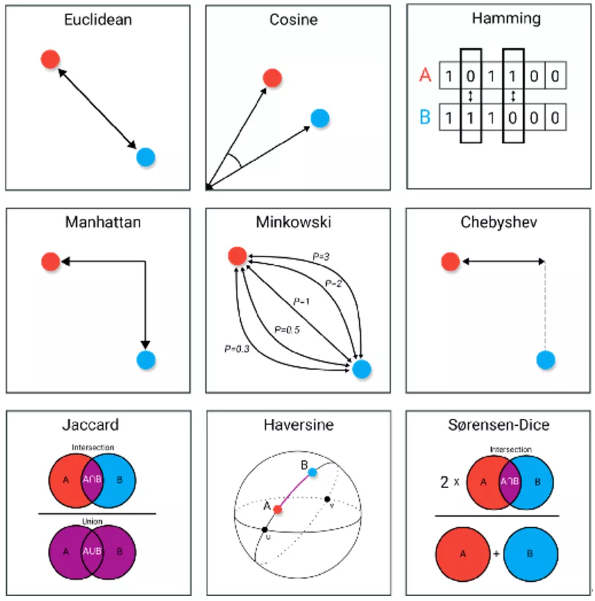
距离度量与相似度算法原理与实践

本文涉及到的距离度量方法：

1. 欧氏距离
2. 曼哈顿距离
3. 闵氏距离
4. 切比雪夫距离
5. 标准化欧氏距离
6. 马氏距离
7. 汉明距离
8. 编辑距离
9. DTW距离
10. 杰卡德相似系数
11. 杰拉德距离
12. Tanimoto系数（广义Jaccard相似系数）
13. 余弦距离
14. 皮尔逊相关系数
15. 斯皮尔曼相关系数
16. 肯德尔相关性系数
17. 布雷柯蒂斯距离
18. Haversine距离
19. Sørensen-Dice指数
20. 卡方检验
21. 信息熵
22. 交叉熵
23. 相对熵
24. 互信息
25. 兰氏距离

限于文章篇幅，文中涉及到的每种距离度量的算法实现，请点击下面链接查看：

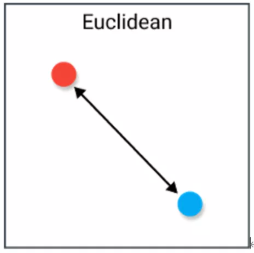
<https://github.com/jpegbert/MachineLearning/tree/master/similarity>



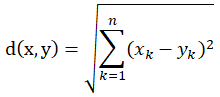
# 欧氏距离

## 原理

欧式距离欧氏距离是最常见、最常用的一种距离计算方式，也叫欧几里得距离、L2距离，可以解释为连接两点的线段长度。



设n维空间中两点a(x1,x2,…,xn)和b(y1,y2,…,yn)，则a和b的欧式距离表示为：



可以看到，欧几里得距离得到的结果是一个非负数，最大值是正无穷大，但是通常情况下相似度结果的取值范围在 [-1, 1] 之间。可以对它求倒数将结果转化到 (0, 1]之间：



分母+1是为了避免遇到被0整除的错误。

## 缺点

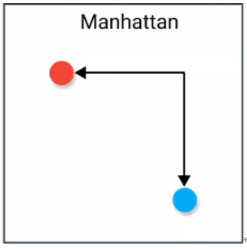
虽然这是一种常见的距离测量方法，但欧几里得距离并不是尺度不变的，这意味着计算出的距离可能会根据特征的单位而有所偏斜。通常情况下，在使用这种距离测量之前，需要对数据进行归一化。

此外，随着数据维度的增加，欧几里得距离的作用就越小。这与维度的诅咒有关，它涉及到高维空间的概念，并不像我们直观地期望的那样，从二维或三维空间中发挥作用。

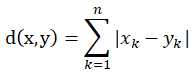
# 曼哈顿距离

## 原理

曼哈顿距离也称为城市街区距离、L1距离，顾名思义，假设在曼哈顿街区从P点到Q点，我们不能直接穿过高楼大厦走直线的距离，而是表示走过的街道的距离。在二维坐标上的表示，即两个点在标准坐标系上的绝对轴距之和。



