A-H 八个两页短篇的简介 导读和相关参考文献

何 炳 生 maths.nju.edu.cn/~hebma

我心目中的最优化理论与方法, 是最接地气的应用数学。在这个领域里我偏重算法研究, 主要兴趣是单调变分不等式的投影收缩 (Projection and Contraction) 算法和凸优化的 (ADMM类) 分裂收缩 (Splitting and contraction) 算法, 组织发表了 (包括一些浅陋之作的) 百多篇文章。这些方法得到了学术界的一些好评, 也被工程界用来成功地解决了一些他们的问题。我曾经写过两页的介绍, 企图概括自己一生的工作, 但显得过于简单, 不尽如人意。退休前后我还做过总结, 用 PPT 写成二十讲的系列讲义, 挂在个人主页, 但篇幅较长, 通读要花不少时间。我若精力允许, 也应该调整充实内容, 修订更新。转眼 10 多年过去, 时而还有年轻朋友向我咨询学术问题。随着年龄的增长, 我常自问, 一生中自己最看重的工作是哪些? 能否从自己零零散散的收获, 以及七旬以后的工作进展中, 选一些每个(包括主要推导过程)都能用 A4 两页讲清的东西, 介绍给那些想对我的工作有所了解的年轻人。写成几篇以后, 请在读博士生看看, 回馈说看懂基本没有问题。目前完成的这 **A–H** 八单元的 16 篇, 对我自己而言, 犹如老人自制的文玩小把件, 每件都能勾起我一段美好的回忆。读懂这些材料所需的预备知识, 是中学的数理基础和普通的大学数学, 同时需要认可一般的优化原理。只是对这 15 年多来热门的 ADMM 类分裂收缩算法感兴趣的学者, 也可以直接从 **C** 单元开始读起。最后的 **G** 和 **H** 两单元, 说明我古稀之后没有马上洗手是对的。

为展示研究路径, 也给出主要参考文献, 但读懂这些短篇, 阅读所列文献并非必需的。

A-从变分不等式的投影收缩算法到凸优化的分裂收缩算法概论. 2013年, 我从南京大学退休, 试图用两页纸概括总结一下职业生涯的科研成果, 戏称要为自己写一篇 "墓志铭", 于是就有了这个短篇 **A**。说起来, 我从单调变分不等式的投影收缩算法做起, 到 (包括ADMM的) 线性约束凸优化的分裂收缩算法, 可以说是一条主线, 一个模式, 作为概论, 短篇 **A** 主要是把这两类算法的关系说清楚。如需进一步了解投影收缩算法的成果, 建议阅读 [1,2]。对于如何从单调变分不等式的投影收缩算法过渡到线性约束凸优化的分裂收缩算法的, 我在论文 [4,5]中做了说明。短篇 **A** 的第二节已经包含了变分不等式意义下凸优化分裂收缩算法的统一框架, 以及方法收敛的关键不等式。这个统一框架, 在 [3–5] 中都有提及。

作为概论的两页 **A**, 大概只适合曾经关注过我工作的学生和朋友有兴趣时重新拾起。后面 **B** 和 **C** 两单元中的短篇, 分别介绍并证明变分不等式的投影收缩算法和热门的 ADMM 方法的主要收敛性质。从单元 **D** 到 单元 **H**, 阐述我们在 **A** 篇 §2 中给出的统一框架思想指导下取得的成果。

B-变分不等式投影收缩算法中的孪生方向和姊妹方法 在单调变分不等式的投影收缩算法中, 我们根据投影得到一对孪生方向建立的姊妹方法, 被一些工程界学者用来解决了长期困扰他们的问题。姊妹方法中用孪生方向和相同步长, 内涵的数学之美常常令我自己陶醉不已! 回头看, 虽然发现这些关系的过程有点漫长, 但其中用到的数学就是投影的基本性质, 犹如平面几何中的余弦定理。用两页纸写成的短篇 **B**, 详细介绍了投影收缩算法中孪生方向的由来, 也给出了姊妹方法关键收敛性质的证明。如果有进一步了解这方面工作的需求, 建议阅读论文 [1,2,4,6], 也可以阅读我系列讲义 [7] 的 1-5 讲。论文 [1] 和 [4] 中提到了三个基本不等式并指出孪生方向也可以从这些基本不等式的不同组合导出。一般情形下求解单调变分不等式的预测－校正框架和一些可供参考的算例可见 [8,9]。将这些想法推广到凸优化问题的求解上, 文献 [10]做了一些探讨。

