9/26/23

Last time

Solving eigenvalue problem $A\vec{v} = \lambda \vec{v}$
for largest eigenvalue using 'power' method

We iterate

$$\vec{z}^{(k)} = A^k \vec{z}^{(0)}$$

starting with a guess that needs to have
a component along $\vec{V}_0$

Rayleigh quotient $\mu(\vec{x}) = \dfrac{\vec{x}^T A \vec{x}}{\vec{x}^T \vec{x}}$

and normalize $\vec{q}^{(k)} = \dfrac{\vec{z}^{(k)}}{\|\vec{z}^{(k)}\|}$

$$\vec{z}^{(k)} = A \vec{q}^{(k-1)}$$

$$\vec{q}^{(k)} = \vec{z}^{(k)} / \|\vec{z}^{(k)}\|$$

$$\mu(\vec{q}^{(k)}) = [\vec{q}^{(k)}]^T A \vec{q}^{(k)}$$

Eventually $\mu(\vec{q}^{(k)}) \to \lambda_0$ $\vec{q}^{(k)} \to \vec{V}_0$

Zeros and Minima (Chapter 5)

We often need to solve equations of the form $f(x) = 0$ where $f(x)$ is non-linear

Van der Waals equation of state $\left(P + a/v^2\right)(v - b) = RT \quad v = \frac{V}{\text{mole}}$

QM particle in finite well of width $2a$, height V_0 $\quad K \tan(Ka) = \left[\frac{2mV_0}{\hbar^2} - K^2\right]^{1/2}$

Or worse: $\vec{f}(\vec{x}) = \vec{0}$

$$f_0(x_0, x_1, \dots x_{n-1}) = 0$$
$$f_1(x_0, x_1, \dots x_{n-1}) = 0$$
$$\vdots$$
$$f_{n-1}(x_0, x_1, \dots x_{n-1}) = 0$$

n non-linear equations w/ n unknowns

Also we want to solve or find minima of scalar ϕ

$$\min \phi(\vec{x}) = \min \phi(x_0, x_1, \dots x_{n-1})$$

Our test problem $f(x) = e^{x - \sqrt{x}} - x = 0$

[See plot]

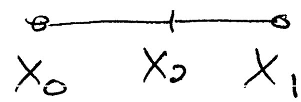
We observe roots around 1, 2.5

(1 is clearly exact)

Bisection Method

Assume you have 'bracketed' a root
$\Rightarrow$ found an X_0 and X_1 such that
$f(X_0)$ and $f(X_1)$ have opposite signs
So the root must exist in interval (X_0, X_1)

Evaluate midpoint $X_2 = \frac{X_0 + X_1}{2}$



Evaluate $f(X_0)\,f(X_2)$

If $f(X_0)\,f(X_2) < 0$ root is in (X_0, X_2)
repeat with X_2 in place of X_1

If $f(X_0)\,f(X_2) > 0$ root is in (X_2, X_1)
repeat with X_2 in place of X_0

If we call $X_0 = a$ and $X_1 = b$ we generate
a series of iterations $X^{(0)} = X^0,\ X^{(1)},\ X^{(2)}, \ldots$
where each new iteration halves the interval
So if X^* is true root

$$|X^{(k)} - X^*| \le \frac{1}{2} |X^{(k-1)} - X^*|$$ (See Eqn 5.16)

our iterations are linearly convergent

We choose to terminate iterations when fractional change is less than tolerance

$$\frac{|x^{(k)} - x^{(k-1)}|}{|x^{(k)}|} \leq \epsilon \sim \text{small } \#$$

[See bisection.py]

Doesn't generalize well to higher dimensional problems

<u>Newton's Method</u>

Need $f(x)$ and $f'(x)$

Let $x^{(k-1)}$ be our last iteration (or first guess)

Taylor expand $f(x)$ around this point

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

where ξ is a point between x and $x^{(k-1)}$

Take $x = x^*$ so $f(x^*) = 0$. If $f(x)$ is linear (it isn't, but if it was)

$$f(x^*) = 0 = f(x^{(k-1)}) + (x^* - x^{(k-1)}) f'(x^{(k-1)})$$

$$\Rightarrow \quad x^* = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})}$$

