CHAPTER 1

Solving Non-Linear Equations $f(x) = 0$

1.1 Introduction to Calculation Errors and Approximations

The word "error" does not mean fault (i.e. the existence of a flaw in reasoning or the algorithm) but rather errors that arise during calculations, for example, when using approximated values or rounding (there is a difference between a pocket calculator and a computer). Even very small initially, these errors can accumulate when a large number of operations are performed and can compromise the precision of the results. Our goal is to minimize the error.

1.1.1 True and Approximated Values

Before quantifying errors, we must distinguish between the two types of values we encounter in scientific computing:

Definition 1.1

The true value, denoted by x, represents the actual magnitude of a quantity. It is the result of a perfect mathematical calculation or an ideal measurement (e.g., $\sqrt{2}$, e, or π).

Definition 1.2

The approximated value denoted by x^*, is the numerical result obtained through an algorithm, a measuring instrument, or a computer. Due to finite precision, x^* is usually a close approximation of x.

Example 1.1

Consider the fraction $\frac{1}{3}$.

- The true value is $x = 0.333333...$ (infinite decimals).
- An approximated value might be $x^* = 0.3333$ (if we use 4 decimal places).

1.1.2 Absolute and Relative Errors

To measure the proximity of an approximated value x^* to the exact value x, we define the following:

Definition 1.3

The absolute error of x^* relative to x, denoted by E_a, is given by:

$$E_a = |x - x^*|$$

Example 1.2

Suppose the exact value is $x = 2026.001$ and the measured approximation is $x^* = 2026$. The absolute error is:

$$E_a = |2026.001 - 2026| = 0.001$$

Definition 1.4

The relative error, denoted by E_r, provides context to the error by comparing it to the exact value:

$$E_r = \frac{|x - x^*|}{|x|}, \quad x \neq 0$$

Example 1.3

Using the previous values $(x = 2026.001, x^* = 2026)$:

$$E_r = \frac{0.001}{2026.001} \approx 4.94 \times 10^{-7}$$

1.1.3 Error Bounds

In practice, the exact value x is often unknown. Therefore, we use an **upper bound** of the error, denoted as Δx.

Definition 1.5

The upper bound Δx is any real number satisfying:

$$|x - x^*| \leq \Delta x \quad \Leftrightarrow \quad x^* - \Delta x \leq x \leq x^* + \Delta x$$

This is commonly written as: $x = x^* \pm \Delta x$.

Remark 1.1

The relative error is often expressed as a percentage (%). The upper bound of the relative error is denoted by $\delta x = \frac{\Delta x}{|x^*|}$.

1.1.4 Significant Figures and Rounding

Definition 1.6

An approximated value x^* is said to have n exact significant digits after the decimal point if:

$$\Delta x \leq 0.5 \times 10^{-n}$$

Example 1.4

Consider the approximation of π using $\frac{22}{7} \approx 3.142857$. The error is $\Delta x = |\pi - \frac{22}{7}| \approx 0.00126$. Since $0.00126 \leq 0.5 \times 10^{-2}$, the hundredths place is significant, and we have 3 total significant digits (3.14).

1.1.5 Rounding Rules

Definition 1.7

Rounding is the process of reducing the digits of a number while trying to keep its value as close as possible to the original. To round a number to n decimal places:

- If the $(n+1)^{th}$ digit is **less than 5**, we keep the n^{th} digit as it is (Round Down).
- If the $(n+1)^{th}$ digit is **5 or greater**, we increase the n^{th} digit by 1 (Round Up).

Example 1.5

Round the following numbers to **two** decimal places ($n = 2$):

1. $x = 12.4532$: The third decimal digit is 3 (< 5), so we round down: $x^* = 12.45$.
2. $y = 12.4578$: The third decimal digit is 7 (≥ 5), so we round up: $y^* = 12.46$.

1.2 Root Isolation

Before applying any numerical method, we must first locate an interval $[a, b]$ that contains the root α. This step allows us to restrict the search area and ensure that numerical methods converge to the correct solution. This process is known as **Root Isolation**.

1.2.1 Analytical Method (Intermediate Value Theorem)

To prove that a solution exists and is unique in an interval $I = [a, b]$, we verify the following two conditions:

Theorem 1.1

Let $f(x)$ be a continuous function on the closed interval $[a, b]$.

