



# Xianling Mao的专栏

Explore the world with my head and feet.

目录视图

摘要视图

**RSS** 订阅

个人资料



xianlingmao

访问：900007次

积分：4578

等级：

排名：第6289名

原创：27篇

转载：13篇

译文：0篇

评论：159条

文章搜索

文章分类

[Emacs](#) (3)

[mathmatics](#) (16)

[other](#) (8)

[QA](#) (0)

[信息抽取](#) (5)

[信息检索](#) (2)

[图模型](#) (5)

[机器学习](#) (14)

[模拟与采样](#) (5)

[topic model](#) (6)

[lda](#) (2)

[gibbs sampling](#) (3)

[latex](#) (1)

[Deep Learning](#) (1)

文章存档

[2013年01月](#) (1)

[2012年09月](#) (1)

[2012年08月](#) (2)

[2012年07月](#) (5)

[2012年06月](#) (3)

展开

阅读排行

[深入理解拉格朗日乘法](#) (231386)

[随机模拟的基本思想和常](#) (64191)

[浅谈深度学习\(Deep Leai](#)

赠书 | [AI专栏（AI圣经！《深度学习》中文版）](#)

[机器学习&数据挖掘 系统实训](#)

[【获奖公布】征文 | 你会为 AI 转型么？](#)

深入理解模拟退火算法（Simulated Annealing）

2012-07-29 12:2042081人阅读评论(1)收藏举报

分类：[模拟与采样（4）](#)[机器学习（13）](#)[mathmatics（15）](#)

版权声明：本文为博主原创文章，未经博主允许不得转载。

本文将对模拟退火算法（Simulated Annealing)进行介绍，深入理解这个算法。

模拟退火算法和上一篇文章[随机模拟算法](#)中的Metropolis算法有着紧密的联系，在这里将详细探讨这种关系。

我们先从这个算法要解决的问题出发，逐步引出相应的算法。（pku，sewm，shinning）

一. 问题

人们经常遇到这样的问题：在某个定义域S内，求某个函数f(x)的最小值，形式化为Min f(x)，x属于S。这是一个优化问题，根据f(x)的形式不同，有很多的优化算法来解决这类问题，简单的有穷举法（适用于定义域小的情况），图解法，数学分析法（求导数法）等精确算法，如果很难精确求得，还有很多的近似求解法，如贪心法（如爬山法，最速下降法，梯度下降法），随机模拟方法（MCMC等）。本文将介绍的模拟退火方法属于随机模拟方法，但是是求得精确解的（概率为1求得全局优化解），神奇吧！

二. 算法的基本思想

那怎么求解f(x)的最小值呢？它的思想是利用Gibbs分布：S中的一个点x的概率满足分布：p(x, T) = EXP(-f(x)/T) / Z，Z是归一化因子。现在假设T-> 0，可以看到f(x)越小，p(x,T)越大，意味着f(x)以非常大的概率取得最小值，这就是基本思想。现在是怎么操作的问题，这个分布是确定的，T已知，f(x)的表达式已知（对一个值，就是可以计算出来量），p(x,T)这个分布的形式就已知了，这时候我们利用[Metropolis模拟算法](#)去对这个分布抽样，在样本中，出现最多的非常有可能就是我们要求的最小值。似乎问题都解决了，然而如果直接把T设为很小，这个算法的计算时间太长了，因为S可能太大！不实用！

因此，我们采用了层层推进的办法，我们先把T设为一个较大的值Ti，然后在这个Ti的情况下采样，这个时候由于约束条件的限制，p(x, Ti)的定义域只是S的子集，因此速度较快，在采样分布稳定之后，再通过一个T的控制函数g(Ti, Tj)（是一个单调不增函数，T>=0）得到一个更小的Tj，然后在Tj的情况下，再对分布p(x, Tj)采样，直到达到稳定分布，我们可以知道这个时候f(x)比起Ti时候的f(x)变得更加小了，直到达到较小T的时候（自己设定）我们可以认为达到了优化解的情况。

