① We hv learnt various ML algo ⟹ Try to evaluate the best possible combination of hyperparameters which will hv min. amt of error.

ML algo

SL Un.SL

Cl. Reg

Optimization Techniques

② Irreducable error i.e. error that remains even after optimization.

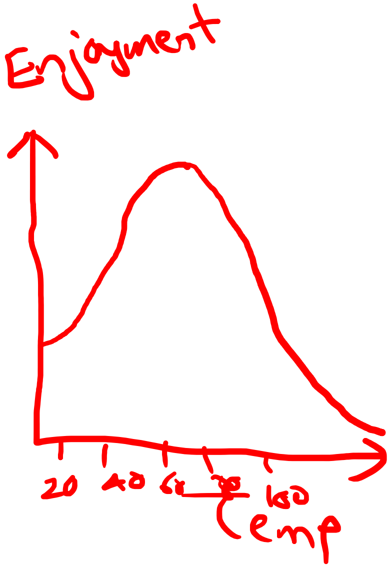
③ Minimization of Error ⟹ Optimization.

④ why do we need Optimization?

To save money i.e. we want our model to give us best possible decision.

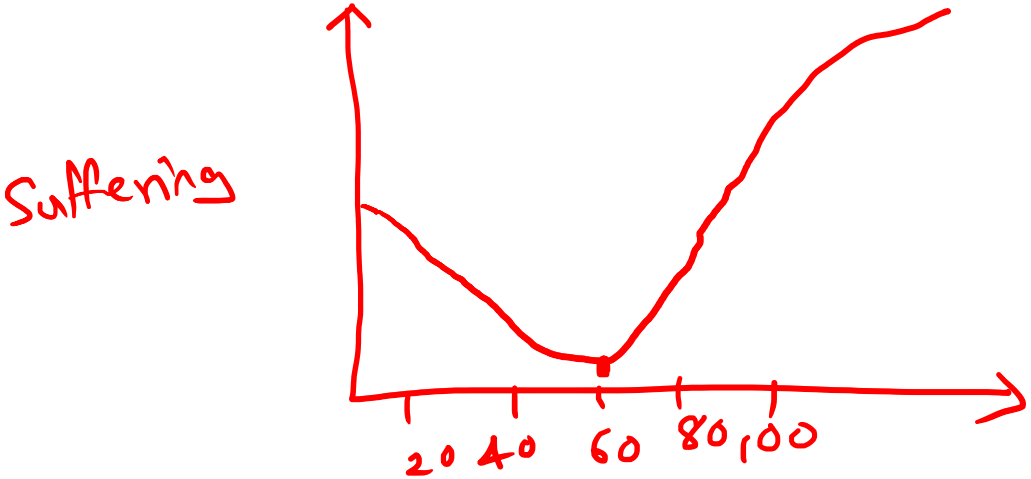
Loved ⇒ Best Temp

Good Morning
Ice89


Enjoyment
20
40
60
80
100
temp

Data Scientist → he flips the problem Upside-down.

