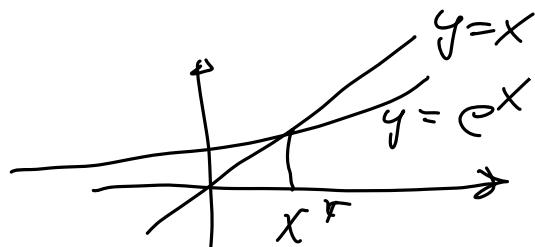


- Course description, organization, etc.
- Part I: Nonlinear Equations and Optimization

§1 Introduction: Concept of Iterative Methods

Nonlinear equations

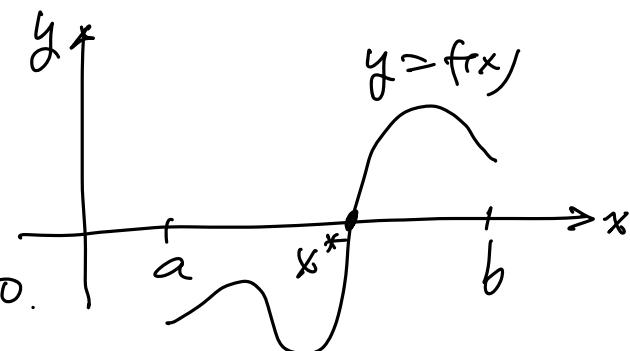
Example 1. $f(x) = x - e^x$



$$2. \begin{cases} x^2y - x + y = 1 \\ xy - e^{x+y} = 0 \end{cases}$$

- Single eq. one-variable.

$$f(x) = 0, \quad x \in [a, b], \quad f(x^*) = 0.$$



- Systems of nonlinear equations: n equations

for n unknowns

$$\begin{cases} f_1(x_1, \dots, x_n) = 0, \\ \vdots \\ f_n(x_1, \dots, x_n) = 0. \end{cases} \quad \text{or} \quad F(x) = 0$$

$$x = (x_1, \dots, x_n), \quad 0 = (0, \dots, 0)$$

$$F(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}, \quad F(x^*) = 0$$

More examples

- 1. Polynomial equations $f(x) = 0$

$$f(x) = a_n x^n + \dots + a_1 x + a_0, \quad a_n \neq 0, \quad n \geq 1.$$

all $a_k \in \mathbb{R}$.

- Simple forms often used in modeling.

① Many functions can be well approximated by polynomials. [2]

② But: there are no general solution formulas for polynomial equation $f(x)=0$ if $\deg f \geq 5$.

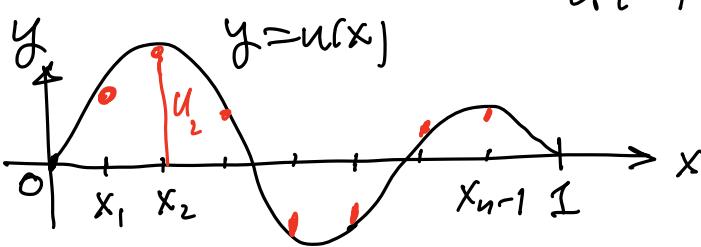
2. System of algebraic equations resulting from discretization of a nonlinear partial differential eq. (PDE).

Example $u=u(x)$, $x \in [0, 1]$. $f=f(x)$, $\alpha=\alpha(u)$.

$$\begin{cases} -u'' + \alpha(u) = f & x \in (0, 1) \\ u(0) = 0, u(1) = 0. \end{cases}$$

Examples of $\alpha(u)$: $\alpha(u) = \sinh u = \frac{e^u - e^{-u}}{2}$.

$$\alpha(u) = \frac{u}{u+1}.$$



Fix a large n , set $h = \frac{1}{n}$.
 $x_0 = 0, x_1 = h, \dots, x_i = ih, \dots, x_n = nh = 1$.

$$u(x_{i \pm 1}) = u(x_i \pm h)$$

Appr. $u(x_i)$ by $u_i \in \mathbb{R}$.

$$= u(x_i) \pm u'(x_i)h + \frac{1}{2}u''(x_i)h^2 \pm \frac{1}{6}u'''(x_i)h^3 + O(h^4)$$

$$\frac{u(x_{i-1}) + u(x_{i+1}) - 2u(x_i)}{h^2} = \underbrace{u''(x_i)}_{\sim u''(x_i)} + O(h^2).$$

$$\frac{-u_{i-1} - u_{i+1} + 2u_i}{h^2} + \alpha(u_i) = f_i, \quad f_i = f(x_i)$$

$$\# \text{eq.} = \# \text{unknowns} = n-1. \quad u_0 = u(0) = 0, \quad u_n = u(1) = 0.$$

