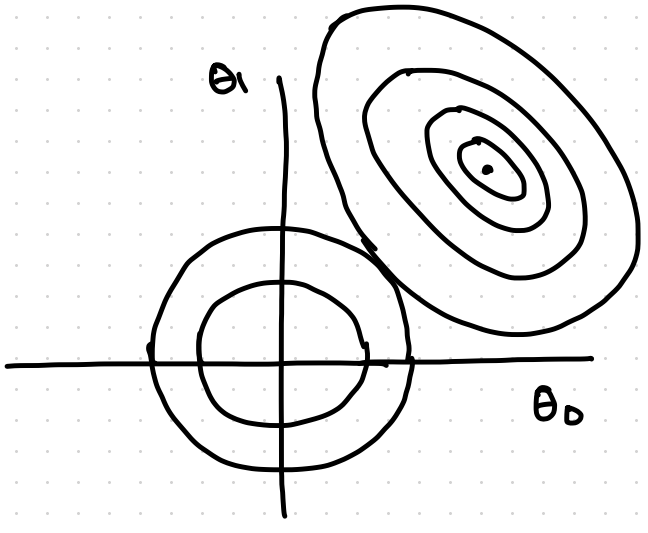
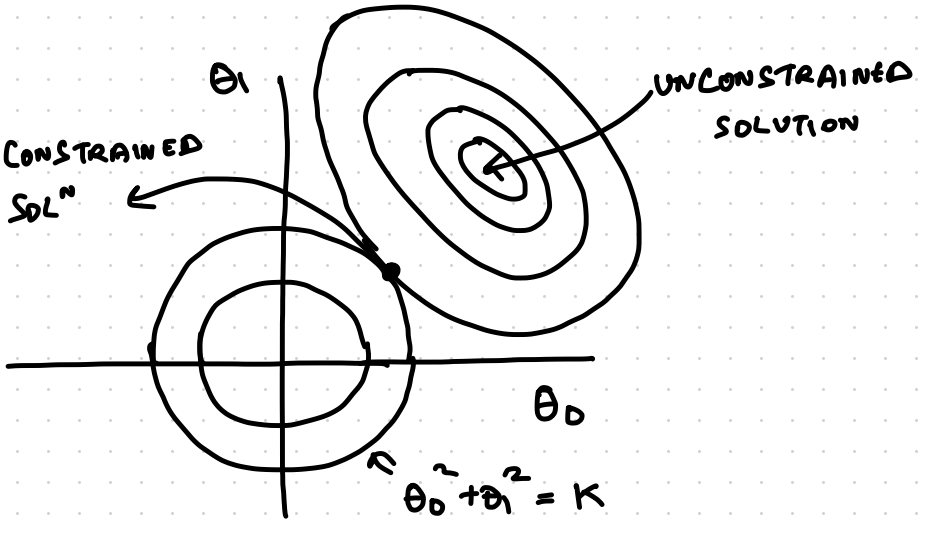
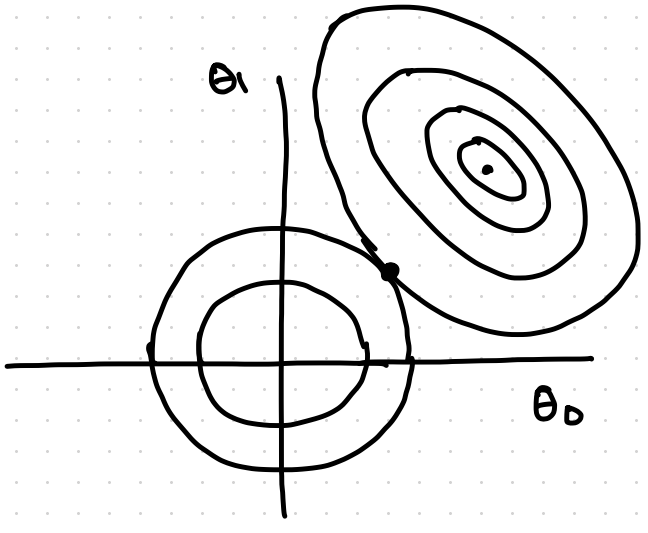
WHY LASSO GIVES SPARSITY