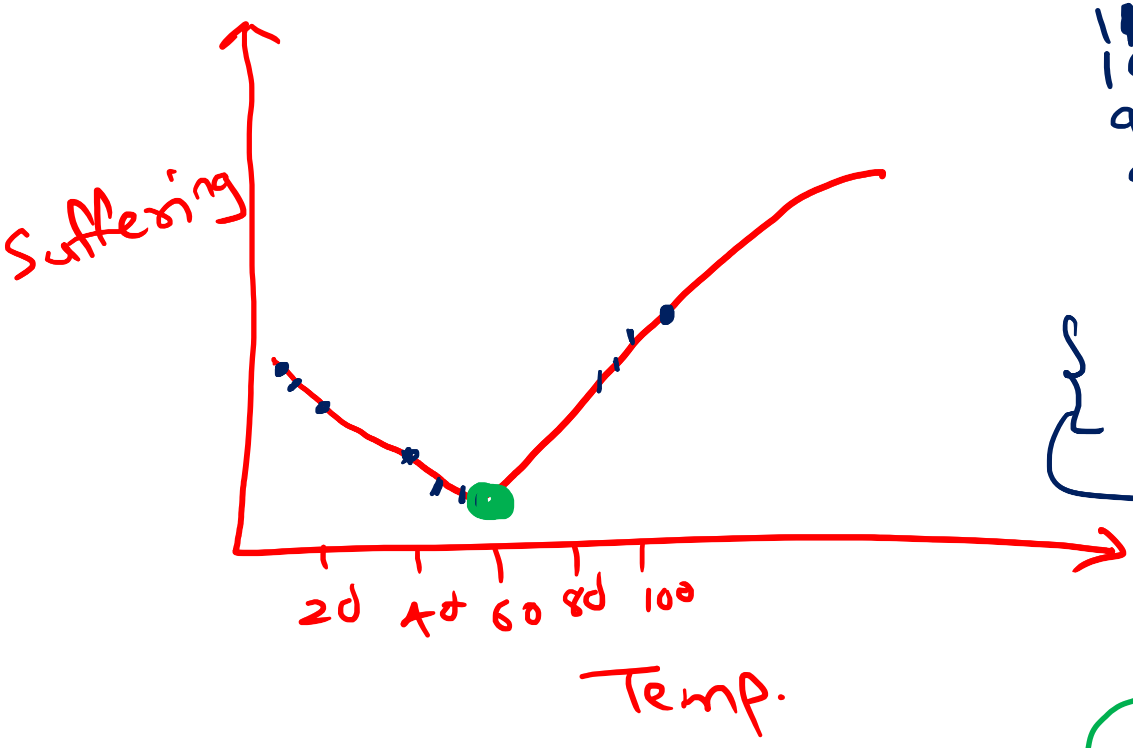
Instead of trying to maximize tea drinking enjoyment,
try to Minimize Suffering while drinking tea.



Ways to Optimize

1. Greedy Search / Exhaustive Search / Grid Search
2. Gradient Descent (3 Types)
3. Robust Methods

Exhaustive Search:



Raghu is at my home

I want ~~to~~ to offer him TEA.

110
100
95
90
⋮
60 — ok to be drunk
55 — it cud be better

59, 58, 57, 56

↓ 55 56,57 ↓ 60

58°

Temp.	Suffering
β	$J(\beta)$
10	–
15	–
20	–
25	–
30	–
⋮	⋮
40	–
45	–
50	–
60	–

Adv: ① Simple
② Global Optima

Disadv:
① Get close to real min bcz of predefined stepsize — never achieve minimum.
② Computationally expensive.