设n维空间中两点a(x1,x2,…,xn)和b(y1,y2,…,yn)，则a与b的曼哈顿距离为：



## 缺点

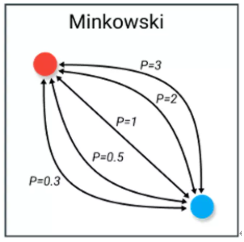
虽然曼哈顿距离对于高维数据似乎还不错，但它是一个比欧几里得距离更不直观的测量方法，尤其是在高维数据中使用时。

而且，它比欧几里得距离更容易给出一个更高的距离值，因为它不可能是最短路径。这不一定会带来问题，但你应该考虑到这一点。

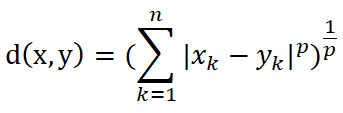
# 闵氏距离

## 原理

闵氏距离，全名闵可夫斯基距离，它不是一种距离，而是一组距离的定义，是对多个距离度量公式的概括性的表述。闵可夫斯基距离是欧氏空间中的一种测度，被看做是欧氏距离和曼哈顿距离的一种推广。



两个n维变量a(x1,x2,…,xn)与 b(y1,y2,…,yn)间的闵可夫斯基距离定义为：



p取1或2时的闵氏距离是最为常用的：

* p = 2即为欧氏距离。
* p = 1时则为曼哈顿距离。
* 当p取无穷时的极限情况下，可以得到切比雪夫距离。

P为不同值时，等距离组成的形状：

图片

注意，当 p<1 时，闵可夫斯基距离不再符合三角形法则，举个例子：当 p<1, (0,0) 到 (1,1) 的距离等于 (1+1)^{1/p}>2, 而(0,1)到这两个点的距离都是1。

闵可夫斯基距离比较直观，但是它与数据的分布无关，具有一定的局限性，如果x方向的幅值远远大于y方向的值，这个距离公式就会过度放大x维度的作用。所以，在计算距离之前，我们可能还需要对数据进行z-transform处理，即减去均值，除以标准差：



其中，μ为该维度上的均值，σ为该维度上的标准差。可以看到，上述处理开始体现数据的统计特性了。这种方法在假设数据各个维度不相关的情况下利用数据分布的特性计算出不同的距离。如果维度相互之间数据相关（例如：身高较高的信息很有可能会带来体重较重的信息，因为两者是有关联的），这时候就要用到马氏距离（Mahalanobis distance）了。

## 缺点

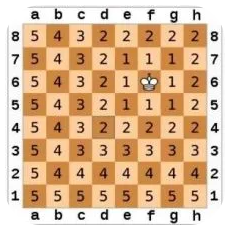
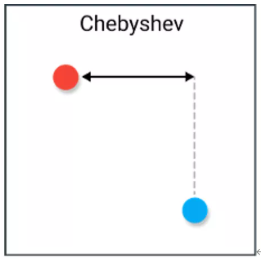
闵氏距离的缺点主要有两个：

* 将各个分量的量纲(scale)，也就是“单位”当作相同看待了
* 没有考虑各个分量的分布（期望，方差等)可能是不同的

# 切比雪夫距离

## 原理

切比雪夫距离，也叫度量，定义为沿任何坐标维度的两个向量之间的最大差异。换句话说，它是沿着一个轴线的最大距离。切比雪夫距离来源于国际象棋，在国际象棋中，国王可以直行、横行、斜行，所以国王走一步可以移动到相邻的8个方格中的任意一个，因此距离国王位置f6最近的8个位置的距离均为1。国王从格子(x1，y1)走到格子(x2，y2)最少需要多少步，你会发现这个距离是max(|x2-x1|,|y2-y1|)，这个距离就叫切比雪夫距离。

对于两个n维变量a(x1,x2,…,xn)与 b(y1,y2,…,yn)间的切比雪夫距离表示为：



## 缺点

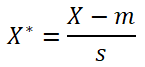
闵氏距离，包括曼哈顿距离、欧氏距离和切比雪夫距离都存在明显的缺点：

* 将各个分量的量纲(scale)，也就是“单位”当作相同的看待，量纲不同时，通常需要对数据做正规化；
* 没有考虑各个分量的分布（期望，方差等)可能是不同的；
* 各个维度必须是互相独立的，也就是“正交”的。

