说说我的主要研究兴趣

—兼谈华罗庚推广优选法对我的影响

何炳生

南京大学数学系, 南京, 210093.

变分不等式求解是最优化方法的一个重要分支。作为描述平衡问题的核心工具, 变分不等式在工程管理、宏观调控等领域具有广阔的应用背景。在单调变分不等式的求解方法中, 我主要对"只用函数值的方法"感兴趣, 并致力于"少用函数值的方法" 之研究。本文叙述了华罗庚推广优选法对我的影响、说明研究工作和个人经历的关系、以及我继续这方面的一些相关研究的理由。对想在最优化方法方面研读的学生, 提了一些建议。

关键词. 优选法, 变分不等式, 单调性, 基于信息的计算

1 我为什么选择最优化方法作为专业方向

我的研究方向是最优化方法, 或者叫做数学规划中的数值方法, 它属于计算数学与运筹学的交叉学科。1966 年我高中毕业回乡务农, 期间"推广优选法小分队"在我家乡县城的一个普及报告, 是我第一次接触"优化"。1977 年恢复高考, 我被南京大学录取。华罗庚教授 1980 年在南京大学礼堂关于"优选法"和普及数学应用的报告以及南京大学何旭初先生在当年国内优化界的学术地位, 对我报考研究生时选择"最优化"作为专业方向产生了很大的影响。和多数做优化的同行一样, 我们最初是奔"应用"来的, 由于我上大学之前就劳动了 10 多年, 年岁大了就特别想学一些"立杆见影"有用的东西。等到念完博士, 当了教授, 能够做一些"研究"时, 我已经或多或少地偏移了当初学习"最优化"的初衷, 不再细究实际应用背景。在熟人眼里, 我当了 10 多年农民以后, 又上了大学, 当了教授, 好像也是实现了人生价值。然而面对学生提出的诸如"应用背景"这样一类问题, 我常常为不能给学生比较简明的答复而感到困惑。后来我对互补问题(变分不等式的特殊情形)有了一定的了解, 发现对这类问题比较容易举例说明它们的应用背景。因此, 我的主要研究兴趣就逐步放到了最优化方法的一个分支—变分不等式的求解上面。

2 什么是变分不等式

设 Ω 是 R^n 中的一个闭凸集, F 是从 Ω 到 R^n 的一个映射。变分不等式是求

$$x \in \Omega, \qquad (x' - x)^T F(x) \geq 0, \qquad \forall\, x' \in \Omega. \tag{1}$$

Ω 和 F 是变分不等式中的二大要素, 我们记它为 $\mathrm{VI}(\Omega, F)$。在变分不等式 $\mathrm{VI}(\Omega, F)$ 中, 如果 Ω 是一个正卦限, 它就等价为一个互补问题 (Complementarity Problem):

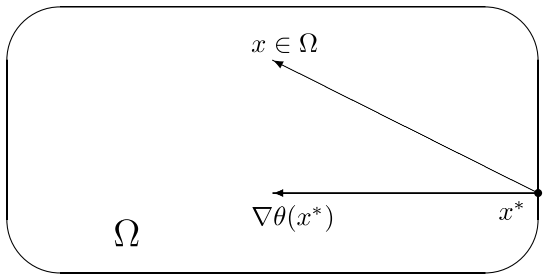
$$x \geq 0, \qquad F(x) \geq 0, \qquad x^T F(x) = 0. \tag{2}$$

变分不等式是包含了一些优化问题的数学问题。设 $\theta(x)$ 是从 R^n 到 R 的可微凸函数, 什么样的点是优化问题

$$\min \{\, \theta(x) \mid x \in \Omega \} \tag{3}$$

的最优点呢？从几何上来看，如果某一点 x^* 是问题 (3) 的解，它必须属于 Ω，并且从这点出发的所有可行方向都不是下降方向。换句话说，可行方向集与下降方向集的交必须是空集。用 $\nabla\theta(x)$ 表示 $\theta(x)$ 的梯度. 我们记

- $S_d(x) = \{s \in R^n \mid s^T \nabla\theta(x) < 0\}$，为点 x 处的下降方向集；
- $S_f(x) = \{s \in R^n \mid s = x' - x,\ x' \in \Omega\}$，为点 x 处的可行方向集.



