Section 2.3 Fixed Point Iteration Schemes

Enclosure Methods
- Guaranteed to converge to a root under mild conditions.
- Rate of convergence is slow; often requires many iterations to achieve a specified level accuracy.

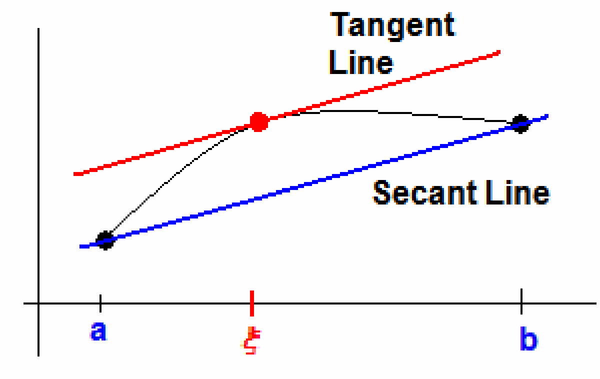
Fixed Point Schemes
- When constructed properly rapid convergence is exhibited.
- NOT guaranteed to converge.

In the development of fixed point schemes the Mean Value Theorem plays an important role.

MEAN VALUE THEOREM If the function f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then there exists a real number $\xi \in (a, b)$ such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Geometric Interpretation. Construct the secant line between point (a, f(a)) and (b, f(b)). Then there exists a point ξ point in (a, b) so that the derivative of f at x = ξ is the same as the slope of the secant line.



From the point of view of numerical methods and analysis, the Mean Value Theorem (MVT) is probably second in importance only to Taylor's Theorem. Why? Consider a slightly reworked form

$$\mathbf{f(b) - f(a) = f'(\xi)(b - a)}$$

Thus, the MVT allows us to replace differences of function values with differences of argument values, if we scale by the derivative of the function.

Example: We can use the MVT to determine that

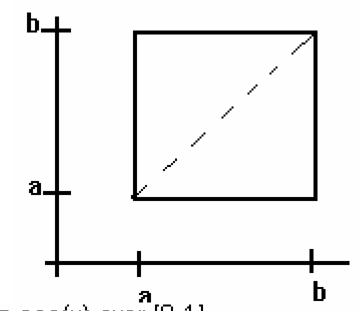
$$|\cos(x_1) - \cos(x_2)| \le |x_1 - x_2| \quad \textbf{EXPLAIN!}$$

++

Definition:
A fixed point for a given function g(x) is a number p for which g(p) = p.

Note, this implies that point (p, p) is on the graph of y = g(x). So intersections of y = g(x) and y = x are fixed points of function g.

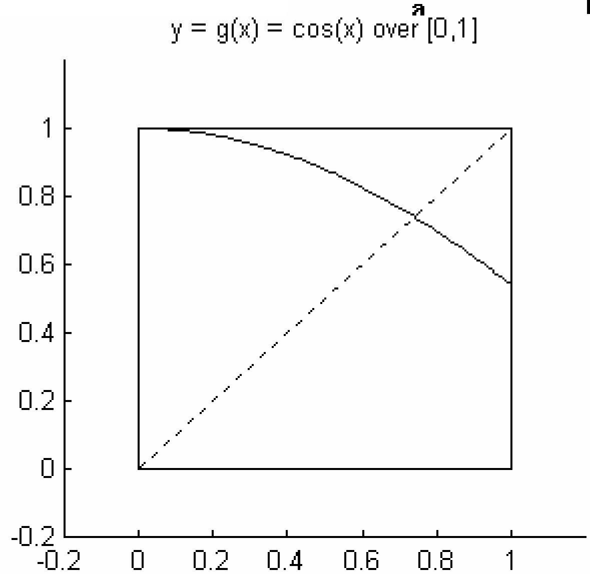
Terminology: We say a function g maps (domain) interval [a,b] into itself if for every x in [a,b], g(x) is in the (range) interval [a,b]. We use the notation **g: [a, b] → [a, b]**.



