

标准化

目录

◆ 什么是标准化

◆ 常见标准化方法

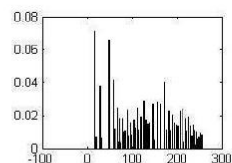
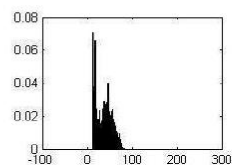
什么是标准化

什么是标准化

- ◆ 数据标准化 (Normalization) 或归一化可以增强信息的辨识度



线性归一化:

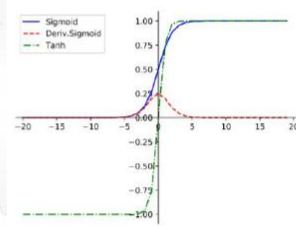


直方图均衡化

神经网络中标准化的作用

◆ 去除量纲干扰，保证数据的有效性，稳定数据分布

- 去除量纲的干扰，防止数值过小的特征被淹没
 - [年龄：30 身高：170 性别：1 年薪：500,000]
- 保证数据的有效性
 - 稳定前向传播激活值和反向传播过程中的梯度
- 稳定数据分布
 - 当深层网络中数据分布有明显的偏移，不利于网络学习

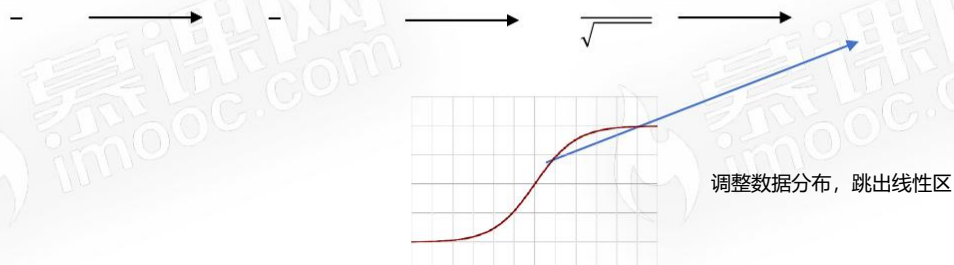


常见标准化方法

BN

◆ 批标准化方法 (Batch Normalization, 简称BN)

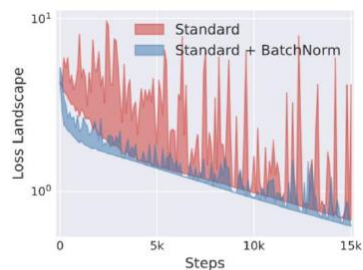
假设神经网络某层一个batch的输入为 $X=[x_1, x_2, \dots, x_n]$, 其中 x_i 代表一个样本, n 为batch size。数据维度为 (N, C, H, W) , 每一层normalization是基于 $N \times H \times W$ 个数值进行求平均以及方差的操作。



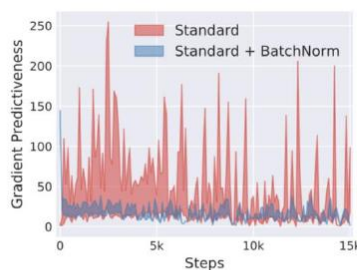
BN的作用

◆ 提高训练速度, 稳定模型训练

- 减轻了对参数初始化的依赖, 前向激活值与反向梯度更加有效。
- 平滑了优化目标函数曲面, 梯度更稳定, 可以使用更高的学习率, 从而跳出局部极值, 增强了泛化能力。



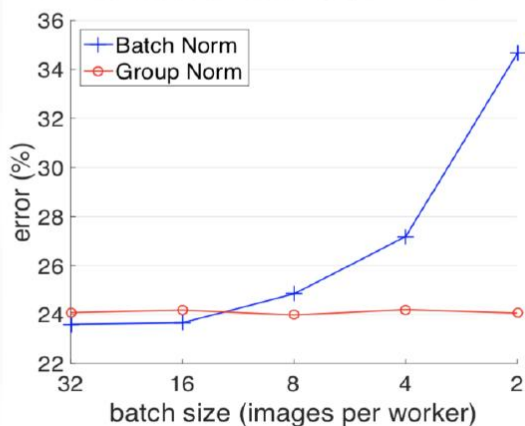
(a) loss landscape



(b) gradient predictiveness

BN的缺点

- ◆ 要求固定的Batch长度与均匀采样；batch过小数值计算不稳定



当batch减小到8以后，误差明显增大

BN的改进

- ◆ Batch Renormalization，增加样本相关的变换，让每个batch的均值和方差逼近真实分布

$$\mu_{\text{batch}} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \sigma_{\text{batch}}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu_{\text{batch}})^2$$

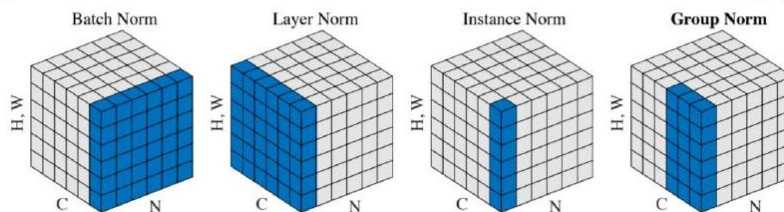
实际使用：

- (1) 先使用BN训练到一个相对稳定的状态
- (2) 稳定后再使用Batch Renormalization，r和d在一定大小范围内迭代

常见标准化方法的对比

◆ 令输入数据维度为(N,C,H,W)，区别在于计算的数据维度不同

方法	归一化的范围	特点
Batch Normalization	$N \times H \times W$	通用
Layer Normalization	$C \times H \times W$	适合非定长输入，NLP应用广泛
Group Normalization	$G \times H \times W$	适合小batch输入
Instance Normalization	$H \times W$	图像生成以及风格迁移类应用



下次预告：泛化与正则化