根据我们对 $S_f(x)$ 和 $S_d(x)$ 的定义，x^* 是最优解必须有 $x^* \in \Omega$ 和 $S_f(x^*) \cap S_d(x^*) = \emptyset$，这句话的等价数学形式就是 x^* 是问题

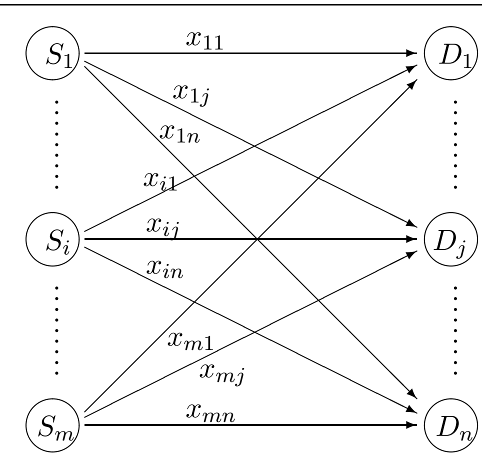
$$x \in \Omega, \quad (x' - x)^T \nabla\theta(x) \geq 0, \quad \forall\ x' \in \Omega \tag{4}$$

的解. 将 (4) 中的 $\nabla\theta(x)$ 看作 (1) 中的 $F(x)$，可微凸优化问题就是一个特殊的单调变分不等式. 说它特殊，是因为当 $\theta(x)$ 二次可微时，它的 Hessian 矩阵 $\nabla^2\theta(x)$ 一定是一个对称矩阵. 而在变分不等式 VI(Ω, F) 中，如果映射 F 可微，我们并不要求它的 Jacobian 矩阵 $\nabla F(x)$ 对称.

3 为什么说变分不等式有较多的应用背景

变分不等式问题，作为描述平衡问题的核心工具，可以用来解释工程管理、经济平衡及宏观调控等领域的许多问题[4]。我们以宏观调控中的保护资源-保障供给问题以及交通流量控制问题为例说明变分不等式容易找到应用背景。

宏观调控问题. 假设某种不可再生资源(例如煤) 由 m 个资源地生产和 n 个需求地消费。如图所示，它由经营者们从资源地采购运到需求地销售。经营者会根据贪婪原理决定自己的经营方案，他们不会自觉顾及生态效益和社会效益。职能部门要采取经济手段来保护资源和保障供给。具体说来，在过度开采的资源地 S_i 通过对经营者征收资源税 y_i 来保护资源，对供货不足的需求地 D_j 通过给经营者价格补贴 z_j 而吸引经营者增加供货。对于政府给出的税收和补贴政策，经营者们又会找到相应的新的经营方案。我们的任务则是用数学方法帮助职能部门找到最优税收和最优补贴。要求:



x_{ij}： 从 S_i 到 D_j 的交易量;
s_i： 在 S_i 的总采购量, $s_i = \sum_{j=1}^n x_{ij}$;
d_j: 在 D_j 的总销售量, $d_j = \sum_{i=1}^m x_{ij}$。

- 通过征收资源税保护资源, 使 $s \le s^{\max}$, 同时又让经济尽可能繁荣;
- 通过财政补贴保障供给, 使 $d \ge d^{\min}$, 同时又尽可能减少财政支出。

换句话说, 职能部门找到最优税收向量 $y^* \in \mathbb{R}^m_+$ 和最优补贴向量 $z^* \in \mathbb{R}^n_+$, 以及经营者们由此产生的 $s(y^*, z^*)$ 和 $d(y^*, z^*)$ 应该满足

$$y^* \ge 0, \quad s^{\max} - s(y^*, z^*) \ge 0, \quad {y^*}^T(s^{\max} - s(y^*, z^*)) = 0,$$

和

$$z^* \ge 0, \quad d(y^*, z^*) - d^{\min} \ge 0, \quad {z^*}^T(d(y^*, z^*) - d^{\min}) = 0.$$

