



# 子午线

## 学习笔记

作者: leekarry

组织: 果壳

时间: October 28, 2019

版本: 0.1



我很快,快到时间都会变慢。而我一慢,时间就会过得飞快。

# 目 录

<b>1</b>	<b>深度前馈网络</b>	<b>1</b>
1.1	实例:学习 XOR	1
1.2	基于梯度的学习	1
1.3	隐藏单元	1
1.4	架构设计	1
1.5	反向传播和其他的微分算法	1
<b>2</b>	<b>深度学习中的正则化</b>	<b>2</b>
2.1	参数范数惩罚	2
2.2	作为约束的范数惩罚	2
2.3	正则化和欠约束问题	2
2.4	数据集增强	2
2.5	噪声鲁棒性	2
2.6	半监督学习	2
2.7	多任务学习	2
2.8	提前终止	2
2.9	参数绑定和参数共享	2
2.10	稀疏表示	2
2.11	Bagging 和其他集成方法	2
2.12	Dropout	2
2.13	切面距离、正切传播和流形正切分类器	2
<b>3</b>	<b>深度模型中的优化</b>	<b>3</b>
3.1	学习和纯优化有什么不同	3
3.2	神经网络优化中的挑战	3
3.3	基本算法	3
3.4	参数初始化策略	3
3.5	自适应学习率算法	3
3.6	二阶近似方法	3
3.7	优化策略和元算法	3
<b>4</b>	<b>卷积网络</b>	<b>4</b>
4.1	卷积运算	4
4.2	动机	4
4.3	池化	4

4.4	卷积与池化作为一种无限强的先验 . . . . .	4
4.5	基本卷积函数的变体 . . . . .	4
4.6	结构化输出 . . . . .	4
4.7	数据类型 . . . . .	4
4.8	高效的卷积算法 . . . . .	4
4.9	随机或无监督的特征 . . . . .	4
4.10	卷积网络的神经科学基础 . . . . .	4
4.11	卷积网络与深度学习的历史 . . . . .	4
<b>5</b>	<b>序列建模:循环和递归网络</b>	<b>5</b>
5.1	展开计算图 . . . . .	5
5.2	循环神经网络 . . . . .	5
5.3	双向 RNN . . . . .	5
5.4	基于编码-解码的序列到序列架构 . . . . .	5
5.5	深度循环网络 . . . . .	5
5.6	递归神经网络 . . . . .	5
5.7	长期依赖的挑战 . . . . .	5
5.8	回声状态网络 . . . . .	5
5.9	渗漏单元和其他多时间尺度的策略 . . . . .	5
5.10	长短期记忆和其他门控 RNN . . . . .	5
5.11	优化长期依赖 . . . . .	5
5.12	外显记忆 . . . . .	5
<b>6</b>	<b>实践方法论</b>	<b>6</b>
6.1	性能度量 . . . . .	6
6.2	默认的基准模型 . . . . .	6
6.3	决定是否收集更多数据 . . . . .	6
6.4	选择超参数 . . . . .	6
6.5	调试策略 . . . . .	6
6.6	示例:多位数字识别 . . . . .	6
<b>7</b>	<b>应用</b>	<b>7</b>
7.1	大规模深度学习 . . . . .	7
7.2	计算机视觉 . . . . .	7
7.3	语音识别 . . . . .	7
7.4	自然语言处理 . . . . .	7
7.5	其他应用 . . . . .	7

<b>8 线性因子模型</b>	<b>8</b>
8.1 概率 PCA 和因子分析	8
8.2 独立成分分析	8
8.3 慢特征分析	8
8.4 稀疏编码	8
8.5 PCA 的流形解释	8
<b>9 自编码器</b>	<b>9</b>
9.1 欠完备自编码器	9
9.2 正则自编码器	9
9.3 表示能力、层的大小和深度	9
9.4 随机编码器和解码器	9
9.5 去噪自编码器详解	9
9.6 使用自编码器学习流形	9
9.7 收缩自编码器	9
9.8 预测稀疏分解	9
9.9 自编译器的应用	9
<b>10 表示学习</b>	<b>10</b>
10.1 贪心逐层无监督预训练	10
10.2 迁移学习和领域自适应	10
10.3 半监督解释因果关系	10
10.4 分布式表示	10
10.5 得益于深度的指数增益	10
10.6 提供发现潜在原因的线索	10
<b>11 深度学习中的结构化概率模型</b>	<b>11</b>
11.1 非结构化建模的挑战	11
11.2 使用图描述模型结构	11
11.3 从图模型中采样	11
11.4 结构化建模的优势	11
11.5 学习依赖关系	11
11.6 推断和近似推断	11
11.7 结构化概率模型的深度学习方法	11
<b>12 深度生成模型</b>	<b>12</b>
12.1 玻尔兹曼机	12
12.2 受限玻尔兹曼机	12
12.3 深度信念网络	12
12.4 深度玻尔兹曼机	12

