

凸优化(七)——牛顿法

Herbert002 (/u/fada150f26bf) [+ 关注](#)

2016.08.08 17:23* 字数 779 阅读 6523 评论 0 喜欢 3

(/u/fada150f26bf)

〇、说明

凸优化主要学习《凸优化》(Stephen Boyd等著, 王书宁等译)[1]这本书。学习过程中, 对其内容的理解时有困惑, 也参考一些其他书籍资料。笔者尽量将这部分知识整理地简洁明了, 成此系列笔记。

如有错误疏漏, 烦请指出。如要转载, 请联系笔者, hpf_2006pyy@163.com。

一、简介

用目标函数的二阶泰勒展开近似该目标函数, 通过求解这个二次函数的极小值来求解凸优化的搜索方向。

二、推导

2.1、牛顿法推导

目标函数 $f(x)$ 在 x 处的二阶 Taylor 展开为

$$f(x+v) \approx f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v,$$

当 x 固定时, v 取多少可以使 $f(x+v)$ 最小呢? 可以考虑上式的右边, 其为 v 的二次函数。

使上式两边对 v 求偏导

$$\frac{\partial}{\partial v} f(x+v) = \nabla f(x) + \nabla^2 f(x) v,$$

当 $f(x+v)$ 最小时, $f(x+v)$ 对 v 的偏导为零, 即

$$\nabla f(x) + \nabla^2 f(x) v = 0$$

则有

$$v = -\nabla^2 f(x)^{-1} \nabla f(x)$$



将这个方向作为搜索方向，这个方向被称作 Newton 步径，即

$$\Delta x_{nt} = -\nabla^2 f(x)^{-1} \nabla f(x),$$

$\nabla^2 f(x)$ 即为函数 $f(x)$ 的 Hessian 矩阵，由凸函数的 Hessian 矩阵的正定性可知，除 $f(x)$ 到达极值点 $\nabla f(x) = 0$ 外，有

$$\nabla f(x)^T \Delta x_{nt} = -\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) < 0$$

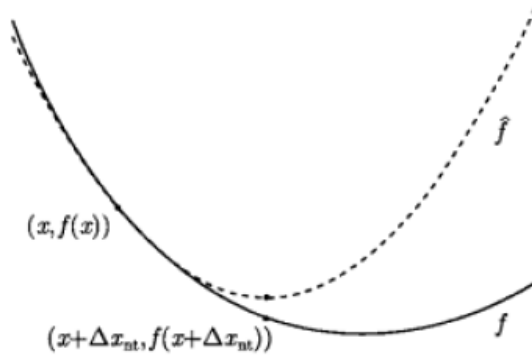


图1[1]

图 1 大概说明了牛顿法的原理，实线是目标函数 $f(x)$ ，虚线是 $f(x)$ 在 x 处的 Taylor 二阶展开。

2.2、Hessian范数下的最速下降方法

Hessian 矩阵定义的二次范数

$$\|u\|_{\nabla^2 f(x)} = (u^T \nabla^2 f(x) u)^{1/2}$$

Newton 步径是采用 Hessian 矩阵定义的二次范数导出的最速下降方法。

这从另一个角度揭示了为什么 Newton 步径是好的搜索方向了。

这里我没有去查找证明过程，我觉得只要知道就可以了，因为这有助于理解最速下降方法(《凸优化(六)——最速下降法》)。

三、优势

在实际应用中，牛顿法往往比梯度下降法有更少的迭代次数。

2.2 已经从一个角度说明了 Newton 步径是好的搜索方向。

知乎问答《最优化问题中，牛顿法为什么比梯度下降法求解需要的迭代次数更少？》[2] 这篇也讲了一些，其中，排名第一的引自 Wiki 的“从几何上说，牛顿法就是用一个二次曲面去拟合你当前所处位置的局部曲面，而梯度下降法是用一个平面去拟合当前的局部曲

(/apps/redi
utm_sourc
banner-clc



面，通常情况下，二次曲面的拟合会比平面更好，所以牛顿法选择的下降路径会更符合真实的最优下降路径”，比较有说服力和概括性。

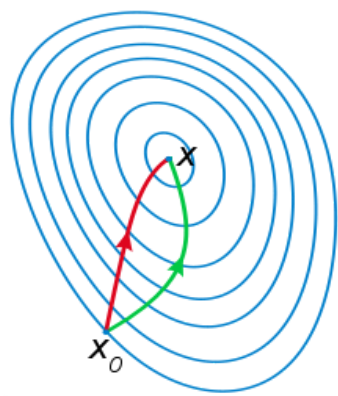


图2[2]

