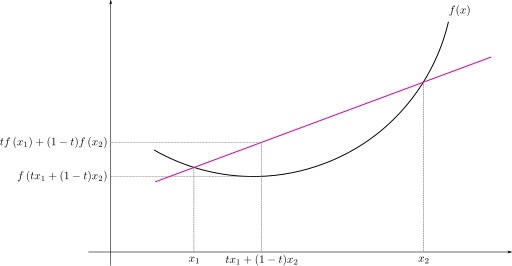
**Jensen不等式**

两点坐标形式（凹函数）



长度

长度，

坐标

直角三角形

长度，

坐标

可以把两点形式拓展到多点，即在定义域上取个数，是它们的权重()则有

取等号的条件是所取的个数相等。因为若个数相等假设为，则左侧就是加权的，就是。右侧是里面的加权，也是。

期望形式（凹函数）

这幅图描述了随机变量的分布以及凹映射函数，通过映射可以得到的分布。（就是一个自变量的变换）

凹映射函数使的分布与比较起来变得“先急后缓”，后面部分（最高点以后）的概率变大，期望也就应该出现在后面部分，则在图像上就表现为

这是关于随机变量的形式，我们可以延伸一下，得到以下关于随机变量的积分形式

其中，，。上面形式其实就是这个形式中的特例。

取等号条件由上面可知，所取的个相等，也就是不论如何改变，都是中的一个常数。

**EM算法**

使用极大似然估计法估计一个含有隐变量的模型参数。

似然函数为：

设观测数据为

这里遇到了极大似然估计法中未曾出现过的问题，首先，是未知数据，其次，式子里出现了的形式，这是不好求到的。EM算法运用迭代的思想，我们虽然不能直接求出原似然函数的，但是针对每个确定的点，都能画出原似然函数的下界函数，通过求这个下界函数的argmax，可以保证新的对应的似然变大。

为了可以套用Jensen不等式，我们可以同时乘以和除以一个关于的分布。于是可以将这个式子对应Jensen不等式，其中，对应了，对应了，对应了，对求和对应了对的积分。

是一个凸函数，所以上面的不等号反向，得到：

我们感兴趣的是等号条件，因为取等号的时候，可以得到一个似然函数的紧的下界函数。（后面讨论）

根据Jensen不等式的等号条件，有

设，由于是一个分布，有

这就是令等式成立的的选取。上面说了EM算法是迭代更新的，那么对于每个确定的点，假设叫做，我们可以选取

然后将

看做一个函数，代入就是

这个函数保证了在点处，似然与原函数的似然相等，这样，任何一个新的能使此函数似然增大的新的点，也能使原似然增大。而一个非紧的下界函数是没有这种性质的。

这个下界函数还可以进一步简化：

后面部分是不随的改变而改变的，所以我们只要求这一项的即可：

我们可以把它写作期望的形式：

解释为：确定了上一步的参数后，对于每种可能的隐变量值，求得其似然然后乘以概率后求和（求其似然的期望）。