Tomi Kis 11-11-19

11 Nonlinear Equations

In this chapter we consider the problem of finding zeros of a continuous function f, i.e. solving $f(x) = 0$ for example $e^x - x = 0$ or a system of nonlinear equations.

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0, \\ f_2(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ f_n(x_1, \ldots, x_n) &= 0 \end{aligned} \tag{11.1}$$

We are going to write this generic system in vector form as

$$\mathbf{f}(\mathbf{x}) = 0 \tag{11.2}$$

where $\mathbf{f} : U \subset \mathbb{R}^n \to \mathbb{R}^n$. Unless otherwise noted the function f is assumed to be smooth in its domain U. We are going to start with the scalar case, $n = 1$ and look a very simple but robust method that relies only on the continuity of the function and the existence of a zero.

11.1 Bisection Method

Suppose we are interested in solving a nonlinear equation in one unknown

$$f(x) = 0 \tag{11.3}$$

Where f is a continuous function on an interval $[a, b]$ and has at least one zero there. Also suppose that f has values of different sign at the end points of the interval, i.e,

$$f(a)f(b) < 0 \tag{11.4}$$

By the Intermediate Value Theorem, f has at least one zero in (a, b). To locate a zero we bisect the interval and check on which subinterval f changes sign. We repeat the process until we bracket a zero within a desired accuracy. The Bisection algorithm to find a zero x^* is shown below.

Algorithm: The Bisection Method

Given f, a and b, $a < b$ and Tolerance = TOL
While (k <max iteration) and $(b - a)$ >TOL
    $c = (a + b)/2$
    if $f(c) = 0$
        $x^* = c$, return
    if $\text{sgn}(f(c)) = \text{sgn}(f(a))$
        $a = c$
    else
        $b = c$
    k = k +1

$x^* = (a+b)/2$

11.1.1 Convergence of Bisection Method

With the bisection method we generate a sequence

$$c_k = \frac{a_k + b_k}{2}, \qquad , k = 1, 2, \ldots \tag{11.5}$$

where each a_k and b_k are the endpoints of the subinterval we select at each bisection step, as f changes sign there. Since

$$b_k - a_k = \frac{b-a}{2^{k-1}}, k = 1, 2, \ldots \tag{11.6}$$

and c_k is the midpoint of the interval, we have

$$|c_k - x^*| \leq \frac{b_k - a_k}{2} \leq \frac{b-a}{2^k} \tag{11.7}$$

and consequently $c_k \to x^*$, a zero of f in $[a, b]$.

11.2 Rate of convergence

We now define in precise terms the rate of convergence of a sequence of approximations to a value x^*.

Definition 1 *Suppose a sequence $\{x_n\}$ converges to x^* as $n \to \infty$. We say that $x_n \to x^*$ of order $p (p \geq 1)$ if there is a positive integer N and a constant C such that*

$$|x_{n+1} - x^*| \leq C|x_n - x^*|^p, \text{ for all } n \geq N \tag{11.8}$$

or equivalently

$$\lim_{n\to\infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^p} \leq C. \tag{11.9}$$

If $p = 1$ then $C < 1$.

Example 1 *The sequence generated by the bisection method converges linearly to x^* because*

$$\frac{|c_{n+1} - x^*|}{|c_n - x^*|} \leq \frac{\frac{b-a}{2^{n+1}}}{\frac{b-a}{2^n}} = \frac{1}{2} \tag{11.10}$$

Let's examine the significance of the rate of convergence. Consider first, $p = 1$, linear convergence. Suppose

$$|x_{n+1} - x^*| \approx C|x_n - x^*|, \qquad n \leq N,. \tag{11.11}$$

Then

$$|x_{n+1} - x^*| \approx C|x_n - x^*|, \tag{11.12}$$
$$|x_{n+2} - x^*| \approx C|x_{n+1} - x^*| \approx C^2|x_n - x^*|, \tag{11.13}$$