图2形象地说明了牛顿法和梯度下降法的区别，红色为牛顿方法搜索路径，绿色为梯度下降法搜索路径。

四、拟牛顿法

牛顿法需要计算目标函数Hessian矩阵的逆矩阵，运算复杂度太高，计算效率很低，尤其维数很大时。拟牛顿算法的核心思想用一个近似矩阵替代逆Hessian矩阵。

五、等式约束的牛顿法

等式约束凸优化问题，

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

其中对应的 $f(x)$ 是凸函数。

等式约束的 Newton 法和无约束 Newton 法有两点不同：

- 1、初始点必须可行，即 $x_0 \in \text{dom } f, \mathbf{Ax}_0 = \mathbf{b}$,
- 2、Newton 方向 Δx_{nt} 是可行方向，即 $\mathbf{A}\Delta x_{nt} = 0$ 。

附录

A、参考

- [1]、《凸优化》，Stephen Boyd等著，王书宁等译
- [2]、《最优化问题中，牛顿法为什么比梯度下降法求解需要的迭代次数更少？》
(<https://link.jianshu.com?t=http://www.zhihu.com/question/19723347>)

B、相关目录

凸优化(一)——概述 (<https://www.jianshu.com/p/fe2e7f0e89e5>)

凸优化(二)——凸集 (<https://www.jianshu.com/p/3204322ec1a2>)

凸优化(三)——凸函数 (<https://www.jianshu.com/p/967e1ea18469>)

凸优化(四)——问题求解 (<https://www.jianshu.com/p/1e327f82ab9a>)

凸优化(五)——回溯直线搜索 (<https://www.jianshu.com/p/86072fcf3bca>)

凸优化(六)——最速下降法
(<https://www.jianshu.com/p/1238602a2720>)

凸优化(七)——牛顿法

凸优化(八)——Lagrange对偶问题 (<https://www.jianshu.com/p/96db9a1d16e9>)

C、时间线

2016-08-08 第一次发布

小礼物走一走，来简书关注我

赞赏支持

机器学习 (/nb/3270818)

举报文章 © 著作权归作者所有



Herbert002 (/u/fada150f26bf)

写了 17508 字，被 261 人关注，获得了 141 个喜欢
(/u/fada150f26bf)

+ 关注

喜欢 | 3



更多分享



下载简书 App ▶
随时随地发现和创作内容



(/apps/redirect?utm_source=note-bottom-click)




登录 (/sign-in?utm_source=desktop&utm_medium=not-signed-in-com)

评论

(/apps/redi
utm_sourc
banner-clic

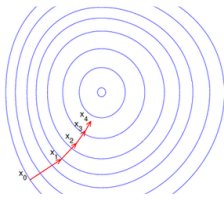
智慧如你，不想发表一点想法 (/sign_in?utm_source=desktop&utm_medium=not-signed-in-nocomments-text)咩~

被以下专题收入，发现更多相似内容




凸优化 (/c/1a466847fcb8?utm_source=desktop&utm_medium=notes-included-collection)

(/p/74ea8f2695dd?




utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendatio
【转】常见的几种最优化方法 (/p/74ea8f2695dd?utm_campaign=maleski...

转自Poll 的笔记 阅读目录 梯度下降法（Gradient Descent） 牛顿法和拟牛顿法（Newton's method & Quasi-Newton Methods） 共轭梯度法（Conjugate Gradient） 启发式优化方法 解决约束优化问题——拉格朗日...

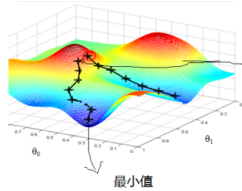
 JSong1122 (/u/94e76e4d33d5?
utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio

机器学习算法小结与收割offer遇到的问题 (/p/ace5051d0023?utm_campai...

机器学习是做NLP和计算机视觉这类应用算法的基础，虽然现在深度学习模型大行其道，但是懂一些传统算法的原理和它们之间的区别还是很有必要的。可以帮助我们做一些模型选择。本篇博文就总结一下各种机...


