Week 4.

One dimensional un-constrained Minimization.

1D problem is defined as

$$\min(f) \quad s.t. \quad f : \mathbb{R} \to \mathbb{R}.$$
$$w/ \ \Omega = \mathbb{R}. \text{ (The whole of real line)}$$

~ The approach is to use an ~~itt~~ iterative search algorithm, also called a "line-search" method.

~ 1D search methods are important

i) They are special cases of search method. used in multivariate case.

ii) They are used as part of general multivariable algorithms.

Iterative algorithm. In an iterative algorithm, we start with an initial candidate solution $x^{(0)}$ and generates a ~~series~~ sequence of iterates $x^{(1)}, x^{(2)}, \dots$ for each iteration $k = 0, 1, 2, \dots$

$x^{(k+1)}$ will depend on $x^{(k)}$ & f (objective function)

The algorithm may use f evaluated at specific points, or perhaps its 1st derivative f' or even its 2nd derivative f''.

Algorithm of interest

i) Golden section method (uses only f)

ii) Fibonacci Method (uses only f)

iii) Bisection method (uses only f')

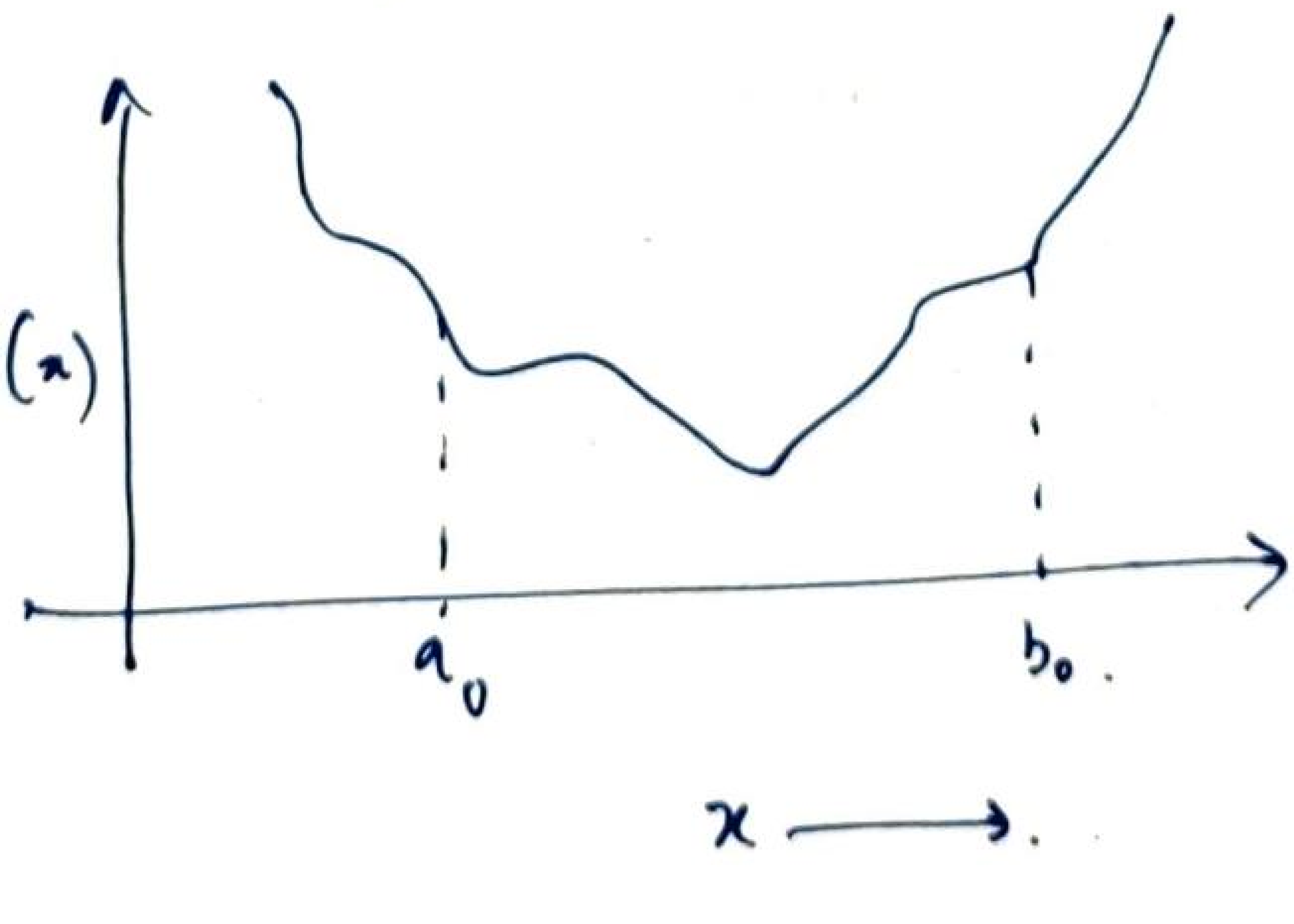
iv) Secant Method (uses only f')

v) Newton's Method (uses f' and f'')

Golden section method.

~~need~~ Target: To find $\arg\min(f)$ over a closed interval $[a_0, b_0]$.

Assumptions i) $f : \mathbb{R} \to \mathbb{R}$. is unimodal, which means that f has only one local minimizer.



method: i) We evaluate f at different points in the interval $[a_0, b_0]$. The choice of these points, in a

~~minimum number of iterations~~. aim to find the approximation to the minimizer 'f' in as ~~a~~ minimum number of iterations as possible.

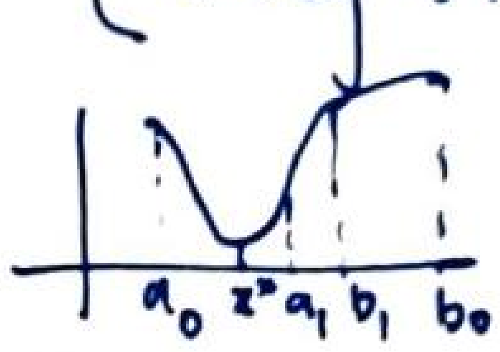
1st iteration.

ii) Chose two intermediate points a_1, b_1 within the interval

$$a_0 < a_1 < b_1 < b_0.$$

now as f is unimodal (only one minimum).

