EM算法

吴锐 李林 李震

EM算法是针对概率模型中含有隐变量（latent variable）时的极大似然估计法。

直接求似然的极大值存在两点困难：

1. 存在有隐变量

2. 包含和（或积分）的对数

EM算法则是一种近似计算含有隐变量概率模型的极大似然估计的方法

从两个方向来阐述EM算法的可行性：

1. EM算法通过迭代逐步近似极大化观测数据关于参数的似然，每一步中均会提升似然函数的值
2. 该迭代具有收敛性

有关EM算法的一些问题讨论：

1. 为什么会有因变量的存在？

隐变量应该是问题驱动的。一方面来说，当前问题中明显地含有隐变量（例如三硬币模型）；另一方面，需要加入隐变量才能够对数据的分布做出合适的假设（例如单个高斯无法描述的数据分布）。

1. 为什么会用迭代的方法来计算？

因为在对观测变量的似然函数进行极大似然估计时，由于隐变量的存在无法直接求得解析解，所以EM考虑通过迭代的方式，不断极大化该似然函数的下界，最终达到收敛。

1. 如何保证EM算法的收敛性？

首先每一步迭代中保证观测数据的似然函数的递增，其次该似然必然是有界的。

有关EM算法的详细过程，如下述笔记所示（编辑公式较为麻烦，所以直接贴图了）：

E 、 1 2 ) = p , 9 
一 P 9 ) 
= ㄥ k ㄥ ( llp) 
q 。 丶 p ( ㄨ 9 ' l) E D 值 p ㄅ 衵 等 

0 袰 す 与 由 ー 些 目 
助 哽 数 に 窺 。 z , セ 褪 通 辻 い , 
) 、 气 ) マ 十 ム ? f グ 司 入 9 ? 
9 [ 2 湜 系 于 角 褫 籌 布 
甲 2 ) 引 長 '753 p ( x 弩 乞 Bo 頃 形 式 
う 、 う 広 見 Bo に 下 っ 0 魯 磔 幗 5 狂 ム ? 
み / p 図 田 。 乙 乙 ダ ( 7 め / 囘 凶 9 ) ) 
反 刻 9 ( 2 ) = p 保 / x. P ⅸ = 釦 
念 , , P ( 刈 の , 乙 0 
賀 肓 竹 松 。 中 , 之 肩 % 旧 ) の = ・ 。 仡 つ 、 
由 等 岩 邵 存 ゴ , ィ 日 ト ・ 
P ( 'X 岶 “ り 彡 p に ⅸ 9 ん 
乙 歩 新 由 易 
み 下 一 弋 , は 以 調 ely “ ) 
一 ゝ 時 ル 延 恥 ト 覊 , 毛 褓 P(x19 ) 臼 
。 角 新 イ 9 臨 
午 、 し ル 从 」 イ + 広 ? 
E* 气 = p 、 り , 硬 乙 0 う 卩 刈 0 。 、 ) ) 塒 
衛 量 3 汚 p ん 9 。 ) 倒 ル セ 乙 与 p 勾 跖 由 及 
mb 、 イ 吏 。 。 大 化 。 , 々 ツ p 図 9 地 廴 