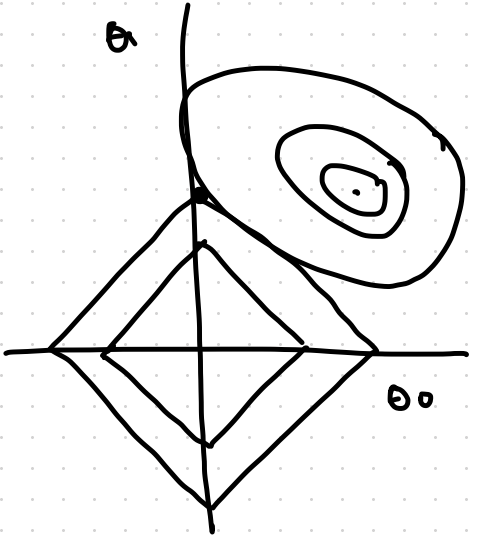
1. GEOMETRIC INTERPRETATION
2. G.D. BASED INTERPRETATION


θ1
θ0

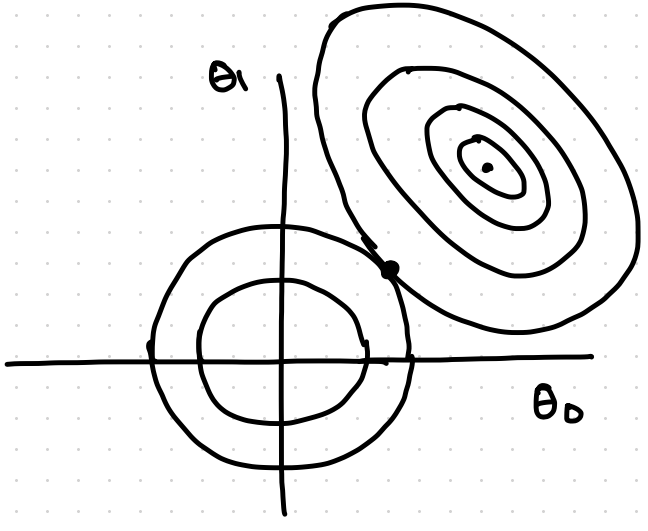

θ1
θ0
UNCONSTRAINED SOLUTION
CONSTRAINED SOL^N
θ0^2 + θ1^2 = K
CONTOUR OF L2 NORM OF θ



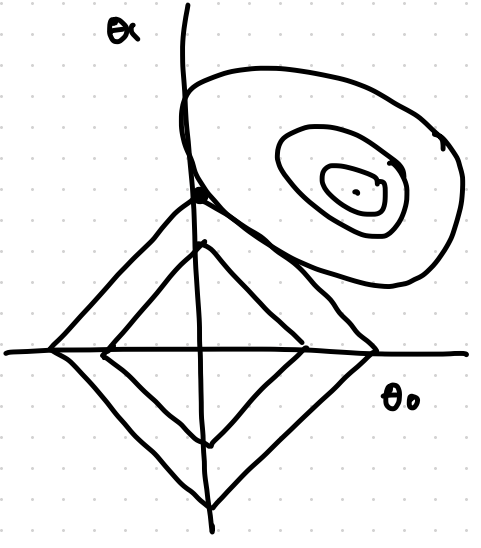
L_2 NORM



L_1 NORM

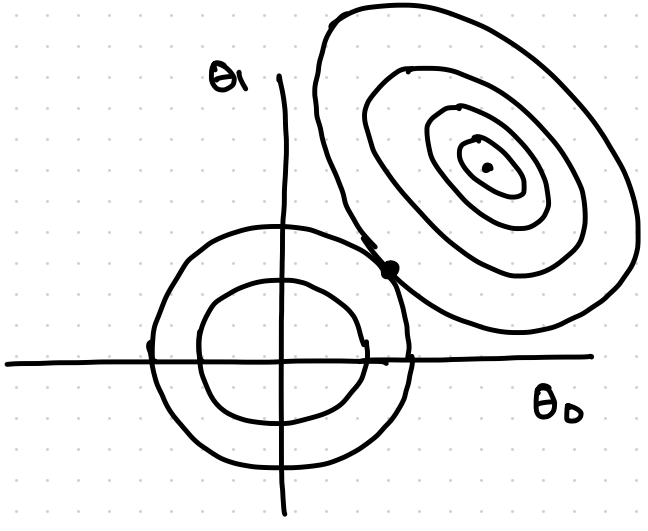


