

Pytorch

深度学习入门与实战

讲师：日月光华



批标准化

讲师：日月光华 讲师QQ：984595060



什么是标准化

传统机器学习中标准化也叫做归一化，

一般是将数据映射到指定的范围，用于去除不同维度数据的量纲以及量纲单位。

什么是标准化

数据标准化让机器学习模型看到的不同样本彼此之间更加相似，这有助于模型的学习与对新数据的泛化

什么是标准化

常见的数据标准化形式：

标准化和归一化

将数据减去其平均值使其中心为 0，然后将数据除以其标准差使其标准差为 1。

什么是批标准化

Batch Normalization, 批标准化, 和普通的数据标准化类似, 是将分散的数据统一的一种做法, 也是优化神经网络的一种方法.

什么是批标准化

批标准化

不仅在将数据输入模型之前对数据做标准化。

在网络的每一次变换之后都应该考虑数据标准化

什么是批标准化

批标准化

即使在训练过程中均值和方差随时间发生变化，它也可以适应性地将数据标准化

为什么要做批标准化

批标准化解决的问题是梯度消失与梯度爆炸。

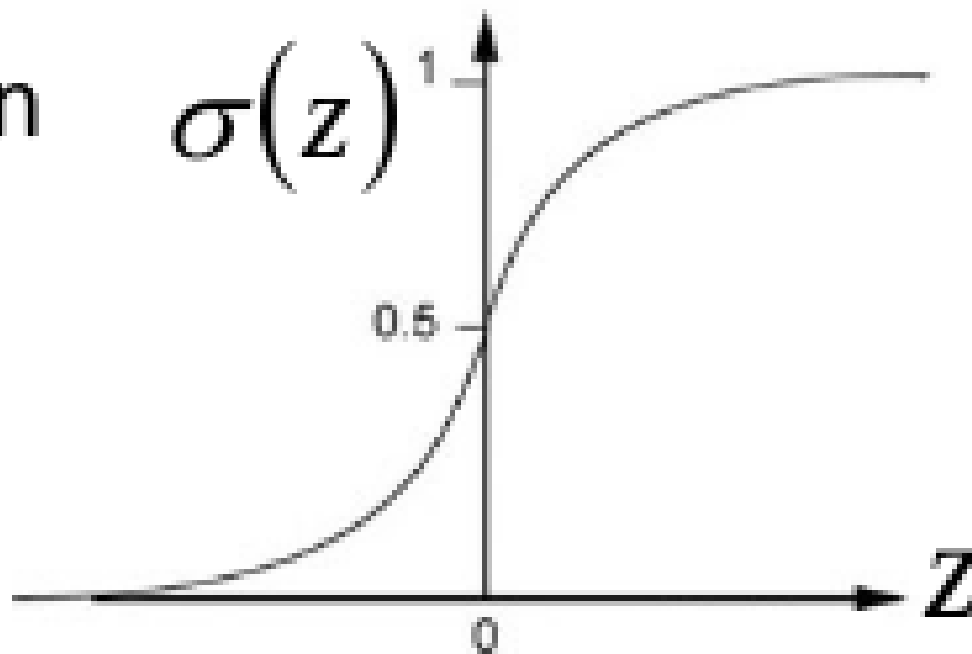
批标准化一种训练优化方法。

为什么要做批标准化

关于梯度消失，以sigmoid函数为例子，sigmoid函数使得输出在[0,1]之间。。

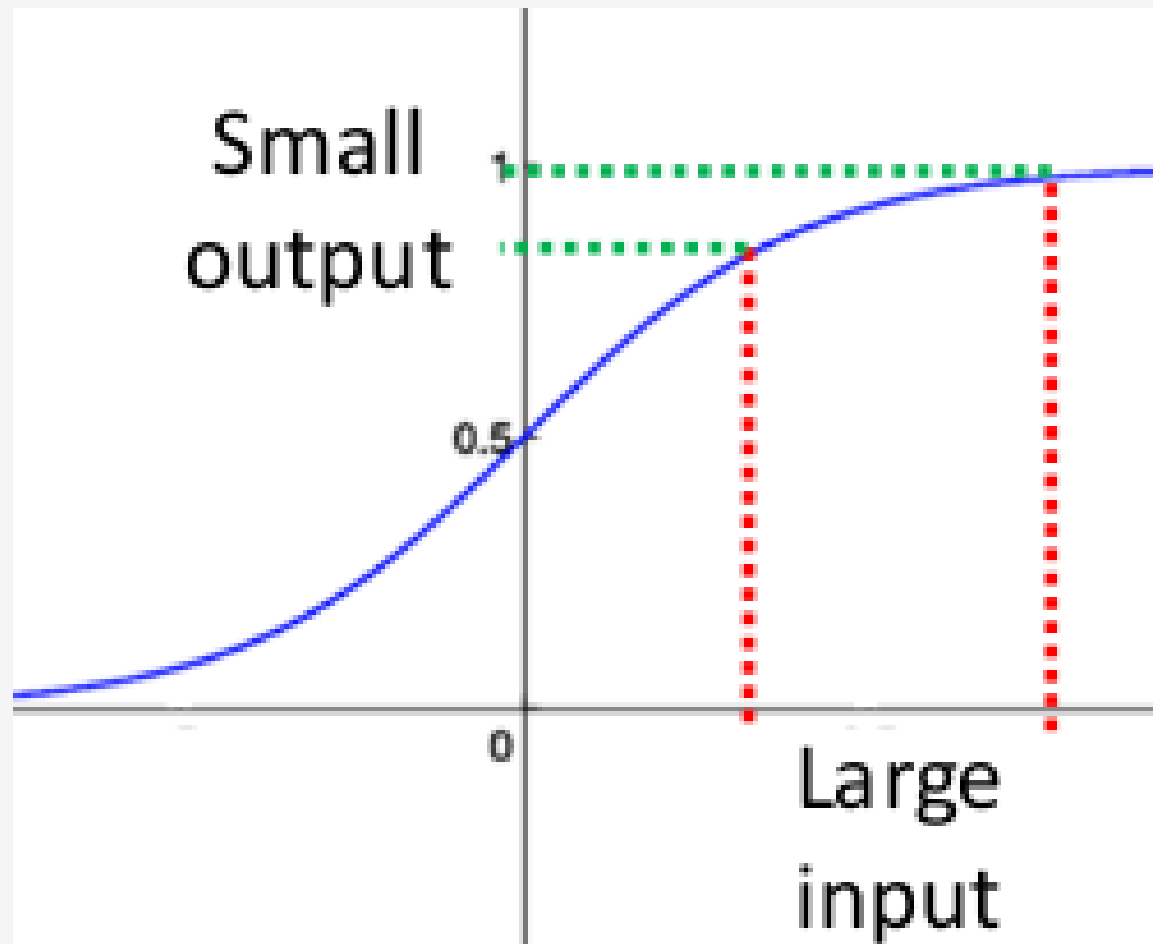
Sigmoid Function

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



为什么要做批标准化

如果输入很大，其对应的斜率就很小，反向传播梯度就很小，学习速率就很慢。



批标准化好处

我们知道数据预处理做标准化可以加速收敛，同理，在神经网络使用标准化也可以加速收敛，而且还有更多好处。

