2/14/2019

→ Finish Secant Method, Newton's Method.
→ Discuss bisection.
→ Begin Linear Algebra

Thm: 1.8 Convergence of Newton's method.

S'pose f is cnts, and also has cnts second derivative f'', defined on the closed interval $I_\delta = [\xi-\delta, \xi+\delta]$,

$\delta > 0$ so that $f(\xi) = 0$ and $f''(\xi) \neq 0$.

S'pose $\dfrac{|f''(x)|}{|f'(y)|} \leq A \quad \forall x, y \in I_\delta$.

If $|\xi - x_0| \leq h$, $h = \min(\delta, \frac{1}{A})$,

then (x_k) defined by Newton's method converges quadratically to ξ.

Pf: Suppose $|\xi - x_k| \leq h = \min(\delta, \frac{1}{A})$,

i.e. $x_k \in I_\delta$. use a Taylor Expansion

$$f(\xi) - f(x_k) = f'(x_k)(\xi - x_k) + \frac{1}{2} f''(\eta_k) * (\xi - x_k)^2$$

For some $\eta_k \in I\delta$.

We know $f(\xi) = 0$. Also, looking at the Newton iteration

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

$$= x_k + \frac{1}{f'(x_k)} \left(f'(x_k)(\xi - x_k) + \frac{1}{2} f''(\eta_k)(\xi - x_k)^2 \right).$$

$$= \xi + \frac{1}{2} \frac{f''(\eta_k)}{f'(x_k)} (\xi - x_k)^2$$

$$\Rightarrow \xi - x_{k+1} = -\frac{1}{2} \frac{f''(\eta_k)}{f'(x_k)} (\xi - x_k)^2. \quad (**)$$

Note that we have by assumption

$$|\xi - x_k| \le \frac{1}{A}$$

$$\left| \frac{1}{2} \frac{f''(\eta_k)}{f'(x_k)} \right| \le \frac{A}{2},$$

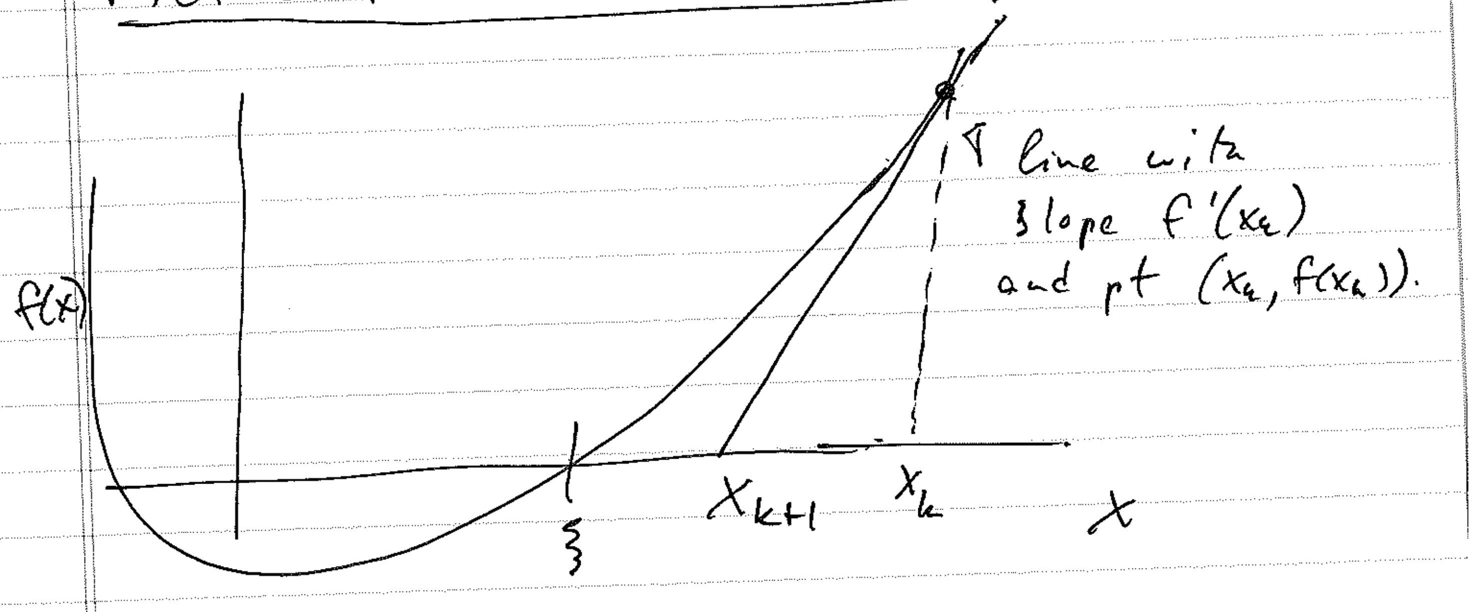
So $|\xi - x_{k+1}| \le \frac{1}{2} |\xi - x_k|. \quad (*)$

$(*)$ shows that $|\xi - x_{k+1}| \le 2^{-k} h \quad k \ge 0,$

so $x_k \to \xi$.