So when $f(x)$ is not linear we will iterate w/

$$x^{(k)} = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})} \qquad k = 1, 2, \ldots$$

Can show from our earlier equation with $x = x^*$ and keeping $f''(\xi)$ term that (lots of algebra!)

$$x^{(k)} - x^* = C\,(x^{(k-1)} - x^*)^2$$

where C depends on $f''(\xi)$ and $f'(x^{(k-1)})$

Method is quadratically convergent

$\underbrace{\text{Geometric interpretation}}$ and trouble on next pages

Approx $f(x)$ with tangent at $x^{(k-1)}, f(x^{(k-1)})$

get $x^{(k)}$ by point where tangent crosses x axis

5-4a

sional problems in a reasonably straightforward manner. At a big-picture level, New method requires more input than the approaches we saw earlier: in addition to being evaluate the function $f(x)$, one must also be able to evaluate its first derivative $f'(x)$ s obviously trivial for our example above, where $f(x)$ is analytically known, but may e so easy to access for the case where $f(x)$ is an externally provided (costly) rou Furthermore, to give the conclusion ahead of time: there are many situations in which on's method can get in trouble, so it always pays to think about your specific prob nstead of blindly trusting a canned routine. Even so, if you already have a reasonable ate of where the root may lie, Newton's method is usually a fast and reliable solution

5.2.5.1 Algorithm and Interpretation

ill assume that $f(x)$ has continuous first and second derivatives. Also, take $x^{(k-1)}$ to e last iterate we've produced (or just an initial guess). Similarly to what we did in 5.26) above, we will now write down a Taylor expansion of $f(x)$ around $x^{(k-1)}$; this we go up to one order higher:

$$f(x) = f(x^{(k-1)}) + \left(x - x^{(k-1)}\right) f'(x^{(k-1)}) + \frac{1}{2}\left(x - x^{(k-1)}\right)^2 f''(\xi) \quad (5.37)$$

e ξ is a point between x and $x^{(k-1)}$. If we now take $x = x^*$ then we have $f(x^*) = 0$. If rther assume that $f(x)$ is linear (in which case $f''(\xi) = 0$), we get:

$$0 = f(x^{(k-1)}) + \left(x^* - x^{(k-1)}\right) f'(x^{(k-1)}) \quad (5.38)$$

n can be re-arranged to give:

$$x^* = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})} \quad (5.39)$$

ords: for a linear function, an initial guess can be combined with the values of th ion and the first derivative (at that initial guess) to locate the root.

is motivates Newton's method: if $f(x)$ is non-linear, we still use the same formula 5.39), this time in order to evaluate not the root but our next iterate (which, we hope, s us closer to the root):

$$x^{(k)} = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})}, \quad k = 1, 2, \ldots \quad (5.40)$$

e did in the previous paragraph, we are here neglecting the second derivative term e Taylor expansion. However if we are, indeed, converging, then $(x^{(k)} - x^{(k-1)})^2$ in 5.37) will actually be smaller than $x^{(k)} - x^{(k-1)}$, so all will be well.

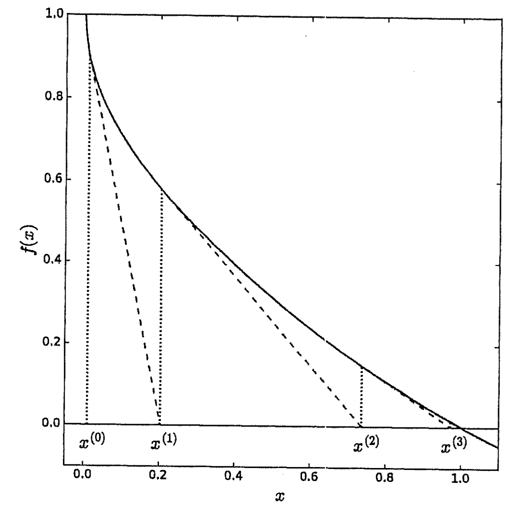


Illustration of Newton's method for our example function

call $x^{(1)}$ the point where that tangent intercepts the x axis and repeat. For our example, this process brings us very close to the root in just a few steps.

