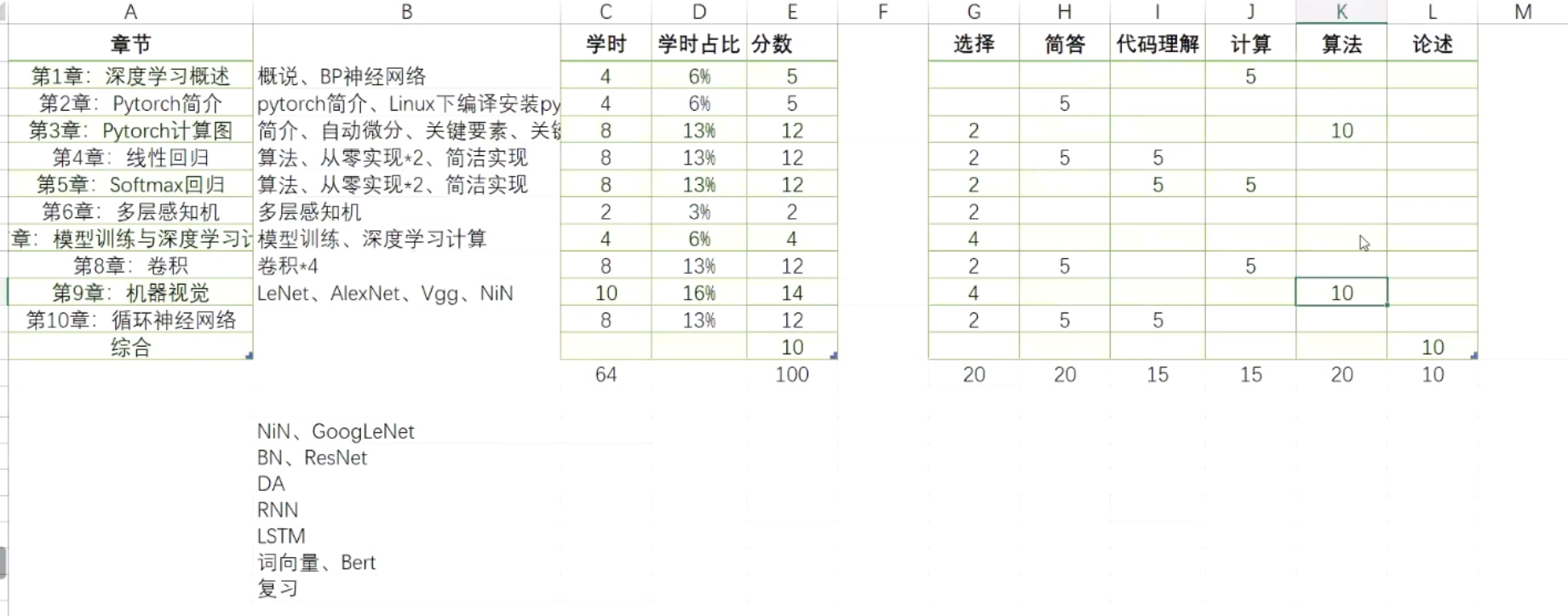
深度学习

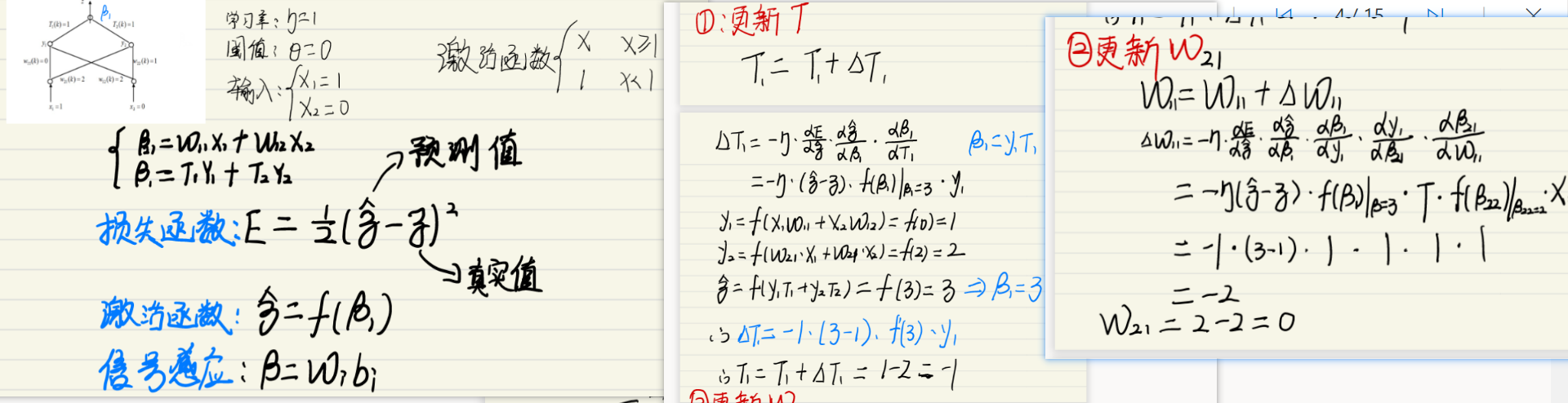
2022年5月28日

12:58

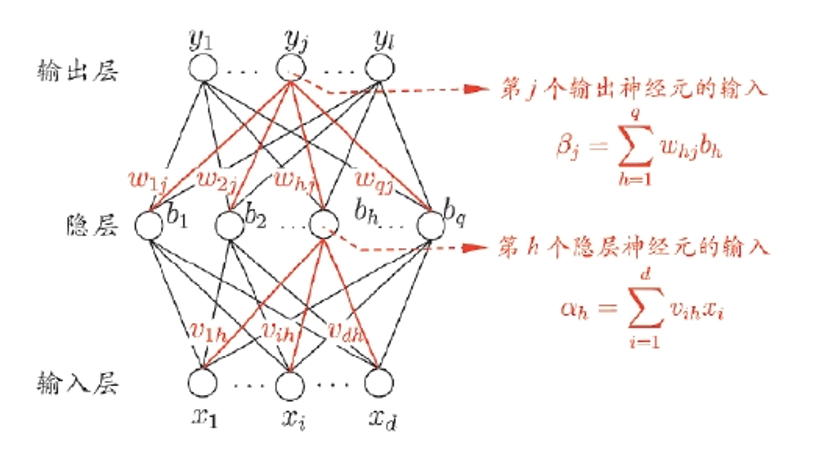


**第1章：深度学习概述**

        BP神经网络的反向传播（《深度学习08\_小复习.pptx》，链式法则：损失函数，激活函数、信号强度；每多一层多一个sigma符号：也是网络不能太深的原因之一）



        网络参数确定（权值个数和偏置个数）



网络不能太深的原因：sigma提高复杂度，难以计算，容易过拟合 ，梯度消失爆炸