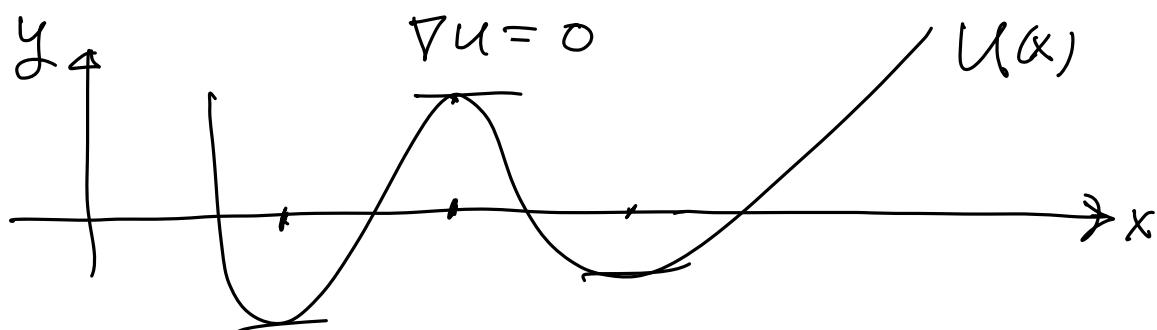
$$F_i(u_1, \dots, u_{n-1}) = -u_{i-1} + 2u_i - u_{i+1} + h^2\alpha(u_i) - h^2f_i \quad (i=1, \dots, n-1)$$

In 3D with (x, y, t) , $h = \frac{1}{N} = \frac{1}{100}$, $n = 10^6$.
A million equations of a million unknowns.

3. Finding (local) minimum or maximum point of a (multivariable) function

$$U: \mathbb{R}^n \rightarrow \mathbb{R}, \min_x U(x) \iff \nabla U(x) = 0$$

$$F(x) = 0, F = \nabla U: \mathbb{R}^n \rightarrow \mathbb{R}^n$$



Solution methods: Iterative methods

Initial guess $x_0 \Rightarrow x_1 \Rightarrow x_2 \Rightarrow \dots x_n = \Phi(x_{n-1})$

Or: initial guesses $x_0, x_1 \Rightarrow x_2 \Rightarrow x_3 \Rightarrow \dots$

$$x_{n+1} = \Phi(x_n, x_{n-1}).$$

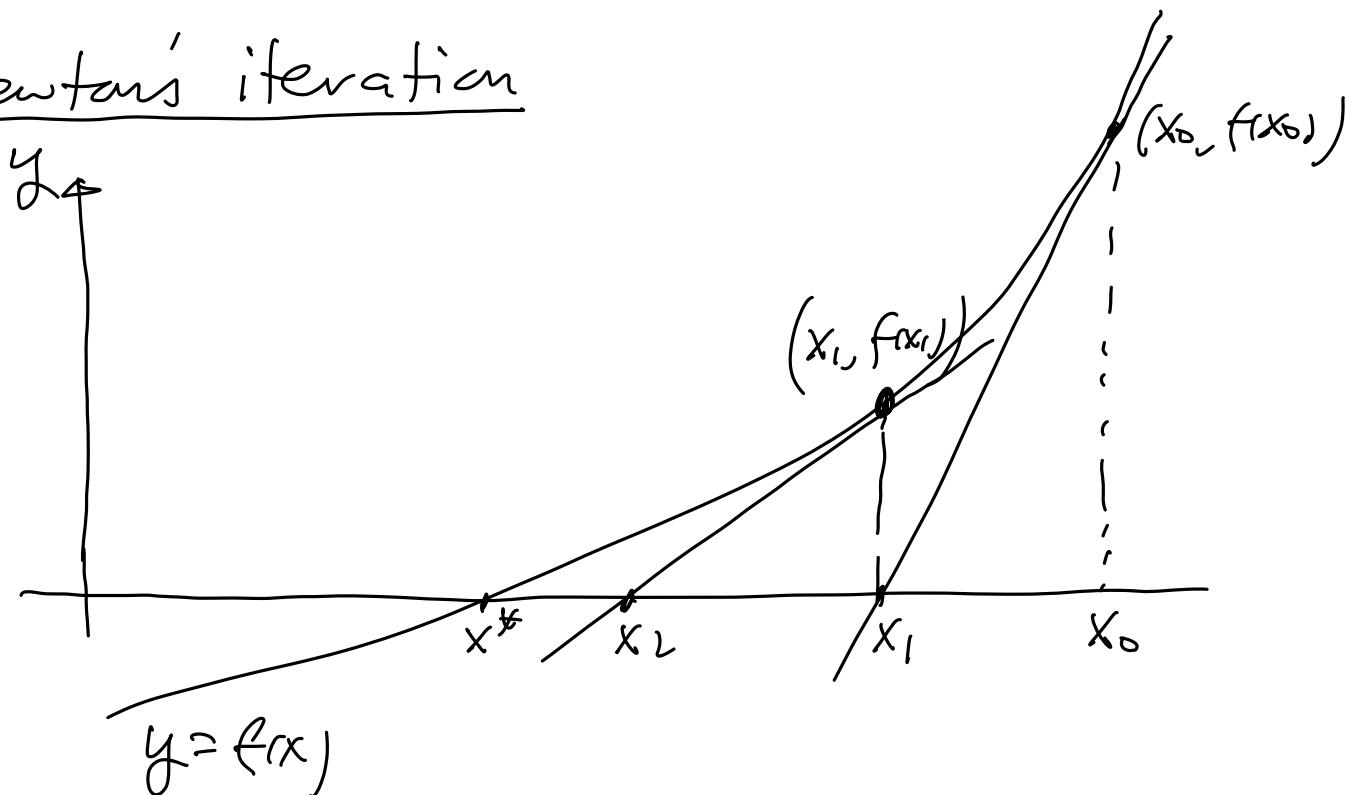
Questions:

