Simple Iteration, cont’d

In general, nonlinear equations cannot be solved in a finite sequence of steps. As linear equations can be solved using direct methods such as Gaussian elimination, nonlinear equations usually require iterative methods. In iterative methods, an approximate solution is refined with each iteration until it is determined to be sufficiently accurate, at which time the iteration terminates. Since it is desirable for iterative methods to converge to the solution as rapidly as possible, it is necessary to be able to measure the speed with which an iterative method converges.

To that end, we assume that an iterative method generates a sequence of iterates $x_0, x_1, x_2, \ldots$ that converges to the exact solution $x^*$. Ideally, we would like the error in a given iterate $x_{k+1}$ to be much smaller than the error in the previous iterate $x_k$. For example, if the error is raised to a power greater than 1 from iteration to iteration, then, because the error is typically less than 1, it will approach zero very rapidly. This leads to the following definition.

**Definition (Order and Rate of Convergence)** Let $\{x_k\}_{k=0}^{\infty}$ be a sequence in $\mathbb{R}^n$ that converges to $x^* \in \mathbb{R}^n$ and assume that $x_k \neq x^*$ for each $k$. We say that the order of convergence of $\{x_k\}$ to $x^*$ is order $r$, with asymptotic error constant $C$, if

$$\lim_{k \to \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^r} = C,$$

where $r \geq 1$. If $r = 1$, then the number $\rho = -\log_{10} C$ is called the asymptotic rate of convergence.

If $r = 1$, and $0 < C < 1$, we say that convergence is linear. If $r = 1$ and $C = 0$, or if $1 < r < 2$ for any positive $C$, then we say that convergence is superlinear. If $r = 2$, then the method converges quadratically, and if $r = 3$, we say it converges cubically, and so on. Note that the value of $C$ need only be bounded above in the case of linear convergence.

When convergence is linear, the asymptotic rate of convergence $\rho$ indicates the number of correct decimal digits obtained in a single iteration. In other words, $\lceil 1/\rho \rceil + 1$ iterations are required to obtain an additional correct decimal digit, where $\lceil x \rceil$ is the “floor” of $x$, which is the largest integer that is less than or equal to $x$.

If $g$ satisfies the conditions of the Contraction Mapping Theorem with Lipschitz constant $L$, then Fixed-point Iteration achieves at least linear convergence, with an asymptotic error constant.
that is bounded above by $L$. This value can be used to estimate the number of iterations needed to obtain an additional correct decimal digit, but it can also be used to estimate the total number of iterations needed for a specified degree of accuracy.

From the Lipschitz condition, we have, for $k \geq 1$,

$$|x_k - x^*| \leq L|x_{k-1} - x^*| \leq L^k|x_0 - x^*|.$$  

From

$$|x_0 - x^*| \leq |x_0 - x_1 + x_1 - x^*| \leq |x_0 - x_1| + |x_1 - x^*| \leq |x_0 - x_1| + L|x_0 - x^*|$$

we obtain

$$|x_k - x^*| \leq \frac{L^k}{1-L}|x_1 - x_0|.$$  

Therefore, in order to satisfy $|x_k - x^*| \leq \epsilon$, the number of iterations, $k$, must satisfy

$$k \geq \frac{\ln |x_1 - x_0| - \ln(\epsilon(1-L))}{\ln(1/L)}.$$  

That is, we can bound the number of iterations after performing a single iteration, as long as the Lipschitz constant $L$ is known.

We know that Fixed-point Iteration will converge to the unique fixed point in $[a,b]$ if $g$ satisfies the conditions of the Contraction Mapping Theorem. However, if $g$ is differentiable on $[a,b]$, its derivative can be used to obtain an alternative criterion for convergence that can be more practical than computing the Lipschitz constant $L$. If we denote the error in $x_k$ by $e_k = x_k - x^*$, we can see from Taylor’s Theorem and the fact that $g(x^*) = x^*$ that

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

where $\xi_k$ lies between $x_k$ and $x^*$. Therefore, if $|g'(x^*)| \leq k$, where $k < 1$, then Fixed-point Iteration is locally convergent; that is, it converges if $x_0$ is chosen sufficiently close to $x^*$. This leads to the following result.

