



1-Intro

NONLINEAR EQUATIONS

Ari Maller





NONLINEAR EQUATIONS

- relaxation method
- bisection method
- Newton's method
- max and mins of functions

NONLINEAR EQUATIONS

- We deal with linear equations so often we can forget that most equations can not be solved algebraically.
- polynomials: $x^5 + 8x^3 + 6x^2 + 3x + 4 = 0$
- exponentials and logs: $2x - \log(x) = 0$, $x - e^{2x} = 5$
- trig functions: $\sin(x) - x^2 = x$, $\tan(x) = \log(x)$

NONLINEAR EQUATIONS

- So far we have focussed on systems of linear equations. In these cases as long as we have as many equations as variables we know they can always be solved.
- Nonlinear equations are much more difficult. Often there is no way to solve them analytically. Numerical solutions are often the only option.
- There are a number of different approaches to solving nonlinear equations, the first one we will look at is the relaxation method.

RELAXATION METHOD

- Let's consider a single nonlinear equation with a single variable that we want to solve. For example,

$$x = 2 - e^{-x}$$

- there is no known analytic method for solving this equation. The relaxation method is to simply iterate over guesses for x until the value 'relaxes' to a solution.
- So we can start with $x=1$, plugging that into the right side of the equation gives $x'=1.632$. Then we plug 1.632 into the right side of the equation which gives $x''=1.804$.
- If we keep on doing this we find that we arrive at a solution $x=1.84140564533$. Plugging this in to the right side and getting the same thing out means this is a solution to the equation.

RELAXATION METHOD

- The relaxation method is a simple and fast method for solving nonlinear equations. There can be some problems with it.
 1. You have to be able to write the equation as $x=f(x)$. If your equation isn't this way you can try to rearrange it, but if not you can't use this method.
 2. The equation can have more than one solution. You can try to find other solutions by changing your initial starting guess. This will work best if you have some idea of where you expect the solution to be. There are some solutions to some equations that you simply won't find.
 3. Sometimes the method will give oscillations instead of relaxing to one solution.

RELAXATION METHOD

- Consider for example the equation

$$x = e^{1-x^2}$$

- this also has no known analytic solution, but you can easily see that $x=1$ is a solution.
- If you try relaxation on this equation and start with $x=1/2$ you'll see that x oscillates between 2.7 and 0.0016 and doesn't move towards one solution.
- You can rearrange the equation by taking the log of both sides and getting
$$x = \sqrt{1 - \log x}$$
- now the relaxation method will work and give you $x=1$.

RELAXATION METHOD

- We can understand why sometimes the relaxation method works and other times it doesn't by performing a Taylor expansion around a point x near a solution x^* . This gives

$$x' = f(x) = f(x^*) + (x - x^*)f'(x^*) + \dots$$

- since $f(x^*) = x^*$ if we neglect the higher order terms we get

$$x' - x^* = (x - x^*)f'(x^*)$$

- this tells us the distance $x-x^*$ to the true solution gets multiplied by $f'(x^*)$ each iterations. If $f'(x^*) < 1$ then our iteration will converge. If $f'(x^*) > 1$ then with each step we will get farther away.

RELAXATION METHOD

- If $f'(x^*) > 1$ we can try to invert the function to get $f^{-1}(x)$, since $f'(x^*) > 1$ the derivative of $f^{-1}(x^*)$ should be less than one and the iteration will converge.
- Of course it is not always possible to invert a function, but in cases where you can the inverted function should work with the relaxation method.
- In summary the relaxation method does not always work. But between the cases where it works on the first try, or it works on the inverted function, or there is some other way to write the function where it will work it is a very useful method for solving nonlinear equations.



EXERCISE 6.10

.....

- Consider the equation $x = 1 - e^{-cx}$, where c is a known parameter. This equation arises in a variety of situations, including the physics of contact processes and mathematical models of epidemics.
- Write a program to solve this equation for x using the relaxation method for $c = 2$.
- Modify your program to calculate the solution for c from 0 to 3 in steps of 0.1 and make a plot of x as a function of c .