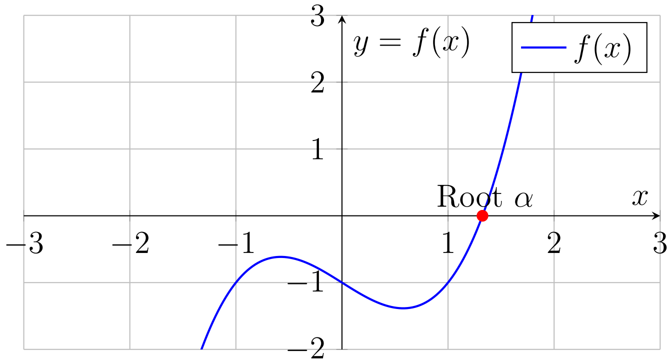
1. **Existence:** If $f(a) \times f(b) < 0$, then there exists at least one root $\alpha \in (a, b)$ such that $f(\alpha) = 0$.

2. **Uniqueness:** If f is differentiable on (a, b) and $f'(x)$ maintains a constant sign (either $f'(x) > 0$ or $f'(x) < 0$) for all $x \in [a, b]$, then the root α is unique.

1.2.2 Graphical Method

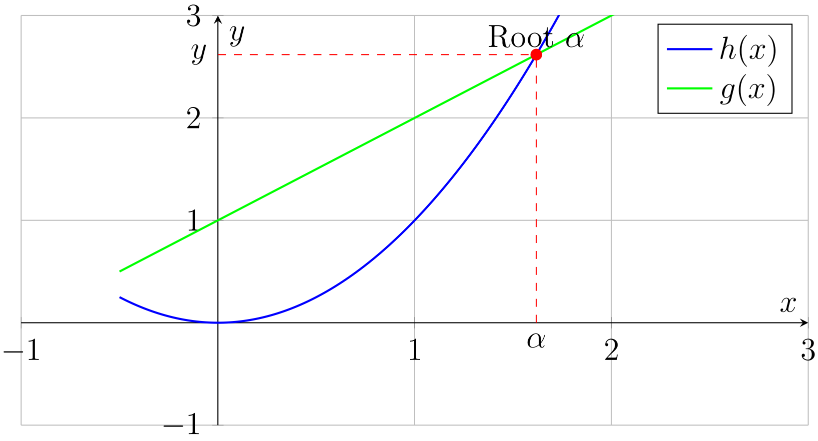
The graphical method provides a visual estimation of the location of roots and is often used as a preliminary step.

- **Single Function:** Plot the graph of $y = f(x)$ and observe where the curve intersects the x-axis. The abscissa of the intersection point gives an approximation of the root.



Here, the root is the point where the curve crosses the x-axis.

- **Function Splitting:** Rewrite the equation $f(x) = 0$ in the form $h(x) = g(x)$. The root corresponds to the x-coordinate of the intersection point of the graphs $y_1 = h(x)$ and $y_2 = g(x)$.



The root α is the x-coordinate of the intersection point of $h(x)$ and $g(x)$.

1.3 Bisection Method (Dichotomy Method)

The objective of the bisection method is to construct a sequence of nested intervals with decreasing lengths, each containing a root of the equation $f(x) = 0$.

1.3.1 Principle of the Method

The bisection method is based on the Intermediate Value Theorem.

Let $f(x) = 0$ and let α be a root of f located in the interval $[a, b]$, where f is continuous and

$$f(a) \times f(b) < 0.$$

The method consists of the following steps:

1. Divide the interval $[a, b]$ into two equal parts and define the midpoint
$$x_0 = \frac{a+b}{2}.$$
2. Two subintervals are obtained: $[a, x_0]$ and $[x_0, b]$. Since α is a root, it must belong to one of them:
   - If $f(a) \times f(x_0) < 0$, then $\alpha \in [a, x_0]$.
   - If $f(x_0) \times f(b) < 0$, then $\alpha \in [x_0, b]$.
3. The new interval $[a_1, b_1]$ is defined as:
$$a_1 = \begin{cases} a & \text{if } \alpha \in [a, x_0], \\ x_0 & \text{if } \alpha \in [x_0, b], \end{cases} \qquad b_1 = \begin{cases} x_0 & \text{if } \alpha \in [a, x_0], \\ b & \text{if } \alpha \in [x_0, b]. \end{cases}$$

By repeating this procedure, we obtain a sequence of midpoints

$$x_1 = \frac{a_1+b_1}{2}, \quad x_2 = \frac{a_2+b_2}{2}, \ \ldots, \ x_n = \frac{a_n+b_n}{2},$$

which converges to the root α. The sequence $(x_n)_{n\in\mathbb{N}}$ provides successive approximations of the root α. As the length of the interval $[a_n, b_n]$ decreases, the approximation becomes more accurate.

