



将这几个分布带入目标函数：

$$\begin{aligned}
 &= \log \prod_{i=1}^N \sqrt{\frac{\lambda}{2\pi}} \cdot e^{-\frac{\lambda(x_i - \mu)^2}{2}} + \log \sqrt{\frac{\kappa_0 \lambda}{2\pi}} \cdot e^{-\frac{\kappa_0 \lambda (\mu - \mu_0)^2}{2}} + \log \frac{\beta^{a_0} \lambda^{a_0-1} e^{-b_0 \lambda}}{\Gamma(a_0)} \\
 &= \frac{N}{2} \log \lambda - \frac{\lambda}{2} \sum_{i=1}^N (x_i - \mu)^2 + \frac{1}{2} \log(\kappa_0 \lambda) - \frac{\kappa_0 \lambda}{2} (\mu - \mu_0)^2 \\
 &\quad + (a_0 - 1) \log \lambda - b_0 \lambda + \text{const}
 \end{aligned}$$

既然得到上边的结果，接下来我们就要进行参数估计了，根据变分思想，我们估计一个参数，是试用其他参数来进行求期望得到的，那么就更新 $\mu$ ：

$$\log \tilde{p}(\mu, \lambda) = \frac{N}{2} \log \lambda - \frac{\lambda}{2} \sum_{i=1}^N (x_i - \mu)^2 + \frac{1}{2} \log(\kappa_0 \lambda) - \frac{\kappa_0 \lambda}{2} (\mu - \mu_0)^2 + (a_0 - 1) \log \lambda - b_0 \lambda + \text{const}$$

$$\begin{aligned}
 \log q_{\mu}(\mu) &= E_{q_{\lambda}}(\log \tilde{p}(\mu, \lambda)) \\
 &= E_{q_{\lambda}}\left(-\frac{\lambda}{2} \sum_{i=1}^N (x_i - \mu)^2 - \frac{\kappa_0 \lambda}{2} (\mu - \mu_0)^2\right) + \text{const} \\
 &= -\frac{E_{q_{\lambda}}(\lambda)}{2} \left(\sum_{i=1}^N (x_i - \mu)^2 + \kappa_0 (\mu - \mu_0)^2\right) + \text{const}
 \end{aligned}$$

我们可以通过对比高斯分布的对数形式：

$$\begin{aligned}
 \ln p(x) &= \ln \frac{1}{\sqrt{2\pi}} - \ln \sigma - \frac{(x - \mu)^2}{2\sigma^2} \\
 \log q_{\mu}(\mu) &= -\frac{E_{q_{\lambda}}(\lambda)}{2} \left(\sum_{i=1}^N (x_i - \mu)^2 + \kappa_0 (\mu - \mu_0)^2\right) + \text{const}
 \end{aligned}$$

计算得到：

$$\begin{cases} \mu_N = \frac{\kappa_0 \mu_0 + N \bar{x}}{\kappa_0 + N} \\ \kappa_N = (\kappa_0 + N) E_{q_{\lambda}}(\lambda) \end{cases} \quad \begin{cases} E_{q(\mu)}(\mu) = \mu_N \\ E_{q(\mu)}(\mu^2) = \frac{1}{\kappa_N} + \mu_N^2 \end{cases}$$

接着更新 $\lambda$ ：

$$\log \tilde{p}(\mu, \lambda) = \frac{N}{2} \log \lambda - \frac{\lambda}{2} \sum_{i=1}^N (x_i - \mu)^2 + \frac{1}{2} \log(\kappa_0 \lambda) - \frac{\kappa_0 \lambda}{2} (\mu - \mu_0)^2 + (a_0 - 1) \log \lambda - b_0 \lambda + \text{const}$$

$$\begin{aligned}
 \log q_{\lambda}(\lambda) &= E_{q_{\mu}}(\log \tilde{p}(\mu, \lambda)) \\
 &= E_{q_{\mu}}\left(\frac{N}{2} \log \lambda - \frac{\lambda}{2} \sum_{i=1}^N (x_i - \mu)^2 + \frac{1}{2} \log(\kappa_0 \lambda) - \frac{\kappa_0 \lambda}{2} (\mu - \mu_0)^2 + (a_0 - 1) \log \lambda - b_0 \lambda\right) + \text{const} \\
 &= \frac{N}{2} \log \lambda + \frac{1}{2} \log \lambda + (a_0 - 1) \log \lambda - b_0 \lambda - \frac{\lambda}{2} E_{q_{\mu}}\left(\kappa_0 (\mu - \mu_0)^2 + \sum_{i=1}^N (x_i - \mu)^2\right) + \text{const}
 \end{aligned}$$