L_2 NORM

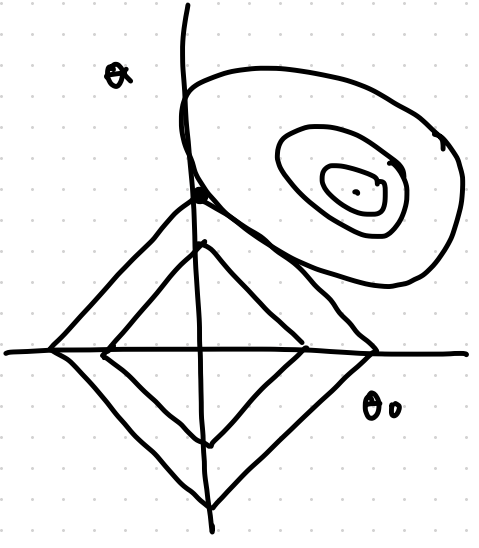


L_1 NORM

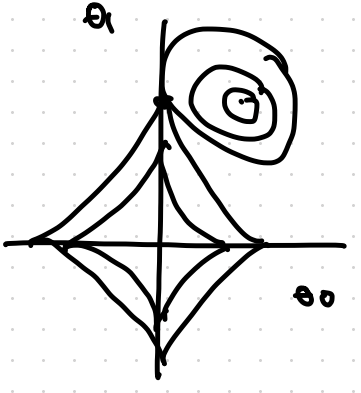
L_p NORM
$(0 < p < 1)$



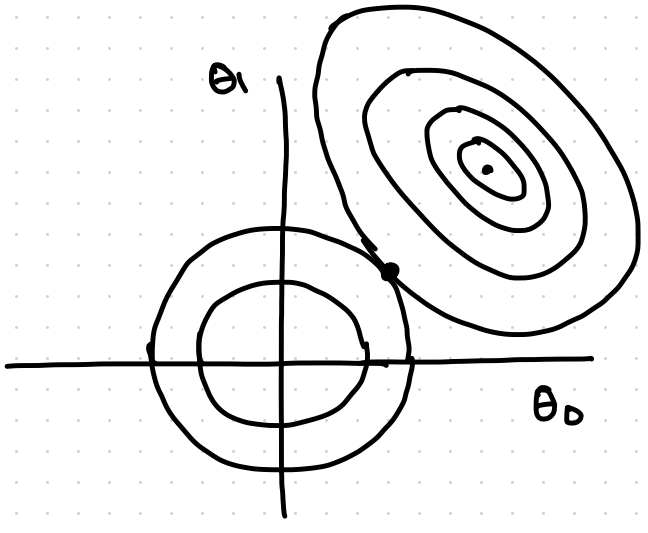
L_2 NORM



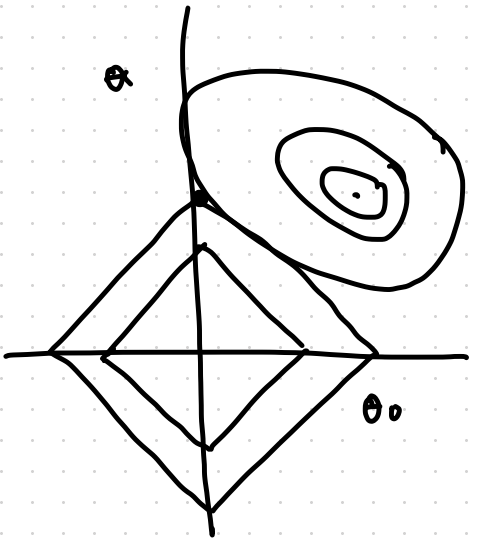
L_1 NORM



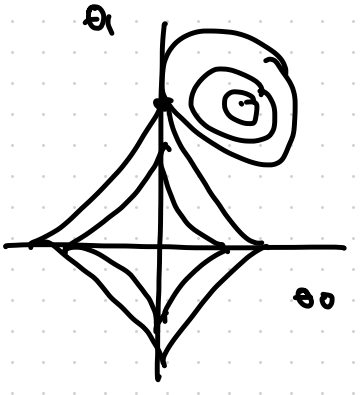
L_p NORM
$(0 < p < 1)$



L_2 NORM



L_1 NORM



L_p NORM
