Record for 1.20-1.26

目标：

1. Use Helinger Distance to repesent the result.

2. 尝试通过EM和GMM去表示各种distributions，然后用cFF和eFF的方法去计算approximate add的结果

对于两个离散的概率分布 P=(p1,p2,…,pk), Q=(q1,q2,…,qk), Hellinger distance is defined as

H(P,Q)=sqrt(sum\_{i=1}^k (sqrt(pi)-sqrt(qi))^2)/sqrt(2)

My method 对于部分分布得到的结果和真实结果的Hellinger distance小于0.05，对于部分情况Hellinger distance偏大，明显不适用

|  |  |  |
| --- | --- | --- |
| A[0] | B[0] | HD |
| 0.5 | 0.5 | 0 |
| 0.4 | 0.4 | 0.0543666 |
| 0.4 | 0.6 | 0.02540254 |
| 0.3 | 0.7 | 0.10321432 |
| 0.3 | 0.5 | 0.05370310 |
| 0.7 | 0.7 | 0.18182379 |
| 0.5 | 0.7 | 0.08863494 |

每组都是测试50次取得平均值

对于全部随机的情况，得到的平均HD是0.248490610

该方法只能适用于某些分布

对于第一位在0.3到0.7之间的随机情况，得到的平均HD是0.05151329

以上HD都是针对cPMF

对于aPMF 0.051347575

0.4~0.6 HD\_aPMF=0.0174841

0.2~0.8 HD\_aPMF=0.1031565

对于16位相加的时候，在适用的情况下得到的HD是0.06

对于所有情况下得到的HD是0.36

EM Algorithm

用EM去求GMM时，高斯分布也是未知的？

所求的高斯分布的参数， theta={(omega\_j,miu\_j,sum\_j)}\_{j=1}^k

https://www.cnblogs.com/sylvanas2012/p/5053798.html