12.5 实值数据上的玻尔兹曼机 . . . . .	12
12.6 卷积玻尔兹曼机 . . . . .	12
12.7 用于结构化或序列输出的玻尔兹曼机 . . . . .	12
12.8 其他玻尔兹曼机 . . . . .	12
12.9 通过随机操作的反向传播 . . . . .	12
12.10 有向生成网络 . . . . .	12
12.11 从自编码器采样 . . . . .	12
12.12 生成随机网络 . . . . .	12
12.13 其他生成方案 . . . . .	12
12.14 评估生成模型 . . . . .	12
12.15 结论 . . . . .	12



# 第 1 章 深度前馈网络



## 1.1 实例:学习 XOR

## 1.2 基于梯度的学习

## 1.3 隐藏单元

## 1.4 架构设计

## 1.5 反向传播和其他的微分算法

## 第 2 章 深度学习中的正则化



### 2.1 参数范数惩罚

### 2.2 作为约束的范数惩罚

### 2.3 正则化和欠约束问题

### 2.4 数据集增强

### 2.5 噪声鲁棒性

### 2.6 半监督学习

### 2.7 多任务学习

### 2.8 提前终止

### 2.9 参数绑定和参数共享

### 2.10 稀疏表示

### 2.11 Bagging 和其他集成方法

### 2.12 Dropout

### 2.13 切面距离、正切传播和流形正切分类器

## 第 3 章 深度模型中的优化



3.1 学习和纯优化有什么不同

3.2 神经网络优化中的挑战

3.3 基本算法

3.4 参数初始化策略

3.5 自适应学习率算法

3.6 二阶近似方法

3.7 优化策略和元算法



## 第 4 章 卷积网络

---

### 4.1 卷积运算

### 4.2 动机

### 4.3 池化

### 4.4 卷积与池化作为一种无限强的先验

### 4.5 基本卷积函数的变体

### 4.6 结构化输出

### 4.7 数据类型

### 4.8 高效的卷积算法

### 4.9 随机或无监督的特征

### 4.10 卷积网络的神经科学基础

### 4.11 卷积网络与深度学习的历史

## 第 5 章 序列建模:循环和递归网络

---

### 5.1 展开计算图

### 5.2 循环神经网络

### 5.3 双向 RNN

### 5.4 基于编码-解码的序列到序列架构

### 5.5 深度循环网络

### 5.6 递归神经网络

### 5.7 长期依赖的挑战

### 5.8 回声状态网络

### 5.9 渗漏单元和其他多时间尺度的策略

### 5.10 长短期记忆和其他门控 RNN

### 5.11 优化长期依赖

### 5.12 外显记忆

## 第 6 章 实践方法论

---

### 6.1 性能度量

### 6.2 默认的基准模型

### 6.3 决定是否收集更多数据

### 6.4 选择超参数

### 6.5 调试策略

### 6.6 示例:多位数字识别

## 第 7 章 应用

---

### 7.1 大规模深度学习

### 7.2 计算机视觉

### 7.3 语音识别

### 7.4 自然语言处理

### 7.5 其他应用

## 第 8 章 线性因子模型



### 8.1 概率 PCA 和因子分析

### 8.2 独立成分分析

### 8.3 慢特征分析

### 8.4 稀疏编码

### 8.5 PCA 的流形解释

## 第 9 章 自编码器

---

### 9.1 欠完备自编码器

### 9.2 正则自编码器

### 9.3 表示能力、层的大小和深度

### 9.4 随机编码器和解码器

### 9.5 去噪自编码器详解

### 9.6 使用自编码器学习流形

### 9.7 收缩自编码器

### 9.8 预测稀疏分解

### 9.9 自编译器的应用

## 第 10 章 表示学习



**10.1 贪心逐层无监督预训练**

**10.2 迁移学习和领域自适应**

**10.3 半监督解释因果关系**

**10.4 分布式表示**

**10.5 得益于深度的指数增益**

**10.6 提供发现潜在原因的线索**

## 第 11 章 深度学习中的结构化概率模型



11.1 非结构化建模的挑战

11.2 使用图描述模型结构

11.3 从图模型中采样

11.4 结构化建模的优势

11.5 学习依赖关系

11.6 推断和近似推断

11.7 结构化概率模型的深度学习方法



## 第 12 章 深度生成模型

---

12.1 玻尔兹曼机

12.2 受限玻尔兹曼机

12.3 深度信念网络

12.4 深度玻尔兹曼机

12.5 实值数据上的玻尔兹曼机

12.6 卷积玻尔兹曼机

12.7 用于结构化或序列输出的玻尔兹曼机

12.8 其他玻尔兹曼机

12.9 通过随机操作的反向传播

12.10 有向生成网络

12.11 从自编码器采样

12.12 生成随机网络

12.13 其他生成方案

12.14 评估生成模型

12.15 结论