网络参数确定（权值个数和偏置个数）

线+圈

**第2章：Pytorch简介**

        Pytorch安装要点：装python、装CUDA、装gpu版的torch

pip3 install torch torchvision torchaudio --extra-index-url <https://download.pytorch.org/whl/cu113>

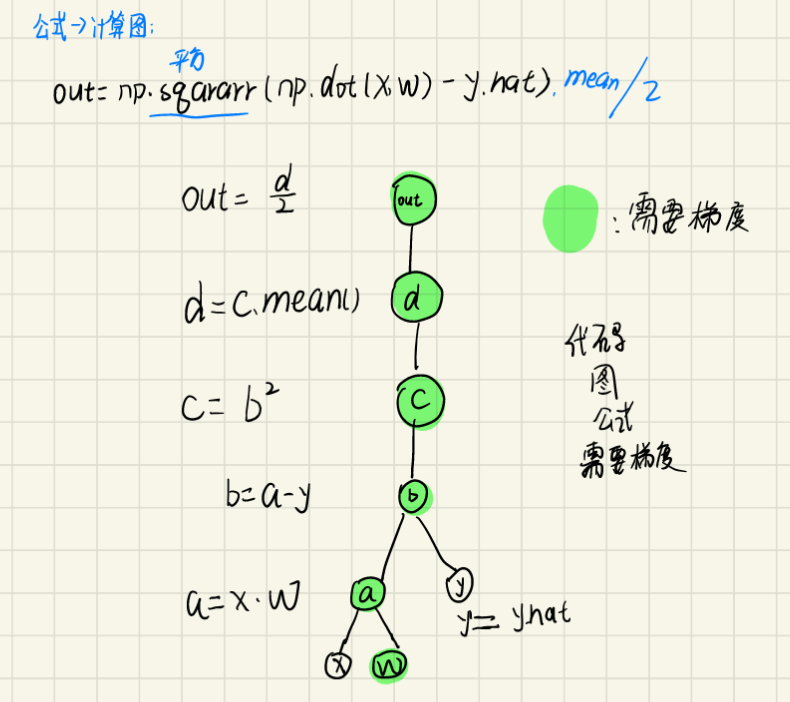
        Linux编译安装Python的主要步骤：下载、解压、配置(./configure)、编译安装(make && make install)

