Fixed-Point Iteration

Thursday, September 3, 2020 9:39 AM

Another variant on root-finding... this one does extend to higher dimension but we'll stick to scalars.

Outline:
- Definition
- Graphical Interpretation
- Theory, part 1 (existence, non-constructive)
- Theory, part 2 (constructive)
- Convergence Rate
- Pictures
- Recap
- Examples

Def A *fixed-point** of a function g is a point p such that

$$g(p) = p$$

* unrelated to "fixed point"/"floating point" representations of number on a computer.

Mathematically, equivalent to root-finding (i.e., p is a root of $f(x) := g(x) - x$)

Most examples can just as naturally be cast as finding zeros or minimization/maximization ... so why discuss?

... leads to very natural algorithm, "fixed-point iteration"

$$p_{n+1} = g(p_n).$$ (no derivatives of g needed)

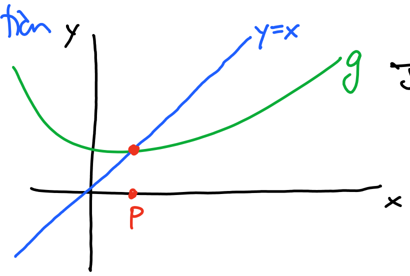
If (p_n) converges, $p_n \to p$, then (still assuming g is continuous) this limit p is a fixed point

(proof: $p = \lim_{n\to\infty} p_{n+1} = \lim_{n\to\infty} g(p_n) \overset{\text{used continuity of } g}{=} g(\lim_{n\to\infty} p_n) = g(p)$)

Graphical Interpretation

Solve $g(p) = p$

or $g(x) = x$

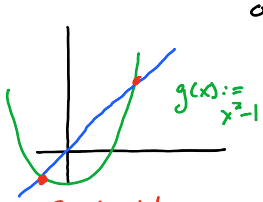


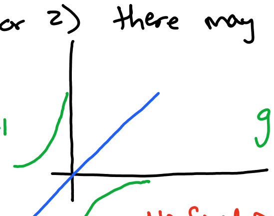
Just find the intersection of $y = f(x)$ and $y = x$

It's immediately clear 1) there may not be any fixed points

or 2) there may be several (or infinite)

ex: define $g(x) := x$





Several Fixed points   No fixed points

Theory, part 1: existence (non-constructive)

Main tool: Intermediate Value Thm (IVT), just like for the bisection method

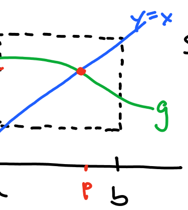
Theorem (Thm 2.3 (i) in Burden + Faires 9th ed.)

Let $g \in C([a,b])$, then if all its output is in $[a,b]$ as well, (i.e. $\forall x \in [a,b],\ g(x) \in [a,b]$)

then g has a fixed point in $[a,b]$ (... may not be unique).

proof

Define $h(x) = g(x) - x$ and apply IVT to show $\exists\, x \in [a,b]$ s.t. $h(x) = 0$.



Specifically, $h(a) = g(a) - a$.

Either $g(a) = a$ (and we're done)

or $g(a) > a$ (since range is in $[a,b]$) $\Rightarrow$ $\underline{h(a) > 0}$

Similarly, $g(b) = b$ (and we're done) or $\underline{h(b) < 0}$

So $h(a)\cdot h(b) < 0$, IVT applies, and $\exists\, x \in [a,b]$ w/ $h(x) = 0$

i.e. $g(x) = x$. ∎

Theory, part 2: uniqueness, and a constructive way to find fixed pts.

Definition A function g is called Lipschitz with constant L on an interval $A \subseteq \mathbb{R}$ if $(\forall x, y \in A)\ \ |g(x) - g(y)| \le L \cdot |x - y|$ $\quad (*)$

(In particular, such a function is uniformly continuous, and hence continuous. hence, sometimes we say "Lipschitz continuous")

"The" Lipschitz constant L of a Lipschitz continuous function is the smallest such L that satisfies $(*)$.