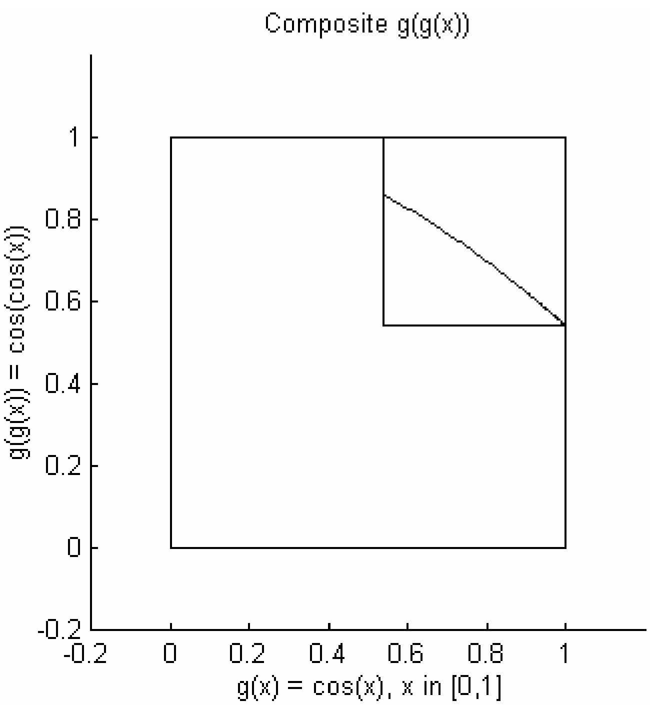
That is, $a \le g(x) \le b$. This can also be expressed by the notation ***g(x)*** $\in$ ***[a,b]*** *for all x* $\in$ *[a,b].* Geometrically this means if you graph y = g(x) over [a, b] that the resulting curve will be in the box in the xy-plane determined by $a \le x \le b$, $a \le y \le b$. (See the figure.) Note that the line y = x goes through the corners (a, a) and (b, b).

Example:

Consider y = g(x) = cos(x) over interval [0, 1].



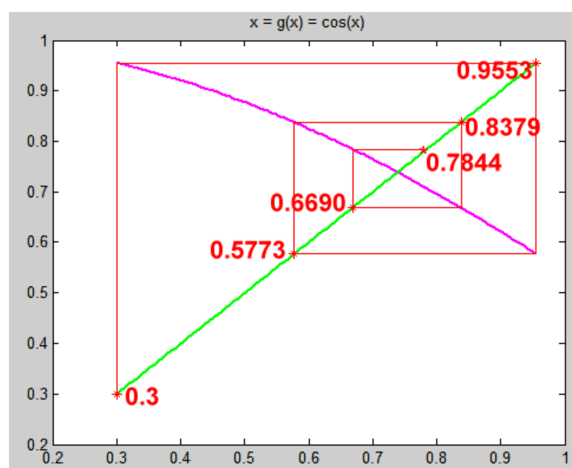
If we next think of computing the composition of g(x) = cos(x) with itself, that is graph g(x) vs. g(g(x)) we get the following picture.



Note that now the box containing the graph of g(g(x)) is smaller. For the function g(x) = cos(x) successive iterations with itself trap the composite function in smaller and smaller boxes. By construction the line y = x goes through corners of the boxes. Hence continuing the composite function should lead to an isolation of the intersection of line y= x and curve y = g(x), which is a fixed point of g(x).

At least intuitively we could approximate the fixed point of g(x) = cos(x) over [0, 1] by starting with initial guess p_0 in [0, 1] and generating a sequence from compositions of g with itself:

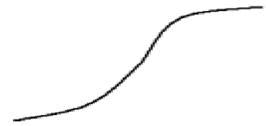
$$p_0, g(p_0), g(g(p_0)), g(g(g(p_0))), \ldots.$$



We can display this geometrically using a web diagram.

