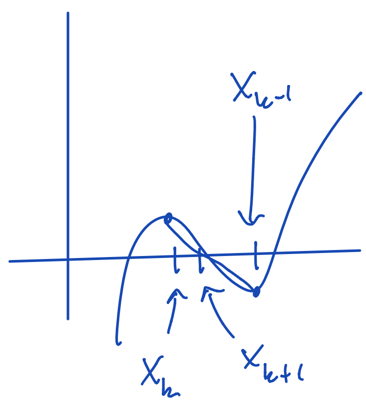
Lecture 3 Numerical Analysis 1/30/18

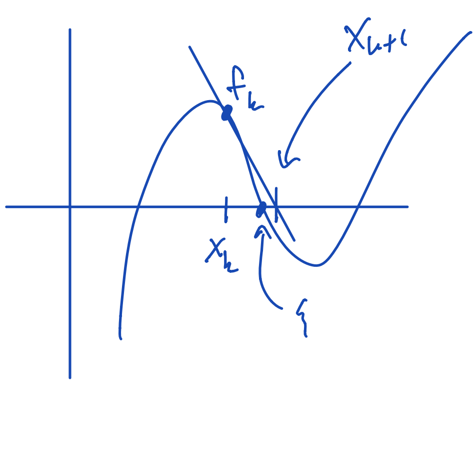
Last time: Methods for solving $\underbrace{f(x)}_{\text{possibly a nonlinear function}}=0$

1- secant method



$$x_{k+1} = x_k - \underbrace{\frac{x_k - x_{k-1}}{f_k - f_{k-1}}}_{\text{approximation of } 1/f'(x_k)} f_k$$

2- Newton's Method



$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Convergence is quadratic:

$$|\xi - x_{k+1}| \sim |\xi - x_k|^2$$

Now: Revisit the secant method:

$$x_{k+1} = x_k - \underbrace{\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}}_{\substack{\text{an approximation} \\ \text{to } 1/f'(x_k)}} f(x_k)$$

In general, the secant method converges slower than Newton's Method, with a rate of:

$$\lim_{k\to\infty} \frac{|\xi - x_{k+1}|}{|\xi - x_k|^q} = \text{const.}$$

$q = \frac{1}{2}(1+\sqrt{5}) \approx 1.6$

(proof omitted)

How about in 2 dimensions? What does Newton's Method look like?

Ex: Solve

$$f_1(x_1, x_2) = 0$$
$$f_2(x_1, x_2) = 0$$

$\iff \underbrace{\vec{f}(\vec{x}) = \vec{0}}_{\text{vector form.}}$

Let's recall the multi-variable Taylor series expansion:

$$f(x_1,x_2) = f(y_1,y_2) + \left[\frac{\partial f}{\partial x_1}(y_1,y_2)\right](x_1-y_1) + \left[\frac{\partial f}{\partial x_2}(y_1,y_2)\right](x_2-y_2)$$

$$\text{2nd-order terms}\left\{\begin{aligned} &+ \left[\frac{\partial^2 f}{\partial x_1 \partial x_2}(y_1,y_2)\right](x_1-y_1)(x_2-y_2) \\ &+ \frac{1}{2}\left[\frac{\partial^2 f}{\partial x_1^2}(y_1,y_2)\right](x_1-y_1)^2 \\ &+ \frac{1}{2}\left[\frac{\partial^2 f}{\partial x_2^2}(y_1,y_2)\right](x_2-y_2)^2 \end{aligned}\right. \quad + \ldots$$

But $\underline{f \text{ was scalar-valued}}$. What is the vector version of this Taylor expansion?

$$\vec{f}(\vec{x}) = \vec{f}(\vec{y}) + D\vec{f}(y)\,(\vec{x}-\vec{y}) + (\vec{x}-\vec{y})^T H\vec{f}(\vec{y})(\vec{x}-\vec{y}) + \ldots$$

where D is the Jacobian (also sometimes called the derivative of $\vec{f}$) and H is the Hessian of $\vec{f}$, involving 2^{nd} derivative: (it is a 3^{rd} order tensor).

$$D\vec{f} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix}$$

$D\vec{f}$ is a matrix

$H\vec{f}$ is not a matrix, but a tensor.