Note Suppose $g \in C^1(A)$, i.e., g' exists.

If $|g'(x)| \le L$ $(\forall x \in A)$ then g is L-Lipschitz

(proof: F.T.C., $|g(x) - g(y)| = \left| \int_y^x g'(s)\,ds \right|$

$\le \int_y^x |g'(s)|\,ds \le L \cdot |x-y|$)

So often we just show $|g'| \le L$. However, you can have a Lipschitz function that isn't differentiable (ex: $g(x) = |x|$ is $L=1$ Lipschitz)

Definition A function g is a contraction on $A \subseteq \mathbb{R}$ if it is Lipschitz with $L < 1$

⚠ Contraction means $\forall x, y\ \ |g(x) - g(y)| \le L \cdot |x - y|$ for some $L < 1$

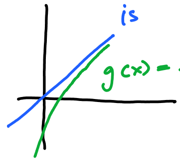
this is not quite the same as $\forall x, y\ \ |g(x) - g(y)| < |x - y|$

Ex: $g(x) = e^{-x} + x$ so $g'(x) = -e^{-x} + 1$

on $A=[0,100]$, g is a contraction since $|g'(x)| < 1-e^{-100}$ on A

on $A=[0,\infty)$, g is not a contraction since $\lim_{x\to\infty} |g'(x)| = 1$

(in neither case is there a fixed point: the issue is $g(x) \notin A$ for $x \in A$)



$g(x) = -e^{-x}+1$ (asymptotes to $y=x$ but doesn't cross)

Theorem (Banach Fixed point theorem aka Contraction Mapping Theorem)

Let g be a contraction on $[a,b]$ and $(\forall x \in [a,b])\ g(x) \in [a,b]$
$\qquad\qquad$ ① $\qquad\qquad\qquad\qquad\qquad\qquad$ ②

then

① there is a unique fixed point of g inside $[a,b]$, and
(existence and uniqueness) $\qquad$ call the limit p

② defining the fixed point iteration by $p_{n+1} = g(p_n)$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad p_0 \in [a,b]$ arbitrary

then

$\lim_{n\to\infty} p_n = p$ and at rate $|p_n - p| \le L^n \cdot |p_0 - p|$ $\qquad$ Recall this is "linear convergence"

(construction... ie., how to find it)

⚠ The theorem gives conditions that guarantee existence and uniqueness, but they are not necessary (in particular, $[a,b]$ might be too big)

Proof First uniqueness, then existence + construction
$\qquad\qquad$ part (I) $\qquad\qquad\qquad\qquad$ part (II)

(I) Suppose there is a fixed point $p \in [a,b]$ with $g(p) = p$, and let $q \in [a,b]$ with $g(q) = q$ also. Then

$$|p-q| \overset{\text{def. fixed pt.}}{=} |g(p)-g(q)| \le L \cdot |p-q| \quad \text{for } L<1$$

so if $|p-q| \ne 0$, divide equation by it, to get $1 \le L$. Contradiction

Hence $|p-q| = 0$, ie $p = q$, meaning there cannot be distinct fixed pts.

(II) To show existence, we show (p_n) is a Cauchy sequence. This isn't in the scope of our course, so don't worry about it. (or use the IVT)

To show construction, ie. $|p_n - p| \le L^n \cdot |p_0 - p|$ and rate

just need to calculate: $|p_n - p| = |g(p_{n-1}) - g(p)|$ $\qquad$ by def'n of fixed pt.

$$\le L \cdot |p_{n-1} - p| \quad \text{since it is a contraction } (L<1)$$

$$\le L^n |p_0 - p|.$$

$0<L<1 \Rightarrow L^n \to 0$

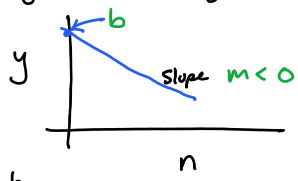
so $p_n \to p$ $\qquad$ □

Convergence rate

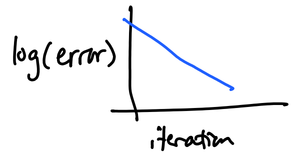
$|p_n - p| \le L^n \cdot |p_0 - p|$, $L<1$ is linear convergence

e_n , e for "error"

$$e_n \le L^n \cdot e_0 \quad \text{i.e.,} \quad \underbrace{\log(e_n)}_{y} \le n \cdot \underbrace{\log(L)}_{m} + \underbrace{\log(e_0)}_{b}$$