记

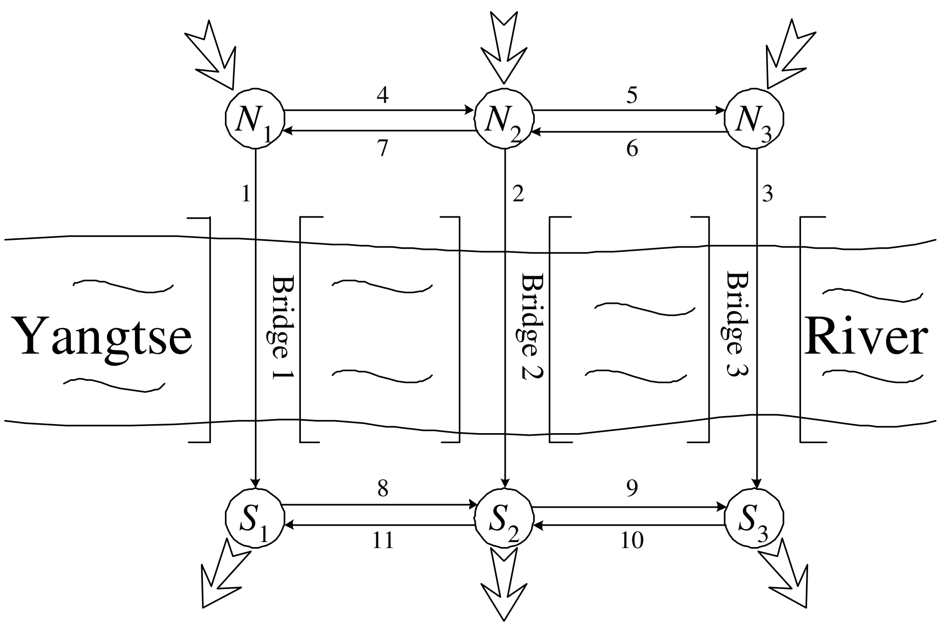
$$u = \begin{pmatrix} y \\ z \end{pmatrix} \quad \text{和} \quad F(u) = \begin{pmatrix} s^{\max} - s(u) \\ d(u) - d^{\min} \end{pmatrix}.$$

通过经济手段实现宏观调控中的保护资源-保障供给问题就归结为求解以下的互补问题

$$u \ge 0, \quad F(u) \ge 0, \quad u^T F(u) = 0.$$

互补问题是约束集合 Ω 为正卦限的变分不等式。

大桥流量调控问题. 随着交通事业的发展, 已有过江大桥的沿江城市又纷纷建起了二桥, 三桥。社会公益要求合理使用交通网络, 避免交通拥挤。我们考察有三座长江大桥的某市由北向南的一个过江运输网络模型. 这个网络图由 6 个点和 11 条有向边组成。



不失一般性, 我们可以将 N_1, N_2 和 N_3 看作由北向南的车辆在江北的出发地, 同时将 S_1, S_2 和 S_3 看作它们在江南的集散地. 担负确定运输任务的车辆可以经过不同的桥从江北到江南. 交通管理部门希望通过征收合理的过桥费, 使得每座桥上的流量在他们希望的控制范围内. 对给定的收费 $0 \le x \in R^3$, 桥上的流量是由驾驶员们根据经济原理所提供的, 是收费的函数, 可以写作 $f(x) \in R^3$. 我们设 $0 < b \in R^3$ 是管理部门希望控制的桥上的流量上界, 通过收费实现合理流量分配问题的数学模型便是求 $0 \le x \in R^3$, 使得

$$x_i \ge 0, \quad f_i(x) \le b_i, \quad x_i[f_i(x) - b_i] = 0, \quad i = 1, 2, 3.$$

如果我们记 $F(x) = b - f(x)$, 则上述问题可以表述成以下的互补问题

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0.$$

类似的问题还可以举出很多。例如, 政府对有限资源(例如水、电)的合理配置使用, 对不同行业实行不同的价格政策, 这样的问题也可以被描述成一个变分不等式。

4 我为什么研究基于信息的计算方法

当年我听“推广优选法小分队”的报告, 最深刻的感受是认识到实际生活中碰到的最优化问题中, 函数一般没有显式表达式。求解这些问题, 是要对给定的自变量, 观测相应的函数值(信息), 使用“只用函数值的方法”。由于信息的观测与获取往往是代价不菲的, 因此要尽可能少用函数值。华罗庚教授推广普及的优选法[17], 就是求解一维单峰函数极值问题的只用函数值且少用函数值的方法。在计算过程中只用并少用函数值的方法, 用现代的术语, 可以称之为基于信息的计算方法。或许是受华罗庚教授推广普及优选法的深刻影响, 或许是我“而立”之前 10 多年的务农经历所形成的一些狭隘的看法一直在起作用, 我对只用函数值的方法比较看重。