通过对比Gamma分布对数形式：

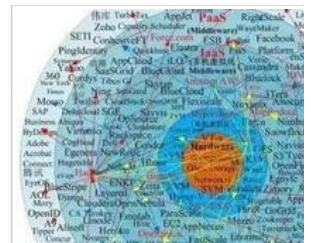
$$\begin{aligned}
 \ln p(x; \alpha, \beta) &= \alpha \ln \beta + (\alpha - 1) \ln x - \beta x - \ln \Gamma(\alpha) \\
 \log q_{\lambda}(\lambda) &= \frac{N}{2} \log \lambda + \frac{1}{2} \log \lambda + (a_0 - 1) \log \lambda - b_0 \lambda - \frac{\lambda}{2} E_{q_{\mu}}\left(\kappa_0 (\mu - \mu_0)^2 + \sum_{i=1}^N (x_i - \mu)^2\right) + \text{const}
 \end{aligned}$$

得到：

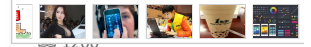
加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册



## 大数据学习路线



## 联系我们



请扫描二维码联系

webmaster@csdn.net

400-660-0111

QQ客服 950523531

关于 招聘 广告服务

©1999-2018 CSDN版权所有

京ICP证09002463号

经营性网站备案信息

网络110报警服务

中国互联网举报中心

北京互联网违法和不良信息举报中心

$$\begin{cases} a_N = a_0 + \frac{N+1}{2} \\ b_N = b_0 + \frac{1}{2} E_{q_{\lambda}} \left( \kappa_0 (\mu - \mu_0)^2 + \sum_{i=1}^N (x_i - \mu)^2 \right) \\ = b_0 + \kappa_0 (E(\mu^2) + \mu_0^2 - 2E(\mu)\mu_0) + \frac{1}{2} \sum_{i=1}^N (x_i^2 + E(\mu^2) - 2E(\mu)x_i) \end{cases}$$

同时：

$$E_{q(\lambda)}(\lambda) = \frac{a_N}{b_N}$$

将结果带入公式，根据：

$$\begin{cases} \mu_N = \frac{\kappa_0 \mu_0 + N \cdot \bar{x}}{\kappa_0 + N} \\ \kappa_N = (\kappa_0 + N) E_{q_{\lambda}}(\lambda) \end{cases} \quad \begin{cases} a_N = a_0 + \frac{N+1}{2} \\ b_N = b_0 + \kappa_0 (E(\mu^2) + \mu_0^2 - 2E(\mu)\mu_0) + \frac{1}{2} \sum_{i=1}^N (x_i^2 + E(\mu^2) - 2E(\mu)x_i) \end{cases}$$

以及：

$$\begin{cases} E_{q(\mu)}(\mu) = \mu_N \\ E_{q(\mu)}(\mu^2) = \frac{1}{\kappa_N} + \mu_N^2 \end{cases} \quad E_{q(\lambda)}(\lambda) = \frac{a_N}{b_N}$$

得到：

$$\begin{cases} \mu_N = \frac{\kappa_0 \mu_0 + N \cdot \bar{x}}{\kappa_0 + N} \\ \kappa_N = (\kappa_0 + N) \frac{a_N}{b_N} \end{cases}$$

这就是更新过程，公式虽然多一点，都没啥新东西，比较简单。

变分总结

变分既能够推断隐变量，也能推断未知参数，是一个有力的参数估计工具。难点在于推导复杂。

相对于采样，变分不容易计算但是效率很高。

平均场方法的变分推导，对离散和连续的隐变量都适用，在平均场方法的框架下，变分推导一次更新一个分布，其本质是坐标上升。

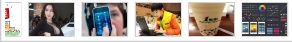
变分除了和贝叶斯结合得到VB，还可以和EM算法结合，得到VBEM，用于带隐变量和未知参数的估计，如GMM、LDA

到此变分算法就全部介绍完了，欢迎批评指正。

版权声明：本文为博主原创文章，未经博主允许不得转载。 <https://blog.csdn.net/u012771351/article/details/53098647>