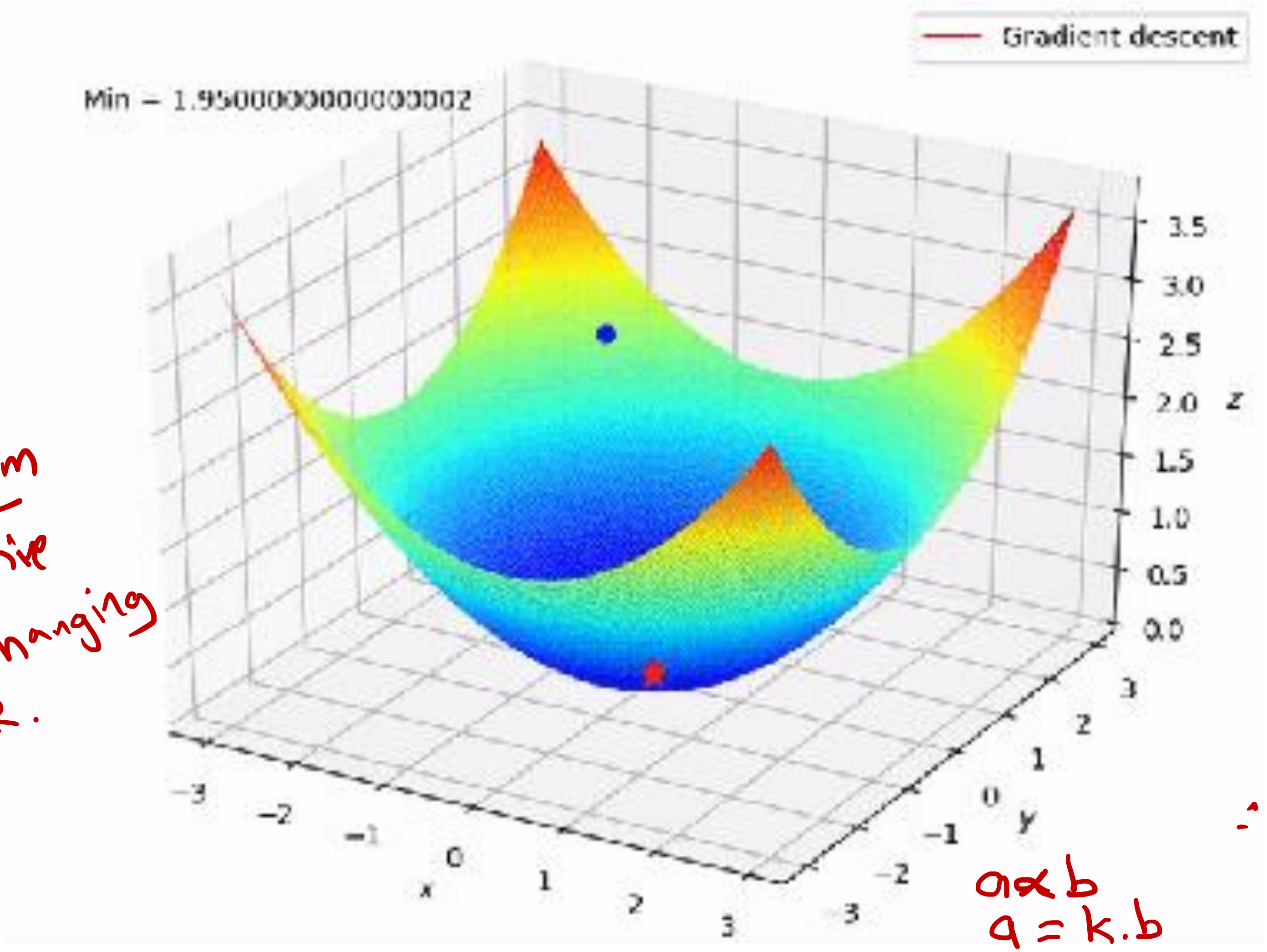
Gradient Descent:

↳ (Slope)

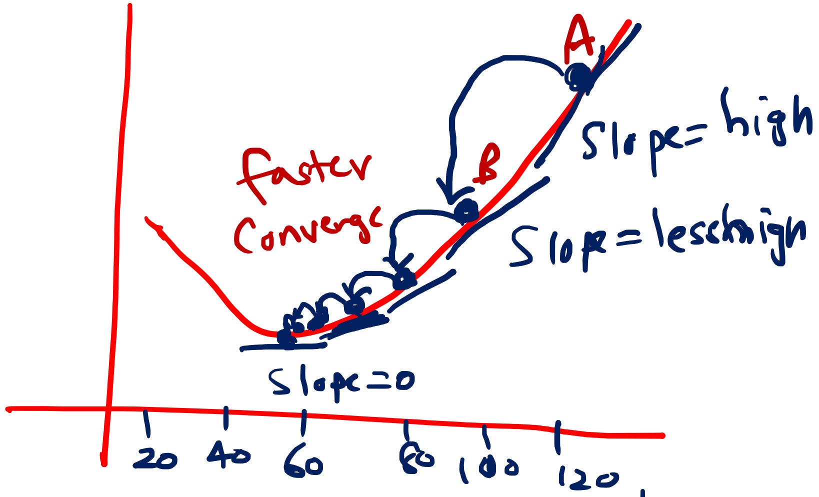
Starting point is always random

Adam

Adaptive
↳ changing
L.R.