So we have

$$|x_{n+k} - x^*| \approx C^k|x_n - x^*|, \qquad n \leq N, \tag{11.14}$$

and this is the reason of the requirement $C < 1$ for $p = 1$. If the error at the N step, $|x_N - x^*|$, is small enough it will be reduced by a factor of C^k after k more steps. Setting $C^k = 10^{-d_k}$, then the error $|x_N - x^*|$ will be reduced approximately

$$d_k = -k\log(C) \tag{11.15}$$

digits. Let us now do a similar analysis for $p = 2$, quadratic convergence. We have

$$|x_{n+1} - x^*| \approx C|x_n - x^*|^2, \tag{11.16}$$
$$|x_{n+2} - x^*| \approx C|x_{n+1} - x^*|^2 \approx C(C|x_n - x^*|^2)^2 = C^3|x_n - x^*|^4 \tag{11.17}$$

So we have

$$|x_{n+k} - x^*| \approx C^{2^k-1}|x_n - x^*|^{2^k}, k = 0, 1, \ldots \tag{11.18}$$

To see how many digits of accuracy we gain in k steps beginning from x_N, we write

$$C_{2^k-1}|x_N - x^*|^{2^k} = 10^{-d_k}|x_N - x^*| \tag{11.19}$$

and solving for d_k we get

$$d_k = -\left(\log(C) + \log(|x_N - x^*|)\right)(2^k - 1) \tag{11.20}$$

Then for $p > 1$ and as $k \to \infty$, we have $d_k = \alpha_p p^k$, where

$$\alpha_p = \frac{1}{1-p}\big(\log(C) + \log(|x_N - x^*|)\big). \tag{11.21}$$

11.2.1 Interpolation-Based Methods

Assuming again that f is a continuous function in $[a, b]$ and $f(a)f(b) < 0$ we can proceed as in the bisection method but instead of using the midpoint $c = (a + b)/2$ to subdivide the interval in question we could use the root of linear polynomial interpolating $(a, f(a))$ and $(b, f(b))$. This is called the *method of false position*. Unfortunately, this method only converges linearly and under stronger assumptions than the Bisection Method.

An alternative approach to use interpolation to obtain numerical methods for $f(x) = 0$ is to proceed as follows: Given $m + 1$ approximations to the zero of $f, x_0, \ldots, x_m$, construct the interpolating polynomial of f, p_m , at those points, and set the root of p_m closest to x_m as the new approximation to the zero of f. In practice, only $m = 1, 2$ are used. The method for $m = 1$ is called the Secant method and we will look at it in some detail shortly. The method for $m = 2$ is called Muller's Method.

11.3 Newton's Method

If the function f is smooth, say at least $C^2[a,b]$, and we have already a good approximation x_0 to a zero x^* of f then the tangent line of f at x_0, $y = f(x_0) + f'(x_0)(x - x_0)$ provides a good approximation to f in a small neighborhood of x_0, i.e.

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \tag{11.22}$$

Then we can define the next approximation as the zero of that tangent line, i.e.

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}, \tag{11.23}$$

At the k step or iteration we get the new approximation x_{k+1} according to:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, \ldots \tag{11.24}$$

This iteration is called Newton's method or Newton-Raphson's method. There are some conditions for this method to work and converge. But when it does converge it does it at least quadratically. Indeed, a Taylor expansion of f around x_k gives

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{f''(\xi_k)}{2}(x - x_k)^2 \tag{11.25}$$

where ξ_k is a point between x and x_k. Evaluating at $x = x^*$ and using that $f(x^*) = 0$ we get

$$0 = f(x_k) + f'(x_k)(x^* - x_k) + \frac{f''(\xi_k)}{2}(x - x_k)^2 \tag{11.26}$$

which we can recast as

$$x^* = x^k - \frac{f(x_k)}{f'(x_k)} + \frac{f''(\xi_k)}{2f'(x_k)}(x - x_k)^2 = x_{k+1} - \frac{f''(\xi_k)}{2f'(x_k)}(x - x_k)^2 \tag{11.27}$$

Thus

$$|x_{k+1} - x^*| = \frac{|f''(\xi_k)|}{2|f'(x_k)|}|x - x_k|^2 \tag{11.28}$$

So if the sequence $\{x_k\}$ generated by Newton's method converges then it does so at least quadratically.