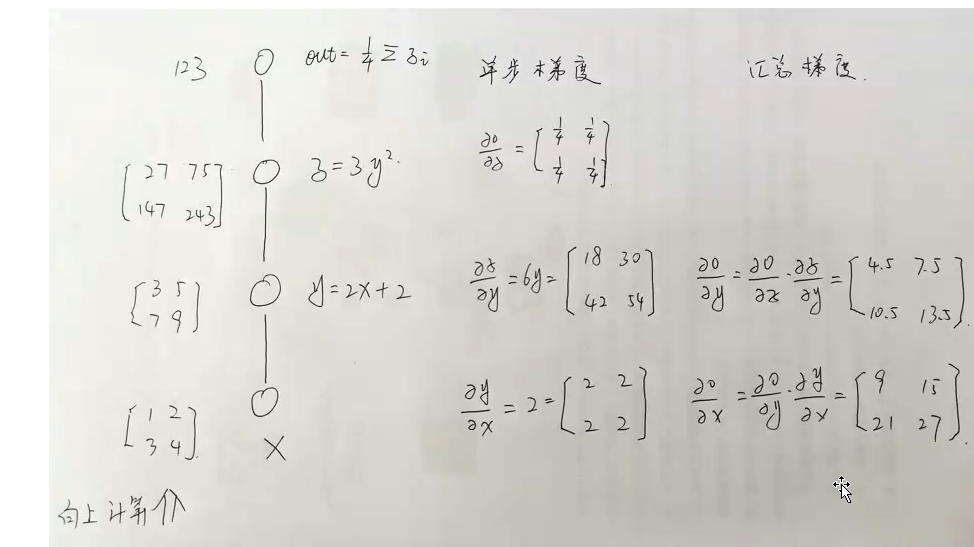
        Tensor的基本知识：数据方面就是numpy，计算图方面是张量的本质。