++

The preceding example had one fixed point.



Example: Fixed Points of the Logistic Equation

One of the most popular mathematical models for the generation-by-generation growth of a population is the logistic equation:

$$p_{n+1} = cp_n(1 - p_n),$$

where $0 < c < 4$ is a constant and p_n denotes the normalized size of the population in the n-th generation, measured relative to the maximum population which the environment can support. The fixed points of the function on the right-hand side of the logistic equation play an important role in the dynamics of the long-term behavior of the population. Using the definition, the fixed points for the logistic equation are the solutions of

$$p = cp(1 - p).$$

Solving this quadratic equation produces $p = 0$ and $p = (c - 1)/c$.

++

Formal Theory:

Existence:

If g is continuous on [a, b] and maps [a, b] into itself, then g has a fixed point in [a, b].

Uniqueness:

If g is continuous on [a, b], maps [a, b] into itself and g'(x) exists on (a, b) and there is a positive constant k < 1 so that

$$|g'(x)| \le k < 1 \text{ for } x \text{ in } (a, b),$$

then there exists a unique fixed point of g in [a, b].

Proof of Existence:

By the hypotheses, $g(a) \ge a$ and $g(b) \le b$. If equality holds in either case then we have a fixed point of g. Hence suppose that $g(a) > a$ and $g(b) < b$. Define $h(x) = g(x) – x$. (Claim: h is continuous on [a, b]. Why?) It follows that $h(a) > 0$ and $h(b) < 0$, so there is a p in [a, b] so that $h(p) = 0$. (Why?) Hence $h(p) = g(p) – p = 0$ ➔ $g(p) = p$ so we have the existence of a fixed point in [a, b].

Proof of Uniqueness:

Here we will use the additional hypothesis that there is a positive constant k < 1 so that

$$|g'(x)| \le k < 1 \text{ for } x \text{ in } (a,b).$$

Assume there are two fixed points, p and q, of g in [a, b]. Then

$$|p-q| = |g(p) - g(q)| = |g'(\alpha)| * |p-q| \quad \textbf{(Why?)}$$
$$\le k * |p-q| < |p-q|$$

But a value can't be less than itself, hence there is a unique fixed point.

It should be noted that the hypotheses of this theorem are sufficient conditions. By themselves, these conditions guarantee the existence and uniqueness of a fixed point. However, the hypotheses are not necessary conditions, meaning that it is possible for a function to violate one or more of the hypotheses, yet still have a (possibly unique) fixed point. For example, consider the function $g(x) = 4x(1-x)$ on the interval $[0.1, \infty)$. Since $\lim_{x\to\infty} g(x) \to -\infty$, g clearly does not map $[0.1, \infty)$ onto itself. Furthermore, $\lim_{x\to\infty} |g'(x)| \to +\infty$, so that g also violates the hypothesis regarding the magnitude of the first derivative. However, g has fixed points at $x = 0$ and $x = 3/4$, so that g does in fact have a unique fixed point on the interval $[0.1, \infty)$.

Definition: FIXED POINT ITERATION SCHEME (also known as a FUNCTIONAL iteration scheme) to approximate the fixed point, p, of a function g, generates the sequence $\{p_n\}$ by the rule $p_n = g(p_{n-1})$ for all $n \ge 1$, given a starting approximation, p_0.

Within a fixed point iteration scheme, the function g is often referred to as the *iteration function*.