$(0 < p < 1)$

SPARSITY →

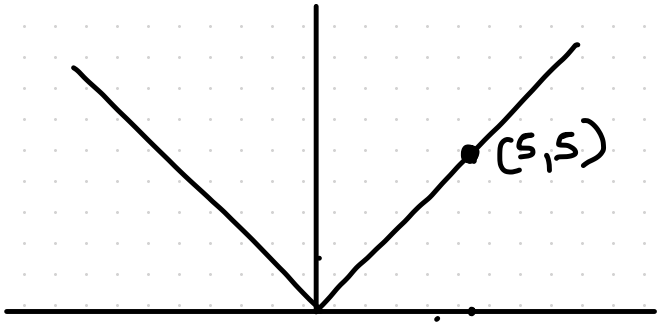
PROB. OF INTERSECTING AXIS →

DIFFICULTY OF SOLVING →

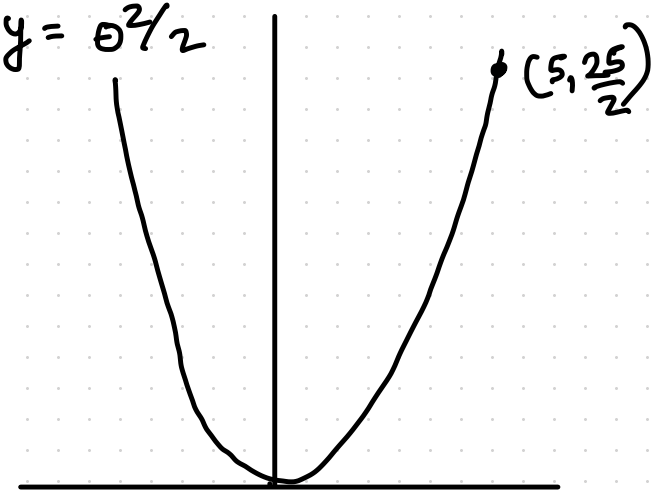
$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)

$y = \theta^2/2$

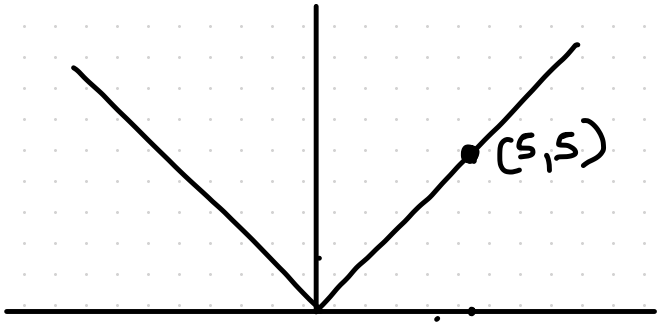
$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



