LASSO REGRESSION (L1 REGULARISED REGRESSION)

$$\underset{\theta^*}{\text{MINIMIZE}} \quad (y - X\theta)^T (y - X\theta)$$

s.t.

$$|\theta| \leq s$$

USING KKT CONDITIONS

$$\text{MINIMIZE} \quad (y - X\theta)^T (y - X\theta) + \delta^2 |\theta|$$

CONVEX

$|\theta|$ is NOT DIFFERENTIABLE

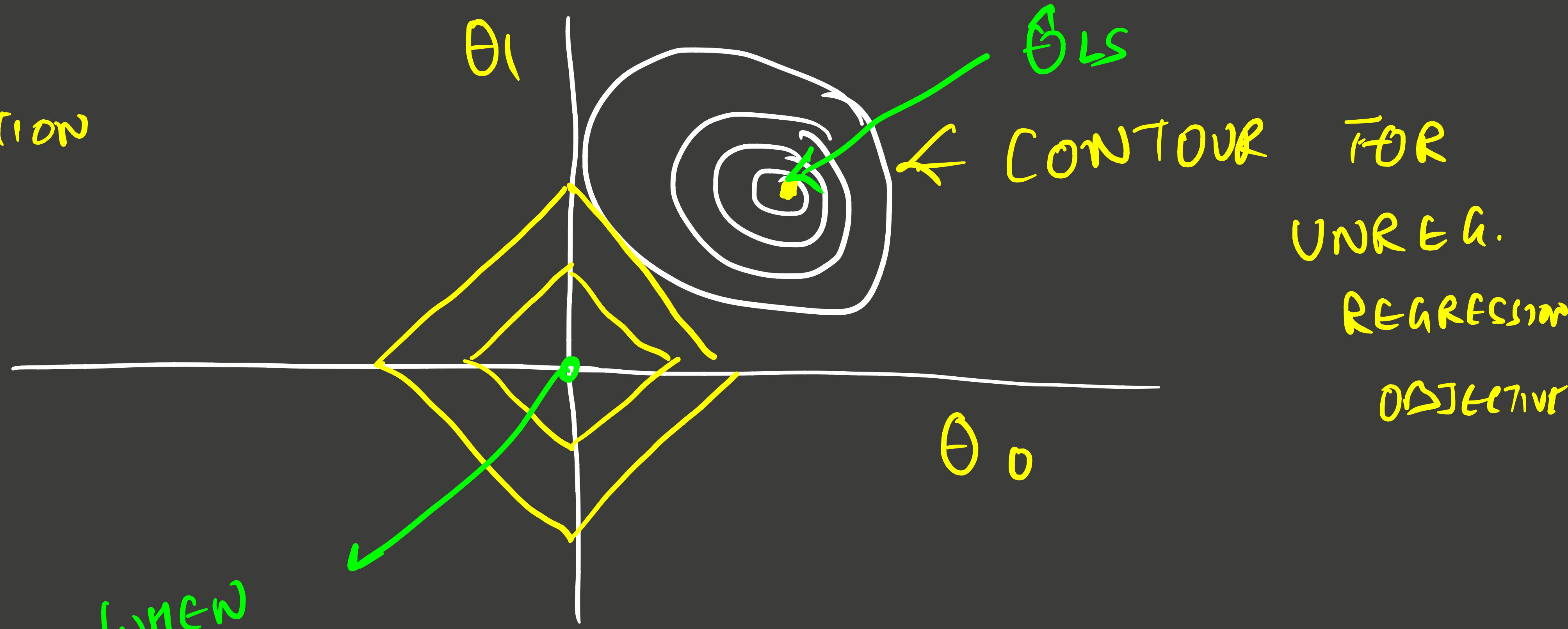
$\therefore$ WE CANNOT SOLVE

$$\frac{\partial}{\partial \theta}\left((y-X\theta)^T(y-X\theta)+\delta^2|\theta|\right)=0$$

HOW TO SOLVE?

