



1. 学术兵法 01

MATH

关于基础代数学习 03

——别品贤

2. 思想纵横 09

阶的估计技巧与应用 11

——马金韬

正切函数的连分函数展开及其应用 19

——陈德豪

3. 匠心独运 25

数值积分浅谈 27

——农孝儒

微分方程数值解的稳定性 35

——卢小石

统计模型在数量遗传学中的应用 44

——管瑞

金融交换产品综述及其投资应用 70

——邱嘉诚

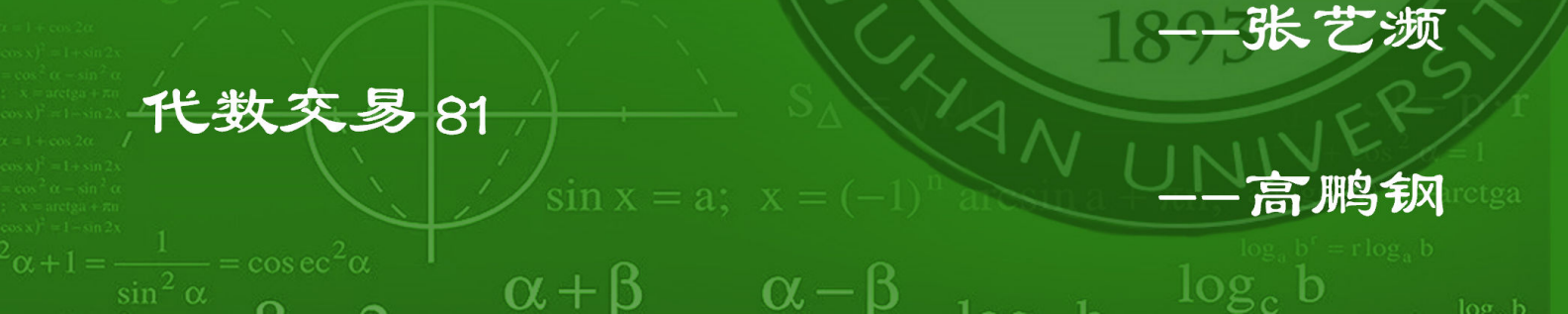
4. 学海掠影 77

榜样珞珈年度人物的数学学习之路 79

——张艺濒

代数交易 81

——高鹏钢



珞珈数学编辑部

指导老师

孙俊 赵晓飞 张继伟
刘研岩 胡亦钧 胡法拉

主编

韦苏原

编委

乔浩博 李坤贤 朱曹阳
但宣昭 张蔚泓

设计

李文月

排版

韦苏原 张蔚泓

审稿

尹鹏

特别鸣谢

武汉大学数学与统计学院

新媒体运营中心

写在本书的最前面

齐民友

数学是为了什么？很难有一个答案。因为各种不同的答案其实都是相互交叉的。但是，大体上有两种：一是为了某个实际的目的。这不一定不好。例如学习某个其他学科需要一点数学知识。我们有针对性地学习有关数学知识，这是很正常的。也就是说，把数学当作一块敲门砖，这不是不可以，问题在于敲什么样的门。大家都在批评应试教育，好像责任全在学校和教师，这就很不公平。我们都看到有的人就是为了混一个什么才来学数学，甚至“动机”很不一样，难道也怪应试教育吗？也还有另外一种学数学的人。目的应该说是很高尚的：为了祖国和人民，为了进一步探索科学真理，等等。还有很多很复杂的情况，不可一概而言。更不能简单地用好坏二字划分。

但还有一个比较简单的划分方法。有一种情况是：把现有的教材学到能 OK 就行了。另一种情况就是，还需要或自己感到有一种要求，要多懂一些，多学一些，也希望自己能参与创造进一步的知识。现在提出建设创新型国家，这里有许多政策性质，经济性质的问题，对于一个学校里的学生和教师，似乎有一点距离。但是，把学校生活，把学数学的过程，变得更充满活泼的创造气氛，是我们共同的愿望。只有这样，才能做到多懂一些，多学一些，自己也能参与数学的创造。而且在大学里生活的更充实一些，更愉快一些，更有意思一些。创造有三种境界：创造的欲望，创造的磨练（说句大家听起来可能有点烦的话，做习题是最简单也最不可少的磨练，数学不是看懂的而是算懂的），创造的愉悦。希望《珞珈数学》这个刊物能在这方面满足读者。

现在很流行一种“傻瓜书”，例如 WINDOWS XP 傻瓜书，英文是 WINDOWS XP for dummies。特点就是，规定要记忆的几条，背得下来，就叫做学会了。对于傻瓜书也不能一概否定。例如，如果有人要学一点线性代数，目的就是例如一门什么课程，出现了矩阵记号。这样，找来一本线性代数 for dummies 就能满足要求，问题在于有一些书，有一些教材，本来不是傻瓜书，而有丰富的内涵，可是把它当作傻瓜书来读，也是常见的事情。总的特点就是，完全不需思考。所以现在也有反对傻瓜书的网站。希望《珞珈数学》这个刊物能帮助读者们，学会不把很好的课程或者好书当作傻瓜书来读，学会把数学作为一门充满创造精神的科学来体验，享受。

编者的话

武汉大学数学与统计学院“珞珈数学研习会”每年出版的期刊都颇受同学的喜爱与肯定，但自新冠肺炎疫情爆发以来，社团成员们的学习任务以及编稿交流，都受到前所未有的挑战，2021年被迫停刊。随着时下疫情形势的好转，为营造良好的学术氛围，在院团委的促进下，“珞珈数学研习会”推出了2022年期刊。

数学作为一门严谨而又天马行空的学科，既有理性的逻辑推导，又有感性的直观感知。作为数学界的巨匠大师、我国函数论研究的主要奠基人李国平院士曾言“真正的学问不是靠老师教出来的，而是靠学生自己钻研出来的。唯有能自学的人，才有可能成为真正的学问家。”自学与钻研的重要性与必要性可见一斑。“珞珈数学研习会”由志同道合的武大数院学子创办，旨在通过以学生为主体的数学研讨精进学习之道，提升思辨能力。本期刊的主体内容均为数学与统计学院2018级、2019级和2020级本科生撰写，或数学方法，或思维延拓，或交叉学科中的数学应用。期刊欲带领读者从各个维度观察数学，调试与数学相处的姿态。刊中的符号和文字的声音能够给杂器以宁静，给燥热以清冽，呈现独特的数学之美。

本期刊的编写得到了武汉大学数学与统计学院的大力支持，得到了孙俊老师、赵晓飞老师、张继伟老师、刘研岩老师、胡亦钧老师对所有投稿文章的建议与指导，得到了在美国深造的管瑞学姐的指点迷津，得到了数模协会主编衣孝儒的经验之谈，得到了数院学子的积极投稿，在此一并表示衷心的感谢！是你们让本期刊得以发光发热。

真诚地欢迎广大读者以及专家提出进一步的批评意见和帮助,投递邮箱:2304018045@qq.com

——2022年8月于珞珈山

part 1

學
術
真
法

MATH

关于基础代数学习

作者：别品贤 2019 级 数学与应用数学班

内容提要

- 线性代数
- 近世代数

- 交换代数与同调代数

1 前言

这篇文章写于笔者大三结束的暑假。笔者应珞数之邀，决定写一些更实用的、对各位读者更有帮助的东西与大家分享。这篇文章的初衷是给 20 级及往后的学弟学妹们对本院开过或没开过的代数方向的必修课与选修课一个大致的了解，并给出一点点浅薄的学习参考。诚然笔者才疏学浅，文字中若有偏颇之处，还请读者们不吝赐教。

2 基础之基础—线性代数

线性代数是数学专业任何专业方向都必须熟练掌握的基础必修课，也是很多非基础数学专业的同学本科学过的唯一一门代数课。不同于数学分析中讲究的是分析学的严谨性，线性代数一开始就展现了代数学独有的抽象性与技巧性，这也给初学线性代数的同学带来了一定的难度。可以说从线性代数的学习中，同学们就可以隐约感觉到自己是否有继续学习代数学的天赋了。如果你能够在代数学习的过程中被抽象的代数结构所吸引，并从看似无从下手的练习题中获得解题的灵感，而且还在这个过程中获得了快乐，那么恭喜你已经拥有了继续深入学习代数学的潜力！

笔者一直倾向于认为学习数学的能力和天赋都是可以后天培养的，研究代数的天赋也是可以从后天的学习与训练中慢慢产生的。以笔者个人为例，我在进大学之前是没有任何代数学基础的，对代数的理解仅仅停留在中学里学习的如何解一元二次方程上，可以说大学的代数对当时的我来说是完全全新的知识。但是通过 3 年的学习，现在的我可以自豪地说代数学是我本科里学的最好的基础数学方向，这很大程度上也归功于我大学一年级线性代数的基础非常扎实。要想打好线性代数的基础，笔者认为主要包含着两个方面。一是要熟悉代数学抽象的语言，并能够习惯用这样的语言“说话”；二是要能够积累一些代数学中的经典 tricks，并能够将这些 tricks 为己所用。对于第一个方面，最好的方法就是多看书多思考，把书中抽象的定义反复体会，自己想想能不能通过一些实例将抽象与具体结合起来（这点对于初学者来说是至关重要的，解析几何就为线性代数提供了大家非常熟悉的具体的舞台）但同时也要能够反过来，将熟悉的、具体的东西中抽象的代数结构提取出来，并将之推而广之，得出更一般性的结论，这样就能将大家熟悉的结论应用到更广阔的空间之中。对于第二个方面，最好的

方法则是多刷题多积累。我知道有些人是不赞成唯刷题论的，但是对于基础课程的学习，笔者认为适量的刷题对线性代数技巧的积累是非常必要的。这里笔者想要强调的是高质量的刷题，而非高数量的刷题。线性代数的内容其实并不能算很多，许多题目的“套路”其实都是相似的，所以这时多积累典型题目就显得尤为重要。从本人的经验来说，反复体会一道经典的题目，有时带来的效果比无目的地刷 10 道题还要大得多！笔者十分推荐初学线性代数的同学准备一本习题集，每周抽空把这一周来见过的有难度、有趣、有价值（这里就因人而异了）的题目抄在习题集上，并且在每道题后面附上一两句简要的 hint（不要抄完整解答！）。等到日后有时间了重新翻阅习题集时，再回头看看自己之前写的 hint，是否能回想起这道题目的关键技巧？能否高屋建瓴地把握住这一类题的解题思想？长此以往，积累的线性代数的 tricks 就会如条件反射一样储存在大脑中，最终达到“看题得解，一题多解”的境界。

推荐书籍：作为本科第一门代数课程，笔者认为还是以中文书籍为主较宜。

1. 复旦姚慕生、谢启鸿 高等代数白皮书：首推经典读物，适合任何层次的学生，建议 n 刷。以例题为主，其中许多经典的例题要反复体会，课后习题难度较低，仅供参考。
2. 线性代数 科大 李炯生：非常经典的教材，观点很高，但是难度也很大。其中课后习题的价值最高，缺点在于没有答案，建议几个同学约好一起刷这本书，可以互相讨论习题解法。
3. 樊启斌 高等代数 900 题：我校樊启斌老师整理的宝贵的线性代数习题集，非常适合课后自主练习。题目由易到难都有，个人建议是遇见难题实在无法想出大可先放下，待到学后面再回头看看，能否用新的观点解决之。若能将整本书做出 80% 以上，线性代数的基础基本上就很扎实了，做出 90% 以上，线性代数就算基本通关了。
4. 樊启斌 高等代数习题解（樊砖）：樊老师去年新作，建议配合 900 题使用。先自主思考 900 题上的习题，如遇困难再在樊砖上寻找答案。樊砖很厚，如果初学者阅读起来比较吃力，还是建议先以白皮书为主要参考书。

3 基础篇—近世代数

近世代数又称为抽象代数，往往会开设在大二，主要讲授群论、环论和域论的基础知识，有时也会介绍一点基础的模论和伽罗瓦理论。作为第一门系统的介绍和引入代数学的思维和方法的学科，近世代数于代数学的学习有着地基般的重要意义。实际上，笔者认为任何一个数学专业的同学都应该对近世代数的基本内容有一定了解，至少也应有所耳闻。由于近世代数本身的内容是极其丰富的，作为一篇介绍性的文章，不可能都展开来说，所以笔者只略笔介绍一下三个较为主要的分支和一些学习经验。

群论顾名思义为研究名为群的代数结构的学科，其在近世代数中具有最为基本的重要地位，因为许多更复杂的代数结构，包括环、域、模和向量空间等等都可以看作是在群的基础上添加新的运算和公理而形成的，进而许多群论的结果和方法可以在新的代数结构中进行平凡的推广。比如说，群的同态基本定理、第一同构定理和第二同构定理，就可以完全照搬到

环、模等代数结构中。而在群论中，通过群的自同构研究其自身的结构、通过研究群作用在不同的代数结构上，也可以用来研究群或其它的代数结构，这些也都是现代代数学中最为基本的研究方法。

环论主要研究的是比群更特殊的代数结构，其主要内容是介绍环的理想的性质，内容包括中国剩余定理的环论版本（你将会发现它将概括你在数论和线性代数中所熟知的版本）一些特殊的理想（素理想、极大理想），然后就是关于多项式环的讨论，最后则是有关几种特殊的整环的介绍。可以说其最核心的内容还是围绕着理想展开的，而其余内容可以看成交换代数、代数数论和域论等领域的铺垫，所以相较于群论，环论“独立”的内容可能比较少，而且与其它学科相关的知识可能初学时有些难以理解，可以在学到更加深入的知识的时候再回过头来复习环论的知识点。以达到温故知新的效果。

模实际上可看成线性空间的推广，因而线性代数的许多结论在模论中都有着相对应的结果。初学者第一次接触模的概念时，可能会觉得有些抽象，这时不妨类比线性代数的结论进行理解，并多掌握一些例子，这会有助于对概念的理解。比较经典的模论包含两个方面，一个是基础概念的介绍，这一部分内容是相对简单的，同时也是学习交换代数与同调代数的基础。另一部分则是关于“主理想的有限生成模结构”的理论，这则可以看成线性空间中线性变换的 *Jordan* 标准型理论的推广，是非常漂亮的经典结果。

域的代数结构相比环更加好，所以我们不再拘泥于一个域本身的代数性质，而是更关心的是域的扩张理论，这是域论的核心内容。域论大致包括有限扩张、代数扩张、分裂域、正规扩张、可分扩张，以及伽罗瓦扩张这些内容。不同扩张之间的性质、关联、应用，是较为繁琐的，笔者这里就不赘述了。尽管域论的概念看起来并非十分繁复，其习题的难度却是较高的，根据个人的经验，如果不进行一定量的习题进行训练，是几乎不可能完全掌握好域论的知识点的。伽罗瓦理论建立了域论和群论的联系，是近世代数最为伟大的结果之一，由于教学课时的原因，很多学校一个学期的代数课程基本上都没能讲到伽罗瓦理论，但本人是非常推荐读者们在假期时间里抽时间研读一下伽罗瓦理论相关的内容，对拓宽视野、锻炼代数素养等都是大有裨益的。

初学者学习近世代数时，一定要耐心理解不同概念及其联系，并且一定要掌握用重要例子去理解一些重要的结论。实际上，对于抽象的代数学而言，例子可以说是学习最重要的“救命稻草”，如果脑海里不能储备大量的例子，大部分情况下学完后很容易快速遗忘。那么例子可以从哪里来呢？笔者的答案是可以从初等数论和线性代数中找到丰富的例子。同时，笔者还认为我们在学习近世代数时，一定要具备把抽象的语言翻译成“人话”的能力，搞清楚书上每一个步骤是想干什么，为什么要这样写，为什么不那样写，其逻辑是什么，是至关重要的。在做练习题时，由于学科的抽象性，习题中会有一些技巧性很强的证明和习题，如果初学时卡住了，可以不用过于较真，但一定要在心底留一个问号，第一遍先着重理解主干的知识和体系，待到读者的代数层次到达一个新的境界之时，再回头看看也许就会恍然大悟。

推荐书籍：

1. 聂灵沼、丁石孙 代数学引论：难度相对而言较大的近世代数书籍，但内容是国内书籍中最为丰富的，尤其是基础的环论和模论知识，习题难度较大，且没有完整答案，但是

只要自己肯静下心来认真研究课后习题，就一定会对代数水平的提升大有帮助。

2. 丘维声 近世代数：本书的一个特色是群论占据了其中相当大的篇幅，且较为全面，而且对初学者比较友好，常常从高等代数、初等数论的一些常识引入新概念，并且整个体系非常清晰，个人推荐将其作为初学群论的读物。
3. *Michael Artin Algebra*: 经典的英文代数教材，讲法与国内的教科书很不一样。但内容非常全面，并且也很适合初学者入门，大多数习题难度不算很大。
4. *J.S.Milne Notes : Group Theory; Fields and Galois Theory*: 可以从网站:中进行下载，是很好的 notes，无论是初学还是复习都非常值得一读。

4 进阶篇—交换代数与同调代数

如果你已经成功拿下了近世代数，想要挑战更加深入的代数，那么稍微进阶一点的便是交换代数与同调代数了。其中，交换代数在现代代数数论与代数几何中有非常广泛的应用，而同调代数则在代数拓扑中扮演了重要的地位。笔者希望简单介绍一下这两门学科，并分别给出一些学习上的建议。

经典的交换代数主要包括交换环的一些基本的性质，分式环(模)和局部化方法，诺特环与阿廷环—两类重要的交换环，赋值环与戴德金整环，完备化和维数理论等等。交换环的性质和模论较为贴近近世代数的内容，可以看做近世代数所学知识的深入和推广，所以笔者认为有志学代数的读者可以在学完近世代数后无缝衔接交换代数。交换代数中关于理想的根、素理想与极大理想、准素理想和支集，环的局部化和局部性质等内容，笔者认为是相当有趣的，而且理解起来不太需要其它更深入的学科，尤其推荐学习完近世代数，对更深入的交换环理论感兴趣的同学尝试着更深入地学习一下。笔者就是在学习交换代数的这些内容时渐渐感受到了代数学的独特魅力的。至于其他内容，比如赋值环和完备化和维数理论等本身有着很强的代数数论或代数几何背景，在初学时可能会觉得晦涩抽象，如果不进行深入的练习，也是很容易遗忘的。所以笔者认为交换代数的部分内容其实可以留到学习代数数论与代数几何的时候再一起学习，这样学习效果会更好一些。若有志于彻底学好交换代数的同学不妨找我们院的薛江维老师寻求指导。

有人说交换代数是本院本科阶段开设的难度最大的课程之一，我个人是认同这一说辞的，但是还请读者们不必对它产生畏难情绪。事实上笔者认为在初学交换代数时也不用过于较真，有些定理的证明自己无法证出来是非常正常的，但必须要做到的是理解证明的核心思想，理解定理和定义并且会使用。习题的训练仍然是十分必要的，任何一门代数课程不做大量的 exercise 是永远没办法掌握牢固的。初学者可以用模仿的方式来演练学过的内容，而实际上，很多 exercise 其实就是让你模仿一定理的证明，以达到增强理解的目的。笔者在学习交换代数的时候也曾经经历过几个小时苦思冥想一道题的折磨，但就是在这样反复的折磨过程中渐渐地体会到了一些交换代数的魅力，渐渐地也慢慢对这门学科上手了不少，所以笔者也希望有志研究代数的读者们都经历一下这样的“磨练”，方能获得实力上的成长。