1.3.2 Stopping Criterion

The following proposition establishes a rigorous mathematical guarantee for the convergence. It allows us to pre-calculate exactly how many steps are required to achieve a specific level of accuracy, ensuring that our numerical approximation is within the desired tolerance ε.

Proposition 1.1

Let $f(x)$ be a continuous function on $[a, b]$ such that $f(a) \times f(b) < 0$. The number of iterations n required to obtain an approximate root x_n with a precision ε, such that $|x_n - \alpha| < \varepsilon$, is given by:

$$n > \frac{\ln(b-a) - \ln(\varepsilon)}{\ln 2} - 1$$

In practice, the minimum number of iterations is:

$$n = \left\lceil \frac{\ln\left(\frac{b-a}{\varepsilon}\right)}{\ln 2} - 1 \right\rceil + 1$$

Example 1.6

Consider the function $f(x) = x^2 - 2$ defined on the interval $I = [1, 2]$.

1. Prove that the equation $f(x) = 0$ has a unique solution α in the interval $[1, 2]$.
2. Determine the number of iterations n required to approximate α with a precision $\varepsilon = 0.2$.
3. Calculate the approximate value of the root using the bisection method.

Solution

1. Existence and Uniqueness

To establish the existence and uniqueness of the root α:

- **Continuity:** $f(x) = x^2 - 2$ is a polynomial function, hence it is continuous on $[1, 2]$.
- **Existence:** We have $f(1) = 1^2 - 2 = -1 < 0$ and $f(2) = 2^2 - 2 = 2 > 0$. Since $f(1) \times f(2) < 0$, by the *Intermediate Value Theorem (IVT)*, there exists at least one root $\alpha \in [1, 2]$.
- **Uniqueness:** The derivative $f'(x) = 2x$ is strictly positive for all $x \in [1, 2]$. Thus, f is strictly increasing, which guarantees the **uniqueness** of the root α.

2. Stopping Criterion Application

The number of iterations n must satisfy:

$$n = \left\lceil \frac{\ln\left(\frac{b-a}{\varepsilon}\right)}{\ln 2} - 1 \right\rceil + 1$$

Substituting $a = 1, b = 2$, and $\varepsilon = 0.2$:

$$\begin{aligned} n &= \left\lceil \frac{\ln\left(\frac{1-0}{0.2}\right)}{\ln 2} - 1 \right\rceil + 1 \\ &= \left\lceil \frac{\ln(5)}{\ln 2} - 1 \right\rceil + 1 \\ &= \lceil 2.3219 - 1 \rceil + 1 \\ &= \lceil 1.3219 \rceil + 1 \\ &= 1 + 1 = 2 \end{aligned}$$

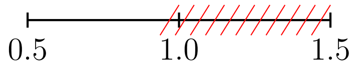
3. Iterative Calculation

We perform the bisection method steps:

1. **n=0:** $a = 0.5$ and $b = 1.5$. $x_0 = \frac{0.5+1.5}{2} = 1.0$. $f(a) = f(1) = 1^2 - 2 = -1$ $f(x_0) = f(1.5) = 1.5^2 - 2 = 0.25$

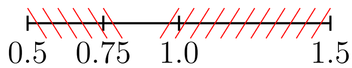
   Since $f(a) \times f(x_0) = (-1) \times (0.25) = -0.25 < 0$, the root lies in :

   $\alpha \in]0.5, 1.0[$

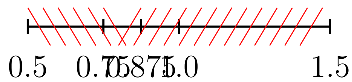
2. **n=1:** $f(1.0) = 0$, update $b_1 = 1.0$. $x_1 = \frac{0.5+1.0}{2} = 0.75$, so $a_2 = 0.75$ and $b_1 = 1.0$. The error: $|x_1 - x_0| = 0.25 > 0.15$. $f(a_1) = f(1) = 1^2 - 2 = -1$ $f(x_1) = f(1.25) = 1.25^2 - 2 = -0.4375$. Since $f(a_1) \times f(x_1) > 0$, then The root α lies in:

   $\alpha \in]0.75, 1.0[$

3. **n=2:** $f(0.75) > 0$, update $a_2 = 0.75$. $x_2 = \frac{0.75+1.0}{2} = 0.875$. Error: $|x_2 - x_1| = 0.125 < 0.15$.

