Chapter 6

Brief introduction on how to find zeros of real functions

Solving explicitly the unconstrained problem

Remark. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a $\mathbb{C}^1$ function. We have proven that to solve the problem

$$\min_{x \in \mathbb{R}^n} f(x)$$

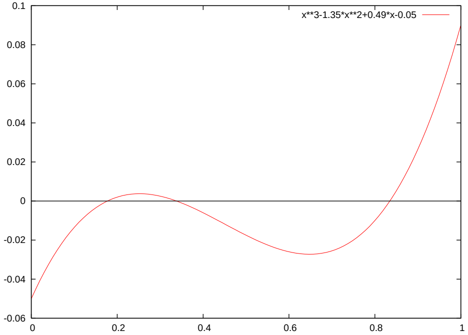
it is necessary to find out points (vectors) $x^\star$ such that $\nabla f(x^\star) = 0$.

Remark. A possible strategy for doing so is to start at a given vector $x_0 \in D$ and construct a sequence x_k such that

$$x_k = \min_{\alpha \in \mathbb{R}} f(x_{k-1} + \alpha p_k), \quad p_k \in \mathbb{R}^n.$$

Consequence. It is worthy to first do a quick overview for the one-dimensional problem of finding zeroes of functions.

Example



$$f(x) = x^3 - 1.35\ x^2 + 0.49\ x - 0.05$$

Strategy

Let $f : \mathcal{I} \subset \mathbb{R} \to \mathbb{R}$.

1. Location: Where the zeros are?
2. Separation (or uniqueness): Determine a domain (interval) with a unique zero.
3. Approximation (root-finding algorithms): Construct a sequence x_k in the domain above such that

$$x_k \to x^\star \quad \text{with} \quad f(x^\star) = 0.$$

Location: Bolzano's Theorem

Theorem (Bolzano). Let $f : [a, b] \to \mathbb{R}$ be a continuous function satisfying $f(a)f(b) < 0$. Then, there exists $x^\star \in (a, b)$ such that $f(x^\star) = 0$.

Proof.

(0) Let $[a_0, b_0] = [a, b]$ and let $n = 0$.

(1) Compute $c_{n+1} = (a_n + b_n)/2$.

(2) One (and only one) below can happen.

(2.a) $f(c_{n+1}) = 0$. Then $x^\star = c_{n+1}$.
(2.b) $f(a_n)f(c_{n+1}) < 0$. Then $[a_{n+1}, b_{n+1}] = [a_n, c_{n+1}]$;
(2.c) $f(c_{n+1})f(b_n) < 0$. Then $[a_{n+1}, b_{n+1}] = [c_{n+1}, b_n]$.

(3) Do $n = n + 1$ and move to step (1).

Location: Bolzano's Theorem

Proof. To conclude the theorem we argue as follows.

(i) If the process stops, i.e., (2.a), we have found $x^\star$.

(ii) Otherwise we have constructed an infinite sequence of nested intervals

$$[a, b] = [a_0, b_0] \supset [a_1, b_1] \supset \dots \supset [a_n, b_n] \supset \dots,$$

(iii) Clearly $\ell([a_n, b_n]) = \frac{|b_0 - a_0|}{2^n} \to 0$

(iv) Let

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = x^\star \ .$$

(v) Since $f(a_n)f(b_n) < 0$ for all $n \geq 0$, we have

$$0 \geq \lim_{n\to\infty} f(a_n)f(b_n) = (f(x^\star))^2 \qquad (\Rightarrow f(x^\star) = 0).$$

Uniqueness: Rolle's Theorem

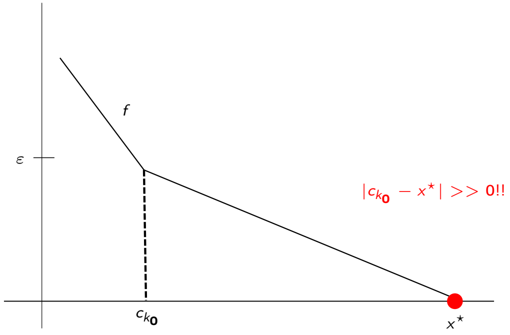
Theorem (Rolle). Let $f : [a, b] \to \mathbb{R}$ be a continuous function in $[a, b]$ and derivable in (a, b). Suppose $f(a) = f(b)$. Then there exists $\zeta^\star \in (a, b)$ such that $f'(\zeta^\star) = 0$.