Theorem 1 *Let x^* be a simple zero of f, i.e. $f(x^*) = 0$ and $f'(x^*) \neq 0$, and suppose $f \in C^2$. Then there's a neighborthood I_ϵ of x^* such that Newton's method converges to x^* for any initial guess in I_ϵ .*

Proof. Since f' is continuous and $f'(x^*) \neq 0$ we can choose $\epsilon > 0$, sufficiently small so that $f'(x) \neq 0$ for all x such that $|x - x^*| \leq \epsilon$, let this be I_ϵ and that $\epsilon M(\epsilon) < 1$ where

$$M(\epsilon) = \frac{\max_{x \in I_\epsilon} |f''(x)|}{2 \min_{x \in I_\epsilon} |f'(x)|}. \tag{11.29}$$

This is possible as $f \in C^2(I_\epsilon)$ and

$$\lim_{\epsilon \to 0} M(\epsilon) = \frac{|f''(x^*)|}{|f'(x^*)|} \tag{11.30}$$

The condition $\epsilon M(\epsilon) < 1$ allows us to guarantee that x^* is the only zero of f in I_ϵ, as we show now. A Taylor expansion of f around x^* gives

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{f''(\xi)}{2}(x - x^*)^2 \tag{11.31}$$

$$f(x) = f'(x^*)(x - x^*)\left(1 + \frac{f''(\xi)}{2f'(x^*)}(x - x^*)\right) \tag{11.32}$$

and since

$$\left|\frac{f''(\xi)}{2f'(x^*)}(x - x^*)\right| = |(x - x^*)|\left|\frac{f''(\xi)}{2f'(x^*)}\right| \leq \epsilon M(\epsilon) < 1 \tag{11.33}$$

for all $f(x) \neq 0$ and for all $x \in I_\epsilon$, unless $x = x^*$. We will now show that Newton's iteration is well defined starting from any initial guess $x_0 \in I_\epsilon$. We prove this by induction. From (11.28) with $k = 0$ it follow that $x_1 \in I_\epsilon$ as

$$|(x_1 - x^*)| \leq |(x_0 - x^*)|^2 \left|\frac{f''(\xi_0)}{2f'(x_0)}\right| \leq \epsilon^2 M(\epsilon) < \epsilon \tag{11.34}$$

Now assume that $x_k \in I_\epsilon$ from (11.28) we have

$$|(x_{k+1} - x^*)| \leq |(x_k - x^*)|^2 \left|\frac{f''(\xi_k)}{2f'(x_k)}\right| \leq \epsilon^2 M(\epsilon) < \epsilon \tag{11.35}$$

so $x_{k+1} \in I_\epsilon$. Now,

$$\begin{aligned}
|(x_{k+1} - x^*)| &\leq |(x_k - x^*)|^2 M(\epsilon) < \epsilon\,|(x_k - x^*)|\,\epsilon M(\epsilon) && (11.36)\\
&\leq |(x_{k-1} - x^*)|\,(\epsilon M(\epsilon))^2 && (11.37)\\
&\ \ \vdots && (11.38)\\
\leq |(x_0 - x^*)|\,(\epsilon M(\epsilon))^{k+1} & && (11.39)\\
& && (11.40)
\end{aligned}$$

as since $\epsilon M(\epsilon) < 1$ is follows that $x_k \to x^*$ as $k \to \infty$. □

The need for a good initial guess x_0 for Newton's method should be emphasized. In practice, this is obtained with another method, like bisection.

11.4 The Secant Method

Sometimes it could be computationally expensive or not possible to evaluate the derivative of f. The following method, known as the secant method, replaces the derivative by the secant:

$$x_{k+1} = x_k - \frac{f(x_k)}{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}}, k = 1, 2, \ldots \tag{11.41}$$

Note that since $f(x^*) = 0$, we have

$$\begin{aligned}
x_{k+1} - x^* =& x_k - x^* - \frac{f(x_k)}{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}} \\
=& x_k - x^* - \frac{f(x_k)}{f[x_k, x_{k-1}]} \\
=& (x_k - x^*) \left(1 - \frac{f[x_k, x^*]}{f[x_k, x_{k-1}]}\right) \\
=& (x_k - x^*) \left(\frac{f[x_k, x_{k-1}] - f[x_k, x^*]}{f[x_k, x_{k-1}]}\right) &(11.42)\\
=& (x_k - x^*)(x_{k-1} - x^*) \left(\frac{f[x_{,k-1}, x_k, x^*]}{f[x_k, x_{k-1}]}\right) &(11.43)
\end{aligned}$$

