



数学建模热的冷思考： 谨防“知”多“识”少、 “学”多“习”少

浙江大学数学科学学院 陈叔平

2017.7.15

一、开场白：从今天的数学教育说起

- 从学校到学校、从书本到书本容易让一个人形成一种思维惯性：只要听课、看书就能够学会自己想学的东西，会做书上的习题和任课教师出的考题就可以证明自己学懂了。碰到新问题，普遍的反应是“这我没学过，我不会”，或者立马上网搜索，去找到现成的答案，缺少好奇的心态、探索的勇气和独立思考的精神。这样的学习实在算不上成功。事实上，人生遇到的事，大多“做的时候不会”，“要的时候没有”。

- 上述思维实际上是一种误区。须知，知识 = 知 + 识，学习 = 学 + 习，经验 = 经 + 验，……。 “识”是指明白，知道很多但都没搞明白是毫无意义的，比如“光盘”；“习”是指自己练，学而不习很难真正明白，很难真正掌握一样东西，比如游泳；经历过而不体验、缺感悟谈不上有收获，比如旅游。学生认识和思维的这种误区，学校和老师是有责任的。在今天，“用人类的全部知识武装自己”、“追求卓越”之类的说法太多，误导作用远大于励志作用。
- 学生认识和思维的这种误区，学校和老师是有责任的。在今天，“用人类的全部知识武装自己”、“追求卓越”之类的说法太多，误导作用远大于励志作用。
- 各门课都有这方面问题，数学建模尤其不能走入这个误区。今天，社会转型、经济发展和科技进步，我们更需要也更有可能是摆脱这个误区，而数学建模的特点恰恰能引导学生走出这个误区。我们应该有这样的信心和这样的追求。

二、数学建模要强化几个意识： 问题意识，分解意识，检验意识

- 数学基础课与数学建模有什么相同与不同？
- 数学基础课主要是传授知识打基础，数学建模则是培养能力用数学。相同之处是两者都需要概念定义、条件结论、分析计算和逻辑推理，都需要理出思路、找到正确的路线。差别在于：数学课后的题目大多直接用数学语言叙述的，条件、结论给得很明确，（除要求举反例外）答案基本上是唯一的，方法也大同小异（很多时候是有“套路”的），做对做错很容易评判，目的是帮助理解和掌握这一部分知识内容；数学建模面对的实际问题一般是用自然语言描述的，条件和结论并不明确、具体，甚至是相当开放的，讨论或解决的路线、方法及用到的知识内容可能是多种多样的、综合的，更重要的是得到的结果要经过实际检验。

- 不重视这种差异会让数学建模走偏。数学建模要先把问题搞清楚：需要弄清楚什么、回答什么或解决什么？这一步错了往下全是无用功。然后作一点定性分析：大致需要什么条件，什么是根本不可能的，主要的困难在哪里，怎样合理地分解？这一步做得好就容易走下去，容易找到一条正确的路线而不至于盲目搬方法、套公式。这一步很重要但仅靠数学知识是不够的，需要学习相关知识或找人合作。接下来是选择变量和建立变量之间的关系，即建立数学模型。这一步需要明确假设：没有假设是无法建立模型的，假设不合理会偏离问题的本源，而模型太复杂又会超越现有的能力。有了模型，就可以通过数学分析和编程计算来回答原始的问题，这基本上是数学知识的运用和计算机相关的能力。事情至此还没有完：我们还必须检验所得的结果是否正确或合适。

- 我们举几个例子。
- 第一个例子是“杭钢1#高炉炼铁优化操作系统”，这是刘祥官老师调入浙大应用数学系后带着几个人完成的第一项工作。企业提出的目标要求是：把高炉利用系数（高炉单位容积单位时间里的出铁量）提上去，把综合焦比（生产一吨生铁需要消耗的焦炭折算后的平均值）降下来。到现场深入考察并请教相关专家（包括相关资料）后把任务归结为：根据采集到的温度传感器、炉顶烟气分析等数据以及目前的操作情况，确定新的操作方案。同时对照先进水平，分析实现目标的可能性。然后找出主要的、挑战性的数学问题是：由高炉外壁的温度信息推断（反演）炉内的温度分布；建立参数与指标之间的函数（统计）关系。解决这两方面问题后，着手编程、调试、检验。