* COORDINATE DESCENT

Geometric Interpretation



When $\delta^2 \to \infty$

$\hat{\theta}_{LASSO} \to 0$

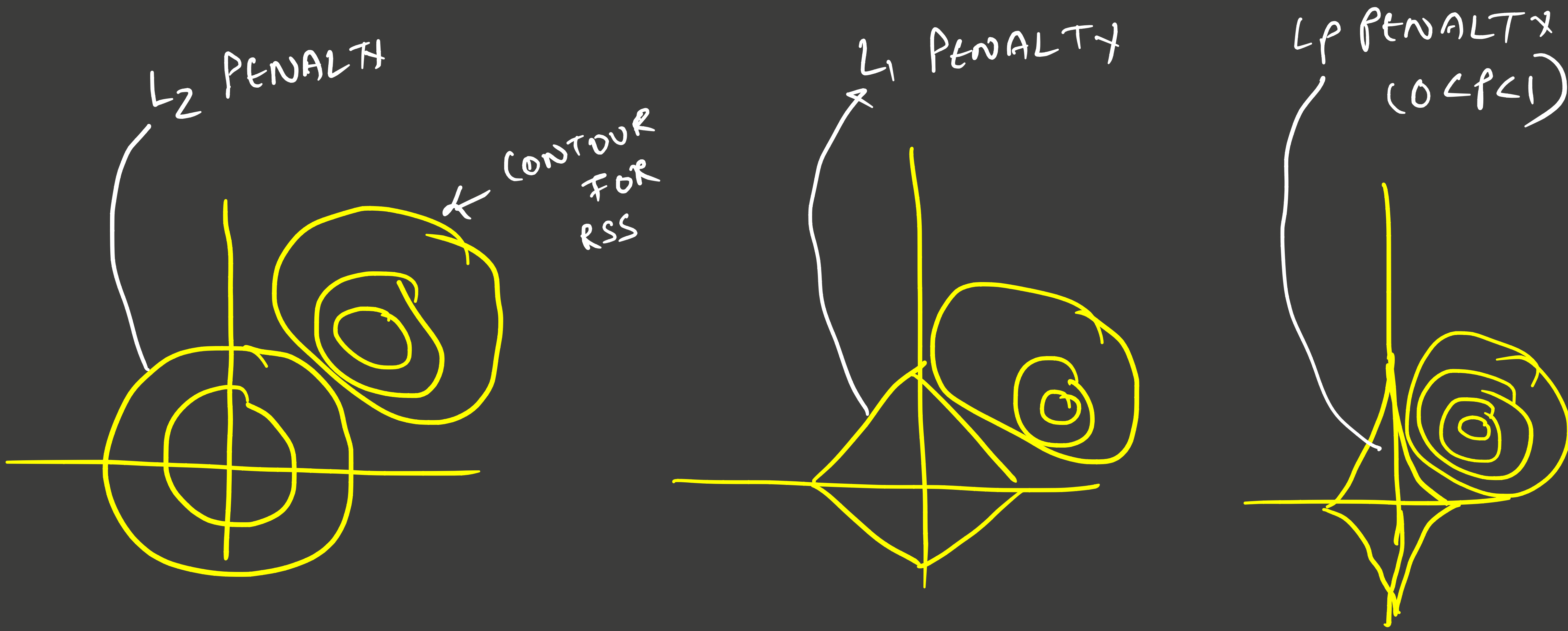
* LASSO → Least Absolute Shrinkage and Selection Operator

* Popular since leads to → sparse solutⁿ

⇓

Many θ_i's are 0!
(or solⁿ lies on some axis)

WHY LASSO GIVES SPARSE SOLN ? (I)



