2.2. **Newton's method for systems of equations.** We consider only the case of two equations, i.e., given two functions $f(x,y)$ and $g(x,y)$, we wish to find values $x = s$, $y = t$ that simultaneously satisfy the equations

$$f(s,t) = 0, \qquad g(s,t) = 0.$$

To obtain Newton's method for this system, we follow the procedure used for a single equation, i.e., we expand these functions in a Taylor series (where the Taylor series needed is the one in two variables). If we already have an approximation (x_n, y_n), then we use a Taylor series expansion about this point, i.e., for some numbers ξ_1, η_1, ξ_2, η_2, we have

$$\begin{aligned}
0 = f(s,t) &= f(x_n,y_n) + \frac{\partial f}{\partial x}(x_n,y_n)(s-x_n) + \frac{\partial f}{\partial y}(x_n,y_n)(t-y_n) \\
&\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\xi_1,\eta_1)(s-x_n)^2 + \frac{\partial^2 f}{\partial x \partial y}(\xi_1,\eta_1)(s-x_n)(t-y_n) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(\xi_1,\eta_1)(t-y_n)^2, \\
0 = g(s,t) &= g(x_n,y_n) + \frac{\partial g}{\partial x}(x_n,y_n)(s-x_n) + \frac{\partial g}{\partial y}(x_n,y_n)(t-y_n) \\
&\quad + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(\xi_2,\eta_2)(s-x_n)^2 + \frac{\partial^2 g}{\partial x \partial y}(\xi_2,\eta_2)(s-x_n)(t-y_n) + \frac{1}{2}\frac{\partial^2 g}{\partial y^2}(\xi_2,\eta_2)(t-y_n)^2.
\end{aligned}$$

As in one variable, if we assume that x_n is close to s and y_n is close to t, then we expect the quadratic terms in $s - x_n$ and $t - y_n$ to be small compared to the linear terms. Neglecting these terms, we define (x_{n+1}, y_{n+1}) as the solution of the linear system of equations

$$\begin{aligned}
\frac{\partial f}{\partial x}(x_n,y_n)(x_{n+1}-x_n) + \frac{\partial f}{\partial y}(x_n,y_n)(y_{n+1}-y_n) &= -f(x_n,y_n), \\
\frac{\partial g}{\partial x}(x_n,y_n)(x_{n+1}-x_n) + \frac{\partial g}{\partial y}(x_n,y_n)(y_{n+1}-y_n) &= -g(x_n,y_n).
\end{aligned}$$

Newton's method: Starting from an initial approximation (x_0, y_0), define a sequence of approximations (x_n, y_n) by the iteration scheme.

$$x_{n+1} = x_n + \delta_n, \qquad y_{n+1} = y_n + \epsilon_n,$$

where (δ_n, ϵ_n) is the solution of the linear system

$$\begin{pmatrix} f_x(x_n,y_n) & f_y(x_n,y_n) \\ g_x(x_n,y_n) & g_y(x_n,y_n) \end{pmatrix} \begin{pmatrix} \delta_n \\ \epsilon_n \end{pmatrix} = - \begin{pmatrix} f(x_n,y_n) \\ g(x_n,y_n) \end{pmatrix}.$$

3. Fixed Point Iteration:

To study the convergence of some of these methods, we next consider a scheme called fixed point iteration. In this method, instead of seeking a root of $f(x) = 0$, we look for a fixed point of a function $g(x)$, i.e., a value of x satisfying $x = g(x)$. We choose $g(x)$ so that fixed points of g are roots of f. One problem is that there are infinite number of ways we can do this. As we shall see, it will be important to choose g so that it has certain desirable properties.

Example: $f(x) = x^2 - x - 2$. Then some choices of $g(x)$ are: a) $f(x) = x^2 - 2$, (b) $g(x) = \sqrt{2+x}$, (c) $g(x) = 1 + 2/x$, (d) $g(x) = x - (x^2 - x - 2)/m$, where $m > 0$ is a constant.

Note that if we define $g(x) = x - f(x)/f'(x)$, then a simple root x^* of f (i.e., $f'(x^*) \neq 0$) will be a fixed point of $g(x)$.

Fixed point iteration algorithm: Given a starting guess x_0, we define the iteration $x_{n+1} = g(x_n)$. We then have the following convergence result for this iteration scheme.

