

CSDN

首页 博客 学院 下载 GitChat TinyMind ...

搜博主文章

写博客 发Chat 登录 注册

peghoty

学习是一种态度!

RSS订阅

个人资料

皮果提

关注

原创	粉丝	喜欢	评论
104	2659	213	897

等级： 博客 7 访问：167万+

积分：2万+ 排名：518

智能停车场

归档

2014年11月	1篇
2014年10月	4篇
2014年8月	1篇
2014年7月	6篇
2014年6月	4篇

展开

- 热门文章
- word2vec 中的数学原理详解（一）目录和前言

阅读量：141844
- word2vec 中的数学原理详解（四）基于 Hierarchical Softmax 的模型

阅读量：83222
- word2vec 中的数学原理详解（三）背景知识

阅读量：59585
- word2vec 中的数学原理详解（五）基于 Negative Sampling 的模型

阅读量：59145
- word2vec 中的数学原理详解（二）预备知

最新评论

word2vec 中的数学原理详解...
LINTEBI：博主，方便的话可以发一份PDF给我吗？感激不尽！感激不尽！604156331@qq.com

word2vec 中的数学原理详解...
Qinjian666：NB闪闪的博主

word2vec 中的数学原理详解...
leayc：#6.9.3: <https://github.com/RaRe-Technologies/gensim>...

word2vec 中的数学原理详解...
p03721：博主可以发一份PDF吗？感谢！875051204@qq.com

受限玻尔兹曼机（RBM）学习笔记（...
jocelynxyq：赞楼主的分享！

原 牛顿法与拟牛顿法学习笔记（二）拟牛顿条件

2014年03月24日 00:51:44

阅读数：26571

机器学习算法中经常碰到非线性优化问题，如 **Sparse Filtering** 算法，其主要工作在于求解一个非线性极小化问题。在具体实现中，大多调用的是成熟的软件包做支撑，其中最常用的一个算法是 **L-BFGS**。为了解这个算法的数学机理，这几天做了一些调研，现把学习过程中理解的一些东西整理出来。

目录链接

- (1) 牛顿法
- (2) 拟牛顿条件
- (3) DFP 算法
- (4) BFGS 算法
- (5) L-BFGS 算法



§2 拟牛顿法

如上节所述，牛顿法虽然收敛速度快，但是计算过程中需要计算目标函数的二阶偏导数，计算复杂度较高，而且有时目标函数的海森矩阵无法保持正定，从而使得牛顿法失效。为了克服这两个问题，人们提出了拟牛顿法。这个方法的基本思想是：不用二阶偏导数而构造出可以近似海森矩阵（或海森矩阵的逆）的正定对称阵，在“拟牛顿”的条件下优化目标函数。不同的构造方法产生了不同的拟牛顿法。

也有人把“拟牛顿法”翻译成“准牛顿法”，其实都是表示“类似于牛顿法”的意思啦，因为只是对用来计算搜索方向的海森矩阵（或海森矩阵的逆）作了近似计算罢了。至此，牛顿法和拟牛顿法之间的关系已经跟大家澄清啦。

在介绍具体的拟牛顿法之前，我们先推导一个拟牛顿条件，或者叫做拟牛顿方程，还有叫做割线条件 (Secant condition)。因为对海森矩阵（或海森矩阵的逆）做近似总不能随便近似吧，我们也需要理论指导，而拟牛顿条件则是用来提供理论指导的，它指出了用来近似的矩阵应该满足的条件。

为明确起见，下文中用 B 表示对海森矩阵 H 本身的近似，而用 D 表示对海森矩阵的逆 H^{-1} 的近似，即 $B \approx H, D \approx H^{-1}$ 。

§2.1 拟牛顿条件

设经过 $k+1$ 次迭代后得到 \mathbf{x}_{k+1} ，此时将目标函数 $f(\mathbf{x})$ 在 \mathbf{x}_{k+1} 附近作泰勒展开，取二阶近似，得到

$$f(\mathbf{x}) \approx f(\mathbf{x}_{k+1}) + \nabla f(\mathbf{x}_{k+1}) \cdot (\mathbf{x} - \mathbf{x}_{k+1}) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{x}_{k+1})^T \cdot \nabla^2 f(\mathbf{x}_{k+1}) \cdot (\mathbf{x} - \mathbf{x}_{k+1}), \quad (2.14)$$