**Theorem (Fixed-point Theorem)** Let $g$ be a continuous function on the interval $[a,b]$, and let $g$ be differentiable on $[a,b]$. If $g(x) \in [a,b]$ for each $x \in [a,b]$, and if there exists a constant $k < 1$ such that

$$|g'(x)| \leq k, \quad x \in (a,b),$$

then the sequence of iterates $\{x_k\}_{k=0}^{\infty}$ converges to the unique fixed point $x^*$ of $g$ in $[a,b]$, for any initial guess $x_0 \in [a,b]$.

It can be seen from the preceding discussion why $g'(x)$ must be bounded away from 1 on $(a,b)$, as opposed to the weaker condition $|g'(x)| < 1$ on $(a,b)$. If $g'(x)$ is allowed to approach 1 as $x$
approaches a point \( c \in (a, b) \), then it is possible that the error \( e_k \) might not approach zero as \( k \) increases, in which case Fixed-point Iteration would not converge.

Suppose that \( g \) satisfies the conditions of the Fixed-Point Theorem, and that \( g \) is also continuously differentiable on \([a, b]\). We can use the Mean Value Theorem to obtain

\[
e_{k+1} = x_{k+1} - x^* = g(x_k) - g(x^*) = g'(\xi_k)(x_k - x^*) = g'(\xi_k)e_k,
\]

where \( \xi_k \) lies between \( x_k \) and \( x^* \). It follows from the continuity of \( g' \) at \( x^* \) that for any initial iterate \( x_0 \in [a, b] \), Fixed-point Iteration converges linearly with asymptotic error constant \( |g'(x^*)| \), since, by the definition of \( \xi_k \) and the continuity of \( g' \),

\[
\lim_{k \to \infty} \frac{|e_{k+1}|}{|e_k|} = \lim_{k \to \infty} |g'(\xi_k)| = |g'(x^*)|.
\]

Recall that the conditions we have stated for linear convergence are nearly identical to the conditions for \( g \) to have a unique fixed point in \([a, b]\). The only difference is that now, we also require \( g' \) to be continuous on \([a, b]\).

The derivative can also be used to indicate why Fixed-point Iteration might not converge.

**Example** The function \( g(x) = x^2 + \frac{3}{16} \) has two fixed points, \( x_1^* = 1/4 \) and \( x_2^* = 3/4 \), as can be determined by solving the quadratic equation \( x^2 + \frac{3}{16} = x \). If we consider the interval \([0, 3/8]\), then \( g \) satisfies the conditions of the Fixed-point Theorem, as \( g'(x) = 2x < 1 \) on this interval, and therefore Fixed-point Iteration will converge to \( x_1^* \) for any \( x_0 \in [0, 3/8] \).

On the other hand, \( g'(3/4) = 2(3/4) = 3/2 > 1 \). Therefore, it is not possible for \( g \) to satisfy the conditions of the Fixed-point Theorem. Furthermore, if \( x_0 \) is chosen so that \( 1/4 < x_0 < 3/4 \), then Fixed-point Iteration will converge to \( x_1^* = 1/4 \), whereas if \( x_0 > 3/4 \), then Fixed-point Iteration diverges. \( \square \)

The fixed point \( x_2^* = 3/4 \) in the preceding example is an unstable fixed point of \( g \), meaning that no choice of \( x_0 \) yields a sequence of iterates that converges to \( x_2^* \). The fixed point \( x_1^* = 1/4 \) is a stable fixed point of \( g \), meaning that any choice of \( x_0 \) that is sufficiently close to \( x_1^* \) yields a sequence of iterates that converges to \( x_1^* \).

