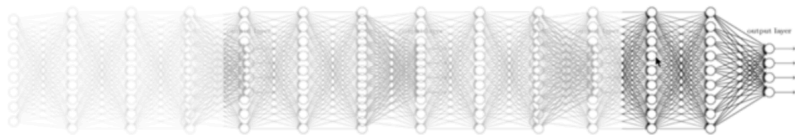


Deep Learning par la Pratique

Optimisation/Régularisation de réseaux profonds

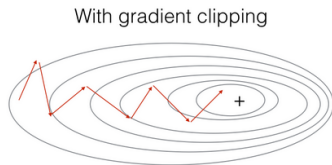
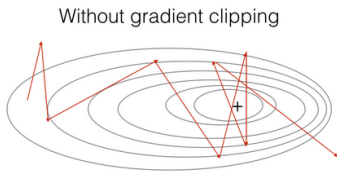
Optimisation de réseaux profonds

Disparition du gradient (ou son explosion)



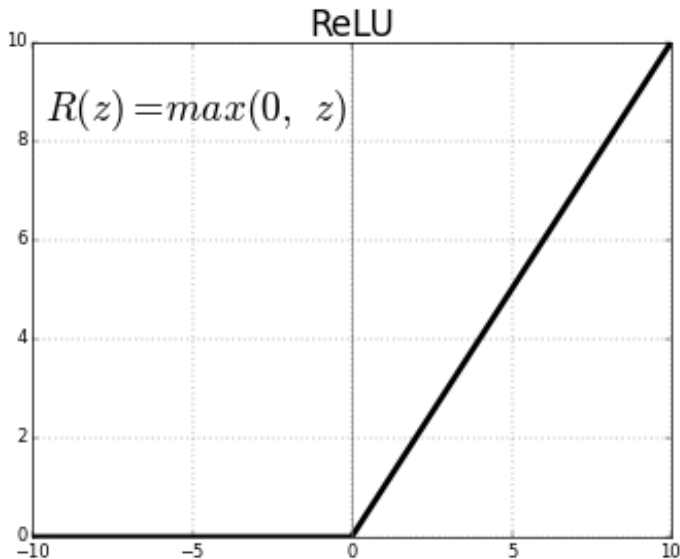
Optimisation de réseaux profonds

Exploding Gradient \Rightarrow Gradient clipping



Optimisation de réseaux profonds

Vanishing Gradient \Rightarrow utilisation de ReLu plutôt que les sigmoïdes



Optimisateur plus rapide que SGD

- Root Mean Square Propagation (RMSProp)
- Adaptive Gradient Algorithm (AdaGrad)
- Adaptive Moment Estimation (Adam) \Leftarrow
- ... AdaBound (2019) ?