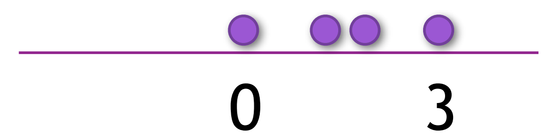
Notation and Math

Regularization intuition

Minimization toy example: Choose a value for λ and then minimize: $\frac{1}{2}(3-w)^2 + \lambda w^2$

The first term $\frac{1}{2}(3-w)^2$ is minimized when $w = 3$

The second term λw^2 is minimized when $w = 0$



- If $\lambda = \frac{1}{2}$ we can use calculus to minimize this function
  $\frac{d\frac{1}{2}(3-w)^2 + \frac{1}{2}w^2}{dw} = -(3-w) + w$, so it is minimized when $w = 3/2$

- If $\lambda = \frac{1}{4}$ we can use calculus to minimize this function
  $\frac{d\frac{1}{2}(3-w)^2 + \frac{1}{4}w^2}{dw} = -(3-w) + \frac{1}{2}w$, so it is minimized when $w = 2$

λ determines how much we prioritize the first term over the second term.
Smaller λ means we are closer to 3, larger λ means we are closer to 0

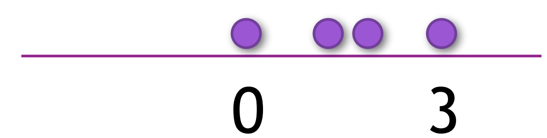
The next slide was not presented in class

Regularization intuition

Maximization toy example. Choose a value for λ and then maximize: $-\frac{1}{2}(3-w)^2 - \lambda w^2$

The first term $-\frac{1}{2}(3-w)^2$ is maximization when $w = 3$

The second term $-\lambda w^2$ is maximization when $w = 0$



- If $\lambda = \frac{1}{2}$ we can use calculus to maximization this function

  $$\frac{d - \frac{1}{2}(3-w)^2 - \frac{1}{2}w^2}{dw} = (3-w) - w,$$ so it is maximization when $w = 3/2$

- If $\lambda = \frac{1}{4}$ we can use calculus to maximization this function

  $$\frac{d - \frac{1}{2}(3-w)^2 - \frac{1}{4}w^2}{dw} = (3-w) - \frac{1}{2}w,$$ so it is maximization when $w = 2$

λ determines how much we prioritize the first term over the second term.
Smaller λ means we are closer to 3, larger λ means we are closer to 0