# 多核学习概述

SVM 是一种最小化结构风险的机器学习算法，克服了经验风险最小化机器学习算法所带来的分类函数推广能力差，即分类器泛化能力差的缺憾。SVM 建立在统计学习理论的VC 维和结构风险最小化原理基础上，能根据有限的样本信息在模型的复杂性（即对特定训练样本的学习精度）和学习能力（即无错误地识别任意样本的能力）之间寻求最佳折衷，获得最好的推广能力也即泛化能力。

作为一种小样本的学习模型，SVM在处理小样本问题、非线性可分数据和高维空间数据集的分类和回归问题上具有独到的优势，是一个具有最下分类误差率和最大范化能力的强有力工具。使得它成为广受欢迎的模式识别和机器学习模型，被应用于诸多领域。然而，很多实际问题当中的样本特征属性之间存在异构性，使得通常所使用的单一核函数的SVM方法效果大打折扣。因而有学者针对此问题，提出了多核学习方法（Mutiple-Kernel-Learning）。

多核学习方法（Mutiple-Kernel-Learning）是在已有SVM学习模型的基础上，针对单一核在处理异构信息的局限性，提出利用多个基本核函数的组合来形成新的核函数，从而将特征空间映射到更易区分的特征空间当中的学习方法。

# 多核学习发展过程

目前为止，针对多核的研究主要集中在两个方面。其一聚焦于基本核的组合形式上；其二则聚焦于多核学习的参数结算问题上。

关于多核学习问题的研究可以追溯到2001年，Crammer等人针对多核学习方法的损失函数作了研究，但是并未将工作重心放到具体的分类和回归问题上来。Lanckriet等人在2004年的研究中解决了如何针对具体的问题构造合适的MKL公式并且利用SDP和QCQP方法加以求解。然而，该论文中的方法却不能解决大规模的学习问题。最早的具有实用意义的MKL算法在Bach2004年发表的论文中被提出。该论文同时还提出了如何利用L1约束生成一个稀疏解。2006年，Sonnenburg等人在保留L1约束的情况下，使用SILP方法求解了同样的问题，成功将多核学习应用推广到百万级的数据集中。2008年，利用梯度下降法，效率更高的算法被提出（Rakotomamonjy et al.,2008;Varma & Ray,2007）。而后，Bach又提出了实现大规模核函数组合的可行性算法。此外，还有人针对超核方法和多类的MKL做了研究。

# 多核学习主要问题

## 3.1 核函数形式问题

### 3.1.1 线性组合核函数



直接将多个基本核进行线性组合，这种方式常常用在对单核作SVM输出后进行加权组合，这时可以连接大量的基本核函数，可以避免最佳参数的手工搜索。或者用于特征的选择上，这时一般一个特征对应一个核函数，每个核函数赋予一个权值di，代表着对应特征的重要性，如果最优化目标函数中的regurilization项选择的是L1的约束，那么最终解出的解将是一个稀疏的解（**d**）。这样，d值为0或者d值非常小的特征就被剔除掉了。

### 3.1.2 通用组合形式核函数





已被证明，核函数的和、核函数的乘积以及它们的组合仍是核函数（Mercer定理）。所以可以根据需要选择自己的多核组合形式。但是必须满足的条件是，组成的多核必须是半正定的并且它对d的导数存在且连续。通用核函数可以把原始特征映射到更高的空间中去。也可用于特征选择问题（More Generality in MKL）。

## 3.2 核参数解算问题

参数结算归结为两大类方法：

1. SMO算法基础下的结算方法

该方法[MKL and the SMO Algorithm]引入拉格朗日乘子，通过最佳化步骤消除d，利用SMO算法解得所有的SVM模型参数（）后，再带入公式算出核函数参数（d）来。



1. 迭代思路下的结算方法

基于这样的基本思路：先给一个d的初值，这样剩下的最优化问题就可以利用已有的标准SVM解算工具解算。但是针对不同的核函数组合形式和对应的约束（L-P），研究者们又做了不同的优化。比较典型的有SimpleMKL、Shogun、SPG-GMKL等。

# 多核问题实际应用实例

## 4.1 多核作用的层次

### 4.1.1 先单核作SVM再后期组合

若是线性组合，这种方式最易于理解，它只是将多个核分别作SVM模型训练，再在模型输出前一步将它们作线性组合，其中的核参数d则代表着对应核的权重。

若是非线性组合核函数，那么对应于将原有特征映射到一个新的综合作用下的特征空间。

### 4.1.2先作用于局部特征后再组合

典型的应用是特征选择问题。

### 4.1.3 局部样本采用不同的核

该种组合多核学习，与考虑特征属性间的异构性不同，它还考虑到样本可能存在的异构性。一种极端的方式是给每一个样本套一个多核组合，然后再进行核矩阵的计算。

详细见论文：

Group-Sensitive Mutiple Kernel Learning For Object Categorization