FIXED POINT ITERATION

We begin with a computational example. Consider solving the two equations

$$\text{E1:} \quad x = 1 + .5 \sin x$$
$$\text{E2:} \quad x = 3 + 2 \sin x$$

Graphs of these two equations are shown on accompanying graphs, with the solutions being

$$\text{E1:} \quad \alpha = 1.49870113351785$$
$$\text{E2:} \quad \alpha = 3.09438341304928$$

We are going to use a numerical scheme called 'fixed point iteration'. It amounts to making an initial guess of x_0 and substituting this into the right side of the equation. The resulting value is denoted by x_1; and then the process is repeated, this time substituting x_1 into the right side. This is repeated until convergence occurs or until the iteration is terminated.

In the above cases, we show the results of the first 10 iterations in the accompanying table. Clearly convergence is occurring with E1, but not with E2. Why?

The above iterations can be written symbolically as

$$\begin{array}{ll} \text{E1:} & x_{n+1} = 1 + .5 \sin x_n \\ \text{E2:} & x_{n+1} = 3 + 2 \sin x_n \end{array}$$

for $n = 0, 1, 2, ...$ Why does one of these iterations converge, but not the other? The graphs show similar behaviour, so why the difference.

As another example, note that the Newton method

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

is also a fixed point iteration, for the equation

$$x = x - \frac{f(x)}{f'(x)}$$

In general, we are interested in solving equations

$$x = g(x)$$

by means of fixed point iteration:

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

It is called 'fixed point iteration' because the root α is a fixed point of the function $g(x)$, meaning that α is a number for which $g(\alpha) = \alpha$.

EXISTENCE THEOREM

We begin by asking whether the equation $x = g(x)$ has a solution. For this to occur, the graphs of $y = x$ and $y = g(x)$ must intersect, as seen on the earlier graphs. The lemmas and theorems in the book give conditions under which we are guaranteed there is a fixed point α.

Lemma: Let $g(x)$ be a continuous function on the interval $[a, b]$, and suppose it satisfies the property

$$a \le x \le b \quad \Rightarrow \quad a \le g(x) \le b \qquad\qquad (\#)$$

Then the equation $x = g(x)$ has at least one solution α in the interval $[a, b]$. See the graphs for examples.

The proof of this is fairly intuitive. Look at the function

$$f(x) = x - g(x), \qquad a \le x \le b$$

Evaluating at the endpoints,

$$f(a) \le 0, \qquad\qquad f(b) \ge 0$$

The function $f(x)$ is continuous on $[a, b]$, and therefore it contains a zero in the interval.

Theorem: Assume $g(x)$ and $g'(x)$ exist and are continuous on the interval $[a, b]$; and further, assume

$$a \le x \le b \quad \Rightarrow \quad a \le g(x) \le b$$

$$\lambda \equiv \max_{a \le x \le b} \left| g'(x) \right| < 1$$

Then:

S1. The equation $x = g(x)$ has a unique solution α in $[a, b]$.

S2. For any initial guess x_0 in $[a, b]$, the iteration

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

will converge to α.

S3.

$$|\alpha - x_n| \le \frac{\lambda^n}{1 - \lambda} |x_1 - x_0|, \quad n \ge 0$$

S4.

$$\lim_{n \to \infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = g'(\alpha)$$

Thus for x_n close to α,

$$\alpha - x_{n+1} \approx g'(\alpha)\,(\alpha - x_n)$$

The proof is given in the text, and I go over only a portion of it here. For S2, note that from (#), if x_0 is in $[a, b]$, then

$$x_1 = g(x_0)$$

is also in $[a, b]$. Repeat the argument to show that

$$x_2 = g(x_1)$$

belongs to $[a, b]$. This can be continued by induction to show that every x_n belongs to $[a, b]$.

We need the following general result. For any two points w and z in $[a, b]$,

$$g(w) - g(z) = g'(c)\,(w - z)$$

for some unknown point c between w and z. Therefore,

$$|g(w) - g(z)| \le \lambda\,|w - z|$$

for any $a \le w, z \le b$.

For S3, subtract $x_{n+1} = g(x_n)$ from $\alpha = g(\alpha)$ to get

$$\begin{aligned} \alpha - x_{n+1} &= g(\alpha) - g(x_n) \\ &= g'(c_n)\,(\alpha - x_n) \qquad\qquad (*) \\ |\alpha - x_{n+1}| &\le \lambda\,|\alpha - x_n| \end{aligned}$$

with c_n between α and x_n. From the last statement, we have that the error is guaranteed to decrease by a factor of λ with each iteration. This leads to

$$|\alpha - x_n| \le \lambda^n\,|\alpha - x_n|\,, \qquad\qquad n \ge 0$$

With some extra manipulation, we can obtain the error bound in S3.

For S4, use (*) to write

$$\frac{\alpha - x_{n+1}}{\alpha - x_n} = g'(c_n)$$

Since $x_n \to \alpha$ and c_n is between α and x_n, we have $g'(c_n) \to g'(\alpha)$.