POINTEDNESS of Lp NORM $\xrightarrow{\text{INCREASING}}$

∴ PROBABILITY OF INTERSECTING AT AN AXIS $\xrightarrow{\text{INCREASING}}$

∴ SPARSITY $\xrightarrow{\text{INCREASING}}$

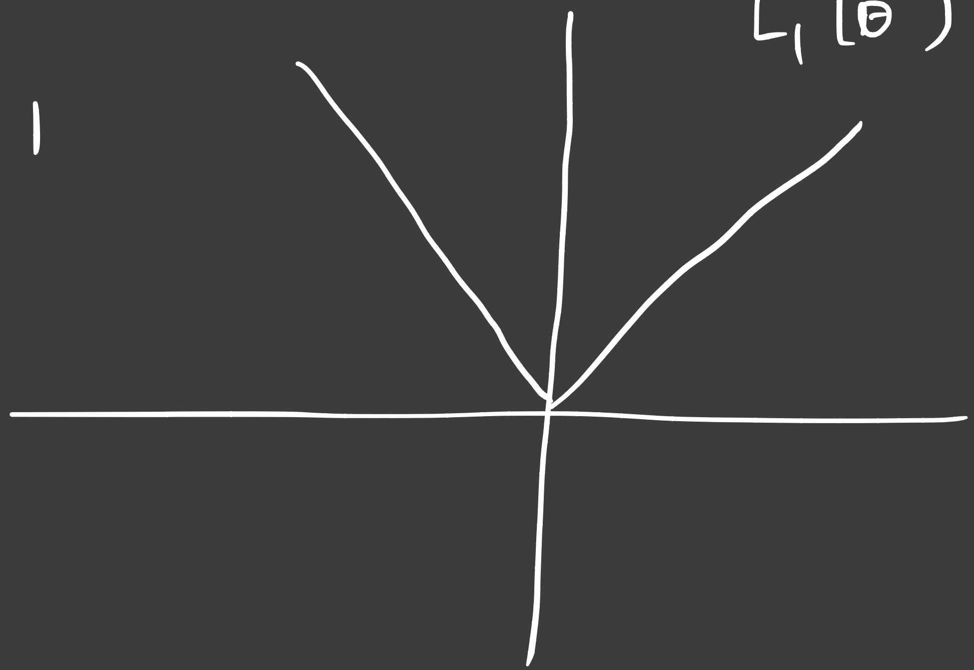
BUT

SOLVING $\xrightarrow{\text{MORE DIFFICULT}}$

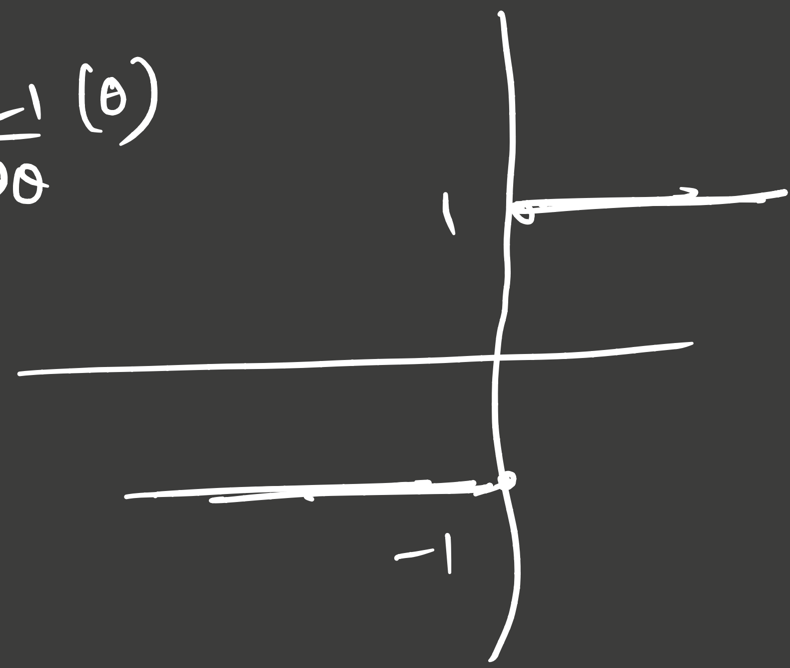
WHY LASSO GIVES SPARSE SOL^N (INTERPRETATION II)

