

Tanh non-linear Layer

Zhiliang Tian

2016.06.04

1 Introduction

1.1 What is Tanh

Tanh is a non-linear function.

- Output range is (-1, 1)
- continuously differentiable, so it can be used in neural network.

2 Feed Forward

2.1 Formula

$$y = \text{Tanh}(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \quad (1)$$

$$\mathbf{y} = \text{Tanh}(\mathbf{x}) = \left\{ \frac{\exp(\mathbf{x}_0) - \exp(-\mathbf{x}_0)}{\exp(\mathbf{x}_0) + \exp(-\mathbf{x}_0)}, \frac{\exp(\mathbf{x}_1) - \exp(-\mathbf{x}_1)}{\exp(\mathbf{x}_1) + \exp(-\mathbf{x}_1)}, \dots, \frac{\exp(\mathbf{x}_N) - \exp(-\mathbf{x}_N)}{\exp(\mathbf{x}_N) + \exp(-\mathbf{x}_N)} \right\}^T \quad (2)$$

$$\mathbf{y}_i = \text{Tanh}(\mathbf{x}_i) = \frac{\exp(\mathbf{x}_i) - \exp(-\mathbf{x}_i)}{\exp(\mathbf{x}_i) + \exp(-\mathbf{x}_i)} \quad (3)$$

2.2 Graph

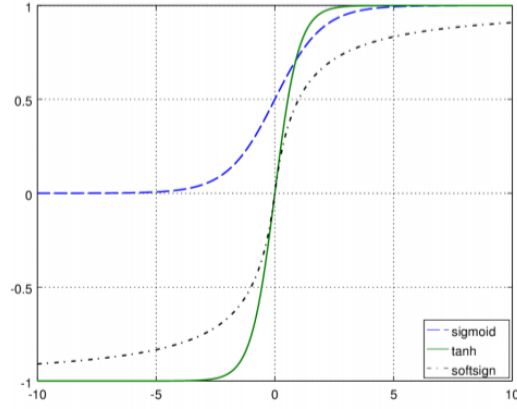


Figure 1: Tanh compared with softsign and sigmoid

2.3 Some Discuss

•

3 Back Propagation

$$\begin{aligned}
& \frac{\alpha(\tanh(x))}{\alpha(x)} \\
&= \frac{(\exp(x) - \exp(-x))'}{\exp(x) + \exp(-x)} + (\exp(x) - \exp(-x)) \left(\frac{1}{\exp(x) + \exp(-x)} \right)' \\
&= \frac{(\exp(x) + \exp(-x))}{\exp(x) + \exp(-x)} + (-1) * \frac{(\exp(x) - \exp(-x))}{(\exp(x) + \exp(-x))^2} * (\exp(x) + \exp(-x))' \quad (4) \\
&= 1 - \frac{(\exp(x) - \exp(-x))^2}{(\exp(x) + \exp(-x))^2} \\
&= \frac{4}{(\exp(x) + \exp(-x))^2}
\end{aligned}$$

$$\frac{\alpha(\tanh(x))}{\alpha(x)} = \frac{4}{(\exp(x) + \exp(-x))^2} \quad (5)$$