 在河之简 (/u/5ff1acaa6334?
utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio

(/p/a46e08a8f7bb?



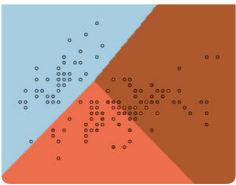
utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendatio
梯度下降法（Gradient Descent） (/p/a46e08a8f7bb?utm_campaign=ma...

转载-刘建平Pinard-www.cnblogs.com/pinard/p/5970503.html 在求解机器学习算法的模型参数，即无约束优化问题时，梯度下降（Gradient Descent）是最常采用的方法之一，另一种常用的方法是最小二乘法。这...

 商三郎 (/u/9230d3e7cbab?
utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio

(/p/bbdeb356057e?






(/apps/redi
utm_sourc
banner-clip

utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendatio
Logistic回归(Logistic Regression)算法笔记(二)-scikit l... (/p/bbdeb3560...

本节不仅介绍了Logistic回归在sklearn中模型应用，还介绍了liblinear、牛顿法、拟牛顿法(DFP算法、BFGS
算法、L-BFGS算法)、梯度下降、随机梯度下降等，正文如下，欢迎围观喔~~（我的字迹请大家别吐槽了...

 keepStriving (/u/afb62bc37c48?

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio


(/p/e976e1853cb9?

1 (Quasi-Newton Methods) 是求解非线性优化问题最有效的方法之一。
美国 Argonne 国家实验室的物理学家 W. C. Davidon 提出。Da
当时看来是非线性优化领域最具创造性的发明之一。不久 R. Fletcher
证实了这种新的算法远比其他方法快速和可靠，使得非线性优化
迅速。在之后的 20 年里，拟牛顿方法得到了蓬勃发展，出现了大量
的相关论文 [10]。
法，BFGS 方法以及 L-BFGS 算法都是重要的拟牛顿法。本文将
当然，在介绍拟牛顿法之前，我们先看看什么是牛顿法，以及拟牛
法。为此，考虑如下无约束的极小化问题
$$\min_{x \in \mathbb{R}^n} f(x),$$

其中 $x_1, \dots, x_n \in \mathbb{R}$ 。由于本文不准备对收敛性理论进行讨论，因此
→ 我们假设 f 为凸函数，且两阶连续
0.1) 的解为 x^* 。

utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendatio
牛顿下降法和梯度下降法基础 (/p/e976e1853cb9?utm_campaign=maleski...

前言 学习机器学习一直是用梯度下降法的，对于牛顿下降法并没有什么了解，但是小学期同学的一个项目用
到了牛顿下降法，同时好多大神同学的夏令营之旅问到了牛顿下降法（我等弱鸡疯狂被刷。。。所以就...

 WZFish0408 (/u/0a357c1fefe2?


utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio

(/p/0c9ccbc61fd7?



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendatio
孤独的“偷影子人” (/p/0c9ccbc61fd7?utm_campaign=maleskine&utm_con...

“有些人只拥吻影子，于是只拥有幸福的幻影。”--莎翁 最近看了马克李维的《偷影子的人》，故事讲述的是一个经常受欺负的小男孩有超能力，可以偷别人的影子，看到别人说不...

 沉栖语墨 (/u/53cba5578515?


utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio

(/p/10903b409319?



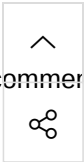
utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendatio
oh my god，你是在替儿子（女儿）选妃（驸马）吗 (/p/10903b409319?ut...

oh my god，你是在替儿子（女儿）选妃（驸马）了 文/林子 金星的新节目《中国式相亲》近日大火！各路口水直指“中国式父母”。相亲时遇到的奇葩事件其实并不止今年才有，只是如今愈演愈烈。倒是挺佩服金星...

 静待花开的林子 (/u/ae1a898fa3bc?

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio


(/p/10845cd97971?





utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendatio
推荐6款貌美又好用的APP (/p/10845cd97971?utm_campaign=maleskine...

只能说认识简书太晚了。简书最大的好处是，这个平台都是爱读书、爱写作的朋友，你在这里可以认识很多优秀的写作者，自己写的东西也会有很多人来给你评论，自己的文章也可以投稿，比你自已闷着头吭哧吭...

 童淇 (/u/5a5c1bacb4b2?)

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendatio

(/apps/redi
utm_sourc
banner-clic