$a \propto b$
$a = k.b$



If slope is high, jump higher

If slope is less, jump Closer

$$\boxed{\text{StepSize} \propto \text{Slope}}$$

∴ Stepsize = Learning Rate × Slope

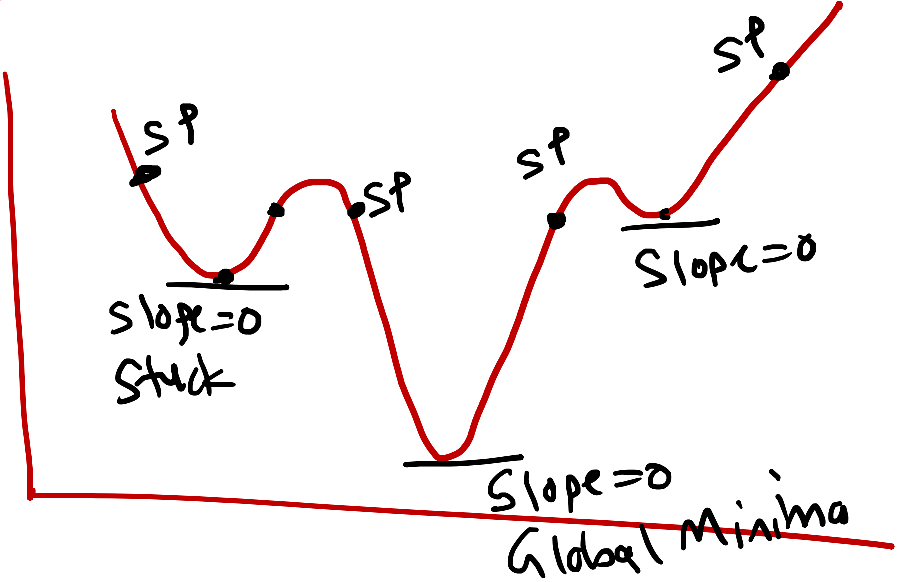
↳ LR = 0.05 (Hyper parameter)

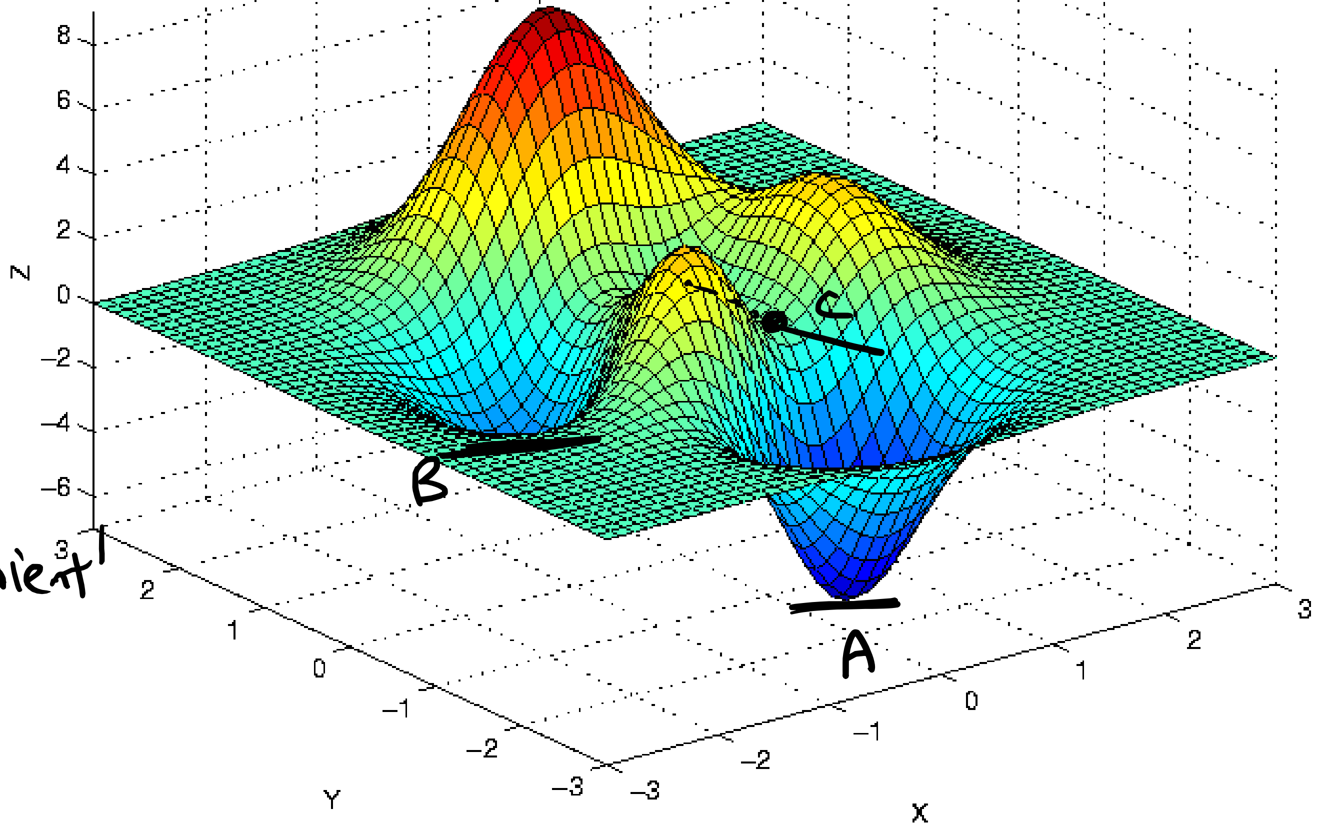
Drawbacks of G.D:

aim → To find ⇒ # A → Global Minima (only one)

B – Local Minima (more than one)

C – Saddle point (more than one)

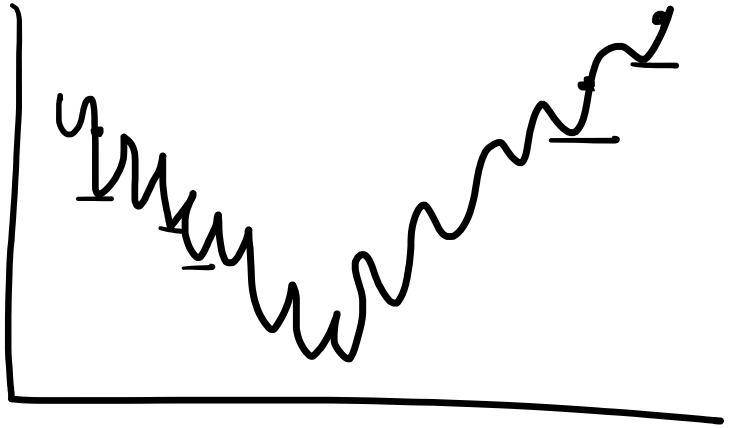




multiple times

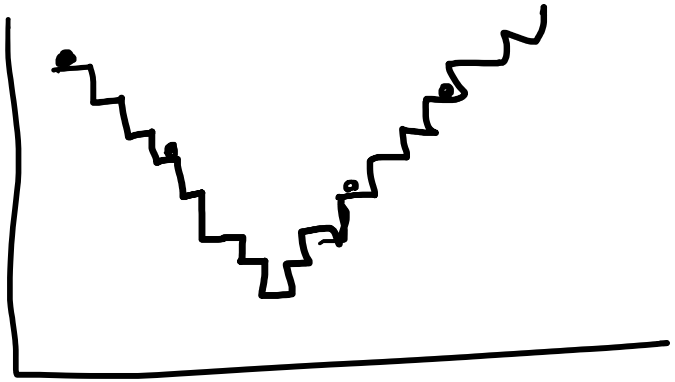
```
from sk.LM import LR
obj = LogR(            method = 'gradient'
obj.fit ( X, y)
```