$y = \theta^2/2$

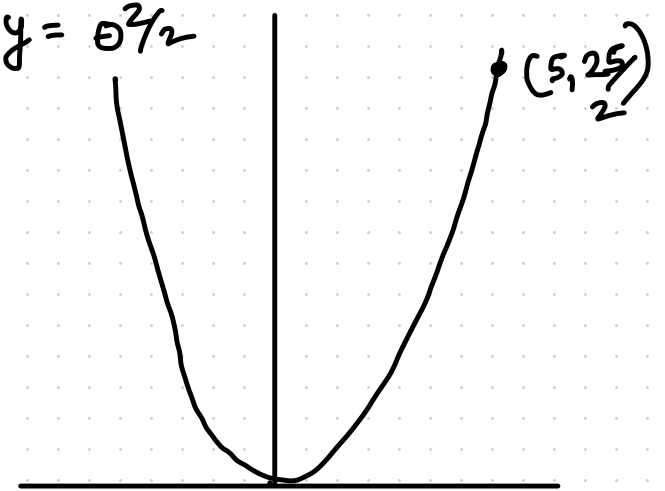


$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



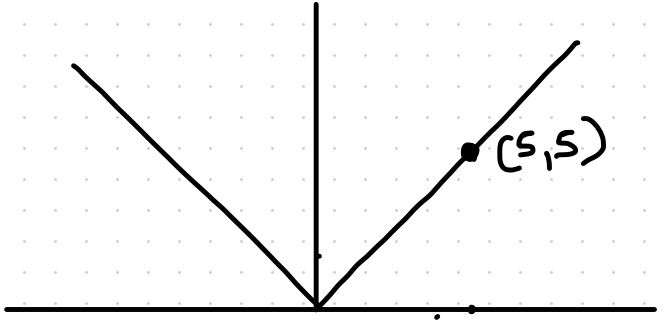
$\frac{\partial y}{\partial \theta} = 1$ (Assume $\theta > 0$)

$y = \theta^2/2$



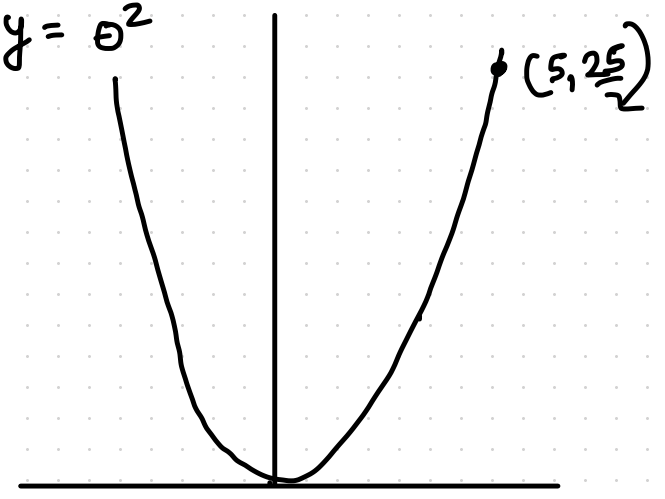
$\frac{\partial y}{\partial \theta} = \frac{2\theta}{2} = \theta$

$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



$\frac{\partial y}{\partial \theta} = 1$ (Assume $\theta > 0$)

$y = \theta^2$



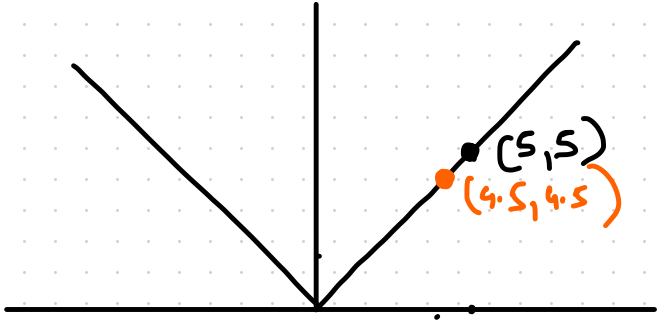
$\frac{\partial y}{\partial \theta} = \frac{2\theta}{2} = \theta$