在 (2.14) 两边同时作用一个梯度算子 ∇ ，可得

$$\nabla f(\mathbf{x}) - \nabla f(\mathbf{x}_{k+1}) \approx \nabla^2 f(\mathbf{x}_{k+1}) \cdot (\mathbf{x} - \mathbf{x}_{k+1}). \quad (2.15)$$

在 (2.15) 中，令 $\mathbf{g}_k = \nabla f(\mathbf{x}_k)$ ，可得

$$\mathbf{g}_{k+1} - \mathbf{g}_k \approx H_{k+1} \cdot (\mathbf{x}_{k+1} - \mathbf{x}_k), \quad (2.16)$$

若引入记号

$$\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k, \quad \mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k, \quad (2.17)$$

则 (2.16) 可紧凑地写成

$$\mathbf{y}_k \approx H_{k+1} \cdot \mathbf{s}_k \quad (2.18)$$

或者

$$\mathbf{s}_k \approx H_{k+1}^{-1} \cdot \mathbf{y}_k. \quad (2.19)$$

这就是所谓的**拟牛顿条件**，它对迭代过程中的海森矩阵 H_{k+1} 作约束，因此，对 H_{k+1} 做近似的 B_{k+1} ，以及对 H_{k+1}^{-1} 做近似的 D_{k+1} 可以将

$$\mathbf{y}_k = B_{k+1} \cdot \mathbf{s}_k \quad (2.20)$$

或者

$$\mathbf{s}_k = D_{k+1} \cdot \mathbf{y}_k. \quad (2.21)$$

作为指导.

接下来，我们依次介绍几种常见的拟牛顿法.

参考文献

- [1] <http://www.materialssimulation.com/node/625>
- [2] <http://www.codelast.com/?p=2780>
- [3] http://www.tydxq.cn/kuai_su/youhuasheji/suanfayuanli/3.1.asp
- [4] http://en.wikipedia.org/wiki/BFGS_method
- [5] http://en.wikipedia.org/wiki/Sherman-Morrison_formula
- [6] Dai Y H. A perfect example for the BFGS method[J]. Mathematical Programming, 2013, 138(1-2): 501-530.
- [7] Liu D C, Nocedal J. On the limited memory BFGS method for large scale optimization[J]. Mathematical programming, 1989, 45(1-3): 503-528.
- [8] Nocedal J. Updating quasi-Newton matrices with limited storage[J]. Mathematics of computation, 1980, 35(151): 773-782.
- [9] <http://baike.baidu.com/view/6062086.htm>

作者: peghoty

出处: <http://blog.csdn.net/itplus/article/details/21896619>

欢迎转载/分享, 但请务必声明文章出处.

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/peghoty/article/details/21896619>

文章标签： [拟牛顿方程](#) [割线条件](#) [海森矩阵](#) [Hessian Matrix](#)

个人分类： [数学天地](#)

为什么说大数据工程师会在2018年越过越滋润？

大数据生态系统在2017年终于实现了火力全开！这次我们总结了，给大家完整的展现下大数据工程师的详情，希望这篇分享可以帮到所有程序员！


[查看更多>>](#)


13596

[查看更多>>](#)

想对作者说点什么？ [我来说一句](#)

 **__hanbingtao__** 2018-01-16 17:15:08 #3楼
写的真好，赞一个！

 **sharpstill** 2015-06-08 22:30:34 #2楼
博主这个文章写的非常不错，是我查到的中文资料里牛顿迭代法里写的最不错的文章了，但是能否提供文字版本啊。我好转换成pdf打印，这个图片实在不方便啊？

 **bj20031960** 2014-09-16 16:43:51 #1楼 [查看回复\(2\)](#)
请问2.14式为什么要从第k+1得到的点 $x_{(k+1)}$ 入手，然后在2.15式带入前一次得到的点 x_k ？而不是顺着来，从 x_k 到 $x_{(k+1)}$ ？

[上一页](#) 1 [下一页](#)

数值优化（ Numerical Optimization ）学习系列-拟牛顿方法（ Quasi-Newton ）

概述 拟牛顿方法类似于最速下降法，在每一步迭代过程中仅仅利用梯度信息，但是通过度量梯度之间的变化，能够产生超线性的收敛效果。本节主要学习一下知识点： 1. 拟牛顿方程推导 2. ...