切比雪夫距离通常用于非常特殊的使用情况，这使得它很难像欧几里得距离或余弦相似性那样作为一个通用的距离度量。出于这个原因，我们建议只有当你绝对确定它适合你的使用情况时才使用它。

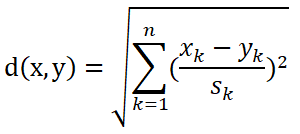
# 标准化欧氏距离

标准化欧氏距离是针对简单欧氏距离的缺点而作的一种改进方案。其实就是将各个分量都标准化，即先将各个分量都“标准化”到均值、方差相等。假设样本集X的均值(mean)为m，标准差(standard deviation)为s，那么X的“标准化变量”表示为：



即标准化后的值 = (标准化前的值－分量的均值) /分量的标准差

经过简单的推导就可以得到两个n维向量X(x1,x2,…,xn)和Y(y1,y2,…,yn)的标准化欧氏距离公式如下：



如果将方差的倒数看成是一个权重，也可称之为加权欧氏距离(Weighted Euclidean distance)。

# 马氏距离

马氏距离（Mahalanobis Distance）是一种有效的计算两个未知样本集的相似度的方法。与欧氏距离不同的是它考虑到各种特性之间的联系（例如：一条关于身高的信息会带来一条关于体重的信息，因为两者是有关联的）并且是尺度无关的（scale-invariant），即独立于测量尺度。修正了欧式距离中各个维度尺度不一致且相关的问题。

一些基本概念：

* 方差：方差是标准差的平方，而标准差的意义是数据集中各个点到均值点距离的平均值。反应的是数据的离散程度。
* 协方差：标准差与方差是描述一维数据，当存在多维数据时，我们通常需要知道每个维数的变量中间是否存在关联。协方差就是衡量多维数据集中，变量之间相关性的统计量。如果两个变量之间的协方差为正值，则这两个变量之间存在正相关，若为负值，则为负相关。

闵氏距离比较直观，但是它与数据的分布无关，具有一定的局限性，标准化欧氏距离涉及到数据分布，但没有考虑到数据的相关性，而马氏距离在计算两个样本之间的距离时，考虑到了样本所在分布造成的影响，主要是因为：

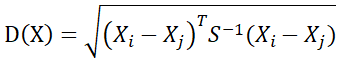
1）不同维度的方差不同，进而不同维度对距离的重要性不同。

2）不同维度可能存在相关性，影响距离的度量。

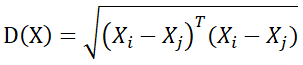
定义：假设有M个样本向量X1,X2,…,Xm，协方差矩阵记为S，均值记为向量u，则其中样本向量X到u的马氏距离表示为：



马氏距离也可以定义为两个服从同一分布并且其协方差矩阵为S的随机变量Xi与Xj的差异程度（马氏距离）：



若协方差矩阵S是单位矩阵（各个样本向量之间独立同分布），则公式就成了：



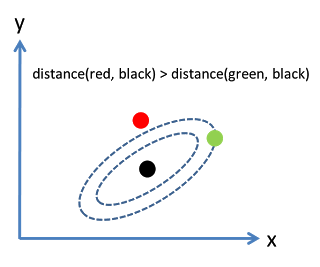
即：欧氏距离

若协方差矩阵是对角矩阵，公式变成了标准化欧氏距离。

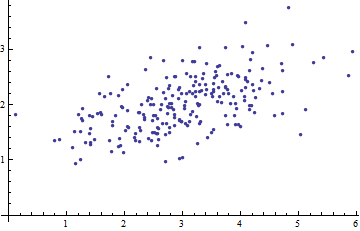
根据马氏距离的定义，可以得到它的几个特点如下：

* 两点之间的马氏距离与原始数据的测量单位无关（不受量纲的影响）
* 标准化数据和中心化数据（即原始数据与均值之差）计算出的二点之间的马氏距离相同
* 可以排除变量之间的相关性的干扰
* 满足距离的四个基本公理：非负性、自反性、对称性和三角不等式
* 缺点是夸大了变化微小的变量的作用

