1. 过拟合产生的原因可能有哪些

数据集规模，模型复杂度，噪声，问题复杂度

1. 深度学习中常见的正则化策略有哪些，分别是什么思路

数据预处理（数据扩增等），参数正则化，Dropout，早停，批标准化

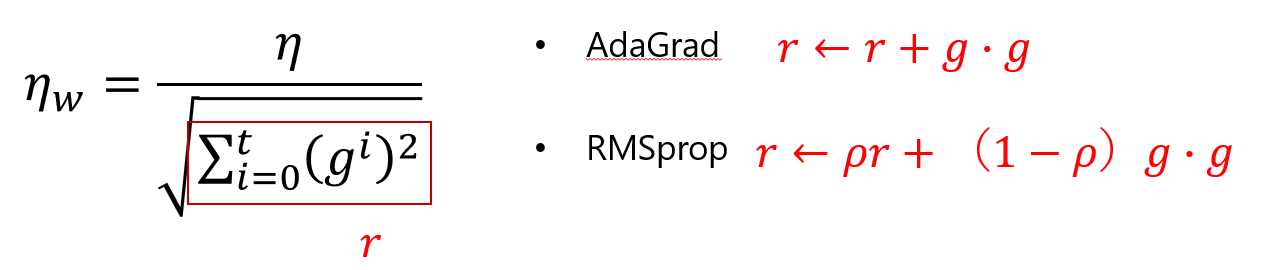
1. 描述Dropout的实现过程，什么时候Drop，什么时候不Drop

训练阶段P%的概率失活，测试阶段训练得到的参数\*（1-p%），不失活，可以看成集成了大量神经网络

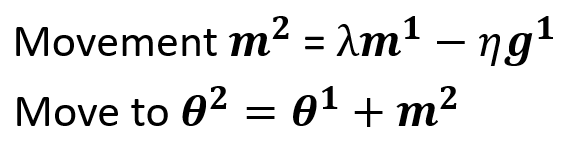
1. BGD，SGD，MBGD分别是什么

批量梯度下降，随机梯度下降，小批量随机梯度下降

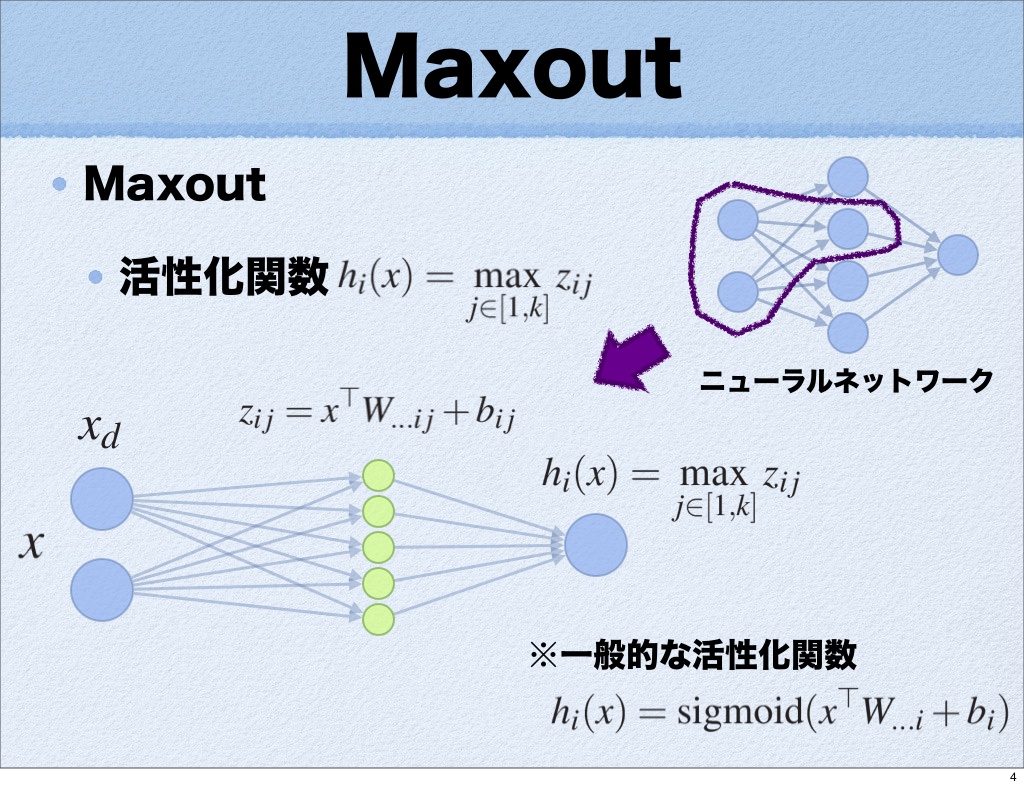
1. AdaGrad和RMSprop的思路、权重更新表达式和区别