三.模拟退火算法（simulated annealing）

在这一节，我们把算法总结在这里：

设定起始值T0，T的控制函数g(Ti, Tj)

while(T还没有达到我们设定的最小值)

{

对这时的T情况下的Gibbs分布p(x, T)用Metropolis模拟算法进行采样x(t)，直到达到稳定状态为止, 然后进入下一次循环（这句话的伪代码描述见下面）

}

循环中对p(x,T)进行Metropolis采样的算法我们放在这里，它也是一个循环

1 of 3

2017年08月07日 08:35

模型选择的几种方法：AI (62319)

狄利克雷过程（dirichlet | (56202)

核方法(kernel method)的 (49746)

什么叫共轭先验或者共轭 (43730)

俄罗斯的数学太牛逼了， (43482)

深入理解模拟退火算法（ (42146)

话题模型（topic model） (42073)

(34547)

评论排行

深入理解拉格朗日乘法 (66)

随机模拟的基本思想和常 (23)

核方法(kernel method)的 (16)

浅谈深度学习(Deep Lear (13)

一. 图模型（graphical m (9)

看懂信息检索和网络数据 (6)

Topic Model的分类和设 (3)

话题模型（topic model） (3)

狄利克雷过程（dirichlet | (3)

EM算法 (3)

推荐文章

\* CSDN日报20170725——《新的开始，从研究生到入职亚马逊》

\* 深入剖析基于并发AQS的重入锁(ReentrantLock)及其Condition实现原理

\* Android版本的"Wannacry"文件加密病毒样本分析(附带锁机)

\* 工作与生活真的可以平衡吗？

\* 《Real-Time Rendering 3rd》提炼总结——高级着色：BRDF及相关技术

\* 《三体》读后思考-泰勒展开/维度打击/黑暗森林

最新评论

随机模拟的基本思想和常用采样: task\_force: 楼主，有一事请教，对g(x)分布进行采样得到样本值x\_i后，和式为什么不是\sum\_1^n？

随机模拟的基本思想和常用采样: task\_force: 楼主，有一点不明白希望请教一下，n个样本按照g(x)的分布进行采样，则积分为什么变为和式{\sum\_...

核方法(kernel method)的主要思: roguesir: 赞

什么叫共轭先验或者共轭分布？ qq\_30330939: 不太对哦，请看一下维基百科，那里明明是说先验和后验共轭

深入理解拉格朗日乘法（Lagr: MaRinQ: 解释得非常不错 赞一个！

核方法(kernel method)的主要思: jason-刘: 请教楼主一个问题，证明一个函数满足Mercer定理时，训练数据集是怎么选取的？大小有规定吗？ 还有...

测度论--长度是怎样炼成的[zz] death\_ray: 谢谢，学实分析感觉怀疑自己的数学体系，看完了之后豁然开朗的感觉。

已知T，当前的样本x(t)(来自上一次循环的最后一个样本)，建议分布h(x(t), x(t+1))

while(未达到稳定状态)

```
{  
  
    先抽取一个建议样本y ~ h(x(t), x(t+1))  
  
    if( f(y) < f(x(t)) { 令x(t+1) = y; }  
  
    else  
  
    {  
  
        令a = min{1 , p(x(t+1), T)/p(x(t), T)};  
  
        然后抽取u ~ Uniform(0, 1);  
  
        if( u < a) {x(t+1) = y};  
  
        else { x(t+1) = x(t) };  
  
    }  
}
```

上面这个循环完全就是Metropolis算法，详细的解释可以看我的上一篇文章：[模拟计算](#)。

（pku，sewm，shinning）

至于，这个算法的名字：模拟退火，来自于我们这个过程和物理学中打造钢铁时候的加热-等温-降温过程，这里的T为温度，f(x)可以令为能量，上述的Metropolis过程就相当于等温过程，p(x, T)就是在T时候系统所处状态的概率分布，可以明显的看到，这个概率与系统的能量和温度密切相关，大部分的资料都是先将这个物理过程，然后再讲算法本身，如果人们不熟悉这个物理过程，简单的介绍反而会使得大家对这个算法搞不清楚。因此，本文为了叙述清楚，没有描述这个过程，反而更加的清楚！

