

## Part I. SVM

SVM学习的基本想法是求解能够正确划分训练数据集并且几何间隔最大的分离超平面，要求数据集先行可分

定义距离  $\gamma_i := y_i (\frac{\omega}{\|\omega\|} \cdot \mathbf{x}_i + \frac{b}{\|\omega\|})$

其中  $y_i$  代表分到的类别 (1/-1)

分类要使得 $\min(\gamma_i)$ 最大

(投影到一直线，使得其中两类中间有一段差距极大)

转化成达到同一极大值的最小  $\omega$

拉格朗日化(发现是凸函数)+计算(求偏导)

$$\begin{aligned} \max_{\lambda} (\sum \lambda_i - \frac{1}{2} \sum \lambda_i \lambda_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j)) \\ s.t. \sum \lambda_i y_i = 0 \end{aligned}$$

再用SMO算法(线性代数)

对于线性不可分的，用映射使其变成高维可分

对于误差较大，引入误差。

## Part II. TMVA

TMVA是集成到分析框架ROOT中，是一个包含大量多元分类算法的工具箱。

含有:

矩形切割优化

射影似然估计

多维似然估计

线性和非线性判别分析

人工神经网络

支持向量机

增加/袋装决策树

通过规则集成的预测学习

一个通用的增强分类器，允许增强上面任何一个分类器

一个通用类别分类器，允许一个人分割训练数据为不相交类别与

独立的mva。

TMVA过程:

