

## Normalizing activations in a network

### Batch Normalization

- Easily train deep networks. Normalizing inputs helps to speed up learning. So, can we normalize  $z()$  or  $a()$  to speed up that layer.
- Normalize each layer of  $z$  separately