对MedLDA模型的改进：

======================================

每个单词，对应参数，是二维向量，它的第一个元素为，第二个元素为

由生成二项分布

==============暂时不用================

假设单词的状态为(其值为0或为1)。为0表示该单词不被选择作为分类的特征；若为1，表示该单词被选择作为分类的特征。并假设服从分布，表示为0或1的potential，为0时是，为1是。





用表示文档可能被赋予的主题。表示文档中单词可能被赋予的状态。

假设

(注意，是向量,而是标量)

 (1)



(2)

从公式(1)可以看出，与(即)是相互独立的。



继续公式(2)







问题变为：





的变为：



 (10)

E-step，优化即通过调整，来达到优化的目的，因为，决定了。因为中的不等式限制不包含,所以它的更新方法和Unsupervised LDA中的一样。



对求导，设置导数为0,得：



M-step,优化参数，其中的优化方法基本没变。看一下参数的优化过程。









所以

采用梯度上升的方法进行优化