lec 26

outline:

- quiz discussion
- nonlin equations
- optimization?
- ICES

Q's: don't know errors - how do we compute them? (not the point!)

but: important question: when do we stop? (when the guess stops changing)

can Newton's method fail? what did I mean by non-conv.

$f(x) = x^2 - 3$

$\tilde{f}(x+h) = f(x) + f'(x)\, h$

$\hat{f}(1+h) = f(1) + f'(1)\, \cancel{h}\ (\tilde{x}-1)$

$\tilde{x} = 1 + h$

$h = \hat{x} - 1$

$\tilde{f}(\tilde{x}) = f(1) + f'(1) \cdot (\tilde{x} - 1)$

$= (-2) + 2 \cdot (\hat{x} - 1)$

$= (-2) + 2\tilde{x} - 2 = 2\hat{x} - 4$

Newton step: $x - \dfrac{f(x)}{f'(x)}$

$f(x) = x^2 - 3 \quad f'(x) = 2x$

$2 - \dfrac{f(2)}{f'(2)} =$

$f(x) = x\exp(x) \quad f'(x) = xe^x + e^x$

$x - \dfrac{e^x}{xe^x + e^x} = x - \dfrac{1}{x+1}$

$= 5 - \dfrac{1}{5+1} = 5 - \dfrac{1}{6}$

$\dfrac{30}{6} - \dfrac{5}{6}$

11

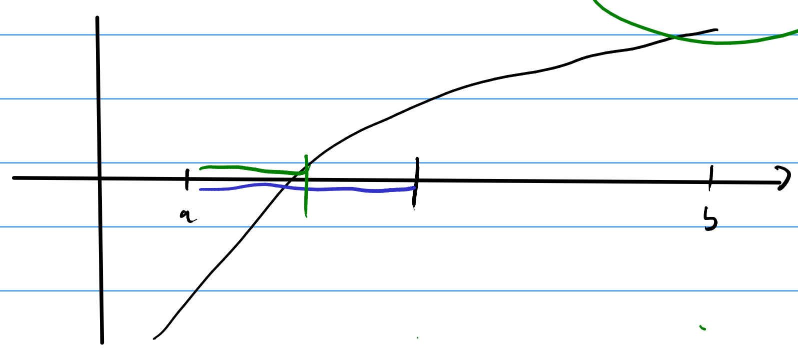
Solving nonlinear equations

Have: $f: \mathbb{R} \to \mathbb{R}$ function

Want: x such that $f(x) = y$

Rewrite the problem so that we only need $g(x) = 0$. (i.e. no explicit right-hand side)

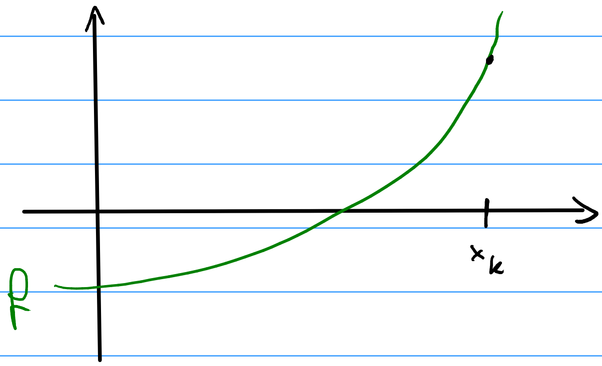
What if we know that f is continuous and $f(a) \cdot f(b) < 0$?



Can we use this "bracket" to track down the zero?

Newton's method

Suppose x_k is our current guess of the zero.



Idea: Build a solvable approximate version of f using $f(x_k)$ and $f'(x_k)$

Find the zero of the approximate version.

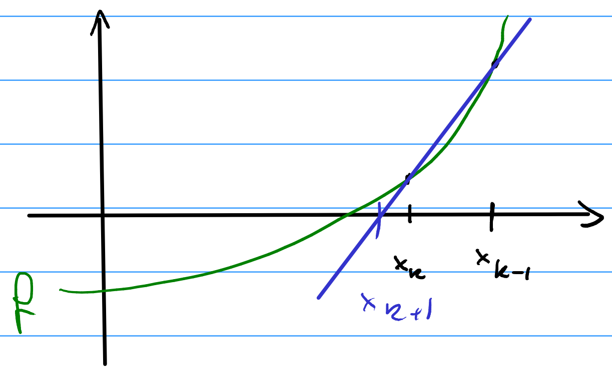
This is called Newton's method.

Name some downsides of Newton's method.

- doesn't always work
- need a derivative

Secant Method

How else could we find a line approximating a function?



Estimate the slope of the approximating line:

$$\frac{\text{Rise}}{\text{Run}} \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

Now use this estimate in Newton's method:

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

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

Solving systems of nonlinear equations

Want to solve $\vec{f}(\vec{x}) = \vec{0}$. $\quad f: \mathbb{R}^n \to \mathbb{R}^n$ why? !

Let's try to carry over our 1-dimensional ideas.

Let's first get an idea of what behavior can occur.

