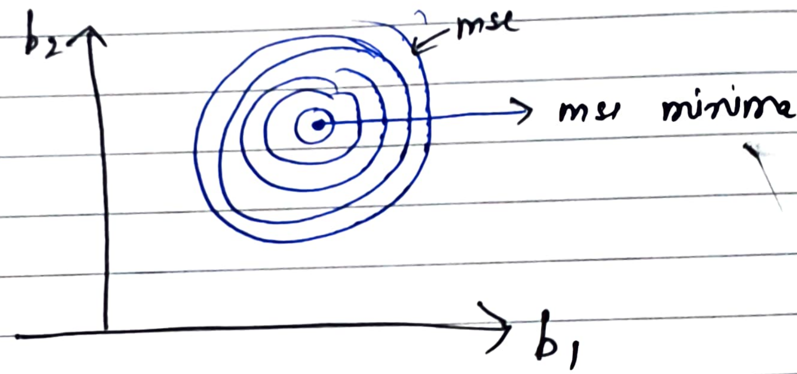
(Non Mathematics)

INTUITION FOR RIDGE

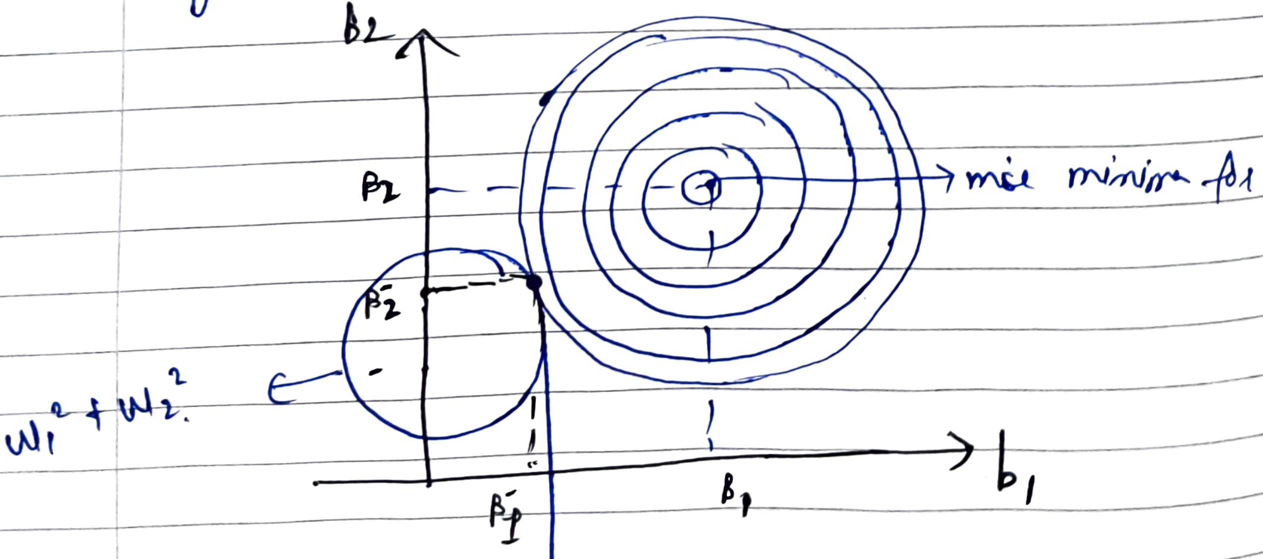
- Loss fn = mse + $\lambda \, ||W||^2$
- let's say we have 2 inputs, ∴ 3 weights → W_0, W_1, W_2
- ∴ Loss fn = mse + $\lambda(W_1^2 + W_2^2)$

<u>plotting b_1 b_2 and mse loss fn</u>



plotting regularized / penalty term of ridge.

→ Ridge term → $\lambda(W_1^2 + W_2^2)$ → make a circle



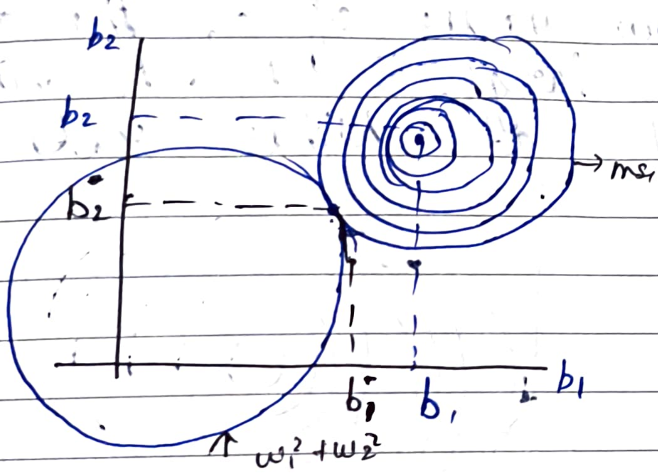
At point where MSE & circle (regularized penalty term) is intersecting will considers the solution (minima), not the mse minima.

∴ And if you notice, where both are interacting ∴ the β_1 & β_2 values are low / decreased.