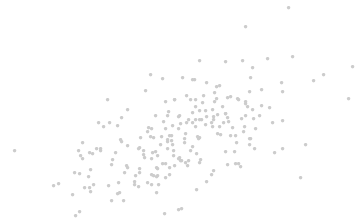
考虑下面这张图，椭圆表示等高线，从欧几里得的距离来算，绿黑距离大于红黑距离，但是从马氏距离，结果恰好相反：



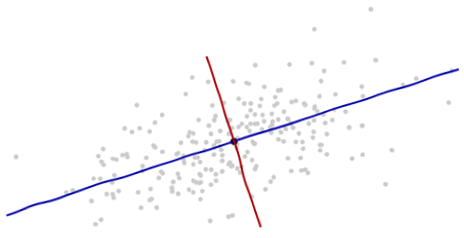
马氏距离实际上是利用Cholesky transformation来消除不同维度之间的相关性和尺度不同的性质。下图是一个二元变量数据的散点图：



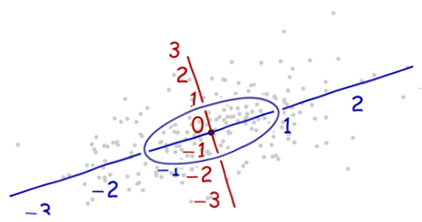
当我们将坐标轴拿掉，如下图：



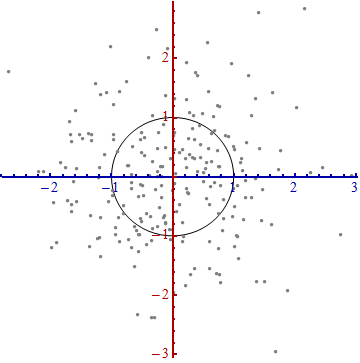
根据数据本身的提示信息来引入新的坐标轴：坐标的原点在这些点的中央（根据点的平均值算得）。第一个坐标轴（下图中蓝色的线）沿着数据点的“脊椎”，并向两端延伸，定义为使得数据方差最大的方向。第二个坐标轴（下图红色的线）会与第一个坐标轴垂直并向两端延伸。如果数据的维度超过了两维，那就选择使得数据方差是第二个最大的方向，以此类推。



我们需要一个比例尺度。沿着每一个坐标轴的标准差来定义一个单位长度。使用“68-95-99.7法则”更容易找到合理的单位。（大约68%的点需要在离原点一个单位长度的范围内；大约95%的点需要在离原点两个单位的长度范围内；99.7的点需要在3个单位程度范围内。）为了以示参考，如下图：



由于每个轴上的单位长度不相等，所以上图中距离原点一个单位的形成的轨迹并不是一个圆形。为了更好的呈现图表，我们将图片进行旋转。同时，并让每个轴方向上的单位长度相同：



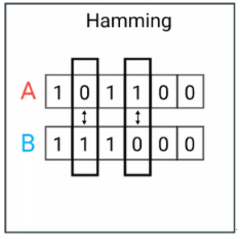
上面就是从散点图中构建坐标系统的过程，为的是方便进行测量。说明：

* 沿着新坐标轴的单位向量是协方差矩阵的特征向量。注意到没有变形的椭圆，变成圆形后沿着特征向量用标准差（协方差的平方根）将距离长度分割。
* 坐标轴扩展的量是协方差矩阵的逆的特征值（平方根），同理的，坐标轴缩小的量是协方差矩阵的特征值。所以，点越分散，需要的将椭圆转成圆的缩小量就越多。
* 尽管上述的操作可以用到任何数据上，但是对于多元正态分布的数据表现更好。在其他情况下，点的平均值或许不能很好的表示数据的中心，或者数据的“脊椎”（数据的大致趋势方向）不能用变量作为概率分布测度来准确的确定。
* 原始坐标系的平移、旋转，以及坐标轴的伸缩一起形成了仿射变换（affine transformation）。除了最开始的平移之外，其余的变换都是基底变换，从原始的一个变为新的一个。
* 在新的坐标系中，多元正态分布像是标准正太分布，当将变量投影到任何一条穿过原点的坐标轴上。特别是，在每一个新的坐标轴上，它就是标准正态分布。从这点出发来看，多元正态分布彼此之实质性的差异就在于它们的维度。

