

- Regularization is the process to discourage complexity and minimize overfitting.

• Common ways

- more data
- cleaner data
- Early stopping
- Pruning
- Dropout
- Resampling data
- Add noise (e.g. data augmentation)
- add a constraint to loss function.

↳ Linear Regression:

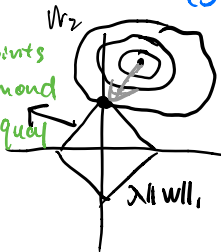
Penalize the cost function by adding a regularization term $\sum |w|$ or $\sum w^2$ to get smaller weights thus simpler model.

So that coefficients will not fit too perfectly to dataset,
→ less overfitting.

L_1 : LASSO

$$\text{loss} = \sum (y - \hat{y})^2 + \lambda \sum_{j=1}^m |w_j|$$

all the points on the diamond will have equal $\sum |w_j|$



cross point: reach the optimal $\min(\text{loss})$.

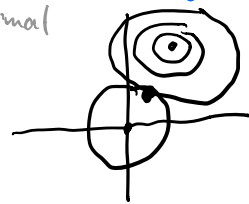
diamond

"lines touching"

- induce sparsity, tend to zero

L_2 : ridge

$$\text{loss} = \sum (y - \hat{y})^2 + \lambda \sum w_j^2$$



' ,