5.2.5.2 Convergence Properties

We now turn to the convergence properties of Newton's method. To orient the reader, what we'll try to do is to relate $x^{(k)} - x^*$ to $x^{(k-1)} - x^*$, as per Eq. (5.16). We will employ our earlier Taylor expansion, Eq. (5.37), and take $x = x^*$, but this time without assuming that the second derivative vanishes. Furthermore, we assume that we are dealing with a simple root x^*, for which we therefore have $f'(x^*) \neq 0$; that means we can also assume $f'(x) \neq 0$ in the vicinity of the root. We have:

$$0 = f(x^{(k-1)}) + \left(x^* - x^{(k-1)}\right) f'(x^{(k-1)}) + \frac{1}{2}\left(x^* - x^{(k-1)}\right)^2 f''(\xi) \quad (5.41)$$

where the left-hand side is the result of $f(x^*) = 0$.

Dividing by $f'(x^{(k-1)})$ and re-arranging, we find:

$$-\frac{f(x^{(k-1)})}{f'(x^{(k-1)})} - x^* + x^{(k-1)} = \frac{\left(x^* - x^{(k-1)}\right)^2 f''(\xi)}{2f'(x^{(k-1)})} \quad (5.42)$$

The first and third terms on the left-hand side can be combined together to give $x^{(k)}$, as per Newton's prescription in Eq. (5.40). This leads to:

5-4b

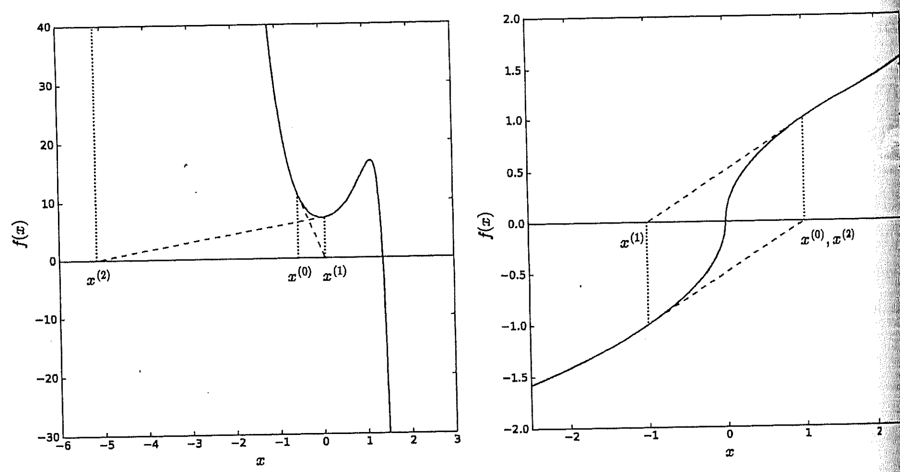


Fig. 5.8 Cases where Newton's method struggles

sides, our result is identical to that in Eq. (5.16), under the assumption that:

$$\frac{f''(\xi)}{2f'(x^{(k-1)})} \leq m \tag{5.44}$$

Since the right-hand side of Eq. (5.43) contains a square, we find that, using the notation of Eq. (5.16), $p = 2$, so if $m < 1$ holds then *Newton's method is quadratically convergent,* sufficiently close to the root.[18] This explains why our iterates approached the root so rapidly in Fig. 5.7, even though we intentionally picked a poor starting point.

5.2.5.3 Multiple Roots and Other Issues

Of course, as already hinted at above, there are situations where Newton's method can misbehave. As it so happens, our starting point in Fig. 5.7 was *near* 0, but not actually 0. For our example function $f(x) = e^{x-\sqrt{x}} - x$, the first derivative has a $\sqrt{x}$ in the denominator, so picking $x^{(0)} = 0$ would have gotten us in trouble. This can easily be avoided by picking another starting point. There are other problems that relate not to our initial guess, but to the behavior of $f(x)$ itself.