- how to construct Φ from F ?
- convergence? $\lim_{n \rightarrow \infty} x_n = x^*$.
- accuracy? Order of convergence?

We will cover

- Ⓐ Newton's iteration
- Ⓑ Fixed-point iteration
- Ⓒ The gradient-descent method for optimization

Newton's iteration



Given x_0 . Tangent line at $(x_0, f(x_0))$.

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Set $y = 0$ to get x_1 .

$$0 = f(x_0) + f'(x_0)(x_1 - x_0)$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Newton's iteration:

$$\text{Given } x_0, \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n=0, 1, \dots)$$

Example $f(x) = x^2 - 2$. $x^* = \sqrt{2} \approx 1.41421356237$, $f'(x) = 2x$.

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n^2 + 2}{2x_n} = \frac{1}{2} \left(x_n + \frac{2}{x_n} \right)$$

$$x_0 = 99, \quad x_1 = 49.5101010, \quad x_2 = 24.7752484$$

$$x_3 = 12.4279871, \quad x_4 = 6.2944571$$

$$x_5 = 3.30609848, \quad x_6 = 1.955521$$

[5]

$$x_7 = 1.4891331, \quad x_8 = 1.416098193$$

$$x_9 = 1.41421482, \quad x_{10} = 1.414213562$$

$$|x_{10} - x_*| \leq 10^{-8}, \quad \frac{|x_{10} - x_*|}{|x_*|} \leq 10^{-7}.$$

Compare with the bisection method.

$$[0, 2], \quad f(0) = -2 < 0, \quad f(2) = 2 > 0, \quad x_1 = \frac{1}{2}(0+2) = 1$$

$$f(1) < 0, \quad [1, 2], \quad x_2 = \frac{1}{2}(1+2) = 1.5$$

$$f(1.5) > 0, \quad [1, 1.5], \quad x_3 = \frac{1}{2}(1+1.5) = 1.25$$

$$f(1.25) < 0, \quad [1.25, 1.5], \quad x_4 = \frac{1}{2}(1.25+1.5) = 1.375$$

$$f(1.375) < 0, \quad [1.375, 1.5], \quad x_5 = \frac{1}{2}(1.375+1.5) = 1.4375$$

$$f(1.4375) > 0, \quad [1.375, 1.4375], \quad x_6 = \frac{1}{2}(1.375+1.4375) = 1.40625$$

