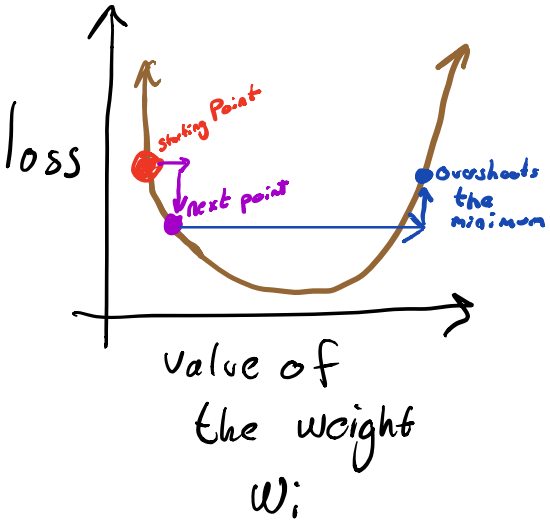
Regularization Parameters

$E(w_i)$ → error "loss" on your loss function "rmse"

$$w_i \leftarrow w_i - \eta \frac{\delta E}{\delta w_i} - \eta \lambda w_i$$

η : Learning rate: determines how big/small steps should be taken to descend



Smaller learning rate
- takes forever to descend "overfitting"

Higher learning rate
- might give exploding gradients "underfitting"

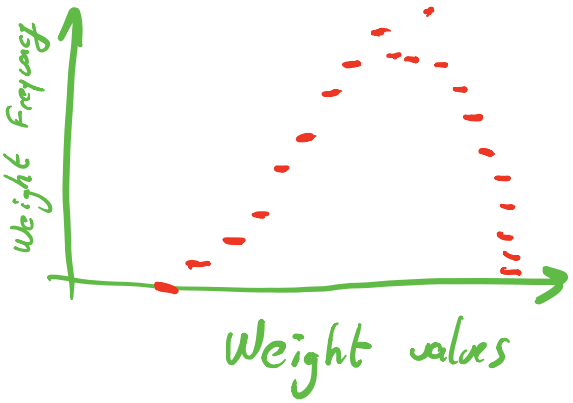
1 cycle policy
- Starts with large learning rate decreases as descending

$\eta \lambda w_i$: Weight decay

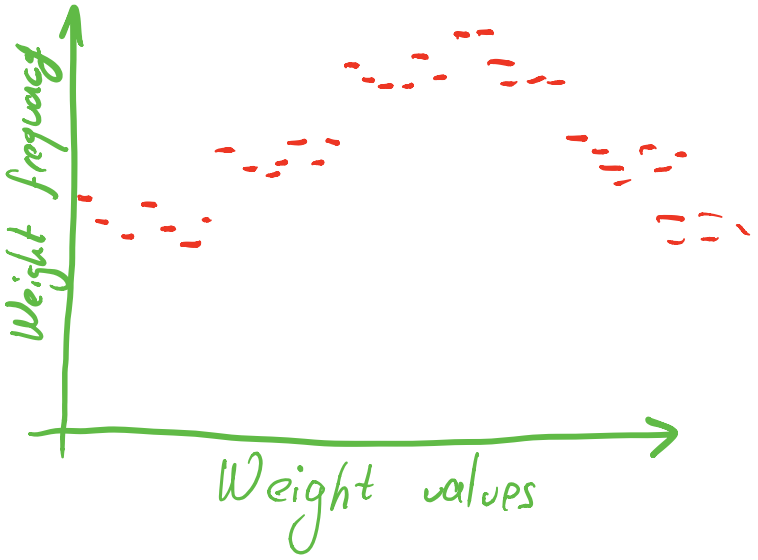
regularizes the amount of change on value of the weight

of the weight

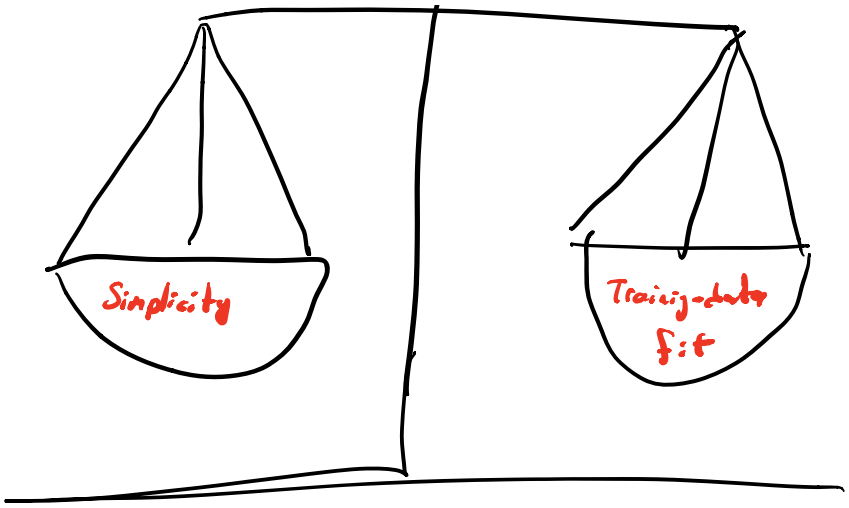
λ : regularization rate (lambda)



high value of lambda



- Choosing the right lambda value



BALANCE THIS

if lambda value is to high

- model will be simple
- risk of underfitting
- won't learn enough to make usefull predictions

if lambda value is to low

- model will be too complex
- risk of overfitting
- will learn too much about

- will learn the particularities of the training dataset, won't generalize to new data set

---

Caps GNN specific parameters

- Theta (Reconstruction loss weight)

Weight for penalizing the GCN layer for creating capsule forms different from the input

Too high	Too low
- underfitting - won't learn	- overfitting - will learn to much

---

Gexf dataset	vs	JSON style
- Industy standard - need to label		- dictionary based "fast access - no need to

as directed or undirected

label as directed or undirected