1. 层次聚类HC成功运用在那些问题上面？缺点是什么？

HC 成功运用在temporal expression patterns（时间表达模式）上[1]，比如预测淋巴瘤患者的病情状况[2]，提供乳腺肿瘤的分子肖像图[3]。

缺点在于，给数据强加了一个树形结构（基因之间的关系可能非常复杂），对于评价两个基因是否属于同一类的评价标准非常敏感，通常需要主观的去定义聚类。

1. 自组织映射SOM成功运用到了那些上？缺点是什么？

SOM成功运用于确定在造血细胞分化过程中参与的pathways,识别白血病的亚类型

缺点在于，结果非常的不稳定，不同的初始条件，能得到不同的结果。

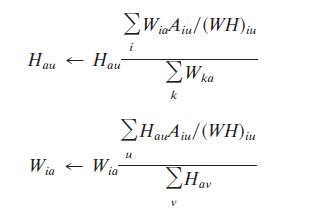
1. 文中非负矩阵分解NMF的提出，是否能解决上面的问题？

文中所提出的的NMF方法，在随机的W × H 上得到稳定的结果，而且可以人为选择聚类的数目，一定程度上可以解释生物学意义。

1. NMF的具体实现方法？







使迭代，D达到比较小的数值即可

1. Model Selection?

M × M 的连接矩阵C，如果i,j 属于同一聚类之中，则=1，如果i,j 不属于同一聚类之中，则=0。

* ：一致性矩阵，许多连接举证的平均值。采用非对角形式

：表示矩阵的分散程度，数值最大为1，数值越大，结果越好。

1. 将NMF运用在急性白血病上边





NMF与HC对比，其稳定性更高；与SOM对比，其分类效果更好。