LET $\alpha = 0.5$

$\theta_0^1 = \theta_0^0 - 0.5 * 1 = 4.5$

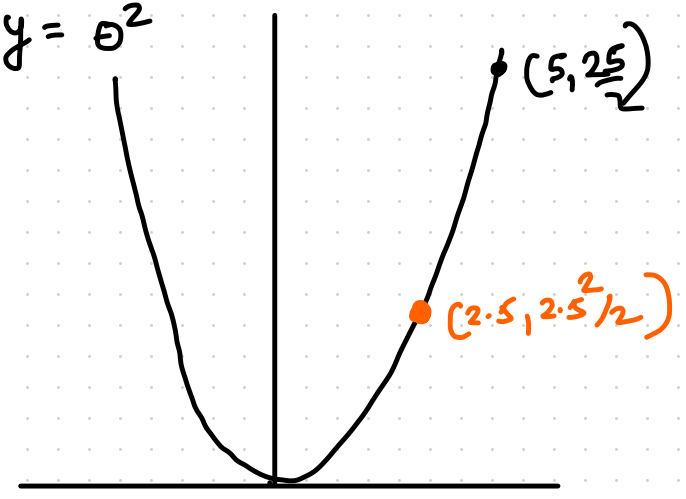
$\theta_0^1 = \theta_0^0 - 0.5 * 5 = 2.5$

$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



$\frac{\partial y}{\partial \theta} = 1$ (Assume $\theta > 0$)

$y = \theta^2$



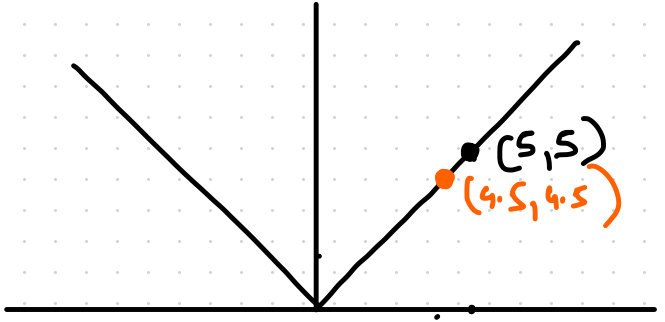
$\frac{\partial y}{\partial \theta} = \frac{2\theta}{2} = \theta$

LET $\alpha = 0.5$

$\theta_0^1 = \theta_0^0 - 0.5 * 1 = 4.5$

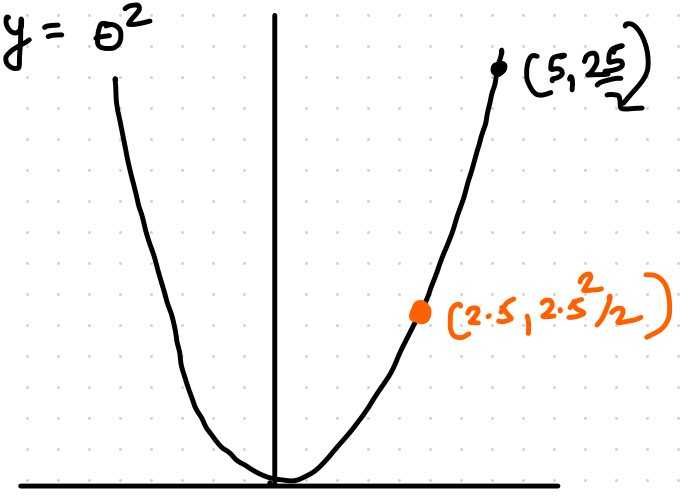
$\theta_0^1 = \theta_0^0 - 0.5 * 5 = 2.5$

$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



$\frac{\partial y}{\partial \theta} = 1$ (Assume $\theta > 0$)