$$f(1.40625) < 0, \quad [1.40625, 1.4375], \quad x_7 = 1.421875$$

$$f(x_7) > 0, \quad [1.40625, 1.421875], \quad x_8 = 1.4140625$$

$$f(x_8) < 0, \quad [1.4140625, 1.421875], \quad x_9 = 1.41796875$$

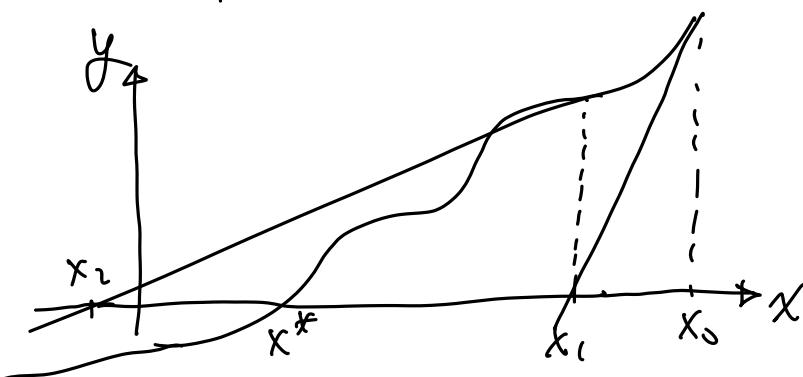
If start with $[\sqrt{2}-\varepsilon, \sqrt{2}+\varepsilon]$ e.g. $\varepsilon = 10^{-20}$.

$$x_k = \sqrt{2} - \varepsilon + 2^{-k} \quad |x_k - x^*| = 2^{-k} - \varepsilon < 10^{-8}$$

$$2^{-k} < 10^{-8}. \quad k \approx 26 \text{ or } 27.$$

Issues

① Behavior of f near x^* .



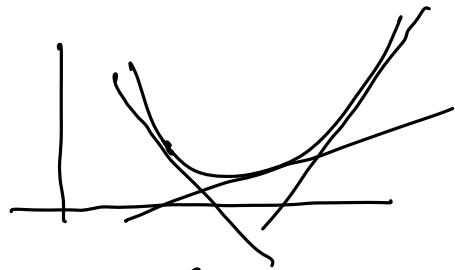
Require:

$f(x)$: smooth

$f'(x)$ increasing
or decreasing
in $[a, b]$.

$$f' \uparrow \Leftrightarrow f'' \geq 0 \Leftrightarrow f \text{ is convex.}$$

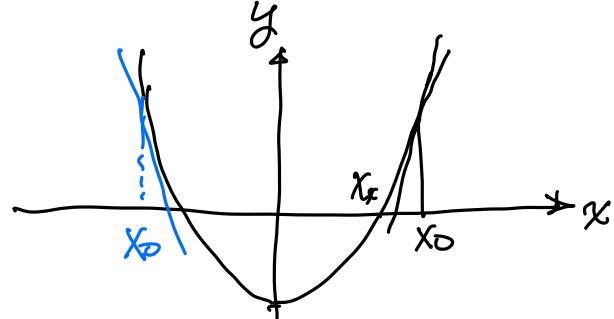
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in [a, b], \quad \forall \lambda \in [0, 1]. \quad [6]$$



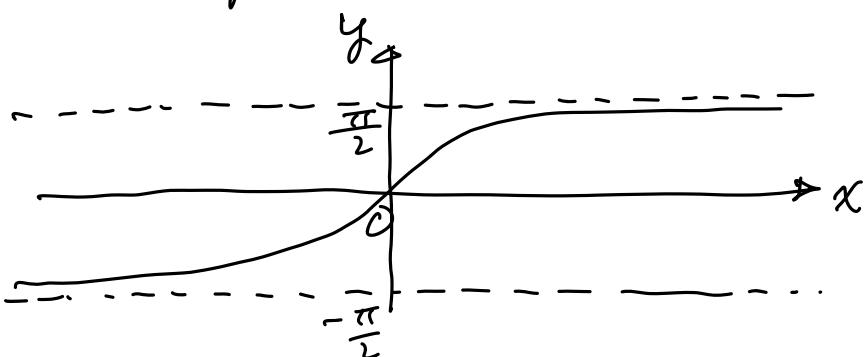
every tangent line is
below the graph of $y = f(x)$.

① Initial point x_0 .

Example $f(x) = x^2 - 2$



Example $f(x) = \arctan x, x \in (-\infty, \infty), x_0 = 0$.



$$f'(x) = \frac{1}{1+x^2}.$$

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

$$\text{Newton's: } x_{n+1} = x_n - (1+x_n^2) \arctan x_n \quad (n=0, 1, \dots)$$

If $|x_0| \ll 1$, i.e., x_0 is very "close" to $x^* = 0$, then $x_n \rightarrow x^*$.

If $|x_0| \gg 1$, e.g., $|x_0| \geq \max(1, \frac{2|x_0|}{1+x_0^2})$ then $|x_n| \rightarrow \infty$.