C-利用变分不等式证明乘子交替方向法 (ADMM) 的关键收敛性质. 我们从 1997 年开始在变分不等式投影收缩算法的基础上开展 ADMM 研究 [11], 直到 2006 年, 我们研究的目标主要针对管理科学中的一些结构型变分不等式, 并在自调比准则 [12]、求解实现 [13,14]等方面做了些工作。2009 年, 我们才发现一大批 ADMM 的最新应用, 适逢第一届华东地区运筹学与控制论博士生学术论坛在上海大学召开, 我在大会报告中以 "信息技术中的凸优化问

题及交替方向法求解”为题呼请年轻人注意。2010 年 Boyd 他们关于 ADMM 的综述文章 [**?**], 让 ADMM 的研究和应用在全世界得到了令人瞩目的重视。受此鼓舞的我们, 凭借已有的基础, 顺势而为, 迅速投入, 做了些包括 ADMM 收敛速率 [15,16], 不精确 ADMM [17] 和一系列方法拓展方面的工作。虽然我们最初是在 Gowinski [18] 工作的基础上研究 ADMM, 早期的论文也采用他们的步长法则 [11,19], 为了让初学者更好地理解 ADMM 的收敛性质, 这里仅对工程界普遍采用的 ADMM 最简版本做论证。**C** 用变分不等式为框架写的 ADMM 关键收敛性质的简短而自洽的证明, 初学者容易入门。此后的所有单元, 讲的都是与 ADMM 研究相关的我们的研究成果。

D-凸优化问题对应的变分不等式和因需定制的PPA算法 短篇 **D** 首先介绍了简单约束凸优化问题的最优性条件之间犹如积分和微分之间的关系。**min-max** 问题的解点和线性约束凸优化问题 Lagrange 函数的鞍点, 都可以等价转换成一个(混合)变分不等式 (VI) 的解点。求解这类变分不等式, 我们提出了一类通过构造适当的分块正定矩阵, 分裂求解优化子问题实现的邻近点算法 [21], 并称其为因需定制的邻近点算法 (Customized PPA) [22,23]。用两页纸写下的这个短篇 **D**, 包含了对一些典型问题如何构造 Customized PPA 的例子。利用 PPA 中正定矩阵的分块结构, 使迭代得以用分裂的形式实现, 是这些方法被称为分裂算法的原因。关于求解变分不等式的正定矩阵模下的邻近点算法, 进一步的资料可以参阅 [3,24]和 [37]。通过合理设计PPA中的正定矩阵, 构造一些分摊子问题难度的均困方法, 称为 Balanced 方法。这些工作可以在 [25,26]和 [39]中找到。

E-效率更高的求解两个可分离块凸优化问题的ADMM类方法 在短篇 **C** 里我们用两页纸给出了 ADMM 的收敛性证明。在常用的 ADMM 的基础上, 做一些简单修改 (包括迭代中改变变量更新顺序的 PPA 型 ADMM 算法 [27]和两次修正乘子的对称 ADMM 算法 [28])。这些想法比较自然的算法, 在几乎不增加工作量的情况下, 效率一般都有30%以上的提高。求解两个可分离块的 ADMM 类方法有很多新的发展, 但想法自然又实现简单, 效率还有明显提高的是这两条。短篇 **E** 介绍了这两个算法, 并用在短篇 **A** 的§2 中已经提及的统一框架下验证了收敛性。建议中文读者参考 [5] 。关于统一框架算法的两条简单而又重要的性质, 我们会在短篇 **G** 和 **H** 中给予系统的证明。

F-求解三个可分离块凸优化问题的 ADMM 类分裂收缩算法 ADMM 是求解两个可分离块的凸优化问题的有效方法, 但直接推广到用来求解三个可分离块的凸优化问题是并不能保证收敛的 [29]。在这个事实得到明确证明之前, 根据社会生活中的一些机理, 我们提出了两个修正的 ADMM 类算法 [30,31], 并且都在统一框架下验证了算法的收敛性。我们将算法及其收敛性证明写成短篇 **F**。读者如需理解这些算法的机理, 建议阅读文献 [32]。处理三个可分离块问题最优正则化的文章 [33], 是处理两个可分离块最优线性化ADMM [20] 同类思想的推广, 也非常值得介绍, 有兴趣的读者可以参阅相关文献。这个短篇的内容, 也已经包含在文献 [5]里边。