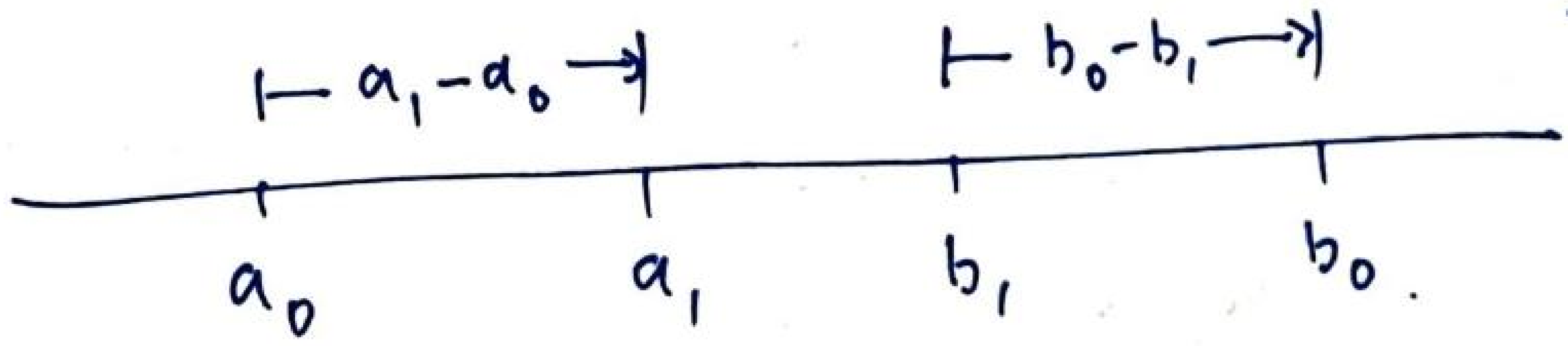
Case I: if $f(a_1) < f(b_1)$

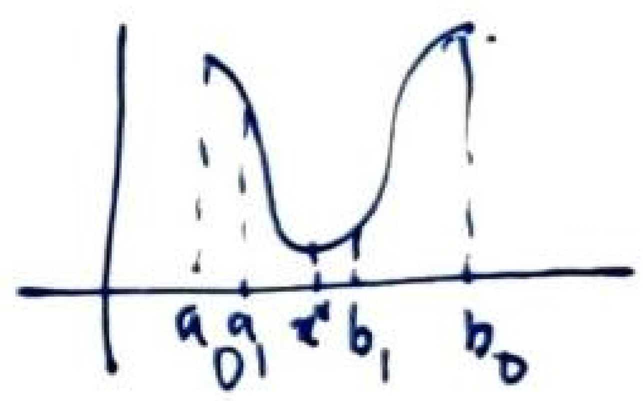


$\Rightarrow f(a_1) < f(b_1)$

$\Rightarrow$ ~~minimum is~~ $\arg\min f = x^* \in [a_0, b_1]$

Case II: If $f(a_1) \geq f(b_1) \Rightarrow \arg\min f = x^* \in [a_1, b_0]$





$\Downarrow$

$f(a_1) > f(b_1)$

Chose these intermediate points in such a way that the reduction in the range is symmetric.

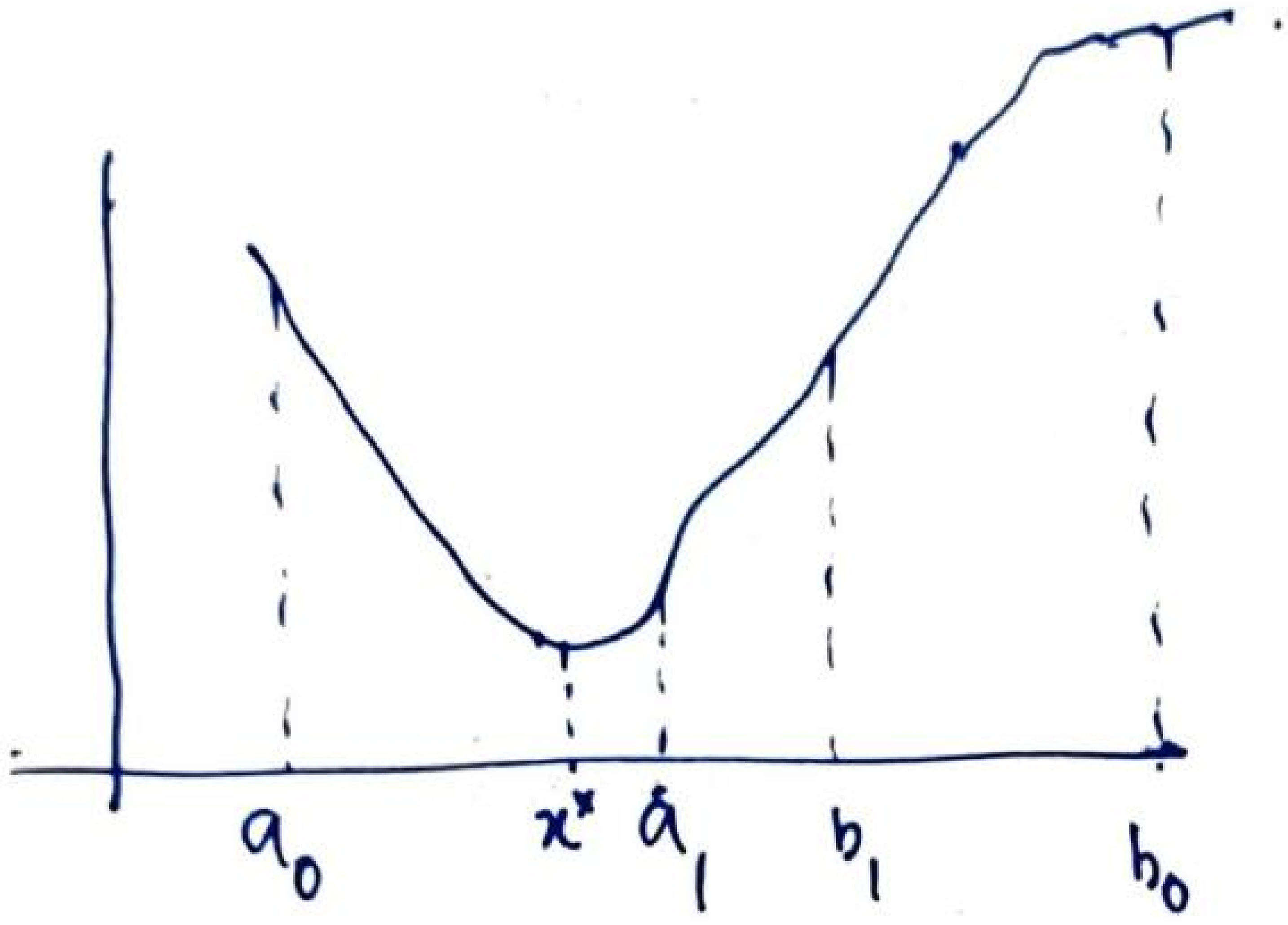
i.e., $$a_1 - a_0 = b_0 - b_1 = \rho(b_0 - a_0).$$

w/ $\rho < \frac{1}{2}$ need to be.

Next iteration

(iii) The new reduced range of uncertainty is $[a_0, b_1]$ OR $[a_1, b_0]$. We can repeate.

the process. and similarly find two new points. a_2 & b_2, using the same value $\rho < \frac{1}{2}$ as before.



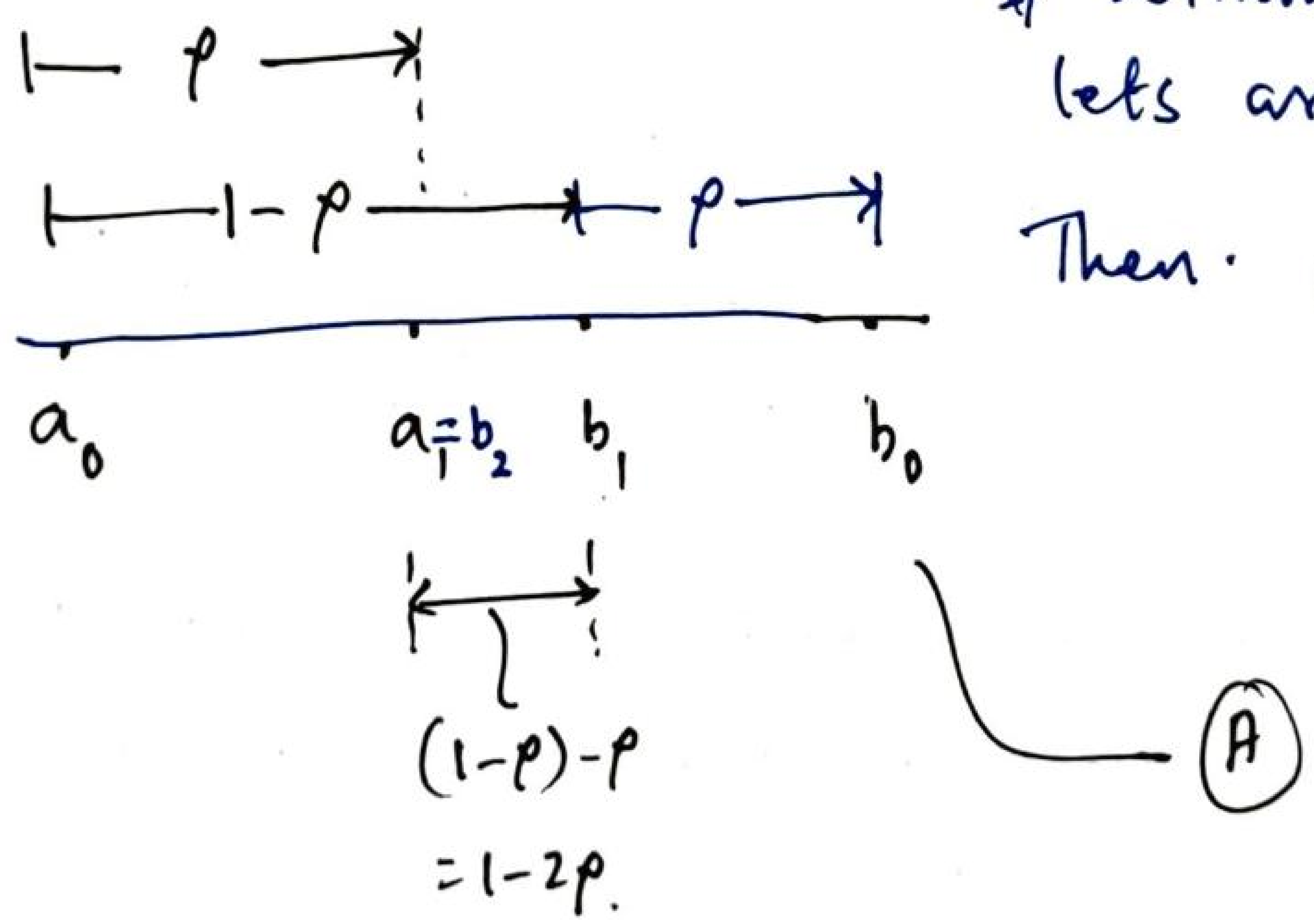
If $f(a_1) < f(b_1)$

Then $x^* \in [a_0, b_1]$

now for the next iteration If we choose $a_1 = b_2$, the we need to find one additional pt a_2.

Without loss of generality lets assume $a_0 - b_0 = 1$.

Then. $a_1 - a_0 = b_1 - b_0 = \rho$



$[a, b_1]$ is ~~not~~ the new uncertainty interval. From the above diagram.

$$\underset{\text{new pt}}{a_2} - a_0 = b_1 - \underset{\text{new pt.}}{b_2} = \rho\,(b_1 - a_0).$$

$$\Rightarrow 1 - 2\rho = \rho\,(1-\rho). \quad [\text{From A above}].$$

$$\Rightarrow \quad \rho_1 = \frac{3+\sqrt{5}}{2} \quad , \quad \rho_2 = \frac{3-\sqrt{5}}{2}.$$

As we require $\rho < \frac{1}{2}$ hence the acceptable solution is.

$$\boxed{\rho = \frac{3-\sqrt{5}}{2} \simeq 0.382}.$$

Observation: I)

$$1-\rho = 1 - \frac{3-\sqrt{5}}{2} = \frac{\sqrt{5}-1}{2}.$$

$$\Rightarrow \quad \frac{\rho}{1-\rho} = \frac{3-\sqrt{5}}{\sqrt{5}-1} = \frac{(3-\sqrt{5})(\sqrt{5}+1)}{5-1} = \frac{3\sqrt{5}-5+3-\sqrt{5}}{4}$$

$$\Rightarrow \quad \frac{\rho}{1-\rho} = \frac{\sqrt{5}-1}{2} = \frac{1-\rho}{1}.$$

$$\Rightarrow \quad \frac{\rho}{1-\rho} = \frac{1-\rho}{1}$$

ρ : shorter segment
$1-\rho$: longer segment

Thus, dividing the range in the ratio of $\rho : (1-\rho)$ has the effect that the ratio of shorter segment to the longer equals the ratio of the longer to the sum of the two.

— This rule was referred to by ancient GREEK geometers as the "Golden Section".

~~The constraint "$\underline{x} \in \Omega$" is called a set constraint. Often, the constraint set Ω takes the form $\Omega = \{ \underline{x} : \underline{h}(\underline{x}) = 0, \ \underline{g}(\underline{x}) \leq 0 \}$, where $\underline{h}$ & $\underline{g}$ are given functions. We refer to ~~sum~~ such constraints as functional constraints. The remainder of this chapter deals with general set constraints, including special case $\Omega = \mathbb{R}^n$ is called unconstrained case.~~

Observation II.

Using the ~~g~~ Golden Section rule means at every stage of the uncertainty range reduction (except the first), the objective function f need only to be evaluated at one new point.

The uncertainty range is reduced by the ratio

$$1 - \rho = 0.61803 \text{ at every stage.}$$

Hence, after N steps (iterations) of reduction using GSM reduces the range by the factor

$$(1-\rho)^N \approx (0.61803)^N$$

Example: Find the minimum of $f(x) = x^4 - 14x^3 + 60x^2 - 70x$ in the interval $[0,2]$. We wish to locate this value of x to within a range of 0.3.

Solun: After N stages the rang $[0,2]$ is reduced by $(0.61803)^N$, so we choose N so that

$$2(0.61803)^N \leq 0.3$$

$$\Rightarrow \quad N \approx 4$$

Iteration 1: We evaluate f at two intermediate points a_1 & b_1. We have.

$$a_1 = a_0 + \rho(b_0 - a_0) = 0.7639$$

$$b_1 = a_0 + (1-\rho)(b_0 - a_0) = 1.236.$$

where $\rho = \dfrac{3-\sqrt{5}}{2}$.