同调代数可能是目前笔者接触的代数课程中，抽象程度最高的一门课了。相较于交换代

数，同调代数可能更偏向于一种 language，而其主要难度其实恰恰就在于其高度抽象性。同调代数的发展与范畴论的出现密不可分，而范畴与函子的理论其实并没有引入什么新的东西，仅仅是将我们之前见过的所有代数结构用了一个更加广泛的语言进行概述。在后面的学习中，我们将会渐渐体会到用范畴的语言来描述代数对象的优越性。回到同调代数本身来说，同调代数是(上)同调函子及其代数结构的研究。“同调”与“上同调”是一对对偶的概念，它们满足的范畴论性质相反。本科阶段的同调代数课程，主要包括模论、范畴论、复形和导出函子、群的同调、谱序列等等内容，其上手难度远高于仅仅需要基础环论知识的交换代数。但同时，有别于交换代数的应用场合主要在代数几何中，同调代数作为一个基本的工具，其在基础数学各方向的应用都是非常广泛的。所以笔者认为，同调代数是所有学基础数学的学生（不仅限于代数方向）都应该花时间了解并掌握的一门重要课程。

那么如何才能学好同调代数呢？这个问题笔者也不敢回答，就让笔者借用自己在一本同调代数的教材中读到的话吧！该教材的作者说，想要学好同调代数很简单，只需要随便找一本同调代数的书，把书中每一个例题和习题中的交换图 (*commutative diagram*) 自己动手跑一遍就完事了。这番话让笔者受益匪浅，笔者认为同调代数的学习是绝不能照本宣科的，对着书上的例子看作者的证明看得津津有味就觉得自己掌握了，这种行为是极其错误的。同调代数的学习一定是要“hands-on”的，你不动手操作一下，不在草稿纸上老实地跑图验证，你就永远没有学会同调代数，对同调代数的理解也仅仅停留在最基本的概念上，甚至可能连基本的概念都没有理解到位。但同时，同调代数的高度抽象性也让许多读者望而却步，这时具体的例子又一次发挥了重要作用。同调代数本身其实是在所谓“阿贝尔范畴”下进行的，而我们之前学过的阿贝尔群、环上的模等等，它们形成的范畴其实就是阿贝尔范畴的具体例子。当读者对抽象的同调代数感到困惑的时候，不妨回到更具体更熟悉的代数结构中，这样既回顾了以往的代数知识，也能更好的理解正在学习的知识，可谓是一箭双雕。

推荐书籍：

1. *Atiyah Introduction to Commutative Algebra*: 最为经典的交换代数教材，内容丰富而简明，正文难度不是很高，但习题难度较大且十分重要，一定要认真做完，其答案在 google 上就可以搜到全套 pdf。这本书中的许多内容在代数几何中有非常丰富的背景，是笔者最为推荐的交换代数教材。
2. *Allen Altman and Steven Kleiman A Term of Commutative Algebra*: 用更为现代化的语言和处理改写 *Atiyah* 的书的一本讲义，在作者主页有 pdf 下载和习题解答，可供读者选择，参考。
3. *Weible An Introduction to Homological Algebra*: 非常经典的教材，上手有些难度，但如果想掌握扎实的同调代数基础，是非常建议阅读的。
4. *Rotman An Introduction to Homological Algebra*: *Rotman* 的书都写的很厚，但厚的好处在于写的内容非常详细，对初学者非常友好，有耐心的读者可以选择这本书来学习同调代数。

5 给新生数学学习的几点小建议

这一段的主要内容是根据笔者三年来的学习经历，对自己在数学学习上的经验教训进行了一些总结，并决定分享给刚入学的新生们，希望能够对他们的本科学习有所帮助。

5.1 养成良好的学习态度

不畏难也不浮躁，良好的学习态度应是“自信对待数学学习，谦虚面对数学学科”。一是不要把数学的学习妖魔化，宣扬“只有高智商才能学好数学”。如果没有足够的自信和勇气，妄自菲薄瞻前顾后，是不可能学好的。勇敢面对学习中遇到的困难并努力克服它们，这个过程本身对于是否继续从事数学工作的人都是十分有益的。二是数学作为一门基础学科，其内容的深度与广度非同一般，无论何人何时都应保持敬畏的态度，一些看似基础的内容往往在后续的学习中起到至关重要的作用。切勿停留在自己的舒适区内，应不断突破自己的上限，否则很可能在后续的学习中遭到“反噬”。

5.2 找到最适合自己的学习方式

许多同学刚进校时都迫切地想知道，怎样学好数学？笔者的评价是因人而异。事实上笔者熟悉的很多本科数学成绩拔尖的朋友，他们学习数学的方式都不尽相同。所以笔者认为刚进校的同学，不妨先把学长学姐们推崇的学习方式尝试一段时间，选择适合自己的方式学习。这里分享笔者的学习方式：笔者是两相结合派，即扎实基础的同时也适当进行自学。一般在寒暑假自学，相对浅略地读一些未来要学习的课程书籍。而在上课中以课堂内容为主，笔记不必一字不差，根据理解简要写在草稿纸上课后整理，课下选一些课外参考书进行补充练习。笔者有每周撰写习题集的习惯，也是个人最习惯的温故知新的方式。

5.3 积极主动的参与数学问题的讨论

笔者曾多次主讲“学习邦”讨论班。大学与中学非常不同的一点在于，要把所学的知识消化好，课下积极主动地参与讨论尤为必要。讨论的对象可以是下课还没走人的任课老师，也可以是学长学姐，或者是同班同学。讨论的范围不必局限于课堂上所讲的内容，更不要将讨论作为应付完成作业的工具。讨论的前提应当是自己对某一问题已经有一定的思考，但是某一方面没弄明白，希望和他人探讨。讨论的结果更应思考能否将问题推而广之，得出一个更深刻的结论。讨论的形式是多样的，有小范围私下讨论，笔者也很推崇同学们参加或自行举办大范围的讨论班，通过讲题，和台下的同学互动交流，往往会产生意想不到的效果。能理解一道题、会做一道题、能讲一道题，这是三个递进的关系。把一道复杂的题目讲好，是综合能力的最好体现，笔者非常建议同学们把握住机会，在翻转课堂或者“学习邦”讨论班等平台上锻炼能力，对日后的数学学习也会有很大的帮助。

PART 2

思想
纵横



阶的估计技巧与应用

作者：马金韬 2019级 基地班

1 绪论

阶的估计本质上是一种更加细致精确刻画某个量的极限过程的方法。比如直观的可以感受到，对于 $0 < a < b$ 而言， x^a 相比 x^b 在 $x \rightarrow 0$ 的过程中减小的要更慢一些，而在 $x \rightarrow \infty$ 的过程中要增大的更慢一些。我们做阶的估计就是要更精确的刻画到底“慢”多少。

我们将在第 2 部分给出所谓 O 记号和 o 记号的定义，并给出关于它们的一些基本的性质。第三部分我们将给出一些例子，读者将发现在有限数的极限过程中，Taylor 多项式会在阶的估计过程中有重要作用。在第 4, 5 部分我们将给出阶的估计方法的应用。

鉴于篇幅，我们这里将只讨论最简单（但仍然十分重要）的一些阶的估计的技巧与应用，我们将只关注实的一元函数。希望进一步学习阶的估计技巧的读者可以参考阶的估计基础，潘承洞，于秀雄，高等教育出版社，2015。

2 记号与定义

定义 1

设 f, g 是定义在 \mathbb{R} 上的两个实值函数，若在某点 $x_0 \in \overline{\mathbb{R}}$ 处（这里 $\overline{\mathbb{R}}$ 表示扩充实数系），存在 $\varepsilon > 0$ 以及 $M > 0$ 使得对任意 $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$ 有

$$|f(x)| \leq M|g(x)|, \quad (1)$$

则我们记当 $x \rightarrow x_0$ 时 $f = O(g)$ 。



定义 2

若在 $(x_0 - \varepsilon, x_0 + \varepsilon)$ 上 $g \neq 0$ ，且有

$$\lim_{x \rightarrow x_0} \left| \frac{f(x)}{g(x)} \right| = 0, \quad (2)$$

则我们记作当 $x \rightarrow x_0$ 时 $f = o(g)$ 。



定义 3

在上面的定义中，若 x_0 是自明的，在不引起歧义的情况下，也简单记作 $f = O(g)$ 和 $f = o(g)$ 。当在某个区间 I 上都有 (1) 式成立的话，我们也称（在 I 上）有 $f = O(g)$ 。



注 定义1中, 特别的, 若 f, g 是定义在 \mathbb{N} 上, x_0 取 $+\infty$, 我们仍然采用相同的记号。不难验证下面关于 O 和 o 记号的性质。

命题 1

设 f, g 和 φ, ψ 是 \mathbb{R} 上的实函数, $\{a_n\}, \{b_n\}$ 是 \mathbb{R} 上的实数列。有以下法则成立:

- (i) 若 $f = O(\varphi)$ 且 $\varphi = O(\psi)$, 则 $f = O(\psi)$;
- (ii) 若 $f = O(\varphi)$ 且 $\varphi = o(\psi)$, 则 $f = o(\psi)$;
- (iii) $O(f) + O(g) = O(f + g)$;
- (iv) $O(f)O(g) = O(fg)$;
- (v) $o(f) + o(g) = o(|f| + |g|)$;
- (vi) $o(f)o(g) = o(fg)$;
- (vii) $o(1)O(f) = o(f)$;
- (viii) $O(1)o(f) = o(f)$;
- (ix) $o(f) + O(f) = O(f)$;
- (x) 若 $a_n = o(b_n)$, 则当 $A \rightarrow \infty$ 时 $\sum_{n=A}^B a_n = o\left(\sum_{n=A}^B b_n\right)$, 其中 $B > A$;

下面进一步设 f, g 可积。

- (xi) 若在某个区间 I 上 $f = O(g)$, 则 $\int_I f(t)dt = O\left(\int_I g(t)dt\right)$;
- (xii) 若当 $x \rightarrow +\infty$ 时 $f(x) = o(g(x))$, 则 $\int_x^y f(t)dt = o\left(\int_x^y g(t)dt\right)$, 其中 $y > x$;

以上的性质某些仍可以做一定的推广, 这里不赘述了。以下使用这些性质时不在特别提出。

定义 4

设有一列 \mathbb{R} 上的实函数列 $\{\varphi_k(x)\}$ 满足 $\varphi_k = o(\varphi_{k-1}), \forall k \in \mathbb{N}^*$ 。若存在 $n \in \mathbb{N}^*$ 以及 $a_0, \dots, a_n \in \mathbb{R}$ 使得函数 f 满足

$$f(x) = \sum_{k=0}^n a_k \varphi_k(x) + o(\varphi_n(x)) \quad (3)$$

则称 (3) 式为 f 的一个渐进多项式。

进一步, 若存在一列实数 a_k , 使得对任意的 $n \in \mathbb{N}^*$ 都有 (3) 式成立, 则称 $\sum_{k=0}^{\infty} a_k \varphi_k(x)$ 为 f 的一个渐进级数。



注 1: 由于 φ_k 的选取不同, 可见一般来说一个函数的渐进级数并不唯一。一般的, 渐进级数也并不一定收敛, 即使收敛, 也不一定收敛到 f 。

注 2: 有时对余项没有很强的要求, 我们也会写成 $O(\varphi_n(x))$ 。

例题 1 我们知道 $\ln(x+1)$ 在 0 处有 Taylor 展开

$$\ln(x+1) = \sum_{k=1}^n \frac{(-1)^{k-1}}{k} x^k + o(x^n), \quad \forall n \in \mathbb{N}^*$$

显然上式给出了 $\ln(x+1)$ 在 $x \rightarrow 0$ 时的一个渐进多项式。

类似的, 对 $e^x - 1$ 在 0 处有

$$e^x - 1 = \sum_{k=1}^n \frac{1}{k!} x^k + o(x^n), \quad \forall n \in \mathbb{N}^*$$

注 求某个函数在某一点的 Taylor 多项式的过程事实上就是在寻找该函数在该点处的一个渐进多项式。

例题 2 熟知在 $x \rightarrow 0$ 时, 有等价无穷小 $(1+x)^\alpha - 1 \sim \alpha x$ 。于是自然可以得出如下的渐进多项式: $(1+x)^\alpha = 1 + \alpha x + o(x)$ 。

3 $\Gamma(x)$ 在 $x \rightarrow +\infty$ 时的一阶渐进多项式

命题 2

在 $x \rightarrow +\infty$ 时, 有

$$\Gamma(x) = x^{x-\frac{1}{2}} e^{-x} \sqrt{2\pi} \left(1 + O\left(\frac{1}{x}\right) \right). \quad (4)$$



为证明这个估计式, 我们先证明如下若干引理。

引理 1

设 n 为正整数, 当 $x \rightarrow +\infty$ 时, 有

$$\ln \Gamma(n) = \left(n - \frac{1}{2} \right) \ln n - n + C_0 + O\left(\frac{1}{n}\right), \quad (5)$$

其中 C_0 是常数。



证明：由 Γ 函数的递推公式，有 $\Gamma(n) = (n-1)!$ ，从而我们有以下计算：

$$\begin{aligned}
 \ln \Gamma(n) &= \ln((n-1)!) = \sum_{k=1}^{n-1} \ln k \\
 &= \sum_{k=1}^{n-1} \left(\int_k^{k+1} \ln t dt + \int_k^{k+1} (\ln k - \ln t) dt \right) \\
 &= \int_1^n \ln t dt - \sum_{k=1}^{n-1} \int_k^{k+1} \ln \left(\frac{t}{k} \right) dt \\
 &= (t \ln t - t) \Big|_1^n - \sum_{k=1}^{n-1} \int_0^1 \ln \left(1 + \frac{t}{k} \right) dt \\
 &= n \ln n - n + 1 - \sum_{k=1}^{n-1} \int_0^1 \left(\frac{t}{k} + O\left(\frac{t^2}{k^2}\right) \right) dt \\
 &= n \ln n - n + 1 - \sum_{k=1}^{n-1} \frac{1}{2k} + O\left(\sum_{k=1}^{n-1} \frac{1}{k^2}\right) \\
 &= n \ln n - n + 1 - \frac{1}{2} \ln n - \gamma + O\left(\sum_{k=1}^{\infty} \frac{1}{k^2}\right) + O\left(\sum_{k=n}^{\infty} \frac{1}{k^2}\right) \\
 &= \left(n - \frac{1}{2}\right) \ln n - n + C_0 + O\left(\frac{1}{n}\right).
 \end{aligned}$$

□

引理 2

设 $a > 0$ ，则在 $x \rightarrow +\infty$ 时，有以下渐进多项式

$$\frac{\Gamma(x)}{\Gamma(x+a)} = x^{-a} + O(x^{-a-1}). \quad (6)$$



证明：先设 $a > 1$ 。

由 Γ 函数和 B 函数的关系，我们有

$$\begin{aligned}
 \frac{\Gamma(x)\Gamma(a)}{\Gamma(x+a)} &= B(a, x) \\
 &= \int_0^1 (1-t)^{a-1} t^{x-1} dt \\
 &= \int_0^{\infty} (1-e^{-s})^{a-1} e^{-sx} ds \quad (t \triangleq e^{-s}) \\
 &= \left(\int_0^1 + \int_1^{\infty} \right) (1-e^{-s})^{a-1} e^{-sx} ds \\
 &= I_1 + I_2,
 \end{aligned}$$

下面我们分别估计 I_1 和 I_2 。

$$\begin{aligned}
 I_1 &= \int_0^1 (1 - e^{-s})^{a-1} e^{-sx} ds \\
 &= \int_0^1 (s + O(s^2))^{a-1} e^{-sx} ds \\
 &= \int_0^1 s^{a-1} (1 + O(s))^{a-1} e^{-sx} ds \\
 &= \int_0^1 s^{a-1} (1 + (a-1)O(s) + o(s)) e^{-sx} ds \\
 &= \int_0^1 s^{a-1} (1 + O(s)) e^{-sx} ds \\
 &= \int_0^x x^{-a} r^{a-1} \left(1 + O\left(\frac{r}{x}\right)\right) e^{-r} dr \\
 &= x^{-a} \int_0^x r^{a-1} e^{-r} dr + x^{-a-1} \int_0^x r^{a-1} O(r) e^{-r} dr \\
 &= x^{-a} \Gamma(a) - x^{-a} \int_x^\infty r^{a-1} e^{-r} dr + O\left(x^{-a-1} \int_0^x r^a e^{-r} dr\right) \\
 &= x^{-a} \Gamma(a) + O\left(x^{-a} \int_x^\infty e^{-r/2} dr\right) + O(x^{-a-1} \Gamma(a+1)) \\
 &= x^{-a} \Gamma(a) + O(x^{-a-1}). \\
 I_2 &= \int_1^\infty (1 - e^{-s})^{a-1} e^{-sx} ds \\
 &= \int_1^\infty O(e^{-sx}) ds \\
 &= O(e^{-x}/x) = O(x^{-a-1}).
 \end{aligned}$$

故

$$\frac{\Gamma(x)}{\Gamma(x+a)} = x^{-a} + O(x^{-a-1}) \Gamma(a) = x^{-a} + O(x^{-a-1}).$$

当 $0 < a \leq 1$ 时, 有

$$\frac{\Gamma(x)}{\Gamma(x+a)} = \frac{(x+a)\Gamma(x)}{\Gamma(x+a+1)} = (x+a)(x^{-a-1} + O(x^{-a-2})) = x^{-a} + O(x^{-a-1}).$$

□

命题2的证明：记 $n = [x]$, $a = \{x\}$ 分别是 x 的整数和小数部分。由 (5) 式和 (6) 式，有

$$\begin{aligned}
 \ln \Gamma(x) &= \ln \Gamma(n) - \ln(x^{-a} + O(x^{-a-1})) \\
 &= \ln \Gamma(n) + a \ln x + \ln(1 + O(1/x)) \\
 &= \left(n - \frac{1}{2}\right) \ln n - n + C_0 + O(1/n) + a \ln x + O(1/x) \\
 &= \left(x - a - \frac{1}{2}\right) \ln(x - a) - x + a + C_0 + a \ln x + O(1/x) \\
 &= \left(x - a - \frac{1}{2}\right) \left(\ln x + \ln\left(1 - \frac{a}{x}\right)\right) - x + a + C_0 + a \ln x + O(1/x) \\
 &= \left(x - \frac{1}{2}\right) \ln x + \left(x - a - \frac{1}{2}\right) \ln\left(1 - \frac{a}{x}\right) - x + a + C_0 + O(1/x) \\
 &= \left(x - \frac{1}{2}\right) \ln x + \left(x - a - \frac{1}{2}\right) \left(-\frac{a}{x} + O\left(\frac{a^2}{x^2}\right)\right) - x + a + C_0 + O(1/x) \\
 &= \left(x - \frac{1}{2}\right) \ln x - x + C_0 + O(1/x). \tag{*}
 \end{aligned}$$