② Drawback

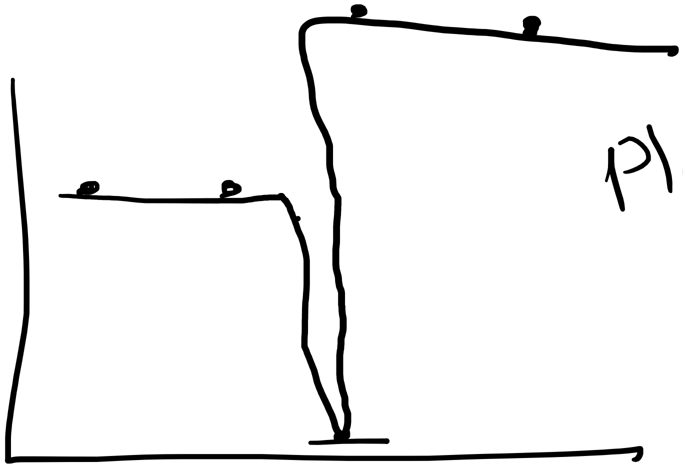


error func. is not smooth

③



④



Plateau

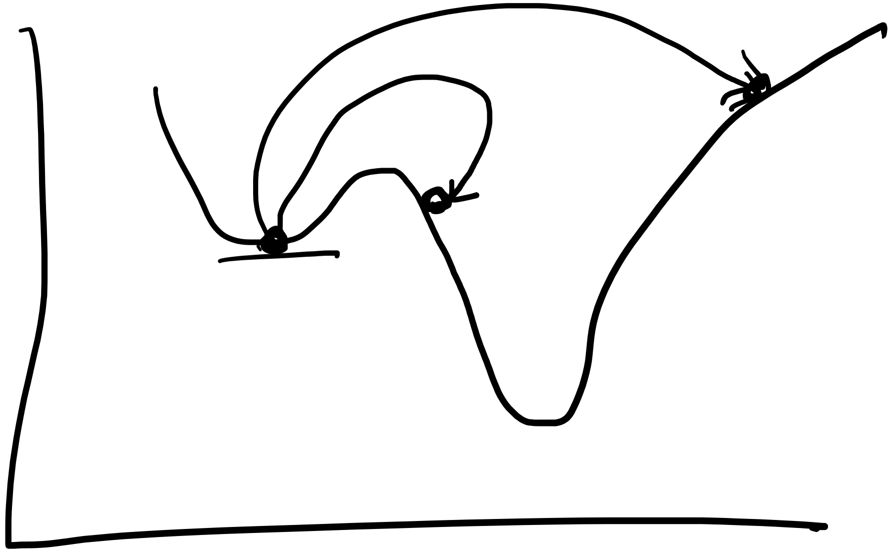
③ Robust Method ! (helpus to come out of local Minima)

Ex.h. Mtd

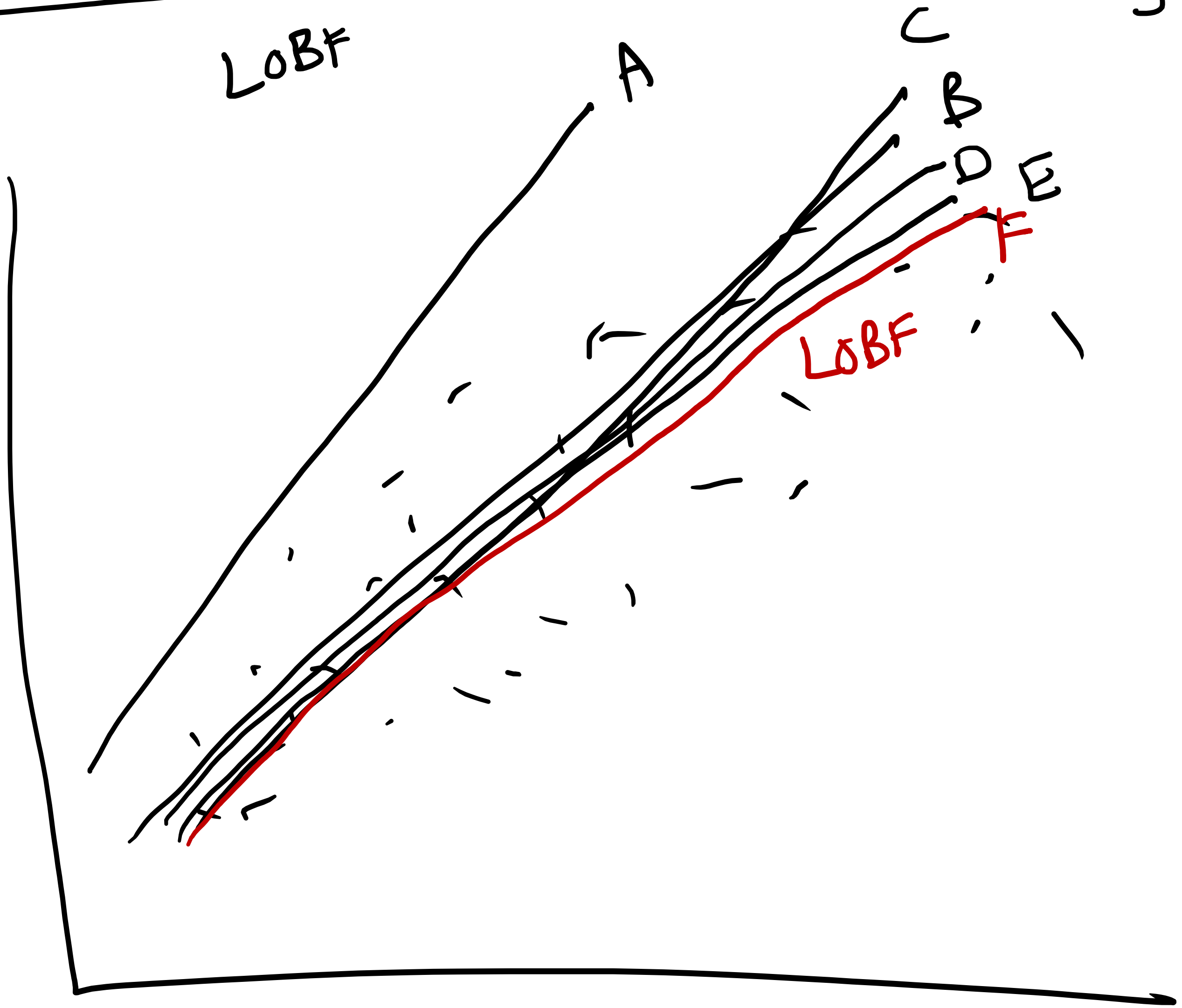
Simulated Annealing / Evolutionary algo)