G-从好不容易凑出一个方法到并不费劲构造一簇算法 线性约束的凸优化问题的拉格朗日函数的鞍点, 都可以转换成一个单调变分不等式的解点。为求解这类变分不等式, 十多年前, 我们就提出了一个预测-校正的统一框架 [3–5], 并给出了这个算法框架下保证收敛的条件。这对验证算法收敛性带来极大的方便, 有时我们也利用这个框架的收敛性条件去凑一些方法 [34,35]。收敛关键性质的证明, 用到一个“积化和差”的恒等式, 帮了我们大忙。这样的证明技巧, 也常常被学者们借用。这些给人一种凑得很巧的神秘感觉, 直到 2022 年, 我们才对算法框架的收敛性条件给出了一个等价表示, 据此便可以并不费劲地构造一簇算法 [36–38]。对这类构造性方法, 包括统一的收敛性证明, 也用两页纸写成了短篇 **G**。这样的成果没有在晚年的我手边溜走, 是值得庆幸的!

H-求解线性约束凸优化问题的预测-校正型广义邻近点算法 对由凸优化转换得来的变分不等式, 短篇 **D** 中已经介绍了因需定制的 PPA 算法。在短篇 **G**中, 我们根据预测-校正统一框架的收敛性条件的等价表示, 对确定的合格预测, 可以选择不同的校正矩阵, 从而并不费劲地构造一簇算法。这里, 在短篇 **G** 的基础上, 采用由预测唯一确定的特殊的校正矩阵, 得到预

测-校正的广义邻近点算法 [38], 具有短篇 **D** 中经典 PPA 算法的所有漂亮性质。我们将个结果及其证明用两页纸写成短篇 **H**。

结语. 利用变分不等式和邻近点算法这两大法宝, 才有了我们自成体系又颇有特色的工作。几位初步了解我们工作的欧美学者, 称赞我们的方法“Very Simple yet Powerful”和“Elegant”, 指的是我们在 VI 框架下构造了 PPA 算法。事实上, 对线性约束的凸优化问题, 我们的统一框架不仅涵盖了 PPA, ALM 和 ADMM 等耳熟能详的基础算法, 还为这些方法无法胜任求解的 (三块和多于三块的) 可分离凸优化问题提供了系列求解方法。有朋友说我用的主要数学工具, 本质上就是中学数学的余弦定理, 我也觉得大差不离。要说我在学术生涯中主导 (和带领学生) 做了点什么, 我有一本专著 [39] 即将出版, 但关键思想多半已经在这八单元16个短篇里边。对希望了解我这一生究竟做了些什么的数学界朋友, 浏览一下我的这些“文玩小把件”, 大体上也就可以! 穷其一生, 我主要也就做成了这些东西。

何炳生　2024 年 7 月

References

[1] B.S. He, A class of projection and contraction methods for monotone variational inequalities, Applied Mathematics and Optimization (1997) 35: 69-76.

[2] B.S. He and L. Z. Liao, Improvements of some projection methods for monotone nonlinear variational inequalities, JOTA (2002) 112: 111-128.

[3] B. S. He, PPA-like contraction methods for convex optimization: a framework using variational inequality approach, J. Oper. Res. Soc. China (2015) 3: 391-420.

[4] 何炳生, 从变分不等式的统一收缩算法到凸优化的分裂收缩算法, 高等学校计算数学学报(2016) 38: 74–96.

[5] 何炳生, 我和乘子交替方向法20 年. 运筹学学报(2018) 22: 1–31.

[6] B.S. He, X.M. Yuan and J.J.Z. Zhang, Comparison of two kinds of prediction-correction methods for monotone variational inequalities, COA (2004) 27: 247-267.

[7] 何炳生. 凸优化和单调变分不等式的收缩算法, 见 maths.nju.edu.cn/~hebma 中讲义.

[8] B.S. He, L.Z. Liao and X. Wang, Proximal-like contraction methods for monotone variational inequalities in a unified framework I: Effective quadruplet and primary methods, COA (2012) 51: 649-679.

[9] B.S. He, L.Z. Liao, and X. Wang, Proximal-like contraction methods for monotone variational inequalities in a unified framework II: General methods and numerical experiments, COA (2012) 51: 681-708.

[10] 何炳生, 凸优化和单调变分不等式收缩算法的统一框架. 中国科学: 数学 2018 年第 48 卷第 2 期: 255－272.

[11] B.S. He and H. Yang, Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities, Operations Research Letters (1998) 23: 151－161.

[12] B.S. He, H. Yang, and S.L. Wang, Alternating directions method with self- adaptive penalty parameters for monotone variational inequalities JOTA (2000) 106: 349-368.

[13] B.S. He, L.Z. Liao, D.R. Han, and H. Yang, A new inexact alternating directions method for monotone variational inequalities, Math. Program (2002) 92: 103–118.