由 Legendre 加倍公式，有

$$\ln \Gamma(2x) = (2x - 1) \ln 2 - \frac{1}{2} \ln \pi + \ln \Gamma(x) + \ln \Gamma\left(x + \frac{1}{2}\right),$$

将 (*) 式代入得

$$\begin{aligned}
 \left(2x - \frac{1}{2}\right) \ln(2x) - 2x + C_0 + O(1/x) &= (2x - 1) \ln 2 - \frac{1}{2} \ln \pi + \left(x - \frac{1}{2}\right) \ln x - x + C_0 \\
 &\quad + x \ln\left(x + \frac{1}{2}\right) - \left(x + \frac{1}{2}\right) + C_0 + O(1/x),
 \end{aligned}$$

整理得

$$\frac{1}{2} \ln(2\pi) + \frac{1}{2} = x \ln\left(1 + \frac{1}{2x}\right) + C_0 + O(1/x).$$

令 $x \rightarrow +\infty$ ，考虑到 $\lim_{x \rightarrow +\infty} x \ln\left(1 + \frac{1}{2x}\right) = \frac{1}{2}$ ，可知 $C_0 = \frac{1}{2} \ln(2\pi)$ 。代入 (*) 式就有

$$\ln \Gamma(x) = \left(x - \frac{1}{2}\right) \ln x - x + \frac{1}{2} \ln(2\pi) + O(1/x),$$

即 (5) 式。 □

4 用阶的估计方法得到正项级数审敛法

记 $\sum_{n=1}^{\infty} a_n$ 是一个正项级数。下面我们利用阶的估计方法得出两个正项级数的审敛法。在正式给出之前，我们先回忆最简单的 d'Alembert 审敛法，事实上我们可以认为该方法的条件是给出了一个阶的估计式 $a_n = O(1/n^d)$ ，从而由 $\sum 1/n^d$ 的敛散性可以推导出 $\sum a_n$ 的敛散性。

命题 3

设 $\lim_{n \rightarrow \infty} a_n = 0$, $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} = 1$, 记 $l = \lim_{n \rightarrow \infty} (1 - \sqrt[n]{a_n}) \frac{n}{\ln n}$ (若存在), 则级数 $\sum a_n$ 在 $l > 1$ 时收敛, 在 $l < 1$ 时发散。

证明: $l > 1$ 时:

存在 $1 < k < l$, $N > 0$ 使得当 $n > N$ 时有 $(1 - \sqrt[n]{a_n}) \frac{n}{\ln n} > k$ 。则

$$\begin{aligned} a_n &< \left(1 - \frac{k \ln n}{n}\right)^n \\ &= \exp \left\{ n \ln \left(1 - \frac{k \ln n}{n}\right) \right\} \\ &= \exp \left\{ n \left(-\frac{k \ln n}{n}\right) + O\left(\left(\frac{k \ln n}{n}\right)^2\right) \right\} \\ &= \exp \left\{ -k \ln n + O\left(\frac{\ln^2 n}{n}\right) \right\} \\ &= n^{-k} O(1) = O(n^{-k}). \end{aligned}$$

而 $\sum 1/n^k$ 收敛, 从而 $\sum a_n$ 收敛。

$l < 1$ 时:

存在 $l < k < 1$, $N > 0$ 使得当 $n > N$ 时有 $(1 - \sqrt[n]{a_n}) \frac{n}{\ln n} < k$ 。则由不等式 $\ln x + \frac{1}{x} \geq 1$ 得

$$\begin{aligned} a_n &> \left(1 - \frac{k \ln n}{n}\right)^n \\ &= \exp \left\{ n \ln \left(1 - \frac{k \ln n}{n}\right) \right\} \\ &\geq \exp \left\{ \frac{kn \ln n}{k \ln n - n} \right\} \\ &= n^{-k/(1-k \ln n/n)}. \end{aligned}$$

存在 $\delta > 1$ 使得 $k\delta < 1$, 且可以取 N 充分大使得 $\frac{\ln n}{n} < \frac{1-k\delta}{k}$ 。则

$$a_n > n^{-k/(1-k \ln n/n)} > n^{-1/\delta},$$

故 $\sum a_n$ 发散。 □

命题 4

设 $\lim_{n \rightarrow \infty} a_n = 0$, $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} = 1$, 记 $l = \lim_{n \rightarrow \infty} \left(1 - \frac{\ln n}{n} - \sqrt[n]{a_n}\right) \frac{n}{\ln \ln n}$ (若存在), 则级数 $\sum a_n$ 在 $l > 1$ 时收敛, 在 $l < 1$ 时发散。

证明: $l > 1$ 时:

存在 $1 < k < l$, $N > 0$ 使得 $n > N$ 时有 $\left(1 - \frac{\ln n}{n} - \sqrt[n]{a_n}\right) \frac{n}{\ln \ln n} > k$ 。则

$$\begin{aligned}
a_n &< \exp \left\{ n \ln \left(1 - \frac{\ln n}{n} - \frac{k \ln \ln n}{n} \right) \right\} \\
&= \exp \left\{ n \left(-\frac{\ln n + k \ln \ln n}{n} + O \left(\left(\frac{\ln n + k \ln \ln n}{n} \right)^2 \right) \right) \right\} \\
&= \exp \{ -\ln n - k \ln \ln n + O(1) \} \\
&= O \left(\frac{1}{n \ln^k n} \right)
\end{aligned}$$

而 $\sum \frac{1}{n \ln^k n}$ 收敛, 从而 $\sum a_n$ 收敛。

$l < 1$ 时:

存在 $l < k < 1$, $N > 0$ 使得 $n > N$ 时有 $\left(1 - \frac{\ln n}{n} - \frac{k \ln \ln n}{n} \right) \frac{n}{\ln \ln n} < k$ 。注意到 $x > 0$ 充分

小时有不等式 $\ln(1-x) > -x - x^2$, 又有 $\lim_{n \rightarrow \infty} \frac{\ln n + k \ln \ln n}{n} = 0$, 故可以取 N 充分大使得

$\ln \left(1 - \frac{\ln n + k \ln \ln n}{n} \right) > -\frac{\ln n + k \ln \ln n}{n} - \left(\frac{\ln n + k \ln \ln n}{n} \right)^2$. 进而有

$$\begin{aligned}
a_n &> \exp \left\{ n \ln \left(1 - \frac{\ln n + k \ln \ln n}{n} \right) \right\} \\
&> \exp \left\{ n \left(-\frac{\ln n + k \ln \ln n}{n} - \left(\frac{\ln n + k \ln \ln n}{n} \right)^2 \right) \right\} \\
&= \exp \left\{ -\ln n - k \ln \ln n - \frac{(\ln n + k \ln \ln n)^2}{n} \right\}.
\end{aligned}$$

注意到 $\frac{(\ln n + k \ln \ln n)^2}{n}$ 有上界, 记为 M , 于是

$$a_n > \exp \{ -\ln n - k \ln \ln n - M \} = \frac{e^{-M}}{n \ln^k n},$$

而 $\sum \frac{1}{n \ln^k n}$ 发散, 从而 $\sum a_n$ 发散。 □

5 结语

事实上, 在分析学的各种分支中, 阶的估计都是十分常见的, 与其说是一种方法, 它更像是我们在研究极限过程中的一种习惯。读者将会体会到, 真正困难的地方在于对各种函数的放缩, 以及求渐近展开式时的需要计算的极限。作为例子, 读者可以自行用阶的估计方法考察如下问题:

1. 通过将引理1, 引理2做更高阶的渐进多项式, 将 Γ 函数展开成更高阶的渐进多项式;
2. 模仿第 4 节的方法给出更精细的审敛法!
3. 函数光滑性对其 Fourier 级数的收敛性和极限的影响;
4. 球面上的等周不等式。

正切函数的连分函数展开及其应用

作者：陈德豪 2020级 数学与应用数学班

定理 1

在 $[-1, 1]$ 上有 $f_m(x)$ 一致收敛于 $\tan x$, 其中

$$f_m(x) = \frac{x}{1 - \frac{x^2}{3 - \frac{x^2}{5 - \frac{x^2}{\ddots \frac{x^2}{2m-1 - \frac{x^2}{2m+1}}}}}}$$



定义 1

定义一类函数

$$\varphi_m(x) = \sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+2m+1)!!} \frac{(-x^2)^n}{(2n+1)!}$$



由于

$$\overline{\lim}_{n \rightarrow \infty} \sqrt[n]{\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+2m+1)!!} \frac{1}{(2n+1)!}} = 0,$$

因此原级数的收敛半径为 $+\infty$, $\varphi_m(x)$ 在 $(-\infty, +\infty)$ 上有定义。

特别地, 有

$$\varphi_0(x) = \frac{\sin x}{x}, \varphi_1(x) = \sum_{n=0}^{\infty} \frac{(-x^2)^n}{(2n+1)!(2n+3)!}.$$

引理 1

若 $x \in [-1, 1]$, 则成立不等式:

$$\frac{2m + \frac{5}{2}}{(2m+3)!!} \leq \varphi_m(x) \leq \frac{1}{(2m+1)!!}.$$



证明 若 $x_0 \in [-1, 1]$, 则

$$\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+2m+1)!!} \frac{(-x^2)^n}{(2n+1)!}$$

是交错级数, 并且

$$\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+2m+1)!!} \frac{x^{2n}}{(2n+1)!}$$

单调递减。

记

$$S_k = \sum_{n=0}^k \frac{(2n+1)!!}{(2n+2m+1)!!} \frac{(-x^2)^n}{(2n+1)!}$$

因此有 $\{S_{2k}\}$ 单调递减, $\{S_{2k+1}\}$ 单调递增, 故

$$\varphi_m(x) = \lim_{k \rightarrow \infty} S_{2k} \leq S_0 = \frac{1}{(2m+1)!!},$$

$$\varphi_m(x) = \lim_{k \rightarrow \infty} S_{2k+1} \geq S_1 = \frac{1}{(2m+1)!!} - \frac{3!!x_0^2}{3!(2m+3)!!} \geq \frac{2m+\frac{5}{2}}{(2m+3)!!}.$$

引理 2

若 $x \in [-1, 1]$, 则成立不等式:

$$\frac{\varphi_m(x)}{\varphi_{m+1}(x)} = 2m+3 - \frac{x^2}{\frac{\varphi_{m+1}(x)}{\varphi_{m+2}(x)}}.$$



证明 由引理证明 1 知 $\varphi_m(x)$ 在 $[-1, 1]$ 上无零点, 故

$$\begin{aligned} \frac{\varphi_m(x)}{\varphi_{m+1}(x)} &= \frac{\varphi_m(x)}{\varphi_{m+1}(x)} = \frac{\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+2m+1)!!} \frac{(-x^2)^n}{(2n+1)!}}{\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+2m+3)!!} \frac{(-x^2)^n}{(2n+1)!}} \\ &= 2m+3 + \frac{\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+1)!} \left[\frac{1}{(2n+2m+1)!!} - \frac{2m+3}{(2n+2m+3)!!} \right] (-x^2)^n}{\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+1)!} \frac{1}{(2n+2m+3)!!} (-x^2)^n} \\ &= 2m+3 + \frac{\sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n-1)!} \frac{(-x^2)^n}{(2n+2m+3)!!}}{\sum_{n=0}^{\infty} \frac{(2n+1)!!}{(2n+1)!} \frac{(-x^2)^n}{(2n+2m+3)!!}} \\ &= 2m+3 - \frac{x^2 \varphi_{m+2}(x)}{\varphi_{m+1}(x)} \\ &= 2m+3 - \frac{x^2}{\frac{\varphi_{m+1}(x)}{\varphi_{m+2}(x)}}. \end{aligned}$$

□

$\tan x$ 可以利用上面构造的 $\varphi_m(x)$ 部分展开:

$$\begin{aligned}\tan x &= x \frac{\frac{\sin x}{x}}{\cos x} = \frac{x}{\sum_{n=0}^{\infty} \frac{(-x^2)^n}{(2n)!}} \\ &= \frac{x}{1 + \frac{\sum_{n=1}^{\infty} \frac{(-x^2)^n}{(2n-1)!(2n+1)}}{\sum_{n=0}^{\infty} \frac{(-x^2)^n}{(2n+1)!}}} \\ &= \frac{x}{1 + \frac{-x^2}{\frac{\varphi_0(x)}{\varphi_1(x)}}}.\end{aligned}$$

反复利用引理 2, 得

$$\tan x = \frac{x}{1 - \frac{x^2}{3 - \frac{x^2}{5 - \frac{x^2}{\ddots - \frac{x^2}{2m+1 - \frac{\varphi_m(x)}{\varphi_{m+1}(x)}}}}}}$$

引理 3

$f_m(x)$ 在 $[-1, 1]$ 上始终有定义, 即分母不为 0.



证明 对 $x_0 \in [-1, 1]$

$$f_m(x_0) = \frac{x_0}{1 - \frac{x_0^2}{3 - \frac{x_0^2}{5 - \frac{x_0^2}{\ddots - \frac{x_0^2}{2m-1 - \frac{x_0^2}{2m+1}}}}}}$$

令

$$m \geq 2, \quad \xi_{m+1} = 2m + 1, \quad \xi_m = 2m - 1 - \frac{x_0^2}{2m + 1} > 1,$$

$$\xi_{m-1} = 2m - 3 - \frac{x_0^2}{\xi_m} > 1,$$

利用

$$\xi_i = 2i - 1 - \frac{x_0^2}{\xi_{i+1}} > 0, i \geq 1,$$

递推可以归纳地证明

$$i \geq 2\xi_i > 1, \xi_1 = 1 - \frac{x_0^2}{\xi_2} > 0.$$

ξ_1, \dots, ξ_{m+1} 为 $f_m(x_0)$ 所表示的连分式中的所有分母, 均不为 0.

注 $m = 1$ 时 $f_m(x_0) = \frac{x_0}{1 - \frac{x_0^2}{3}}$ 分母也不为 0, 但对于函数列的极限而言 $m \geq 2$ 的情况足矣.

下面来证明本节的定理:

证明 为了方便, 仍用 ξ_1, \dots, ξ_{m+1} 表示 $f_m(x_0)$ 的分母.

$$|\tan x - f_m(x)| = \left| \frac{x}{1 - \frac{\frac{x^2}{\varphi_1(x)}}{\xi_2}} - \frac{x}{1 - \frac{x^2}{\xi_2}} \right|,$$

$$|x| \leq 1, \frac{5}{6} \leq \varphi_0(x) \leq 1, \frac{3}{10} \leq \varphi_1(x) \leq \frac{1}{3}, 0 \leq \frac{x_2}{\frac{\varphi_0(x)}{\varphi_1(x)}} \leq \frac{2}{5},$$

$$\xi_2 = 3 - \frac{x_0^2}{\xi_1} > 2, \quad 0 < \frac{x^2}{\xi_2} < \frac{1}{2},$$

令 $y(t) = \frac{|x|}{1-t}$, 利用 Lagrange 中值定理:

存在 η_1 介于 $\frac{x^2}{\frac{\varphi_0(x)}{\varphi_1(x)}}$ 与 $\frac{x^2}{\xi_2}$ 之间, 显然 $\eta_1 \in (0, \frac{1}{2})$,

原式

$$= \frac{|x|}{|1 - \eta_1|^2} \left| \frac{x^2}{\frac{\varphi_0(x)}{\varphi_1(x)}} - \frac{x^2}{\xi_2} \right|$$

$$= \frac{|x|^3}{|1 - \eta_1|^2} \left| \frac{1}{3 - \frac{x^2}{\frac{\varphi_1(x)}{\varphi_2(x)}}} - \frac{1}{3 - \frac{x^2}{\xi_3}} \right|.$$

反复利用 Lagrange 中值定理可得: 存在 η_{k+1} 介于 $\frac{x^2}{\xi_{k+2}}$ 与 $\frac{x^2}{\frac{\varphi_k(x)}{\varphi_{k+1}(x)}}$ 之间, 显然 $0 <$

$\eta_{k+1} < 1$,

使得

$$\left| \frac{1}{2k+1 - \frac{x^2}{\frac{\varphi_k(x)}{\varphi_{k+1}(x)}}} - \frac{1}{2k+1 - \frac{x^2}{\xi_{k+2}}} \right| = \frac{1}{(2k+1 - \eta_{k+1})^2} \left| \frac{x^2}{\frac{\varphi_k(x)}{\varphi_{k+1}(x)}} - \frac{x^2}{\xi_{k+2}} \right|,$$

原式

$$= |x| \frac{|x|^2}{(1-\eta_1)^2} \cdots \frac{|x|^2}{(2m-1-\eta_m)^2} \left| \frac{\varphi_m(x)}{\varphi_{m-1}(x) - \frac{1}{2m+1}} \right|$$

$$\leq 4 \cdot \frac{1}{4} \cdot \frac{1}{16} \cdots \frac{1}{(2m-2)^2} \cdot \frac{1}{m}.$$

显然 $m \rightarrow \infty$ 时, 这个值以 0 为极限。

注 就最后放缩的结果看得出来这个连分函数展开计算 $\tan x$ 收敛速度很快。

它有一个更重要的应用:

考虑 $u, v \in \mathbb{N}^*, u \leq v$

$$\tan\left(\frac{u}{v}\right) = \frac{\frac{u}{v}}{1 - \frac{\frac{u^2}{v^2}}{3 - \frac{\frac{u^2}{v^2}}{5 - \frac{\frac{u^2}{v^2}}{\ddots - \frac{\frac{u^2}{v^2}}{2m+1 - \frac{\frac{\varphi_m(u/v)}{\varphi_{m+1}(u/v)}}}}}}}}$$

$$= \frac{u}{v - \frac{u^2}{3v - \frac{u^2}{5v - \frac{u^2}{\ddots - \frac{u^2}{(2m+1)v - \frac{u^2/v \varphi_{m+1}(u/v)}{\varphi_m(u/v)}}}}}}}}.$$

由引理

$$\frac{\varphi_{m+1}(x)}{\varphi_m(x)} \leq \frac{1}{2m + \frac{5}{2}},$$

可得:

$$Mm \geq M, \frac{u^2}{v} \frac{\varphi_{m+1}\left(\frac{u}{v}\right)}{\varphi_m\left(\frac{u}{v}\right)} < 1.$$

下面来证明 $\frac{u^2}{v} \frac{\varphi_{M+1}(u/v)}{\varphi_M(u/v)}$ 不是有理数:

证明 若

$$\frac{p_1}{p_0} = \frac{u^2 \varphi_{M+1}(u/v)}{v \varphi_M(u/v)} = \frac{u^2/v}{2M+3 - \frac{u^2 \varphi_{M+2}(u/v)}{v^2 \varphi_{M+1}(u/v)}}, p_0, p_1 \in N^*,$$

$$\frac{(2M+3)vp_1 - u^2p_0}{p_1} = \frac{u^2 \varphi_{M+2}(u/v)}{v \varphi_{M+1}(u/v)}, p_2 = (2M+3)vp_1 - u^2p_0,$$

$$0 < \frac{u^2 \varphi_{M+1}(u/v)}{v \varphi_M(u/v)} < 1, p_1 < p_0,$$

又有

$$0 < \frac{u^2 \varphi_{M+2}(u/v)}{v \varphi_{M+1}(u/v)} < 1, 0 < p_2 < p_1,$$

$$p_{k+1} = (2M+1+2k)p_k - u^2p_{k-1},$$

归纳的证明:

$$\frac{p_{k+1}}{p_k} = \frac{u^2 \varphi_{M+k+1}(u/v)}{v \varphi_{M+k}(u/v)} \in (0, 1),$$

从而 $p_{k+1} < p_k$, 且 $\{p_k\}$ 是正整数列. 不能存在 $\{p_k\}$ 这样严格单调减的正整数列. 因此假设错误, $\frac{u^2 \varphi_{M+1}(u/v)}{v \varphi_M(u/v)}$ 是无理数.