与实际应用有密切联系的, 特别是源自管理科学的变分不等式, 其函数也通常没有显式表达式。以前面提到的保护资源-保障供给问题为例, 对给定的资源地的税收量 $y \geq 0$ 和需求地的价格补贴量 $z \geq 0$, 经营者们根据贪婪原理找到相应的经营量 $x(y, z)$. 我们通常能够观察到的只是那些依赖于 y 和 z 的经营者们在资源地的采购量 $s(y, z)$ 和在需求地的供货量 $d(y, z)$. 同样, 在大桥流量调控的互补问题中, 流量 $f(x)$ 是收费 x 的函数, 这个函数肯定存在, 却没有显式表达式. 对给定的 $x \geq 0$, 我们能够观测到相应的流量 $f(x)$. 求解这些问题, 同样需要只用函数值且少用函数值的求解方法。针对各种不同结构单调变分不等式提供只用函数值且少用函数值的求解的方法一直是我的一个欲罢不能的研究课题[5, 8, 9, 10]。

近 10 年来, 最优化研究领域领袖人物, 英国剑桥大学 Powell 教授在多个场合的报告都是关于求解极小化问题的只用函数值(Minimization without Derivatives)的方法。并称之为求解多元极小化问题的最有用的方法(the most useful algorithm)。对 Powell 教授的报告的理解, 也成了我至今没有割舍“基于信息的求解变分不等式的方法”这一研究的一个原因。

李大潜院士最近多次呼吁大力提倡和推动以问题驱动的应用数学研究[18], 并将“既不是纯基础, 也不是纯应用, 但却是沟通基础与应用的重要的关键环节”作为对“以问题驱动的应用数学研究”的一种界定。与实际应用有密切联系的, 特别是源自管理科学的变分不等式, 其自变量通常不止一个, 函数也一般没有显式表达式。求解这样的变分不等式, 同样需要只用函数值的方法。读李大潜院士的报告, 掺和一些为我所用的理解, 我们将“基於信息”的求解方法当作一个“以问题驱动”的求解方法, 又多了些继续这方面研究的理由。

5 我为什么主要研究单调变分不等式的求解方法

上世纪八十年代以来, 变分不等式的研究得到很快的发展, 新的方法层出不穷[3]。这个时期, 也是“内点法”蓬勃发展的时期。内点法的“优秀品质”主要表现在处理凸优化问题上。在学习内点法的时候, 我领教了一些看视简单的经典非凸问题, 认识到非凸优化问题的特殊难度。

由于深知了非凸优化问题的难度以及“可微非凸优化等价于非单调变分不等式”这一事实, 我在参与变分不等式研究时, 只将变分不等式中比较简单的一类单调变分不等式的求解作为主

要研究课题。而描述现实生活中平衡问题的变分不等式中, 像降价会促进消费, 道路拥挤会增加通行时间, 其函数也确实具有单调的性质。

既然求解一维优化问题的(只用函数值的)黄金分割法也只能处理单峰问题, 研究变分不等式而只用函数值的求解方法时, 理智的做法应该将研究对象限定单调变分不等式上。由于受限于只用函数值, 我们很容易为这些算法人为地构造收敛很慢的算例。尽管如此, 这些算法在对上面提到的宏观调控、交通平衡中的一些代表性问题的算例中也有不俗的表现。

6 我为什么把这些方法称为投影收缩方法

我们研究的这些方法, 它们的要素与投影和收缩 (Projection and Contraction) 有关。到简单凸集的投影是方法要用到的一些基本运算。

虽然解集是未知的, 方法却保证迭代点到解集的距离越来越近, 这是算法被称为收缩算法的原因。投影收缩算法的定义及算法框架在 Springer 出版社1975 年出版的 Blum 和 Oettli 的德文专著[1] 中就说清楚了。Springer 是著名的出版社, 作者 E. Blum 和 W. Oettli 两位是苏黎世大学毕业的博士, 都已作古, Oettli 教授生前我还访问过他。我们现在研究的算法, 完全符合他们的专著中关于投影收缩算法的定义与算法框架, 所不同的只是构造了未知距离函数的不同下降方向。出于对历史的尊重和对前人工作的一种敬畏, 我们还是把这些方法称为投影收缩方法 (Projection and Contraction Methods)。