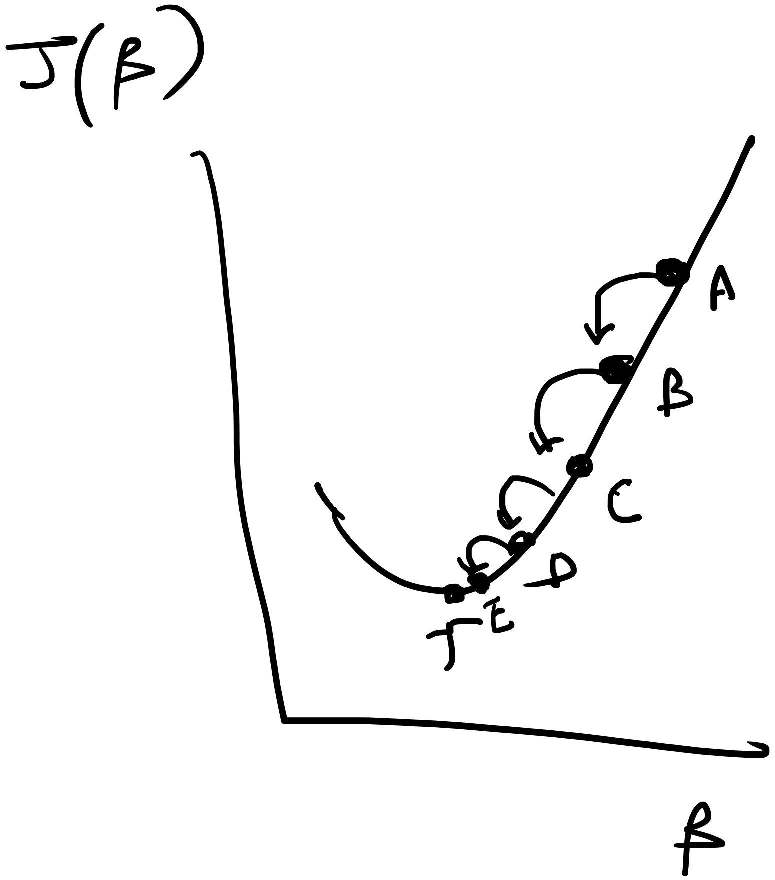
Genetic algo.

Grad . Des.