这显然意味着 $\tan\left(\frac{u}{v}\right)$ 不是有理数, 这样就得到了如下定理:

定理 2

若 $q \in [-1, 1]$, 那么 q 与 $\tan q$ 至多有一个是有理数。



它有一个重要的推论, 即:

推论 1

π 是无理数。



事实上, 这是显然的, 因为 $\frac{\pi}{4}$ 与 $\tan \frac{\pi}{4} = 1$ 不能同时为有理数。

上面这个证明是最早给出的 π 是无理数的证明之一. 比起主要定理而言, $\tan x$ 的部分展开显然更具理论价值, 为了把无理数的结果推到更大的区间上去, 需要对引理进行推广, 另外主要定理推广到复平面上也不失为一种尝试。

part 3

匠心獨運



数值积分浅谈

作者：衣孝儒 2019级 信息与计算科学班

1 从定积分谈起

在正式开始这篇文章之前，我们先来回顾一下定积分的定义：

定义 1

对于一个定义于 $[a, b]$ 上的黎曼可积的函数 f ，其定积分的定义式为：

$$\int_a^b f(x)dx = \lim_{\|T\| \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i. \quad (1)$$



简单概括来说就是：分割、求和、取极限。

我们知道，定积分作为广泛应用于生产生活实践的一种数学工具，在各个领域中都显示出了极大的威力，为了解决定积分问题，我们发展出了各种巧妙的积分方式，扩展了我们求解积分的边界。但是在现实运用中，需要进行积分的式子本身往往十分复杂，其精确的积分结果往往具有十分复杂的表达形式，更进一步地，大量的积分表达式甚至不存在初等表示方法。

与此同时，在现实生产生活运用中，我们需要的往往又并不是定积分完全精确的代数表示，而是一个相对精确的数值表示，因此，很多人就开始思考，我们如何绕过代数的方法，用数值的方式求解定积分，而这一想法直接催生了数值积分这一领域。在本文中，我将用尽可能简洁的语言介绍各种不同的数值积分手段，为大家建立起数值积分的基本思想。

2 定积分的启发

应该有部分读者从定积分的定义可以自然地想到一种数值积分的手段，在定积分的定义式中，只要我们将区间 $[a, b]$ 划分得足够细密，那么右侧的表达式就可以自然地作为原始积分的一个良好近似，从而得出可用的积分结果。简单起见，我们采用等距分割，并且在每个小区间内采用右端点作为函数值的近似点，这样我们就得到了数值积分的一个可用的公式：

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} \sum_{k=1}^n f(a + k\Delta x) \Delta x \approx \sum_{k=1}^N f(a + k\Delta x) \Delta x. \quad (2)$$

按照数值积分的分类，这是一种复化长方形插值公式。事实上，数值积分的方法远不止于此，本文主要讲解插值型积分公式。按照插值点是否可以自由选取，插值型积分公式分为 Newton-Cotes 积分公式和 Gauss 积分公式两大类，而在 Newton-Cotes 积分公式的实际运用中，又常常使用复化的积分公式。这里只做一个简单的分类，下面我们将展开具体的讲解。

3 插值型积分公式

插值型积分公式，顾名思义，就是在求积区间 $[a, b]$ 上对被积函数 $f(x)$ 做插值，利用得到的插值多项式 $p_n(x)$ 代替被积函数 $f(x)$ 导出的公式，这一类求积公式中的求积节点 $\{x_k\}_{k=0}^n$ 就是对原函数做插值时的插值节点，所以这类积分公式称为插值型求积公式。关于插值的具体知识，在此就不再赘述。

插值型积分公式分为两类，Newton-Cotes 积分公式与 Gauss 积分公式，对于不同的区间 $[a, b]$ ，其都可通过变量线性变换变换至 $[-1, 1]$ 内，从而写为统一的形式：

$$\int_{-1}^1 f(x)dx \approx \sum_{k=0}^N A_k f(x_k). \quad (3)$$

其中 A_k 为第 k 个积分点处函数值对应的权重，为待定系数。下文为讨论及叙述的方便，在给出具体积分点权重及位置时均以 $[-1, 1]$ 区间为例展开讨论，其他区间的积分公式可以直接通过变量代换获得。两种积分方法的主要区别在于积分点可不可以自由选取，更大的自由度意味着更高的积分精度，在具体比较二者区别之前，我们先引入代数精度的概念：

定义 2

若对任意不高于 n 次的多项式 $p_n(x) \in \mathbf{P}_n[a, b]$ ，求积公式的误差均为 0，而对 $n+1$ 次项的积分误差不为零，则称该求积公式具有 n 次代数精度。

但是这个定义在实际使用中并不方便，为此，我们给出等价定义如下：

定义 3

若对函数 $f(x) = 1, x, x^2, \dots, x^n$ ，求积公式精确成立，而对 $f(x) = x^{n+1}$ 求积公式的误差不为 0，则称其具有 n 次代数精度。

一般地，对于有 n 个积分点的数值积分，Newton-Cotes 积分公式中积分点已经事先指定无法自由选取，其最高只能实现 $n+1$ 阶代数精度。而对于 Gauss 积分公式，其积分点可以自由选取，更高的自由度带来了更高的代数精度，其最高可以实现 $2n+1$ 阶代数精度。下面我们将对这两种数值积分格式分别进行讲解。

4 Newton-Cotes 积分公式

Newton-Cotes 积分公式的核心特点是积分点已经事先选取好，唯一能调整的是每个积分点处的权重。不失一般性，也考虑到实际运用中最普遍的情况，我们此处只考虑求积区间 $[-1, 1]$ 被 n 等分的情形。此时设求积节点为 $\{x_k\}_{k=0}^n$ ，那么 $x_0 = -1, x_n = 1, x_j = a + jh, j = 0, 1, \dots, n; h = \frac{2}{n}$ ，那么我们的积分公式就变为：

$$\int_{-1}^1 f(x)dx \approx \sum_{k=0}^N A_k f(-1 + kh). \quad (4)$$

由上文所叙述的定义 3, 及 Newton-Cotes 积分公式的代数精度定理, 我们可以直接依次选取 $f(x) = 1, x, x^2, \dots, x^n$ 并逐个代入, 即可获得关于积分点权重的线性方程组, 求解方程组即可获得 Newton-Cotes 积分公式的不同阶次精度表达形式。在此我们将系数表格列出如下:

n	积分点权重 (由左至右)
1	$\frac{1}{2}, \frac{1}{2}$
2	$\frac{1}{6}, \frac{4}{6}, \frac{1}{6}$
3	$\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$
4	$\frac{7}{90}, \frac{16}{45}, \frac{2}{15}, \frac{16}{45}, \frac{7}{90}$
5	$\frac{19}{288}, \frac{25}{96}, \frac{25}{144}, \frac{25}{144}, \frac{25}{96}, \frac{19}{288}$

表 1: Newton-Cotes 积分公式系数表

可以看到, 随着区间等分份数的增大, 积分点的权重系数迅速复杂化, 并不方便实际使用, 另一方面, 随着等分份数的增大, 插值的精度也会有不同程度的下降, 具体地, 一般有如下两个影响插值精度的主要问题:

问题 1 函数性质本身不佳导致龙格现象。

问题 2 舍入误差导致的结果偏差。

接下来我们将对这两个问题分别展开进行一定的说明。

4.1 龙格现象

一般来说, 随着插值阶次的升高, 多项式对被插函数的性质往往逼近得越好。但是有些函数是例外, 它们随着插值多项式阶次的升高, 插值效果反而越差。一个典型的例子是:

$$f(x) = \frac{1}{1+x^2}. \quad (5)$$

我们有如下定理:

定理 1

上述函数 $f(x)$ 的等距节点插值函数 $p_n(x)$ 在 $[-r, r]$ 外不收敛, 其中 $r \approx 3.63$.



具体的证明我们这里略去不表, 有兴趣的同学可以查阅这本书: Analysis of Numerical Methods by Eugene Isaacson, Herbert Bishop Keller 275-278. 这里我们只是借用这一结论, 来支持我们后续的说明。究其原因, 是由于 $f(x)$ 这一函数虽然在实数域内连续, 但是在复平面内有奇点, 复平面奇点的奇异性传导到边缘, 影响了远处的拟合准确程度。我们把这类函数称为龙格函数, 这类函数在插值过程中出现的问题我们称之为龙格现象, 由于龙格现象的存在, 我们对很大一类函数都无法通过高次多项式直接插值获得。

4.2 勒贝格常数

另一方面,对于具有较好性质,能随着多项式阶次升高,插值效果不断增强的函数,存在着计算精度的问题。众所周知,对于绝大多数的插值问题,我们都需要通过计算机获取最终的结果,这里便存在着计算过程中的误差积累传播的问题。我们知道,在实数域中,我们已经获得了 $\sin x$ 一个非常精确的展开式:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \quad (6)$$

但是在浮点数体系中,这一精确展开式却会不断产生并放大误差,最终造成结果的不稳定。那么,如何刻画这种不稳定程度呢?我们提出了勒贝格常数的概念,这里提供其三个等价定义:

定义 4

对于原函数 $f(x)$ 与对应的插值多项式 $p_n^f(x)$, 设其插值点为 $\{x_i\}_{i=0}^n \subseteq [a, b]$, 我们定义其勒贝格常数为

$$A_n(x) = \sup_{f \neq 0} \frac{\|p_n^f\|_{C[a,b]}}{\|f\|_{C[a,b]}} = \sup_{|y_i| \leq 1} \|p_n^{\{y_i\}}\|_{C[a,b]} = \left\| \sum_{i=0}^n |l_i(x)| \right\|_{C[a,b]} \quad (7)$$



可以证明,三者之间彼此等价。在这一基础上对于函数 $\sin x$, 我们可以证明其勒贝格常数的一个估计为:

$$A_{2n}(x) \geq \frac{4^{n-2}}{n^2} \quad (8)$$

从这一估计可以看出,随着插值多项式次数的增大,函数 $\sin x$ 的误差迅速增大。

综合上述两段的讨论,我们可以得出结论:

结论 对于一般的函数插值问题,使用高次多项式插值会在精度上带来巨大的损失,所以在实际运用中,我们往往避免高阶插值,而更常使用低阶的 Newton-Cotes 积分公式。

但是如何保证低阶的 Newton-Cotes 积分公式依然有较好的精度呢?答案是复化的 Newton-Cotes 积分公式。

4.3 复化 Newton-Cotes 积分公式

从之前的讨论中,我们可以较为容易地获取低阶的 Newton-Cotes 积分公式, $n=1$ 时我们一般称之为梯形求积公式, $n=2$ 时我们一般称之为 Simpson 求积公式。如果采用 $I(f)$ 表示函数 $f(x)$ 在 $[a, b]$ 上的数值积分,那我们可以将公式分别写为:

$$n = 1, I(f) = \frac{b-a}{2}(f(a) + f(b)). \quad (9)$$

$$n = 2, I(f) = \frac{b-a}{6}(f(a) + 4f(\frac{a+b}{2}) + f(b)). \quad (10)$$

只要将积分区间本身进行细分,在细分后的每一个小区间上使用梯形求积公式或者 Simpson 求积公式,并将结果加和,就可以得到复化的梯形求积公式或者复化的 Simpson 求积公式。同样地,其他求积公式也可以用同样的方法进行复化,这里我们只列出等距情形的复化梯形

求积公式与复化 Simpson 求积公式:

$$I(f) = \frac{h}{2} (f(a) + 2 \sum_{i=1}^{n-1} f(a + ih) + f(b)). \quad (11)$$

$$I(f) = \sum_{j=1}^n \frac{h}{3} (f(x_{2j-2}) + 4f(x_{2j-1}) + f(x_{2j})). \quad (12)$$

复化的积分公式是一种低阶的积分手段, 可以避免高次插值带来的龙格现象及其他不稳定问题, 同时, 对于一般的分段多项式插值, 我们可以证明如下的误差估计公式:

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x - x_i). \quad (13)$$

这一公式说明, 对于任意分片多项式插值, 我们可以直接利用缩小分片区间的方法来减小误差, 从这一思路出发, 我们可以获得常见分片插值方法的误差估计。

定理 2

对于在区间 $[a, b]$ 上的分片线性插值, 我们有如下误差估计

$$\|f(x) - p(x)\|_{C[a,b]} \leq \frac{h^2}{8} \|f''\|_{C[a,b]}. \quad (14)$$



由此可见, 在原函数性质较好的情况下, 只要缩小每个分片区间 $[a, b]$ 的长度, 即可获得更好的插值效果。

5 Gauss 积分公式的构建

Gauss 积分公式的核心特点是积分点事先未知, 可以按照需求自由选取, 每个积分点的位置与积分点处的权重都可以自由选取。此时设求积节点为 $\{x_k\}_{k=0}^n$, 那么我们的积分公式就变为:

$$\int_{-1}^1 f(x) dx \approx \sum_{k=0}^N A_k f(x_k) \quad (15)$$

为了求解出积分点权重与积分点位置, 我们依次选取 $f(x) = 1, x, x^2, \dots, x^{2n+1}$ 并逐个代入, 即可获得关于积分点权重与积分点位置的非线性方程组, 对其进行求解即可获得积分点的位置与对应权重。在此, 我们以一个例题辅助说明:

5.1 问题呈现

对定积分 $\int_{-1}^1 f(x) dx$ 求出具有 3 次、5 次、7 次代数精度的高斯型积分公式

$$\int_{-1}^1 f(x) dx \approx \sum_{i=0}^n A_i f(x_i)$$

用表格列出其系数 $A_i, x_i, i = 0, 1, \dots, n$, 并给出数值算例积分 $\int_{-1}^1 e^{x^2} dx$.

5.2 问题求解

利用高斯积分的基本公式

$$\int_{-1}^1 f(x)dx \approx \sum_{i=0}^n A_i f(x_i)$$

只要我们依次将 $1, x, x^2, x^3, \dots, x^n$ 分别在 $n = 3, 5, 7$ 时代入上述方程组，方程个数等于未知数个数，以 $n = 5$ 时为例可以获得方程组如下：

$$\begin{aligned} A_1 + A_2 + A_3 &= 2 \\ A_1 x_1 + A_2 x_2 + A_3 x_3 &= 0 \\ A_1 x_1^2 + A_2 x_2^2 + A_3 x_3^2 &= \frac{2}{3} \\ A_1 x_1^3 + A_2 x_2^3 + A_3 x_3^3 &= 0 \\ A_1 x_1^4 + A_2 x_2^4 + A_3 x_3^4 &= \frac{2}{5} \\ A_1 x_1^5 + A_2 x_2^5 + A_3 x_3^5 &= 0 \end{aligned}$$

对此方程组恰当选取初值进行求解即可得到系数 $A_i, x_i, i = 0, 1, \dots, n$ ，列表展示如下：(小数保留至四位)

	i=1	i=2
A_i	1.0000	1.0000
x_i	-0.5774	0.5774

表 2: 3 次精度的高斯型积分系数

	i=1	i=2	i=3
A_i	0.5556	0.8889	0.5556
x_i	-0.7746	0.0000	0.7746

表 3: 5 次精度的高斯型积分系数

	i=1	i=2	i=3	i=4
A_i	0.6521	0.6521	0.3479	0.3479
x_i	0.3400	-0.3400	0.8611	-0.8611

表 4: 7 次精度的高斯型积分系数

5.3 数值算例

本节我们选取一个具体的积分问题来验证上述求解所得系数的合理性。

$$\int_{-1}^1 e^{x^2} dx$$

将其代入求解格式

$$\int_{-1}^1 f(x)dx \approx \sum_{i=0}^n A_i f(x_i)$$

可得三种计算格式积分值与 MATLAB 自带的积分功能值的比较为：

	3次精度	5次精度	7次精度	MATLAB 积分解
$\int_{-1}^1 e^{x^2} dx$	2.791224	2.913465	2.924540	2.925303

可见高斯积分具有非常优秀的精度，且相较于 Newton-Cotes 积分公式，达到相同精度所需的运算量减少了一半，因此其在大型仿真问题中往往是最常用的数值积分手段。

6 正交多项式与高斯节点

在上面的讨论中，看似已经给出了高斯节点完整的权重与位置计算手段，在实际应用中也可以查表直接获取高斯节点的信息，但事实上，细心的读者可能已经发现，在 5.2 中我们获得的方程组是一个非线性方程组，在实际的求解中其收敛性往往十分复杂，往往要通过大量的尝试才能找到一个能正确收敛到真实解的初值，并且随着维数的升高，其寻找恰当初值的难度成倍增加。因此，接下来我们将介绍一种将非线性方程组线性化的方法，以便快速求出高斯积分的积分点和各点的权重。

6.1 正交多项式

在数学中有一类非常重要的多项式，我们称之为正交多项式，其定义如下：

定义 5

设 $\phi_k(x), k = 0, 1, 2, \dots$ 为 k 次多项式，若 $\int_a^b \rho(x)\phi_i(x)\phi_j(x)dx = \delta_{ij}$ ，则称 $\{\phi_k(x)\}_{k=0}^{\infty}$ 为带权 $\rho(x)$ 正交多项式族。带权的 gauss 积分定义为：

$$\int_a^b \rho(x)f(x)dx \approx \sum_{i=0}^n A_i f(x_i) \quad (16)$$



同时，我们有两个与此相关的重要定理：

定理 3

x_0, x_1, \dots, x_n 为高斯点等价于 $w(x) = \prod_{k=0}^n (x - x_k)$ 与任意次数不大于 n 的多项式 $P_n(x)$ 带权正交。



定理 4

正交多项式族 $\{\phi_0, \phi_1, \dots, \phi_n, \dots\}$ 有性质：任意次数不大于 n 的多项式 $P(x)$ 必与 ϕ_{n+1} 正交。