Now if $x_k \to x^*$, then

$$\frac{f[x_{,k-1}, x_k, x^*]}{f[x_k, x_{k-1}]} \to \frac{f''(x^*)}{2f'(x^*)} \tag{11.44}$$

and

$$\lim_{k \to \infty} \frac{x_k - x^*}{x_{k-1} - x^*} = 0, \tag{11.45}$$

i.e. the sequence generated by the secant method would converge faster than linear. Define $e_k = |x_k - x^*|$, the above calculation shows that

$$e_{k+1} \approx c e_k e_{k-1}. \tag{11.46}$$

Let's try to determine the rate of convergence of the secant method. Starting with the ansatz $e_k \approx A e_{k-1}^p$ or equivalently $e_{k-1} = \left(\frac{e_k}{A}\right)^{1/p}$ we have

$$e_{k+1} \approx c e_k \left(\frac{e_k}{A}\right)^{1/p}, \tag{11.47}$$

which implies

$$\frac{A^{1+\frac{1}{p}}}{c} \approx c e_k^{1-p+\frac{1}{p}} \tag{11.48}$$

Since the left hand side is a constant we must have $1 - p + \frac{1}{p} = 0$ which gives $p = \frac{1 \pm \sqrt{5}}{2}$, thus

$$p = \frac{1 + \sqrt{5}}{2} \approx 1.61803. \tag{11.49}$$

gives the rate of convergence of the secant method. It is better than linear, but worse than quadratic. Sufficient conditions for local convergence are as in Newton's method. In general, this is called *super linear* convergence.

11.5 Fixed Point Iteration

Newton's method is a particular example of a functional iteration of the form

$$x_{k+1} = g(x_k), \qquad k = 0, 1, \dots \tag{11.50}$$

with the particular choice of $g(x) = x - \frac{f(x)}{f'(x)}$. Clearly, if x^* is a zero of f, then x^*. is a fixed point of g, i.e. $g(x^*) = x^*$. We will look at fixed point iterations as a tool for solving $f(x) = 0$.

Example 2 *Suppose we want to solve for $f(x) = x - e^{-x} = 0$ in $[0, 1]$. Then if we take $g(x) = e^{-x}$, a fixed point of g corresponds to a zero of f.*

Definition 2 *Let g be defined in an interval $[a, b]$. We say that g is a contraction or a contractive map if there is a constant L with $0 \leq L < 1$ such that*

$$|g(x) - g(y)| \leq L|x - y|, \text{ for all } x, y \in [a, b] \tag{11.51}$$

If x^* is a fixed point of g in $[a, b]$ then

$$\begin{aligned} |x_k - x^*| &= |g(x_{k-1}) - g(x^*)| \\ &\leq L|x_{k-1} - x^*| \\ &\vdots \\ &\leq L^k|x_0 - x^*| \to 0, \text{ as } k \to \infty \end{aligned} \tag{11.52}$$

Theorem 2 *If g is a contraction on $[a, b]$ and maps $[a, b]$ into $[a, b]$ then g has a unique fixed point x^* in $[a, b]$ and the fixed point iteration converges to x^* for any $[a, b]$. Moreover*

$$(1) \qquad |x_k - x^*| \leq \frac{L}{1 - L}|x_1 - x_0| \tag{11.53}$$

$$(2) \qquad |x_k - x^*| \leq L^k|x_0 - x^*| \tag{11.54}$$

Proof. (2) has been shown already. Since $g : [a, b] \to [a, b]$, the fixed point iteration $x_{k+1} = g(x_k), k = 0, 1, \dots$ is well-defined and

$$\begin{aligned} |x_{k+1} - x_k| &= |g(x_k) - g(x_{k-1}| \\ &\leq L|x_k - x_{k-1}| \qquad (11.55) \\ &\vdots \\ &\leq L^k|x_1 - x_0|. \qquad (11.56) \end{aligned}$$

