1. Fixed Points

The notion of fixed points is central to numerical analysis. It appears quite often in convergence proofs of iterative algorithms, such as the Jacobi method, the Gauss-Seidel method, and Successive over-Relaxation (SOR), and later, Newton's method.

Definition. The number p is a fixed point for a given function g if $g(p) = p$.

Example. Consider an iterative scheme with matrix operator T, e.g., Jacobi method has $T_J = D^{-1}(L+U)$. The iterative algorithm satisfies

$$\boldsymbol{x}^{(k+1)} = T\boldsymbol{x}^{(k)} + \boldsymbol{c}.$$

Consider a function

$$g\left(\boldsymbol{x}^{(k)}\right) = T\boldsymbol{x}^{(k)} + \boldsymbol{c}.$$

Clearly, if $\boldsymbol{x}^{(k+1)}$ converges to $\boldsymbol{x}$, we must achieve

$$\boldsymbol{x} = T\boldsymbol{x} + c \iff \boldsymbol{x} = g(\boldsymbol{x}) = T\boldsymbol{x} + \boldsymbol{c},$$

that is, the limit is a fixed point of the function g.

In this section, we consider the problem of finding solutions to fixed-point problems and the connection between the fixed-point problems and root-finding problems we wish to solve. Root-finding problems and fixed-point problems are equivalent classes in the following sense:

- Given a root-finding problem $f(p) = 0$, we can define a function g with a fixed point at p **in a number of ways**, for example, as
$$g(x) = x - f(x), \text{or } g(x) = x + 3f(x).$$
- Conversely, if the function g has a fixed point at p, then the function defined by
$$f(x) = x - g(x)$$
has a zero at p.

Remark. **One must note that for a single root-finding problem $f(p) = 0$, there are more than one way to convert to a fixed-point problem, i.e. many many choices of $g(x)$.**

The reason why we put root-finding problems in the form of fixed-point form is because they are then easier to analyze (such as the convergence analysis of iterative methods for linear systems).

Example. Determine any fixed points of the function $g(x) = x^2 - 2$.

Solution. A fixed point p for g satisfies

$$p = g(p) = p^2 - 2$$

which implies that $p^2 - p - 2 = 0 \implies p = 2, -1$.

Remark. The fixed point(s) of a function $g(x)$ occur at the intersection between itself and the line $y = x$.

Is there a special class of functions or conditions such that a fixed point always exists? Let's draw any function that maps between the same set $[a, b]$. Putting the line $y = x$ through it seems to always intersect the function.

Theorem. *(Existence and Uniqueness of Fixed Points)*

(1) *(existence) If $g \in C[a,b]$ and $g(x) \in [a,b]$ for all $x \in [a,b]$, i.e.*
$$g : [a,b] \to [a,b],$$
then g has at least one fixed point in $[a,b]$.

(2) *(uniqueness) If, in addition, $g'(x)$ exists on (a,b) and a positive constant $k < 1$ exists with*
$$|g'(x)| \leq k, \quad \text{for all } x \in (a,b),$$
then there is exactly one fixed point in $[a,b]$.

Proof. We begin with the first statement.

(1) If $g(a) = a$ or $g(b) = b$, then g has a fixed point at an endpoint, and we are done. If not, then $g(a) > a$ and $g(b) < b$ (because if $g(a) < a$ or $g(b) > b$, then $g : [a,b] \to [a_-, b_+]$ which is not our assumption).
Consider function $h(x) = g(x) - x$. It is continuous on $[a,b]$, with
$$h(a) = g(a) - a > 0,$$
$$h(b) = g(b) - b < 0.$$
Intermediate value theorem then informs us that there exists (at least one) $p \in (a,b)$ such that $h(p) = 0$. This p is a fixed point of g because
$$0 = h(p) = g(p) - p \implies g(p) = p.$$

(2) We have just proved that there is at least one fixed point of $g(x)$. Now with further regularity (smoothness) on g such that $|g'(x)| \leq k < 1$, we proceed to prove that the fixed point, if exists, is unique.
We suppose that there are **two fixed points** p and q that both lie in $[a,b]$ for g. If $p \neq q$, say $p > q$, then the Mean Value Theorem implies that there exists $\xi \in (q,p) \subset [a,b]$ such that

$$\frac{g(p) - g(q)}{p-q} = g'(\xi).$$

Thus, using the fact that $g(p) = p$ and $g(q) = q$, we find

$$|p-q| = |g(p) - g(q)| = |g'(\xi)| \, |p-q| \leq k \, |p-q| < |p-q|,$$

which is impossible. This contradiction must come from the only supposition that $p \neq q$. Hence, $p = q$, and the fixed point in $[a,b]$ is unique.

□

Remark. Note that this proof is very similar to the use of Rolle's Theorem when proving that there exists exactly one root for certain function $f(x)$ on some interval $[a,b]$. In fact, these two procedures are equivalent.

