Application to Convex Functions

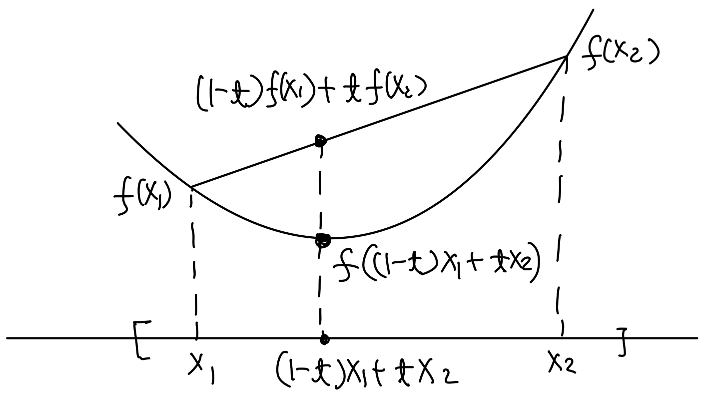
Def 6.4.5 Let I be an interval.

A function $f: I \to \mathbb{R}$ is said to be convex on I if $\forall\, t \in [0,1]$ and any $x_1, x_2 \in I$, we have

$$f((1-t)x_1 + t x_2) \le (1-t) f(x_1) + t f(x_2)$$

Geometric meaning:

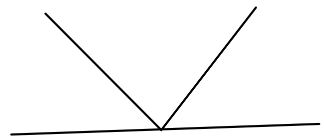
Graph always below (at most up to) chord (with same end pts)



Remark: Convex function need not be differentiable

eg: $f(x) = |x|$ is convex

but not differentiable



Thm 6.4.6 Let
- $f: I \to \mathbb{R}$ (I (open) interval)
- $f''(x)$ exists $\forall x \in I$

Then f is convex on $I \iff f''(x) \geq 0, \ \forall x \in I$

Pf ($\Rightarrow$) (Ex 16 of §6.4)

$$f''(x) \text{ exists} \Rightarrow f''(x) = \lim_{h \to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

Now, f convex on $I \Rightarrow$

$\forall x \in I$ and $h \in \mathbb{R}$ such that $x \pm h \in I$, we have

$$f\left(\tfrac{1}{2}(x+h) + \tfrac{1}{2}(x-h)\right) \leq \tfrac{1}{2} f(x+h) + \tfrac{1}{2} f(x-h)$$

i.e. $$2f(x) \leq f(x+h) + f(x-h)$$

Therefore $$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \geq 0$$

$$\Rightarrow \quad f''(x) = \lim_{h \to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \geq 0, \quad \forall x \in I.$$

$(\Leftarrow)$ Assuming $f''(x) \geq 0$, $\forall x \in I$.

$\forall t \in [0,1]$ & $\forall x_1, x_2 \in I$,

let $x_0 = (1-t)x_1 + t x_2$ (clearly $\in I$)

Then Taylor's Thm $\Rightarrow$

$$f(x_1) = f(x_0) + f'(x_0)(x_1 - x_0) + \frac{1}{2} f''(c_1)(x_1 - x_0)^2$$

$$\geq f(x_0) + f'(x_0)(x_1 - x_0) \qquad \left(\text{for some } c_1 \text{ between } x_1 \text{ \& } x_0\right)$$

and

$$f(x_2) = f(x_0) + f'(x_0)(x_2 - x_0) + \frac{1}{2} f''(c_2)(x_2 - x_0)^2$$

$$\geq f(x_0) + f'(x_0)(x_2 - x_0) \qquad \left(\text{for some } c_2 \text{ between } x_2 \text{ \& } x_0\right)$$

$$\Rightarrow (1-t) f(x_1) + t f(x_2)$$

$$\geq (1-t) f(x_0) + t f(x_0) + f'(x_0)\Big[(1-t)(x_1 - x_0) + t(x_2 - x_0)\Big]$$