定理 3 将高斯点的求解等价于与任意次数不大于 n 的多项式 $P_n(x)$ 带权正交的多项式的求解，而定理 4 进一步将这一多项式直接确定为 ϕ_{n+1} ，也就是说，这两个定理将高斯积分公式的节点计算从非线性方程组里剥离出来单独计算，将复杂的非线性方程组求解简化为先求解多项式求根问题，再进行线性方程组求解，极大地减少了求解的难度。

由于线性方程组已经产生出一整套高效快速的求解算法，已经有许多成熟的求解包可以使用，我们现在关注的主题就转移到正交多项式 ϕ_{n+1} 的确定及其根求解，事实上，我们完全可以对任意给定的权函数 $\rho(x)$ 及初始函数 $\phi_0(x)$ 通过正交多项式的定义递推得到 ϕ_{n+1} 的表达式，在这里我们以一个简例辅助说明：

6.2 数值算例

例题 1 求解如下数值积分具有三阶精度的高斯积分格式：

$$\int_0^1 \sqrt{x}f(x)dx$$

解 三阶精度的高斯积分格式我们可以表示为：

$$\int_0^1 \sqrt{x}f(x)dx \approx A_0f(x_0) + A_1f(x_1)$$

根据定理 3、4，我们只需确定正交多项式 $\phi_2(x)$ 即可。我们可设

$$\phi_0(x) = 1$$

$$\phi_1(x) = x + a$$

$$\phi_2(x) = x^2 + bx + c$$

利用正交多项式的内积性质，我们可以获得：

$$(\phi_0, \phi_1) = 0 \Rightarrow \int_0^1 \sqrt{x}(x+a)dx = 0$$

$$(\phi_0, \phi_2) = 0 \Rightarrow \int_0^1 \sqrt{x}(x^2 + bx + c)dx = 0$$

$$(\phi_1, \phi_2) = 0 \Rightarrow \int_0^1 \sqrt{x}(x+a)(x^2 + bx + c)dx = 0$$

从上述方程组中我们可以求得 $a = -\frac{3}{5}, b = -\frac{10}{9}, c = \frac{5}{21}$ ，从而我们有

$$\phi_2(x) = x^2 - \frac{10}{9}x + \frac{5}{21}$$

求解其根我们就可以获得两个高斯积分点 x_0, x_1 的值，再将 $f(x) = 1, x$ 依次代入，求解线性方程组即可获得完整的积分表达式。

7 结语

本文我们从定积分的定义出发，介绍了两大类插值型积分公式，Newton-Cotes 积分公式和高斯积分公式。Newton-Cotes 积分公式中我们介绍了龙格现象和舍入误差两类主要的误差来源，并给出了复化 Newton-Cotes 积分公式作为比较可行合理的精度提升方案；高斯积分公式中我们介绍了高斯积分节点和权重的确定手段，介绍了正交多项式在确定高斯节点中的应用，并给出了具体的数值算例。总的来说给出了数值积分一个初步而概括的介绍。文章避免了复杂的推导，侧重于脉络的梳理与介绍，希望这篇文章能让读者产生对计算数学的兴趣。

微分方程数值解的稳定性

作者：卢小石 2019 级 信息与计算科学班

摘要

我们在利用数值格式数值求解偏微分方程时，自然而然地会提出这样的问题：什么样的数值格式是可以用的？如何确定数值仿真结果可以作为精确解的近似？为此我们要进行一些理论上的分析，介绍数值格式收敛性、相容性、稳定性的定义。并且我们将主要关注求解包含时间的偏微分方程的数值格式的稳定性，并给出一些简单的例子。

关键词：有限差分 冯诺依曼稳定性分析

1 背景知识

微分方程被广泛应用于各个领域的建模，在电磁学、热力学、经济、生物等领域都能看到微分方程组的身影。因此求解微分方程组是非常重要的，但是解析解通常难以获得，因此我们尝试通过其他方法寻找近似解。其中，利用离散的数值点来近似表示微分方程的解是一种常见的方法，不仅便于实现，也非常易于理解，这种方法也有不错的精确度。

通常，求解线性微分方程的有限差分方法 (finite difference method) 可以分为以下四个步骤：

- 网格划分：给定一个目标区间 (a, b) ，以及正整数 n ，我们产生如下的网格划分

$$x_i = a + ih, \quad i = 0, 1, \dots, n, \quad h = \frac{1}{n}.$$

参数 n 可以根据精确度的需求来进行选择；

- 用有限差分格式来表示微分项：我们在这里介绍最常用的几种有限差分格式

$$\text{向前差分: } u'(x_i) = \frac{u(x_i) - u(x_{i-1}))}{h} + O(h)$$

$$\text{向后差分: } u'(x_{i+1}) = \frac{u(x_{i+1}) - u(x_i)}{h} + O(h)$$

$$\text{中心差分: } u'(x_i) = \frac{u(x_{i+1}) - u(x_{i-1}))}{2h} + O(h^2)$$

$$\text{二阶微分的中心差分: } u''(x_i) = \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2} + O(h^2)$$

我们舍去末尾的截断误差就可以得到相应的格式。在解的取值未知的格点反复使用相应的格式，并且用 U_i 代替 $u(x_i)$ 作为未知变量，我们可以得到一个线性方程组；

- 编程求解上述线性方程组；
- 误差分析。

如果我们想进行时间方向上的有限差分，我们只需要仿照以上的步骤，将空间区间 (a, b) 替换为时间区间 $(0, T)$ ，将空间步长 h 替换为时间步长 Δt 即可。接下来我们主要讨论误差分析部分 (对某个固定时间点进行讨论)，为此我们先给出一些相关定义。

定义 1

令 $\mathbf{U} = [U_1, U_2, \dots, U_n]^T$ 表示时间 T 时通过有限差分方法求解得到的近似解， $\mathbf{u} = [u(x_1, T), \dots, u(x_n, T)]^T$ 表示精确解。称 $\mathbf{E} = \mathbf{U} - \mathbf{u}$ 为全局误差向量。我们通常选定某个向量范数来度量全局误差。



定义 2

若 $\lim_{h \rightarrow 0} \|E\| = 0$ ，则称该有限差分格式是收敛的。



定义 3

设 $u(x, t)$ 是时间 t 时微分方程的精确解。令 P 表示在某个线性微分方程中作用于 u 的微分算子，令 P_h 表示作用于 u 的对应的有限差分算子。我们定义局部截断误差如下

$$T(x, t) = P_h u - P u.$$

我们称一个有限差分格式是相容的，如果

$$\lim_{h \rightarrow 0, \Delta t \rightarrow 0} T(x, t) = \lim_{h \rightarrow 0, \Delta t \rightarrow 0} (P_h u - P u) = 0.$$



为了帮助大家更好地理解局部截断误差和相容性的概念，我们给出一个例子。

例题 1 我们考虑如下的微分方程

$$u_x(x, t) + u_t(x, t) = f(x, t).$$

在这个例子中

$$P = \frac{\partial}{\partial x} + \frac{\partial}{\partial t}, \quad P u = u_x + u_t = f.$$

对于空间方向的一阶微分，我们使用中心差分来近似；对于时间上的一阶微分，我们用向前差分来近似。于是

$$P_h u = \frac{u(x+h, t) - u(x-h, t)}{2h} + \frac{u(x, t+\Delta t) - u(x, t)}{\Delta t},$$

从而

$$T(x, t) = \frac{u(x+h, t) - u(x-h, t)}{2h} + \frac{u(x, t+\Delta t) - u(x, t)}{\Delta t} - f(x, t) = O(h^2) + O(\Delta t).$$

从定义不难得出，该有限差分格式是相容的。

定义 4

我们将在有限差分方法的第二步得到的线性方程组表示为 $AU = F$ 。我们称一个有限差分格式对于线性边值问题是稳定的，如果系数矩阵 A 是可逆的，并且存在与 h 无关的常数 C 和 h_0 ，使得

$$\|A^{-1}\| \leq C, \quad \forall h \in (0, h_0).$$



注 如果待求解函数中有时间变量，则对每一固定时间点，对空间方向进行离散都可以得到这样的线性方程组。

由于精确解并不总是已知的，因此从定义入手来分析有限差分方法的收敛性常常会遇到困难，接下来我们给出一个判断收敛性的重要定理，但出于篇幅考虑不给出证明，感兴趣的读者可以自行查询相关资料。

定理 1

一个相容并且稳定的有限差分格式是收敛的。



通常，有限差分格式的相容性比稳定性更容易证明，因此在接下来的讨论中，我们着重讨论有限差分格式的稳定性判别方法，并且我们的讨论范围将限制在包含时间微分的微分方程。

2 问题引入

为了更加直观地展示数值解的效果，我们先给出一个简单的例子。

$$u_t + au_x = 0, \quad 0 < x < 1$$
$$u(0, x) = x, \quad u(t, 0) = -at.$$

该问题的精确解为 $u(t, x) = x - at$ 。我们取 $a = 1$ ，并且使用迎风格式 (upwind method)：

$$\frac{U_j^{k+1} - U_j^k}{\Delta t} = -a \frac{U_j^k - U_{j-1}^k}{h}$$
$$U_j^{k+1} = (1 - \mu)U_j^k + \mu U_{j-1}^k, \quad \mu = \frac{a\Delta t}{h}.$$

接下来我们先看看数值求解的结果。

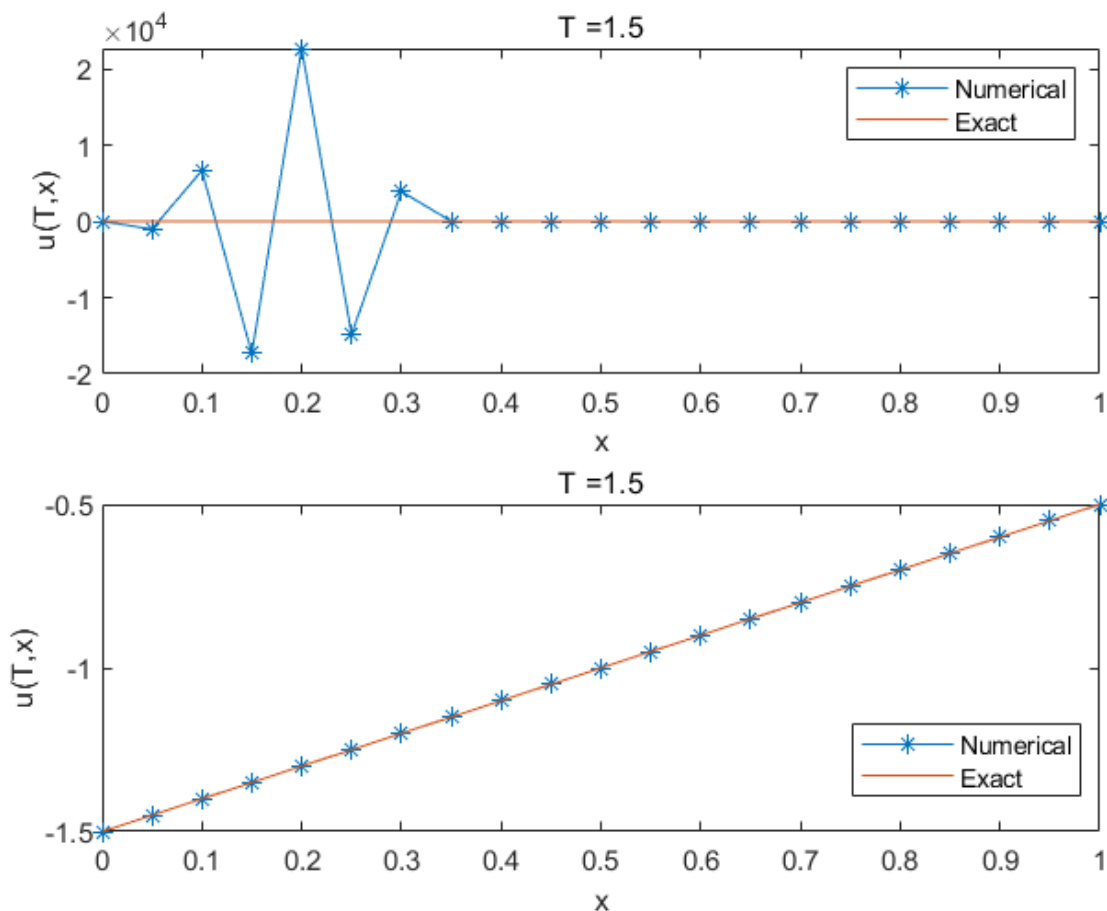


图 1.1: 迎风格式求解传输方程

两种情况我们均取 $h = 0.05$. 第一种情况下, 我们令 $\Delta t = 5h/a$, 第二种情况下我们令 $\Delta t = h/(2a)$. 我们可以很明显的看到当时间前进到 1.5s 时, 第一种情况出现了很明显的误差, 而第二种情况数值解则很好地逼近了精确解。

这是为什么呢? 实际上是由于第二种时间步长的选取可以令该格式符合稳定性要求。稳定性的判别方法有许多种, 接下来我们介绍冯诺依曼稳定性检验 (The von Neumann Stability Analysis) 方法。

3 离散的傅里叶变换

在正式进入稳定性的讨论之前, 我们需要简单的了解一下稳定性分析所需要使用的工具——傅里叶变换以及它的离散形式。

定义 5

我们考虑 $u(x) \in L^2(-\infty, +\infty)$, 即 $\int_{-\infty}^{+\infty} u^2 dx < \infty$ 或 $\|u\| < \infty$. 我们定义 $u(x)$ 的傅里叶变换为

$$\hat{u}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-i\omega x} u(x) dx.$$

其中 $i = \sqrt{-1}$ 为虚数单位。对应的傅里叶逆变换为

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega x} \hat{u}(\omega) d\omega.$$



有一个重要的关系我们需要在此强调

$$\|\hat{u}\|_2 = \|u\|_2, \text{ i.e. } \int_{-\infty}^{+\infty} |\hat{u}|^2 d\omega = \int_{-\infty}^{+\infty} |u|^2 dx,$$

该等式关系被称为帕西瓦尔等式 (Parseval's relation).

我们知道积分可以通过分割、求和、取极限来计算, 于是我们通过将积分符号转化为求和符号来得到傅里叶变换和逆变换对应的离散形式。

定义 6

若用 $\dots, v_{-1}, v_0, v_1, \dots$ 来表示连续函数 $v(x)$ 在离散点 $x_i = ih$ 的取值, 则离散的傅里叶变换可以被定义如下

$$\hat{v}(\xi) = \frac{1}{\sqrt{2\pi}} \sum_{j=-\infty}^{\infty} h e^{-i\xi j h} v_j.$$



通过如上定义, 我们不难证明 $\hat{v}(\xi)$ 是一个以 $2\pi/h$ 为周期的周期函数, 所以我们可以仅仅关注区间 $[-\pi/h, \pi/h]$, 由此我们有如下定义。

定义 7

离散的傅里叶逆变换定义如下

$$v_j = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{i\xi j h} \hat{v}(\xi) d\xi.$$



定义 8

为了方便后续定理叙述, 我们引入离散范数如下

$$\|\mathbf{v}\|_h = \sqrt{\sum_{j=-\infty}^{\infty} v_j^2 h}$$



我们需要指出的是, 帕西瓦尔等式在离散场合仍然成立, 即

$$\|\hat{v}\|_h^2 = \int_{-\pi/h}^{\pi/h} |\hat{v}(\xi)|^2 d\xi = \sum_{j=-\infty}^{\infty} |v_j|^2 h = \|\mathbf{v}\|_h^2.$$

4 冯诺依曼稳定性讨论

我们用 h 表示空间的划分, 用 Δt 表示时间上的划分, 同时我们用上标表示时间水平的变化, 用下标表示空间水平的变化。对于包含时间微分的微分方程有限差分格式, 我们给出一种新的稳定性定义。

定义 9

若对任意正的时间 T , 都存在与时间划分 Δt 和空间划分 h 均无关的整数 J 和常数 C_T 使得

$$\|\mathbf{v}^n\|_h \leq C_T \sum_{j=0}^J \|\mathbf{v}^j\|_h,$$

对任意的 n 满足 $0 \leq n\Delta t \leq T$ 并且 $(\Delta t, h) \in \Gamma$ 均成立, 则称有限差分格式 $P_{\Delta t, h} v_j^k = 0$ 在区域 Γ 是稳定的。



我们要指明的是, 这里定义的稳定性通常与 $Pu = f$ 中的 f 无关。根据上述定义我们很容易得到一个稳定性的充分条件。

定理 2

如果 $\|\mathbf{v}^{k+1}\|_h \leq \|\mathbf{v}^k\|_h$ 对任意的非负整数 k 均成立, 那么对应的有限差分格式是稳定的。♥

接下来为了避免理论叙述过于抽象, 我们借助一个例子展开, 来定义增长因子并且展示冯诺依曼稳定性分析的步骤。

例题 2 对于一阶微分方程 $u_t + au_x = 0$, 我们对时间一阶偏导数用向前差分格式 (forward FD method), 对空间一阶偏导数用中心差分格式 (central FD method), 于是得到

$$\frac{U_j^{k+1} - U_j^k}{\Delta t} + \frac{a}{2h}(U_{j+1}^k - U_{j-1}^k) = 0,$$

$$U_j^{k+1} = U_j^k - \mu(U_{j+1}^k - U_{j-1}^k), \quad \mu = \frac{a\Delta t}{2h}.$$

我们借助离散的傅里叶逆变换来表示上述格式中出现的各项

$$U_j^k = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{i\xi jh} \hat{U}^k(\xi) d\xi,$$

$$U_{j+1}^k = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{i\xi(j+1)h} \hat{U}^k(\xi) d\xi = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{i\xi h} e^{i\xi jh} \hat{U}^k(\xi) d\xi,$$

$$U_{j-1}^k = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{-i\xi h} e^{i\xi jh} \hat{U}^k(\xi) d\xi,$$

$$U_j^{k+1} = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{i\xi jh} \hat{U}^{k+1}(\xi) d\xi,$$

我们将上述表达带入到有限差分格式中, 还可以得到关系式

$$U_j^{k+1} = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{i\xi jh} (1 + \mu(e^{i\xi h} - e^{-i\xi h})) \hat{U}^k(\xi) d\xi.$$

离散的傅里叶变换唯一，于是通过比较最后两式，我们可以得到

$$\hat{U}^{k+1}(\xi) = (1 + \mu(e^{i\xi h} - e^{-i\xi h})) \hat{U}^k(\xi) = g(\xi) \hat{U}^k(\xi),$$

其中

$$g(\xi) = 1 + \mu(e^{i\xi h} - e^{-i\xi h})$$

被称为增长因子 (growth factor)。如果 $|g(\xi)| \leq 1$ ，我们可以得到 $|\hat{U}^{k+1}| \leq |\hat{U}^k|$ ，于是由离散的帕西瓦尔等式即有 $\|\hat{U}^{k+1}\|_h \leq \|\hat{U}^k\|_h$ ，因而该有限差分格式就是稳定的。