（pku，sewm，shinning）

顶

踩

5

2

上一篇

随机模拟的基本思想和常用采样方法（sampling）

下一篇

梯度、Hessian矩阵、平面方程的法线以及函数导数的含义

相关文章推荐	
<div><div></div><div>模拟退火算法的个人理解</div></div>	<div><div></div><div>深入理解拉格朗日乘法（Lagrange Multiplier）...</div></div>
<div><div></div><div>深入理解拉格朗日乘法（Lagrange Multiplier）...</div></div>	<div><div></div><div>深入理解拉格朗日乘法（Lagrange Multiplier）...</div></div>
<div><div></div><div>拉格朗日乘法</div></div>	<div><div></div><div>解密SVM</div></div>
<div><div></div><div>拉格朗日乘法</div></div>	<div><div></div><div>拉格朗日乘子法和KKT条件</div></div>
<div><div></div><div>拉格朗日乘法</div></div>	<div><div></div><div>机器学习数学原理（5）——广泛拉格朗日乘法</div></div>



看懂信息检索和网络数据挖掘领  
任蒙蒙-Angel: 认真琢磨，对于提高写作能力很有帮助

深入理解拉格朗日乘子法（Lagrangian multiplier）  
yandoudoujiayou: 请问博主，在求解带有1范数和核范数的模型中使用增广拉格朗日函数，等式约束为矩阵Z3=Z1,Z3=Z2...

随机模拟的基本思想和常用采样方法  
friskit: 博主你好，公式中的inf貌似应该换成int。\\int是积分，\\inf是下确界

猜你在找

- 【直播】机器学习&数据挖掘7周实训--韦玮

【直播】3小时掌握Docker最佳实战-徐西宁

【直播】计算机视觉原理及实战--屈教授

【直播】机器学习之矩阵--黄博士

【直播】机器学习之凸优化--马博士
- 【套餐】系统集成项目管理工程师顺利通关--徐朋

【套餐】机器学习系列套餐（算法+实战）--唐宇迪


【套餐】微信订阅号+服务号Java版 v2.0--翟东平

【套餐】微信订阅号+服务号Java版 v2.0--翟东平

【套餐】Javascript 设计模式实战--曾亮

查看评论

1楼 [zhangyuereal](#) 2013-08-25 21:11发表



博主写的真好，但是我还有些疑问。"我们采用了层层推进的办法，我们先把T设为一个较大的值Ti，然后在采样，这个时候由于约束条件的限制，p(x，Ti)的定义域只是S的子集，因此速度较快，在采样分布稳定之后，我们得到一个更小的Tj，然后在Tj的情况下，再对分布p(x, Tj)采样，我们可以知道这个时候f(x)比起Ti时候的f(x)变得更加小了，直到达达到较小T的时候（自己设定）我们可以知道达到了最优解的情况"

①为何说，当T 是一个较大值得时候，p(x，Ti)的定义域只是S的子集？②如何判断MCMC采样已经收敛？③如何设定？直接用最小值，进行采样，可以么？④别的资料讲模拟退火的时候，没有提到MCMC.我一直在找这方面的资料，这篇文章给了我很多启发。有没有也是讲这两者联系的资料，请推荐我看一下。

发表评论

用 户 名：

haijunz

评论内容：



提交

\* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

公司简介 | 招贤纳士 | 广告服务 | 联系方式 | 版权声明 | 法律顾问 | 问题报告 | 合作伙伴 | 论坛反馈

网站客服 杂志客服 微博客服 webmaster@csdn.net 400-660-0108 | 北京创新乐知信息技术有限公司 版权所有 | 江苏知之为计算机有限公司 |

江苏乐知网络技术有限公司

京 ICP 证 09002463 号 | Copyright © 1999-2017, CSDN.NET, All Rights Reserved



关闭