_/_/__

Problem Classification and Sol^n concepts —

[Week - 2]

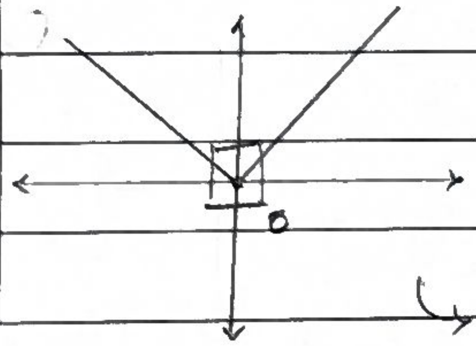
Problem Setup :

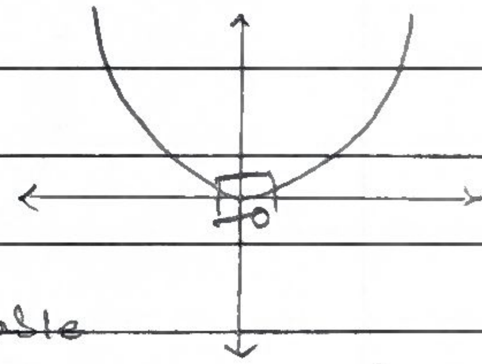
Unconstraint optimization :

$$\min_{x \in R^n} f(x), \quad f : R^n \to \mathbb{R}$$

$y = x^2$

$y = |x|$





↳ f is not differentiable at minimizer.

f is differentiable.

For next few lectures we will asume :

→ f is differentiable at x^*

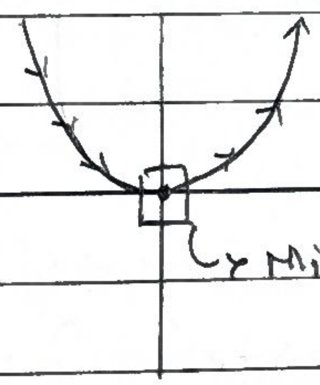
f' is continuous.

Fermat's theorem (1D): If f is differentiable at x^*

$$\boxed{f'(x^*) = 0}$$

- $f'(x) > 0$
  ↳ f is increasing

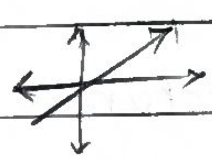
$f'(x) < 0$
↳ f is decreasing



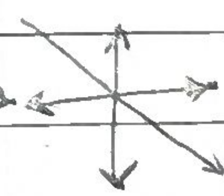
$y = x^2$

$\frac{d}{dx}(y) = 2x \begin{cases} - & x < 0 \\ + & x > 0 \end{cases}$

↳ Minimizer
↳ $\begin{cases} f' \text{ is } -ve \text{ in one side} \\ f' \text{ is } +ve \text{ in other side} \end{cases}$

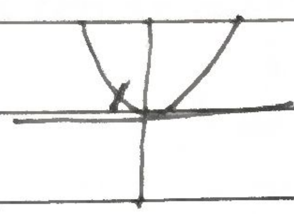
$f(x) = x$ → 

$f'(x) = 1$

$f(x) = -x$ → 

$f'(x) = 1$

$$f'(x^*) = \lim_{t \to 0} \frac{f(x^* + t) - f(x^*)}{t}$$



We know
$f(x^* + t) \geq f(x^*)$

$$\frac{f(x^* + t) - f(x^*)}{\boxed{t}}$$

-ve +ve

if,
$t < 0$

$$\lim_{t \to 0^-} \frac{f(x^* + t) - f(x^*)}{t} \leq 0$$

if,
$t > 0$

$$\lim_{t \to 0^+} \frac{f(x^* + t) - f(x^*)}{t} \geq 0$$

But we know,
RHL = LHL,
and it only happens when

$$\boxed{F(x^*) = 0}$$

But Fermat's rule is not sufficient.

Eg:

$$f(x^*) = (x-1)^2$$
$$f'(x^*) = 2x - 2$$
$$f'(1) = 0 \qquad f'(0) = -2$$

$f'(x^*) = 0$ may be min, max or neither

Here,

$f'(x^*) = 0$, x^* is called critical point or stationary point

As we are unable to conclude anything

__/__/__

So, we now need ~~suf~~ sufficiency criteria,
Here comes the power of convexity.