Based on the demo: Does bisection stand a chance?

Let's try Newton's method then. What's the linear approximation of f ?

$$1D: \quad \tilde{f}(x+h) = f(x) + f'(x)\cdot h$$

$$nD: \quad \tilde{\vec{f}}(\vec{x}+\vec{h}) = \vec{f}(\vec{x}) + J_f(\vec{x})\cdot\vec{h}$$

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & & \frac{\partial f_1}{\partial x_n} \\ & & \\ \frac{\partial f_n}{\partial x_1} & & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

OK, now solve that for h.

$$\vec{0} = \vec{f}(\vec{x}+\vec{h}) \approx \vec{f}(\vec{x}) + J_f(\vec{x})\cdot\vec{h}$$

linear system $\rightarrow$ solve for h $\quad\rightarrow\quad -\vec{f}(\vec{x}) = J_f(\vec{x})\cdot\vec{h} \qquad -\frac{f}{f'}$

$\rightarrow\ x_{k+1} = x_k + h = x_k - J_f^{-1}(x_k)\cdot\vec{f}(\vec{x}_k)$

Let's do an example of that:

$$f(x,y) = \begin{pmatrix} x + 2y - 2 \\ x^2 + 4y^2 - 4 \end{pmatrix} \begin{matrix} = f_1 \\ = f_2 \end{matrix}$$

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2x & 8y \end{pmatrix}$$

What are the downsides of this method?

- locally
- need deriv.

So how about (an n-dimensional analog of) the secant method?

$$\begin{matrix} f(x_n) & f(x_{k-1}) & & J_f(x) \\ n & n & & n^2 \end{matrix}$$

Broyden's method

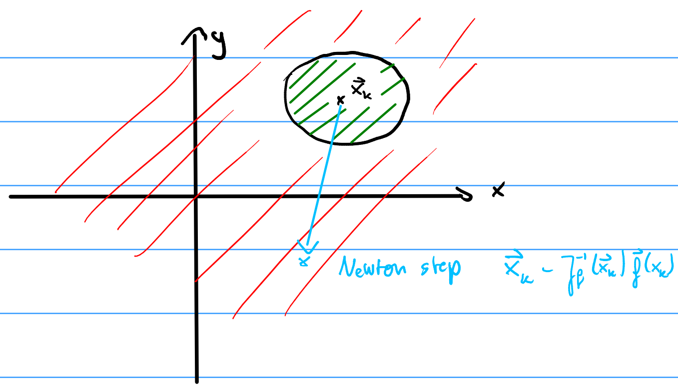
So carrying over the secant method to n dimensions is not easy.

It's possible, but beyond the scope of our class.

Here are two starting points to search:

- Broyden's method
- Secant updating methods

Here's one more idea: If we could figure out where the linear approximation in Newton is 'trustworthy', would that buy us anything?



12

Optimization

Let's try to weaken the requirement $f(\vec{x}) = \vec{0}$. $(f: \mathbb{R}^n \to \mathbb{R}^n)$

Create a problem statement for "optimization".

What if I'm interested in the largest possible value of a function g instead?

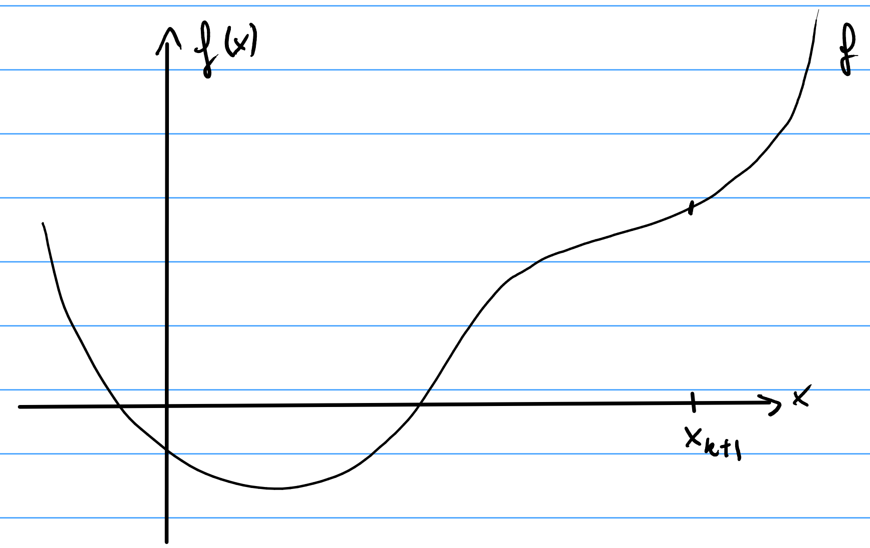
What could go wrong?

How can we tell if we've got a (local) minimum in 1D? Remember calculus!

And in n dimensions?

Let's steal the idea from Newton's method for equation solving.

Build a simple version of f and minimize that. Let's try in 1D first.



Does a linear approximation (a line) help at all?

$\tilde{f}(x+h) =$

Now minimize that.

Does that look at all familiar?