**目标函数**

其中，X表示训练样本，每行是一个样本，Y是样本的类别标签，为我们要求的系数矩阵。

**目标函数**第一项表示多元线性回归，第二项是F范数惩罚项，增加方法的鲁棒性，第一项和第二项构成岭回归，第三项为核范数，起低秩约束作用，称这个模型为核范数低秩岭回归，用NNLR表示。

使用Xβ的核函数是基于一个结论，即，对线性回归模型施加

和是等价的，定理的证明见论文[1]。这样做的优点是模型可以直接给出解析解，避免了迭代求解，因此计算速度更快。

%{因为X中的向量通常存在非常多的相关性，因此X通常是低秩的，所以，参数也应该是低秩的，因为,又因为核范数是秩的凸近似，所以增加第三项来进行低秩约束。}%

**低秩回归模型提出的背景**是在数据特征维度高，标签类别个数多的情况，数据维度高会产生维度灾难问题（本质是样本数量少，难以刻画数据的分布），通常的解决方法是进行降维，而低秩回归恰好可以实现降维的作用，理论证明可见论文[2].从另一个角度看是因为类别之间存在相关性，使得系数矩阵具有了低维结构。

**低秩回归的定义**为β的秩小于p和c，其中p是特征维度，c是标签类别个数。

**模型求解**

1. **迭代求解：**

**增加变量**，使得核函数中的β和F范数损失项中的β去耦合，以便进行迭代求解。然后使用**增广拉格朗日方法**将目标函数变温无约束问题，最后，**逐变量进行迭代求解**

首先，增加约束条件，则原问题变为

根据增广拉格朗日乘子法【附录1】，可将上式化为如下形式

（1）固定和A，求解M

上式等价于

这种核范数与F范数之和的形式可以直接给出最优解的形式；

其中，。

（2）固定M和A，求

令0，可得：

（3）A是拉格朗日乘子项，以如下形式进行迭代

一般的，初始取0.1， 取1.1-1.2。

M，和A矩阵可初始化为0矩阵，依次迭代即可求得最优解。

**2.解析解**

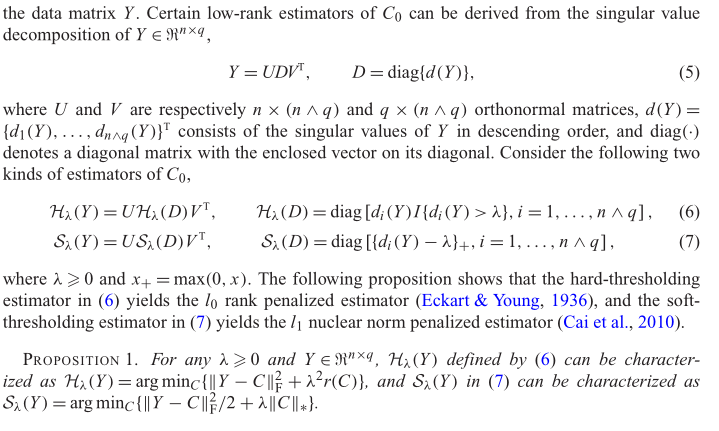
将岭回归项按照F范数的性质放到第一项中，然后将核函数项中的X替换为第一项加入岭回归项后的X的形式，进而用软值域奇异值收缩估计方法直接求解。