- 第二个例子是公共自行车运营管理。杭州市是国内最早推出公共自行车的城市，每辆车骑行一小时之内免费（可归还后再借，从头算起，依然免费），市区范围内共设有4000多个站点，投放近九万辆车，年借用一亿多次，杭州市政府每年根据公共自行车的借用量核定财政拨款。该系统已复制到国内几十个城市，最近获得了重要的国际大奖。由于自行车借还都需要在锁桩上操作，而每个站点只有20余个锁桩，市民经常会遇到“借不到”或“还不了”的尴尬，影响了使用体验。同时，公共自行车公司只有150余名维修工，要维修保养分布在4000多个站点的自行车，即使自行车损坏率不高，，也会疲于奔命、捉襟见肘、力不从心。正因为如此，“共享单车”就应运而生，有了市场机会（按下不提）。公共自行车公司的信息化工作做得非常好，使得上述困难有了解决的需要和研究的条件。我们简单说一下“维修保养”问题。从数据发现，有少部分站点在45天内到过的不同车号的自行车数目占全部自行车的90%以上。由此入手，我们就可以尝试、探索各种可能的解决方案，并与基层员工切磋、实践。

- 因此，相对于数学基础课，数学建模需要强化几方面意识和能力。
- 1. 问题意识：数学建模面对的是实际问题而不是数学问题，要先问用什么模型，再问用什么方法。倘若现有的方法和问题之间已经可以建立1-1对应，那世界上就不会有困难的事情了，也不需要创新发展，更轮不到我们施展才华。合适的模型需要合理的假设，合理的假设需要把实际问题搞清楚。问题比答案重要，实践比理论在先。
- 2. 分解问题的意识和能力：一步到位、一揽子解决、毕其功于一役的想法和心态难以面对富有挑战性的问题。有些困难的问题我们做不了但却能读懂别人的工作，是因为别人把问题巧妙地分解成一系列有办法处理的子问题。合理分解的标准是：子问题的难度大大降低了；子问题解决的结果拼起来就能回答原问题。
- 3. 检验意识：数学问题的解答正确与否可以靠逻辑和证明来判定，但逻辑与证明不能作为实际问题是否得到了正确解决的唯一准则，而必须一开始就考虑检验的标准和途径。
- 从事数学建模除了必要的数学基础外，更需要充分的好奇，敏锐的捕捉，虚心地学习，合理的假设，恰当地分解，规范地研究，科学地检验。

三、别挖个坑把自己埋了：

几头不落好，里外不是人

1. 数学建模不应该是传授数学知识的一门新课。姜伯驹院士曾说过：“别把数学建模搞成‘建模数学’”。
2. 找准定位、补别人的短板、彰显特色应该是我们的目标，证明定理、发表论文非我们的优势所在，同质化竞争难有胜算。一件事如果其中几乎没有数学问题或数学困难，就不需要也不适合我们去做；如果已经是纯粹的数学问题，我们也犯不着抢别人的饭碗。我们不能指望别人把数学模型建好后提出标准的数学问题交给我们来研究，也不能局限于写篇论文或研究报告交账。那样很难提高能力，很难得不到机会，很难打开局面。
3. 方向自信、道路自信、文化自信、体制自信

四、数学建模贵在实践，难在实践

1. 许多实际问题已经被抽象成标准的数学问题，经过多年研究已经有大量的学术文献和成熟的方法，未解决的难点大多是让成果落地，如组合优化问题。实际问题的真正困难在于其过于庞大、过于复杂。上网检索有价值的东西很少。这恰恰为我们留下了空间。
2. 计算机模拟很容易被误用、滥用。真正好的东西我们了解得还不够多。
3. 问题比答案更重要，实践并不比理论更容易。