Examples:

$x = \cos(x)$	$x = e^{-x}$	
0.300000000000000	0.600000000000000	🡐 p_0 = initial guesses
0.955336489125606	0.548811636094027	🡐 $p_1 = g(p_0)$
0.577334044471186	0.577635844258916	🡐 $p_2 = g(p_1)$
0.837920683127127	0.561223619437973	🡐 etc.
0.669009730822383	0.570510548780605	
0.784436224742356	0.565236784068813	
0.707786647275637	0.568225584061220	
0.759802755285230	0.566529806871672	
0.724971882753918	0.567491330229635	
0.748518067212715	0.566945936306695	

Connection between Fixed Points and Root Problems

Given a root problem f(x) = 0. **To get a fixed point formulation we use algebra to rearrange the root problem into the form x = g(x).**

Once we have the equivalent form x = g(x) we use the fixed point method which makes initial guess p_0 then uses composite functions as in

$p_1 = g(p_0)$, $p_2 = g(p_1) = g(g(p_0))$, $p_3 = g(p_2) = g(g(g(p_0)))$,
$p_4 = g(p_3)$, ... etc. (This is called the fixed point sequence.)

Example: $\mathbf{f(x) = x^2 - x - 2 = 0}$

Fixed point formulations:

i) $x(x-1) = 2 \Leftrightarrow x - 1 = \frac{2}{x} \Leftrightarrow x = 1 + \frac{2}{x}$; define $g(x) = 1 + \frac{2}{x}$

ii) $x^2 = x + 2 \Leftrightarrow x = \sqrt{x+2}$; define $g(x) = \sqrt{x+2}$

iii) $x = x^2 - 2$; define $g(x) = x^2 - 2$

Warning: Not every formulation is guaranteed to produce a convergent fixed point sequence for an arbitrary initial guess p_0.

Fixed Point Theorem:

If g is continuous on [a, b], maps [a, b] into itself and g'(x) exists on (a, b) and there is a positive constant k < 1 so that $|g'(x)| \le k < 1$ for x in (a, b), then there exists a unique fixed point of g in [a, b] which can be computed by choosing an initial guess p_0 in [a, b] and generating the sequence

$p_1 = g(p_0)$, $p_2 = g(p_1) = g(g(p_0))$,
$p_3 = g(p_2) = g(g(g(p_0)))$, $p_4 = g(p_3)$, ...

which will converge to the unique fixed point in [a, b].

We have already shown Existence and Uniqueness of a fixed point under the stated hypotheses. Thus we need only show that the sequence

$p_1 = g(p_0)$, $p_2 = g(p_1) = g(g(p_0))$,
$p_3 = g(p_2) = g(g(g(p_0)))$, $p_4 = g(p_3)$, ...

which will converge to the unique fixed point in [a, b]. **We show that** $\lim_{n\to\infty} p_n = p$.

$$\begin{aligned}
|p_n - p| &= |g(p_{n-1}) - g(p)| = |g'(\xi_{n-1})||p_{n-1} - p| \\
&\le k|p_{n-1} - p| = k|g(p_{n-2}) - g(p)| = k|g'(\xi_{n-2})||p_{n-2} - p| \\
&\le k^2 |p_{n-2} - p| \\
&\vdots \\
&\le k^n |p_0 - p|
\end{aligned}$$

Take the limit of both sides of the string of inequalities and since $0 < k < 1$, we have
$\lim_{n\to\infty} |p_n - p| = 0 \Leftrightarrow \lim_{n\to\infty} p_n = p$.

Other results:

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

This follows since p_0 is in [a, b]. This gives us a measure of the error in the nth approximation to the fixed point that doesn't require knowing p.

Also, to determine how many iterations to perform to get a desired accuracy tol we can determine n from the inequality $k^n \max\{p_0 - a,\ b - p_0\} < tol$. (Hint: Use logs.)

2. If we go back to the proof of the Fixed Point Thm we can reformulate expressions as follows:

$$|p_0 - p| = |p_0 - g(p_0) + p_1 - p| \le |g(p) - g(p_0)| + |p_1 - p_0| \le k|p_0 - p| + |p_1 - p_0|$$

then we get $|p_0 - p| \le \frac{1}{1-k}|p_1 - p_0|$. Teaming this with $|p_n - p| \le k^n |p_0 - p|$ we get that

$$|p_n - p| \le \frac{k^n}{1-k}|p_1 - p_0|$$

This gives us a measure of the error in the nth approximation to the fixed point. Also, to determine how many iterations to perform to get a desired accuracy tol we can determine n from the inequality
$|p_n - p| \le \frac{k^n}{1-k}|p_1 - p_0| < tol$. (Hint: Use logs.)