An example is given in the left panel of Fig. 5.8. Our initial guess $x^{(0)}$ is perfectly normal and does not suffer from discontinuities, or other problems. However, by taking the tangent and finding the intercept with the x axis, our $x^{(1)}$ happens to be near a local extremum (minimum in this case); since $f'(x^{(k-1)})$ appears in the denominator in Newton's prescription in Eq. (5.40), a small derivative leads to a large step, considerably away from the root. Note that, for this misbehavior to arise, our previous iterate doesn't even need to be "at" the extremum, only in its neighborhood. It's worth observing that, for this case, the root $x^* \approx 1.4$ actually *is* found by Newton's method after a few dozen iterations: after slowly

Things are even worse in the case of the right panel of Fig. 5.8, enters an infinite cycle. Here we start with a positive $x^{(0)}$, then the t negative iterate (which happens to be the exact opposite of where we and then taking the tangent and finding the intercept brings us back $x^{(2)} = x^{(0)}$. After that, the entire cycle keeps repeating, without ever the root which we can see is at $x^* = 0$. Of course, this example is som real-world application, the function would likely not be perfectly sym

Another problem arises when we are faced with a root that is *not* recall, in our study of the convergence properties for Newton's metho that $f'(x^*) \neq 0$. This is not the case when you are faced with a mul problem not only for the convergence study but also for the prescrip contains the derivative in the denominator, so $f'(x^*) = 0$ can obvious

As an example, in Fig. 5.9 we are plotting the function:

$$f(x) = x^4 - 9x^3 + 25x^2 - 24x + 4$$

which has one root at $x^* \approx 0.21$, another one at $x^* \approx 4.79$, as we $x^* = 2$. This is easier to see if we rewrite our function as:

$$f(x) = (x-2)^2\,(x^2 - 5x + 1)$$

As it turns out, Newton's method can find the double root, but it ta you'd expect to do so. Intriguingly, this is an example where the bise bracketing methods) cannot help us: the function does not change s (i.e., the root is not bracketed). As you will find out when you solve p method for this case converges linearly! If you would like to have it c (as is usually the case), you need to implement Newton's prescription form, namely:

$$x^{(k)} = x^{(k-1)} - 2\frac{f(x^{(k-1)})}{f'(x^{(k-1)})}, \quad k = 1, 2, \ldots$$

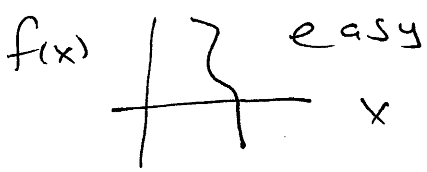
as opposed to Eq. (5.40). Problem 5.9 discusses the source of this m

Unfortunately, in practice we don't know ahead of time that we are root (or a triple root, etc.), so we cannot pick such a prefactor that convergence. As you will discover in problem 5.10, a useful trick (wh of your knowledge of the multiplicity of the root) is to apply New this time not on the function $f(x)$ but on the function:

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

As you will show in that problem, $w(x)$ has a simple root at x^*, regard of the root of $f(x)$ at x^*. Of course, there's no free lunch, and yo cancellations (which may cause headaches) when evaluating the d

Note: We are trying to find where $f(x)=0$. Obviously easier if crossing of y-axis occurs when f is 'steep'



$$x_{est} - x_{true} \approx \underbrace{\frac{1}{f'(x_{true})}}\left[f(x_{est}) - f(x_{true}) \right]$$

$K = 1/f'(x)$ condition number

when $f'(x)$ large, K is small and we should do well

$f'(x)$ small $\rightarrow K$ large $\rightarrow$ hard problem

Also double roots can be problematic (problem 5.10)

Secant Method: modification to Newton algorithm

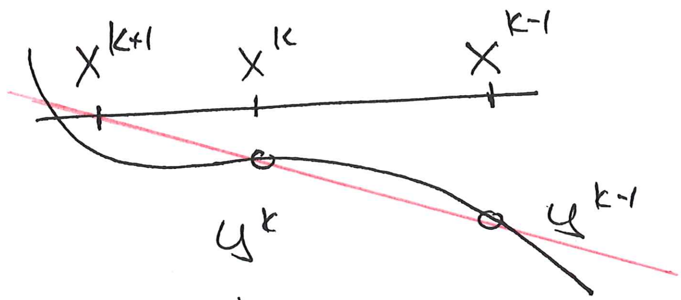
estimate $f'(x)$ a la finite diff

$$x^{(k)} = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})} \qquad k = 1, 2, \ldots$$

$$f'(x^{(k-1)}) \approx \frac{f(x^{(k-1)}) - f(x^{(k-2)})}{x^{(k-1)} - x^{(k-2)}} \qquad k = 2, 3, \ldots$$

Like a finite diff but spacing is not const.

