

## 《最优化方法》课程的实践环节2

2017.11.16

- **目标:** 加深学生对所学知识的理解, 锻炼学生使用优化模型和软件解决实际应用问题的能力。
- **成绩:** 占总成绩的15%。
- **晚提交的减分方式:** 晚一天(24小时之内), 减30%; 晚两天(48小时之内), 减60%; 两天之后(48小时之后)的提交无效, 该项成绩为零。
- **提交指导:**
  - ◆ 截止时间: 2018. 12. 29(周五), 晚上11:00。
  - ◆ 提交文件的名称为OPT Student ID (例如, 您的学号为 1108143, 您提交的文件名应为OPT1108143.\*, 如果有多个文件, 请先放在该名称的文件夹内, 压缩该文件夹后提交)。
  - ◆ 提交作业方式: 登录课程中心该课程的“作业”页面, 提交作业。(建议在截止日期前一天提交, 防止网络出现问题。提交后, 在截止日期前可以再重新提交一次)。
  - ◆ 在2018.1.4(周四)的课堂上发布提交不成功(提交文件有问题或未按要求提交)的学生或者未提交学生的信息。
- **要求:** 以下四种方式中任选一种方式, 按要求完成; 选择 4 的同学可以组队完成, 但每队不超过两个人, 并在作业开始部分说明 1) 题目来源, 2) 每个人完成的工作内容及占总工作量的百分比。

### 1 使用优化软件、编写重要算法的程序(文档8分, 程序7分)

编写部分重要算法的程序、观察并解释算法的行为和性能。具体完成教材上的习题5.6, 5.7, 5.8, 5.9, 5.19, 5.27, 6.4, 要求画出算法流程图或者给出**算法的伪代码**, 给出其中**重要数量(比如梯度、牛顿法或者信赖域法中的Hessian阵)**的表达式; 以恰当的方式**(比如图片和表格等)**展示计算结果(需要说明与结果相对应的算法的参数设置), 并对计算结果给予充分的**说明和分析**。

### 2 完成两个无约束优化的数值实验(文档8分, 程序7分)

完成提供的两个无约束优化的实验(见文件two\_uncon. pdf), 并按要求提交实验报告。实验报告包括问题重述, 给出其中重要量的表达式(比如梯度、牛顿法或者信赖域法中的Hessian阵等), 以恰当的方式(比如图片和表格等)展示计算结果, 并对计算结果给予充分的说明和分析。最后陈述计算过程遇到的问题及解决办法。

### 3 撰写一个数值报告(文档8分, 程序7分)

在提供的问题集(见文件“continous\_problems. pdf”)中任选一个问题, 完整阐述原问题、相应的优化问题(包括离散化过程)、利用优化软件求解所得到的具体问题、解释所得结果; 并根据所用软件的帮助文档, 简要阐述所调用函数使用的优化

算法的原理。

#### 4 撰写课程小论文(文档8分, 仿真实验7分)

请寻找所学知识的相关应用(可以是自己感兴趣的, 也可以是附录二提供的某个主题中的一篇, 各种数学建模竞赛题除外), 撰写一篇小论文, 要求给出问题的背景、问题的数学描述、解决方法和计算结果。**计算结果部分必须**是自己仿真的结果, 具体可以是自己编程求解的, 也可以是自己利用优化软件求解的; 另外必须要有参考文献及其引用, 并按照附录一的说明在文中恰当地引用参考文献(否则起评分为10分)。具体可参考阳志明的作业(见文件yangzm\_paper.pdf, yangzm\_simulation.pdf)。

如果是学生自己寻找的相关应用撰写的课程小论文, 请写200字左右的说明, 包括题目的来源和难点、自己的收获等。

#### 附录一: 参考文献及其引用说明

参考文献按照其在正文中出现的先后以阿拉伯数字连续编码, 序号置于方括号内。一种文献被反复引用者, 在正文中用同一序号标示。一般来说, 引用一次的文献的页码(或页码范围)在参考文献中列出。格式为“著作的名称, 作者, 出版年”或者“论文名, 作者, 期刊名, 年卷(期)”+“: 页码(或者页码范围)”。需要仔细核对顺序编码下的参考文献序号, 做到序号与其所指示的文献同文后参考文献列表一致。另外, 参考文献页码或页码范围也要准确无误。