Now for $n \geq m$

$$x_n - x_m = x_n - x_{n-1} + x_{n-1} + \cdots + x_{m+1} - x_m \tag{11.57}$$

and so

$$
\begin{aligned}
|x_n - x_m| \leq& |x_n - x_{n-1}| + |x_{n-1} - x_{n-2}| + \cdots + |x_{m+1} - x_m| && (11.58)\\
\leq& L^{n-1}|x_1 - x_0| + L^{n-2}|x_1 - x_0| + \cdots + L^m|x_1 - x_0| && (11.59)\\
\leq& L^m|x_1 - x_0|(1 + L + L^2 + \cdots + L^{n-1-m}) && (11.60)\\
\leq& L^m|x_1 - x_0|\sum_{j=0}^{\infty} L^j = \frac{L^m}{1-L}|x_1 - x_0|. && (11.61)
\end{aligned}
$$

Thus given $\epsilon > 0$ there is an N such that

$$\frac{L^m}{1-L}|x_1 - x_0| \leq \epsilon \tag{11.62}$$

and thus for $N \geq m \geq N$, $|x_n - x_m| \leq \epsilon$ i.e., $\{x_n\}$ is a Cauchy sequence in $[a, b]$ so it converges to a point $x^* \in [a, b]$. But

$$|x_k - g(x^*)| = |g(x_{k-1}) - g(x^*)| \leq L|x_{k-1} - g(x^*)| \tag{11.63}$$

and so $x_k \to g(x^*)$ as $k \to \infty$, that is, x^* is a fixed point of g.

Now suppose that there are two fixed points, $x_1, x_2 \in [a, b]$.

$$|x_1 - x_2| = |g(x_1) - g(x_2)| \leq L|x_1 - x_2| \tag{11.64}$$

This implies we have

$$(1 - L)|x_1 - x_2| \leq 0 \tag{11.65}$$

but as $0 \leq L < 1$, then $|x_1 - x_2| = 0$ implies that $x_1 = x_2$, i.e. the fixed point is unique. □

If g is differentiable in (a, b), then by the mean value theorem

$$g(x) - g(y) = g'(\xi)(x - y), \qquad \text{for some}\xi \in [a, b] \tag{11.66}$$

and if the derivative is bounded by a constant $L < 1$ i.e. $|g'(x)| \leq L$ for all $x \in (a, b)$, then we have

$$|g(x) - g(y)| \leq L|x - y| \tag{11.67}$$

with $0 \leq L < 1$, i.e. g is contractive in $[a, b]$.

Example 3 *let $g(x) = (x^2 + 3)/4$ for $x \in [0, 1]$. Then $0 \leq g(x) << 1$ and $g'(x) \leq 1/2$ for all $x \in [0, 1]$. So g is contractive in $[0, 1]$ and the fixed point iteration will converge to the unique fixed point of g in $[0, 1]$.*

Note that

$$x_{k+1} - x^* = g(x_k) - g(x^*) = g'(\xi_k)(x_k - x^*), \text{ for some}\xi_k \in [x_k, x^*] \tag{11.68}$$

Thus

$$\frac{x_{k+1} - x^*}{(x_k - x^*)} = g'(\xi_k) \tag{11.69}$$

and unless $g'(x^*) \neq 0$, the fixed point iteration converges linearly, when it does converge.

11.6 Systems of Nonlinear Equations

We now look at the problem of finding numerical approximation to the solutions of a nonlinear system of equations $\mathbf{f}(\mathbf{x}) = 0$, where $\mathbf{f} : U \subset \mathbb{R}^n \to \mathbb{R}^n$. The main approach to solve a nonlinear system is fixed point iteration

$$\mathbf{x}_{k+1} = \mathbf{G}(\mathbf{x}_k), \qquad k = 0, 1, \ldots \tag{11.70}$$

where we assume that $\mathbf{G}$ is defined on a closed set $B \subset \mathbb{R}^n$ and $\mathbf{G} : B \to B$. The map $\mathbf{G}$ is a contraction, with respect to some norm $\|\cdot\|$, if there is a constant L with $0 \leq L < 1$ and

$$\|\mathbf{G}(\mathbf{x}) - \mathbf{G}(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \qquad \text{for all } \mathbf{x}, \mathbf{y} \in B. \tag{11.71}$$

Then, as we know, by the contraction map principle, $\mathbf{G}$ has a unique fixed point and the sequence generated by the fixed point iteration converges to it.

