Weight Decay

A parameter norm penalty using the L2 norm:

$$\text{Loss} = \text{Loss} + \frac{1}{2} \|w\|_2^2$$

L2 norm penalty

Loss function   Loss function   weights

In regression this use of the L2 norm penalty is called ridge regression.

Chris Albon