_0)$ принимается за следующее приближение к решению. Итак, метод Ньютона состоит в следующем. Выбирается некоторое начальное приближение t_0 и строится последовательность

$$t_{s+1} = t_s - \frac{f(t_s)}{f'(t_s)}. \tag{1}$$

Метод Ньютона имеет простой геометрический смысл: t_{s+1} есть абсцисса точки пересечения касательной к графику функции f, построенной в точке $\big(t_s, f(t_s)\big)$, с осью абсцисс.

Иногда оказывается удобным не пересчитывать на каждом шаге производную и пользоваться упрощенной формулой

$$\tilde{t}_{s+1} = \tilde{t}_s - \frac{f(\tilde{t}_s)}{f'(t_0)}.$$

Этот метод называется *модифицированным методом Ньютона.* Очевидно, что если $\tilde{t}_0 = t_0$, то и $\tilde{t}_1 = t_1$. Модифицированный метод Ньютона также имеет очевидный геометрический смысл.

В случае трудностей с вычислением производной ее значения можно заменять, используя численное дифференцирование. Пусть взяты два близких начальных приближения t_0 и t_1. Тогда можно построить следующее приближение по формуле $t_2 = t_1 - f(t_1)/f(t_0, t_1)$, заменив в формуле (1) при $s = 1$ производную $f'(t_1)$ на разделенную разность $f(t_0, t_1)$. Это приводит к последовательности

$$t_{s+1} = t_s - \frac{f(t_s)}{f(t_{s-1}, t_s)} = t_s - \frac{t_s - t_{s-1}}{f(t_s) - f(t_{s-1})} f(t_s). \tag{2}$$

Этот метод называется *методом секущих.*

Иногда употребляют также *метод хорд.* Предполагается, что известны начальные приближения t_0 и t_1, такие что $f(t_0)$ и $f(t_1)$ имеют противоположные знаки. Тогда построенное по формуле (2) (при $s = 1$) приближение t_2 лежит между t_0 и t_1. Из промежутков $[t_0, t_2]$ и $[t_2, t_1]$ выбирается тот, на концах которого функция f принимает значения разных знаков, и делается новый шаг, аналогичный предыдущему.

Вернемся к методу Ньютона. Формулу (1) можно рассматривать как формулу метода итераций для уравнения $t = \varphi(t)$, где $\varphi(t) = t - f(t)/f'(t)$. Легко проверить, что $\varphi'(t^*) = 0$. Поэтому, в связи с результатами предыдущего параграфа, следует ожидать квадратичную сходимость метода.

Аналогично модифицированный метод Ньютона можно рассматривать как метод итерации для уравнения $t = \tilde{\varphi}(t)$, где $\tilde{\varphi}(t) = t - f(t)/f(t_0)$ Здесь $\tilde{\varphi}'(t_0) = 0$, и при достаточно хорошем начальном приближении можно ожидать сходимость этого метода с быстротой геометрической прогрессии с малым знаменателем.

Исследование методов Ньютона и модифицированного будем проводить для случая системы уравнений

$$F(x) = 0 \tag{3}$$

($x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$, $F(x) = \big(f_1(x), \ldots, f_n(x)\big)$). Согласно формуле Тейлора

$$\mathbb{O} = F(x^*) \approx F(x_0) + F'(x_0)(x^* - x_0).$$

Так что исходя из начального приближения x_0 последующие по методу Ньютона строятся по формуле, вполне аналогичной (1):

$$x_{s+1} = x_s - [F'(x_s)]^{-1} F(x_s). \tag{4}$$

Формула модифицированного метода:

$$\tilde{x}_{s+1} = \tilde{x}_s - [F'(x_0)]^{-1} F(\tilde{x}_s). \tag{5}$$

Реально в методе Ньютона нахождение x_{s+1} требует не обращения матрицы (что нерационально), а решения системы уравнений:

$$x_{s+1} = x_s + \Delta x_s, \qquad F'(x_s)\Delta x_s = -F(x_s).$$

В модифицированном методе обращение матрицы может оказаться оправданным.

Теорема 1 (о методе Ньютона). Пусть x^* – решение системы уравнений (1). Пусть в некоторой окрестности точки x^* отображение $F(x)$ трижды непрерывно дифференцируемо и $\det F'(x^*) \neq 0$. Тогда метод Ньютона сходится для x^* квадратически.

Доказательство. $F'(x) \to F'(x^*)$ при $x \to x^*$. Но при условии $\|F'(x) - F'(x^*)\| \cdot \|[F'(x^*)]^{-1}\| < 1$ матрица Якоби $F'(x)$ также обратима (мы используем теорему 5 §2 главы 3). Поэтому для некоторого $r > 0$ при всех тех x, для которых $\|x - x^*\| \leq r$, существует матрица $\Gamma(x) = [F'(x)]^{-1} = \{\gamma_{kj}(x)\}$. Легко видеть, что функции $\gamma_{kj}(x)$ дважды дифференцируемы. Метод Ньютона теперь запишется так: $x_{s+1} = x_s - \Gamma(x_s)F(x_s) = \Phi(x_s)$. Отображение $\Phi(x)$ дважды непрерывно дифференцируемо в окрестности точки x^*, и чтобы воспользоваться теоремой 2 из предыдущего параграфа, достаточно убедиться, что $\Phi'(x^*) = \mathbb{O}$. Для компонент отображения Φ имеем

$$\varphi_k(x) = \xi_k - \sum_{j=1}^{n} \gamma_{kj}(x) \cdot f_j(x),$$

$$\frac{\partial \varphi_k}{\partial \xi_l} = \delta_{kl} - \sum_{j=1}^{n} \left[\frac{\partial \gamma_{kj}}{\partial \xi_l}(x) f_j(x) + \gamma_{kj}(x) \frac{\partial f_j}{\partial \xi_l}(x) \right].$$

При $x = x^*$ первое слагаемое в квадратных скобках обращается в ноль, а второе есть (k, l)-ый элемент матрицы $\Gamma(x^*)F'(x^*) = E$, т.е. δ_{kj}. Итак, все элементы матрицы $\Phi'(x^*)$ равны нулю, $\Phi'(x^*) = \mathbb{O}$, и для завершения доказательства остается воспользоваться теоремой 2 из §2. ■

Замечание 1. В условиях этой теоремы требование на гладкость отображения F завышено — в действительности достаточно двукратной дифференцируемости.

