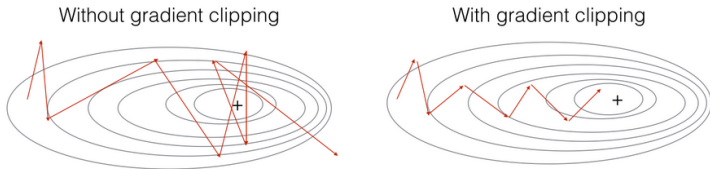


Machine Learning

Régularisation

Régularisation

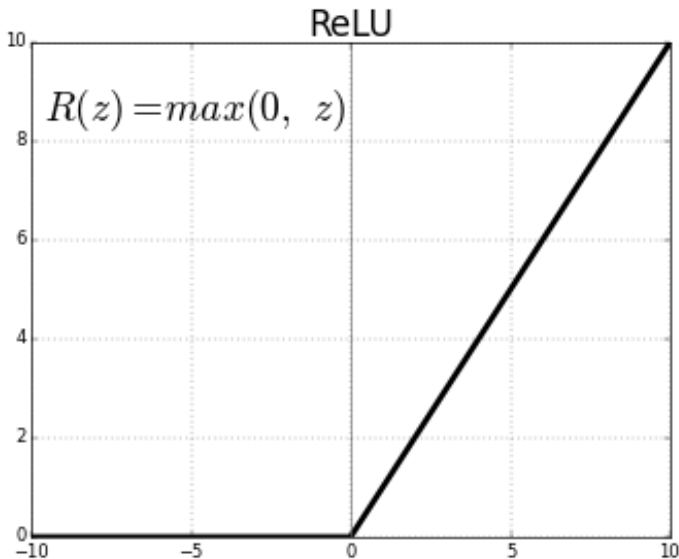
Exploding Gradient \Rightarrow Gradient clipping



Dans le cadre des réseaux récurrents \Rightarrow initialisation avec des matrices orthogonales

Régularisation

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) ?

Régularisation

Dropout

