

KL散度

2015年7月27日 星期一

下午10:33

一，简介

KL-散度（KL-divergence，KL-距离，KLD，Kullback–Leibler divergence，相对熵，relative entropy，信息散度（information divergence）），衡量两个概率分布的距离/差异的量。

二，详述

1，定义（百度百科）

KL散度是两个概率分布P和Q差别的非对称性的度量。

KL散度是用来度量使用基于Q的编码来编码来自P的样本平均所需的额外的比特个数。

典型情况下，P表示数据的真实分布，Q表示数据的理论分布，模型分布，或P的近似分布。

2，公式

$$D(P||Q) = \sum P(i) \log(P(i)/Q(i))$$

（对于离散的情况）

而对于连续的变量，取积分即可

3，版本

博客里说，适合做个算术平均，使得消除方向性，百科里说KLD就是带方向的。

三，例子

见博客里

举一个实际的例子吧：比如有四个类别，一个方法A得到四个类别的概率分别是0.1,0.2,0.3,0.4。另一种方法B（或者说是事实情况）是得到四个类别的概率分别是0.4,0.3,0.2,0.1,那么这两个分布的KL-Distance(A,B)=0.1*log(0.1/0.4)+0.2*log(0.2/0.3)+0.3*log(0.3/0.2)+0.4*log(0.4/0.1)

这个里面有正的，有负的，可以证明KL-Distance() \geq 0.从上面可以看出，KL散度是不对称的。即KL-Distance(A,B) \neq KL-Distance(B,A)

四，参考资料

<http://baike.baidu.com/view/951299.htm?fr=aladdin> 百科 不错

<http://blog.csdn.net/caohao2008/article/details/6910794> 简介（质量一般）

http://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler_divergence wiki

<http://zh.wikipedia.org/zh-tw/%E7%9B%B8%E5%AF%B9%E7%86%B5> 中文 wiki

数学之美中也有专门的介绍