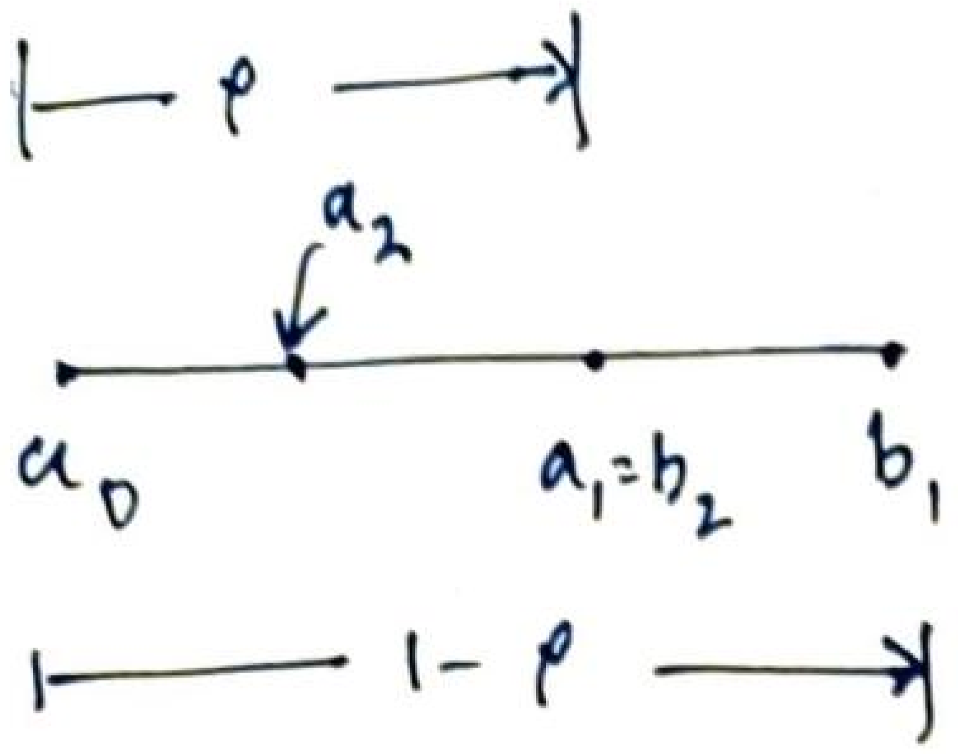
$$f(a_1) = -24.36$$

$$f(b_1) = -18.96 \quad \Rightarrow \quad f(a_1) < f(b_1)$$

So the uncertainty interval is reduced to.

$$[a_0, b_1] = [0, 1.236]$$

Iteration 2. We choose b_2 to coinside with a_1, and so f need only be evaluted at one new point.



$$a_2 = a_0 + \rho(b_1 - a_0) = 0.4721$$
$$b_2 = a_1 = 0.7639.$$

$a_0 - b_0 = 1$

We have.

$$f(a_2) = -21.10$$
$$f(b_2) = f(a_1) = -24.36$$

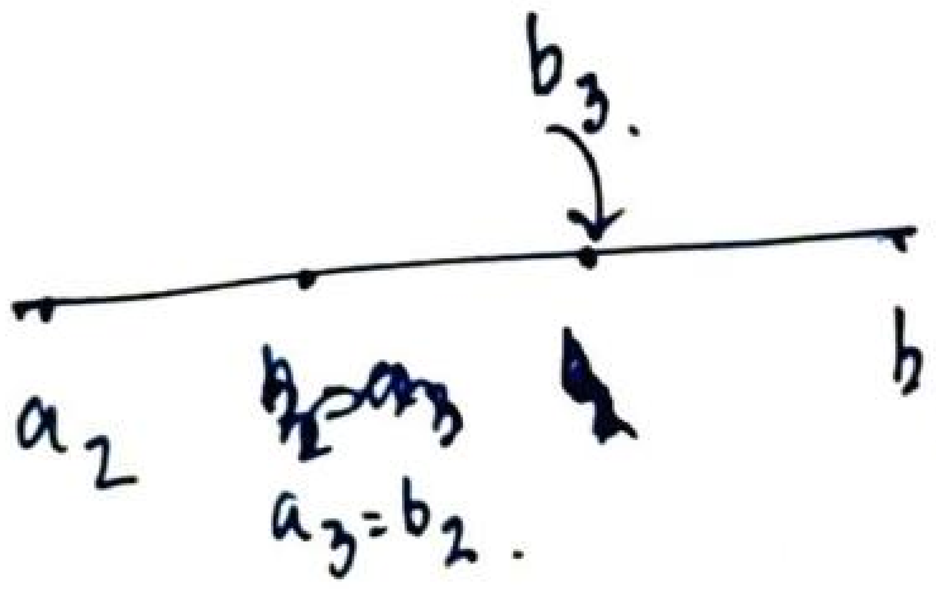
Now $f(b_2) < f(a_2)$, so the uncertainty interval is reduced to.

$$[a_2, b_1] = [0.4721, 1.236]$$

Iteration 3

lets set $a_3 = b_2$.

$$b_3 = a_2 + (1-\rho)(b_1 - a_2) = 0.9443.$$



We have, $f(a_3) = f(b_2) = -24.36$

$$f(b_3) = -23.59$$

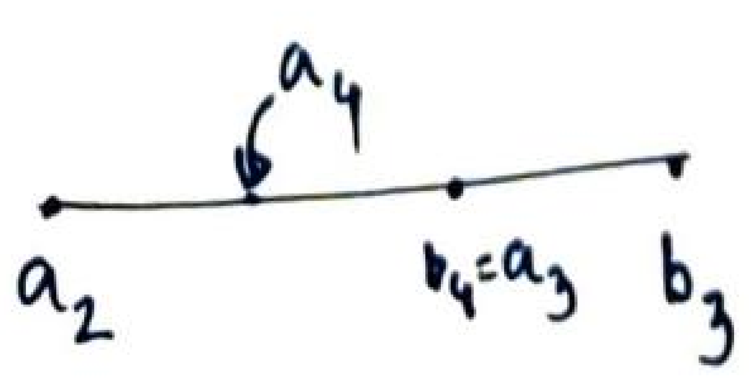
so, $f(b_3) > f(a_3)$. Hence the uncertainty interval is further reduced to.

$$[a_2, b_3] = [0.4721, 0.9443]$$

Iteration 4

We set $b_4 = a_3$ and.

$$a_4 = a_2 + \rho(b_3 - a_2) = 0.6525$$



We have, $f(a_4) = -23.84$

$$f(b_4) = f(a_3) = -24.36.$$

hence, $f(a_4) > f(b_4)$. Thus the $x^* = \arg\min(f)$ is located.

, in the interval

$$[a_4, b_3] = [0.6525, 0.9443].$$

Note that $\cancel{b_4}\ b_3 - a_4 = 0.292 < 0.3$.

Bisection Method.

Optimization prob: Find $\text{argmin}(f) = x^*$

Assumptions: i) $f \to$ unimodal w/ $x \in [a_0, b_0]$

ii) f is continuously differentiable i.e., f' exists. In other words. ~~the~~ $f \in C^1$

Method: The "bisection method" is a simple algorithm for successively reducing the uncertainty interval based on evaluations of the derivatives.

i) To begin with, let $x^{(0)} = \left(\frac{a_0 + b_2}{2}\right)$ be the midpt

$[a_0, b_0]$: uncertainty interval.

ii) Next evaluate $f'(x^{(0)})$.

iii) a) If $f'(x^{(0)}) > 0$, then we infer the $x^* = \text{argmin}(f)$ lies left of $x^{(0)}$. The uncertainty interval is reduced to $[a_0, x^{(0)}]$

b) If $f'(x^{(0)}) < 0$ then we infer x^* lies right of $x^{(0)}$. In other words the uncertainty interval is $[x^{(0)}, b_0]$

c) In the case $f'(\alpha x^{(0)}) = 0$, then we declare $x^* = x^{(0)}$ and terminate the search.

We repeate these a, b, c steps to compute the uncertainty interval.

Bisection method Vs. Golden section method.

i) Instead of values of "f", byse bisection method uses value of f'

ii) At each iteration the length of the uncertainty interval is reducing by 0.5, hence after N-steps the range is reduced by a factor $(1/2)^N$.

iii) The factor is smaller than GSM.

Example: Deduct $f(x) = x^4 - 14x^3 + 60x^2 - 70x$ in $x \in [0, 2]$ to within a range of uncertainty 0.3. & how many iteration steps required?

Ans: i) GSM requires atleast 4

ii) The bisection requires atleas 3

as $$(0.5)^N \leq 0.3/2$$

$$\Rightarrow \boxed{N \approx 3}.$$

Scanned by TapScanner

Newton's Method:

Optimization prob: Find argmin(f), given that $f''(x)$, $f'(x)$ exists at every pt or iteration pt.