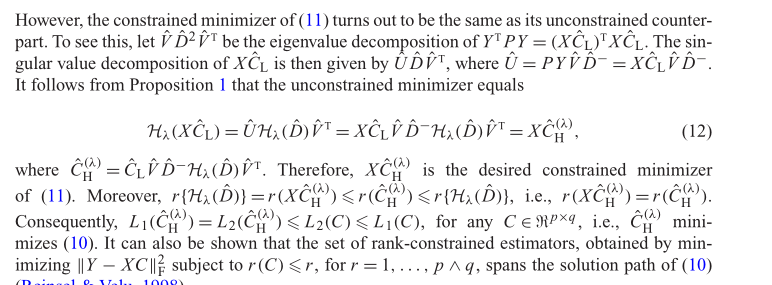
其中

上式进一步化简为

其中 表示标准线性回归的解，这一步化简的依据是三角不等式。

再根据软值域奇异值收缩估计方法,参考论文[3]。直接给出解。





的特征值分解为

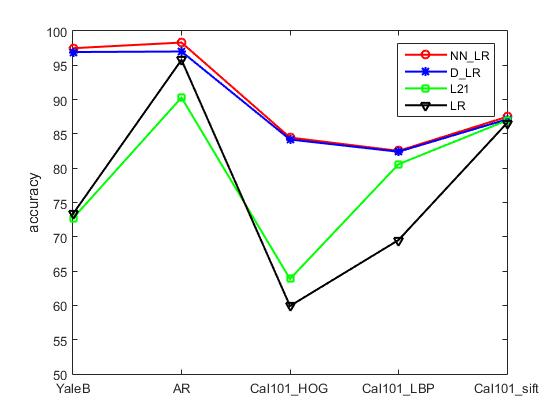
**多视觉形式**

表示第v个视觉的数据，表示第v个视觉的参数矩阵。求解时，每个视觉分别按照单视觉模型求解，完成后，将所有视觉的结果求和，即为最终的结果。

**实验**

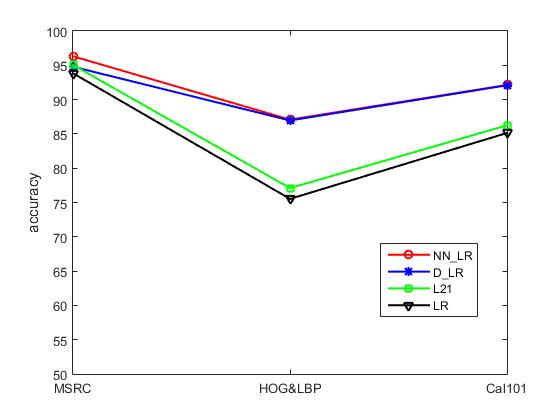
1. 单视觉实验

实验数据库为YaleB，AR和Cal101，对比算法选择低秩分解岭回归D\_LR，L21岭回归和线性回归。对于YaleB和AR库，每类选择50%的数据作为训练数据，剩余的作为测试数据，Cal101库每类随机选择40个样本做训练数据，剩余的作为测试数据，随机的选取15组数据，完成实验，结果求均值，结果如下：



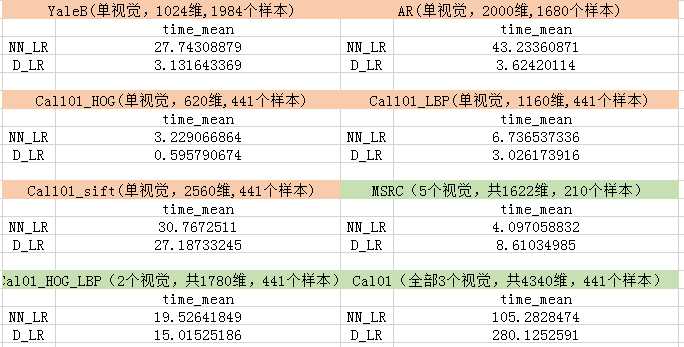
1. 多视觉

实验数据库选择MSRC和Cal101，分别用MSRC，Cal01中的HOG和LBP两个视觉，Cal101全部3个视觉做三组对比实验。对比算法与单视觉相同，MSRC每类选择50%的数据作为训练数据，剩余的作为测试数据，Cal101的选择与单视觉相同，随机的选取15组数据，完成实验，结果求均值，结果如下：



可以看出，无论是单视觉还是多视觉，增加低秩约束的NN\_LR和D\_LR方法相较线性回归和L21有更好的效果，NN\_LR方法比D\_LR方法效果更好。

从算法效率上来说，NN\_LR使用迭代求解，D\_LR有直接的解析解，所以当数据量较大，维度较低时，D\_LR方法效率更高，当数据量较小，维度较大时，NN\_LR方法效率更高。



附录

1.增广拉格朗日方法，wiki百科https://en.wikipedia.org/wiki/Augmented\_Lagrangian\_method