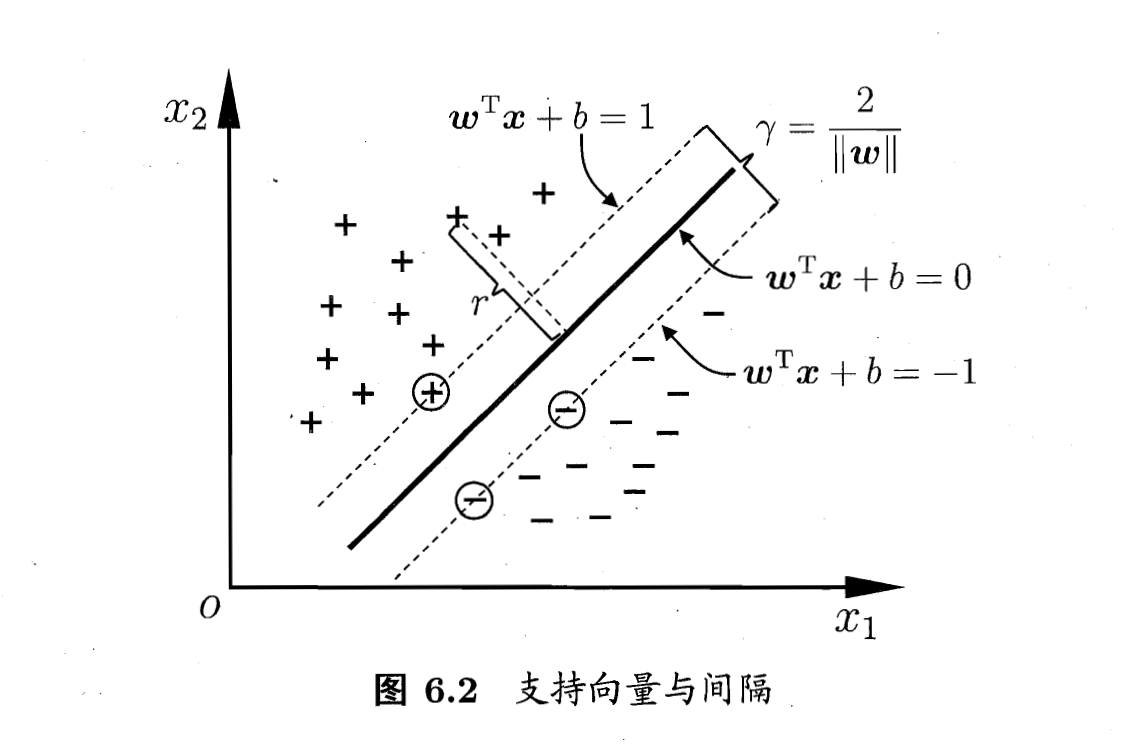
# 6.1间隔与支持向量机

分类学习最基本的思想是在样本空间中找到一个划分超平面，将不同类别分开。在诸多满足条件的超平面中选择**最鲁棒的**，**泛化能力最强的**。

**支持向量**：距离超平面**最近的几个样本点**，且上述点被超平面正确划分。

（每个样本点都对应一个特征向量）

**间隔**：两个异类支持向量到超平面的距离之和，拥有**最大间隔**的划分超平面才是最鲁棒的。



# 6.3核函数

对于在**原始样本空间**中线性不可分问题（无法用一个超平面划分），可以将样本从原始空间映射到一个**更高维度的空间**，使得样本在该空间内线性可分。

在更高维空间中求解最鲁棒超平面的过程中需要计算**两个样本在该高维空间内的内积**，该计算步骤通常是困难的，因此引入**核函数**，两个样本在高维空间中的内积等于在**原始空间中通过核函数计算的结果**。当我们已知到高维空间的映射形式，则可以写出对应核函数。任何一个核函数都隐式地定义了一个特征空间。核函数选择的不合适意味着将样本映射到了一个不合适的特征空间，很可能导致性能不佳。

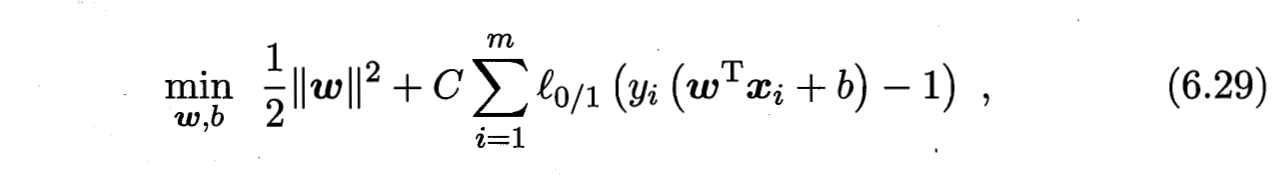
常见核函数：线性核、多项式核、高斯核、拉普拉斯核、Sigmoid核。

将上述核函数通过函数组合也能得到核函数。

# 6.4软间隔和正则化

对于样本原始空间中**线性不可分**时，前述映射到高维特征空间的方法中**核函数很难找到**，并且即使找到了在某高维特征空间中样本线性可分，也不能确定是否发生了**过拟合**。因此提出了区别于映射到高维空间的另一种缓解方案。

软间隔：区别于之前的硬间隔，软间隔**允许某些样本不满足约束**，拥有最大间隔且不满足的样本最少的超平面是最好的超平面（计算公式中将两种代价简单相加）。



由于l0/1函数非凸、非连续，数学性质不好，往往采用一些函数替代，称为“**替代损失函数**”。

# 6.5支持向量回归

SVR的特别之处在于当f(x)与y的差别大于某一**阈值**时才计算损失，而不是f(x)与y不相等就计算损失。

# 6.6核方法

核方法：基于核函数的学习方法如通过引入核函数（核化）来将线性学习器拓展为非线性学习器。