 **fangqingan_java** 2015-12-27 18:48:32 阅读数：3487

牛顿法与拟牛顿法

牛顿法求函数的根牛顿法的最初提出是用来求解方程的根的。我们假设点 x^* 为函数 $f(x)$ 的根，那么有 $f(x^*)=0$ 。现在我们把函数 $f(x)$ 在点 x_k 处一阶泰勒展...


 **batuwuhanpei** 2016-07-21 11:49:43 阅读数：8316

新出的微投资赚钱方式，2018聪明的苏坡人都在闷声发大财！

仕永投资管理 · 顶新

拟牛顿法 分析与推导

转自 <http://www.cnblogs.com/liuwu265/p/4714396.html>，侵权 网上查拟牛顿法的推导，找到一个感觉比较容易读懂的，保存下来 针对牛顿法...

 **wenzishou** 2017-06-16 19:36:35 阅读数：521

梯度下降、牛顿法、拟牛顿法

介绍 在向量微积分中，标量场的梯度是一个向量场。标量场中某一点上的梯度指向标量场增长最快的方向，梯度的长度是这个最大的变化率。更严格的说，从欧几里得空间 R^n 到 R 的函数的梯度是在 R^n 某一点最佳的线性...

 **a819825294** 2016-08-10 18:50:23 阅读数：7144

拟牛顿法

转自:ACdreamer 今天，我来讲一种在机器学习中常用到的优化算法，叫做BFGS算法。BFGS算法被认为是数值效果最好的拟牛顿法，并且具有全局收敛性和超线性收敛速度。那么接下...

 **lishuandao** 2016-03-31 16:10:35 阅读数：1917

【机器学习详解】解无约束优化问题：梯度下降、牛顿法、拟牛顿法

无约束优化问题是机器学习中最普遍、最简单的优化问题。 $x^* = \min_{x \in \mathbb{R}^n} f(x)$, $x \in \mathbb{R}^n$, $f: \mathbb{R}^n \rightarrow \mathbb{R}$. 梯度下降梯度下降是最简单的迭代优化算法，每一次迭代需求...

luoshixian099 2016-07-06 20:58:22 阅读数：6100

苏坡90后小伙在家无聊玩微信，存款惊呆父母

威能科技·顶新

牛顿法与拟牛顿法学习笔记（一）牛顿法

机器学习算法中经常碰到非线性优化问题，如 Sparse Filtering 算法，其主要工作在于求解一个非线性极小化问题。在具体实现中，大多调用的是成熟的软件包做支撑，其中最常用的一个算法是 L-BF...

peghoty 2014-03-24 00:51:18 阅读数：57081

拟牛顿法公式推导以及python代码实现（一）

目录 拟牛顿法 1.1拟牛顿法的导出与优点 1.2 算法步骤与特点 对称秩一校正公式 DFP算法 3.1 DFP公式推导 3.2 要求解的问题 3.3 python实现 1.拟牛...

qq_39422642 2017-12-22 19:11:16 阅读数：429

拟牛顿算法

最近在看条件随机场中的优化算法。其中就设计到了无约束化的最优化方法，也就是牛顿法。在CRF (conditional random field) 中，使用的是L-BFGS法。费了好大的劲把算法的原理及推...

seaserm 2015-11-05 14:34:58 阅读数：1454

牛顿法与拟牛顿法学习笔记（五）L-BFGS 算法

机器学习算法中经常碰到非线性优化问题，如 Sparse Filtering 算法，其主要工作在于求解一个非线性极小化问题。在具体实现中，大多调用的是成熟的软件包做支撑，其中最常用的一个算法是 L-BF...