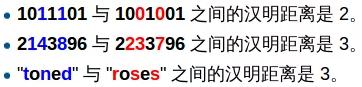
# 汉明距离

## 原理

汉明距离是两个等长字符串对应位置的不同字符的个数。汉明距离有一个最为鲜明的特点就是它比较的两个字符串必须等长，否则距离不成立。它也可以用来比较字符串之间的相似度，计算彼此不同的字符数。



例如：



还可以用简单的匹配系数来表示两点之间的相似度，即：匹配字符数/总字符数。

## 缺点

当两个向量的长度不相等时，汉明距离很难使用。汉明距离只能用于将相同长度的向量相互比较，以了解哪些位置不匹配。

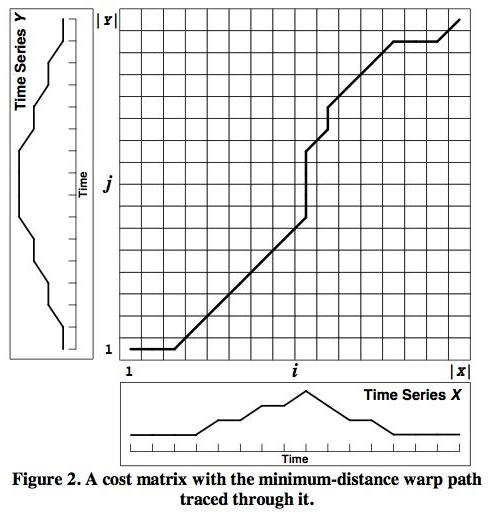
# 编辑距离

汉明距离可以度量两个相同长度的字符串之间的相似度，如果比较的两个字符串长度不同，则不仅要进行替换，还要进行插入与删除的操作，这种情况下下，通常使用更加复杂的编辑距离进行相似度判断，即通过字符编辑（插入、删除或替换），将字符串A转换成字符串B所需要的最少操作数。

# DTW距离

DTW（Dynamic Time Warping，动态时间归整）是一种衡量两个长度不等的时间序列间相似度的方法，主要应用在语音识别领域，识别两段语音是否表示同一个单词。

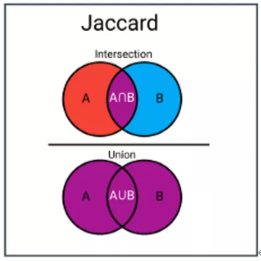
大部分情况下，需要比较的两个序列在整体上具有非常相似的形状，但是这些形状并不是对一一对应的。所以我们在比较他们的相似度之前，需要将其中一个（或者两个）序列在时间轴下warping扭曲，以达到更好的对齐。而DTW就是实现这种warping扭曲的一种有效方法。DTW通过把时间序列进行延伸和缩短，来计算两个时间序列性之间的相似性。



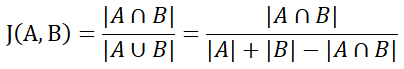
# 杰卡德相似系数

## 原理

Jaccard指数（或称交集比联合）是一种用于计算样本集相似性和差异性的度量。它是交集的大小除以样本集的联合大小。即两个集合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。杰卡德系数值越大，样本的相似度越高。



用符号J(A,B)表示：



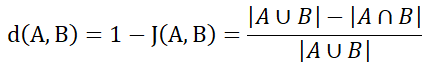
在实践中，它是集合之间相似实体的总数除以实体的总数。例如，如果两个集合有1个共同的实体，而总共有5个不同的实体，那么Jaccard指数将是1/5=0.2。

## 缺点

Jaccard指数（杰卡德相似系数）的一个主要缺点是，它受数据大小的影响很大。大的数据集会对指数产生很大的影响，因为它可以在保持相似的交叉点的同时显著增加联合。

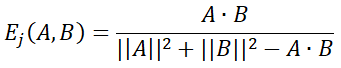
# 杰卡德距离(区分度)

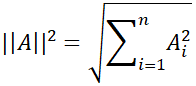
杰拉德距离(Jaccard distance)是用来衡量两个集合差异性的一种指标，它是杰卡德相似系数的补集，被定义为1减去Jaccard相似系数。适用于集合相似性度量，字符串相似性度量。杰卡德距离可用如下公式表示：