3. **Order of Convergence**

From the Fixed point Thm we have

$$|p_n - p| = |g'(\xi_{n-1})||p_{n-1} - p|$$

thus $\lim_{n\to\infty} \frac{|e_n|}{|e_{n-1}|} \lim_{n\to\infty} \frac{|p_n - p|}{|p_{n-1} - p|} = \lim_{n\to\infty} |g'(\xi_{n-1})| = |g'(p)|$

If $|g'(p)| \ne 0$, then the fixed point method is linearly convergent. (Note: you really need to know the exact value of the fixed point p.)

To obtain higher order convergence, it is clear that the iteration function must have a zero derivative at the fixed point. The next theorem indicates that the more derivatives of the iteration function which are zero at the fixed point, the higher will be the order of convergence of the generated sequence.

Theorem:

Let g be a continuous function on the closed interval $[a, b]$ with $\alpha > 1$ continuous derivatives on the open interval (a, b). Further, let $p \in (a, b)$ be a fixed point of g. If

$$g'(p) = g''(p) = \cdots = g^{(\alpha-1)}(p) = 0,$$

but $g^{(\alpha)}(p) \neq 0$, then there exists a $\delta > 0$ such that for any $p_0 \in [p - \delta, p + \delta]$, the sequence $p_n = g(p_{n-1})$ converges to the fixed point p of order α with asymptotic error constant

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^\alpha} = \frac{|g^{(\alpha)}(p)|}{\alpha!}.$$

4. **Stopping Condition**

For a fixed point iteration scheme which produces a linearly convergent sequence (*i.e.*, $g'(p) \neq 0$), a stopping condition can be formulated in much the same manner as one was formulated for the method of false position. In particular, an estimate for $|e_n|$ can be constructed from terms in the sequence $\{p_n\}$. In the case of fixed point iteration, the relevant formulas are

$$|e_n| \approx \left| \frac{g'(p)}{g'(p) - 1} \right| |p_n - p_{n-1}| \qquad (1)$$

and

$$g'(p) \approx \frac{p_n - p_{n-1}}{p_{n-1} - p_{n-2}}. \qquad (2)$$

The details are left as an exercise. Thus an appropriate stopping condition would involve estimating $g'(p)$ and $|e_n|$ using the formulas given above and terminating the iteration when $|e_n|$ falls below the convergence tolerance ϵ.

Example: Error Estimate and Stopping Condition

We know that the function $g(x) = e^{-x}$ has a unique fixed point somewhere near $x = 0.6$. To ten decimal places, the fixed point happens to be $p = 0.5671432904$. The absolute error in the first ten approximations obtained from fixed point iteration with $p_0 = 0$ is listed below, along with the error estimate as obtained from equations (1) and (2).

		error estimate
$p_1 = 1.0000000000$	$\lvert e_1\rvert = 0.4328567096$	
$p_2 = 0.3678794412$	$\lvert e_2\rvert = 0.1992638492$	
$p_3 = 0.6922006276$	$\lvert e_3\rvert = 0.1250573371$	0.1099745306
$p_4 = 0.5004735006$	$\lvert e_4\rvert = 0.0666697898$	0.0712322670
$p_5 = 0.6062435351$	$\lvert e_5\rvert = 0.0391002447$	0.0376047292
$p_6 = 0.5453957860$	$\lvert e_6\rvert = 0.0217475044$	0.0222212089
$p_7 = 0.5796123355$	$\lvert e_7\rvert = 0.0124690451$	0.0123155830
$p_8 = 0.5601154614$	$\lvert e_8\rvert = 0.0070278290$	0.0070769665
$p_9 = 0.5711431151$	$\lvert e_9\rvert = 0.0039998247$	0.0039839812
$p_{10} = 0.5648793474$	$\lvert e_{10}\rvert = 0.0022639430$	0.0022690318

The first few error estimates differ from the actual errors by as much as about 12%. After p_7, however, the error estimates differ from the actual errors by less than one percent.