RELAXATION METHOD FOR TWO OR MORE VARIABLES

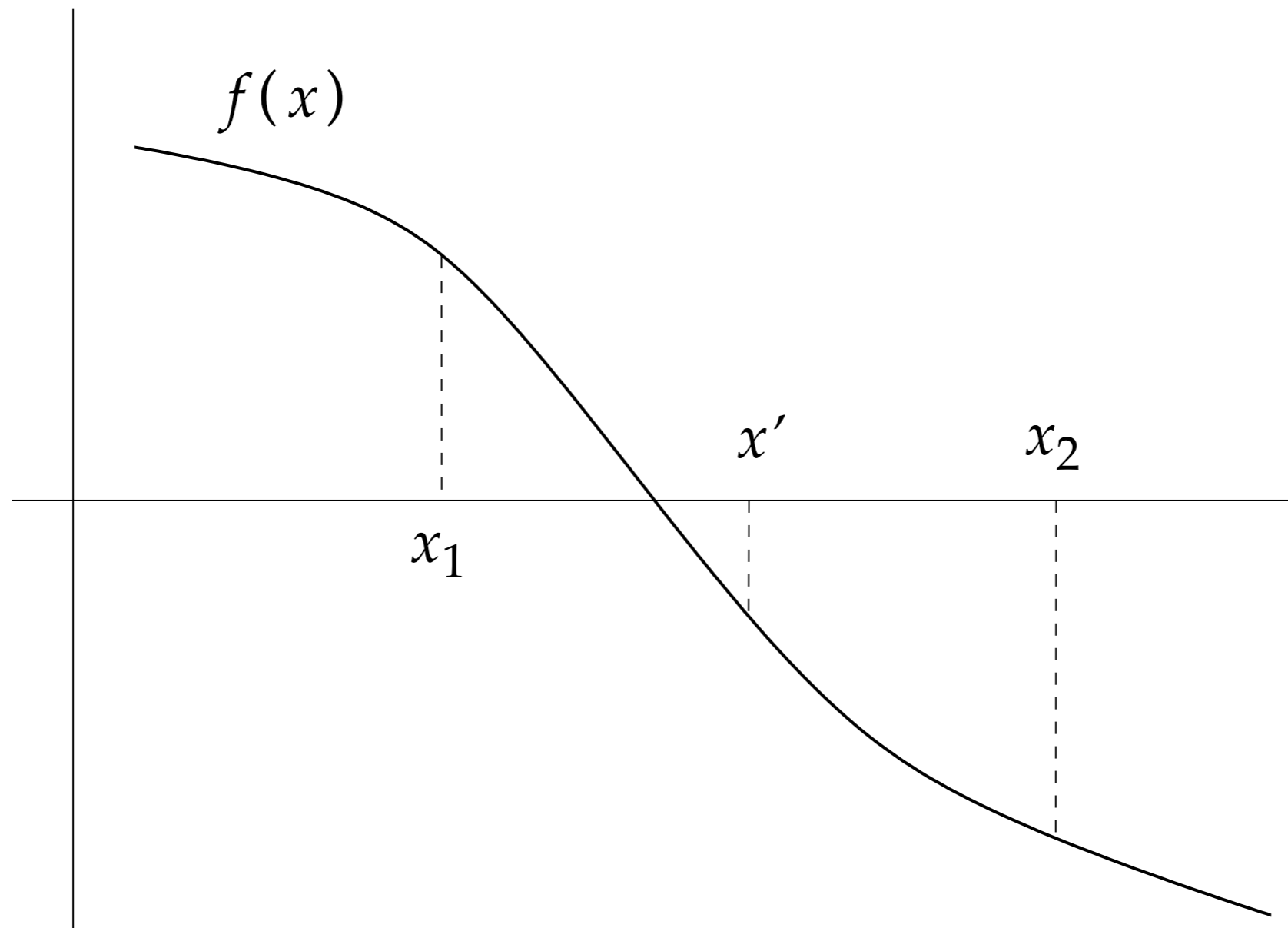
- If instead we have equations for two variables

$$x = f(x, y) \qquad y = g(x, y)$$

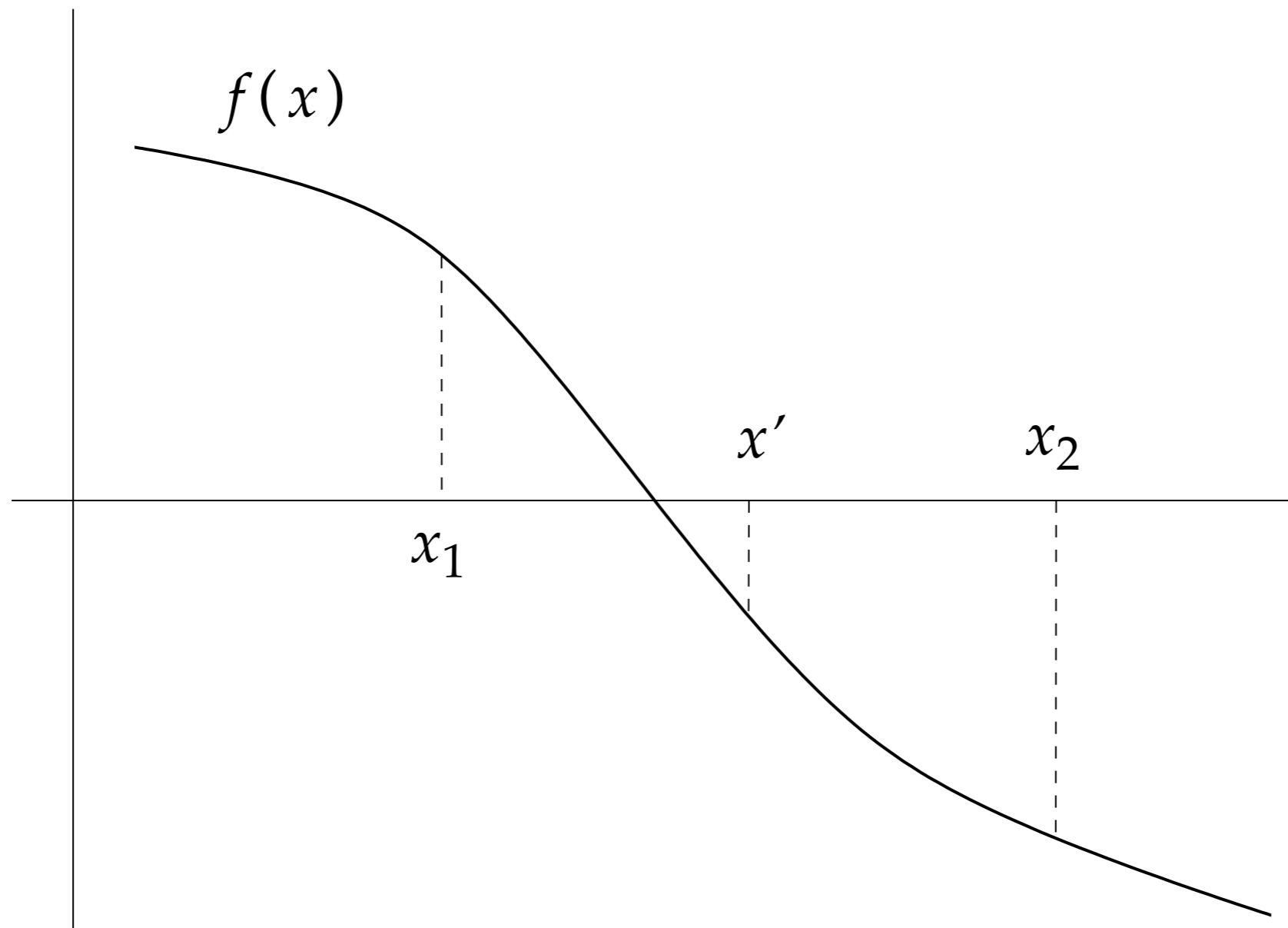
- the relaxation method works just as well as long as each equation can be written in a form that gives back the variable.
- This is one of the strengths of the relaxation method, it is easily extended to many nonlinear equations of multiple variables.
- We'll see other methods are not so easily extended to multiple variables.