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。

# Tanimoto系数（广义Jaccard相似系数）

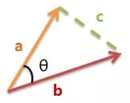


其中A、B分别表示为两个向量，集合中每个元素表示为向量中的一个维度，在每个维度上，取值通常是[0, 1]之间的值（如果取值是二值向量0或1，那么Tanimoto系数就等同Jaccard距离），表示向量乘积，表示向量的模。

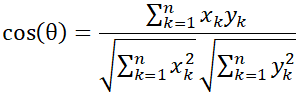
# 余弦距离

## 原理

余弦距离，也称为余弦相似度，是用向量空间中两个向量夹角的余弦值作为衡量两个个体间差异的大小的度量方法。余弦值越接近1，就表明夹角越接近0度，也就表示两个向量越相似。



余弦距离计算公式：



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个向量的夹角越大，当两个向量的方向重合时夹角取值最大值1，方向相反时，夹角余弦取最小值-1。

根据欧氏距离和余弦相似度各自的计算方式和衡量特征，分别适用于不同的数据分析模型：欧氏距离能够体现个体数值特征的绝对差异，所以更多的用于需要从维度的数值大小中体现差异的分析，如使用用户行为指标分析用户价值的相似度或差异；而余弦相似度更多的是从方向上区分差异，而对绝对的数值不敏感，更多的用于使用用户对内容评分来区分用户兴趣的相似度和差异，同时修正了用户间可能存在的度量标准不统一的问题（因为余弦相似度对绝对数值不敏感）。

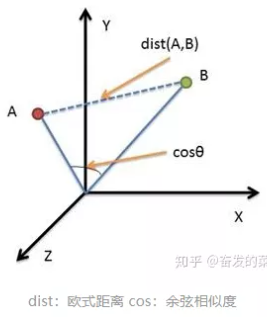
简单说，就是余弦相似性经常被用来抵消欧几里得距离的高维度问题。余弦相似性只是两个向量之间角度的余弦。如果将它们归一化为都有长度为1的向量，它的内积也相同。

## 缺点

余弦相似性的一个主要缺点是不考虑向量的大小，只考虑其方向。在实际应用中，这意味着值的差异没有被完全考虑。以推荐系统为例，那么余弦相似性并没有考虑到不同用户之间的评分等级差异。

比如用户对内容评分，5分制。A和B两个用户对两个商品的评分分别为A：(1,2)和B：(4,5)。我们分别用两种方法计算相似度。使用余弦相似度得出的结果是0.98，看起来两者极为相似，但从评分上看X似乎不喜欢这两个东西，而Y比较喜欢。造成这个现象的原因就在于，余弦相似度没法衡量每个维数值的差异，对数值的不敏感导致了结果的误差。

## 欧式距离vs余弦距离

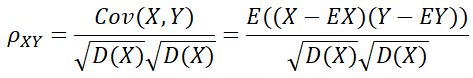


欧氏距离衡量的是空间各点的绝对距离，跟各个点所在的位置坐标直接相关；而余弦距离衡量的是空间向量的夹角，更加体现在方向上的差异，而不是位置。如果保持A点位置不变，B点朝原方向远离坐标轴原点，那么这个时候余弦距离是保持不变的，而A、B两点的距离显然在发生改变，这就是欧氏距离和余弦距离之间的不同之处。

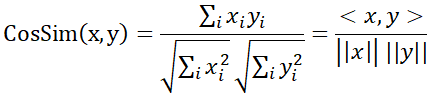
# 皮尔逊相关系数与相似距离

## 皮尔逊相关系数

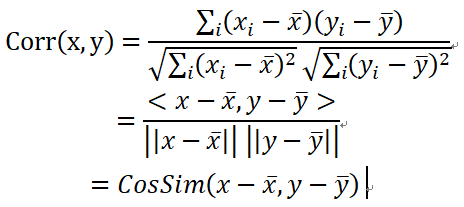
皮尔逊相关系数（Pearson correlation），也叫相关系数，相比于余弦相似度只与向量方向有关，受向量的平移影响，皮尔逊相关系数具有平移不变性和尺度不变性，在夹角余弦公式中，如果将 x 平移到 x+1, 余弦值就会改变。皮尔逊相关系数值在-1.0到1.0之间，接近0的表示无相关性，接近1或者-1被称为具有强相关性，负数表示负相关，不过皮尔逊相关系数只能衡量两个随机变量间的线性相关性。