Suppose that $\mathbf{G} \in C^1(b)$ on some convex set $B \subset \mathbb{R}^n$, for example a ball. Consider the linear segment $\mathbf{x} + t(\mathbf{y} - \mathbf{x})$ for $t \in [0, 1]$ with $\mathbf{x}, \mathbf{y}$ fixed in B. Define the one-variable function

$$\mathbf{h}(t) = \mathbf{G}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})). \tag{11.72}$$

Then by the chain rule,

$$\mathbf{h}'(t) = D\mathbf{G}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) \tag{11.73}$$

Where $D\mathbf{G}$ is the derivative matrix of $\mathbf{G}$. Then, using the definition of $\mathbf{h}$ and the Fundamental Theorem of Calculus we have

$$\mathbf{G}(\mathbf{x}) - \mathbf{G}(\mathbf{y}) = \mathbf{h}(1) - \mathbf{h}(0) - \int_0^1 \mathbf{h}'(t)\, dt \tag{11.74}$$

$$= (\mathbf{y} - \mathbf{x}) \int_0^1 D\mathbf{G}(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\, dt \tag{11.75}$$

Thus if there is an $0 \leq L < 1$ and a norm $\|\cdot\|$ such that

$$\|D\mathbf{G}\| \leq L, \qquad \text{for all } \mathbf{x} \in B \tag{11.76}$$

$\mathbf{G}$ is a contraction, in that norm. The spectral radius of $D\mathbf{G}$, $\rho(D\mathbf{G})$ will determine the rate of convergence of the corresponding fixed point iteration.

11.6.1 Newton's Method

By Taylor theorem

$$\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\mathbf{x_0}) + D\mathbf{f}(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0}) \tag{11.77}$$

so if we take $\mathbf{x}_1$ as the zero of the right hand side of we get

$$\mathbf{x}_1 = \mathbf{x}_0 - [D\mathbf{f}(\mathbf{x_0})]^{-1}\mathbf{f}(\mathbf{x}_0). \tag{11.78}$$

Continuing this way, Newton's method for the system of equations $\mathbf{f}(\mathbf{x}) = 0$ can be written as

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [D\mathbf{f}(\mathbf{x_k})]^{-1}\mathbf{f}(\mathbf{x}_k). \tag{11.79}$$

In the implementation of Newton's method for a system of equations we solve the linear system $D\mathbf{f}(\mathbf{x}_k)\mathbf{v} = -\mathbf{f}(\mathbf{x}_k)$ at each iteration and update $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{v}$.

11.7 Exercises

Exercise 1 *Consider the function $f(x) = x^2 - 2$ in the interval $[1, 2]$.*

1. *Show that that the Bisection Method converges to the zero of $f(x)$ in $[1, 2]$.*
2. *Find x_3 and hence an approximate to $x^* = \sqrt{2}$.*
3. *Compute x_2 in Newton's method to find an approximation of zero of $f(x)$ beginning with $x_0 = 1$.*
4. *Which iteration converges the fastest to $x^* = \sqrt{2}$, Bisection Method or Newton's method? Explain.*

Exercise 2 *The following two methods are proposed to compute $5^{1/3}$*

$$x_k = x_{k-1} - \frac{x_{k-1}^3 - 5}{3x_{k-1}^2}, \qquad x_k = \frac{4x_{k-1} + 5/x_{k-1}^2}{5} \tag{11.80}$$

Explain, based on the theory in the notes, which method is expected to converge the fastest for a sufficiently good initial guess x_0.

Exercise 3 *Consider the equation $x^2 + \cos(x) - 10x = 0$ for $x \in [0, 1]$.*

1. *Show that a solution of this equation is a fixed point of $g(x) = (x^2 + \cos(x))/10$.*
2. *Prove that there is a unique fixed point x^* of g in $[0, 1]$ and hence a unique solution, also x^*, to $x^2 + \cos(x) - 10x = 0$ in $[0, 1]$.*
3. *Show that this fixed point iteration can only converge linearly to x^*.*