大数据学习路线



联系我们



请扫描二维码联系  
✉ webmaster@csdn.net  
☎ 400-660-0111  
🗣 QQ客服

关于 招聘 广告服务  
©1999-2018 CSDN版权所有  
京ICP证09002463号

经营性网站备案信息  
网络110报警服务  
中国互联网举报中心  
北京互联网违法和不良信息举报中心

目前您尚未登录，请 [登录](#) 或 [注册](#) 后进行评论

### Variational Bayes ( 变分贝叶斯 )


( 转载于 ) <http://www.blog.huajh7.com/variational-bayes/> [关键词] 贝叶斯推断, 平均场理论, 变分估计, 贝叶斯推断, KL散度, 确定性估计...






Roning\_hit 2014年10月22日 16:32 5078

### EM算法的九层境界：Hinton和Jordan理解的EM算法

原文链接：点击打开链接 摘要：前言 为什么说EM算法是他们强强发力的领域呢？ 这里我们讨论Hinton和统计大神Jordan的强强发力的领域。当Bayes网络发展到高级阶段， 概率图模型使...



qq\_40954115 2017年12月18日 14:44 3




### 程序猿不会英语怎么行？英语文档都看不懂！

软件工程出身的英语老师，教你用数学公式读懂天下英文→



### Expectation-Maximum ( EM算法 )


EM算法原理，深度理解。首先它是解决含隐变量（ latent variable ）情况下的参数估计问题，而求模型的参数时一般采用最大似然估计，由于含有隐变量，所以对似然函数参数求导是求不出来的，虽然通...



qq\_39388410 2017年10月14日 17:44 796

### 变分贝叶斯推断(Variational Bayes Inference)简介


通常在研究贝叶斯模型中，很多情况下我们关注的是如何求解后验概率(Posterior)，不幸的是，在实际模型中我们很难通过简单的贝叶斯理论求得后验概率的公式解，但是这并不影响我们对贝叶斯模型的爱——既然...



aws3217150 2017年02月25日 16:42 6778

### 无向图的几个基本算法应用

简介 最近在看一些图相关的问题。实际上关于图相关的研究和问题已经非常多了。在前面的几篇文章里，我也谈到过图的定义、遍历法，扩展树生成和最短路径等问题。除了这些问题及应用以外，还有一...



yafeichang 2016年12月26日 20:03 3289


### 一秒创造无法计算的价值

每满2000返200，最高返5000元代金券



### 电商推荐算法应用

一、 电商推荐算法简述 目前比较多的电商模式为B2B，B2C，O2O，在本文介绍和需要举例说明的地方都以B2B电商模式为主。 电商推荐根据推荐内容不同分为物品推荐、商家推荐；流行的推荐应用主要有三个方...



christmasxu 2017年07月30日 16:48 1725

### 数据挖掘十大算法总结--核心思想，算法优缺点，应用领域


参考了许多资料加上个人理解，对十大算法进行如下分类： •分类算法：C4.5,CART,Adaboost,NaiveBayes,KNN,SVM •聚类算法：KMeans •统计学习：EM •关联分析：A...



xie20112480233 2014年11月03日 10:43 9082

### 变分推断 ( variational inference )


大家对贝叶斯公式应该都很熟悉  $P(Z|X)=\frac{p(X,Z)}{\int p(X,Z)dz}$   $P(X,Z)=\frac{p(X,Z)}{\int p(X,Z)dz}$   $P(X,Z)=\frac{p(X,Z)}{\int p(X,Z)dz}$



u012436149 2017年02月12日 12:13 5027

### 变分算法

注：本文中所有公式和思路来自于邹博先生的《机器学习升级版》，我只是为了加深记忆和理解写的本文。 本文介绍的变分算法是机器学习中的参数估计算法，跟数学中的变分法是有一些不一样的。在搜索引擎中搜索变分...



u012771351 2016年11月09日 08:56 1620



### 大数据学习路线



### 联系我们



请扫描二维码联系  
webmaster@csdn.net  
400-660-0111  
QQ客服

关于 招聘 广告服务

©1999-2018 CSDN版权所有  
京ICP证09002463号

经营性网站备案信息  
网络110报警服务  
中国互联网举报中心  
北京互联网违法和不良信息举报中心

Machine Learning读书会，面试&算法讲座，算法公开课，创业活动，算法班集锦 近期活动： 2014年9月3日，第8次西安面试&算法讲座视频 + PPT 的下载地址：...