Corollary. Let $f : [a, b] \to \mathbb{R}$ be a continuous function in $[a, b]$ and derivable in (a, b). Assume that $f(a)f(b) < 0$ and $f'(x) \neq 0$ for all $x \in (a, b)$. Then there exists a unique $x^\star \in (a, b)$ such that $f(x^\star) = 0$.

Proof. From Bolzano's Theorem it is clear that there exists $x^\star \in (a, b)$ such that $f(x^\star) = 0$. Assume there exists $y^\star \neq x^\star$ such that $f(y^\star) = 0$. W.l.o.g take $y^\star > x^\star$ Then $f|_{[x^\star, y^\star]}$ is in the hypothesis of Rolle's Theorem and there exists $\zeta^\star \in (x^\star, y^\star) \subset (a, b)$ such that $f'(\zeta^\star) = 0$, a contradiction.

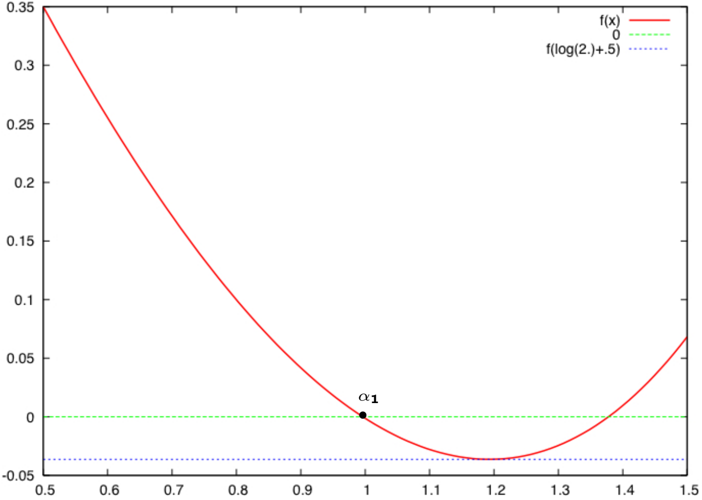
Root-finding algorithm: Bisection method

Assume there is a unique zero $x^\star$ of f in $[a, b]$. From Bolzano's Theorem we can construct a sequence $\{c_n\}_{n\geq 0}$ such that $c_n \to x^\star$. The method always converge. We stop to process at $k = k_0$ with $|f(c_{k_0})| < \varepsilon$ for a given $\varepsilon > 0$.



The stop condition might be incorrect.

Bisection method



Bisection method

Computing $\alpha_1 = 0.99639033$ with $\varepsilon < 10^{-8}$.

n	a_n	b_n	c_{n+1}	$f(c_{n+1})$
0	0.950000000	1.050000000	1.000000000	-1.3e-03
1	0.950000000	1.000000000	0.975000000	8.0e-03
2	0.975000000	1.000000000	0.987500000	3.2e-03
3	0.987500000	1.000000000	0.993750000	9.5e-04
4	0.993750000	1.000000000	0.996875000	-1.7e-04
5	0.993750000	0.996875000	0.995312500	3.9e-04
10	0.996386719	0.996484375	0.996435547	-1.6e-05
15	0.996389771	0.996392822	0.996391296	-3.5e-07
20	0.996390247	0.996390343	0.996390295	1.2e-08
24	0.996390325	0.996390331	0.996390328	1.1e-09

Newton's method (The analysis)

Problem. Improve the efficiency of the bisection method under the extra hypothesis that f is derivable.

Newton's method. Instead of solving $f(x) = 0$ we argue as follows.

(a) Assume $x_0 \approx x^\star$ where $x^\star$ is a solution of $f(x) = 0$.

(b) Consider the linear function which better approximate f near x_0; that is, $L(f, x_0) = f(x_0) + f(x_0)(x - x_0)$.

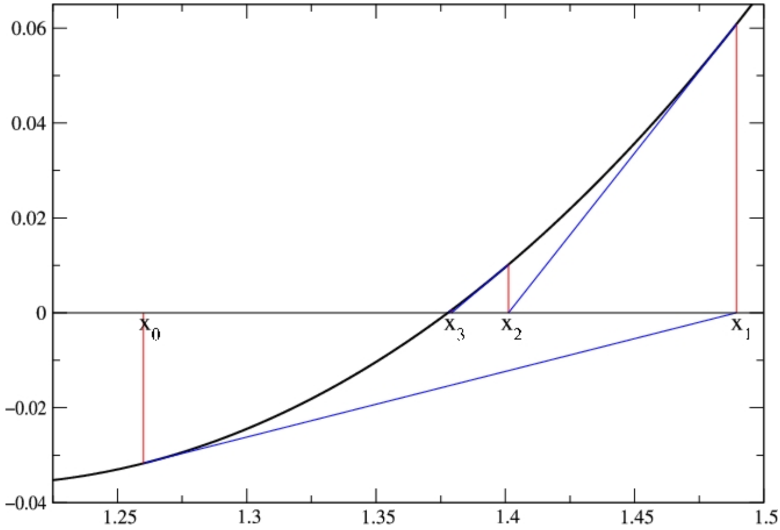
(c) Then consider $L(f, x_0) = 0$.