   $\alpha \in]0.75, 0.875[$

Exercise 1.1:

Consider the function $f(x)$ defined on the interval $I = [0.1, 0.5]$ by:

$$f(x) = \ln(x) + x^2 - 2 = 0$$

1. Prove that the equation $f(x) = 0$ has a unique solution α in the interval $[01, 0.5]$.
2. Determine the number of iterations n required to approximate α with a precision of $\varepsilon = 0.13$ using the bisection method.
3. Calculate the approximate value of the root α after 3 iterations of the bisection method.
4. Apply the Matlab function provided below `bissection.m` to find the approximate solution of $f(x) = 0$, then compare the results with your manual calculations.

```
function [x, y, niter] = bissection(f, a, b, eps)
    if f(a) * f(b) > 0
        error('Inappropriate interval');
        return
    end

    niter = 0;
    while abs(b - a) > eps
        x = (a + b) / 2;

        if f(x) == 0
            niter = niter + 1;
            y = 0;
            return
        elseif f(a) * f(x) > 0
            a = x;
        else
            b = x;
        end

        niter = niter + 1;
    end
    y = f(x);
end
```

1.4 Fixed Point Method

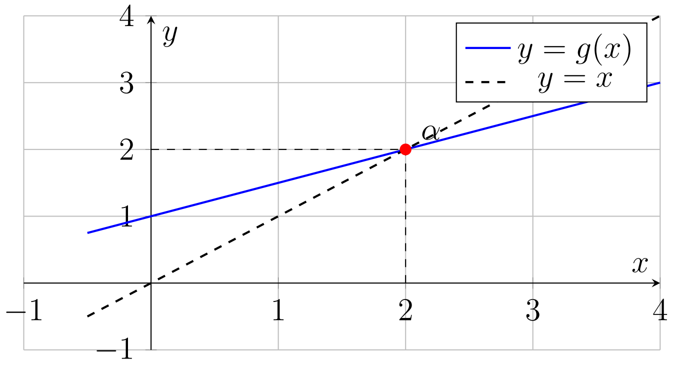
This method is based on the principle of the fixed point of a function.

Definition 1.8

Let g be a continuous function defined on a closed interval $[a, b]$. A point $\bar{x}$ is called a fixed point of the function g if it satisfies the following condition:

$$\bar{x} = g(\bar{x}), \quad \text{where } \bar{x} \in [a, b] \tag{1.1}$$

Geometrically, the fixed point is the intersection of the curve $y = g(x)$ and the line $y = x$.



Geometric interpretation of a fixed point

1.4.1 Principle of the method

The Fixed-Point Iteration method is a numerical technique used to find the roots of an equation $f(x) = 0$ by transforming it into a fixed-point problem.

1. The first step is to rewrite the original equation $f(x) = 0$ in the form:

$$x = g(x) \tag{1.2}$$

   where x is isolated on one side of the equation.

2. Choose an initial approximation x_0 for the root $\bar{x}$. This is the starting point of the sequence.

3. We define the sequence $(x_n)_{n\geq 0}$ using the recursive formula:

$$\begin{cases} x_0 & \text{given initial approximation} \\ x_{n+1} = g(x_n), & n = 0, 1, 2, \ldots \end{cases} \tag{1.3}$$

1.4.2 Convergence Criterion

Proposition 1.2

Let $g : [a, b] \to [a, b]$ be a differentiable function such that

$$|g'(x)| \leq k < 1 \quad \forall x \in [a, b].$$

Then:

1. g has a unique fixed point $\bar{x} \in [a, b]$.
2. For any initial value $x_0 \in [a, b]$, the sequence

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, \ldots$$

converges to $\bar{x}$.

Remark 1.2

If multiple forms of g satisfy this condition, we will have several values of k. We choose the one with the smallest value of k. In practice, we compute $k = \max_{x \in [a,b]} |g'(x)|$, which must be less than 1 for the method to converge.

