模式识别期末论文 张祎扬 181840326

论文：Power Mean SVM for Large Scale Visual Classification

**Observe the problem:**

在目前的视觉分类任务中，处理非常庞大的多维度数据集是非常重要的。在最近的研究进展中，线性分类器可以高效地处理这种庞大的数据集，但是，在视觉任务中，线性分类器通常具有较低的精度。但是另一方面，非线性分类器虽然可能在精度上有更好的表现，但是通常需要花费大量时间来训练。因此，在庞大数据集的视觉分类任务上，缺乏一种同时兼顾精度和高效的分类算法。

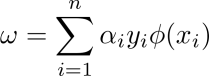
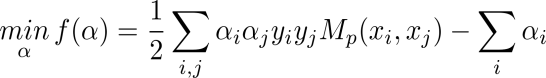
这篇论文主要提出了幂平均值支持向量机算法（即PmSVM），当视觉分类任务的数据集非常大时，和最先进的线性/非线性SVM方法相比，它不仅运算速度快得多，而且有更高的精确度，所以可以为上述问题提供一种解决方案。

**Find the idea:**

在计算机视觉任务的SVM分类器中，大多数情况下，加性核的准确率比点积核高的多，并且加性核实现了高精度、训练速度和测试速度的一个平衡。通常使用的HI内核和卡方内核都是加性核。所以，想要得到一个较为合理的解决方案，寻找一个较为普适的加性核是比较合理的。因此，在这个基础上，作者提出了一个新颖的幂平均值内核，并且将HI和卡方内核作为幂平均值内核的两种特殊情况（即参数P取特定值时的情况）。

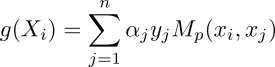
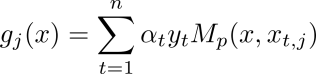
**Formalize the idea:**

作者将PM函数定义为wpsoffice,那么卡方核就是p取-1的情况，HIK就是P取负无穷的情况。作者在论文中提出了一个定理：当p≤0且p≥负无穷时，PM函数的向量形式（因为是加性核）是正定的。作者在文中对该定理进行了证明，因为过程较为易懂，所以这里不加赘述。该定理说明了PM核函数满足Mercer条件。

在课上学过的知识中，我们知道在核SVM的对偶形式中，当核函数满足Mercer条件，可以得到,所以学习过程可以形式化为

**Simplify the formulation:**

为了解决形式化之后的问题，可以使用坐标下降法，用梯度G来度量wpsoffice随着wpsoffice的改变，并且每次更新参数的值，直到迭代找到最优解。这种方法的复杂度是O(nd).为了简化问题，作者提出了梯度估计的方法，它可以把复杂度降到O(d).

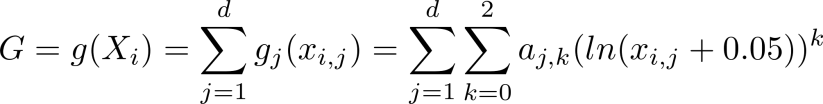
在坐标下降法中，求梯度中的主要成分是.将xi的第j个维度的元素表示为xi,j，并且将第j维度的和表示为单一标量输入的函数.

所以，问题就简化为对所有的j快速计算这种标量函数，并且它的复杂度是O(1),大大提高了性能。根据一些测试，作者发现曲线可以很好地描述标量函数的性能：对所有的j，它都是光滑的单调函数。因此，作者提出可以用低阶的多项式回归来估计曲线。事实上，二次函数已经可以非常精确地估计原函数了。此外，还必须强调两点：

1. 通过最小二乘思想可以解出最优参数/private/var/folders/q1/mr8y2cs53gggq1tjshxm61sh0000gn/T/com.kingsoft.wpsoffice.mac/wpsoffice.wRmzgBwpsoffice
2. 作者从实验中观察到用ln(c+0.05)来代替c作为输入会得到更好的估计。

**Solve the problem:**

至此，PmSVM算法的所有细节都解释完毕了。在PmSVM算法中，其主要思想仍然是梯度下降，不过作者巧妙地定义了标量函数，并且用二次多项式对函数进行拟合，只需要三个点，就可以很精确地估计梯度的值，将算法的复杂度从O(nd),降为O(d).通过这样巧妙的操作，就在加性核高精确度的基础上，改进算法，克服了加性核复杂度高的缺点，实现了精度和效率兼得的效果。

最终改良后的梯度计算公式是,在循环中不断计算梯度，更新参数，得到最优解就是训练分类器的全过程。在该算法中，预测的结果就是sgn(g(q))(q为测试样本).论文中都详细写出了算法步骤，因此在这里只对关键步骤作了解释。

PmSVM的训练和测试都是十分高效的。在PmSVM中，当p=-1时，就是卡方内核。不过实际上我们并不需要将p取到负无穷，-8已经可以足够精确地模拟HIK。同时需要指出，由于PmSVM的计算特性，它需要使用更多的内存。

最后，作者分别在三个数据集上对不同的方法进行了测试，并通过分析结果证明了PmSVM算法的高效性和准确性都是非常出色的。不过，当数据集越来越大而不能储存在主存的时候，PmSVM算法就会遇到一些问题，因此，压缩并利用压缩后的数据集以及使用硬件驱动中的存储空间都是该算法可以继续深入研究的方向。

**Discussion:**

因为篇幅有限，实在很难三言两语就讲清楚论文的精髓。而且由于word文章格式的限制，插入latex公式会稍稍扰乱排版，因此我尽量避免了繁琐的计算细节的解释，只解释了最重要的几个公式。

作者看到了视觉分类任务中传统SVM算法的缺点，并且提出了PmSVM算法来综合各种方法的优点，达到了效率和准确率都比较理想的一个目标。不过PmSVM在内存上比传统的线性SVM要消耗更多的内存，这在数据集越来越大的时候会更为明显。因此，我觉得也许还可以在PmSVM算法上适当改进，以缩小使用内存的规模，这也许可以通过合理优化一些参数的存储来实现。

其次，论文中基本上是直接提出了用二次函数拟合，我觉得可以在效率允许范围内适当讨论一下不同阶数的多项式对拟合精确度的影响。

我能力有限，论文已经极大地开阔了我的眼界了，实在是提不出别的有价值的想法了。

注：我不知道用latex插入的公式在别的电脑上会不会显示错误，所以特意将关键部分截图放在论文的下一页，以保证排版的清晰。希望老师不要误以为我的论文有三页，其实只有两页不到:）