 yansmile1 2015年09月10日 17:02  2292

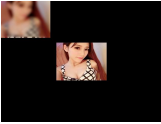
【机器学习】Tensorflow概率编程: 贝叶斯线性回归、变分贝叶斯与黑盒变分推断

我们首先看看Tensorflow概率编程最简单的实例——贝叶斯线性回归。以及背后的数学知识：变分贝叶斯和黑盒推断。...

 lpsl1882 2017年06月30日 23:46  1560

技术外文文献看不懂？教你一个神器，秒懂英语

不背单词和语法，一个公式学好英语



变分推断(variational inference)学习笔记(1)——概念介绍

ref : <http://www.crescentmoon.info/?p=709#more-709> 问题描述 变分推断是一类用于贝叶斯估计和机器学习领域中近似计算复杂 ( intractable...

 AiTODD1 2014年11月13日 21:04  7047

三硬币模型的变分贝叶斯EM解法

 Wind\_Desert 2016年10月29日 00:27  849

将三硬币模型推广为贝叶斯概率并利用变分贝叶斯EM算法来估计参数

贝叶斯推断--Gibbs Sampling

 u010533386 2016年05月19日 03:33  18151

贝叶斯推断 gibbs sampling初识

[转]VB6里的两个你可能没看懂的东西

 GnuDoyng 2004年10月06日 23:45  1253

Q :--请问VB中hwnd属性的含义是什么 Re: Microsoft Windows 操作系统通过给每一个窗口指定一个唯一的标识号(窗口句柄或 h Wnd)来管理所有的窗口。操作系统连续地监视每...

Variational Bayes

 fennvde007 2013年11月20日 09:59  1393

关键词| 贝叶斯推断，平均场理论，变分估计，贝叶斯推断，KL散度，确定性估计 一、前言 变分贝叶斯方法最早由Matthew J.Beal在他的博士论文《Variational Algorithms...

机器学习(2) 变分推断

 GZHermit 2017年03月26日 10:55  905

变分 对于普通的函数 $f(x)f(x)$ ，我们可以认为 $ff$ 是一个关于 $xx$ 的一个实数算子，其作用是将实数 $xx$ 映射到实数 $f(x)f(x)$ 。那么类比这种模式，假设存在函数算子 $FF$ ，它是关于 $f(x)f(x)$ 的...

自动变速箱维修大概要多少钱

变速箱维修价格

百度广告



PRML读书会第十章 Approximate Inference ( 近似推断，变分推断，KL散度，平均...

第十章的主要内容是变分推断 ( Variational Inference )，由中科院自动化所戴玮博士 ( 前后分三次讲完。精彩内容有：为什么需要近似推断、变分推断用到的KL散度、根据平均场 ( Mean Fie...

 Nietzsche2015 2015年02月03日 15:51  9201

Variational Inference入门：variational bayesian EM

实现了variational bayesian EM进行聚类，并与之前的k-means和EM求解高斯混合模型 ( GMM ) 的结果进行了对比。 ...

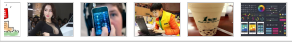
加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！

登录

注册



大数据学习路线



联系我们



请扫描二维码联系  
✉ webmaster@csdn.net  
☎ 400-660-0111  
💬 QQ客服

关于 招聘 广告服务  
©1999-2018 CSDN版权所有  
京ICP证09002463号

经营性网站备案信息  
网络110报警服务  
中国互联网举报中心  
北京互联网违法和不良信息举报中心



GMM的EM算法实现

abcjennifer 2012年11月19日 11:03 106080

在 聚类算法K-Means, K-Medoids, GMM, Spectral clustering , Ncut一文中我们给出了GMM算法的基本模型与似然函数，在EM算法原理中对EM算法的实现与收敛性证...

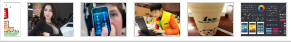
  
1







大数据学习路线



联系我们



请扫描二维码联系  
✉ webmaster@csdn.net  
☎ 400-660-0111  
🗣 QQ客服 95511

关于 招聘 广告服务

©1999-2018 CSDN版权所有  
京ICP证09002463号

经营性网站备案信息  
网络110报警服务  
中国互联网举报中心  
北京互联网违法和不良信息举报中心