If iterations are continued until the error estimate falls below $\epsilon = 5 \times 10^{-6}$, the final approximation to p is $p_{21} = 0.5671477143$. The estimate of the error in this approximation is 4.42383×10^{-6}, which is in excellent agreement with the actual error of 4.42385×10^{-6}.

++

When the sequence produced by fixed point iteration has order of convergence $\alpha > 1$, a simpler stopping condition can be used. Recall that any order of convergence $\alpha > 1$ corresponds to superlinear convergence and that all superlinearly convergent sequences satisfy the limit

$$\lim_{n\to\infty} \frac{|p_n - p_{n-1}|}{|p_{n-1} - p|} = 1$$

(see Exercises 14 and 15 in Section 1-2). This limit implies that

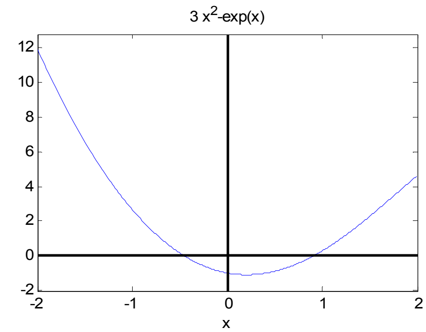
$$|p_{n-1} - p| \approx |p_n - p_{n-1}|$$

for any superlinearly convergent sequence. Since p_n is supposed to be a better approximation to p than p_{n-1}, it follows that $|p_n - p_{n-1}|$ should be a conservative estimate of the error $|e_n| = |p_n - p|$. Consequently, whenever the order of convergence is greater than one, an appropriate stopping condition would be to terminate the iteration as soon as $|p_n - p_{n-1}|$ falls below the convergence tolerance ϵ.

Example:

Let $f(x) = 3x^2 - e^x$. Create a fixed point formulation to estimate a root of f(x) = 0. Determine an interval over which the fixed point method will converge, select a starting value p_0, generate a sequence of approximations and the corresponding error estimates. Predict the number of iterations required so that you can obtain an approximation accurate within 10^{-5}.

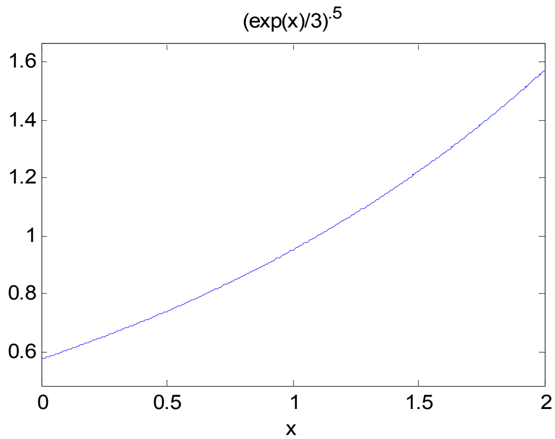
We see there are two roots.



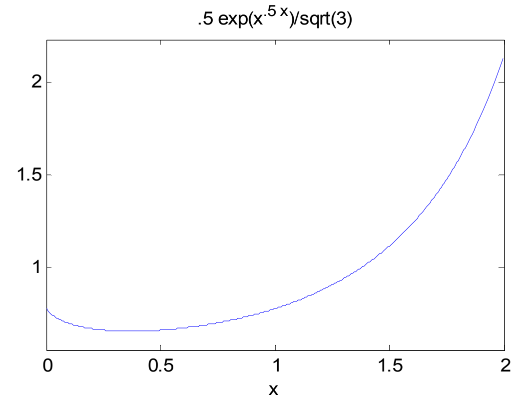
One fixed point formulation is $\mathbf{x} = \mathbf{g(x)} = \left(\frac{\mathbf{e^x}}{\mathbf{3}}\right)^{\mathbf{1/2}}$.