如果将夹角余弦公式写成：



则皮尔逊相关系数则可表示为：



因此，皮尔逊相关系数可以看作中心化后变量的余弦距离。

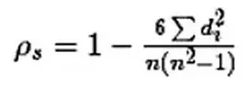
皮尔逊相关系数具有平移不变性和尺度不变性，计算出了两个向量（维度）的相关性。在各个领域都应用广泛，例如，在推荐系统根据为某一用户查找喜好相似的用户，进而提供推荐，优点是可以不受每个用户评分标准不同和观看影片数量不一样的影响。

## 相似距离



# 斯皮尔曼相关系数

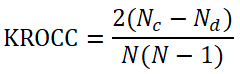
斯皮尔曼相关系数又称斯皮尔曼秩相关系数，是利用两变量的秩次大小作线性相关分析，是根据原始数据的排序位置进行求解，对原始变量的分布不作要求，属于非参数统计方法，适用范围要广些。



计算过程：先对两个变量（X, Y）的数据进行排序，然后记下排序以后的位置（X’, Y’），（X’, Y’）的值就称为秩次，秩次的差值就是上面公式中的di，n是变量中数据的个数。

# 肯德尔相关性系数

肯德尔相关性系数又称肯德尔秩相关系数，它所计算的对象是分类变量，因此它需要的数据集必须是分类变量。

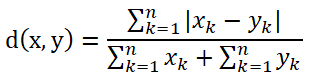


其中，Nc表示主客观评价值中一致的值的个数，Nd则表示了主观评估值和客观评估值不一样的个数。

适用案例：评委对选手的评分（优、中、差等），我们想看两个（或者多个）评委对几位选手的评价标准是否一致；或者医院的尿糖化验报告，想检验各个医院对尿糖的化验结果是否一致。

# 布雷柯蒂斯距离

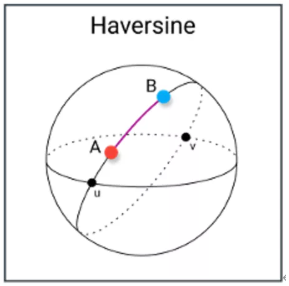
布雷柯蒂斯距离（Bray Curtis Distance）主要用于生态学和环境科学，计算坐标之间的距离。该距离取值在[0,1]之间。它也可以用来计算样本之间的差异。



# Haversine距离

## 原理

Haversine距离是指球面上两点之间的经度和纬度距离。



它与欧几里得距离非常相似，因为它计算的是两点之间的最短线。主要的区别是不可能有直线，因为这里的假设是两点在一个球体上。两点间的Haversine距离公式为：

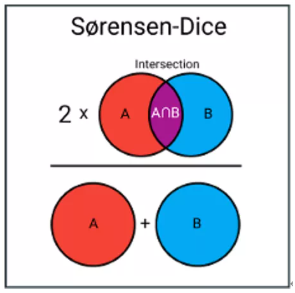


## 缺点

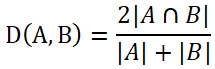
这种距离测量方法的一个缺点是，它假定各点位于一个球体上。在实践中，这种情况很少发生，例如，地球并不是完全的圆形，这可能会使计算在某些情况下变得困难。相反，如果能采用Vincenty距离，则会很有趣，因为它假设的是一个椭圆体。

# Sørensen-Dice指数

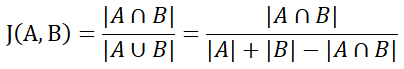
Sørensen-Dice指数与Jaccard指数（杰拉德相似系数）非常相似，因为它衡量样本集的相似性和多样性，因为可以把字符串理解为一种集合，因此Dice距离也会用于度量字符串的相似性。此外，Dice系数的一个非常著名的使用即实验性能评测的F1值。



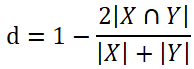
虽然它们的计算方法相似，但Sørensen-Dice指数更直观一些，因为它可以被看作是两组之间的重叠百分比，这个数值在0和1之间。Sørensen–Dice指数公式为：



