1. 闵可夫斯基距离(Minkowski Distance)

严格意义上，闵氏距离不是一种距离，而是一组距离的定义。

(1)闵氏距离的定义:

两个n维变量A(x11,x12,…,x1n)与 B(x21,x22,…,x2n)间的闵可夫斯基距离定义为

其中p是一个变参数。

当p=1时，就是曼哈顿距离

当p=2时，就是欧氏距离

当p→∞时，就是切比雪夫距离

根据变参数的不同，闵氏距离可以表示一类的距离。

2.欧氏距离(Euclidean Distance)

欧氏距离（L2范数）是最易于理解的一种距离计算方法，源自欧氏空间中两点间的距离公式

1. 二维平面上两点a(x1,y1) 与 b(x2,y2)间的欧氏距离：
2. 三维平面上两点A(x1,y1,z1) 与 B(x2,y2,z2)间的欧氏距离：
3. 两个n维向量A(x11,x12,…,x1n)与 B(x21,x22,…,x2n)间的欧氏距离：

表示成向量运算的形式：

3.曼哈顿距离(Manhattan Distance)

从名字就可以猜出这种距离的计算方法了。想象你在曼哈顿要从一个十字路口开车到另外一个十字路口，驾驶距离是两点间的直线距离吗？显然不是，除非你能穿越大楼。实际驾驶距离就是这个“曼哈顿距离”(L1范数)。而这也是曼哈顿距离名称的来源，曼哈顿距离也称为城市街区距离(City Block distance)

1. 二维平面两点A(x1,y1) 与 B(x2, y2)间的曼哈顿距离：
2. 两个n维向量A(x11,x12,…,x1n) 与 B(x21,x22,…,x2n)间的曼哈顿距离：

4.切比雪夫距离(Chebyshev Distance)

国际象棋玩过么？国王走一步能够移动到相邻的8个方格中的任意一个（如图1.11）。那么国王从格子(x1,y1)走到格子(x2,y2)最少需要多少步？自己走走试试。你会发现最少步数总是max(| x2-x1| , |y2-y1| ) 步。有一种类似的一种距离度量方法叫切比雪夫距离(L∞范数)。

1. 二维平面两点A(x1,y1)与B(x2,y2)间的切比雪夫距离：
2. 两个n维向量A(x11,x12,…,x1n) 与 B(x21,x22,…,x2n)间的切比雪夫距离：
3. 这个公式的另一种等价形式是：

 5. 夹角余弦(Cosine)

几何中夹角余弦可用来衡量两个向量方向的差异，机器学习中借用这一概念来衡量样本向量之间的差异。  
(1)在二维空间中向量A(x1,y1)与向量B(x2,y2)的夹角余弦公式：

1. 两个n维样本点A (x11,x12,…,x1n)与 B(x21,x22,…,x2n)的夹角余弦  
   类似的，对于两个n维样本点A(x11,x12,…,x1n)与 B(x21,x22,…,x2n)，可以使用类似于夹角余弦的概念来衡量它们间的相似程度。

即

夹角余弦取值范围为[-1,1]。夹角余弦越大表示两个向量的夹角越小，夹角余弦越小表示两向量的夹角越大。当两个向量的方向重合时夹角余弦取最大值1，当两个向量的方向完全相反夹角余弦取最小值-1。  
6. 汉明距离(Hamming distance)  
(1)汉明距离的定义  
两个等长字符串s1与s2之间的汉明距离定义为将其中一个变为另外一个所需要作的最小替换次数。例如字符串“1111”与“1001”之间的汉明距离为2。  
应用：信息编码（为了增强容错性，应使得编码间的最小汉明距离尽可能大）。  
7. 杰卡德相似系数(Jaccard similarity coefficient)  
(1) 杰卡德相似系数  
两个集合A和B的交集元素在A，B的并集中所占的比例，称为两个集合的杰卡德相似系数，用符号J(A,B)表示。

杰卡德相似系数是衡量两个集合的相似度一种指标。

(2) 杰卡德距离  
与杰卡德相似系数相反的概念是杰卡德距离(Jaccard distance)。杰卡德距离可用如下公式表示：

杰卡德距离用两个集合中不同元素占所有元素的比例来衡量两个集合的区分度。  
  
(3) 杰卡德相似系数与杰卡德距离的应用  
  
可将杰卡德相似系数用在衡量样本的相似度上。  
  
样本A与样本B是两个n维向量，而且所有维度的取值都是0或1。例如：A(0111)和B(1011)。我们将样本看成是一个集合，1表示集合包含该元素，0表示集合不包含该元素。  
  
P：样本A与B都是1的维度的个数  
  
q：样本A是1，样本B是0的维度的个数  
  
r：样本A是0，样本B是1的维度的个数  
  
s：样本A与B都是0的维度的个数

那么样本A与B的杰卡德相似系数可以表示为：  
这里p+q+r可理解为A与B的并集的元素个数，而p是A与B的交集的元素个数。  
而样本A与B的杰卡德距离表示为：

现在，我们有能力为矩阵中对象间的相似程度(接近与远离)提供各种度量方法，以及编码实现。通过计算对象间的距离，我们就可以轻松地得到表2.8中的四个对象所属的类别：以克、天为单位的苹果是水果类别的一个实例; 以吨、年为单位鲨鱼是大型动物的一个实例。这种区别是明显的，但是，如果我们考察颜色这个特征，情况可能会有所不同，苹果和梨都有黄色这个特征，像这种情况我们如何区分呢？