Example. Show that $g(x) = \frac{x^2-1}{3}$ has a unique fixed point on the interval $[-1,1]$.
It suffices to check if g is a function from $[-1,1]$ to $[-1,1]$, and that $|g'(x)| \leq k < 1$ for all $x \in [-1,1]$. If so, g satisfies the premise of the Theorem, and we can immediately conclude that g has a unique fixed point on $[-1,1]$.
To check if g satisfies that $g : [-1,1] \to [-1,1]$, we do optimization.
To check if g satisfies $|g'(x)| \leq k < 1$, we find that

$$|g'(x)| = \left|\frac{2x}{3}\right| \leq \frac{2}{3}, \quad \text{for all } x \in (-1,1).$$

2. Fixed Point Iteration Algorithm

Now, though the theorem is useful in the sense that functions with certain properties are guaranteed to have fixed point, it didn't tell us how to find it. In addition, for functions like $g(p) = 3^{-p}$, it is pretty hard to determine the fixed point by hand, that is, solving $p = 3^{-p}$.

Instead, we iterate like the following

$$p_n = g(p_{n-1})$$

by providing an initial point p_0. If the sequence p_n converges to p and g is continuous, then

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} g(p_{n-1}) \overset{\text{continuity}}{=} g\left(\lim_{n\to\infty} p_{n-1}\right) = g(p),$$

and a solution to $x = g(x)$ is obtained. This technique is called **fixed-point**, or **functional iteration**.

3. Convergence

An algorithm is never a good one until one proves its convergence. The rate of convergence is the icing on the top.

Theorem. *Let $g \in C[a,b]$ be such that $g(x) \in [a,b]$, for all $x \in [a,b]$. Suppose in addition that g' exists on (a,b) and there is a constant $0 < k < 1$ such that*

$$|g'(x)| \leq k, \quad \text{for all } x \in (a,b).$$

Then for any number $p_0 \in [a,b]$, the sequence defined by

$$p_n = g(p_{n-1}), \quad n \geq 1$$

converges to the unique fixed p in $[a,b]$.

Proof. We use a very similar technique as in the uniqueness proof, since we are already provided with the conditions that p is unique. By Mean Value Theorem, there exists ξ_n between p_n and p such that

$$|p_n - p| = |g(p_{n-1}) - g(p)| = |g'(\xi_n)| \, |p_{n-1} - p| \leq k \, |p_{n-1} - p|.$$

Now, we repeat the same argument to $|p_{n-1} - p|$ where we further ξ_{n-1} between p_{n-1} and p such that

$$|p_{n-1} - p| = |g(p_{n-2}) - g(p)| = |g'(\xi_{n-1})| \, |p_{n-2} - p| \leq k \, |p_{n-2} - p|$$

which then implies, inductively, that

$$|p_n - p| \leq k \, |p_{n-1} - p| \leq k^2 \, |p_{n-2} - p| \leq \ldots \leq k^n \, |p_0 - p|.$$

Since $0 < k < 1$, we must have $\lim_{n\to\infty} k^n = 0$ and thus

$$\lim_{n\to\infty} |p_n - p| = \lim_{n\to\infty} k^n \, |p_0 - p| = |p_0 - p| \lim_{n\to\infty} k^n = 0,$$

that is, $p_n \to p$ as $n \to \infty$. □

The following corollary provides the error bounds.

Corollary. *If g satisfies the hypotheses of the Theorem, then bounds for the error invovled using p_n to approximate p are given by*

$$|p_n - p| \leq k^n \max\{p_0 - a, b - p_0\}$$

and

$$|p_n - p| \leq \frac{k^n}{1-k} |p_1 - p_0|, \textit{for all } n \geq 1.$$

Proof. The first inequality is trivial because $p, p_0 \in [a, b]$, so from a step in the proof of the theorem, we readily have

$$|p_n - p| \leq k^n |p_0 - p| \leq k^n \max\{p_0 - a, b - p_0\},$$

by being conservative with the distance between p_1 and p.

The second inequality is more involved, which we don't present here. The central idea is to study the sequence difference $|p_{n+1} - p_n|$, and then relate it to $|p_m - p_n|$. Then, we bound $|p_m - p_n|$ in terms of $|p_1 - p_0|$. Lastly, we utilize geometric series to express the final inequality. More details are in Theorem 2.5. □

Remark. The most important takeaway from the convergence theorem is that the fixed-point iteration converges the fastest when k is very small, namely, when the derivative $g'(x)$ is very small in magnitude. Therefore, when setting up a fixed-point iteration $p_n = g(p_{n-1})$ for certain root-finding problems $f(p) = 0$, one would be wise to choose a fixed-point map g such that it has the smallest derivative in the search region. This remark should help you interpret the results of HW6 Q3.