$\Rightarrow$ $f \in C^2$, & we can have $f(x^k), f'(x^k), f''(x^k)$

I) We can fit a quadratic ~~form~~ function through $x^{(k)}$ that matches its first and second derivatives with that of function f. This quadratic has the form.

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

Note $q(x^{(k)}) = f(x^k)$, $q'(x^k) = f'(x^k)$, $q''(x^k) = f''(x^k)$

Then instead of minimizing $f(x)$ we minimize it approximation $q(x)$. The First Order Necessar. Condition (FONC) for a minimizer of $q(x)$ yields.

$$0 = q'(x) = f'(x^k) + f''(x^k)(x - x^k)$$

$\Rightarrow$ Setting $x = x^{k+1}$

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

Example: Using Newton's method, find the minimizes.

$$f(x) = \frac{1}{2}x^2 - \sin x.$$

Solution: let guess an intial value $x^{(0)} = 0.5$ & a required tolerance or accuracy $\epsilon = 10^{-5}$

The tolleracen gives a "stopping" criteria. i.e.,

$$|x^{k+1} - x^k| < \epsilon$$

lets compute $f'(x) = x - \cos x$

$$f''(x) = 1 + \sin x.$$

Hence,

$$x^{(1)} = 0.5 - \frac{0.5 - \cos(0.5)}{1 + \sin(0.5)}$$

$$= 0.7552.$$

proceeding in a similar manner, we obtain.

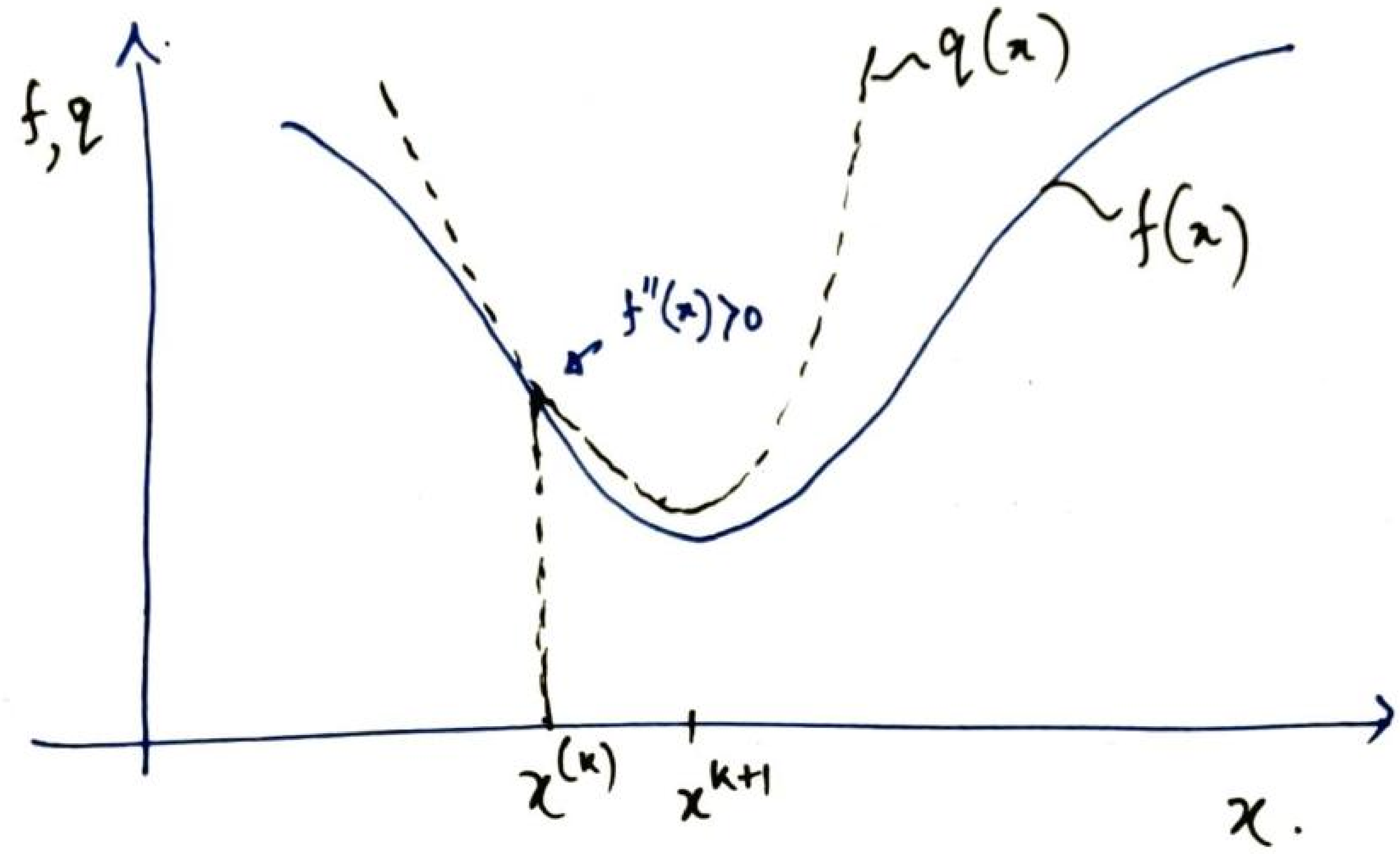
$$x^{(2)} = x^{(1)} - \frac{f'(x^{(1)})}{f''(x^{(1)})} = x^{(1)} - \frac{0.02710}{1.685} = 0.7391$$

$$x^{(3)} = x^{(2)} - \frac{f'(x^{(2)})}{f''(x^{(2)})} = x^{(2)} - \frac{9.461 \times 10^{-5}}{1.673} = 0.7390$$

$$x^{(4)} = x^{(3)} - \frac{f'(x^{(3)})}{f''(x^{(3)})} = x^{(3)} - \frac{1.17 \times 10^{-9}}{1.673} = 0.7390$$

Note. $\left.\begin{array}{l} |x^{(4)} - x^{(3)}| < \epsilon = 10^{-5} \\ f'(x^{(4)}) = -8.6 \times 10^{-6} \approx 0 \end{array}\right\}$ & $f''(x^{(4)}) = 1.673 > 0 \Rightarrow x^* = x^{(4)}$ [Minimizer]

