PARAMETER NORM

Penalties for Regularization

Adds a parameter norm penalty to loss function:

$$Loss' = Loss + \lambda\,\Omega(\theta)$$

Loss function

Hyperparameter (positive number)

Norm Penalty

Parameters

- Most often not applied to parameters of biases.
- Technically each layer can have its own hyperparameter and/or norm penalty.

Chris Albon