必须指出, Korperlevich-Khobotov [14, 15]的外梯度方法也是求解变分不等式的只用函数值的方法。在相应的算法中, 我们的主要贡献是选择了最优步长[9]。外梯度方法中改用我们的最优步长以后, 对前面提到的宏观调控问题的算例, 收敛速度提高 7 倍以上。在求解实际问题的只用函数值的算法中, 当函数值的获取代价不菲时, 节约80%以上的工作量, 是很有经济意义的。

7 我为什么又把这些方法称为预测校正方法

邻近点算法— 也称 PPA 算法(Proximal Point Algorithm)是求解变分不等式的一类经典方法, 它通过系列求解条件较好的同类子问题而得到原问题的解。方法在合适的假设下具有线性收敛性。精确形式的 PPA 算法每步要求解一个只是条件好一些的同样规模的变分不等式, 而方法只具有线性收敛性, 因此在计算实践上是不值得提倡的。

我们研究的大多数算法都可以看作从邻近点算法演化而来。我们将经典的邻近点算法中子问题求解的不精确准则放松到极致得到的点称作为“预测点”, 新的迭代点只利用“预测点”的函数值并通过投影就能轻而易举地“校正”得到。无论是预测公式还是校正公式, 都是可以通过一个简单的投影公式就可以实现的, 我们的主要工作在于选取合适的参数, 使得算法的线性收敛性仍然得到保证并且在计算实践中收敛得尽可能快些。这个线性收敛的方法仍有价值的原因是因为它每步都容易实现, 计算代价很低。

预测-校正, 是计算数学中的基本术语。将这些算法看作是基于邻近点算法的一种预测-校正方法 (Prediction-Correction Method) 是最贴切不过的。对一个总体线性收敛的算法降低子问题的求解代价并保持原有的线性收敛性是贯穿我们研究的主题思想和一贯做法。

8 我为什么对变分不等式的交替方向法感兴趣

结构型变分不等式的交替方向法源自数值优化中的增广 Lagrangian 乘子法(Augmented Lagrangian Methods)。增广 Lagrangian 乘子法和序列二次规划法 (SQP 方法) 是约束最优化方法中二个公认的比较成熟的方法(可见 Nocedal 和 Wright 的 Numerical Optimization [16] 最后二章)。变分不等式是与优化问题有密切联系的。实际问题中的算子往往是可分离的, 当算子是可分离的时候, 就有一系列的基于增广 Lagrangian 乘子法的分裂算法可以考虑。

结构型变分不等式可以看作一个 Leader 和若干个与之合作的 Followers 去共同求解一个系统平衡问题。Leader 的任务是指明航向和分解任务, Followers 是去实施, 做具体工作的。一次迭代中的子问题求解可以看作是 Followers 根据 Leader 给的数据指令去完成一项不太轻松的任务。一个好的 Leader 对与其合作的下级应该不是太苛刻, 在算法设计中的体现就是 Leader 对 Followers 求解子问题的精度要求不能太高, 并且能利用自己的学识与智慧, 发挥聪明才智, 通过对 Followers 反馈信息的加工处理, 为下一次迭代给出更合理的数据指令。在现实社会生活中, 这样的 Leader 才是受群众欢迎的。我们将社会生活中被大家都认可的理念应用到算法的研究与设计中[11]。基于这样的考虑, 我们对不精确准则特别宽松而又保持线性收敛的方法感兴趣。

9 我为什么还在这个领域做研究

眼下, 我的研究兴趣还在这个领域, 就像搞地质勘探, 我对这一块的地质情况比较熟悉, 说不定还能发现一些新的有用的东西。譬如说, 我们把求 $x \in R^n$, 使得

$$f(x) \in \Omega, \qquad (y - f(x))^T x \le 0, \qquad \forall\, y \in \Omega \tag{5}$$