接下来我们来计算检验增长因子的模。

$$\begin{aligned} g(\xi) &= 1 + \mu(e^{i\xi h} - e^{-i\xi h}) \\ &= 1 - 2\mu i \sin(h\xi) \\ |g(\xi)|^2 &= 1 + 4\mu^2 \sin^2(h\xi) \geq 1. \end{aligned}$$

这和我们的预想相悖。但我们注意到 FW-CT 差分方法满足 $\mathbf{U}^{k+1} = f(\mathbf{U}^k, \mathbf{U}^{k+1})$ (满足该式的有限差分方法称为单步时间步进法)，下面的定理给我们提供了一个判断稳定性的简单的方式。

定理 3

令 $\theta = h\xi$. 单步有限差分格式 (常系数) 是稳定的，当且仅当存在常数 K (与 $\theta, \Delta t, h$ 无关) 以及时间空间的步进上限 Δt_0 和 h_0 使得

$$|g(\theta, \Delta t, h)| \leq 1 + K\Delta t, \quad \forall \theta, \forall h \in (0, h_0).$$

如果 $|g(\theta, \Delta t, h)|$ 与 h 和 Δt 无关，那么上述稳定性条件可以用下式代替

$$|g(\theta)| \leq 1.$$



由该定理知，之前讨论的 FW-CT 差分方法是无条件不稳定的。

上面对增长因子的讨论有些繁杂，我们通过观察可以对冯诺依曼稳定性检验做出一些简化，通常可以包含以下四个步骤：

1. 令 $U_j^k = e^{ijh\xi}$ 并在有限差分格式中进行相应替换；
2. 将 U_j^{k+1} 表示为 $U_j^{k+1} = g(\theta)e^{ijh\xi}$ ；
3. 整理求解 $g(\theta)$ 的表达式，计算是否或者何时 $|g(\theta)| \leq 1$ ；
4. 但是需要注意，如果存在某个 θ 使得 $|g(\theta)| > 1$ ，那么该方法是不稳定的。

下面我们使用一个例子来说明简化的冯诺依曼稳定性检验。

例题 3 热方程 $u_t = \beta u_{xx}$ 的向前欧拉方法 (FW-CT) 为

$$\begin{aligned} \frac{U_i^{k+1} - U_i^k}{\Delta t} &= \beta \frac{U_{i-1}^k - 2U_i^k + U_{i+1}^k}{h^2}, \quad i.e. \\ U_i^{k+1} &= U_i^k + \mu (U_{i-1}^k - 2U_i^k + U_{i+1}^k), \quad \mu = \frac{\beta \Delta t}{h^2}. \end{aligned}$$

根据上述的四个步骤，我们可以得到

$$\begin{aligned}
 g(\theta)e^{ijh\xi} &= e^{ijh\xi} + \mu(e^{i(j-1)h\xi} - 2e^{ijh\xi} + e^{i(j+1)h\xi}), \\
 g(\theta) &= 1 + \mu(e^{-i\xi h} - 2 + e^{i\xi h}) \\
 &= 1 + \mu(\cos(\theta) - i\sin(\theta) - 2 + \cos(\theta) + i\sin(\theta)) \\
 &= 1 + 2\mu(\cos(\theta) - 1) \\
 &= 1 - 4\mu\sin^2(\theta/2).
 \end{aligned}$$

有了增长因子的表达式，我们希望知道什么时候 $|g(\theta)| \leq 1$ 或者 $-1 \leq g(\theta) \leq 1$ 。注意到

$$-1 \leq 1 - 4\mu \leq 1 - 4\mu\sin^2(\theta/2) = g(\theta) \leq 1$$

如果令 $-1 \leq 1 - 4\mu$ 我们可以保证 $|g(\theta)| \leq 1$ ，于是稳定的一个充分条件是

$$4\mu \leq 2, \quad i.e. \Delta t \leq \frac{h^2}{2\beta}.$$

即使我们没有办法预测上述条件不满足时会出现什么情况，但是该条件为我们提供了一个 Δt 的可选择上界。

5 实例再分析

对于第二节的问题

$$\begin{aligned}
 u_t + au_x &= 0, \quad 0 < x < 1 \\
 u(0, x) &= x, \quad u(t, 0) = -at
 \end{aligned}$$

其中 $a = 1$ 。我们现在就可以利用冯诺依曼稳定性分析方法对结果进行解释了。通过冯诺依曼稳定性分析，我们得到增长因子为

$$\begin{aligned}
 g(\theta) &= 1 - \mu(1 - e^{-ih\xi}) \\
 &= 1 - \mu(1 - \cos(\theta)) - i\mu\sin(\theta), \\
 |g(\theta)|^2 &= (1 - \mu(1 - \cos(\theta)))^2 + \mu^2\sin^2(\theta)^2 \\
 &= 1 - 2(1 - \mu)\mu(1 - \cos(\theta)).
 \end{aligned}$$

如果 $1 - \mu \geq 0$ (i.e. $\Delta t \leq h/a$)，那么我们有 $|g(\theta)| \leq 1$ 。我们再来回顾一下数值求解的效果。

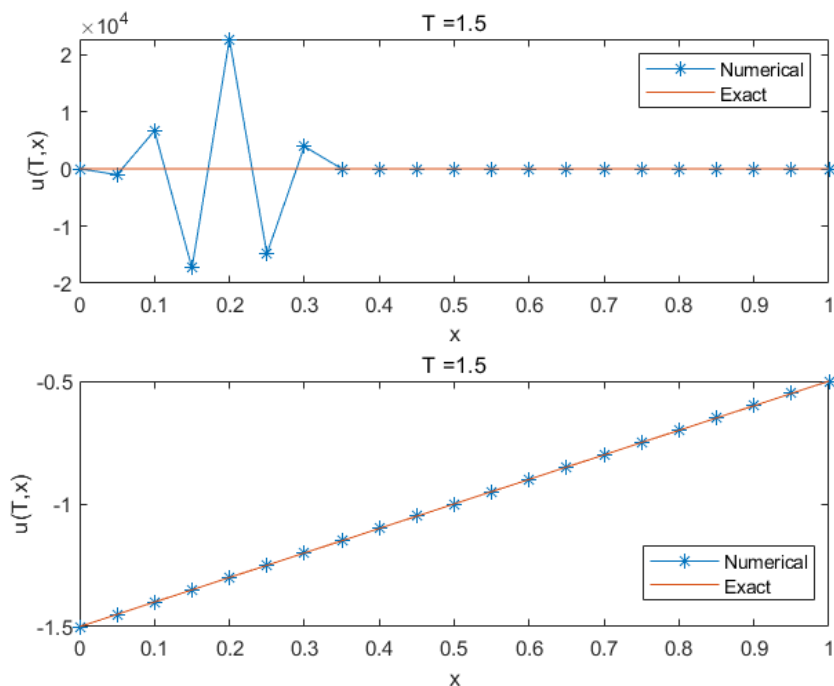


图 1: 求解结果回顾

因此第一种情况下, $\Delta t = 5h/a$ 不能满足稳定性要求, 而第二种情况下 $\Delta t = h/(2a)$ 可以很好地保证数值格式的稳定性。

6 总结

确保设计的格式是稳定的, 这对于数值方法十分重要。除了文中提到的冯诺依曼稳定性检验, 还有能量方法等可以检验稳定性的手段, 在此不做过多展开。此外, 不仅仅是 (含时间微分的) 微分方程数值解格式的稳定性, 数值解中还有许多有意思的问题。本文中所举的例子仅仅是基本例子, 但是所用的方法却是有广泛的适用性的。实际上, 数值格式的设计有许多讲究, 这些也和方程本身的性质有关。如果大家有兴趣可以自己多多阅读一些相关书籍, 并且动手实践。

统计模型在数量遗传学中的应用

作者：管瑞 2018 级 统计班

1 数量遗传学基础

在数量遗传学中，数量性状位点（quantitative trait locus, QTL）定位和全基因组关联分析（Genome-Wide Association Study）已经成为人类基因组学研究的热点之一。许多复杂的人类疾病和具有生物和经济上重要意义的动植物育种都是由多个基因位点和环境因素相互作用决定的，因此需要建立统计模型对基因型与表现型之间的关联情况进行统计推断。随着高密度遗传图谱的建立，传统的数量性状位点定位不能处理高维建模问题。因此，为了追求更高精度的定位以及更优的预测效果，研究者利用贝叶斯统计方法建模。同时，随着大量接近真实人类基因组的实验品系的产生，统计模型也需要综合考虑各种复杂情形，从而更切合实际运用中的需要。本文介绍如何使用混合线性模型联合分析多个复杂性状，建立多响应变量模型，最大化数量性状位点研究的能力和预测精度。通过模拟实验和真实数据实验证明该类方法的有效性，并且在生物学中回答诸如一因多效等实际问题。

1.1 数量性状位点（QTL）定位

数量性状（比如血压和空腹血糖，植物产量，鸡蛋大小等连续变量）一般受到多种遗传因素的影响，这些遗传因素的基因位点叫做数量性状位点（QTL）。QTL 定位，即根据某个实验群体的遗传标记（比如单核苷酸，SNP）片段和表型数据，运用统计模型推测基因型中的标记片段与表型之间的相互关联，从而识别与表型相关的基因组区域，并精确定位这些区域。QTL 定位的前提假设是与数量性状关联的基因标记会更频繁的与表型同时变动，但是与之不相关的基因标记将不会与表型有显著的关联性。

1.1.1 单 QTL 定位

由于技术和算力限制，从 20 世纪初开始到 21 世纪初，QTL 定位方法精度有限，并不能找到准确的 QTL 位置，只能把 QTL 限制在两个标记之间的区间内。最简单的统计方法，比如 t 检验，方差分析，似然比检验，回归分析等都曾被用于 QTL 定位。Lander 和 Botstein[23] 在 1981 年提出区间定位（IM）方法，用相邻两个标记位点之间区域的基因型与表型的关联程度检验统计量来推测 QTL 在此区间存在的证据，成为基于连锁映射 QTL 的基本框架。这种区间定位法假设只有一个 QTL 存在，而忽略了连锁不平衡性导致所得区间受到相邻其他区间的影响，从而降低定位准确度。

1.1.2 多 QTL 定位

单 QTL 模型没有考虑到基因位点中多个 QTL 相互影响的情况，多 QTL 定位解决了这一缺陷。Zeng[57] 和 Janson[18] 提出复合区间定位法 (CIM) 联合考虑多个 QTL 作为协变量，使得每个 QTL 效应相互独立而不受其他区域影响，从而增加检测效力。Zhu 和 Weir 进一步提出基于混合线性模型的复合区间作图方法。Li 等 [24] 在 2007 年提出完备区间定位法 (ICIM)，把 CIM 推广到上位性的 QTL 定位中。多 QTL 定位从统计学角度来看是一个变量选择问题 [29],[1]。由于 QTL 的数目定义了模型维数，导致搜索空间过大，需要一个有效的搜索算法作为支撑，包括前向选择、后向消除、逐步选择和随机算法，模拟退火等。

1.1.3 贝叶斯多 QTL 定位

频率学派对于多 QTL 定位问题建模过程中，无法确定纳入几个自变量，而复杂的变量选择过程消耗大量的算力和时间。贝叶斯学派主要解决了频率学派难以解决的 QTL 数量估计问题 [16] [54]，将 QTL 的数量看作一个随机变量和其他参数一起估计，再用 MCMC 计算框迭代生成马尔可夫链的平稳分布，达到同时估计 QTL 数量与 QTL 位置的目的。RJ-MCMC 算法 [41] 把不同参数空间之间的 Metropolis 跳跃与 MCMC 相结合 [49]，然而 RJ-MCMC 方法需要选择维数改变的接受概率，在实践中这一超参数的选择会增加许多麻烦，并且参数空间的不确定性会引起马尔可夫链的混合。为此，随机搜索变量选 SSVS 方法固定了参数空间的维数，假设每个效应服从两个方差不同的正态分布的混合 (spike-slab 混合先验)，通过后验概率选择效应 [52]。在此基础上，还发展出复合空间表示法 CMS[51] 统一了 SSVS 与 RJ-MCMC。除此之外，还有一类贝叶斯变量选择方法，被称为贝叶斯压缩估计，将不重要的变量效应减小到 0 附近 [49], [48]。综上所述，贝叶斯多 QTL 定位就是贝叶斯子群分析在数量遗传学领域的直接运用。

1.2 全基因组关联分析 (GWAS)

在整个基因组中，单核苷酸多态性 (SNP) 是指核苷酸变异形成的遗传标记 (包括转换, 颠倒, 缺失和插入)。这些变异往往成为新的表现型出现的重要来源，识别 SNP 对表型的影响可以帮助人类培育新的动植物品种，也可以解决医学领域一些复杂综合症的遗传学基础问题。全基因组关联分析 (GWAS) 假设一个样本量很大的实验总体在遗传过程中已经对全基因组范围内的 SNP 变异进行基因分型，利用统计工具分析相关性状之间关联程度，从而筛选出显著影响性状的 SNP 的数量遗传学方法。

从统计学角度来看，QTL 和 GWAS 在模型上没有显著区别。二者之间的区别表现在以下几点：

- 采用的试验源不同. GWAS 一般采用样本量很大的总体，并且假设 SNP 已经在实验群体中大量存在，基因分型在实验总体中是自然且随机发生的，它的目的更加贴近实际应用，大量 GWAS 是在人类基因组中进行的。QTL 定位一般采用实验室培育的动物或植物群

体，实验群体的交配情况经过精心实验设计，样本量相比 GWAS 很小。QTL 一般需要结合理论分析，是理想化的验证材料。

- 定位粒度不同. GWAS 的关联分析是以 SNP 为单位的，需要精确到具体哪一个基因变异影响了表型的变化。QTL 定位的连锁映射是以区间为单位的，一个区间可能包括上万个基因变异位点。在粒度层面，GWAS 比 QTL 定位精准很多，可以把 GWAS 看作 QTL 定位的发展与延续。之前在 QTL 定位上验证有效的方法，都可以运用在 GWAS 上。
- 对算力的依赖程度不同. GWAS 的流行得益于计算机技术的高速发展，而 QTL 定位的样本量，表型个数，基因位点数量都远远不及 GWAS。

1.3 多性状联合分析

全基因组关联分析 (GWAS) 在识别与复杂性状相关的遗传变异方面非常成功。GWAS 的抽样过程中，通常是在总体中收集一组相关性状的信息，然后单独分析这些性状，通过单一响应变量的方式分析全基因组 DNA 标记与复杂性状之间的关联。但是，在许多育种项目中，测量的数量性状之间普遍存在遗传相关性。比如人类疾病综合症哮喘，就由大量高度相关的临床表现型共同影响。这就表明了联合分析多个复杂性状对于 GWAS 和 QTL 定位更有利。从统计建模角度考虑，在不同性状之间存在基因关联的情况下，多变量联合分析可以增加统计效力：由性状间协方差矩阵提供的额外信息在单变量分析中往往会被忽略。从实验设计角度考虑，多变量分析对于一组关联性状只需要进行一次测试，与单一分析相比减少了测试次数，不需要对多重测试进行控制变量对比分析。从遗传学基础考虑，一个遗传变异位点与多个复杂形状相关联的情况时常出现，这种现象在遗传学中叫做“**一因多效性**”。联合分析多个复杂形状更加符合生物学原理，也更加有效。

2 数量遗传学统计建模方法综述

2.1 多响应变量线性混合模型

2.1.1 线性模型

假设自变量和因变量之间是线性关系，同时考虑一个服从正态分布的误差项，线性模型作为统计学中最基础和有效的建模手段，通过对自变量和因变量之间的相互依赖关系进行统计建模，从而进行参数估计，区间估计，假设检验等。线性模型在各种领域都有广泛的应用，例如生物医药，经济金融，计算机科学，工程学，社会科学等。根据具体应用场景中涌现出的不同场景和数据结构，统计领域又需要基于线性模型建立新的变种来反应数据中的结构化特征，丰富统计学的理论内容。

定义 1 (线性模型)

标准形式为：

$$Y = X\beta + \epsilon \quad (1)$$

(1)中 X 是设计矩阵 (自变量), Y 是观察值 (因变量), ϵ 是随机误差项, 一般假设 $\epsilon \sim \mathcal{N}(0, \sigma^2)$ 。

根据参数 β 的性质, 线性模型可以被分成固定效应模型, 随机效应模型和混合效应模型。当 β 作为固定参数时, 模型被称为固定效应模型; 当 β 作为随机变量时, 模型被称为随机效应模型; 而当 β 可以拆分成固定参数和随机变量时, 模型同时包含固定效应, 故被称为混合效应模型, 也称作线性混合模型 (LMM), 是本研究主要讨论的对象。♣

2.2 线性混合模型 (LMM)

定义 2 (线性混合模型)

标准形式定义如下：

$$y = X\beta + Z\mu + \epsilon \quad (2)$$

(2)中 y 是观察值, X 是固定效应设计矩阵, Z 是随机效应设计矩阵, ϵ 是随机误差项, β 是固定效应参数, μ 是随机效应参数, $\mu \sim \mathcal{N}(0, \sigma_\mu^2 \Sigma)$, $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2 I)$, $cov(\mu, \epsilon) = 0$ ♣

根据多元正态分布及其基本性质可知 $y \sim \mathcal{N}(X\beta, V)$, 故 y 的概率密度函数为：

$$f(y) = 2\pi^{-\frac{p}{2}} |V|^{-\frac{1}{2}} \exp -\frac{1}{2}(y - X\beta)^T V^{-1}(y - X\beta) \quad (3)$$

其中 $V = Var(Y) = \sigma_\mu^2 Z\Sigma Z^T + \sigma_\epsilon^2 I$ 。由(3)推出未知参数的最大似然函数 (ML)：

$$L(\beta, V|Y) = -\frac{p}{2} \ln 2\pi - \frac{1}{2} \ln |V| - \frac{1}{2}(Y - X\beta)^T V^{-1}(Y - X\beta) \quad (4)$$

由(4)得到参数的 β 和 μ 的极大似然估计值：

$$\begin{aligned} \hat{\beta} &= (X^T V^{-1} X)^{-1} (X^T V^{-1} Y) \\ \hat{\mu} &= \sigma_\mu^2 Z^T V^{-1} (Y - X\hat{\beta}) \end{aligned}$$

从(3)也可以得到限制性最大似然函数 (REML), 用来解决引入固定效应导致自由度损失带来的有偏估计问题：

$$L(\beta, \Sigma|Y) = -\frac{n-r}{2} \ln 2\pi - \frac{1}{2} \ln |V| - \frac{1}{2}(Y - X\beta)^T V^{-1}(Y - X\beta) - \frac{1}{2} \ln |X^T V^{-1} X| \quad (5)$$

注 在实际应用中, ML 和 REML 往往没有显式解, 可以通过 Newton-Raphson 方法或者 EM 算法进行迭代求解。

2.2.1 多响应变量混合线性模型 (mvLMM)