1.4.3 Stopping Criterion

Proposition 1.3

Let $(x_n)_{n \geq 0}$ be a sequence generated by the fixed-point iteration

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, \ldots$$

and let $\varepsilon > 0$ be a prescribed tolerance. The iteration is stopped at the first index n such that

$$|x_{n+1} - x_n| < \varepsilon.$$

Then x_{n+1} can be taken as an approximate solution to the fixed-point equation $x = g(x)$ within the precision ε.

Exapmle 1.7

Consider the function

$$f(x) = \ln(x) - x^2 + 2$$

defined on the interval $[0.1, 0.5]$ with a precision $\varepsilon = 0.01$.

1. Rewrite the equation $f(x) = 0$ in the form $x = g(x)$.
2. Show that the function

$$g(x) = e^{x^2 - 2}$$

satisfies the convergence conditions ($g([0.1, 0.5]) \subset [0.1, 0.5]$ and $|g'(x)| < 1$ on $[0.1, 0.5]$.)

3. Compute the approximate solution using the fixed-point iteration method starting from $x_0 = 0.3$.

Solution

1. Rewrite the equation in the form $x = g(x)$:

$$x = g_1(x) = \ln(x) - x^2 + 2 + x,$$
$$x = g_2(x) = e^{x^2-2}.$$

2. Verify the convergence condition for each g:

   1. $g_1(x) = \ln(x) - x^2 + 2 + x$:

$$g_1(0.1) = -0.21259 \notin [0.1, 0.5] \quad \Rightarrow \text{not suitable.}$$

   2. $g_2(x) = e^{x^2-2}$: - Check interval inclusion:

$$g_2(0.1) = 0.136 > 0.1, \quad g_2(0.5) = 0.173 < 0.5$$
$$\Rightarrow g_2([0.1, 0.5]) \subset [0.1, 0.5].$$

   - Check derivative for convergence:

$$g_2'(x) = 2xe^{x^2-2} \quad \text{and} \quad g_2'(x) > 0 \text{ on } [0.1, 0.5]$$
$$k = \max_{x\in[0.1,0.5]} |g_2'(x)| = g_2'(0.5) = 0.173 < 1$$

   $\Rightarrow$ Convergence criterion satisfied.

3. Iterative scheme:

$$\begin{cases} x_0 \in [0.1, 0.5], \\ x_{n+1} = g_2(x_n) = e^{x_n^2-2}, \quad n = 0, 1, 2, \ldots \end{cases}$$

Choose $x_0 = 0.3$ (midpoint of the interval).
4. Computation with stopping criterion $|x_{n+1} - x_n| < \varepsilon$

n	x_n	x_{n+1}	$e_n = \|x_{n+1} - x_n\|$	$e_n < 10^{-2}$?
0	0.30000	0.14808	0.15192	No
1	0.14808	0.13834	0.00974	Yes

5. Result:

$$x^* \approx 0.13834, \quad \text{with } \alpha \in (0.12834, 0.14834).$$

Exercise 1.2: C

onsider the equation $f(x) = x^3 + 4x - 10 = 0$ on the interval $I = [1, 2]$. We aim to find the root α using the fixed-point method.

1. Determine the possible functions $g(x)$ that can be derived from the equation $f(x) = 0$.

2. Show that $g(x) = \frac{10}{x^2+4}$ satisfies the convergence conditions on I.

3. Write the iterative scheme and calculate the approximate value of the root α using a precision of $\varepsilon = 0.1$.

4. Apply the Matlab function `FixedPoint.m` provided below to find the approximate solution, then compare the results with your manual calculations.

```
function [alpha, error, niter] = pointFixe(g, x0, itmax, eps)
    alpha = x0;
    for niter = 1 : itmax
        x = alpha;
        alpha = g(x);
        error = abs(alpha - x);
        if error < eps
            return
        end
    end
    warning('The fixed-point method did not converge');
end
```

1.5 Newton-Raphson Method

The Newton-Raphson method is the most efficient and widely used. It relies on the Taylor series expansion.

1.5.1 Principle of the method

Proposition 1.4

Let f be a real-valued function such that $f \in C^2[a,b]$ and $f(\bar{x}) = 0$ for some root $\bar{x} \in (a,b)$. If $f'(x) \neq 0$ for all $x \in [a,b]$, then the iterative sequence $\{x_n\}_{n=0}^{\infty}$ defined by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots \tag{1.4}$$

converges to $\bar{x}$, provided the initial guess x_0 is sufficiently close to $\bar{x}$.