批标准化好处

具有正则化的效果

提高模型的泛化能力

允许更高的学习速率从而加速收敛

批标准化好处

批标准化有助于梯度传播，因此允许更深的网络。对于有些特别深的网络，只有包含多个 Batch Normalization 层时才能进行训练。

批标准化

BatchNormalization 广泛用于pytorch内置的许多高级卷积神经网络架构，比如 ResNet、Inception V3 等。

实现批标准化

BatchNormalization 层通常在卷积层或密集连接层之后使用。

`nn.BatchNorm1d()`

`nn.BatchNorm2d()`

批标准化的实现过程

1. 求每一个训练批次数据的均值
2. 求每一个训练批次数据的方差
3. 数据进行标准化
4. 训练参数 γ , β

批标准化的实现过程

5. 输出 y 通过 γ 与 β 的线性变换得到原来的数值

在训练的正向传播中，不会改变当前输出，只记录下 γ 与 β 。

在反向传播的时候，根据求得的 γ 与 β 通过链式求导方式，求出学习速率以至改变权值。

批标准化的预测过程

对于预测阶段时所使用的均值和方差，其实也是来源于训练集。比如我们在模型训练时我们就记录下每个batch下的均值和方差，待训练完毕后，我们求整个训练样本的均值和方差期望值，作为我们进行预测时进行BN的的均值和方差

批标准化的使用位置

`model.train()`和`model.eval()`

指示模型应在训练模式还是在推理模式下运行。

训练模式 将使用当前批输入的均值和方差对其输入进行标准化。

推理模式 将使用在训练期间学习的移动统计数据的均值和方差来标准化其输入。

批标准化的使用位置

原始论文讲在CNN中一般应作用与非线性激活函数之前，但是，实际上放在激活函数之后效果可能会更好。

谢谢大家

讲师：日月光华 讲师微信：louhh01