**第3章：Pytorch计算图**

给几个公式，可以画出计算图，给计算图，可以复原出公式。

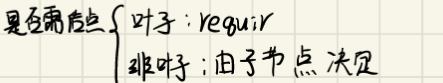
        给计算图，可以向上算数据，可以算回传梯度





Pytorch的计算图结构：

1、是否所有的叶子节点都可以设置为需要计算梯度，是否所有的叶子节点都需要计算梯度。叶子结点初始化梯度



是否

2、是否默认所有的计算图都可以回传多次。不磨人

3、计算图是否一个有向无环图。对

4、是否默认可以让中间节点保存梯度。 默认不保存

5、非叶子节点的导函数是怎么来的？ 计算来

**第4章：线性回归**

分类与回归，线性回归的基本概念，可用的解法

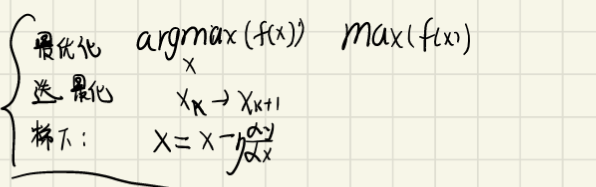
分类和回归：

**回归：*有督监*，给定特征值，预测的目标值是连续的**

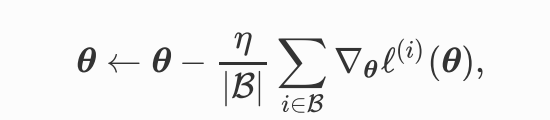
**分类：有监督，给定特征值，预测的目标值是离散的**

**回归：名义量，顺序量，可加量，可比量**

小批量随机梯度下降法

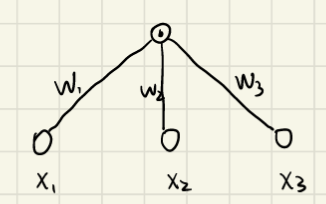


小批量：多了装不下，一个一个太慢



看网络写方程，看方程画网络。

神经网络画图：y = wx1+wx2+wx3+b

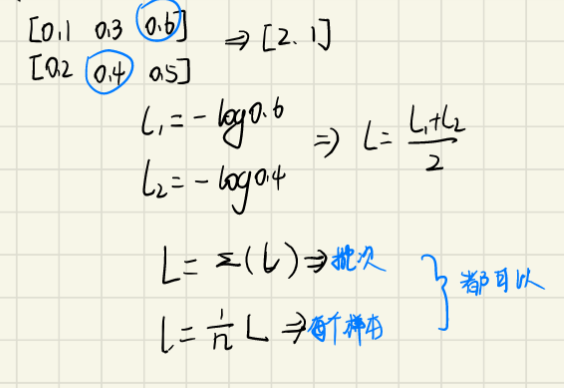


数据加载、损失函数

最优化，迭代最优化，梯度下降法

**第5章：Softmax回归**

        Softmax和交叉熵，代码、计算、业务场景，



        权值、偏置初始化

        .view()的业务意义

**第6章：多层感知机**

单层和多层的区别（非线性激活函数的意义）

没有本质区别：每一层都是线性激活函数，就没区别