BISECTION

- Binary search, or the bisection method is an alternative and more robust method for solving nonlinear equations for a single variable x .
- With this method if a solution exists in a specified interval the method will always find it.
- You can always rearrange an equation into the form $f(x)=0$.
- Thus finding solutions to $f(x)$ is the same as finding the zeros or roots of $f(x)$.
- Bisection works by finding roots.



- Suppose we want to find a root of $f(x)$ between x_1 and x_2 .
- We start by evaluating $f(x_1)$ and $f(x_2)$. If we discover that one of those values is positive and the other negative then so long as $f(x)$ is continuous, there must be at least one point (but possibly more) between x_1 and x_2 where $f(x) = 0$.



- Now let's take a new point half way between x_1 and x_2 , $x' = 1/2(x_1 + x_2)$ and evaluate $f(x')$. $f(x')$ will either be positive or negative (unless it is zero in which case we have our root).
- If we now use $f(x')$ in place of our previous value with the same sign and repeat the procedure we will move closer and closer to the solution we are searching for.

OUTLINE OF BISECTION

1. Given an initial pair of points, x_1, x_2 , check that $f(x_1)$ and $f(x_2)$ have opposite signs. Choose an accuracy, ϵ .
 2. Calculate the midpoint $x' = 1/2(x_1 + x_2)$ and $f(x')$.
 3. If $f(x')$ has the same sign as $f(x_1)$ replace $x_1 = x'$. Otherwise $x_2 = x'$.
 4. If $|x_1 - x_2| > \epsilon$ repeat from step 2. Otherwise calculate $1/2(x_1 + x_2)$ again and that is your answer.
- As with relaxation, bisection improves exponentially. With each step you are twice as close to a solution.

BISECTION

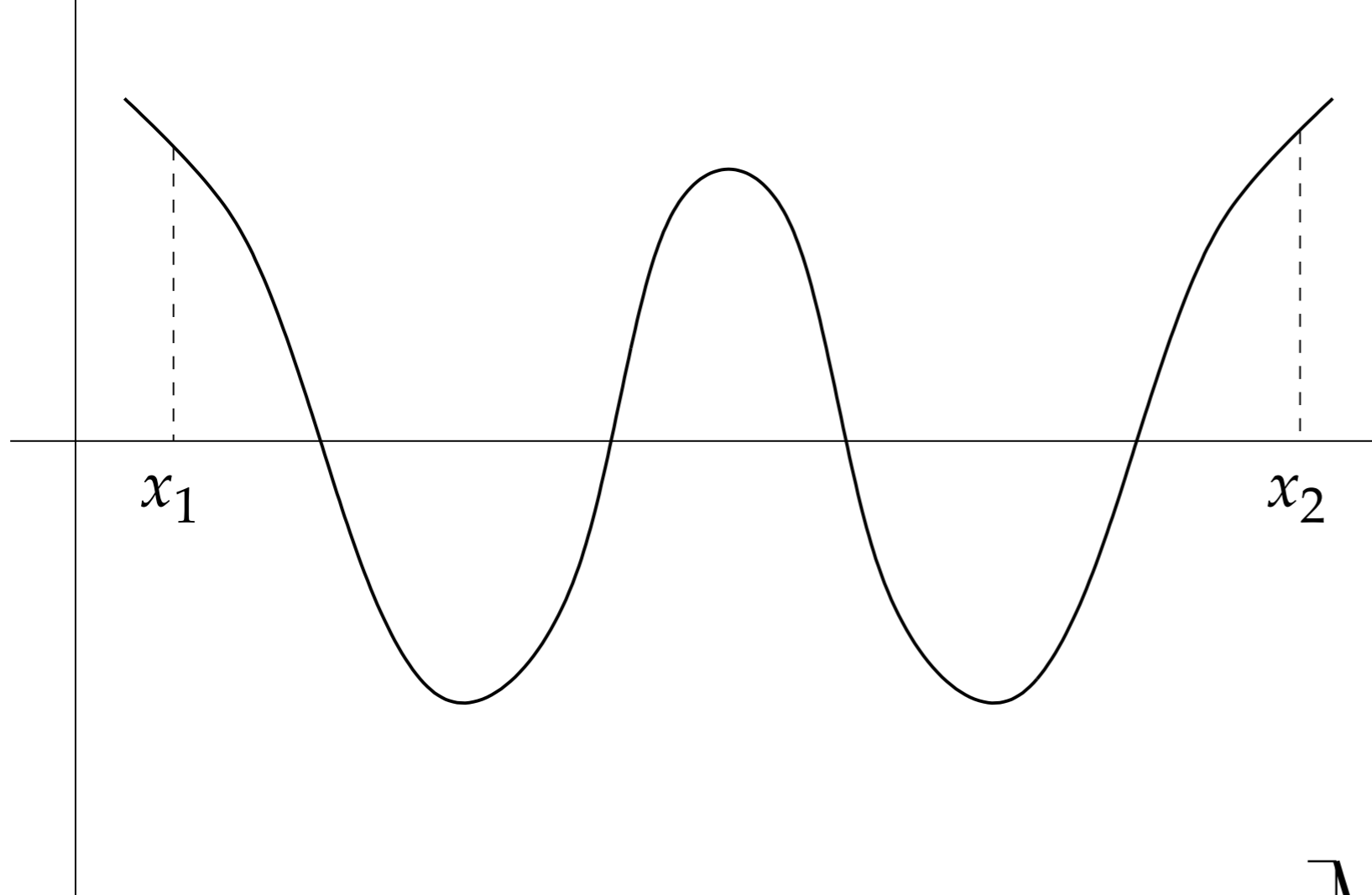
- Suppose the initial distance between our points is Δ . Then with each step the distance goes down by $\Delta/2^N$. We end when the distance is our desired accuracy. So the number of steps needed is