联合考虑多个响应变量的混合线性模型, 可以建立多响应变量混合线性模型 (mvLMM)。

假设有 P 个响应变量, N 个样本。令 $y_p \in \mathbb{R}^N$ 表示第 p 个响应变量的 N 个观察值组成的向量, $X \in \mathbb{R}^{N \times K}$ 表示 K 个固定效应的设计矩阵, $Z \in \mathbb{R}^{N \times Q}$ 表示 Q 个随机效应的设计矩阵, β_p 表示固定效应, $\mu_p \sim \mathcal{N}(0, \sigma_{\mu,p}^2 \Sigma)$, $\epsilon_p \sim \mathcal{N}(0, \sigma_{\epsilon,p}^2 \mathbf{I})$ 。对于第 p 个响应变量, 建立模型:

$$y_p = X\beta_p + Z\mu_p + \epsilon_p \quad (6)$$

故有:

$$y_p | \mu_p \sim \mathcal{N}(X\beta_p + Z\mu_p, \sigma_{\epsilon,p}^2 \mathbf{I}) \quad (7)$$

对(7)中随机变量 μ_p 求积分, 得到响应变量的分布:

$$y_p \sim \mathcal{N}(X\beta_p, \sigma_{\mu,p}^2 R + \sigma_{\epsilon,p}^2)$$

$$R = Z\Sigma Z^T$$

多响应变量线性混合模型还要对响应变量 y_p 和 y_q 之间的协方差建模:

$$\text{Cov}(y_p, y_q) = \rho_{\mu}^{pq} \sigma_{\mu,p} \sigma_{\mu,q} R + \rho_{\epsilon}^{pq} \sigma_{\epsilon,p} \sigma_{\epsilon,q} \mathbf{I}$$

其中 ρ_{μ} 反映了两个响应变量由自变量导致的相关性, ρ_{ϵ} 反映了两个响应变量由其他没有被纳入模型的影响因素导致的相关性。

引入

$$Y = [y_1, \dots, y_P] \in \mathbb{R}^{N \times P}$$

$$B = [\beta_1, \dots, \beta_P] \in \mathbb{R}^{K \times P}$$

$$U = [\mu_1, \dots, \mu_P] \in \mathbb{R}^{Q \times P}$$

$$E = [\epsilon_1, \dots, \epsilon_P] \in \mathbb{R}^{N \times P}$$

定义协方差:

$$C_{\mu} = \begin{bmatrix} \sigma_{\mu,1}^2 & \rho_{\mu} \sigma_{\mu,1} \sigma_{\mu,2} & \cdots & \rho_{\mu} \sigma_{\mu,1} \sigma_{\mu,P} \\ \rho_{\mu} \sigma_{\mu,1} \sigma_{\mu,2} & \sigma_{\mu,2}^2 & \cdots & \rho_{\mu} \sigma_{\mu,2} \sigma_{\mu,P} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{\mu} \sigma_{\mu,1} \sigma_{\mu,P} & \rho_{\mu} \sigma_{\mu,2} \sigma_{\mu,P} & \cdots & \sigma_{\mu,P}^2 \end{bmatrix}$$

$$C_{\epsilon} = \begin{bmatrix} \sigma_{\epsilon,1}^2 & \rho_{\epsilon} \sigma_{\epsilon,1} \sigma_{\epsilon,2} & \cdots & \rho_{\epsilon} \sigma_{\epsilon,1} \sigma_{\epsilon,P} \\ \rho_{\epsilon} \sigma_{\epsilon,1} \sigma_{\epsilon,2} & \sigma_{\epsilon,2}^2 & \cdots & \rho_{\epsilon} \sigma_{\epsilon,2} \sigma_{\epsilon,P} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{\epsilon} \sigma_{\epsilon,1} \sigma_{\epsilon,P} & \rho_{\epsilon} \sigma_{\epsilon,2} \sigma_{\epsilon,P} & \cdots & \sigma_{\epsilon,P}^2 \end{bmatrix}$$

则多响应变量线性混合模型可以写成矩阵形式:

$$Y = XB + ZU + E \quad (8)$$

其中 U 和 E 服从矩阵正态分布:

$$U \sim \text{MVN}(0, R, C_{\mu})$$

$$E \sim MVN(0, I, C_e)$$

定义 3 (多响应变量线性混合模型)

用 $\text{vec}(Y)$ 表示一个堆叠操作，把表型矩阵纵向扩展成一个表型向量。这种表达方式可以将矩阵正态分布 MVN 表达为具有 Kronrcker 结构的标准正态分布。有关操作符 \otimes , $\text{vec}(\cdot)$ 和矩阵正态分布的性质 MVN 详见附录A.2,A.4。可以得到多响应变量情形下的正态分布表示：

$$\text{vec}(Y) \sim \mathcal{N}(\text{vec}(XB), C_\mu \otimes R + C_e \otimes I) \quad (9) \quad \clubsuit$$

注 由以上表示可以推出极大似然函数，限制性极大似然函数，继而进行参数估计等。具体过程与单一响应变量情形相同，故不再赘述。

2.3 贝叶斯统计推断

贝叶斯统计推断采用联合概率模型来描述观察到的（数据和知识）和未观察到的（参数和假设）之间的关系。给定观测数据和未知参数预先指定的先验分布，联合似然函数可以推导出未知参数的后验分布。再结合蒙特卡洛马尔可夫链（MCMC）迭代更新参数直至达到平衡分布。理论上，贝叶斯框架具有灵活性和通用性，原则上可以处理任意复杂的模型。贝叶斯框架也可以充分考虑到所有与未知量相关的不确定性，推断任何感兴趣的参数。框架的高度统一性，通用性成为贝叶斯学派对比频率学派的一大优势。

具体来说，假设模型中的待估参数 θ 构成集合 Θ 。贝叶斯统计学派把 θ 看成随机变量，服从一个先验分布 $\pi(\theta)$ 。假设 X 为总体， (x_1, \dots, x_n) 为样本（观测值），那么在样本已知的情況下可以计算后验分布：

$$p(x|\theta) = \frac{p(x|\theta)\pi(\theta)}{\int_{\Theta} p(x|\theta)\pi(\theta)d\theta}$$

其中 $p(x|\theta)$ 即为似然函数。

2.3.1 先验分布

先验分布的指定在贝叶斯统计推断及其应用中非常关键，会直接影响到参数估计的性能。在缺乏先验信息的情况下，通常可以采用以下几种先验分布：

- 双指数先验分布.
- Jeffery 先验分布. 当缺乏先验信息时，可以指定 Jeffery 先验。
- 共轭先验分布 这种先验分布和后验分布同属于某个分布族的情形具有显示表达式，利于迭代更新。
- spike-slab 混合先验[10] 由两个不同方差，零均值的正态混合分布分布组合而成，方差大的称为slab，方差小的称为spike。由于spike-slab混合先验具有稀疏性，在贝叶斯变量选择的 SSVS 方法中常用，在基因学中由于大多数遗传变异位点对性状没有影响，故同样广泛使用。

2.3.2 MCMC 算法

在实际问题中，参数 θ 的后验分布 $p(x|\theta)$ 通常没有明确的表达式，需要运用 MCMC 算法。MCMC 中，马尔可夫链是一个随机过程，每个状态仅与之前的一个状态相关。蒙特卡洛方法是针对没有解析解的复杂问题，提出的高效数值算法。经过足够多次循环迭代之后，马尔可夫链将达到平衡分布，即作为待估参数的后验分布。贝叶斯统计框架中，由于计算过程中有难以计算的高维积分存在，使得 MCMC 算法等模拟算法应用广泛。在这一类模拟算法中，Gibbs 采样和 Metropolis-Hastings 抽样又最为有效和流行。

- Gibbs 采样. 将多个参数的联合概率密度这一复杂函数简化为单个参数的简单概率密度函数迭代的过程。具体过程如下：

假设 k 个待估参数构成参数向量 $\theta = (\theta_1, \dots, \theta_k)$, $\theta \in p(\theta)$ ，其中 $p(\theta)$ 为联合概率密度函数。转移密度函数 $p_i(\theta) = p(\theta|\theta_{-i})$ ，其中 θ_{-i} 为 θ 中除去第 i 个分量的 $(k-1)$ 为向量。选取 $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$ 为迭代初始点，第 t 迭代进行如下步骤：

- 1) 从 $p_1(\theta|\theta_2^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_k^{(t-1)})$ 中抽取 $\theta_1^{(t)}$
- 2) 从 $p_2(\theta|\theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_k^{(t-1)})$ 中抽取 $\theta_2^{(t)}$

...

- k) 从 $p_k(\theta|\theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{k-1}^{(t)})$ 中抽取 $\theta_k^{(t)}$
- 得到 $\theta^{(k)} = (\theta_1^{(k)}, \theta_2^{(k)}, \dots, \theta_k^{(k)})$

- Metropolis-Hastings 抽样. 每次更新参数时，以概率 α 接受参数更新。

2.3.3 贝叶斯子群分析

假设一个线性模型 $E(y) = \beta_0 + X_1\beta_1 + \dots + X_p\beta_p + \epsilon$ 。贝叶斯子群分析通过计算 X_i 被包含在线性模型里 ($\beta_i \neq 0$) 的后验概率来评估变量 X_i 的价值，从而确定 X_1, \dots, X_p 中的哪些变量会对预测结果起关键作用。所以在贝叶斯框架中，模型选择问题和参数估计问题可以同时进行，我们把这个过程叫做贝叶斯子群分析 [31]。

由于准确计算后验概率涉及大量模型求和几乎不可实现，贝叶斯子群分析往往需要开发高效与高速的算法，且需要和 MCMC 算法相适应。下面简单介绍几种常用的贝叶斯子群分析方法及其优缺点：

- 随机搜索. 1993 年 George 等人提出随机搜索变量选择 (SSVS) 方法 [10]，该方法利用 spike-slab 混合先验分布，计算后验概率从而选择变量。SSVS 算法的缺点是需要人工选择超参数。
- 引入指示变量[22]. 通过额外引入参数表示某个自变量是否被纳入模型里来进行变量选择，这个参数叫做指示变量 (取值 0 或 1)。不仅可以解决超参数选择的问题，而且没有被纳入模型的参数在 MCMC 参数更新的过程中不需要复杂的计算步骤，节省时间。
- 贝叶斯压缩估计. 给未知参数指定 Jeffery 先验 $p(\theta) \propto \frac{1}{\theta}$ ，当观察值显示某个自变量的影响不大时，它的参数向 0 压缩 [49]。如果假设每个自变量都在模型中且服从正态分布，正态分布的方差指定逆卡方分布作为先验分布，等同于对回归系数本身指定较正态分布

后尾的 t 分布作为先验，即为 BayesA[29]。如果进一步假设只有少数几个自变量被纳入模型，通过参数压缩实现变量选择，即为 BayesB[29]。如果用正态分布分布代替 BayesB 中的 t 分布，即为 BayesC[15]。若指定指数先验，等同于对回归系数本身指定双指数分布或者 Laplace 先验，即为 Bayesian LASSO[53]。Bayesian LASSO 允许自变量的极值以更大的概率出现。这类方法的缺点是无法确定最终的模型选择出的自变量。

- 模型空间方法 直接把模型空间的维数作为未知参数估计，彻底避免了对协变量设置先决条件的问题。可逆跳跃 MCMC (RJ-MCMC) 可以让马尔可夫链在不同维度的空间之间跳跃。复合空间表示 (CMS) 进一步解决了 RJ-MCMC 由于维度变化带来的算法过于复杂的问题。

从以上的分析可以看出，贝叶斯子群分析的重点和难点在于如何指定合理的先验分布，以及如何指定合适的超参数。和频率学派的模型选择相比，贝叶斯子群分析具有更多的待估参数，且 MCMC 过程需要几万次的迭代，并且每次迭代之后要对参数重新估计，这个过程消耗大量时间与算力。然而贝叶斯方法的准确度更高，并且贝叶斯框架的泛化性更好。

3 多响应变量线性混合模型在数量遗传学中的应用

3.1 研究现状

在植物，禽畜以及人类复杂数量性状遗传基础的研究中，探究群体结构对于复杂性状的影响一直是主要课题。群体样本中，个体间的遗传相似性水平差异巨大，群体结构会在涉及关联分析的遗传学研究中造成大量的虚假关联。LMM 主要解决的就是群体结构的问题，使得数量遗传学的模型可以综合考虑实验群体的亲缘关系对于表现型的影响，从而控制了**假阳性**问题 [36]。数量遗传学中经常采用**微效多基因假说**，即数量性状除了受到一个（或多个）主效基因 (QTL) 的影响之外，还受到多个微效基因的影响，而这多个微效基因就由亲缘关系控制，它们虽然不是 QTL，但聚合效应对数量性状的表现型影响极大。综合考虑主效基因与微效基因对性状的影响以及基因间的交互作用，可以从残差项中分离出许多对建模有用的关键信息，使得模型对性状的解释性提高，也能获得多个基因位点之间相互关联的情况。

在 21 世纪初，对于微效多基因假说的处理方式是主成分分析，在一般多元线性模型中加入 Q 矩阵。

定义 4 (Q 模型)

$$Y = Q + M + \epsilon \quad (10)$$

其中 Y 是性状， Q 是主成分（作为潜在的 QTL 位点或者 QWAS 中的 SNP 标记）， M 是其他基因标记， ϵ 是随机误差项。



然而 Q 模型仍然缺乏对于群体结构的细致考察，没有充分利用亲缘关系信息。Yu 等人在 2006 年提出 $Q + K$ 模型 [55]。

定义 5 (Q+K 模型)

在微效多基因假说的前提下，采用 LMM 进行建模，并且引入亲缘关系矩阵说明随机效应的协方差结构：

$$Y = Q + M + K + e \quad (11)$$

模型 6.20 假设待检测的 SNP 标记或者 QTL 位点为固定效应，引入随机效应 K 用来表示多个微效基因，通过构建亲缘关系矩阵确定的方差结构，把真正的函数关联信号从群体结构产生的大量虚假信号中分离出来，修正群体结构和基因关联问题。LMM 是数量遗传学中目前公认的最有效的修正群体结构的手段 [58], [17]。



LMM 运用于大规模基因数据时，涉及大量矩阵 V 的逆运算（见(3)式），对算力要求很高。传统迭代的计算复杂度为 $\mathcal{O}(n^3p^3)$ ，其中 n 为样本个数， p 为性状个数，所以只能用于小型数据集，分析单一性状。为了适应大型数据集与多个表型，一系列高效算法采用不同的方法降低计算复杂度，比如 GEMMEA[60] 使用矩阵转换降低表型协方差矩阵的计算复杂度，FaST-LMM[25] 对亲缘关系矩阵进行奇异值分解，MLMM[37] 基于向前选择算法进行多响应变量 LMM 的模型选择。计算复杂度已经被降低到 $\mathcal{O}(n^3 + p^3)$ ，使得大样本情形下，表型与基因型之间的关联分析在实际计算上可行。

把统计模型运用于人类复杂疾病的研究进行回归与关联分析提高了我们对于疾病的认知，发现了上千个与疾病相关的基因位点。但是由于人类群体中实验数据太少，实验费用高昂，且缺乏可复制性，遗传学基础不明确，需要寻找廉价且替代性好的实验群体。除人类外，分析其他有机物（如禽畜，植物）的遗传关系可以极大的推动人类疾病相关风险因素的识别。其中小鼠实验品系由于实验费用低，基因组和人类重合度高，实验材料丰富，而且实验环境因素可控，成为人类基因学研究的重要试验源 [8]。

3.2 面临的问题与挑战

数量遗传学中大多数关联分析只考虑单一性状（响应变量）与基因型（自变量）之间的关系，而近年来多性状联合分析逐渐流行 [2][34]。这种转变的遗传学基础在于，复杂性状除了受到基因与基因间，基因与环境间的交互作用影响之外，还可能出现**多因一效性**（单个表现型受到多个基因的影响）和**一因多效性**（单个基因位点影响多个表现型）。采用 LMM 建模，以及在模型中考虑多个基因位点都可以把多因一效性纳入模型的考虑范畴（详见 1.1）。一因多效性则需要对多性状进行联合分析（详见 1.3），把 LMM 模型推广到 mvLMM 就是针对联合分析多个复杂性状这一需求。研究发现 [33] 多性状关联分析与单一性状相比，发现遗传位点的比例大幅度提高。mv-BIMBAM[42] 把基因型对表型的影响划分为**直接**，**间接**和**无**，并且采用贝叶斯多元回归评估模型的统计功效。mv-SNPTESTS 指定 Wishart 分布和矩阵正态，进行贝叶斯多元回归。mv-PLINK[7] 采用典型关联分析，识别和测试两组变量线性组合的关联性，从而联合检验多个基因变异位点和多个性状之间的关联。MultiPhen[32] 进行“反向回归”，把多个表型当作自变量，单个基因变异位点作为因变量。

虽然多性状联合分析已经对基因遗传学做出了巨大贡献，但是这些工作都只使用简单的贝叶斯多响应变量线性模型，目前没有针对实验群体中的群体结构与数据异质性设计的多响应变量混合线性模型。

3.3 基于 RIX 小鼠品系的模型理论阐述

3.3.1 实验材料

杂交遗传一直小鼠群体中是最常见的 QTL 定位实验设计方案，这种实验设计可以识别出基因序列中与表型变异相关的位点。在杂交遗传实验设计中，两个近交品系交配，它们的 F_1 子代进行互交或者回交，然后对 F_2 自带进行表型和基因分型建立杂交遗传群体。然而这些群体的遗传基础不丰富，不能提供足够的遗传基因库，从而导致 QTL 区域过大，不能精准定位 SNP 的位置。为此，在 RIL 的基础上发展出了基因修饰系，选择繁殖系，重组近亲杂交系，共生系，同源系，远交系等等丰富的实验群体，用于满足不同的实验需要 [3][12]。

重组近交系 (RI) 由两个近交系通过 F_2 子代及其后代的兄弟姐妹交配而来，直到每个独立品系达到近交状态 (至少 20 代)。使用 RI 品系不需要进行基因分型，因为 RI 品系中每只小鼠的基因型都是相通的。RI 品系由于可重复性高，为 QTL 定位提供了简单有效的实验源，从而被广泛应用于遗传研究 [45]。

随着数量遗传学的不断发展，RI 品系作为实验源弊端逐渐展露。品系中样本数量较少，不满足 GWAS 实验的需求。此外，由于 RI 品系的培育方式，基因组中有很一部分有相同的血统，这个线性被称为基因世代一致 (IBD)，没有大量的变异位点可供分析，也没有性状上的差异。RI 品系的 IBD 区域实际上是遗传图谱的盲点[50]。