$y = \theta^2$



$\frac{\partial y}{\partial \theta} = \frac{2\theta}{2} = \theta$

LET $\alpha = 0.5$

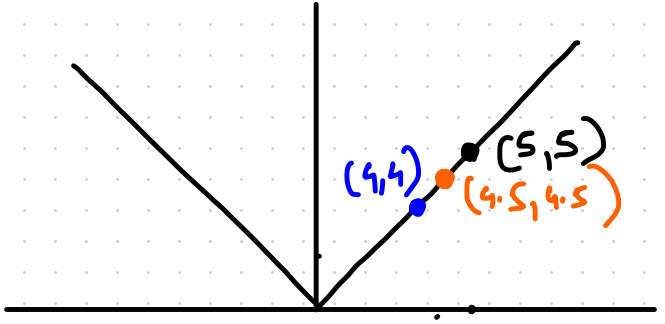
$\theta_0^1 = \theta_0^0 - 0.5 * 1 = 4.5$

$\theta_0^2 = \theta_0^1 - 0.5 * 1 = 4.0$

$\theta_0^1 = \theta_0^0 - 0.5 * 5 = 2.5$

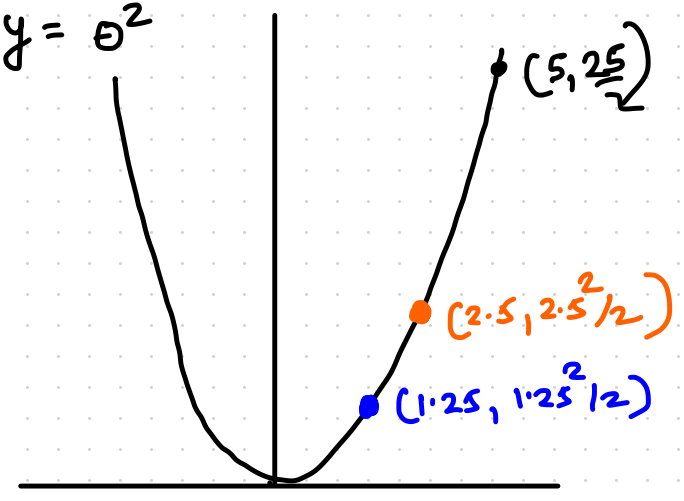
$\theta_0^2 = \theta_0^1 - 0.5 * 2.5 = 1.25$

$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



$\frac{\partial y}{\partial \theta} = 1$ (Assume $\theta > 0$)

$y = \theta^2$



$\frac{\partial y}{\partial \theta} = \frac{2\theta}{2} = \theta$

LET $\alpha = 0.5$

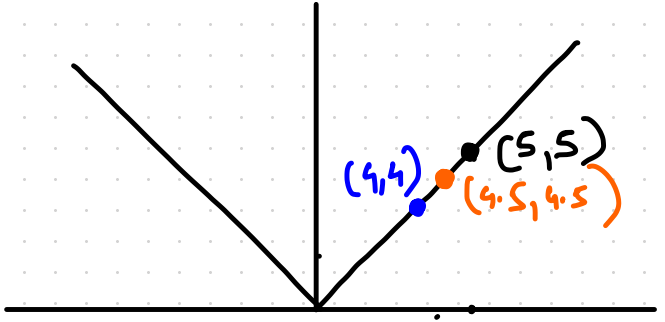
$\theta_0^1 = \theta_0^0 - 0.5 * 1 = 4.5$

