

The Reversible Residual Network: Backpropagation Without Storing Activations

ソース

- <https://arxiv.org/pdf/1707.04585>

概要

- ResNetを可逆的にしたRevNet(Reversible Residual Network)を提案
- 可逆的にすることで学習時のメモリを節約できる
- 計算グラフで誤差逆伝搬の際に次の層から現在の層を計算できるから
- ResNetではそれができず状態を保存しておかないといけなくてメモリを使うことになる

RevNet

- 普通のResNetは以下

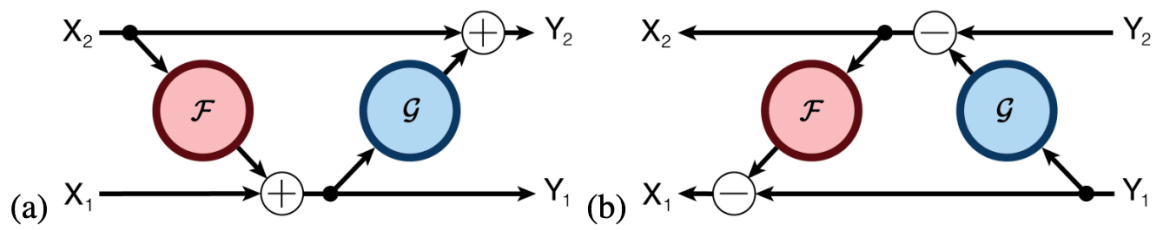
$$y = x + \mathcal{F}(x)$$

- これは不可逆的である
 - y から x がもとまらない
- そこでRevNetでは以下のように入力として x_1, x_2 の2つを受け取り y_1, y_2 の2つを出力する

$$\begin{aligned} y_1 &= x_1 + \mathcal{F}(x_2) \\ y_2 &= x_2 + \mathcal{F}(y_1) \end{aligned} \tag{1}$$

- これは以下のように変形できるため可逆的である

$$\begin{aligned} x_2 &= y_2 - \mathcal{G}(y_1) \\ x_1 &= y_1 - \mathcal{F}(x_2) \end{aligned} \tag{2}$$



英語

- surjection : 全射
- injection : 単射
- bijection : 全単射
- infinitesimal : 無限小