We get quadratic convergence immediately from $(**)$, assuming f'' and f' are continuous.

Picture for Newtons method:



The line above is $y = f'(x_k)\, x + b$.

and contains the point $(x_k, f(x_k))$,

so $b = f(x_k) - f'(x_k) x_k$.

If we solve for the point at which $y = 0$, we obtain

$$0 = f'(x_k)\, x + \left(f(x_k) - f'(x_k) x_k \right)$$

$$\Leftrightarrow \quad x = x_k - \frac{f(x_k)}{f'(x_k)},$$

which is precisely x_{k+1}!.

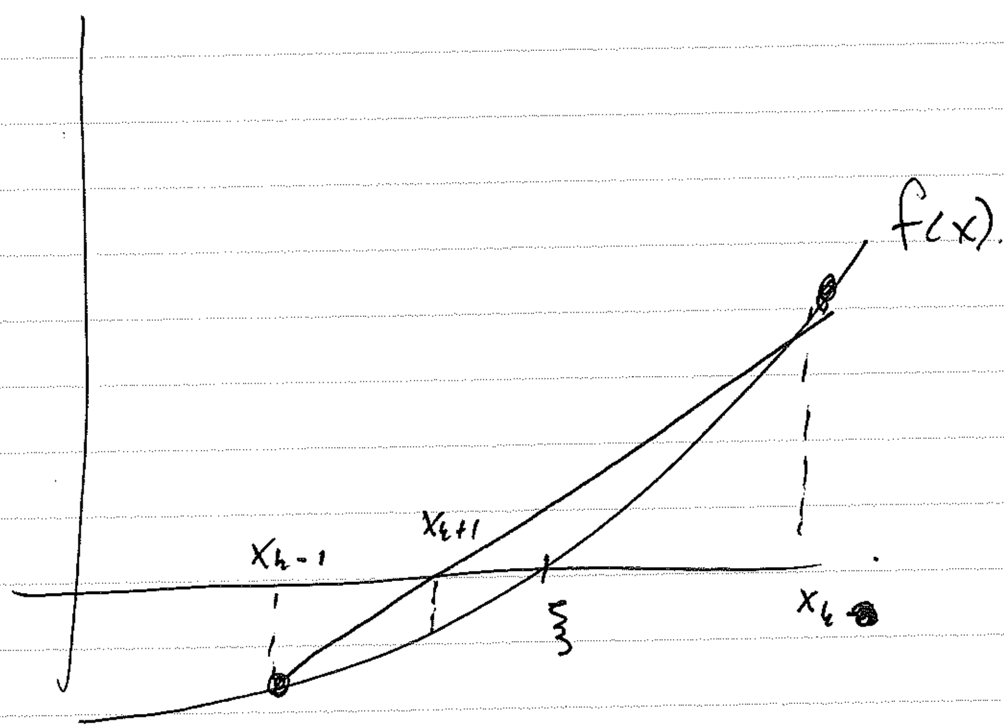
Secant Method:

The idea here is that maybe we don't know the derivative of f, or it is too costly to compute. Then try replacing the derivative with an approximation:

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}},$$

And then the Secant method is:

$$x_{k+1} = x_k - f(x_k)\left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}\right).$$



Thm 1.10:

Suppose f is cntsly differentiable on $I = [\xi - h, \xi + h]$, $h > 0$. Suppose that $f(\xi) = 0$, $f'(\xi) \neq 0$. Then (x_k) defined by the secant method converges at least linearly to ξ provided x_0 and x_1 are sufficiently close to ξ.

Pf: Since $f'(\xi)$, we may suppose $f'(\xi) = \alpha > 0$, WLOG.

Since f' is continuous on I, $\exists$ $I_\delta = [\xi - \delta, \xi + \delta]$, with $0 < \delta \leq h$ so that

$$|f'(x) - \alpha| < \varepsilon, \quad x \in I_\delta.$$

Take $\varepsilon = \frac{1}{4}\alpha$, so we get

$$0 < \frac{3}{4}\alpha < f'(x) < \frac{5}{4}\alpha \quad \forall x \in I_\delta.$$

Now, consider the secant iteration:

$$x_{k+1} = x_k - f(x_k)\left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}\right).$$

We can <u>introduce 0</u> in several places:

$$x_{k+1} - \xi = x_k - \xi - \left(f(x_k) - f(\xi)\right)\left(\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}\right).$$

using the mean value theorem:

$$\begin{cases} f(x_k) - f(\xi) = f'(\eta_1)\,(x_k - \xi) \\ f(x_k) - f(x_{k-1}) = f'(\eta_2)\,(x_k - x_{k-1}). \end{cases}$$

Plug these in to above:

$$x_{k+1} - \xi = x_k - \xi - \frac{f'(\eta_1)\,(x_k - \xi)}{f'(\eta_2)}.$$

So we can bound in the following way:

$$\begin{aligned} |x_{k+1} - \xi| &= \left| 1 - \frac{f'(\eta_1)}{f'(\eta_2)} \right| |x_k - \xi| \\ &\le \left| 1 - \frac{5\alpha/4}{3\alpha/4} \right| |x_k - \xi| \\ &= \frac{2}{3} \, |x_k - \xi|. \end{aligned}$$