Geometric interpretation



Draw straight line between

x^k, y^k and x^{k-1}, y^{k-1} and find x-int.

For points (x_0, y_0) (x_1, y_1)

$$y - y_0 = \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$$

Set $y = 0$

solve for $x \equiv x^{k+1}$

with $x_0 = x^{k-1}$ $x_1 = x^k$

$y_0 = y^{k-1}$ $y_1 = y^k$

[See secant.py]

Other methods available

Ridders: bracketing; superlinear convergence

Fixed point: non-bracketing; simple; may not converge

Look at text examples for details

Systems of non-linear equation (5.4)

n simultaneous non-linear equations w/ n unknowns

$$\underbrace{\vec{f}}_{n \text{ functions}}(\underbrace{\vec{x}}_{n \text{ components}}) = \vec{0}$$

ex: $n=2$

$$f_0(x_0, x_1) = x_0^2 - 2x_0 + x_1^4 - 2x_1^2 + x_1 = 0$$

$$f_1(x_0, x_1) = x_0^2 + x_0 + 2x_1^3 - 2x_1^2 - 3/2\, x_1 - .05 = 0$$

unknowns are x_0, x_1

we will generalize Newton's method

iterate over vectors $\vec{x}^{(k)}$

First consider a single function component f_i

and we taylor expand around iterate $\vec{x}^{(k-1)}$

$$f_i(\vec{x}) = f_i(\vec{x}^{(k-1)}) + (\nabla f_i(\vec{x}^{(k-1)}))^T (\vec{x} - \vec{x}^{(k-1)}) + O(\|\vec{x} - \vec{x}^{(k-1)}\|^2) \quad i = 0, 1, \ldots n-1$$

$$\nabla f_i(\vec{x}) = \begin{bmatrix} \partial f_i / \partial x_0 \\ \partial f_i / \partial x_1 \\ \vdots \\ \partial f_i / \partial x_{n-1} \end{bmatrix}$$

evaluated at $\vec{x}$ in this case

$$\left[\nabla f_i(\vec{x}^{(k-1)})\right]^T (\vec{x} - \vec{x}^{(k-1)}) = \sum_{j=0}^{n-1} \frac{\partial f_i}{\partial x_j}\bigg|_{x_j^{(k-1)}} (x_j - x_j^{(k-1)})$$

$$= \frac{\partial f_i}{\partial x_0}\bigg|_{x_0^{(k-1)}} (x_0 - x_0^{(k-1)}) + \frac{\partial f_i}{\partial x_1}\bigg|_{x_1^{(k-1)}} (x_1 - x_1^{(k-1)})$$

$$+ \dots \frac{\partial f_i}{\partial x_{n-1}}\bigg|_{x_{n-1}^{(k-1)}} (x_{n-1} - x_{n-1}^{(k-1)})$$

For all components we use Jacobian matrix $J(\vec{x})$

$$\vec{J}(x) = \Big[\underbrace{\nabla f_0(\vec{x}) \ \nabla f_1(\vec{x}) \dots \nabla f_{n-1}(\vec{x})}_{\text{column vectors}}\Big]^T$$

$$J_{ij} = \partial f_i / \partial x_j$$

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}^{(k-1)}) + J(\vec{x}^{(k-1)})(\vec{x} - \vec{x}^{(k-1)}) + \underset{\text{shorthand}}{O(\vec{x}^2)}$$

Assume we know solution $\vec{f}(\vec{x}^*) = \vec{0}$

$$\vec{0} = \vec{f}(\vec{x}^{(k-1)}) + J(\vec{x}^{(k-1)})(\vec{x}^* - \vec{x}^{(k-1)})$$

iterative procedure

$$\underbrace{J(\vec{x}^{(k-1)})}(\vec{x}^{(k)} - \vec{x}^{(k-1)}) = \underbrace{-\vec{f}(\vec{x}^{(k-1)})}_{\text{we can evaluate these}}$$

$$\Rightarrow \quad \vec{A}\,\vec{x} = \vec{b}$$

we can evaluate these
$\Rightarrow$ just #ers