|  |
| --- |
|  |
| LSTM |
|  |

本文介绍LSTM网络的 各种模型的理论推导

**\*\*\*\*\*\***

2017-4-8

作者: 钟超

LSTM

目录

[1 LSTM 2](#_Toc486586891)

[1.1 求解门参数 3](#_Toc486586892)

[1.1.1 求解 3](#_Toc486586893)

[1.1.2 求解 4](#_Toc486586894)

[1.2 求解门参数 5](#_Toc486586895)

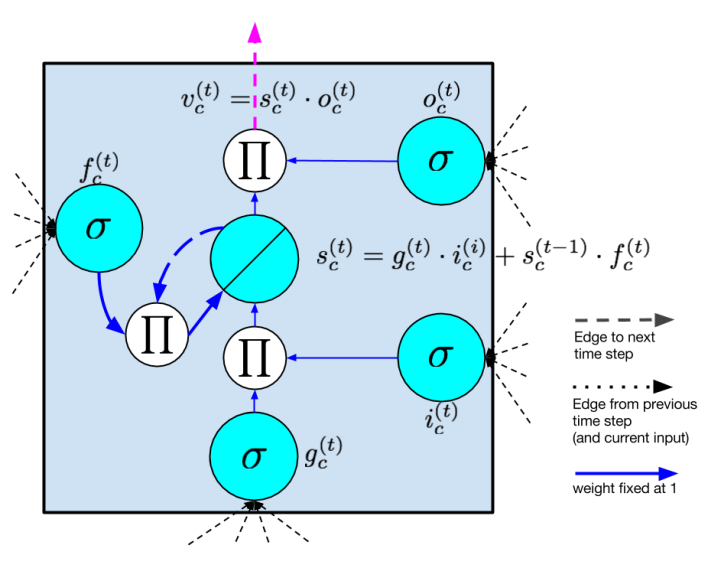
[1.3 求解输出层导数 5](#_Toc486586896)

# 1 LSTM

ref：

A Critical Review of Recurrent Neural Networks for Sequence Learning

http://nicodjimenez.github.io/2014/08/08/lstm.html



（1-1）

（1-2）

（1-3）

（1-4）

（1-5）

（1-6）

（1-7）

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| *，，，* |  |  |  |
| *，，，* |  |  |  |
| *，，，* |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| M | 输入序列到隐藏层维度 |  |  |
| N | 维度 |  |  |
| Q | 维度 |  |  |
| R |  |  |  |

（1-8）

（1-9）

（1-10）

（1-11）

## 1.1 求解门参数

### 1.1.1 求解

因为网络前向传输的时候对时刻之前的损失函数没有影响，所以有

（1-12）

用表示从时刻到的累积损失

（1-13）

表示整个序列的损失。所以可以重新写成

（1-14）

所以

， （1-15）

为了求解门相关的参数需要先求解，求解时，除了要考虑当前时刻对应的的还需要下一个时刻对应的，即如下

（1-16）

其中

（1-17）

**当时，**

（1-18）

（1-19）

（1-20）

（1-21）

当时，为了计算反向传输，除了计算当前时间外，还需要上一时刻的和，所以需要在当前时刻计算出和供下一时刻使用

（1-22）

（1-23）

**整个计算过程**

**时：**

**http://m.blog.csdn.net/u014313009/article/details/51045303**

（1-24）

（1-25）

同时计算

（1-26）

（1-27）

**当t<T时：**

由于

所以

（1-28）

（1-29）

同时计算

（1-30）

（1-31）

### 1.1.2 求解

（1-32）

（1-33）

（1-34）

（1-35）

（1-36）

（1-37）

（1-38）

（1-39）

（1-40）

总结：

按照式（1-24）至式（1-40）用反向传输算法就可以计算出梯度了，反向传输中会用到正向计算的信息，由于正向计算比较简单，在此不做述及。

## 1.2 求解门参数

， （1-41）

（1-42）

（1-43）

（1-44）

（1-45）

（1-46）

## 1.3 求解输出层导数

http://m.blog.csdn.net/u014313009/article/details/51045303

（1-47）

（1-48）