想要有区别：加非线性激活函数

Relu激活函数，非线性

权值、偏置初始化，MLP的24完备性

权值不可以都初始化0，每个神经元都一样，b无约束

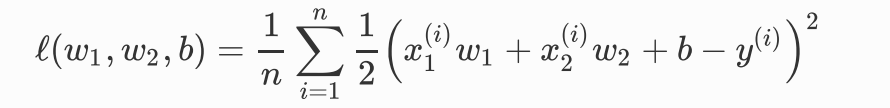
完备性：合理的神经元个数和激活函数，对任意的模型的无限接近，能拟合所有函数

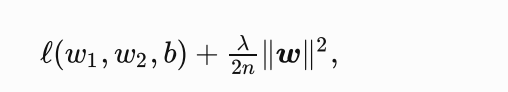
**第7章：模型训练与深度学习计算**

        为什么网络不能过深

过拟合，不容易训练，梯度爆梯度炸，消失

        Wd的原因和表现





加二阶范数，权值衰减

        Dropout的原理和实现

训练有的加班（权值放大），预测都干活

        模型的读写

模型本地生成，参数读写进来，

（迁移学习：下面学习率小，上面大，正则化项：上面小，下面打）

GPU计算的特点和历史

GPU在CPU的指导下干活，协处理器

英伟达，黄仁勋

**第8章：卷积**

        卷积的思想基础，能够解释局部性和平移不变性

                对于某个像素的信息，应当只与其附近的像素有关系，超出一定的距离以后则无关 卷积核体积小

                对某个区域进行的特征提取所得到的输出并不会由于平移操作而发生变化 一个核就行

有生物学基础，简化计算，减少模型参数，方便运算，减少过拟合，

卷积、池化的代码、计算

        步幅、填充、多通道

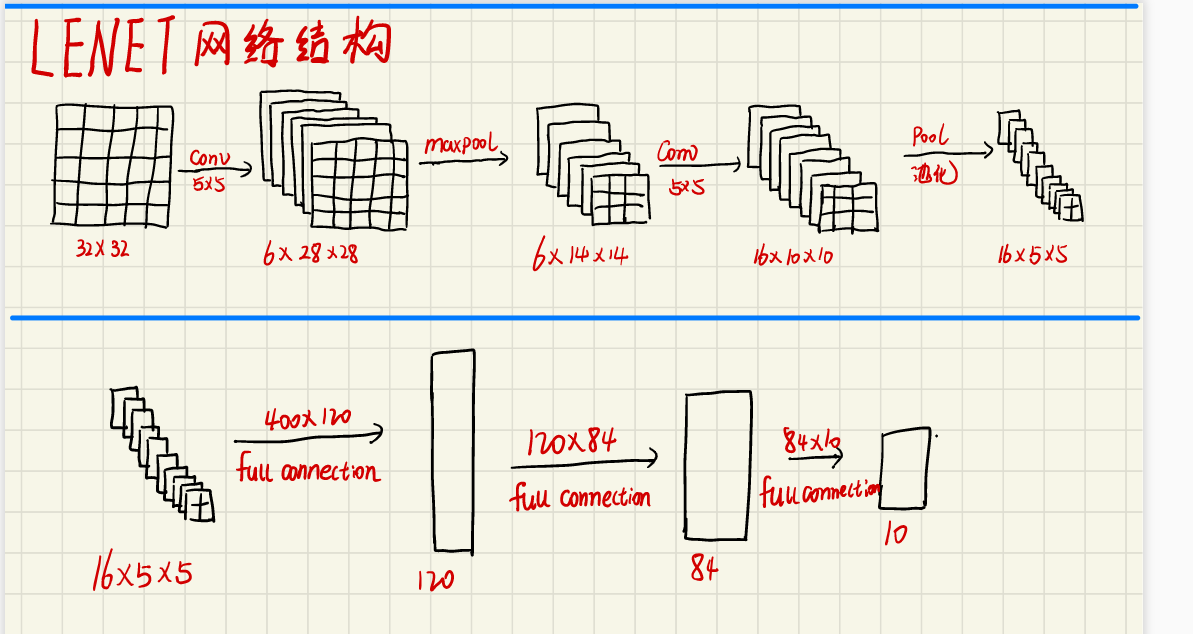
        1\*1卷积

**第9章：机器视觉初步**

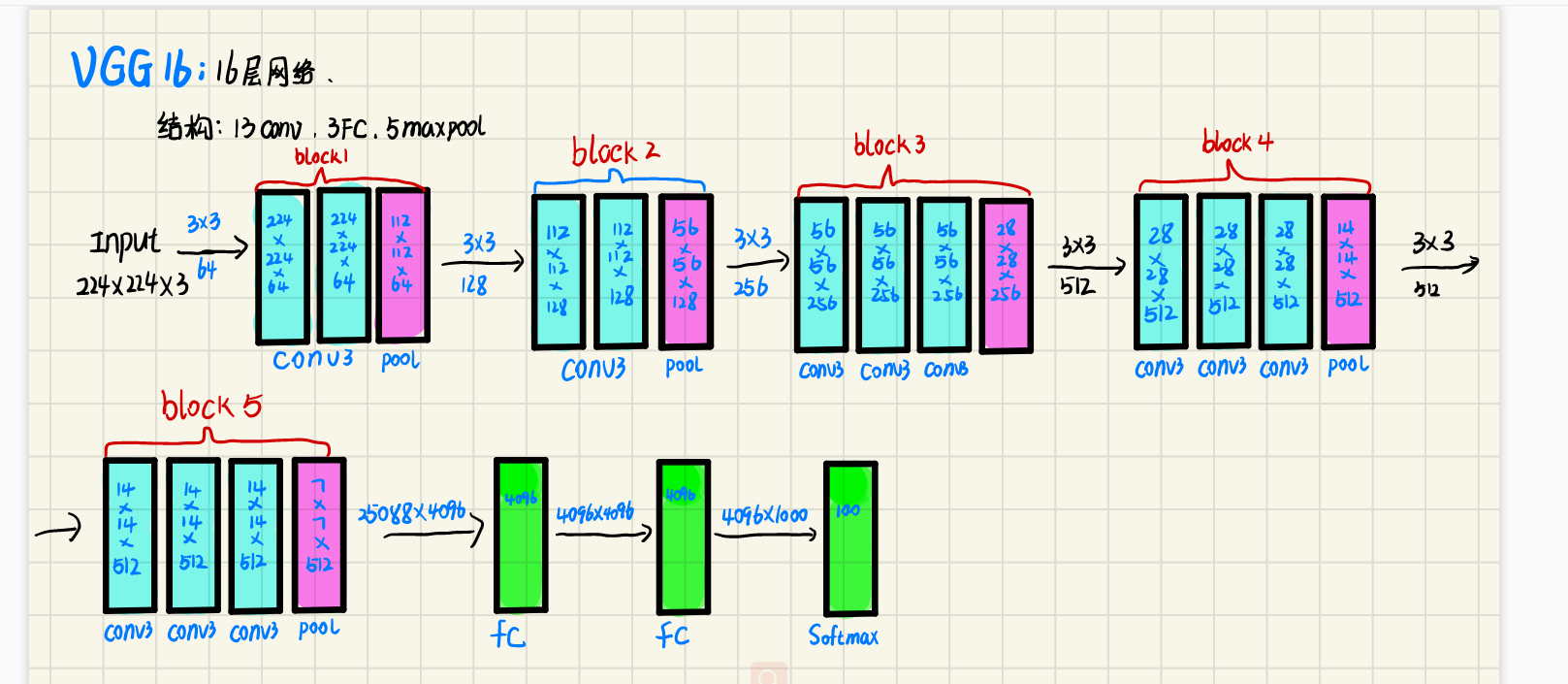
LeNet（姓名命名）、AlexNet（姓名命名）、Vgg、NiN（简写）、GoogLeNet（公司名，人名）、ResNet（残差）