The statement

$$\alpha - x_{n+1} \approx g'(\alpha)\,(\alpha - x_n)$$

tells us that when near to the root α, the errors will decrease by a constant factor of $g'(\alpha)$. If this is negative, then the errors will oscillate between positive and negative, and the iterates will be approaching from both sides. When $g'(\alpha)$ is positive, the iterates will approach α from only one side.

The statements

$$\alpha - x_{n+1} = g'(c_n)\,(\alpha - x_n)$$

$$\alpha - x_{n+1} \approx g'(\alpha)\,(\alpha - x_n)$$

also tell us a bit more of what happens when

$$\left| g'(\alpha) \right| > 1$$

Then the errors will increase as we approach the root rather than decrease in size.

Look at the earlier examples

$$\begin{array}{ll}\text{E1:} & x = 1 + .5\sin x \\ \text{E2:} & x = 3 + 2\sin x\end{array}$$

In the first case E1,

$$\begin{aligned} g(x) &= 1 + .5\sin x \\ g'(x) &= .5\cos x \\ |g'(\alpha| &\le \tfrac{1}{2} \end{aligned}$$

Therefore the fixed point iteration

$$x_{n+1} = 1 + .5\sin x_n$$

will converge for E1.

For the second case E2,

$$\begin{aligned} g(x) &= 3 + 2\sin x \\ g'(x) &= 2\cos x \\ g'(\alpha) &= 2\cos(3.09438341304928) \doteq -1.998 \end{aligned}$$

Therefore the fixed point iteration

$$x_{n+1} = 3 + 2\sin x_n$$

will diverge for E2.

Corollary: Assume $x = g(x)$ has a solution α, and further assume that both $g(x)$ and $g'(x)$ are continuous for all x in some interval about α. In addition, assume

$$\left|g'(\alpha)\right| < 1 \qquad (**)$$

Then any sufficiently small number $\varepsilon > 0$, the interval $[a, b] = [\alpha - \varepsilon, \alpha + \varepsilon]$ will satisfy the hypotheses of the preceding theorem.

This means that if (**) is true, and if we choose x_0 sufficiently close to α, then the fixed point iteration $x_{n+1} = g(x_n)$ will converge and the earlier results S1-S4 will all hold. The corollary does not tell us how close we need to be to α in order to have convergence.

NEWTON'S METHOD

For Newton's method

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

we have it is a fixed point iteration with

$$g(x) = x - \frac{f(x)}{f'(x)}$$

Check its convergence by checking the condition (**).

$$\begin{aligned} g'(x) &= 1 - \frac{f'(x)}{f'(x)} + \frac{f(x)f''(x)}{[f'(x)]^2} \\ &= \frac{f(x)f''(x)}{[f'(x)]^2} \\ g'(\alpha) &= 0 \end{aligned}$$

Therefore the Newton method will converge if x_0 is chosen sufficiently close to α.

HIGHER ORDER METHODS

What happens when $g'(\alpha) = 0$? We use Taylor's theorem to answer this question.
Begin by writing

$$g(x) = g(\alpha) + g'(\alpha)\,(x - \alpha) + \frac{1}{2}g''(c)\,(x - \alpha)^2$$

with c between x and α. Substitute $x = x_n$ and recall that $g(x_n) = x_{n+1}$ and $g(\alpha) = \alpha$. Also assume $g'(\alpha) = 0$.
Then

$$\begin{aligned} x_{n+1} &= \alpha + \tfrac{1}{2}g''(c_n)\,(x_n - \alpha)^2 \\ \alpha - x_{n+1} &= -\tfrac{1}{2}g''(c_n)\,(x_n - \alpha)^2 \end{aligned}$$

with c_n between α and x_n. Thus if $g'(\alpha) = 0$, the fixed point iteration is quadratically convergent or better. In fact, if $g''(\alpha) \neq 0$, then the iteration is exactly quadratically convergent.

ANOTHER RAPID ITERATION

Newton's method is rapid, but requires use of the derivative $f'(x)$. Can we get by without this. The answer is yes! Consider the method

$$\begin{aligned} D_n &= \frac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)} \\ x_{n+1} &= x_n - \frac{f(x_n)}{D_n} \end{aligned}$$

This is an approximation to Newton's method, with $f'(x_n) \approx D_n$. To analyze its convergence, regard it as a fixed point iteration with

$$\begin{aligned} D(x) &= \frac{f(x + f(x)) - f(x)}{f(x)} \\ g(x) &= x - \frac{f(x)}{D(x)} \end{aligned}$$

Then we can, with some difficulty, show $g'(\alpha) = 0$ and $g''(\alpha) \neq 0$. This will prove this new iteration is quadratically convergent.

FIXED POINT INTERATION: ERROR

Recall the result

$$\lim_{n\to\infty} \frac{\alpha - x_n}{\alpha - x_{n-1}} = g'(\alpha)$$

for the iteration

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

Thus

$$\alpha - x_n \approx \lambda\,(\alpha - x_{n-1}) \qquad (*)$$

with $\lambda = g'(\alpha)$ and $|\lambda| < 1$.