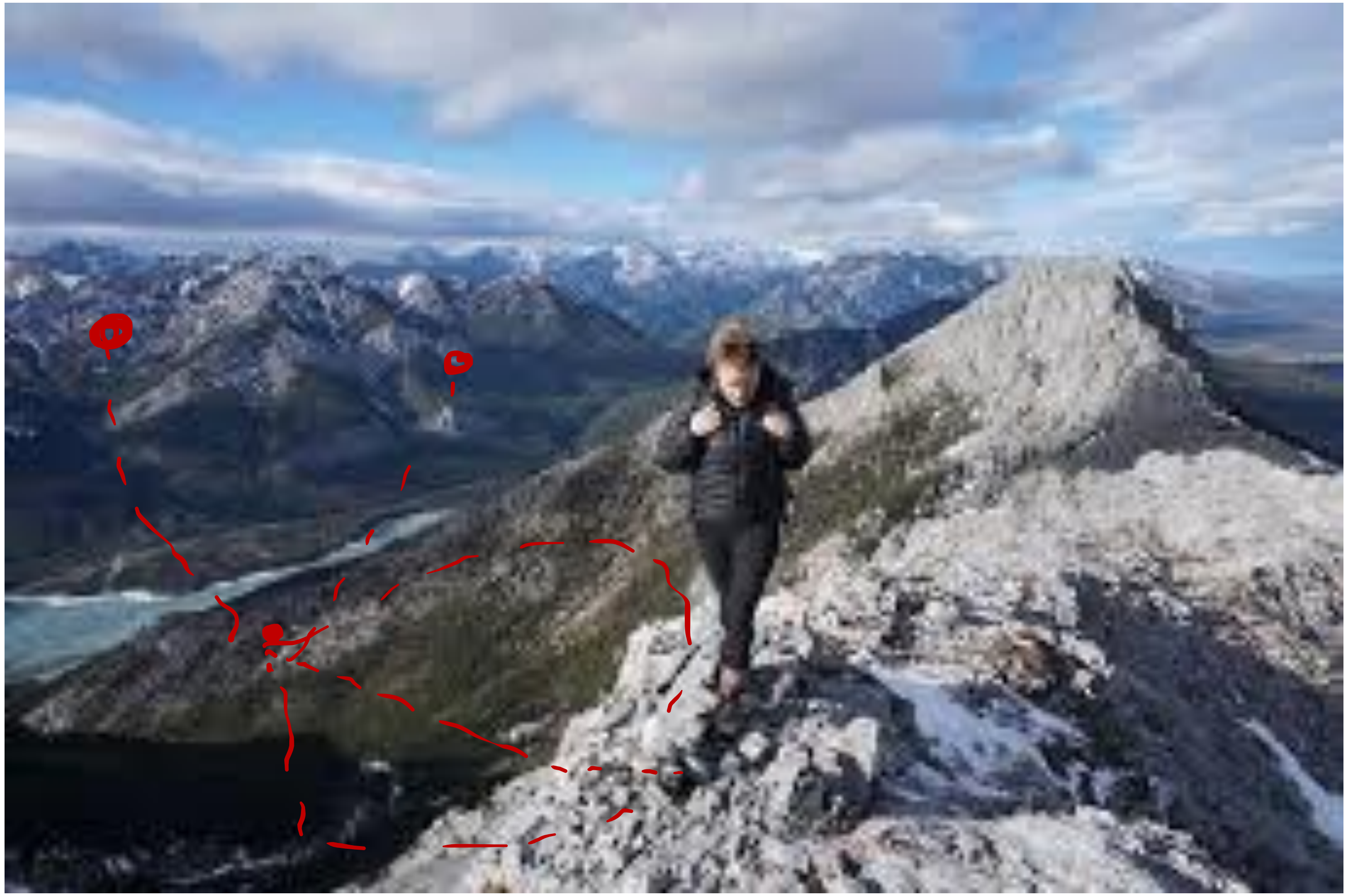


Gradient Descent:





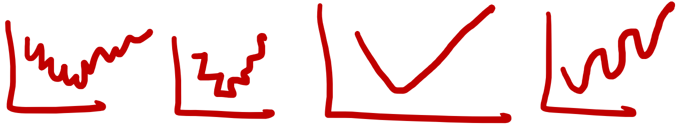
Exhaustive Search:



when u r traveling on foot, you can absolutely get anywhere.

But Time ↑

Gradient Descent: Fast & Efficient



① Smooth & well behaved track

~~GA~~ Robust Mtd: Pickup Trucks

① They dont converge as fast as GD but are definitely better than Ex. Search