LeNet、典型的VGG块、inception块、Res块

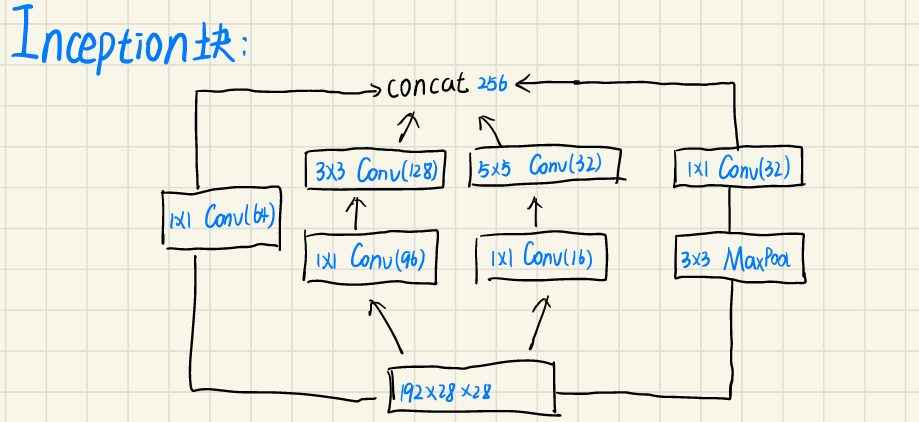
LeNet



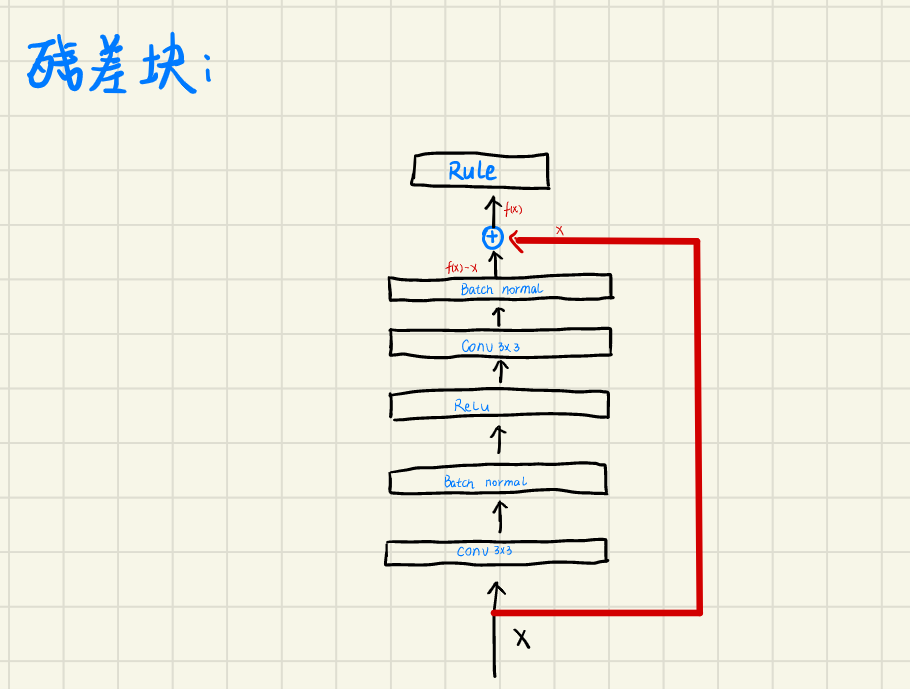
典型的VGG块



Inception



Res块



背

过拟与DA，常用的DA技术

裁剪，翻转，调颜色，角度

每个网络的名字的意义，历史脉络、网络深了是更容易过拟还是会减轻过拟。

**第10章：循环神经网络**

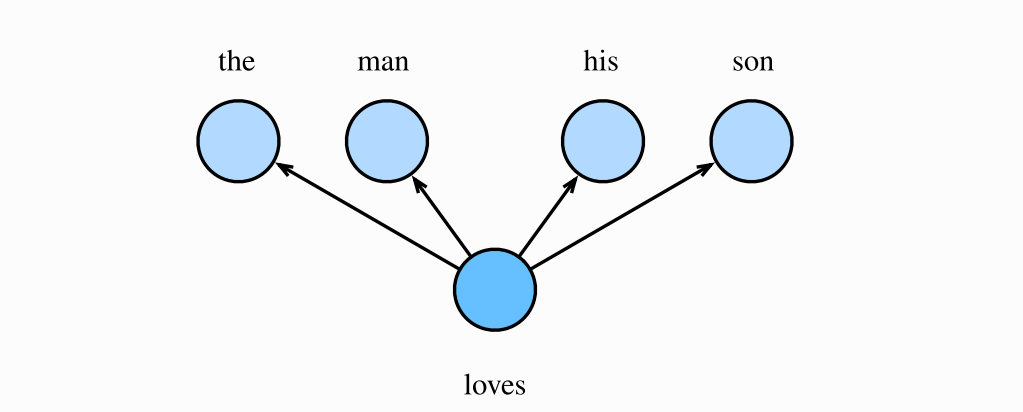