Observation

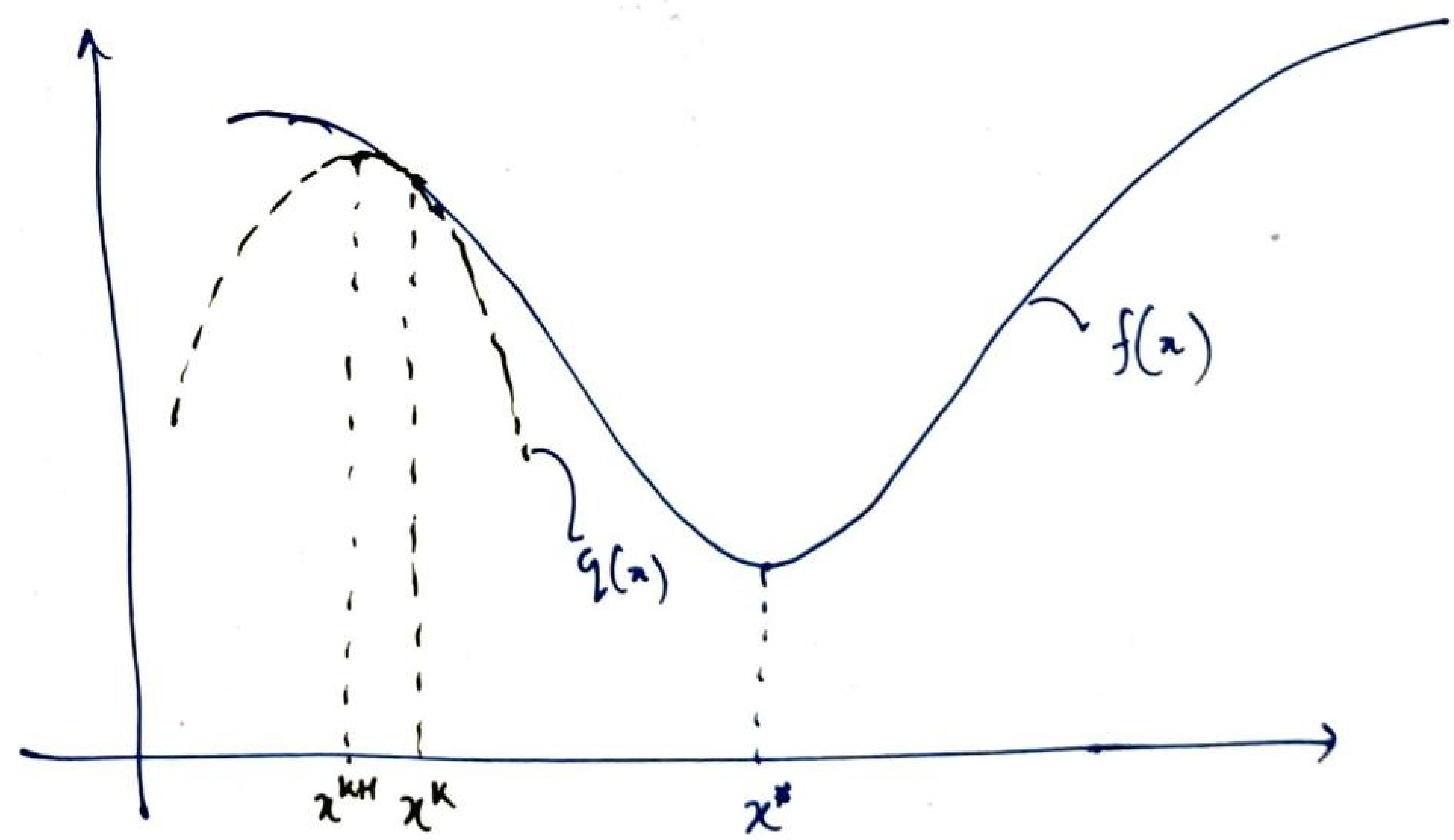


— (A)

Case I: Newton's algo with $f''(x) > 0$.

Newton's method works well if $f''(x) > 0$ everywhere.

Fig (A): However, if $f''(x) < 0$ for some x. Newton's method may fail to converge to the minimizer.

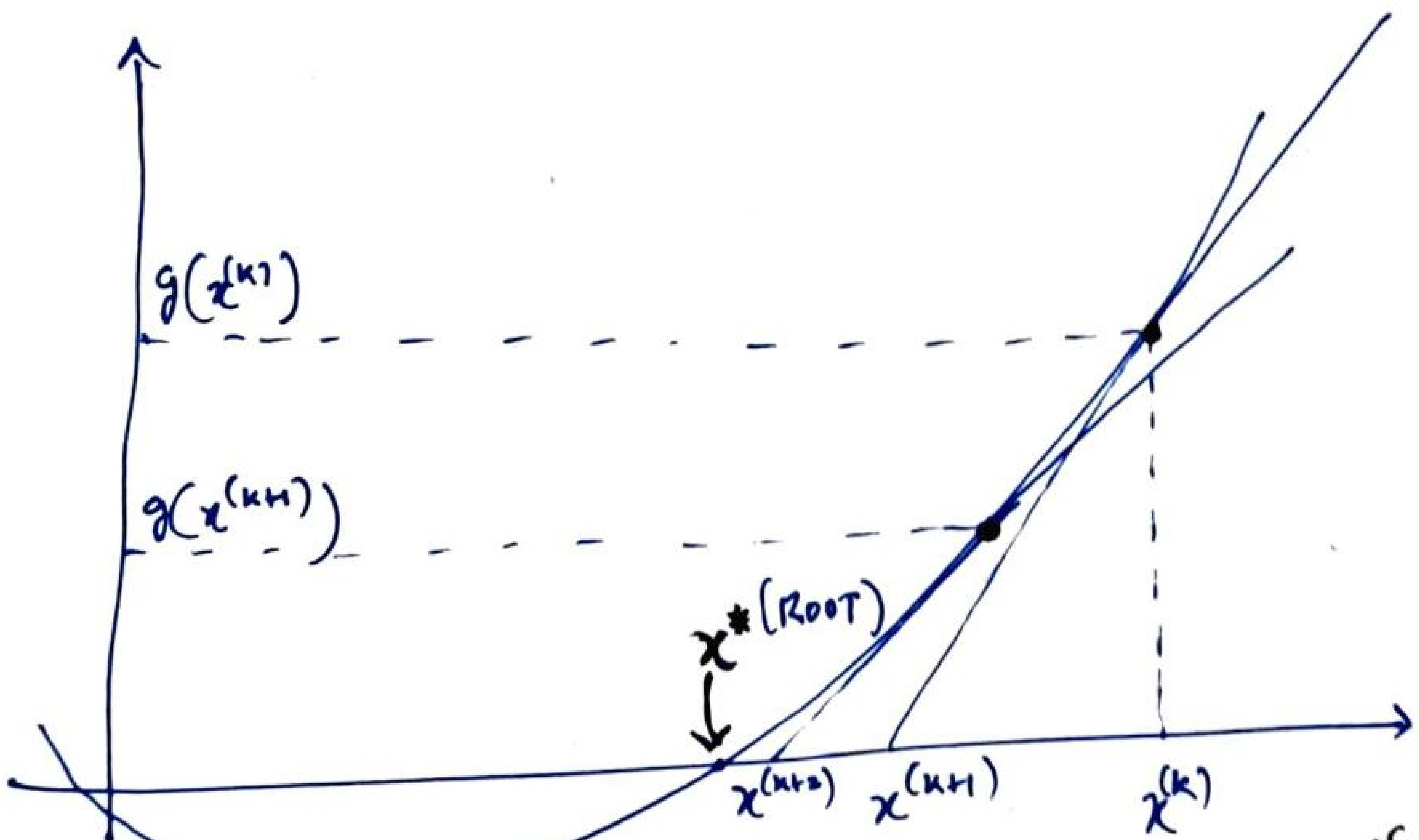


Newton's algo with $f''(x) < 0$.

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

5) Newton's method can also be viewed as a way to derive the 1st derivative of f to zero. Indeed If we get $g(x) = f'(x)$, then we obtain a formula for iterative solution of the equation $\boxed{g(x)=0}$ $\Downarrow$ x^*

$$x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}.$$



Note: If we draw the tangent to $g(x)$ @ $x^{(k)}$, then tangent intersect x-axis @ $x^{(k+1)}$, which is expected to be closure to x^* of $g(x)=0$. The slope of $g(x)$ @ $x^{(k)}$ is $g'(x^{(k)}) = \frac{g(x^{(k)})}{x^{(k)} - x^{(k+1)}} \Rightarrow \boxed{x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}}$

"Newton's method of tangents." geometric meaning.

