**感知机**：

**距离公式**：空间中任意一点到超平面的距离：

其中为的范式。（范式指向量各元素平方和的平方根）

推导：

取空间中任意一点，超平面S：，其中、、均为N维向量。设点到平面的距离为，点在平面S上的投影点为，则满足。因为向量平行于平面的法向量，故有

**损失函数**：误分类点到超平面的距离之和。

**梯度**：

**更新**：随机选择一个误分类点()，对进行更新：

其中为样本标签

**对偶感知机**

初始时设置均为0，逐步修改n次，则关于()的增量分别是和，这里。即最后学习到的分别表示为：

这里，当时，表示第个实例点由于误分而进行更新的次数。实例点更新次数越多，意味着它距离分离超平面越近，也就越难正确分类。换句话说，这样的实例对学习结果影响最大。

**误分条件**：

其中，为常数，因此需要计算

…

**引入Gram矩阵**：

即上述计算为：

[

定理2.1(2)：感知机误分类的次数是有上界的。（证明跳过）