

# 最大期望算法

维基百科，自由的百科全书

**最大期望演算法**（**Expectation-maximization algorithm**，又譯**期望最大化算法**）在统计中被用于寻找，依赖于不可观察的隐性变量的概率模型中，参数的最大似然估计。

在统计计算中，**最大期望（EM）算法**是在**概率模型**中寻找**参数最大似然估计**或者**最大后验估计**的算法，其中概率模型依赖于无法观测的隐性变量。最大期望算法经常用在机器学习和计算机视觉的数据聚类（Data Clustering）领域。最大期望算法经过两个步骤交替进行计算，第一步是计算期望（**E**），利用对隐藏变量的现有估计值，计算其最大似然估计值；第二步是最大化（**M**），最大化在E步上求得的**最大似然值**来计算参数的值。M步上找到的参数估计值被用于下一个E步计算中，这个过程不断交替进行。

## 目录

历史

EM简单教程

最大期望过程说明

估计无法观测的数据

参见

参考文献

## 历史

最大期望值算法由Arthur P. Dempster，Nan Laird和Donald Rubin在他们1977年发表的经典论文中提出。他们指出此方法之前其实已经被很多作者在他们特定的研究领域多次提出过。

## EM简单教程

EM是一个在已知部分相关变量的情况下，估计未知变量的迭代技术。EM的算法流程如下：

1. 初始化分布参数
2. 重复直到收敛：
  1. E步骤：估计未知参数的期望值，给出当前的参数估计。
  2. M步骤：重新估计分布参数，以使得数据的似然性最大，给出未知变量的期望估计。应用于缺失值。

## 最大期望过程说明

我们用**y**表示能够观察到的不完整的变量值，用**x**表示无法观察到的变量值，这样**x**和**y**一起组成了完整的数据。**x**可能是实际测量丢失的数据，也可能是能够简化问题的隐藏变量，如果它的值能够知道的话。例如，在**混合模型**（**Mixture Model**）中，如果“产生”样本的混合元素成分已知的话最大似然公式将变得更加便利（参见下面的例子）。

### 估计无法观测的数据

让 $p$  代表矢量 $\theta$ :  $p(\mathbf{y}, \mathbf{x}|\theta)$  定义的参数的全部数据的概率分布（连续情况下）或者概率聚类函数（离散情况下），那么从这个函数就可以得到全部数据的最大似然值，另外，在给定的观察到的数据条件下未知数据的条件分布可以表示为：

$$p(\mathbf{x}|\mathbf{y}, \theta) = \frac{p(\mathbf{y}, \mathbf{x}|\theta)}{p(\mathbf{y}|\theta)} = \frac{p(\mathbf{y}|\mathbf{x}, \theta)p(\mathbf{x}|\theta)}{\int p(\mathbf{y}|\mathbf{x}, \theta)p(\mathbf{x}|\theta)d\mathbf{x}}$$

## 参见

---

- 估计理论
- 数据聚类

## 参考文献

---

- Arthur Dempster, Nan Laird, and Donald Rubin. "Maximum likelihood from incomplete data via the EM algorithm". *Journal of the Royal Statistical Society Series B*, 39 (1):1–38, 1977[1].
  - Robert Hogg, Joseph McKean and Allen Craig *Introduction to Mathematical Statistics* pp. 359-364. Upper Saddle River, NJ: Pearson Prentice Hall, 2005.
  - Radford Neal, Geoffrey Hinton "A view of the EM algorithm that justifies incremental, sparse, and other variants". In Michael I. Jordan (editor), *Learning in Graphical Models* pp 355-368. Cambridge, MA: MIT Press, 1999.
  - The on-line textbook: Information TheoryInference, and Learning Algorithms, by David J.C. MacKay includes simple examples of the E-M algorithm such as clustering using the soft K-means algorithm, and emphasizes the variational view of the E-M algorithm.
  - A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models , by J. Bilmes includes a simplified derivation of the EM equations for Gaussian Mixtures and Gaussian Mixture Hidden Markov Models.
  - Information Geometry of the EM and em Algorithms for Neural Networks by Shun-Ichi Amari give a view of EM algorithm from geometry view point.
- 

取自“<https://zh.wikipedia.org/w/index.php?title=最大期望算法&oldid=47677504>”

---

本页面最后修订于2018年1月5日 (星期五) 12:54。

本站的全部文字在知识共享 署名-相同方式共享 3.0协议之条款下提供，附加条款亦可能应用。（请参阅使用条款）  
Wikipedia®和维基百科标志是维基媒体基金会的注册商标；维基™是维基媒体基金会的商标。  
维基媒体基金会是在美国佛罗里达州登记的501(c)(3)免税、非营利、慈善机构。