$$N = \log_2 \frac{\Delta}{\epsilon}$$

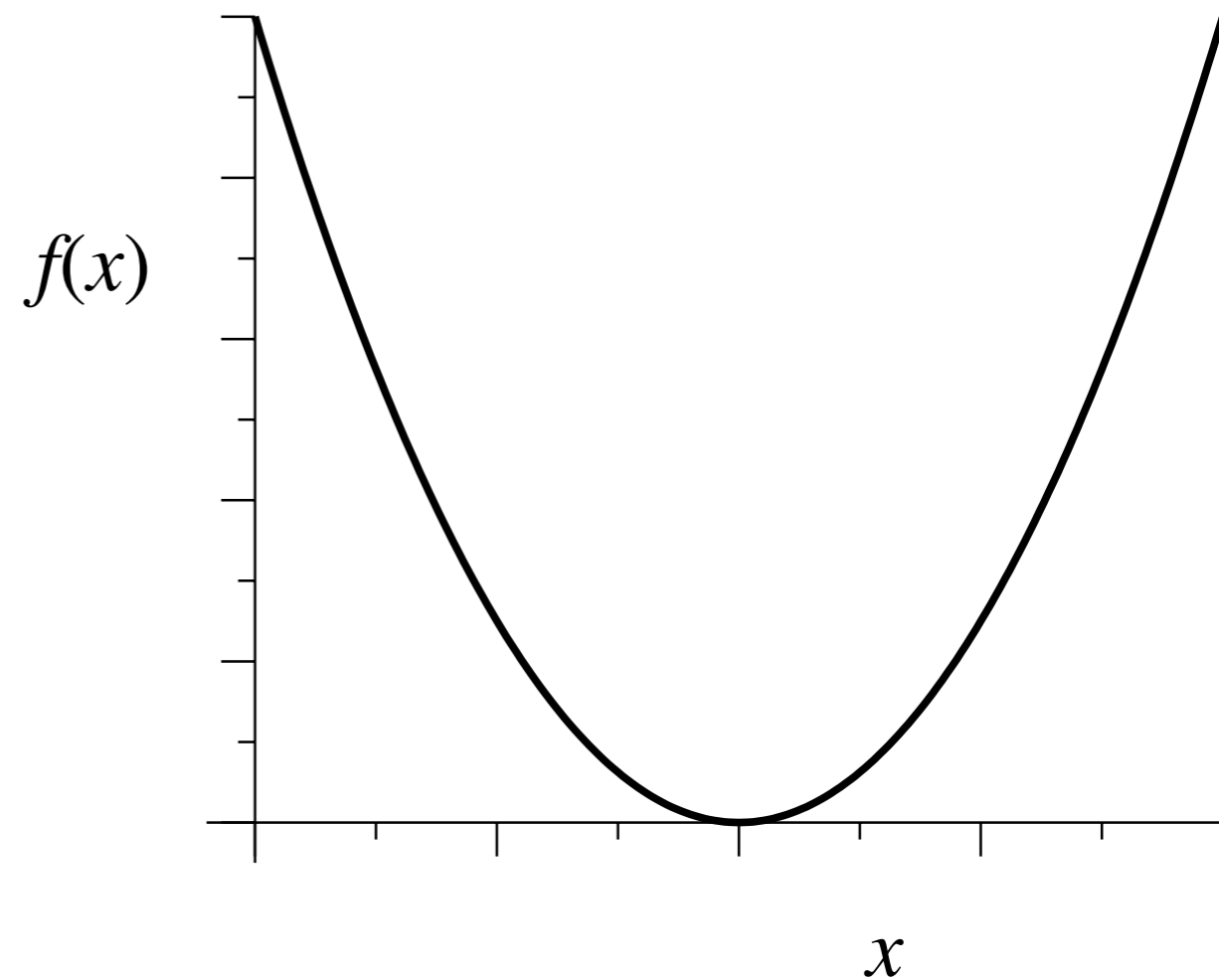
- Log is a very slowly growing function, even if Δ is much larger than ϵ this is not a log of steps. If we start with a $\Delta = 10^{10}$ and we require an accuracy of $\epsilon = 10^{-10}$, that still only requires $\log_2(10^{20}) \sim 66$ steps.
- In other words if $f(x_1)$ and $f(x_2)$ are of different signs you will always find a root in a short number of steps.

BISECTION

- Bisection does have a number of shortcomings, most obviously if $f(x_1)$ and $f(x_2)$ have the same sign then you can't find a root. This may be because there is no root between these values, or because there are an even number, or because the function just touches zero but doesn't cross it.
- Thus bisection works best if you know enough about your function to know some places where it is positive and negative.
- If you are looking for a root near a particular value of x , start with x_1 and x_2 near that value. If they don't bracket the root, then keep doubling the distance to x_1 and x_2 till they do bracket the root.
- A final problem with bisection is that it only work for single equations of one variable.



