**最大期望算法**

**（Expectation-Maximization Algorithm, EM）**

# 凸函数

设是定义域为实数的函数，如果对于所有的实数，，那么称是凸函数。当是向量时，如果其Hessian矩阵H是半正定的，那么也称为凸函数。如果或者，那么称是严格凸函数。

# Jensen不等式

如果是凸函数，是随机变量，那么



特别地，如果是严格凸函数，

那么当且仅当。

# EM算法推导

联合概率密度下某个变量的边缘概率密度函数：注意这里也是随机变量。对每一个样本的所有可能类别求等式右边的联合概率密度函数和，也就得到等式左边为随机变量的边缘概率密度。



因为是随机变量的分布列（概率密度函数），，

其中；

是的期望

函数是凹函数，

其中：









## EM算法（Expectation-Maximization）

输入：观察数据，联合分布，条件分布。

初始化分布参数；

重复以下步骤直到收敛：

E步骤：根据参数初始值或上一次迭代的模型参数来计算出隐性变量的后验概率，其实就是隐性变量的期望。作为隐藏变量的现估计值：



M步骤：将似然函数最大化以获得新的参数值：