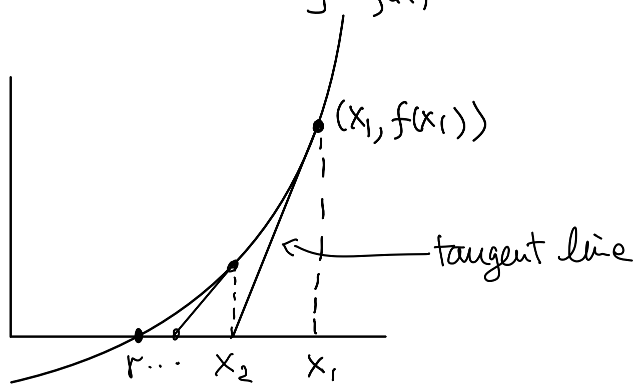
$$= f(x_0) + f'(x_0)\Big[(1-t)x_1 + t x_2 - x_0\Big]$$

$$= f((1-t)x_1 + t x_2), \quad \text{since} \quad x_0 = (1-t)x_1 + t x_2$$

$\#$

Newton's Method

Idea:



$f(r)=0$

Equation of tangent line: $y - f(x_1) = f'(x_1)(x - x_1)$

$\therefore$ its intersection with x-axis, x_2, satisfies

$$-f(x_1) = f'(x_1)(x_2 - x_1)$$

$$\therefore \quad x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \qquad (\text{provided } f'(x_1) \neq 0)$$

Successively

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \left(\begin{array}{c}\text{provided } f'(x_k) \neq 0 \\ \text{for } k = 1, \cdots, n\end{array}\right)$$

Hope (to find condition such that)

$$x_n \to r \qquad (\text{a zero of } f).$$

Thm 6.4.7 (Newton's Method)

Let
- $f:[a,b]\to\mathbb{R}$ twice differentiable $(a<b)$
- $f(a)f(b)<0$ (ie $f(a), f(b)$ have opposite signs)
- $\exists$ constants $m>0$, $M\geq 0$ such that

$$|f'(x)|\geq m>0 \quad \& \quad |f''(x)|\leq M, \qquad \forall x\in[a,b].$$

Then $\exists$ a subinterval $I^*\subset[a,b]$
- containing a zero r of f, such that
- $\forall\, x_1\in I^*$, the sequence (x_n) defined by

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)} \qquad \forall\, n=1,2,3\cdots$$

belongs to I^* and
- $\lim_{n\to\infty} x_n=r$

Moreover

$$|x_{n+1}-r|\leq K|x_n-r|^2 \quad \forall\, n=1,2,3\cdots$$

where $K=M/2m$.

Pf: Since $f(a)f(b)<0$, $f(a), f(b)$ have opposite signs (& nonzero)

f twice differentiable $\Rightarrow$ f cts. on $[a,b]$.

Intermediate Thm $\Rightarrow$ $\exists\, r\in(a,b)$ such that $f(r)=0$.

Note that $|f'(x)|\geq m>0$, $\forall\, x\in[a,b]$, Rolle's Thm

$\Rightarrow$ r is the unique zero of f in $[a,b]$.

i.e. $f(x) \neq 0 \ \forall x \in [a,b] \setminus \{r\}$.

Now $\forall x' \in I$, Taylor's Thm $\Rightarrow$

$$0 = f(r) = f(x') + f'(x')(r-x') + \frac{f''(c')}{2}(r-x')^2$$

for some c' between r & x'.

(since f is twice diff.)

If $x'' = x' - \dfrac{f(x')}{f'(x')}$, we have

$$x'' = x' + \frac{f'(x')(r-x') + \frac{f''(c')}{2}(r-x')^2}{f'(x')}$$

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

$$\Rightarrow \quad |x''-r| \leq \frac{1}{2}\frac{|f''(c')|}{|f'(x')|}|x'-r|^2$$

$$\leq \frac{1}{2}\frac{M}{m}|x'-r|^2 = K|x'-r|^2. \qquad (*)$$

Choose $\delta > 0$ such that

$$\delta < 1/K \quad \& \quad [r-\delta, r+\delta] \subset [a,b],$$

and let $I^* = [r-\delta, r+\delta]$

Then, if $x_n \in I^* \ (\subset [a,b])$ for some $n = 1,2,3,\cdots$,

we have, from $(*)$,

$$|x_{n+1} - r| \leq K|x_n - r|^2 < K\delta^2 < \delta$$

$\therefore \quad x_{n+1} \in I^*$.

i.e. $\quad x_n \in I^* \Rightarrow x_{n+1} \in I^*$.