- If you can find local maxima and minima for your function then these issues can be solved.



NEWTON'S METHOD

- The fastest way to find roots of an equation is Newton's method.
- Newton's method is similar to bisection, except we use the functions derivative to estimate where we should make our next evaluation instead of the halfway point.
- Using the derivative we should end up with a closer guess, though in some cases this doesn't work.
- Note that if the function was a line then evaluating it at any point and using the slope to determine where it crosses zero would work exactly.

NEWTON'S METHOD

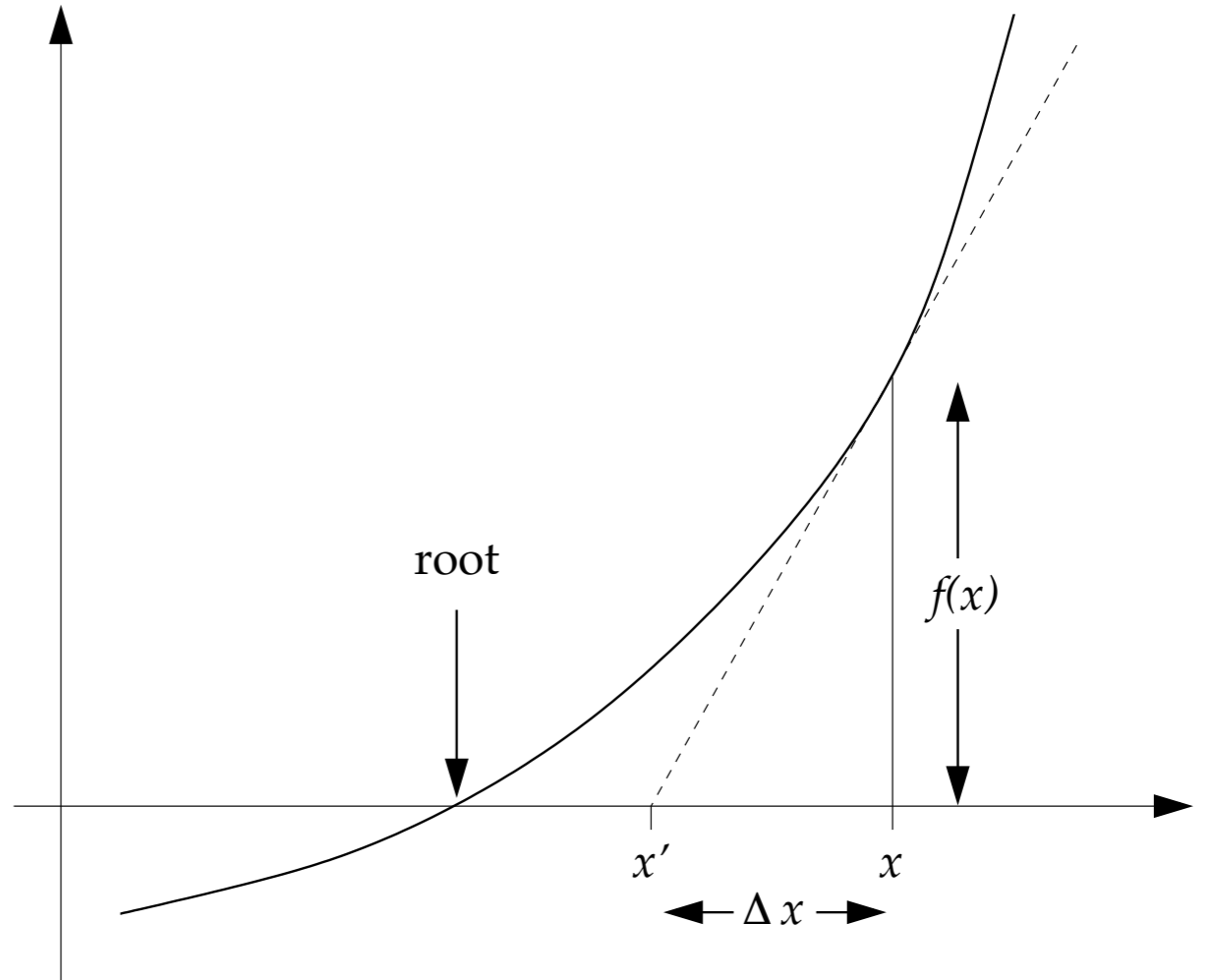
- We can see from the plot that

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

- so our new guess for x' would be

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

- The only catch with this approach is that it requires us to be able to evaluate $f'(x)$.