这样的问题称为*逆变分不等式*, 记作 IVI(Ω, f). 它是由问题 (1) 中的自变量与映射互换位置并将 '$\ge$' 换成 '$\le$' 得到的。在逆变分不等式中, 我们也会遇到类似的结构型逆变分不等式 IVI(S, f), 求 $x \in R^n$, 使得

$$f(x) \in S(f), \quad (y - f(x))^T x \le 0, \quad \forall\, y \in S(f), \tag{6}$$

其中

$$S(f) = \{f \in R^n \mid A^T f \le (\text{or} =) b,\ f \in \mathcal{F}\}, \tag{7}$$

$A \in R^{n\times m}$, $b \in \mathbb{R}^m$, $\mathcal{F} \subset \mathbb{R}^n$ 是一个简单闭凸集.

我们考虑以分时电价实现电网'削峰填谷' 的问题. 假设将一天用电负荷的'峰'、'谷' 分成 n 个时段. 为保障电网安全和合理用电, 电力部门希望通过对不同时段采用不同的电价将用电负荷控制在 $\mathcal{F} = [c, d]$ 这个范围内, 其中 $c, d \in \mathbb{R}^n$. 设现行分段电价是 $x_0 \in \mathbb{R}^n$, 电网负荷是 $q(x_0) \in \mathbb{R}^n$. 可以假设电网负荷是电价的不增(或者递减)函数. 我们记电价的变动量为 $x \in \mathbb{R}^n$, 并记 $f(x) = q(x_0 + x)$. 在通过调控电价实现'削峰填谷' 的同时我们希望电价变动尽可能小. 电力部门要找的最优的分时段的电价调节量 x^*, 是形如 (5)的逆变分不等式

$$f(x) \in \mathcal{F}, \qquad (y - f(x))^T x \le 0, \qquad \forall\, y \in \mathcal{F}$$

的解. 在以分时电价实现电网'削峰填谷'这个问题中, 用电负荷 $f(x)$是功率, 我们将 A 看作一个 n 维的列向量, 它的每个分量为分时段的时长, $A^T f(x)$ 便是一天总用电量(千瓦时). 如果在分时

电价实现电网'削峰填谷' 这个问题中我们对每天的总用电量有所要求, 这样的问题便可以描述成一个形如 (6)-(7) 的结构型的逆变分不等式 IVI(S, f).

与变分不等式一样, 逆变分不等式问题大量出现在现实生活中. 同样, 现实生活中的逆变分不等式, 其函数一般没有显式表达式, 观测函数值的代价往往是不菲的。冯康先生生前曾反复强调: "物理问题可以有多种数学模型, 它们在理论上等价, 但在实践中未必等效。它们可能导致不同的数值计算方法, 有不同的计算效果"。虽然逆变分不等式可以转化为一个扩大了规模的变分不等式, 由于"理论上等价的, 在实践中未必等效", 针对不同结构的逆变分不等式问题, 还是有许多基于信息的计算方法值得研究的[12, 13]。

基于信息的计算方法, 是解决实际问题的需要。在优化研究的这支队伍中, 有少部分人坚持做这方面的研究, 应该是不失有意义的。

10 我为什么总要求学生不是只听我的一家之言

我 30 岁才上大学, 又念数学, 实在是与"扬长避短"相悖。在选择了最优化作为研究方向以后, 我以单调变分不等式求解中"只用函数值的方法"作为主要研究课题, 在"少用函数值的方法"上下功夫。前面已经说到我的这些做法既与我上大学之前的工作经历有关, 也与我刚接触高等数学时所听的(特别是华罗庚和他的推广优选法小分队的)一些报告有关。华罗庚博大精深, 或许推广优选法在他的数学成就与贡献中不占重要地位。然而, 在我长知识的重要阶段能够从媒体接触到的只是他"将数学理论研究和生产实践紧密结合"的那些工作, 以及"只用函数值"和"少用函数值"的方法。就我而言, 年轻时接受的一些观点一直影响着我的思维方式直至今日的最优化方法研究。

因此, 当我自己也能带研究生, 能在讲台上主张一些学术观点时, 就感到责任很大, 生怕学生被自己误导。越民义先生在他的 2006 年的一个报告中[19] 就说, 学生"很容易将老师所说作为数学的前沿作品来接受, 工作后又将这些东西作为衣钵往下传。若因循守旧, 则易造成谬种流传, 误人不浅"。因为总怕自己"点子少路子窄" 而贻误学生, 多年来我没有主动找学生读我的研究生。虽然我也向学生们介绍一些自己的想法与研究, 但总提醒学生不能一叶障目。因为学生"很容易" 将老师的工作当作前沿作品, 我给研究生的第一堂课总是以介绍国内外同行名家和他们的主要方向为主, 要求他们耳听八面眼观四方。