Effect of λ on circle $(w_1^2 + w_2^2)$

① ~~high~~ $\rightarrow$ low λ

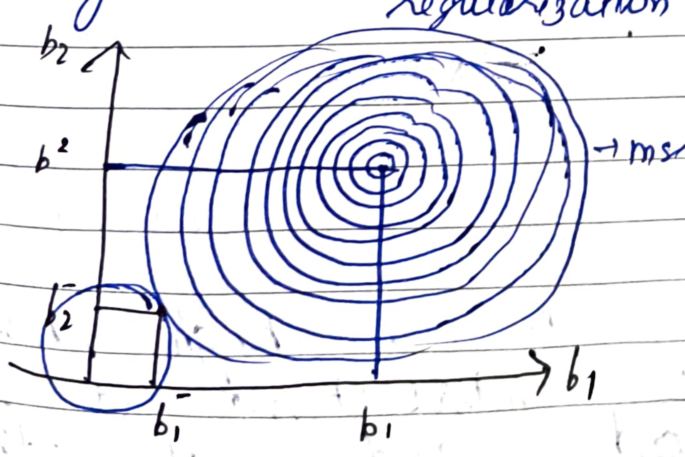
- For low value of λ, the circle will be large, ∴ value of β_1 & β_2 will not decreased at such rate



For high value of λ

② High λ

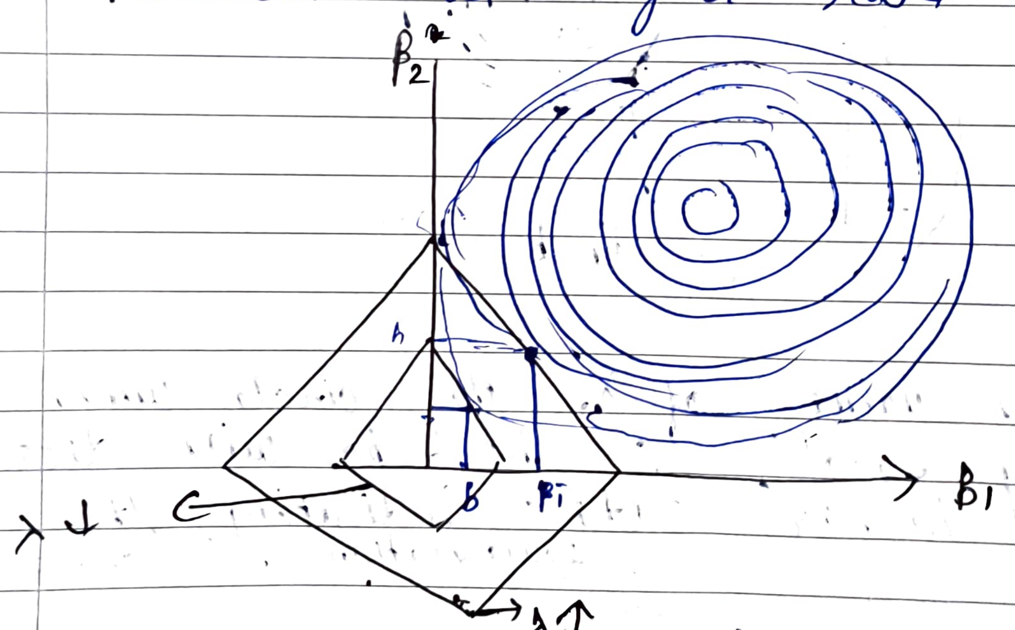
- For high value of λ, the circle will be very small, ∴ b_1 & b_2 will decrease at higher rate ∴ indicating the regularization is strong.



As you can, see for large value of λ, β_1 & β_2 decreases too much.

INTUITION FOR LASSO

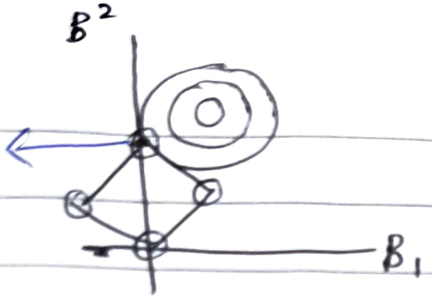
- Loss f^n = mse + $|W_1 + W_2|$
  penalty term
- Unlike Ridge, Lasso makes dimond like structure
- For high value of λ, the diamond is small, $\therefore$ β_1 & β_2 will reduced at higher rate.
- For low value of λ, the diamond will be large. $\therefore$ β_1 & β_2 will not reduced at higher rate



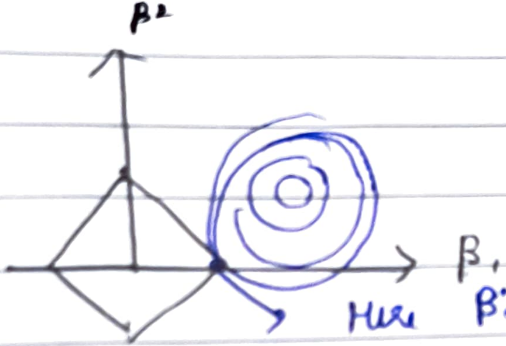
Why lasso creates sparcity?

As, there is a high chance of mse intersecting at edges of diamond. and at edges β_1 & β_2 are zero.

Ex) Here β_1 is zero.



Ex2



Here β_2 is zero.