[14] B.S. He, L.Z. Liao and M.J. Qian, Alternating projection based prediction-correction method for structured variational inequalities, JCM (2006) 24: 693-710.

[15] B.S. He, X.M. Yuan, On the O(1/n) convergence rate of the Douglas-Rachford Alternating Direction Method, SIAM J. Numerical Analysis (2012) 50: 700-709.

[16] B.S. He and X.M. Yuan, On non-ergodic convergence rate of Douglas-Rachford alternating directions method of multipliers, Numerische Mathematik (2015) 130: 567-577.

[17] G.Y. Gu, B. S. He and J. F. Yang, Inexact alternating-direction-based contraction methods for separable linearly constrained convex optimization. JOTA (2014) 163: 105 – 129.

[18] Glowinski R. *Numerical methods for nonlinear variational problems*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

[19] M.H. Xu, Proximal alternating directions method for structured variational inequalities, JOTA (2007) 134: 107-117。

[20] B.S. He, F. Ma and X.M. Yuan, Optimally linearizing the alternating direction method of multipliers for convex programming, COA (2020) 75: 361-388.

[21] B.S. He and X.M. Yuan, Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective. SIAM J. Imaging Science (2012) 5: 119-149.

[22] B.S. He, X.M. Yuan and W.X. Zhang, A customized proximal point algorithm for convex minimization with linear constraints, COA (2013) 56: 559-572.

[23] G.Y. Gu, B.S. He and X.M. Yuan, Customized proximal point algorithms for linearly constrained convex minimization and saddle-point problems: a unified approach, COA (2014) 59: 135-161

[24] 何炳生, 申远, 求解凸规划及鞍点问题定制的PPA算法及其收敛速率, 中国科学: 数学2012年第42卷第5期: 515–525.

[25] B.S. He and X.M. Yuan, Balanced Augmented Lagrangian Method for Convex Optimization. manusript, 2021. arXiv:2108.08554

[26] S. J. Xu, A dual-primal balanced augmented Lagrangian method for lineraly constrained convex programming, J. Appl. Math. and Computing (2023) 69: 1015-1035

[27] X.J. Cai, G.Y. Gu, B.S. He and X.M. Yuan, A proximal point algorithms revisit on the alternating direction method of multipliers, Science China Mathematics (2013) 56: 2179-2186.

[28] B.S. He, H. Liu, Z.R. Wang and X. M. Yuan, A strictly Peaceman-Rachford splitting method for convex programming, SIAM J. Optim. (2014) 24: 1011-1040.

[29] C.H. Chen, B.S. He, Y.Y. Ye and X. M. Yuan, The direct extension of ADMM for multi-block convex minimization problems is not necessary convergent, Math. Program. (2016) 155: 57-79.

[30] B.S. He, M. Tao and X.M. Yuan, Alternating Direction Method with Gaussian Back Substitution for Separable Convex Programming, SIAM J. Optim. (2012) 22: 313-340.

[31] B.S. He, M. Tao and X.M. Yuan, A splitting method for separable convex programming, IMA J. Numerical Analysis, (2015) 31: 394-426.

[32] 何炳生, 修正乘子交替方向法求解三个可分离算子的凸优化. 运筹学学报, 2015, 19(3): 57–70.

[33] B.S. He and X.M. Yuan, On the optimal proximal parameter of an ADMM-like splitting method for separable convex programming. Mathematical methods in image processing and inverse problems, 139-163, Springer Proc. Math. Stat., 360. Springer, Singapore, 2021.

[34] B.S. He and X. M. Yuan, A class of ADMM-based algorithms for three-block separable convex programming. COA (2018) 70:, 791-826.

[35] B.S. He, S.J. Xu and X.M. Yuan, Extensions of ADMM for separable convex optimization problems with linear equality or inequality constraints, Handbook of Numerical Analysis, 24 (2023) 511-557. arXiv:2107.01897v2[math.OC].

[36] B.S He and X.M Yuan, On construction of splitting contraction algorithms in a prediction- correction framework for separable convex optimization, arXiv 2204.11522 [math.OC]

[37] 何炳生, 利用统一框架设计凸优化的分裂收缩算法, 高等学校计算数学学报, 2022, 44: 1-35.

[38] 何炳生, 凸优化分裂收缩算法统一框架的新进展–从好不容易凑出一个方法到并不费劲构造一族算法, 高等学校计算数学学报, 2024, 46: 1-24.

[39] 何炳生, 变分不等式框架下凸优化的分裂收缩算法. 科学出版社 2025 待出版.