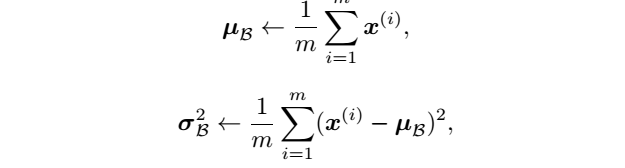
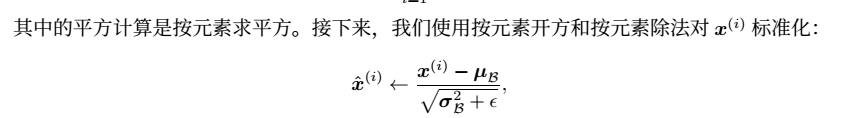
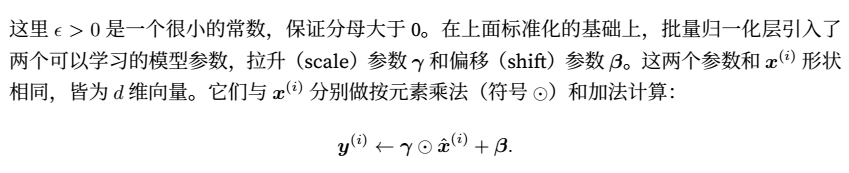
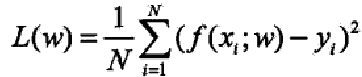
1. CNN中，对全连接层做批量归一化（BN）**。**设全连接层的输⼊为 u，权重参数和偏差参数分别为 W 和 b，仿射变换为。考虑由 m 个样本组成的小批量，   ，

则求小批量 B 的均值、⽅差。如果按元素开⽅和按元素除法对 x(i) 标准化结果为？如果引入拉升和偏移，参数 分别为γ和 β，求BN层输出？







1. 正则化的作用主要是什么？如果已知经验风险是， ，

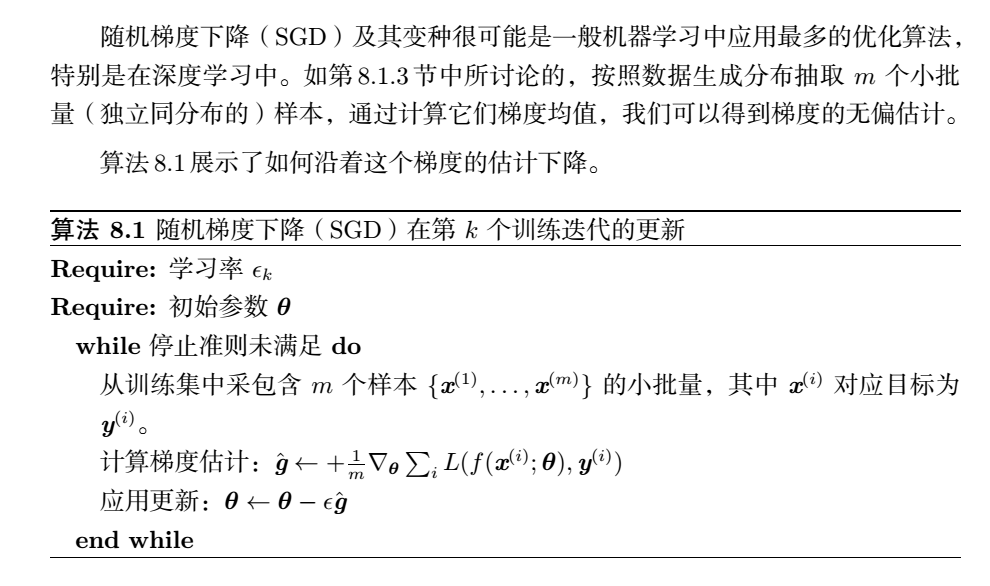
请分别给出和范数的正则化形式。

答：

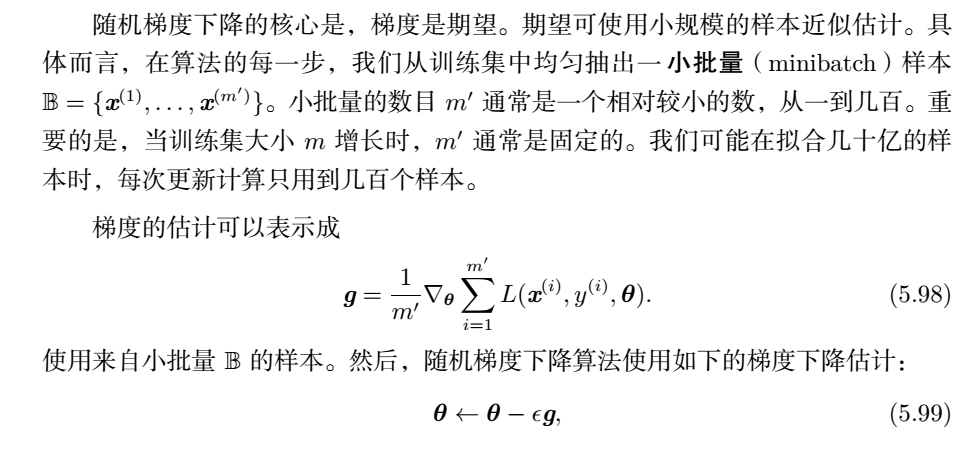
* 正则化通过为模型损失函数添加惩罚项使得学出的模型参数值较小，是应对过拟合的常⽤⼿段。/正则化是指我们修改学习算法，使其降低泛化误差而非训练误差。(**以上两种解法都ok**)
* 范数正则化
* 范数正则化 

1. 请简单给出随机梯度下降SGD算法的原理，给出它的梯度下降估计。

答：



或者



**以上两种解法都ok**。