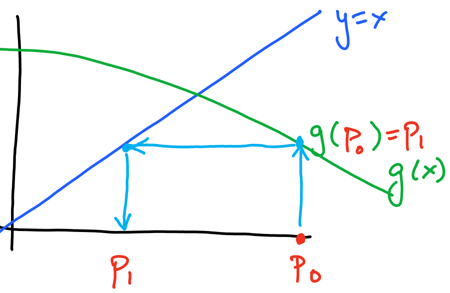
i.e.,

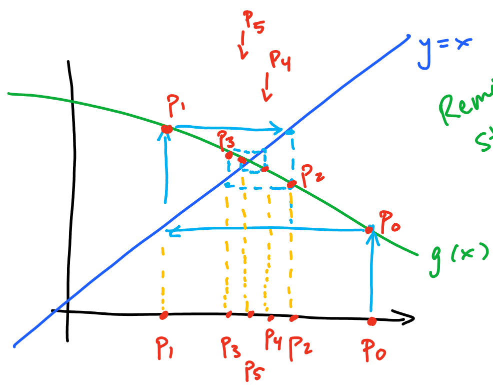


You can estimate slopes by running a least-squares fit to $\log(e_n)$, eg polyfit in Matlab

(in practice, often exclude some of the small n points if visually they don't appear to be following the trend yet)

Pictures of fixed-pt. iterations

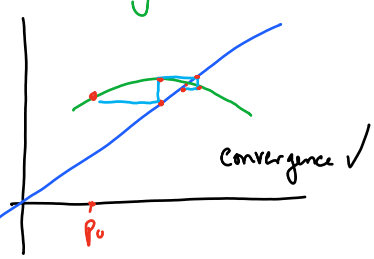


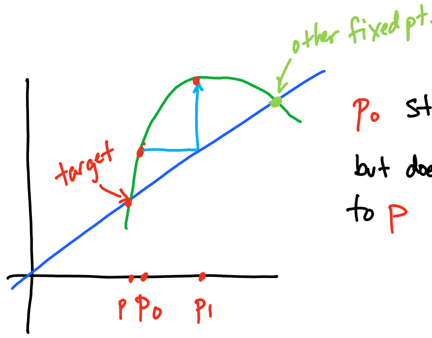


Remind you of stable equilibrium you learned about in APPM 2360?

⚠ Things don't always work!

$g(x) = x^2 - 4x + 3.5$





p_0 starts near p but doesn't converge to p

Recap,

$|g'(x)| \le L < 1$ on $[a,b]$, and $g(x) \in [a,b]$ $\forall x \in [a,b]$,

is sufficient for guaranteeing fixed-pt. iterations will find the unique fixed point on $[a,b]$.

(linear convergence)

Error converges to 0 at least as fast as $|e_n| \le \text{const} \cdot L^n$

As $p_n \to p$, we can shrink $[a,b]$, and find a tighter bound on $|g'(x)|$

i.e., $|g'(p)|$ will determine final convergence rate

If $g'(p) = 0$, we get superlinear convergence!

Examples

① Show $g(x) = (x^2-1)$ has a unique fixed pt. on $[-1,1]$

g ... $\frac{\quad}{3}$ has a single fixed pt on $[-1,1]$

First, check that $g(x) \in [-1,1]$ for $x \in [-1,1]$

(why do we do this step? If $g(x)$ is outside the domain, how can we have $\underset{\text{out of domain}}{x} = \underset{\text{in domain}}{g(x)}$?)