1. SGDM（动量随机梯度下降）的思想和权重更新表达式



1. Maxout：

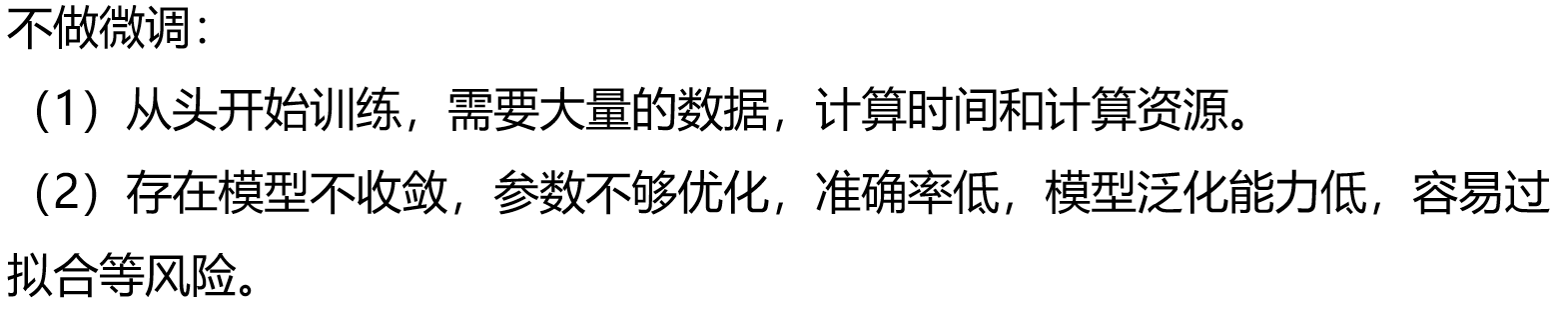


1. 预训练模型是什么，有什么特点

已经用数据集训练好了的模型，比如VGG16/19，Resnet等模型，这些通常是用大型数据集，如Imagenet, COCO等训练好模型参数

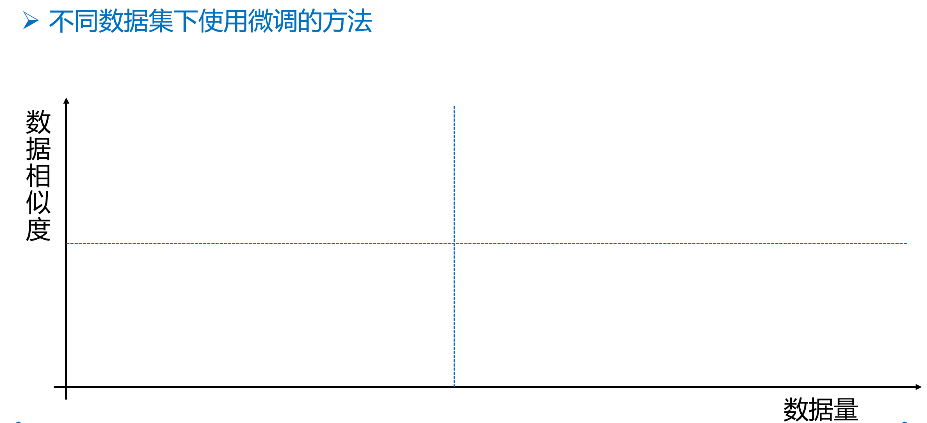
具备了提取浅层基础特征和深层抽象特征的能力

1. 不用微调有什么后果



1. 什么情况下使用微调

数据集相似度高，计算资源少

1. 

右上：保留初始权重，重新训练

左上：修改后面几层或者输出类别

右下：从头开始训练

左下：固定前面几层，训练后面几层

1. 描述模型微调的步骤