学生们常常问到什么是当前优化研究的热点或者主流, 我无法回答这个超出我水平的问题。一般认为, 上世纪 40-50 年代线性规划的单纯形算法曾经是热点; 60-70年代非线性规划的拟牛顿法成了热点; 80 年代开始则内点法成了研究热点。下一个热点是什么, 是还没有到, 还是到了我们没有觉察到? 可以提及一下的是 2006年国际数学家大会中的线性与非线性规划主要内容仍然是与内点法有关的凸优化。大会一小时报告中有一个最优化方法方面的报告, 报告人是 A. Nemirovski, 题目是 Advances in convex optimization: conic programming. 两个 45 分钟报告的题目分别为 Convex optimization of graph Laplacian eigenvalues 和 Designing telecommunication networks by integer programming. 2007 年的国际工业与应用数学会议中, 有二个最优化方法方面的报告, 一个是 A. Nemirovski 的 Recent trends in robust convex optimization, 另一个是 M. Todd 的 Minimum volume ellipsoids, applications, duality, and algorithms.

对于想在最优化方法方面研读的学生，我建议把 Boyd 和 Vandenberghe 的 Convex Optimization [2] 及 Nocedal 和 Wright 的 Numerical Optimization [16] 两本书先通读一下, 希望能在读书的过程中找到自己感兴趣的研究方向, 做出既有理论意义又有实用价值的工作。

References

[1] E. Blum and W. Oettli, Mathematische Optimierung, Econometrics and Operations Research XX, Springer Verlag, 1975.

[2] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[3] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Series in Operations Research, Spinger-Verlag, 2003.

[4] M.C. Ferris and J.S. Pang, Engineering and economic applications of complementarity problems, *SIAM Rev.*, 39, 669-713, 1997.

[5] B.S. He, *A new method for a class of linear variational inequalities*, Mathematical Programming **66**, 137-144, 1994.

[6] B.S. He, *Solving a class of linear projection equations*, Numerische Mathematik **68**, 71-80, 1994.

[7] B.S. He, A modified projection and contraction method for a class of linear complementarity problem, Journal of Computational Mathematics, 14: 54-63, 1996.

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

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

[10] B.S. He, X.M. Yuan and J.J.Z. Zhang, *Comparison of two kinds of prediction-correction methods for monotone variational inequalities*, Comput. Optim. and Appl. **27**, 247-267, 2004

[11] B.S. He and L-Z Liao and M.J. Qian, Alternating projection based prediction-correction methods for structured variational inequalities, Journal of Computational Mathematics, 24(6), 693-710, 2006.

[12] B.S. He, Henry X. Liu, X.L. Fu and X.Z. He, Solving a class of constrained 'black-box' inverse variational inequalities, Sept. 2006, Sciencepaper Online (http://www.paper.edu.cn/process/download.jsp?file=200609-103)

[13] B.S. He, Henry X. Liu, M. Li and X.Z. He, PPA-based methods for monotone inverse variational inequalities, June 2006. Sciencepaper Online (http://www.paper.edu.cn/process/download.jsp?file=200606-219)

[14] E. N. Khobotov, *Modification of the extragradient method for solving variational inequalities and certain optimization problems*, U.S.S.R. Comput. Math. Phys. **27** pp. 120-127, 1987.

[15] G. M. Korpelevich, *The Extragradient Method for Finding Saddle Points and Other Problems*, Ekonomika i Matematchskie Metody **12** 747-756, 1976.

[16] J. Nocedal and S.J. Wright, Numerical Optimization, Springer Verlag, 1999.

[17] 华罗庚, 优选学, 科学出版社, 1981.

[18] 李大潜, 关于大力提倡和推动以问题驱动的应用数学研究的建议, 在数理科学部第三次科学基金项目研究成果报告会上的发言, 2006.

[19] 越民义, 关于数学发展之我见. 中国应用数学研究所, 2006.