ACCURACY OF NEWTON'S METHOD

- As always Taylor expand around the solution x^*

$$f(x^*) = f(x) + (x^* - x)f'(x) + \frac{1}{2}(x^* - x)^2 f''(x) + \dots$$

- now since x^* is a root, $f(x^*) = 0$ so we get

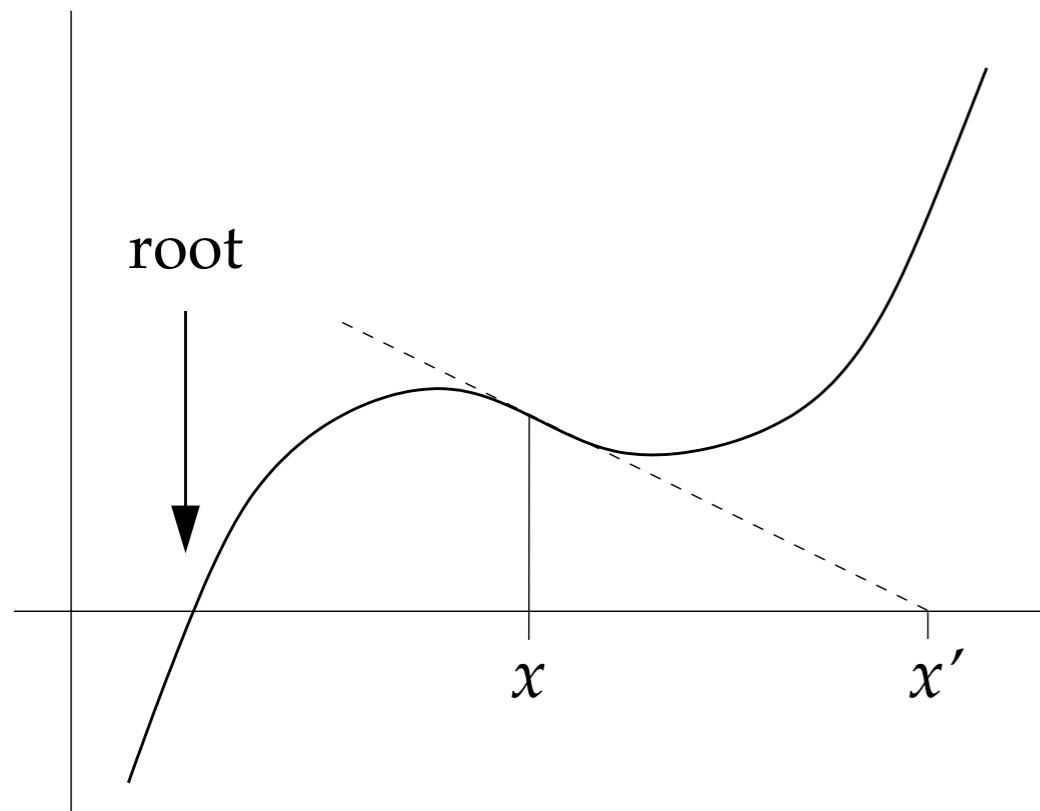
$$x^* = \left[x - \frac{f(x)}{f'(x)} \right] - \frac{1}{2}(x - x^*) \frac{f''(x)}{f'(x)} + \dots$$

- the quantity in square brackets is our estimate x' . If we define our error $x^* = x + \epsilon$ and $x^* = x' + \epsilon'$ the above equation tells us

$$\epsilon' = \frac{-f''(x)}{2f'(x)} \epsilon^2$$

- That is assuming $f''(x) \sim f'(x)$ the error will decrease by ϵ^2 each step. That is incredibly fast convergence, what is called *quadratic convergence*.

ACCURACY OF NEWTON'S METHOD



- If Newton's method behaves strangely it may be because of one of these issues.
- The convergence is so fast that we can assume the error on the last iteration is negligible compared to the previous one and thus know our total error is less than that difference of the two.
- Newton's method has two main issues. We need to be able to calculate derivatives and more importantly Newton's method doesn't always converge. If $f'(x)$ is very small then the error can get larger with each iteration, and if the function slopes the wrong way then we can get farther away instead of closer to the root.

EXERCISE 6.15

$$P(x) = 924x^6 - 2772x^5 + 3150x^4 - 1680x^3 + 420x^2 - 42x + 1$$

- ▶ There is no general formula for the roots of a sixth-order polynomial, but one can find them easily enough using a computer. Make a plot of $P(x)$ from $x=0$ to $x=1$ and by inspecting it find rough values for the six roots of the polynomial - the points at which the function is zero.
- ▶ Write a Python program to solve for the positions of all six roots to at least ten decimal places of accuracy, using Newton's method.
- ▶ Note that the polynomial in this example is just the sixth Legendre polynomial (mapped onto the interval from zero to one), so the calculation performed here is the same as finding the integration points for 6-point Gaussian quadrature (see Section ~5.6.2), and indeed Newton's method is the method of choice for calculating Gaussian quadrature points.

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

SECANT METHOD

- As mentioned earlier one of the problems with Newton's method is that we have to be able to calculate the derivative.
- If we have an analytic expression for the function that is straightforward, but if the values come from an experiment or a numerical calculation we won't have an analytic expression.
- In that case we can always take a numerical derivative. In fact to be efficient we can just use two of our guesses for the root and take a numerical derivative between them.
- This is called the secant method we approximate the derivative as

$$f'(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

- so the next point is given by

$$x_3 = x_2 - f(x_2) \frac{x_2 - x_1}{f(x_2) - f(x_1)}$$

NEWTON'S METHOD FOR TWO OR MORE VARIABLES

- ▶ Newton's method can also be used to solve systems of nonlinear equations. If we have N equations with N variables

$$f_1(x_1, \dots, x_n) = 0$$

⋮

$$f_N(x_1, \dots, x_n) = 0$$

- ▶ If we do our Taylor expansion, but lets write $\mathbf{x} = x_1, \dots, x_n$ to give

$$\mathbf{f}(\mathbf{x}^*) = \mathbf{f}(\mathbf{x}) + \nabla \mathbf{f} \cdot (\mathbf{x}^* - \mathbf{x})$$