Convexity

↳ Functions where
first order derivatives
are also sattisfied

But for now we focus on second order cond^{n}s.

~~First~~ first step always be $f'(x^*) = 0$

Now,

- $f''(x^*) > 0 \to$ +ve curvature
  $f''(x^*) < 0 \longrightarrow$ -ve curvature.

if $f''(x^*) > 0 \implies$ strict local minimum
if $f''(x^*) < 0 \implies$ strict local maximum
if $f''(x^*) = 0 \to$ Inclusive (check higher derivatives / Taylor expansion)

strict means either '<' or '>'
not '≥' or '≤'

Multivariable : Fermat's Rule

similar to previous only in terms of gradient.

$$\boxed{\Delta f(x^*) = 0}$$

Intuition: [redacted]

We can break it into multiple 1D fn.

$$\frac{\partial f(x_0)}{\partial x_i} = \lim_{t \to 0} \frac{f(x_0 + t e_i) - f(x_0)}{t}$$

ith codinate

At any, dirn;

$$f(x_0, d) = \lim_{t \to 0} \frac{f(x_0 + t d) - f(x_0)}{t}$$

vector

|dirn|.

__/__/__

And, any directional derivative can be calc. using the following formula,

$$\boxed{\nabla f(x^*)^T d = 0}$$

It is nothing but inner product.

$$\langle \nabla f(x^*)\ d \rangle \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = a_1 d_1 + \cdots a_n d_n.$$

Illustrations in R^2:

① ~~~~ $f = (x-1)^2 + 2(y+0.5)^2$

$$\frac{\partial f}{\partial x} = 2(x-1)$$

$$\frac{\partial f}{\partial y} = 4(y+0.5)$$

$\nabla f = 0$ at ?

if $\boxed{x = 1,\ y = -0.5}$

$$\frac{\partial f}{\partial x} = 0$$

$$\frac{\partial f}{\partial y} = 0$$

/ /

② $f = -(x^2 + y^2)$

$\frac{\partial f}{\partial x} = -2x$

$\frac{\partial f}{\partial y} = -2y$

$\nabla f = 0$ at

$\boxed{x = 0 \quad , y = 0}$

Multivariable : Second-order sufficiency

let x^* be a stationary point ($\nabla f(x^*) = 0$) and $H = \nabla^2 f(x^*)$

↓ Hessian.

$H \succ 0$ strict minimum local.

$H \prec 0$ strict local maximum.

$H = 0 / \infty$ Saddle point

or

Quadratic model :

$$f(x) \approx f(x^*) + \frac{1}{2}(x - x^*)^T H(x - x^*)$$

$A_{n \times n}$ → Positive definite when,

$$\begin{cases} x^T A x \geq 0 \quad \forall x \\ \quad \hookrightarrow \text{Positive semidefinite} \\ x^T A x > 0 \quad \forall x \in x - \{0\} \\ \quad \hookrightarrow \text{positive definite} \end{cases}$$

$[\quad]_{1 \times n} \quad [\quad]_{n \times n} \quad [\quad]_{n \times 1}$

↳ resultant $\boxed{1 \times 1}$ matrix

$$x^T f''(x_0) x \geq 0$$
$$f''(x_0) x^2 \geq 0$$
$$f''(x_0) \geq 0$$

Basically
positive definite → strict local minimum
negative definite → strict local maximum

__/__/__

Eg:

$F(x,y) = x^2 + xy + y^2$. Then

$$\nabla F = \begin{bmatrix} 2x + y \\ x + 2y \end{bmatrix} = 0 \Rightarrow (x^*, y^*) = (0,0)$$

$$H = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Eigen vector:

$A_{n \times n}$ $\lambda \to$ Eigen vector $\vec{v}$.

$$\det [A - \lambda I] = 0$$

simple just do,

$$\left| \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} \right|$$

det

then equate it to '0'

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix}$$

$\Rightarrow (2-\lambda)^2 - 1 = 0$

$\Rightarrow (2-\lambda)^2 = 1$

$\Rightarrow 2-\lambda = \pm 1$

$\Rightarrow 2-\lambda = +1$	$2-\lambda = -1$
$\lambda = 1$	$\lambda = 3$

Eigen values of H : $\{1, 3\} \succ 0$

$$H \succ 0$$

$\therefore$ strictly local minimum at $(0,0)$