If we were to know λ, then we could solve (*) for α:

$$\alpha \approx \frac{x_n - \lambda x_{n-1}}{1 - \lambda}$$

Usually, we write this as a modification of the currently computed iterate x_n:

$$\begin{aligned}\alpha &\approx \frac{x_n - \lambda x_{n-1}}{1-\lambda} \\ &= \frac{x_n - \lambda x_n}{1-\lambda} + \frac{\lambda x_n - \lambda x_{n-1}}{1-\lambda} \\ &= x_n + \frac{\lambda}{1-\lambda}\left[x_n - x_{n-1}\right]\end{aligned}$$

The formula

$$x_n + \frac{\lambda}{1-\lambda}\left[x_n - x_{n-1}\right]$$

is said to be an extrapolation of the numbers x_{n-1} and x_n. But what is λ?

From

$$\lim_{n\to\infty} \frac{\alpha - x_n}{\alpha - x_{n-1}} = g'(\alpha)$$

we have

$$\lambda \approx \frac{\alpha - x_n}{\alpha - x_{n-1}}$$

Unfortunately this also involves the unknown root α which we seek; and we must find some other way of estimating λ.

To calculate λ consider the ratio

$$\lambda_n = \frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}}$$

To see this is approximately λ as x_n approaches α, write

$$\frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}} = \frac{g(x_{n-1}) - g(x_{n-2})}{x_{n-1} - x_{n-2}} = g'(c_n)$$

with c_n between x_{n-1} and x_{n-2}. As the iterates approach α, the number c_n must also approach α. Thus λ_n approaches λ as $x_n \to \alpha$.

We combine these results to obtain the estimation

$$\widehat{x}_n = x_n + \frac{\lambda_n}{1-\lambda_n}\left[x_n - x_{n-1}\right], \qquad \lambda_n = \frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}}$$

We call $\widehat{x}_n$ the <u>Aitken extrapolate</u> of $\{x_{n-2}, x_{n-1}, x_n\}$; and $\alpha \approx \widehat{x}_n$.

We can also rewrite this as

$$\alpha - x_n \approx \widehat{x}_n - x_n = \frac{\lambda_n}{1-\lambda_n}\left[x_n - x_{n-1}\right]$$

This is called <u>Aitken's error estimation formula</u>.

The accuracy of these procedures is tied directly to the accuracy of the formulas

$$\alpha - x_n \approx \lambda\left(\alpha - x_{n-1}\right), \qquad \alpha - x_{n-1} \approx \lambda\left(\alpha - x_{n-2}\right)$$

If this is accurate, then so are the above extrapolation and error estimation formulas.

EXAMPLE

Consider the iteration

$$x_{n+1} = 6.28 + \sin(x_n), \qquad n = 0, 1, 2, ...$$

for solving

$$x = 6.28 + \sin x$$

Iterates are shown on the accompanying sheet, including calculations of λ_n, the error estimate

$$\alpha - x_n \approx \widehat{x}_n - x_n = \frac{\lambda_n}{1 - \lambda_n} [x_n - x_{n-1}]$$

The latter is called "Estimate (2.6.5)" in the table. In this instance,

$$g'(\alpha) \doteq .9644$$

and therefore the convergence is very slow. This is apparent in the table.

AITKEN'S ALGORITHM

Step 1: Select x_0
Step 2: Calculate

$$x_1 = g(x_0), \quad x_2 = g(x_1)$$

Step3: Calculate

$$x_3 = x_2 + \frac{\lambda_2}{1 - \lambda_2}\left[x_2 - x_1\right], \qquad \lambda_2 = \frac{x_2 - x_1}{x_1 - x_0}$$

Step 4: Calculate

$$x_4 = g(x_3), \quad x_5 = g(x_4)$$

and calculate x_6 as the extrapolate of $\{x_3, x_4, x_5\}$. Continue this procedure, ad infinatum.

Of course in practice we will have some kind of error test to stop this procedure when believe we have sufficient accuracy.

EXAMPLE

Consider again the iteration

$$x_{n+1} = 6.28 + \sin(x_n), \qquad n = 0, 1, 2, ...$$

for solving

$$x = 6.28 + \sin x$$

Now we use the Aitken method, and the results are shown in the accompanying table. With this we have

$$\alpha - x_3 = 7.98 \times 10^{-4}, \quad \alpha - x_6 = 2.27 \times 10^{-6}$$

In comparison, the original iteration had

$$\alpha - x_6 = 1.23 \times 10^{-2}$$

GENERAL COMMENTS

Aitken extrapolation can greatly accelerate the convergence of a linearly convergent iteration

$$x_{n+1} = g(x_n)$$

This shows the power of understanding the behaviour of the error in a numerical process. From that understanding, we can often improve the accuracy, thru extrapolation or some other procedure.

This is a justification for using mathematical analyses to understand numerical methods. We will see this repeated at later points in the course, and it holds with many different types of problems and numerical methods for their solution.