#### 附录二: 备选问题

##### 一、统计、信号/图像处理和压缩感知

文献[1]–[3]是稀疏恢复在三个典型应用领域的经典文献, 分别是在信号处理领域的图像去噪[1], 统计领域的回归分析的特征选择[2]和逼近论领域的过完备化表示中的应用。文献[4]–[6]设计了求解该问题的梯度算法, 可以求解大规模问题, 其中[6]是一种典型的加速梯度法。

##### 二、计算机通信网

优化在计算机通信网领域有重要而深刻的应用。文献[7]–[11]是优化理论, 尤其是对偶理论和最优解的刻画, 在通信网的速率控制方面应用的典型范例, 精读其中的任何一篇相信都会收获颇丰。

##### 三、多速率信号处理

余弦调制的有限脉冲响应滤波器组在数据压缩(语音、音频、图像和视频)、消噪、特征检测和抽取中有广泛应用。它们都是  $M$  通道最大抽取滤波器组的特殊子类, 其中设计余弦调制滤波器组(CMFB)问题最后归结为设计原型滤波器, 而原型滤波器的设计可以建模为优化问题。文献[12]是一篇采用半定规划松弛设计原型滤波器的典型方法。如果能掌握该方法, 并学会使用Sedumi软件, 将会为以后使用锥优化模型解决问题打下坚实的基础。

##### 四、信息论

经典论文[13]中的信道容量问题是一个典型的等式泛函优化问题(特别是VII.

The capacity of a channel in the presence of white thermal noise, 和 IX. Arbitrary Gaussian noise), 课本附录B.1.2 KKT条件的应用实例一节中的功率分配问题(B.1.4)中的容量分配恰好是这个问题的离散化版本。结合该论文[13]与课本附录的相关内容, 读懂材料, 理解连续和离散问题的联系, 如果可以, 请给出仿真结果, 之后总结成一篇报告。该选题对于泛函优化与数学规划的关系及理解最优解的刻画很有帮助。

## 五、组合优化

关于最大割问题而言, 经典论文[14]提出了目前近似率最好的一个逼近算法。读懂论文, 仿出结果, 之后总结成文。该选题对于了解利用凸优化作为工具研究组合问题(如最大割和2-Sat问题等)很有启发。

## 六、鲁棒的Parks–McClellan 等波纹滤波器设计

Parks–McClellan 算法是一个迭代算法, 用以设计最优的有限脉冲响应滤波器, 由 James McClellan 和 Thomas Parks 于 1972 年的著作[15,16]中提出。

此算法的主要精神, 在于利用迭代的方式最小化滤波器在通带(pass band)和阻带(stop band)的最大误差, 因此有时也称为最小化最大误差算法(Mini-max filter design)。由于 Parks–McClellan 算法也可看作 Remez 交换算法在设计有限脉冲响应滤波器时的一种变形, 因此也有人称其为 Remez 交换算法。

阅读文献[15,16], 理解Parks–McClellan算法, 参照文献[17]自行实现算法。总结问题描述、算法原理和步骤, 给出仿真结果, 并对结果进行分析。

## 七、大规模机器学习的最优化方法

随机梯度法[18]是机器学习中非常重要的训练算法, 对于机器学习中出现的大规模训练问题, 随机梯度法又演化出迷你批量法(mini-batch)及其各种各样的用以减小梯度中噪声的一阶方法[19], 为了提高随机梯度法的收敛速度, 研究者也尝试了各种的二阶方法[21]。[22]也讨论了深度学习中的非凸最优化问题的算法设计和目前的研究工作。借助文献[19]和[22], 请描述用于分类的Logistic回归模型, 并就文本分类数据RCV1[20]测试随机梯度法、mini-batch方法和LBFGS方法的效率。