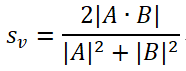
其中分子是A与B的交集数量的两倍，分母为X和Y的长度之和，所以他的范围也在0到1之间。从公式看，Dice系数和Jaccard非常的类似。Jaccard是在分子和分母上都减去了|A∩B|。



与Jaccard不同的是，相应的差异函数



与Jaccard类似, 集合操作可以用两个向量A和B的操作来表示：



## 缺点

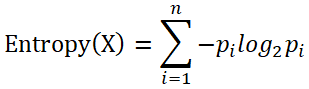
与Jaccard指数一样，它们都高估了集合的重要性，只有很少或没有TP（Truth Positive）值的正集合。因此，它可以求得多盘的平均分数。它将每个项目与相关集合的大小成反比加权，而不是平等对待它们。

# 卡方检验

卡方检验应用统计样本的实际观测值与理论推断值之间的偏离程度，常用来检验某一种观测分布是不是符合某一类典型的理论分布（如二项分布，正态分布等)。如果卡方值越大，二者偏差程度越大；反之，二者偏差越小；若两个值完全相等时，卡方值就为0，表明理论值完全符合。

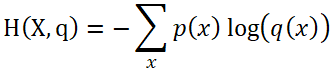
# 信息熵

信息熵是衡量分布的混乱程度或分散程度的一种度量。分布越分散(或者说分布越平均)，信息熵就越大。分布越有序（或者说分布越集中），信息熵就越小。



# 交叉熵

如果一个随机变量X 服从p(x)分布，q(x)用于近似p(x)的概率分布，那么随机变量和模型q之间的交叉熵定义为：

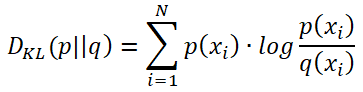


交叉熵在CNN分类中经常用到，用来作为预测值和真实标签值的距离度量。经过卷积操作后，最后一层出来的特征经过softmax函数后会变成一个概率向量，我们可以看作为是概率分布q，而真实标签我们可以看作是概率分布p，因此真实分布p和预测分布q的交叉熵就是我们要求的loss损失值。

# 相对熵

相对熵又称KL散度（Kullback–Leibler divergence，简称KLD），信息散度（information divergence），信息增益（information gain），相对熵是交叉熵与信息熵的差值。

相对熵是衡量两个分布(P、Q)之间的距离；越小越相似。

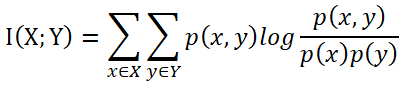


表示的就是概率q与概率p之间的差异，很显然，散度越小，说明概率q与概率p之间越接近，那么估计的概率分布于真实的概率分布也就越接近。

# 互信息(Mutual Information)

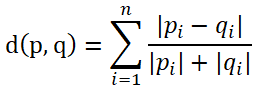
互信息(Mutual Information)是信息论里一种有用的信息度量，它可以看成是一个随机变量中包含的关于另一个随机变量的信息量，或者说是一个随机变量由于已知另一个随机变量而减少的不肯定性。衡量随机变量之间相互依赖程度的度量。

设两个随机变量（X,Y)的联合分布为p(x,y)，边缘分布分别为p(x),p(y)，互信息I(X;Y)是联合分布p(x,y)与边缘分布p(x)p(y)的相对熵，即：



# 兰氏距离

兰氏距离(Lance and Williams distance)堪培拉距离（Canberra Distance），被认为是曼哈顿距离的加权版本。其定义公式为：



通常兰氏距离对于接近于0（大于等于0）的值的变化非常敏感。与马氏距离一样，兰氏距离对数据的量纲不敏感。不过兰氏距离假定变量之间相互独立，没有考虑变量之间的相关性。

参考

<https://mp.weixin.qq.com/s/YPwo5JFlIsvkOi6QJGvRqg>

<https://mp.weixin.qq.com/s/9fioWus3UH2gvEtDLZjOQQ>

<https://mp.weixin.qq.com/s/FVIR3-ygHAQp2wzRRSx4Lw>

<https://mp.weixin.qq.com/s/3-_TzRvXmqdRov1W5_-F7A>

<https://mp.weixin.qq.com/s/RsIVVqoEHTlz6dxzRYba0Q>

<https://mp.weixin.qq.com/s/_s7sOlWeLpHPxbqQk-EFeg>