peghoty 2014-03-24 00:53:27 阅读数：45952

梯度法（SGD）、拟牛顿法（LBFGS）与共轭梯度法（CG）

一、基本原理 梯度法：由一阶泰勒级数展开式， $f(x+dx) = f(x) + \sum(i, df/dx(i) * dx(i)) + r(x, dx)$ 。其中 $r(x, dx)$ 为余项，当 dx 很小时，可忽略余项。...

student1218 2015-07-14 18:15:26 阅读数：3579

牛顿法与拟牛顿法学习笔记（四）BFGS 算法

机器学习算法中经常碰到非线性优化问题，如 Sparse Filtering 算法，其主要工作在于求解一个非线性极小化问题。在具体实现中，大多调用的是成熟的软件包做支撑，其中最常用的一个算法是 L-BF...

peghoty 2014-03-24 00:53:04 阅读数：44637

算法细节系列（3）：梯度下降法，牛顿法，拟牛顿法

算法细节系列（3）：梯度下降法，牛顿法，拟牛顿法迭代算法原型话不多说，直接进入主题。在我看来，不管是梯度下降法还是牛顿法，它们都可以归结为一个式子，即 $x = \phi(x)$ 也就是...

u014688145 2016-12-16 09:17:14 阅读数：3941

5分钟完成加壳，防止代码反编译

Virbox Protector加壳工具，堪比VMP安全强度的加壳工具



【数学】梯度下降，牛顿法与拟牛顿法

这三个优化算法，实在是太过经典，以至于很多文章都在说这个算法。这里主要就写一写我自己的感悟吧。剩下的再集成一下别的感觉比较好的微博 梯度下降 牛顿法 拟牛顿法 参考文献...

haolexiao 2017-03-16 00:20:18 阅读数：872

统计学习方法-牛顿法和拟牛顿法

牛顿法和拟牛顿法 牛顿法和拟牛顿法是求解无约束最优化问题的常用方法，有收敛速度快的优点。牛顿法是迭代算法，每一步需要求解目标函数的海赛矩阵的逆矩阵，计算比较复杂。拟牛顿法通过正定矩阵近似海赛矩阵的...

Jack_lyp2017 2017-12-01 16:19:05 阅读数：264

优化算法——拟牛顿法之BFGS算法

一、BFGS算法简介 BFGS算法是使用较多的一种拟牛顿方法，是由Broyden，Fletcher，Goldfarb，Shanno四个人分别提出的，故称为BFGS校正。 同DFP校...

google19890102 2015-05-20 11:31:14 阅读数：13885

最优化学习笔记(十九)——拟牛顿法(5)BFGS算法

一、BFGS算法的更新公式 为了推导BFGS算法，需要用到对偶或者互补的概念，前边已经讨论过hessian矩阵逆矩阵的近似矩阵需要满足以下条件： $H_{k+1}\Delta g(i) = \Delta x(i) 0 \leq i \leq k$...

chunyun0716 2017-02-12 10:34:08 阅读数：2428

优化算法——牛顿法(Newton Method)

一、牛顿法概述

google19890102 2014-11-13 22:00:59 阅读数：25935

免费云主机试用一年

云服务器免费试用

百度广告



（斯坦福机器学习课程笔记）牛顿法算法学习

代码均使用PYTHON3.x 牛顿法算法复杂些，作以下笔记，免忘。下图（来自<http://www.myexception.cn/cloud/1987100.html>，侵删）为了找到f(...

qq_32231743 2016-12-20 19:18:05 阅读数：2048

牛顿法

平时经常看到牛顿法怎样怎样，一直不得要领，今天下午查了一下维基百科，写写我的认识，很多地方是直观理解，并没有严谨的证明。在我看来，牛顿法至少有两个应用方向，1、求方程的根，2、最优化。牛顿法涉及到方程...

luoleicn 2011-06-05 17:11:00 阅读数：64737