Therefore, if $x_1 \in I^*$, induction $\Rightarrow$

the sequence $(x_n) \subset I^*$.

and satisfies the required inequality

$$|x_{n+1} - r| \leq K|x_n - r|^2, \quad \forall\, n = 1, 2, 3, \cdots$$

Finally, to see limit, we note 1st that

$$|x_{n+1} - r| \leq K|x_n - r|^2 \leq K\delta\, |x_n - r| \qquad \text{---} \; (*)_2$$

Then iterate $(*)_2$:

$$|x_{n+1} - r| \leq (K\delta)|x_n - r| \leq (K\delta)\big(K\delta\, |x_{n-1} - r|\big)$$

$$= (K\delta)^2 |x_{n-1} - r| \leq \cdots$$

$$\leq (K\delta)^n |x_1 - r|$$

Since $K\delta < 1$, $(K\delta)^n \to 0$ as $n \to \infty$,

and $|x_1 - r|$ is a constant, we have

$$|x_{n+1} - r| \to 0 \quad \text{as } n \to \infty$$

i.e. $\quad \lim_{n\to\infty} x_n = r \qquad$ ⋇

eg 6.4.8 Using Newton's Method to approximate $\sqrt{2}$.

Soln: Convert the problem to a problem of finding root in order to use Newton's Method:

Consider $f(x) = x^2 - 2 \quad \forall x \in \mathbb{R}$.

Calculation: $f'(x) = 2x$ ($\neq 0$ near the root, as 0 is not a root)

(f'' exists and satisfies the condition, but we don't need it in the approximation.)

One need to guess an <u>initial point</u> x_1.

Since $1^2 = 1$, $2^2 = 4$, ($f(1) = -1$, $f(2) = 2$)

it seems reasonable to try $x_1 = 1$.

Note that
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$
$$= x_n - \frac{x_n^2 - 2}{2x_n}$$
$$= x_n - \frac{1}{2}x_n + \frac{1}{x_n}$$
$$= \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

$\therefore x_1 = 1 \Rightarrow$
$$x_2 = \frac{1}{2}\left(1 + \frac{2}{1}\right) = \frac{3}{2} = 1.5$$
$$x_3 = \frac{1}{2}\left(\frac{3}{2} + \frac{2}{3/2}\right) = \frac{17}{12} \simeq 1.416666$$
$$\vdots$$

(Check!) $x_5 \approx 1.414213562372$ (correct to 11 places).

Remarks

(a) (*) can be written as $\left(K|x_{n+1}-r|\right) \leq \left(K|x_n-r|\right)^2$

Hence if $K|x_n-r| < 10^{-m}$,

then $K|x_{n+1}-r| < 10^{-2m}$

$\therefore$ number of significant digits in $K|x_n-r|$

has been <u>doubled</u>.

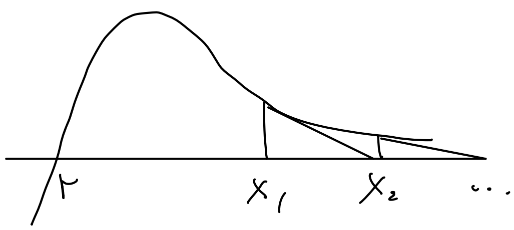
And hence, the sequence (x_n) generated by Newton's method

is said to "<u>converge quadratically</u>".

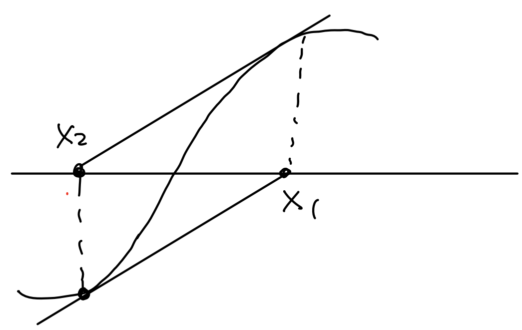
(b) Choose of initial x_1 is <u>important</u> (ie. has to be in I^*),

otherwise (x_n) may not converge to the zero.

Possible situations



$(x_n \to \infty)$



(seq is $(x_1, x_2, x_1, x_2, x_1, x_2, \cdots)$)