Theorem 1. *Let $I = [a, b]$, where a and b are finite and assume that g satisfies the following conditions: (i) g is continuous on I and differentiable on (a, b), (ii) $g(x) \in I$ for all $x \in I$, and (iii) There is a constant L, with $0 < L < 1$ such that $|g'(x)| \leq L$ for all $x \in (a, b)$. Then there is a unique fixed point s of g (i.e., $s = g(s)$) in the interval I and for any choice of $x_0 \in I$, the sequence $\{x_n\}$ defined by the iteration $x_{n+1} = g(x_n)$ converges to s.*

Proof. To prove existence of a fixed point, we set $f(x) = x - g(x)$. Since by (ii), $a \leq g(a) \leq b$ and $a \leq g(b) \leq b$, $f(a) = a - g(a) \leq 0$ and $f(b) = b - g(b) \geq 0$. Since g is continuous on I, so is f. Hence, by the Intermediate Value Theorem, there exists at least one point s in $[a, b]$ such that $f(s) = 0$, i.e., $s = g(s)$. To see there can be only one such point, we suppose there are two fixed points s_1 and s_2. Then using (iii), and the Mean Value Theorem, there exists a point $c \in (a, b)$ such that

$$|s_2 - s_1| = |g(s_2) - g(s_1)| = |g'(c)(s_2 - s_1)| \leq |g'(c)||s_2 - s_1| \leq L|s_2 - s_1|.$$

Since $L < 1$, we must have $s_2 = s_1$. To establish convergence, we again use the Mean Value Theorem to write

$$|s - x_n| = |g(s) - g(x_{n-1})| = |g'(c_n)(s - x_{n-1})| \leq L|s - x_{n-1}| \leq L^2|s - x_{n-2}| \leq \ldots \leq L^n|s - x_0|.$$

Since $L < 1$, $\lim_{n\to\infty} L^n = 0$ and so $\lim_{n\to\infty} |s - x_n| = 0$, i.e., $\lim_{n\to\infty} x_n = s$. □

We can also derive error bounds on the approximation that do not depend on the unknown solution.

Corollary 1.

$$|s - x_n| \leq L^n \max\{b - x_0, x_0 - a\}.$$

From the proof of the theorem, we know that $|s - x_n| \leq L^n|s - x_0|$. Since both x_0 and s belong to I, either $s \in [a, x_0]$ or $s \in [x_0, b]$. Hence, $|s - x_0| \leq L^n \max\{b - x_0, x_0 - a\}$.

It is also possible to establish the following result.

Corollary 2.

$$|s - x_n| \leq \frac{L^n}{1 - L}|x_1 - x_0|.$$

Note that the rate of convergence of the method depends on the constant L. The smaller the value of L, the faster the convergence.

Example: $g(x) = (x^2 - 1)/3$, $I = [-1, 1]$. (i) Since g is a polynomial, it is continuous and differentiable everywhere. (ii) We find the maximum and minimum of $g(x)$ on $I = [-1, 1]$. Now $g'(x) = 2x/3 = 0$ only for $x = 0$. Hence the max and min can occur only at $x = -1, 0, 1$. Since $g(-1) = 0$, $g(1) = 0$ and $g(0) = -1/3$, we get $-1/3 \leq g(x) \leq 0$. Hence $g(x) \in [-1, 1]$ for all $x \in [-1, 1]$and (ii) is satisfied. (iii) is also satisfied since $|g'(x)| = |2x/3| \leq 2/3 = L < 1$ for $x \in [-1, 1]$. Hence, the iteration $x_{n+1} = (x_n^2 - 1)/3$ converges to the unique fixed point of g in $[-1, 1]$.

Example: Suppose we wish to calculate the root $s = 2$ of the function $f(x) = x^2 - x - 2$ by fixed point iteration. If we define $g(x) = x^2 - 2$, then $g'(x) = 2x$ and so $|g'(x)| > 1$ for $x > 1/2$. Hence assumption (iii) is not satisfied for any interval (a, b) containing the root $s = 2$ and the convergence theorem does not apply.

If we try $g(x) = \sqrt{2 + x}$, then $g'(x) = 1/(2\sqrt{2 + x})$. Now for $x \geq 0$, $g(x) \geq 0$ and

$$0 \leq g'(x) \leq 1/(2\sqrt{2}) < 1.$$

Also for $0 \leq x \leq 7$, $g(x) = \sqrt{2 + x} \leq \sqrt{2 + 7} = 3$. Hence, with $I = [0, 7]$, all the assumptions of the theorem are satisfied and so the iteration $x_{n+1} = \sqrt{2 + x_n}$ converges to 2 for any $x_0 \in [0, 7]$.