The preceding example shows that Fixed-point Iteration applied to an equation of the form \( x = g(x) \) can fail to converge to a fixed point \( x^* \) if \( |g'(x^*)| > 1 \). We wish to determine whether this condition indicates non-convergence in general. If \( |g'(x^*)| > 1 \), and \( g' \) is continuous in a neighborhood of \( x^* \), then there exists an interval \( |x - x^*| \leq \delta \) such that \( |g'(x)| > 1 \) on the interval. If \( x_k \) lies within this interval, it follows from the Mean Value Theorem that

\[
|x_{k+1} - x^*| = |g(x_k) - g(x^*)| = |g'(\eta)||x_k - x^*|,
\]

where \( \eta \) lies between \( x_k \) and \( x^* \). Because \( \eta \) is also in this interval, we have

\[
|x_{k+1} - x^*| > |x_k - x^*|.
\]
In other words, the error in the iterates increases whenever they fall within a sufficiently small interval containing the fixed point. Because of this increase, the iterates must eventually fall outside of the interval. Therefore, it is not possible to find a $k_0$, for any given $\delta$, such that $|x_k - x^*| \leq \delta$ for all $k \geq k_0$. We have thus proven the following result.

**Theorem** Let $g$ have a fixed point at $x^*$, and let $g'$ be continuous in a neighborhood of $x^*$. If $|g'(x^*)| > 1$, then Fixed-point Iteration does not converge to $x^*$ for any initial guess $x_0$ except in a finite number of iterations. □

Now, suppose that in addition to the conditions of the Fixed-point Theorem, we assume that $g'(x^*) = 0$, and that $g$ is twice continuously differentiable on $[a, b]$. Then, using Taylor’s Theorem, we obtain

$$e_{k+1} = g(x_k) - g(x^*) = g'(x^*)(x_k - x^*) + \frac{1}{2}g''(\xi_k)(x_k - x^*)^2 = \frac{1}{2}g''(\xi_k)e_k^2,$$

where $\xi_k$ lies between $x_k$ and $x^*$. It follows that for any initial iterate $x_0 \in [a, b]$, Fixed-point Iteration converges at least quadratically, with asymptotic error constant $|g''(x^*)/2|$. Later, this will be exploited to obtain a quadratically convergent method for solving nonlinear equations of the form $f(x) = 0$.

**Iterative Solution of Equations**

Now that we understand the convergence behavior of Fixed-point Iteration, we consider the application of Fixed-point Iteration to the solution of an equation of the form $f(x) = 0$. When rewriting this equation in the form $x = g(x)$, it is essential to choose the function $g$ wisely. One guideline is to choose $g(x) = x + \phi(x)f(x)$, where the function $\phi(x)$ is, ideally, nonzero except possibly at a solution of $f(x) = 0$. This can be satisfied by choosing $\phi(x)$ to be constant, but this can fail, as the following example illustrates.

**Example** Consider the equation

$$x + \ln x = 0.$$

By the Intermediate Value Theorem, this equation has a solution in the interval $[0.5, 0.6]$. Furthermore, this solution is unique. To see this, let $f(x) = x + \ln x$. Then $f'(x) = x + 1/x > 0$ on the domain of $f$, which means that $f$ is increasing on its entire domain. Therefore, it is not possible for $f(x) = 0$ to have more than one solution.

We consider using Fixed-point Iteration to solve the equivalent equation

$$x = x + (−1)(x + \ln x) = −\ln x.$$

However, with $g(x) = −\ln x$, we have $|g'(x)| = |−1/x| > 1$ on the interval $[0.5, 0.6]$. Therefore, by the preceding theorem, Fixed-point Iteration will fail to converge for any initial guess in this
interval. We therefore apply $g^{-1}(x) = e^{-x}$ to both sides of the equation $x = g(x)$ to obtain

$$g^{-1}(x) = g^{-1}(g(x)) = x,$$

which simplifies to

$$x = e^{-x}.$$ 

The function $g(x) = e^{-x}$ satisfies $|g'(x)| < 1$ on $[0.5, 0.6]$, as $g'(x) = -e^{-x}$, and $e^{-x} < 1$ when the argument $x$ is positive. By narrowing this interval to $[0.52, 0.6]$, which is mapped into itself by this choice of $g$, we can apply the Fixed-point Theorem to conclude that Fixed-point Iteration will converge to the unique fixed point of $g$ for any choice of $x_0$ in the interval. □