

$$W \leftarrow \frac{W}{(1 - \lambda)w}$$

$$0.8W - \lambda \text{ gr}$$

$$- \lambda (\text{gradient})$$

↓
LR

$$\min \text{loss} + \underbrace{\lambda \|w\|^2}_{\text{regularization}}$$