(d) Set

$$x_1 := x_0 - \frac{f(x_0)}{f'(x_0)} \qquad \text{and } x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$$

(e) We have $x_n \to x^\star$ as $n \to \infty$.

Newton's method (The geometry)



Newton's method (Remarks)

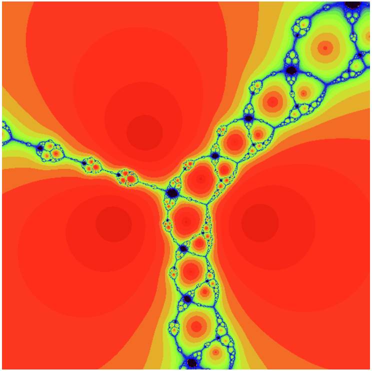
- We have assumed $x_0 \approx x^\star$.
- Morally, we are assuming $f'(x) \neq 0$ near $x^\star$.
- Stop criteria. We might use either $|x_{n+1} - x_n| < \varepsilon$ or $|f(x_{n+1})| < \delta$, on $\varepsilon > 0, \delta > 0$. Both criteria might have problems.
- Convergence. The speed of convergence is much better, locally, than the bisection method.
- The idea can be generalized to higher dimension.

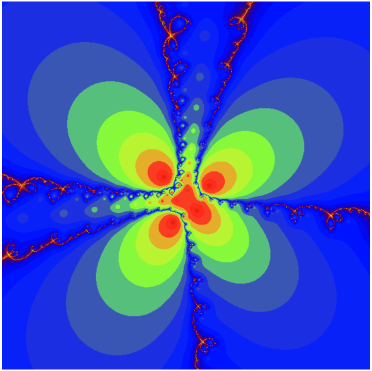
Newton's method (Example)

Computing α_1, a zero of $f(x) = \exp(x - 0.5) - 2x + 0.35$.

n	x_n	$f(x_n)$	$f'(x_n)$	$\lvert x_n - x_{n-1} \rvert$
0	0.95000000000	1.8312185490e-02	-4.3168781451e-01	
1	0.99241997313	1.4312185890e-03	-3.6372883516e-01	0.4e-01
2	0.99635482363	1.2683863782e-05	-3.5727766888e-01	0.4e-02
3	0.99639032504	1.0352153579e-09	-3.5721934888e-01	0.4e-04
4	0.99639032794	1.1102230246e-16	-3.5721934412e-01	0.3e-08

Newton's method (Global approach)





Newton's method as a fixed point method

Newton's method can be viewed as a fixed point method. Indeed we have found a function g such that

$$f(x^\star) = 0 \iff g(x^\star) = x^\star.$$

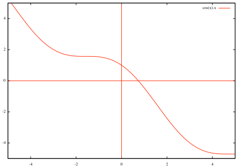
In other words our unknown value $x = x^\star$ instead of a zero of f becomes a fixed point of g. Precisely

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

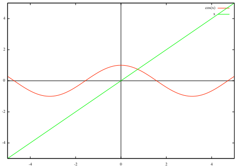
Remark. Newton's method is just a particular case of the fixed point theory to find out zeros of functions. The advantage (fixed points instead of zeros) is that, under certain conditions, the map g gives a natural path to create $x_n \to x^\star$.

Fixed point method (example)

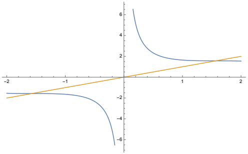
Exercise. Suppose we want to compute $x - \cos(x) = 0$. Then we might consider to iterate $g(x) = \cos(x)$, starting with a value close to a solution.