Apply the fixed point theorem.

$\mathbf{g(x)} = \frac{\mathbf{e^{0.5x}}}{\sqrt{\mathbf{3}}}$, so $\mathbf{g'(x)} = \frac{\mathbf{0.5e^{0.5x}}}{\sqrt{\mathbf{3}}}$



It appears that g maps [0, 2] into [0, 2]



Let's look at the derivative.
Note that | g'(x) | is not less than 1 over [0, 2].

So consider [0, 1]. Does g map [0, 1] into [0, 1]?
Is | g'(x) | < k < 1 over [0, 1]?

$$\left|\mathbf{g'(x)}\right| = \left|\frac{\mathbf{0.5e^{0.5x}}}{\sqrt{\mathbf{3}}}\right| < \left|\frac{\mathbf{0.5e^{0.5*1}}}{\sqrt{\mathbf{3}}}\right| \approx \mathbf{0.4759448347} < \mathbf{0.476} \text{ since } \mathbf{e^{0.5x}} \text{ is increasing on [0, 1]}$$

So we can take k = 0.476.

Let's generate fixed point approximations using **g** with $p_0 = 0.5$.

5.000000000000000e-001	← p_0
7.413324199709890e-001	← p_1
8.364070066183053e-001	← p_2
8.771277404802832e-001	← p_3
8.951694275536428e-001	
9.032811431491092e-001	
9.069521625510804e-001	
9.086184107833576e-001	
9.093757181154459e-001	
9.097201217656593e-001	
9.098767907199163e-001	
9.099480682342758e-001	
9.099804982304047e-001	
9.099952536820517e-001	
9.100019674023000e-001	

Estimates of g'(p)	Error Estimates	
0.39396	0.061803	← e_2
0.4283	0.030507	← e_3
0.44306	0.014353	
0.44961	0.0066264	
0.45256	0.0030347	
0.45389	0.0013849	
0.4545	0.00063097	
0.45477	0.00028727	
0.4549	0.00013074	
0.45496	5.9496e-005	
0.45498	2.7073e-005	

Use the following approximations to estimate the error at each step.

$$|e_n| \approx \left|\frac{g'(p)}{g'(p)-1}\right| |p_n - p_{n-1}|$$

$$g'(p) \approx \frac{p_n - p_{n-1}}{p_{n-1} - p_{n-2}}$$

$$\mathbf{g'(p) \approx \frac{p_2 - p_1}{p_1 - p_0} \approx 0.39396}$$

$$\mathbf{|e_2| \approx \left|\frac{0.39396}{0.39396 - 1}\right| |0.8364 - 0.7413| \approx 0.061803}$$

To estimate the number of iteration to obtain an approximation accurate within 10^{-5}, we can use

$k^n \max\{p_0 - a,\ b - p_0\} < tol$ or $|p_n - p| \le \frac{k^n}{1-k}|p_1 - p_0| < tol$ using k = 0.476.

$p_0 = 0.5$, $a = 0$, $b = 1$

$$\mathbf{(0.476)^n \frac{1}{2} < 10^{-5}}$$

$$\mathbf{n\ln(0.476) < \ln(2 * 10^{-5})}$$

$$\mathbf{n > 15}$$

$p_0 = 0.5$, $p_1 = 0.7413$

$$\mathbf{\frac{(0.476)^n}{1-0.476}|0.7413 - 0.5| < 10^{-5}}$$

$$\mathbf{(0.476)^n < \frac{10^{-5}}{0.4605} = 2.172 \times 10^{-5}}$$

$$\mathbf{n\ln(0.476) < \ln(21.72 \times 10^{-5})}$$

$$\mathbf{n > 15}$$

The function $\mathbf{f(x) = 3x^2 - e^x}$ has a second root. How would you approximate it?