## 参考文献

- [1] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” Phys. D, vol. 60, no. 1-4, pp. 259–268, 1992. (引用次数: 3667)}
- [2] R. Tibshirani, “Regression shrinkage and selection via the lasso,” J. Roy. Stat. Soc. Ser. B, vol. 58, no. 1, pp. 267-288, 1996.(引用次数: 4236)}
- [3] S. Chen, D. Donoho, and M. Saunders, “Atomic decomposition by basis pursuit,” SIAM J. Sci. Comput., vol. 20, no. 1, pp. 33-61, 1998. (引用次数: 2203)
- [4] Figueiredo, R. Nowak, and S. Wright, “Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems,” IEEE Journal of Selected Topics in Signal Processing, vol.1, no.4, pp.586-597, December 2007. (引用次数: 549)
- [5] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, Sparse reconstruction by separable

- approximation, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008), 2008, pp. 3373–3376; IEEE Transactions on Signal processing, 2009, 57(7):2479-2493.
- [6] A. Beck, M. Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems, SIAM J. IMAGING SCIENCES, 2009, 2(1):183–202
  - [7] F. Kelly, A. Maulloo and D. Tan. Rate control in communication networks: Shadow prices, proportional fairness and stability. J. Oper. Res. Soc., vol. 49, pp. 237-252, 1998.
  - [8] M. Chiang, S. H. Low, A. R. Calderbank, J. C. Doyle. Layering as optimization decomposition: a mathematical theory of network architectures. Proceedings of the IEEE, 2007, 95(1):255-312.
  - [9] F. Kelly. The Mathematics of Traffic in Network. Princeton University Press, 2005.
  - [10] S. Low, L. Peterson and L. Wang. Understanding Vegas: a duality model. Journal of the ACM, 2002, 49(2):207-235.
  - [11] S. Low. A duality model of TCP and queue management algorithms. IEEE/ACM Trans. Networking, 2003, 11(4):525-536.
  - [12] Ha Hoang Kha, Hoang Duong Tuan, Truong Q. Nguyen. Efficient Design of Cosine-Modulated Filter Banks via Convex Optimization. IEEE Trans. Signal Processing, vol.. 57, no. 3, MARCH , 2009.
  - [13] Claude E. Shannon, Communication in the Presence of Noise, Proceedings of the IEEE, vol. 86, no. 2, Feb. 1998.
  - [14] Goemans, Michel X.; Williamson, David P. (1995), "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming", *Journal of the ACM* **42** (6): 1115–1145, doi:10.1145/227683.227684.
  - [15] T. W. Parks and J. H. McClellan, “Chebyshev approximation for nonrecursive digital filters with linear phase,” IEEE Trans. Circuit Theory, vol. CT-19, pp. 189–194, Mar. 1972.
  - [16] T. W. Parks and J. H. McClellan, “A program for the design of linear phase finite impulse response filters,” IEEE Trans. Audio- Electroacoustics, vol. AU-20, no. 3, pp. 195–199, Aug. 1972.
  - [17] [Matlab code](#); Matlab code for "Equiripple FIR filter design by the FFT algorithm" by A. Enis Cetin, O. N. Gerek and Y. Yardimci, IEEE Signal Processing Magazine, 1997.
  - [18] H. Robbins and S. Monro, A Stochastic Approximation Method, The Annals of Mathematical Statistics, 1951, vo.22, pp.400-407.
  - [19] L. Bottou, F. E. Curtis, J. Nocedal, Optimization Methods for large-Scale Machine Learning, <http://arxiv.org/abs/1606.04838> Submit to SIAM Review.
  - [20] D. D. Lewis, Y. Yang, T. G. Rose, and F. Li, RCV1: A New Benchmark Collection of Text Categorization Research, Journal of Machine Learning Research, 2004, vo.5, pp. 361-397.
  - [21] R. H. Byrd, G. M. Chin, J. Nocedal, F. Oztoprak A Family of Second-Order Methods for Convex  $l_1$ -Regularized Optimization, Mathematical Programming, 2016, vo. 159, no. 1, pp. 435-467.
  - [22] Frank E. Curtis\_ Katya Scheinberg, Optimization Methods for Supervised Machine Learning: From Linear Models to Deep Learning, <https://arxiv.org/abs/1706.10207>