L_1

$L_1(\theta) = |\theta|$

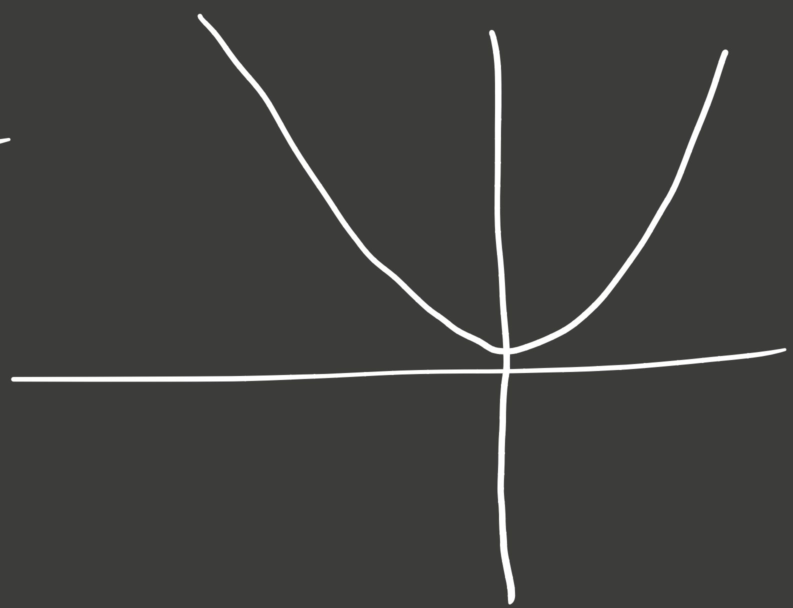


$\frac{\partial L_1(\theta)}{\partial \theta}$

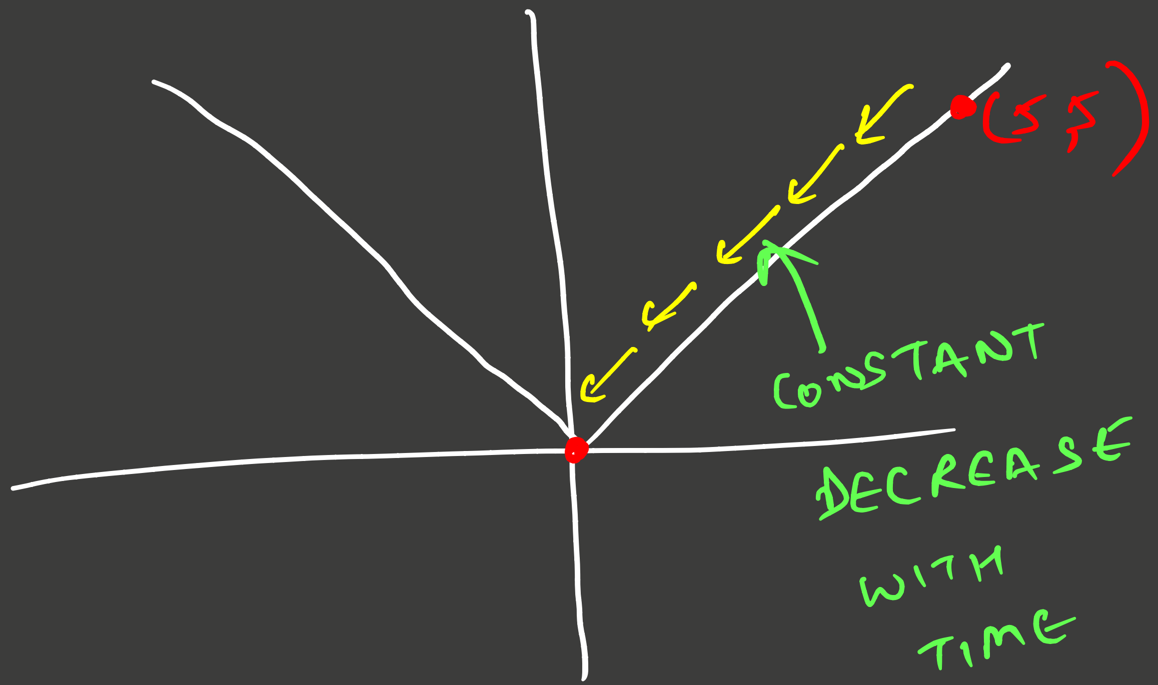


L_2

$L_2(\theta) = \theta^2$

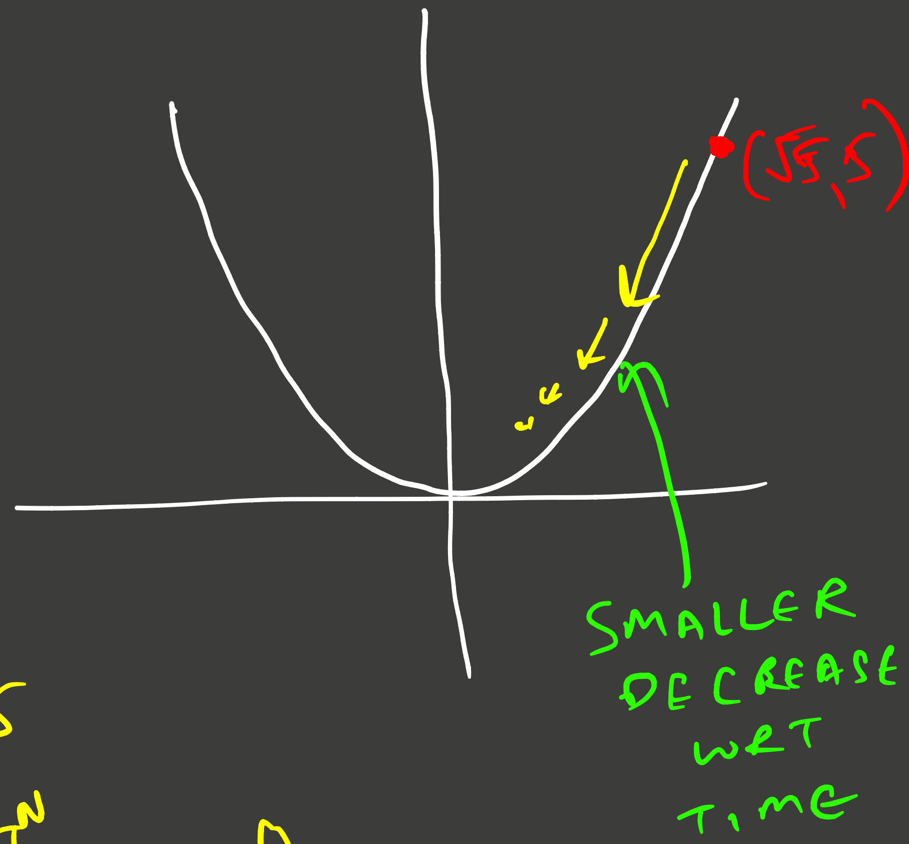


$\frac{\partial L_2(\theta)}{\partial \theta}$