        N元语法、词向量

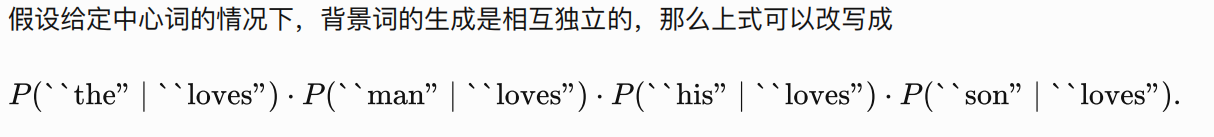
富不过三代：三元语法，2阶马尔科夫链

词向量：词向量的意思就是通过一个数字组成的向量来表示一个词（用实项量表示词，Ont-hot或标号不足以表现词的特征）

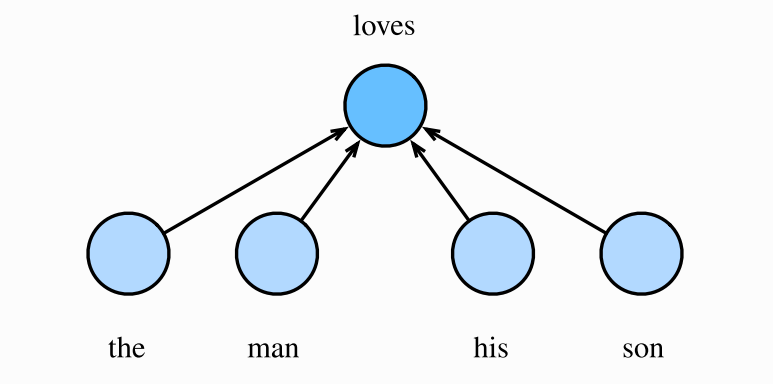
词向量的两种训练方式FF1A

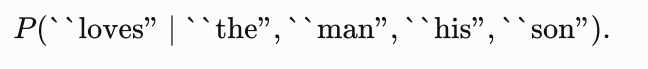
跳字模型：





连续词袋模型





        RNN、LSTM、GRU