Example: "Newton's method of tangent" is extensively used for root finding algo. For e.g.,

$$g(x) = x^3 - 12.2x^2 + 7.45x + 42 = 0. \quad \text{Find root}$$

we have $g'(x) = 3x^2 - 24.4x + 7.45$

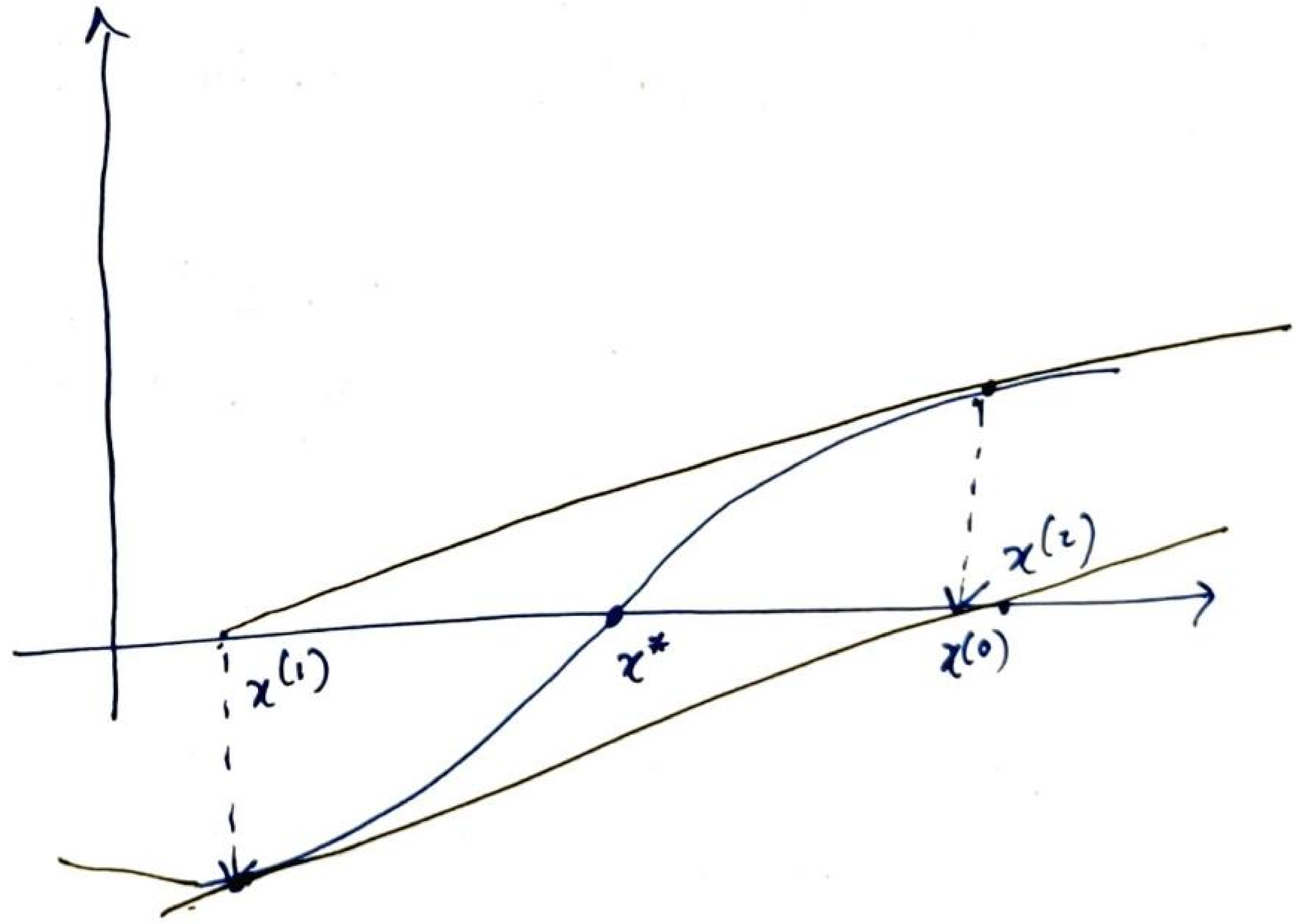
Scanned by TapScanner

Performing two iterations yields.

$$x^{(1)} = x^{(0)} - \frac{g(x)}{g'(x)} = 12 - \frac{102.6}{146.65} = 11.33$$

$$x^{(2)} = x^{(1)} - \frac{g(x^{(1)})}{g'(x^{(1)})} = 11.33 - \frac{14.73}{116.11} = 11.21$$

<u>Observation</u>: Newton's method of tangent fails if the 1st approximation to the root is such that $g(x^{(0)})/g'(x^{(0)})$ is not small enough. Initial approximation to root $x^{(0)}$ is IMP.



The root finding algo fail to converge above.

Secant Method.

Newton's method for minimizing f uses second derivative of "f"

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

If the 2nd derivative is not available, we need to approximate it by 1st derivative information as.

$$f''(x^{(k)}) \approx \frac{f'(x^{(k)}) - f'(x^{(k-1)})}{(x^{(k)} - x^{(k-1)})}$$

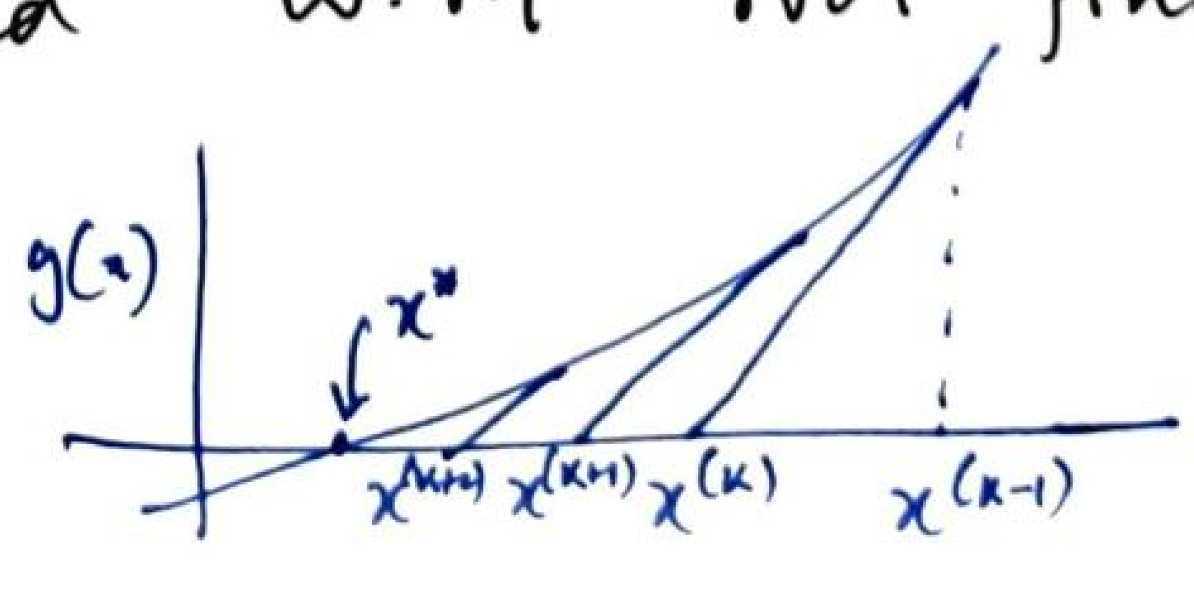
Using the above ~~expression~~ ~~we can~~ approximation of the 2nd derivative, we obtain the algorithm.