Замечание 2. Условие $\det F'(x^*) \neq 0$ теоремы существенно – без него квадратичной сходимости может и не быть. Приведем соответствующий пример в одномерном случае. Пусть $f(t) = t^2$. Корень $t^* = 0$ уравнения $f(t) = 0$ таков, что $f'(t^*) = 0$. Для метода Ньютона легко получаем $t_{s+1} = \frac{1}{2}t_s$, и при $t_0 \neq 0$ сходимость метода всего лишь линейная.

Теорема 2 (о модифицированном методе Ньютона). Пусть $F(x^*) = 0$ и в некоторой окрестности x^* отображение F дважды непрерывно дифференцируемо. Пусть $\det F'(x^*) \neq 0$. Тогда найдутся такие $\delta > 0$ и $c > 0$, что при $x_0 \in S_\delta(x^*)$ модифицированный метод Ньютона сходится с быстротой геометрической прогрессии: $\|\tilde{x}_s - x^*\| \leq q^s\|x_0 - x^*\|$, где $q \leq c\|x_0 - x^*\| < 1$.

Доказательство. Обозначим r радиус того шара $S_r(x^*)$, в котором отображение F дважды дифференцируемо. Найдется такая постоянная c_1, что для $x', x'' \in S_r(x^*)$ выполняется неравенство $\|F'(x') - F'(x'')\| \leq c_1\|x' - x''\|$. Выберем $r_0 < r$ так, что $\|[F'(x^*)]^{-1}\|c_1 r_0 < \frac{1}{2}$. По теореме 5 из §2 главы 4 для всех $x \in S_{r_0}(x^*)$ существует обратная матрица $\Gamma(x) = [F'(x)]^{-1}$ и $\|\Gamma(x)\| \leq 2\|\Gamma(x^*)\| = c_2$. Положим $c = 2c_1c_2$ и выберем δ так, что $\delta \leq r_0$ и $c\delta < 1$. Итак, c и δ выбраны. Покажем, что они требуемые. Пусть $x_0 \in S_\delta(x^*)$. В модифицированном методе Ньютона $\tilde{x}_0 = x_0$ и $\tilde{x}_{s+1} = \Phi(\tilde{x}_s)$, где $\Phi(x) = x - \Gamma_0 F(x)$, $\Gamma_0 = \Gamma(x_0)$. Так как $\Phi'(x) = E - \Gamma_0 F'(x) = \Gamma_0\big(F'(x_0) - F'(x)\big)$,то при $\|x - x^*\| \leq \|x_0 - x^*\|$ будет

$$\|\Phi'(x)\| \leq c_2c_1\|x_0 - x\| \leq c_1c_2(\|x_0 - x^*\| + \|x - x^*\|) \leq c\|x_0 - x^*\| = q < 1.$$

Поэтому по теореме 2 из §1 для любых точек x', x'', таких что $\|x' - x^*\|, \|x'' - x^*\| \le \|x_0 - x^*\|$ имеем

$$\|\Phi(x') - \Phi(x'')\| \le q\|x' - x''\|.$$

Отсюда

$$\|\tilde{x}_s - x^*\| = \|\Phi(\tilde{x}_{s-1}) - \Phi(x^*)\| \le q\|\tilde{x}_{s-1} - x^*\| \le \cdots \le q^s\|x_0 - x^*\|.$$

Этим теорема доказана. ■

Известны методы и более высокого, чем второй, порядка сходимости. Например, методами третьего порядка являются метод касательных гипербол и Чебышева, вычислительные формулы которых в случае одного вещественного уравнения $f(t) = 0$ выглядят соответственно так:

$$t_{s+1} = t_s - \frac{1}{1 - \frac{f''(t_s)f(t_s)}{2[f'(t_s)]^2}} \frac{f(t_s)}{f'(t_s)},$$

$$t_{s+1} = t_s - \left(1 + \frac{f(t_s)f''(t_s)}{2[f'(t_s)]^2}\right) \frac{f(t_s)}{f'(t_s)}.$$

Методы высших порядков обычно менее эффективны, чем метод Ньютона, особенно в случае систем уравнений. Например, при решении систем указанные методы третьего порядка на одном шаге требуют вычисления самих функций φ_k (их n), их первых производных (их n^2) и вторых (их n^3) и решения двух систем линейных уравнений порядка n. Так что один шаг такого метода более трудоемок, чем два шага метода Ньютона. В то же время, грубо говоря, один шаг такого метода возводит погрешность в третью степень, в то время как два шага метода Ньютона – в четвертую.

Но бывают случаи, когда начальное приближение чем-то особенно просто – в этой точке уже известны все производные, среди них много нулей и т.п. Если одного шага метода Ньютона здесь не хватает для достижения нужной точности, то применение методов высшего порядка оправдано. Возможно, методы высших порядков имеют некоторое преимущество при распараллеливании вычислительных процессов.

Задача. Показать, что в случае одного уравнения для четырежды непрерывно дифференцируемой функции f методы Чебышева и касательных гипербол имеют третий порядок сходимости.