$\alpha = 0.5$

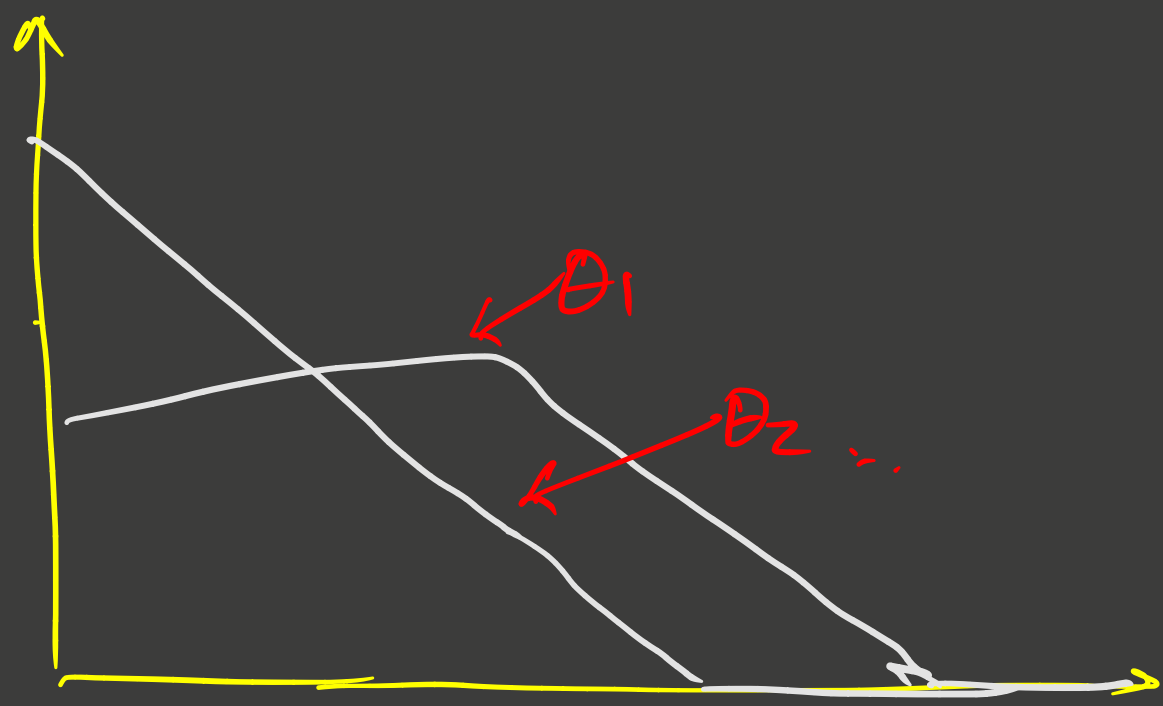
ITERATN	θ
0	5
1	$5 - 0.5 = 4.5$
2	4
3	3.5
⋮	
10	0



$\alpha = 0.5$

ITERATN	θ
0	5
1	$5 - (0.5) * 5 = 2.5$
2	$2.5 - (0.5)(2.5) = 1.25$
3	.625
⋮	
10	$\neq 0$

REGULARISATION PATH



$$\theta_i^2 \to 0$$

as $\delta^2 \to \infty$