Drop the H term, then we have:

$$\vec{f}(\vec{x}) \approx \vec{f}(\vec{y}) + D\vec{f}(\vec{y})\,(\vec{x}-\vec{y})$$

At $\vec{f}(\vec{x}) = \vec{0}$, we have that (i.e. solve for $\vec{x}$)

$$\vec{x} \approx \vec{y} - \underbrace{\left(D\vec{f}(\vec{y})\right)^{-1}}_{\text{the inverse of the Jacobian, evaluated at } \vec{y}.} \vec{f}(\vec{y})$$

Therefore Newton's method in higher dimensions is given by:

$$\vec{x}_{k+1} = \vec{x}_k - \left(D\vec{f}(\vec{x}_k) \right)^{-1} \vec{f}(\vec{x}_k)$$

and convergence is quadratic, assuming that $D\vec{f}$ is invertible, given by

$$\| \vec{\xi} - \vec{x}_{k+1} \| \sim \| \vec{\xi} - \vec{x}_k \|^2$$

$\| \cdot \|$ denotes the usual Euclidean norm.

The above formula is the same for n-dimensional root finding:

$$\vec{f}(\vec{x}) = \vec{0} \longrightarrow \begin{pmatrix} f_1(x_1, x_2, \dots x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

All of these methods for solving $f(x)=0$ have the following in common: the true root ξ is a <u>fixed-point</u> of the iteration.

<u>Ex:</u> Newton's Method:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

If $x_k = \xi$, then

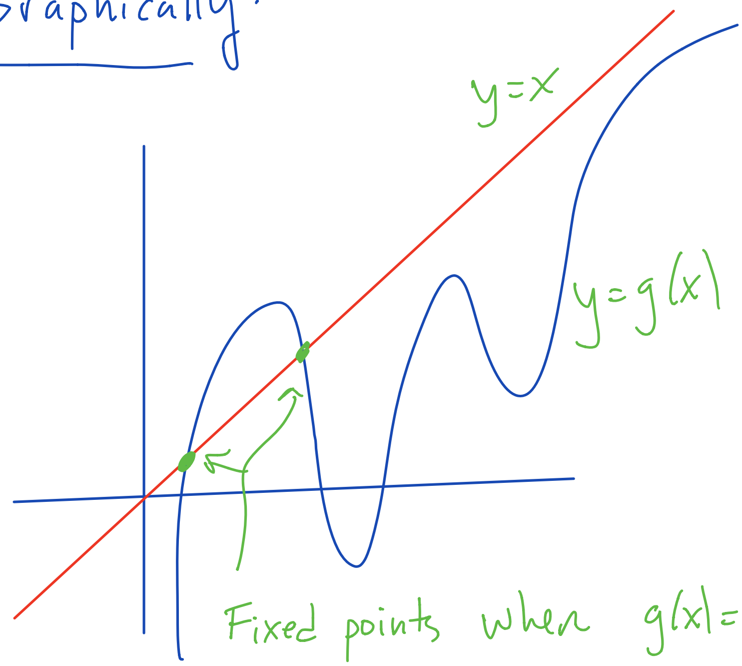
$$x_{k+1} = \xi - \frac{f(\xi)}{f'(\xi)} = \xi - \frac{0}{f'(\xi)} = \xi .$$

So $x_k = \xi \Rightarrow x_{k+1} = \xi$.

This way of thinking can be generalized to the question: Are there any values of x such that for some function g, $g(x)=x$?

$\Rightarrow$ Examine the iteration $x_{k+1} = g(x_k)$.

Graphically:



Fixed points when $g(x) = x$.

Note that solving $f(x)=0$ is equivalent to finding a fixed point of $g(x) = f(x)+x$

since $g(x) = x$

$\Rightarrow f(x)+x = x$

Example Newton's Method looks for a fixed point of the function

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

Definition A simple iteration is given by

$$x_{k+1} = g(x_k)$$

when we assume g is continuous on an interval $[a,b]$, and that $g(x) \in [a,b]$ for all $x \in [a,b]$ (this condition guarantees the existence of a fixed point by the following theorem):

Brouwer's Fixed Point Thm: If g is as above then there exists $\xi \in [a,b]$ such that $\xi = g(\xi)$.

Proof: If $g(x) \in [a,b]$ for all $x \in [a,b]$, then it is easy to show that $g(x) - x = h(x)$ has at least one root in $[a,b]$:

Show that $h(a)h(b) < 0$.

Next: without looking at the range of g on $[a,b]$, examine how it maps nearby points:

Definition (Contraction) Let g be continuous on $[a,b]$. The function g is a contraction on $[a,b]$ if there exists a number L with $0 < L < 1$ such that

$$\underbrace{|g(x) - g(y)| < L|x-y|}_{\text{This means that } g \text{ maps points to values which are closer together.}} \quad \text{for all } x,y \in [a,b].$$

If L is allowed to be any positive number, this is known as a Lipschitz condition.