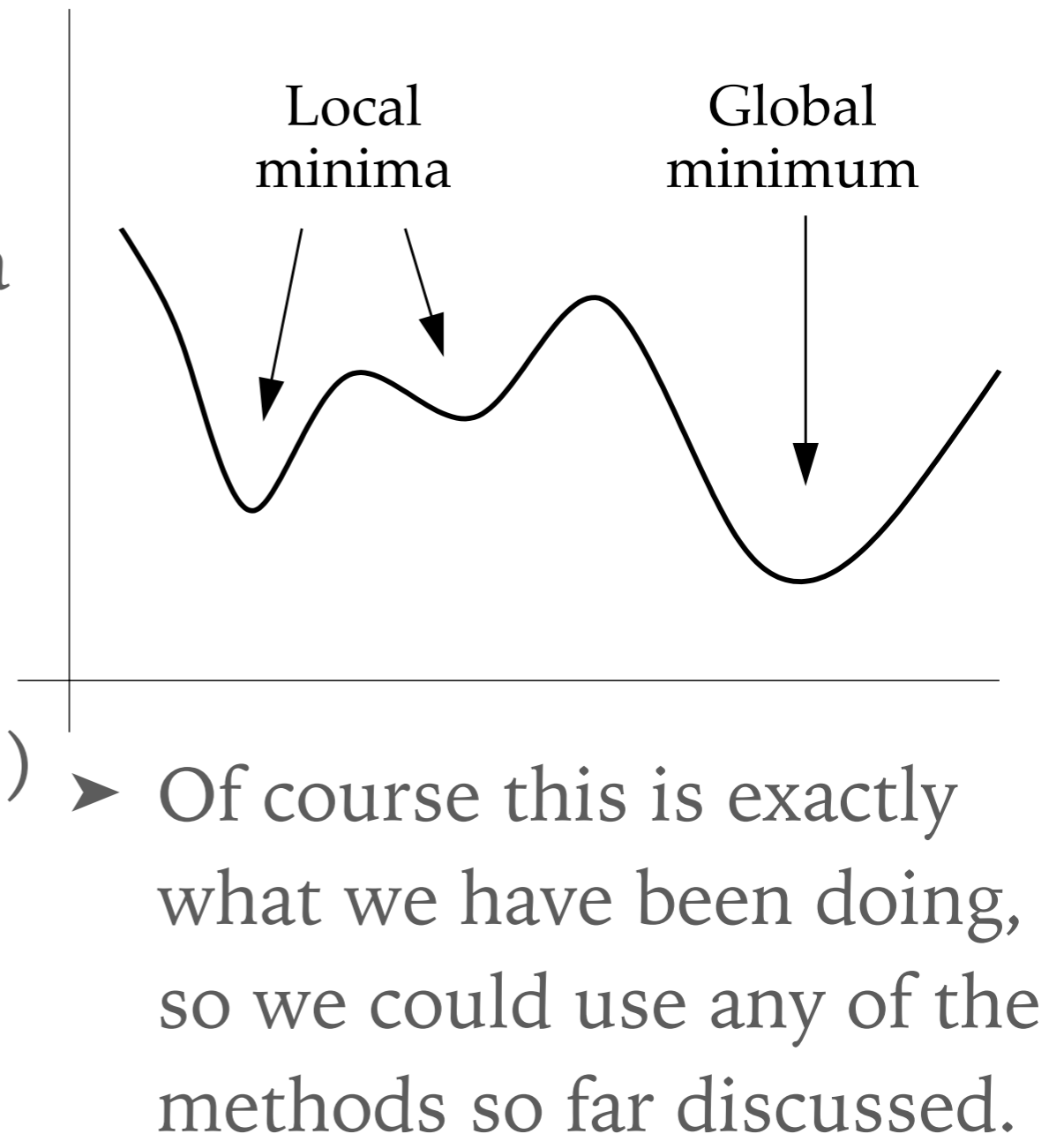
- ▶ where $\nabla \mathbf{f}$ is a $N \times N$ matrix whose elements are $\partial f_i / \partial x_j$. Since \mathbf{x}^* is a root of the equation $\mathbf{f}(\mathbf{x}^*) = 0$ and we have

$$\nabla \mathbf{f} \cdot \Delta \mathbf{x} = \mathbf{f}(\mathbf{x})$$

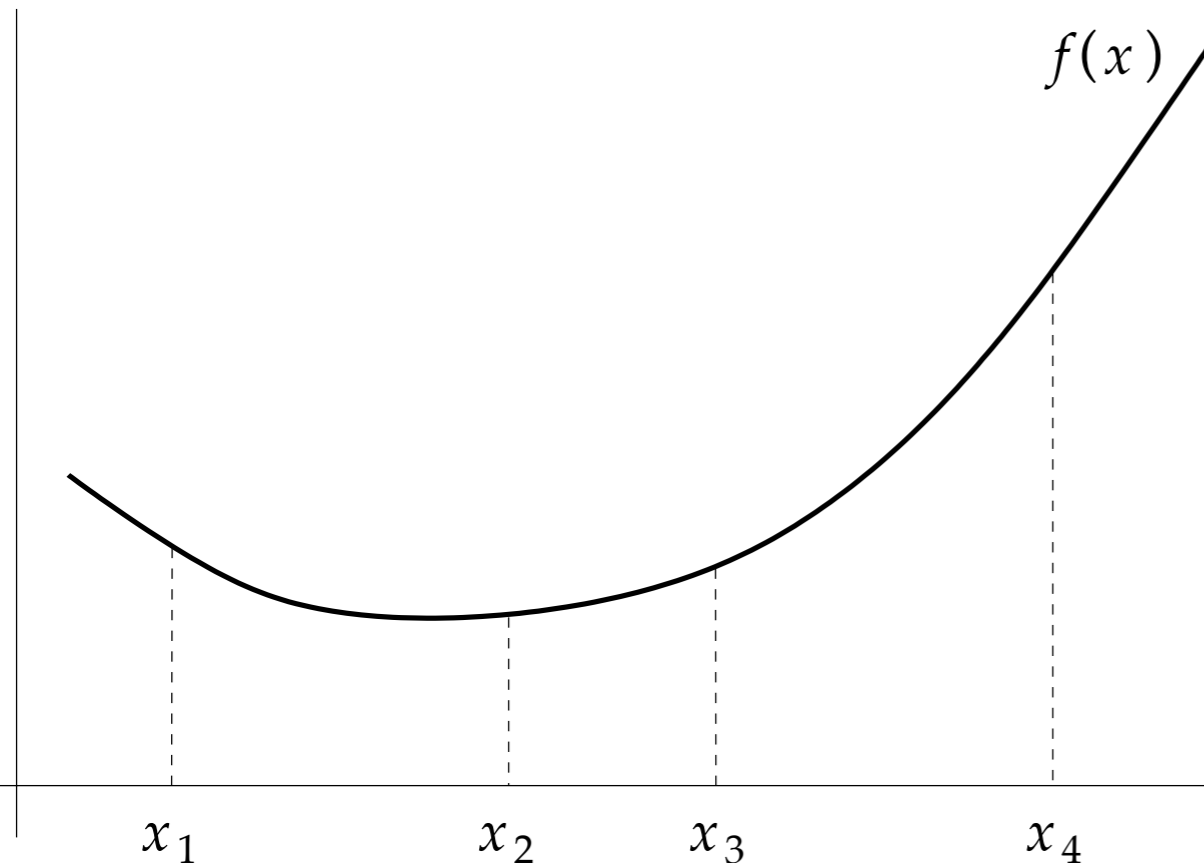
- ▶ but this is just a set of linear equation like $\mathbf{A}\mathbf{x} = \mathbf{v}$ which we can solve by Gaussian elimination. Once we solve for $\Delta \mathbf{x}$ we get our new guess $\mathbf{x}' = \mathbf{x} - \Delta \mathbf{x}$ and we just keep iterating to we are happy with our \mathbf{x}' .