Proof 1.1

Consider the Taylor series expansion of $f(x)$ at the root $\bar{x}$ centered around the current estimate x_n:

$$f(\bar{x}) = f(x_n) + (\bar{x} - x_n)f'(x_n) + \frac{(\bar{x} - x_n)^2}{2!}f''(\xi)$$

where ξ lies between x_n and $\bar{x}$. Since $\bar{x}$ is a root, $f(\bar{x}) = 0$. Substituting this into the expansion gives:

$$0 = f(x_n) + (\bar{x} - x_n)f'(x_n) + \frac{(\bar{x} - x_n)^2}{2!}f''(\xi)$$

If x_n is sufficiently close to $\bar{x}$, the second-order term $\frac{(\bar{x}-x_n)^2}{2!}f''(\xi)$ is negligible. Thus, we obtain the linear approximation:

$$0 \approx f(x_n) + (\bar{x} - x_n)f'(x_n)$$

Solving for $\bar{x}$ yields:

$$\bar{x} \approx x_n - \frac{f(x_n)}{f'(x_n)}$$

By setting the next approximation x_{n+1} to this value, we arrive at the recursive formula:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

1.5.2 Convergence Criterion

Proposition 1.5

Let f be a function defined on $[a, b]$ such that:

- i. $f(a) \times f(b) < 0$.
- ii. $f'(x)$ and $f''(x)$ are nonzero and have a constant sign on the given interval.

Then the Newton-Raphson method converges.

Remark 1.3

If the convergence condition is satisfied, the iterative process of the Newton-Raphson method converges. This means that each new iteration is closer to the solution than the previous one.

1.5.3 Stopping Criterion

Proposition 1.6

Let $\varepsilon > 0$ be a given precision. We stop the iteration when the absolute difference between two successive approximations satisfies:

$$|x_{n+1} - x_n| \leq \varepsilon.$$

In this case, we take x_{n+1} as the approximate solution of the equation $f(x) = 0$ using the Newton-Raphson method.

Example 1.8

Let the function be:

$$f(x) = x(1+e^x) - e^x$$

find the real root of f using the Newton-Raphson method on $[0.5, 1]$ with a precision of $\varepsilon = 10^{-3}$.

1. Convergence Conditions:

$$\begin{aligned}
&f(x) = x(1+e^x) - e^x = 0 \\
&f'(x) = 1 + xe^x \quad \Rightarrow \quad f'(x) > 0 \ \ \forall x \in [0.5, 1] \\
&f''(x) = e^x(1+x) \quad \Rightarrow \quad f''(x) \neq 0 \ \ \forall x \in [0.5, 1] \\
&\text{At } x_0 = 0.5 : f(0.5) \cdot f''(0.5) = (-0.3243) \cdot (2.473) < 0 \\
&\text{At } x_0 = 0.8 : f(0.8) \cdot f''(0.8) = (0.355) \cdot (4.006) > 0
\end{aligned}$$

2. Newton-Raphson Iteration:

   The iteration formula is:

$$\begin{cases} x_0 = 0.8 \\ x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)} \end{cases}$$

n	x_n	x_{n+1}	$\lvert x_{n+1} - x_n \rvert$
0	0.8000	0.6724	$0.1276 > 10^{-3}$
1	0.6724	0.6592	$0.0132 > 10^{-3}$
2	0.6592	0.6590	$0.0001 < 10^{-3}$

3. Result:

$$x^* \approx 0.6590, \quad \text{with } x^* \in [0.6580, 0.6600] \quad (\text{precision } \varepsilon = 10^{-3}).$$

Exercise 1.3: C

onsider the nonlinear equation $f(x) = x^3 - x - 2$.

1. Show that the equation $f(x) = 0$ has a unique root in the interval $[1, 2]$.
2. Derive the expression of $f'(x)$.
3. Write the Newton-Raphson iterative formula:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

4. Starting with $x_0 = 1.5$, compute the approximate solution with precision 10^{-3}.

5. Apply the Matlab function `NewtonRaphson.m` provided below to find the approximate solution, then compare the results with your manual calculations.

```
function [root, error, niter] = NewtonRaphson(f, df, x0, tol, maxIter
)
    root = x0;
    for niter = 1:maxIter
        fx = f(root);
        dfx = df(root);
        root = root - fx / dfx;
        error = abs(f(root));
        if error < tol
            return
        end
    end
end
```