五、数学建模可以让我们想想今天的数学该怎么教，从两个例子说起

- **例1. 某大学微积分考试有这样一道题：**桌上放了两块饼，试证：必可以一刀切下去将两块饼（面积）同时平分（假定刀足够长，饼厚薄是均匀的）。这是一道很有趣的数学建模练习题。用数学语言来叙述就是：对平面上任意两个有界区域，必定存在一条直线将他们的面积同时平分。很容易想到，这道题应该利用连续函数的介值定理来证明：一条直线在平移或绕定点旋转（这两种情况都是单参数的）的过程中必然会与给定的某个有界区域相交，且该区域被直线分割成的两部分的面积都是连续变化的，其中一部分的面积连续增大而另一部分面积连续减少。因此，问题变成（1）怎样把这个想法严格地数学化；（2）同时考虑两个区域时，应该怎么做？这里需要用到的数学基础课相关知识有：* 直线的表示；平面区域（边界）的表示；直线与平面区域（边界）交点的表示；平面有界区域面积的计算，对区域边界有什么要求；等等。我们的第一步是要计算并证明平面有界区域被直线分割得到的（某一侧）面积是直线参数（变动时）的连续函数。然后进入第二步：怎样才能再次用介值定理？

- **例2.** 学生有一次问我，从玉泉校区宿舍出发乘89路公交车到紫金港校区教室去上课一般需要预留多少时间最合理？这里自己无法掌控的是到达上车地点后等车和公交车行驶到目的地（总共 k 公里，中间停靠 m 个站，会遇到 n 个红绿灯）所需要的时间。我们可以让学生把他乘车的时间段里多次乘车的各类情况分别记录下来，然后用学过的统计方法来进行处理。这时就产生一个问题：等车时间、各站上下客时间、遇交通信号灯的等待时间和平均车速这几个随机变量的期望之和是否就等于赶路所需时间之期望值？或者简单地把每次总耗时记录下来进行统计处理。这时，我想起了读过的一个故事。

- 
- 华罗庚和王元所著的《数学模型选谈》这本书很值得读也很适合大学生读。其中讲“简单统计方法”时举了如下一个例子：

以 $x_1, \dots, x_{20}$ 表示某一试验独立重复20次的观察值。命 $\bar{x} = \frac{x_1 + \dots + x_{20}}{20}$ (均值)

与 $S = \sqrt{\frac{\sum_{i=1}^{20} (\bar{x} - x_i)^2}{19}}$ (标准离差)。这时, 做实验的人可以声称, 区间 $(\bar{x} -$

$\frac{1.96s}{\sqrt{20}}, \bar{x} + \frac{1.96s}{\sqrt{20}})$ 是某种“均值”的“置信”估计。这样复杂的方法似乎不易

- 为我国的普通工人所理解。此外, 基本的Gauss假设在这里很可能不成立!

实际上, 我倾向于用如下的简便方法。将观察值排好次序, 记为

$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(20)}$ 。我们可以如实地说, 观察值落在区间

$(\frac{x_{(1)} + x_{(2)}}{2}, \frac{x_{(19)} + x_{(20)}}{2})$ 中的可能性大于 $\frac{18}{19} = 0.9$ 。这个方法可以很快地被我们的

工人和农民接受。

- 我让学生思考如何证明华老的结论，需要什么条件？一些学生不知如何下手，也有学生告诉我说此结论是不对的，他用计算机模拟了几千次，符合这个结论的远不到百分之九十。其实，这个问题也就是要从已经得到的数据中产生两个数构成一个区间，然后极大化落在这个区间中的概率。
- 这件事让我感到学生学统计没有真正学懂，印证了华老在上述例子前说的这样一段话：“在实验科学中，我们常常应用统计方法。当然不能否认，这些方法是重要的。然而，我个人认为某些方法太复杂繁琐，而且很容易被滥用、误用。”
- 我们教数学基础课应该开宗明义、贯穿始终地说清楚几件事：这门课讲什么，为什么要讲这些东西，它们与以前学过的东西有什么关系，其中最主要、最核心的概念、记号、定理有哪些，怎样与实际联系起来直观地理解。



谢谢各位！