Example:

Find a root, or roots, of

$$\underline{f}(x,y) = \begin{bmatrix} (x-3)^2 y \\ (y-2)\, x^4 \end{bmatrix} = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \end{bmatrix}$$

Newton's method for a vector valued function looks like:

$$\underline{x}_{n+1} = \underline{x}_n - \left(\underline{\underline{\nabla f}}(\underline{x}_n)\right)^{-1} \underline{f}(\underline{x}_n)$$

$$\underline{\underline{\nabla f}} = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}.$$

$$f_1 = x^2 y - 6xy + 9y$$

$$f_2 = x^4 y - 2x^4$$

$$\underline{\underline{\nabla f}} = \begin{bmatrix} 2xy - 6y & x^2 - 6x + 9 \\ 4x^3 y - 8x^3 & x^4 \end{bmatrix}$$

What are the roots?

$$(x, y) = (3, 2)$$

$$(x, y) = (0, 0).$$

Let's evaluate the Jacobian $\underline{\underline{\nabla f}}$

at the roots:

$$\underline{\underline{\nabla f}}(3,2) = \begin{bmatrix} 0 & 0 \\ 0 & 81 \end{bmatrix}.$$

$$\underline{\underline{\nabla f}}(0,0) = \begin{bmatrix} 0 & 9 \\ 0 & 0 \end{bmatrix}.$$

i.e. $\underline{\underline{\nabla f}}$ is not invertible at

the roots, so Newton's method

converges with order 1.

Bisection method:

Suppose we have $f(\xi)=0$, and $\xi \in [a,b]$. Also, we know $f(a)f(b)<0$.

Then, we can define an iteration to compute ξ numerically by progressively shrinking the interval $[a,b]$.

For a_k, b_k, define $c_k = \frac{1}{2}(a_k+b_k)$.

Define a new interval

$$(a_{k+1}, b_{k+1}) = \begin{cases} (a_k, c_k) & f(c_k)f(b_k)>0 \\ (c_k, b_k) & f(c_k)f(b_k)<0. \end{cases}$$

Gaussian Elimination

History: Introduced by Gauss in 1809 to solve six linear equations with six unknowns. This was a model for the motion of the asteroid Pallas.

<u>Example:</u>

$$x_1 + x_2 + x_3 = 6$$

$$2x_1 + 4x_2 + 2x_3 = 16$$

$$-x_1 + 5x_2 - 4x_3 = -3.$$

We can rewrite this in matrix form

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix}.$$

~~Multiply 3rd row by 2:~~

$$\cancel{\begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -2 & 10 & -8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 6 \end{pmatrix}}$$

Multiply 1st row by -2:

$$\begin{pmatrix} -2 & -2 & -2 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -12 \\ 16 \\ -3 \end{pmatrix}$$

Add 1st and second rows, and put result in the second row

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ -3 \end{pmatrix}$$

Note that above we went ahead and also multiplied the 1st row by $-1/2$.

Now add the 1st row to the 3rd row, and put result in 3rd row:

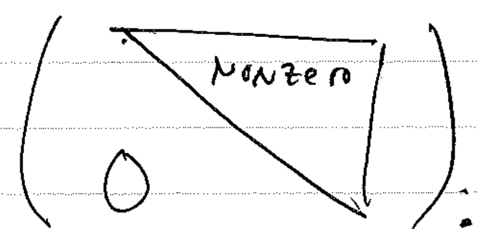
$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 6 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 3 \end{pmatrix}.$$

Finally, take the second row, multiply by -3, add to the third row, and put the result in the 3^{rd} row:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ -9 \end{pmatrix}.$$

To solve, we can now use "back substitution."

First, notice the structure of the resulting matrix: it is called "upper triangular"



Let's go back and look at the first operation we did: Multiply First row by -2 and add to second row: This can be expessed in terms of a matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

i.e. $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}$

$$= \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ -1 & 5 & -4 \end{pmatrix}.$$

Notice that $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is

"lower triangular."

We can write this sequence of row operations as a product of lower Δ matrices, $L_1, L_2, L_3, \ldots$

$$L_1 L_2 L_3 A x = L_1 L_2 L_3 b$$

~o

where we showed in our example that

$$L_1 L_2 L_3 A$$

is an upper triangular matrix.

We can move the $L_1 L_2 L_3$ multiplying b to the left hand side by multiplying by their inverses:

$$L_3^{-1} L_2^{-1} L_1^{-1} L_1 L_2 L_3 A x = b$$

$$\iff (L_1 L_2 L_3)^{-1} (L_1 L_2 L_3 A) x = b.$$

What is the structure of $(L_1 L_2 L_3)^{-1}$?

<u>Exercise</u>: Show $(L_1 L_2 L_3)^{-1}$ is lower triangular.

We've factored $A = (L_1 L_2 L_3)^{-1} (L_1 L_2 L_3 A)$

Into the product of a lower Δ and upper Δ matrix in the process of solving the linear system. This is called the LU Factorization of A.