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

Observation:

Unlike Newton's method, the secant method does not directly involve values of $f(x^{(k)})$. Instead it tries to drive the derivative f' to zero.

It is easy to interpret, when secant method is analysed w.r.t root finding of $g(x)=0$.



For $g(x) = 0 \Rightarrow x^*$ is the root to find

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{g(x^{(k)}) - g(x^{(k-1)})} g(x^{(k)})$$

Unlike Newton's method which uses the slope of g to determine the next pt, secant method uses "secant" betn (k-1)th and kth point to determine the (k+1)th point.

Example: Use, secant method to find root of

$$g(x) = x^3 - 12.2x^2 + 7.45x + 42 = 0.$$

We perform two iterations., with starting pt $x^{(-1)} = 13$ and $x^{(0)} = 12$.

$$x^{(1)} = 11.40$$
$$x^{(2)} = 11.25.$$

Bracketing : # Golden section, Fibonacci, bisection et.
Unimodality of the function is assumed.

Extension of line-search method to Multidimensional Optimization

Observation: In multidimension optimization 1D. search is involved at every iteration level.

to be specific $f: \mathbb{R}^n \to \mathbb{R}$.

Iterative algorithm to finding a minimum of the form

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha_k \underline{d}^{(k)}.$$

where $k=0$, $\underline{x}^{(0)}$ is the "initial guess" and. $\alpha_k \geq 0$ is chosen to minimize.

$$\underset{\text{1D function of } \alpha}{\phi_k(\alpha)} = f\left(\underline{x}^{(k)} + \alpha \underline{d}^{(k)}\right)$$

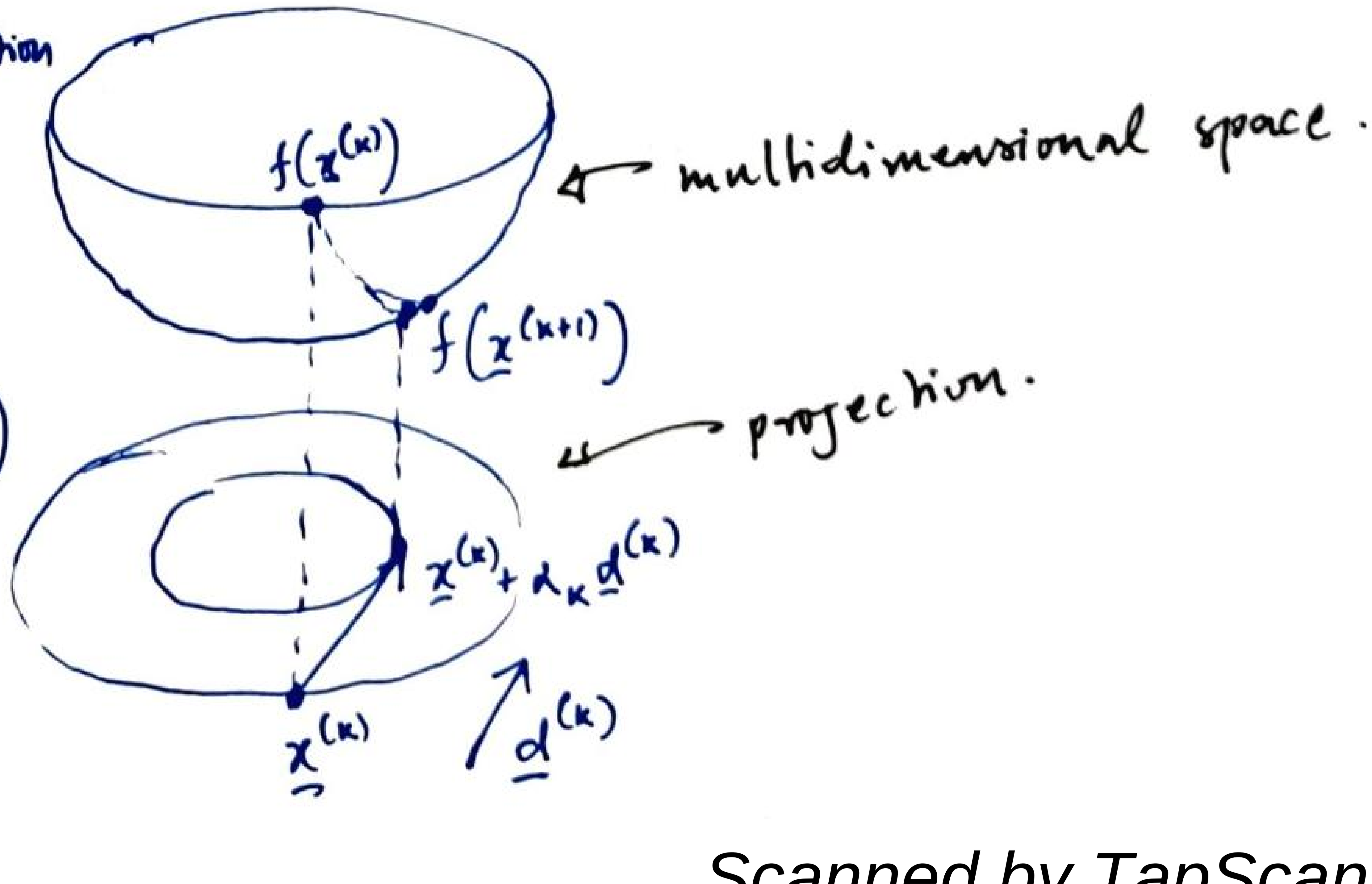
The vector $d^{(k)}$ is SEARCH DIRECTION.

α_k is STEP SIZE

The choice of α_k ensures 1D minimization

This choice ensures under appropriate condition

$$f(\underline{x}^{(k+1)}) < f(\underline{x}^{(k)})$$



Any of the 1D methods discuss i) Golden section ii) secant iii) Newton's Method can be used to. minimize. ϕ_k, we may infact use secant method to find α_k. In this case. derivative of ~~ϕ_k~~ ϕ_k is obtained which is

$$\phi_k'(\alpha) = \nabla f\left(\underline{x}^{(k)} + \alpha_k \underline{d}^{(k)}\right)^T \underline{d}^{(k)}.$$

$$\Rightarrow \quad \alpha_{k+1} = \alpha_k - \frac{\phi_k(\alpha)}{\phi_k'(\alpha)}.$$

later we will find it is better from computational perspective to us. multidimensional algorithm rather using 1D algorithm.

Exercise .

1) ~~Qs~~ let $f(x) = x^2 + 4\cos x$, $x \in \mathbb{R}$. We wish to find minimizer $\underline{x}^* = \arg\min(f)$, $x \in \mathbb{R}$ $\supset [1,2]$.

a) Plot $f(x)$ Vs x in $[1,2]$. to check unimodality

b) Use golden section method to locate x^* to within an ~~unsequity~~ uncertainty of 0.2. Display the intermediate steps. using a table.

Iteration(k)	a_k	b_k	$f(a_k)$	$f(b_k)$	New uncertainty level

ii) Let $f(x) = 8e^{1-x} + 7\log(x)$ w/ natural ~~Logrith~~ logarithm function.

a) Use MATLAB program to implement the ~~gole~~ golden section method., that locates its. minimizer of f over $[1,2]$ to within ~~a certain~~ an uncertainty of 0.23. Display intermediate stage.

.ion

SET

SET