$f(x) = \cos(x) - x$



$g(x) = \cos(x)$



$N_f(x) = x + \tan^{-1}(x)$

Fixed point theory not always work

Exercise. We want to find out the unique zero of

$$f(x) = x - \exp(-x) = 0 \qquad (x^\star = 0.567143)$$

- As a fixed point of $g_1(x) = \exp(-x)$.
- As a fixed point of $g_2(x) = -\log(x)$.
- As a fixed point of $g_3(x) := N_f(x) = x - \dfrac{x - \exp(-x)}{1 + \exp(x)}$

$g_2(x) = -\log(x) \quad \rightarrow$

Not convergent

n	x_n	n	x_n
0	0.55	5	0.895394
1	0.597837	6	0.110492
2	0.514437	7	2.202816
3	0.664682	8	–0.789737
4	0.408447		

Fixed point Theorem

Theorem (fixed point). Let $g : [a, b] \longrightarrow [a, b]$ a continuous function. Suppose f is derivable in (a, b) and it satisfies

$$|g'(x)| \leq k < 1 \quad \forall x \in (a, b) \ .$$

Then, for all $x_0 \in (a, b)$ we have

$$x_n := g(x_{n-1}) \longrightarrow x^\star$$

with $g(x^\star) = x^\star$. Moreover the following inequalities hold

$$|x_n - \alpha| \leq \frac{k^n}{1-k}|x_0 - x_1| \quad \text{and} \quad |x_n - \alpha| \leq \frac{k}{1-k}|x_n - x_{n-1}|.$$

Truncate conditions

- The blue inequality

$$|x_n - \alpha| \leq \frac{k^n}{1-k}|x_0 - x_1|$$

is a priori estimate of the number of iterates.

- The red inequality

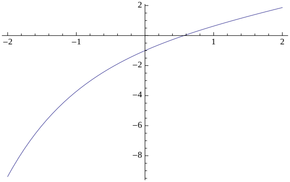
$$|x_n - \alpha| \leq \frac{k}{1-k}|x_n - x_{n-1}|$$

is a key information to decide the stop condition.

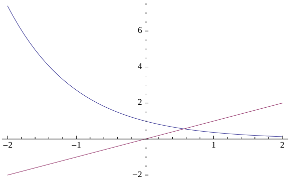
Fixed point Theorem (example)

Exercise: We are calculating the unique zero of $f(x) = x - e^{-x}$.
We use (see above) $g(x) = \exp(-x)$.

- $g : [0.2, 1] \mapsto [0.2, 1]$.
- $|g'(x)| < k = 0.82$ for all $x \in [0.2, 1]$.
- Fixed point Theorem The function g has a unique fixed point $x^\star \in [0.2, 1]$ (i.e., $x = x^\star$ is the unique zero of f in $[0.2, 1]$).



$f(x) = x - e^{-x}$



$g(x) = e^{-x}$

Fixed point Theorem (example)

The a priori bound is

$$\frac{0.82^n}{0.18}|0.5 - 0.60653| < 10^{-8} \quad \Rightarrow n > 90.$$

n	x_n	$\lvert x_n - x_{n-1}\rvert$	$\frac{k}{1-k}\lvert x_n - x_{n-1}\rvert$
0	5.00000000e-01		
1	6.06530660e-01	1.0653066e-01	4.8530633e-1
2	5.45239212e-01	6.12914478e-02	2.792166e-1
3	5.79703095e-01	3.44638830e-02	1.570021334e-1
10	5.66907213e-01	6.52421327e-04	2.9721416e-3
20	5.67142478e-01	2.24611113e-06	1.02323e-5
30	5.67143288e-01	7.73319819e-09	3.5229e-8

Remark. Observe that $n \approx 30$ its enough. This is so because of the bounds of the derivative.

Fixed point Theorem (example)

Remark. We want to find out the unique zero of

$$f(x) = x - \exp(-x) = 0 \qquad (x^\star = 0.567143).$$

- As a fixed point of $g_1(x) = \exp(-x)$.
- As a fixed point of $g_2(x) = -\log(x)$.
- As a fixed point of $g_3(x) := N_f(x) = x - \dfrac{x - \exp(-x)}{1 + \exp(x)}$

Remark. We just noticed above that the method fails. Observe that $|g_2'(x^\star)| > 1$.

Order of convergence

Definition. Assume the above notation. Let $(x_n)_n \to x^\star$ with $g(x^\star) = x^\star$. Denote by $(\varepsilon_n)_n := x_n - x^\star$. We say that the (fixed point) iterative method has order of convergence $m > 0$ if

$$\lim_{n\to\infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|^m} = C > 0.$$

Remark/Exercise.

- If $0 < |g'(|(x^\star)| < 1$ then $m = 1$ (linear convergence).
- If $g'(x^\star) = \cdots = g^{(k-1)}(x^\star) = 0$ and $g^{(k)}(x^\star) \neq 0$ then $m = k$.
- Newton's method $g = N_f$. If $f'(\alpha) \neq 0$ (simple zeros) then $m \geq 2$ (quadràtic convergence)