MAXIMA AND MINIMA

- Closely related to the problem of finding roots is the problem of finding maxima and minima of a function.
- Functions can have local and global minima.
- If we have a function $f(x_1, x_2, \dots)$ and we take its derivatives then solving $\partial f / \partial x_i = 0$ for all i will give us the maxima and minima.



GOLDEN RATIO SEARCH



- A method to find a minima without derivatives is the golden ratio search which is similar to bisection, but we use four unevenly spaced points.
- Suppose at least one of $f(x_2)$ or $f(x_3)$ is less than $f(x_1)$ and $f(x_4)$. Then we know there must be a minimum between x_1 and x_4 .
- If $f(x_2)$ is less than $f(x_3)$ we know the minimum lies between x_1 and x_3 .
- We can now choose a new point between them and narrow our range for minimum is even more.

GOLDEN RATIO SEARCH

- How should we choose where to put our points?
- x_2 and x_3 should be symmetric around the center of x_1 and x_4 since we don't know on which side of the minimum it will land.
- In order to make the method as efficient as possible we want to range in which the minimum is located to decrease as quickly as possible. If we choose x_2 and x_3 close to the center then it makes this step efficient, but the next step less efficient. Let us define z as the ratio between the intervals in two steps.

$$z = \frac{x_4 - x_1}{x_3 - x_1} = \frac{x_2 - x_1}{x_3 - x_1} + 1$$

GOLDEN RATIO SEARCH

- For the next step the ratio will be

$$z = \frac{x_3 - x_1}{x_2 - x_1}$$

- setting these two equal we find $z^2 - z - 1 = 0$

- whose solution that is greater than 1 is $z = \frac{1 + \sqrt{5}}{2} = 1.618\dots$

- this value is called the golden ratio and it pops up repeatedly in physics and mathematics as well as art and architecture.

- so we find that the interior points should be $1/z$ or 0.618 of the way between the exterior points.

- So the golden ratio search algorithm is the following

GOLDEN RATIO SEARCH ALGORITHM

1. Choose two initial outside points x_1 and x_4 , then calculate two interior points using the golden ratio. Evaluate $f(x)$ at all the points. Choose an ε .
2. If $f(x_2) < f(x_3)$ then the minimum lies between x_1 and x_3 . In this case x_3 becomes the new x_4 , x_2 becomes the new x_3 and there is a new x_2 chosen by the golden ratio. Evaluate new $f(x)$.
3. If $f(x_2) > f(x_3)$ then the minimum lies between x_2 and x_4 , x_2 becomes the new x_1 , x_3 becomes the new x_2 and new value for x_3 . Evaluate new $f(x)$.
4. If $x_4 - x_1 > \varepsilon$, repeat from step 2. Otherwise $1/2(x_2 + x_3)$ is your solution.

GAUSS-NEWTON METHOD

- The golden ratio method suffers from the same issues as bisection. You have to bracket the maximum/minimum to find it and it only work for single variables.
- An alternative approach is to used $f'(x)=0$ and look for its roots. If we use Newton's method to find the roots then we have

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

- which is the fundamental formula for the Gauss-Newton method. This can be generalized for systems of nonlinear equations with many variables.

GRADIENT DESCENT

- However, it is often the case that it will not be possible to evaluate the second derivative of our function. It turns out that just using a constant γ as a rough guess of $1/f''(x)$ works in practice very well. This is called gradient descent, where we advance based only on the first derivative

$$x' = x - \gamma f'(x)$$

- this method depends on a good guess for γ , if it is too small then you will converge very slowly. If it is too big you will overshoot the solution. In cases where you can't evaluate the first derivative you can use a numerical approximation,

$$x_3 = x_2 - \gamma \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

SCIPY.OPTIMIZE

- Most of the algorithms we have discussed as well as many many others can be found in the `scipy.optimize` sub package.
- Optimization is a very important numerical technique that we will discuss again later.
- One dimensional root finding uses the `root_scalar(func, args, method=“”)` where the method can be set to ‘bisection’, ‘newton’, ‘secant’ and many other choices.
- The function `minimize(func, x0, method=“”)` performs multidimensional or 1d minimization with the specified method, most of which we have not yet discussed.
- Note that methods that only work in 1d must use the `root_scalar()` function.