So, what is $\max_{x\in[-1,1]} g(x)$ and $\min_{x\in[-1,1]} g(x)$?

- check endpts, $g(-1) = 0$ and $g(1) = 0$
- check critical pts, ie., where $g'(x) = 0$. $g'(x) = \frac{2x}{3}$ so $g'(x) = 0 \Rightarrow x = 0$

so $g(0) = -1/3$

So $g(x) \in [-1/3, 0] \subseteq [-1,1]$ for $x \in [-1,1]$ ✓

Second, check if it's contractive. Since g' exists, just check $|g'(x)|$ on $x \in [-1,1]$. Since $g'(x) = 2/3\, x$

then $|g'(x)| \leq 2/3$, so it's a contraction ($2/3 < 1$).

② Show $g(x) = 3^{-x}$ has a fixed pt. on $[0,1]$

Note $g'(x) = -3^{-x} \cdot \ln(3)$, so $g'(x) < 0$, so g is decreasing.

So $\max_{x\in[0,1]} g(x) = g(0) = 1$ and $\min_{x\in[0,1]} g(x) = g(1) = 1/3$

So $g(x) \in [0,1]$ on $x \in [0,1]$. Since g is continuous, our first theorem guarantees there is a fixed pt.

Note we can't use the contraction mapping to prove uniqueness, since g isn't a contraction on $[0,1]$ since $g'(0) = -\ln(3) = -1.0986$, $|g'(0)| > 1$

③ Root-finding $\Rightarrow$ Fixed pt.

Solve for a root of $x^3 + 4x^2 - 10 = 0$ in $[1,2]$ ($p \approx 1.36$)

Contraction on $[1,2]$?

A) $\underbrace{x^3 + 4x^2 - 10 + x}_{g_1(x)} = x$ So solve $x = g_1(x)$

No, doesn't map $[1,2]$ to $[1,2]$ so may fail

B) $4x^2 = 10 - x^3$

$x = \pm \frac{1}{2}\sqrt{10 - x^3}$

So solve $x = \underbrace{\tfrac{1}{2}\sqrt{10 - x^3}}_{g_2(x)}$

Not on $[1,2]$, $g'(2) = 2.12 > 1$ though $[1, 1.5]$ works

C) $x^3 + 4x^2 - 10 = 0$ ($x = 0$ isn't a root)

$\frac{x^3}{x} = \frac{10 - 4x^2}{x}$, $x^2 = \frac{10}{x} - 4x$, So solve $x = \underbrace{\sqrt{10/x - 4x}}_{g_3(x)}$

No, doesn't map $[1,2]$ to $[1,2]$ and $g'(p) = 3.4 > 1$

D) $x^3 + 4x^2 = 10$

$x^2(x+4) = 10$, $x^2 = \frac{10}{x+4}$ So solve $x = \underbrace{\sqrt{\frac{10}{x+4}}}_{g_4(x)}$

Yes, $|g'(x)| < 0.15$ Fast convergence!

E) $-\frac{x^3 + 4x^2 - 10}{3x^2 + 8x} + x = x$, so solve $x = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}$

Yes, $|g'(x)| < .5$ and $g'(p) = 0$

nonzero at a root
(we know this since a simple root and denominator is f')

$g_S(x)$

$g(p)=0$

= Advanced, optional topic =

For fun... we discuss solving $x=g(x)$, $x\in\mathbb{R}$, a 1D vector space

This contraction-mapping idea extends to vector spaces $\mathbb{R}^n$ and even infinite-dimensional vector spaces! For example, you can think of a function f as a "point" or "vector" in a function space.

From the exam for 1st year Applied Math PhD students in Aug '19:

Let h be a continuous function on $[0,1]$. Show there exists a unique continuous function f on $[0,1]$ satisfying

$$f(x) = \underbrace{h(x) + \int_0^x f(x-t)\,e^{-t^2}\,dt}_{G(f)}$$

proof sketch

Rewrite as $f=G(f)$ and show G is a contraction, so apply Banach fixed pt. theorem.