$\theta_0^2 = \theta_0^1 - 0.5 * 1 = 4.0$

$\theta_0^1 = \theta_0^0 - 0.5 * 5 = 2.5$

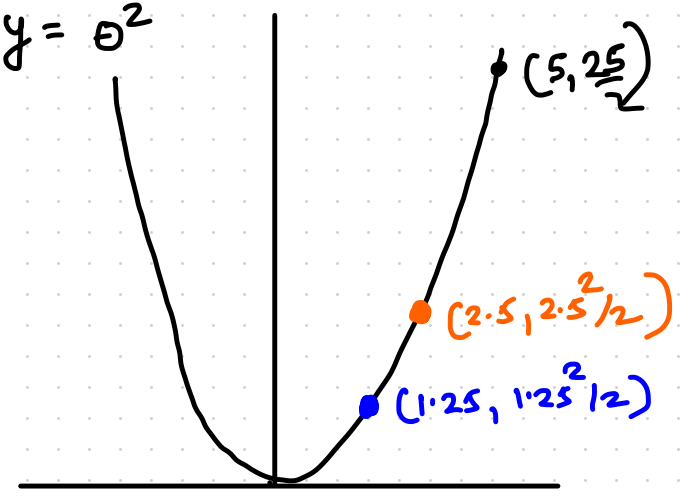
$\theta_0^2 = \theta_0^1 - 0.5 * 2.5 = 1.25$

$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



$\frac{\partial y}{\partial \theta} = 1$ (Assume $\theta > 0$)

$y = \theta^2$



$\frac{\partial y}{\partial \theta} = \frac{2\theta}{2} = \theta$

LET $\alpha = 0.5$

$\theta_0^1 = \theta_0^0 - 0.5 \times 1 = 4.5$

$\theta_0^2 = \theta_0^1 - 0.5 \times 1 = 4.0$

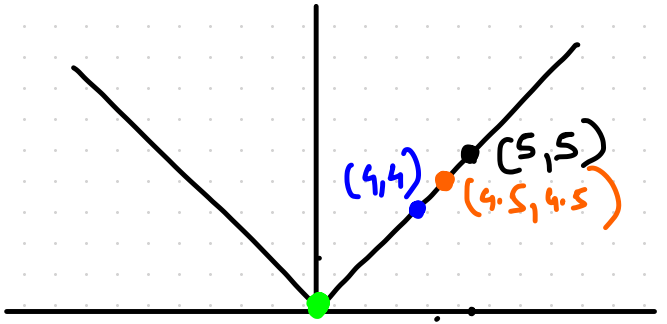
$\theta_0^t = \theta_0^{t-1} - 0.5$

$\theta_0^1 = \theta_0^0 - 0.5 * 5 = 2.5$

$\theta_0^2 = \theta_0^1 - 0.5 \times 2.5 = 1.25$

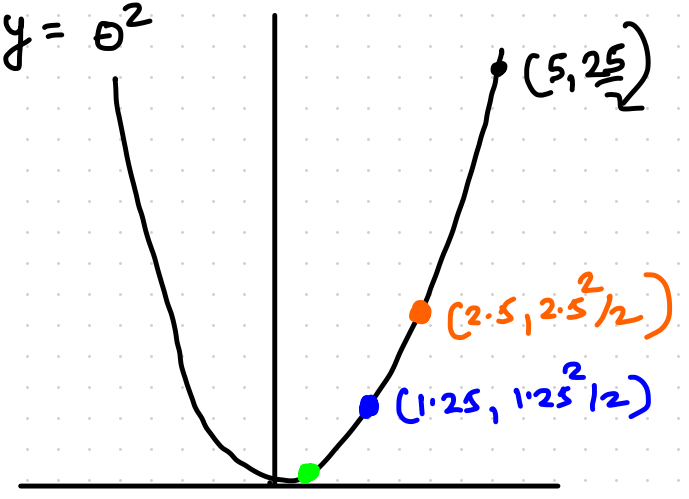
$\theta_0^t = \theta_0^{t-1} - 0.5\,\theta_0^{t-1} = 0.5\,\theta_0^{t-1}$

$y = |\theta|$ (FOR NOW ASSUME $\theta > 0$)



$\frac{\partial y}{\partial \theta} = 1$ (Assume $\theta > 0$)

$y = \theta^2$



$\frac{\partial y}{\partial \theta} = \frac{2\theta}{2} = \theta$

LET $\alpha = 0.5$

$\theta_0^{10} = 0$

$\theta_0^{10} = 5 * (0.5)^{10}$

$= 0.0048$

(Approaching 0 but not exactly zero)