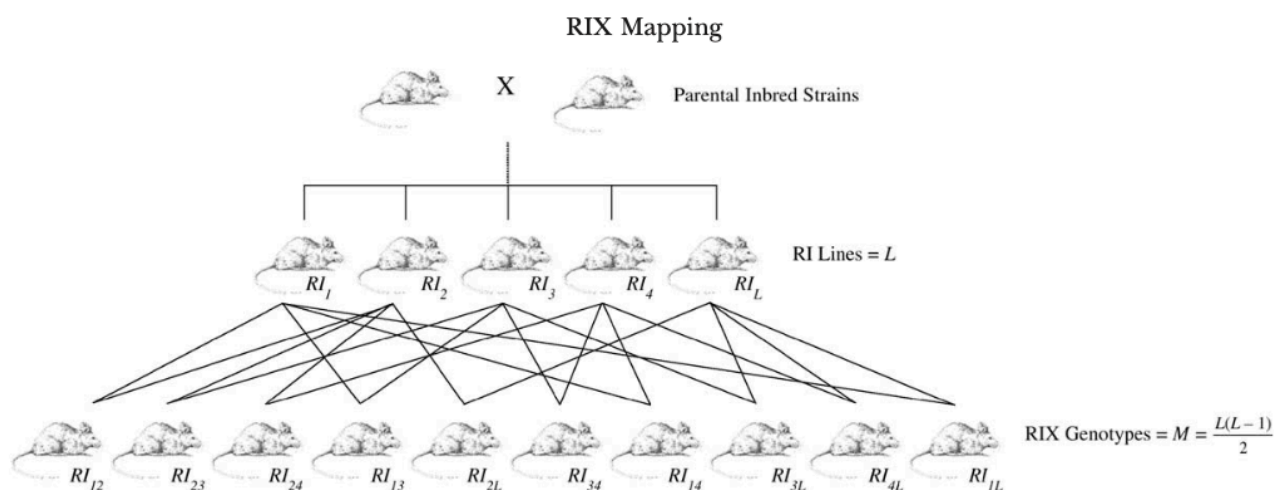


图 1: 重组近交系杂交系 (RIX) 小鼠实验品系的生成

重组近交系杂交 (RIX) 是由 Threadgill 等人在 2002 年首次提出的 [45]，通过同源 RI 小鼠之间杂交产生的 F_1 子代称为 RIX 群体 (如图1)。RIX 实验数据在结构上有明显的群体遗传

结构，具体表现为个体间的亲缘相似性不同，比如共享同一个亲代的 RIX 就不比不共享同一个亲代的 RIX 密切得多。这种特殊数据结构也叫**数据异质性**，导致针对 F_2 杂交实验数据的统计模型不能直接运用在 RIX 实验源上，需要开发为 RIX 群体量身定制的统计建模与分析方法来充分考虑数据异质性带来的其他基因间弱关联。

3.3.2 模型建立

本节联系统计理论与具体实际运用情况，首先根据 RIX 实验设计建立多响应变量混合线性模型，然后给出贝叶斯子群分析框架，包括具体的先验分布指定方案与 MCMC 抽样过程。

假设有 L 个 RI 近交品系，总共可以产生 $N = L(L - 1)/2$ 个 RIX 样本。联合考虑的数量性状总共有 P 个，分别记为 Y_1, \dots, Y_P ，其中 Y_p 是 $N \times 1$ 维关于特性 $p \in \{1, \dots, P\}$ 的观察值向量， $Y_p = (y_{1p}, \dots, y_{Np})'$

基于数量性状的**微效多基因假设**，即数量性状由一个主要 QTL 和多个影响较小的微效基因控制。为了结合群体结构，充分考虑微效基因聚合起来对数量性状的影响，我们使用线性混合模型对性状与基因之间的关系建模。

从 QTL 定位的角度出发，假设 QTL 和多个微效基因对于性状的影响是相互独立的，并且这些微效基因服从正态分布。具体来说，我们的模型考虑 J 个基因标记，认为 QTL 在这些标记之中，这个假设在密集映射的情形下是很合理的，在稀疏映射的情况下，我们可以采用 BayesLASSO 等技术对模型进行微调和改进。 $X \in \mathbb{R}^{N \times J}$ 表示 N 个样本的基因型，矩阵 $Z \in \mathbb{R}^{N \times L}$ 表示 N 个 RIX 样本与其亲本 RI 之间的亲缘关系。由于每个 RIX 样本有两个 RI 亲本，对于任意 $n = 1, \dots, N$ ，有 $\sum_{l=1}^L z_{nl} \equiv 2$ 。

多响应变量混合线性模型可以用如下形式建立：

$$Y = \mu \otimes \mathbf{1}_N + XB + ZA + E \quad (12)$$

其中 $\mu = [\mu_1, \dots, \mu_P]$ ， $B = (\beta_{jp}) \in \mathbb{R}^{J \times P}$ 是性状 p 与基因位点 j 之间的固定效应参数， $A = [\alpha_1, \dots, \alpha_P] \in \mathbb{R}^{L \times P}$ 代表随机效应， $\alpha_p \sim \mathcal{N}(0, \sigma_{\alpha,p}^2 K)$ 群体结构造成的随机微效基因效应参数。由于 RIX 的群体结构， $\alpha_p = (\alpha_{1p}, \dots, \alpha_{Lp})$ 中各个随机分量互相独立，故 K 是对角矩阵，记为 $K = \text{diag}(K^{(1)}, \dots, K^{(L)})$ 。记 $\Sigma_\alpha = (\sigma_{\alpha,p}^2 K^{(l)}) \in \mathbb{R}^{P \times L}$ 。 $E = [\epsilon_1, \dots, \epsilon_P] \in \mathbb{R}^{N \times P}$ 是随机误差项，其中 $\epsilon_p \sim \mathcal{N}(0, \sigma_{\epsilon,p}^2 \mathbf{I})$

与单一性状情形最大的区别在于，多性状需要对性状之间的关联建模。本文假设性状之间的相关性仅在残差项 E 中存在。即残差项在样本之间独立，但是在性状之间相互关联，用 $\Sigma \in \mathbb{R}^{P \times P}$ 表示这个关联矩阵。令 $E = [\epsilon^{(1)}, \dots, \epsilon^{(N)}]'$ ，则 $\forall n \in \{1, \dots, N\}$ ， $\epsilon^{(n)} \sim \mathcal{N}(0, \Sigma)$ 。为了计算方便，本文直接对 Σ^{-1} 指定先验以及进行 MCMC 抽样。

这个模型中可能包含了大量与性状无关的标记位点，这些自变量需要从模型中剔除，即变量选择的过程。为此，本文从贝叶斯子群分析的角度解决这个问题：引入指示变量 $\Gamma = (\gamma_{jp}) \in \mathbb{R}^{J \times P}$ ，用于表示对于第 p 个性状， $\beta^{(p)} = [\beta_{1p}, \dots, \beta_{Jp}]'$ 中哪些效应被包含的模型中 ($\gamma_{jp} = 1$)，哪些需要从模型中剔除 ($\gamma_{jp} = 0$)。

由上述模型的各种假设可知,

$$\text{vec}(Y) \sim \mathcal{N}(\text{vec}(\mu \otimes \mathbf{1}_N + \mathbf{XB}), \mathbf{I}_P \otimes \mathbf{ZZ}' + \Sigma \otimes \mathbf{I}_N)$$

3.3.3 贝叶斯子群分析

可观测的数据有: 性状观察值矩阵 Y , 基因型矩阵 X , 亲缘关系矩阵 Z 。

令 $\Theta = \{\mu, B, A, \Gamma, \Sigma_\alpha, \Sigma^{-1}\}$ 表示模型未知参数的集合, 则联合后验分布可以表示为:

$$p(\mu, B, A, \Gamma, \Sigma_\alpha, \Sigma^{-1} | Y) \propto \prod_{n=1}^N p(Y^{(n)} | \mu, B, A, \Gamma, \Sigma_\alpha, \Sigma^{-1}) \cdot p(\mu, B, A, \Gamma, \Sigma_\alpha, \Sigma^{-1}) \quad (13)$$

其中

$$p(Y^{(n)} | \mu, B, A, \Gamma, \Sigma_\alpha, \Sigma^{-1}) \propto \frac{1}{|\Sigma|^{1/2}} \cdot \exp \left\{ -\frac{\sum_{n=1}^N (Y^{(n)} - \mu - X^{(n)}B - Z^{(n)}A)^2}{2\Sigma} \right\} \quad (14)$$

3.3.4 先验分布指定

公式(13)中, $p(\mu, B, A, \Gamma, \Sigma_\alpha, \Sigma^{-1})$ 最初需要由先验分布指定。

- 1 对于 μ : 指定常数先验, $p(\mu) \propto 1$, 由各个性状的样本均值和样本协方差矩阵决定。
- 2 对于 B : 对 β_{jp} 引入指示变量 γ_{jp} , *s.t.* $\beta_{jp} | \gamma_{jp} = (1 - \gamma_{jp})I_0 + \gamma_{jp}\mathcal{N}(0, v_{jp}^2)$ 。
- 3 对于 Γ : $\gamma_{jp} \sim \text{Bernoulli}(w)$, $w \sim \mathcal{U}(0, 1)$ 。
- 4 对于 v_{jp}^2 : 指定 Jeffery 无信息先验: $p(v_{jp}^2) \propto \frac{1}{v_{jp}^2}$ 。
- 5 对于 Σ_α : 指定 Jeffery 无信息先验 $\Sigma_{\alpha, pl}^2 \propto (\sigma_{\alpha, jpl}^2)^\delta$, 其中 $\delta \in (0, 1/2]$ 是为了保证后验分布正则而引入的超参数。
- 6 对于 Σ^{-1} , 指定无信息先验 $p(\Sigma^{-1}) \sim |\Sigma^{-1}|^{-(1+P)/2}$ 。

结合以上先验分布, 则有:

$$p(\mu, B, A, \Gamma, \Sigma_\alpha, \Sigma^{-1}) = p(\mu)p(B|\Gamma, \Sigma_\alpha)p(\Sigma)p(\Gamma) \quad (15)$$

3.3.5 MCMC 算法与后验分布计算

联合后验分布可以用 Gibbs 采样和 Metropolis 算法来模拟, 基于观测数据和其他参数的条件分布, 交替更新所有未知参数。常用的更新方案是, 对于向量或者矩阵, 采取联合更新的方式。然而对于多基因组多性状 QTL 分析, 这种方法在每次模拟迭代时涉及大量矩阵运算。为了避免计算负担过大, 我们一次只更新一个参数, 这种高效算法不需要矩阵运算, 且当 $\gamma_{jp} = 0$ 时, 可以不再需要对 β_{jp} 进行抽样。

步骤 1. 更新 μ_p :

$$\mu_p | \Theta_{-\mu}, Y \sim \mathcal{N}\left(\frac{\sum_{n=1}^N (Y^{(n)} - \mu_{-p} - X^{(n)}\beta_p - Z^{(n)}\alpha_p)\Sigma_p^{-1}}{N\Sigma_{pp}^{-1}}, \frac{1}{N\Sigma_{pp}^{-1}}\right) \quad (16)$$

步骤 2. 对于未被纳入模型的 β_{jp} (即 $\gamma_{jp} = 0$), 从先验分布 $\mathcal{N}(0, v_{jp}^2)$ 中重新抽样; 对于模型中的 β_{jp} (即 $\gamma_{jp} = 1$), 按照以下方法更新参数:

$$\beta_{jp} | \Theta_{-B}, Y \sim \mathcal{N}\left(\frac{\sum_{n=1}^N x_{nj}(Y^{(n)} - \mu - X^{(n)}\beta_{-jp} - Z^{(n)}\alpha_p)\Sigma_p^{-1}}{\Sigma_{pp}^{-1} \sum_{n=1}^N x_{nj}^2 + \sigma_{\alpha,pj}^{-2}}, \frac{1}{\Sigma_{pp}^{-1} \sum_{n=1}^N x_{nj}^2 + \sigma_{\alpha,pj}^{-2}}\right) \quad (17)$$

步骤 3. 更新 α_{lp} :

$$\alpha_{jl} | \Theta_{-A}, Y \sim \mathcal{N}\left(\frac{\sum_{n=1}^N z_{nl}(Y^{(n)} - \mu - X^{(n)}\beta_p - Z^{(n)}\alpha_{-lp})\Sigma_p^{-1}}{\Sigma_{pp}^{-1} \sum_{n=1}^N z_{nj}^2 + \sigma_{\alpha,pj}^{-2}}, \frac{1}{\Sigma_{pp}^{-1} \sum_{n=1}^N z_{nj}^2 + \sigma_{\alpha,pj}^{-2}}\right) \quad (18)$$

步骤 4. 从逆卡方分布中重新更新 $\sigma_{\alpha,pl}^2$, 组成 Σ_α :

$$\sigma_{\alpha,pl}^2 | \Theta_{-\sigma_{\alpha,pl}}, Y \sim Inv - \chi^2\left(\sum_{l=1}^L \alpha_{pl}^2 / \chi_{L-2\delta}^2\right) \quad (19)$$

步骤 5. 从逆卡方分布中更新 v_{jp}^2 :

$$v_{jp}^2 | \Theta_{-v_{jp}}, Y \sim Inv - \chi^2\left(\sum_{j=1}^J v_{jp}^2 / \chi_{J-2\delta}^2\right) \quad (20)$$

步骤 6. 从 Wishart 分布中更新 Σ^{-1} :

$$\begin{aligned} \Sigma^{-1} | \Theta_{-\Sigma}, Y &\sim Wishart\left(\left(\sum_{n=1}^N (Y^{(n)} - \mu X^{(n)}B - Z^{(n)}A)^T (Y^{(n)} - \mu X^{(n)}B - Z^{(n)}A)\right)^T, n\right) \\ &= |\Sigma^{-1}|^{1/2(n-P-1)} \\ &\cdot \exp\left\{-\frac{1}{2}tr\left\{\sum_{n=1}^N (Y^{(n)} - \mu X^{(n)}B - Z^{(n)}A)^T (Y^{(n)} - \mu X^{(n)}B - Z^{(n)}A)\Sigma^{-1}\right\}\right\} \end{aligned} \quad (21)$$

步骤 7. 更新 Γ :

$$p(\gamma_{jp} = 1 | \Theta_{-\Gamma}, Y) = \frac{p(\gamma_{jp} = 1)p(y | \Theta_{-\Gamma}, \gamma_{jp} = 1)}{p(\gamma_{jp} = 0)p(y | \Theta_{-\Gamma}, \gamma_{jp} = 0) + p(\gamma_{jp} = 1)p(y | \Theta_{-\Gamma}, \gamma_{jp} = 1)} \quad (22)$$

当这一轮抽样结束之后, 用更新的参数值重复以上步骤进行下一轮抽样, 直至马尔可夫链迭代至收敛, 此时采样的参数近似服从后验分布。

4 实际数据分析

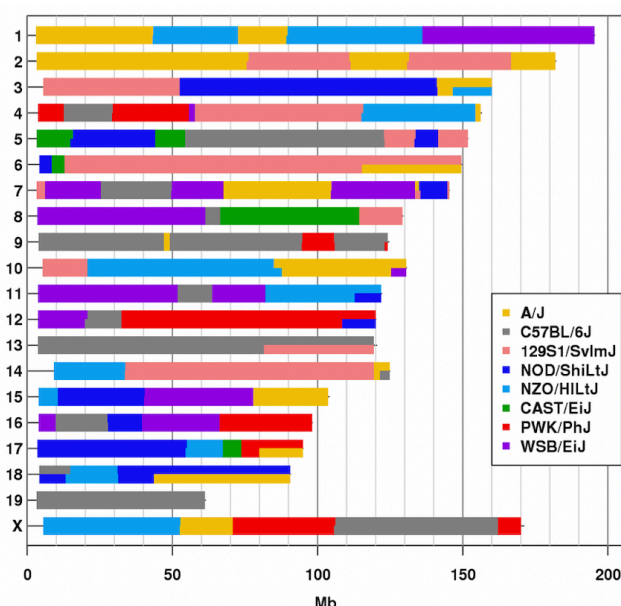
本章采用真实数据多样性远交 (DO) 小鼠群体, 给出数量性状位点分析与全基因组关联分析的最终结果, 并对结果进行解释, 体现模型的实际有效性。

4.1 数据来源

传统的单代杂交遗传图谱研究的分辨率很低（一个 QTL 平均跨越 30Mbp，包括几百个候选基因，无法精确定位），而且只能发现两个亲代品系之间的多态性，不能满足遗传多样性的实验要求。多亲交配种群通过增加基因重组和引入多个创始人品系的等位基因变异位点来解决这一问题。CC[5] 是由 Oak Ridge National Laboratory (ORNL) 开发的一个大规模项目，旨在创造多亲本重组近交系小鼠 [6]，利用 8 个遗传学上差异明显 RI 系小鼠作为亲本品系，与它们的杂交后代进行近亲杂交，从而进行基因重组 [4]。8 个创始人品系分别为 A/J, C57BL/6J, 129S1/SvImJ, NOD/ShiLtJ, NZO/hILtJ, CAST/EiJ, PWK/PhJ 和 WSB/EiJ（颜色表示如图2，按顺序记为 A-H），其中后三个是野生品系。传统近交品系缺乏基因变异，因此野生品系为 CC 群体引入大量遗传多样性。

CC-RIX 小鼠品系是 CC 品系的 F_1 子代按照 RIX 实验设计杂交而成的后代，它既有 CC 群体的 8 个创始人等位基因，又具有 RIX 实验设计的结构，可以很好的验证统计模型对真实情况的抽象能力。

Collaborative Cross Line



Diversity Outbred Mouse

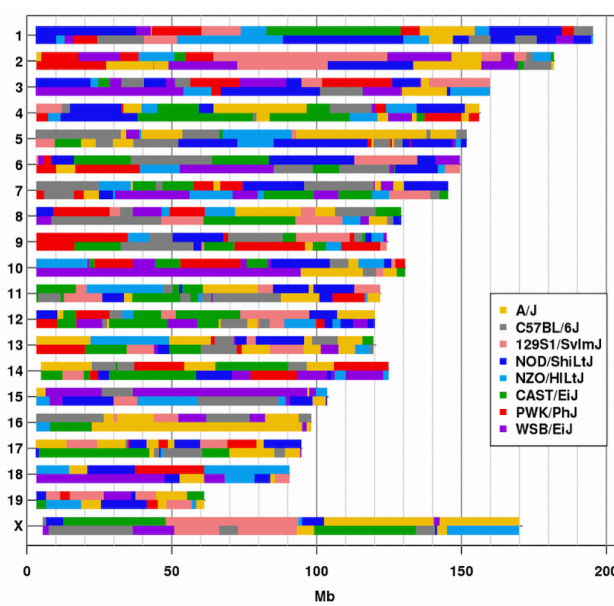


图 2: CC 和 DO 小鼠的基因异质性对比

异质种群是一类新开发的小鼠品系，由多个创始人品系杂交而衍生，作为近交种群维持多代。这种培育方法引入多个创始人品系中的等位基因，增加基因重组发生的概率（大约有 4500 万个不同的 SNP，是传统小鼠品系中的四倍），捕获更大的遗传多样性，构建更加多样化的性状，解决了原本实验品系的缺点。多样性远交（DO）是由 CC 群体的八个创始人种群随机远交（outcrossing）所产生的异质种群 [4]。DO 小鼠虽然和 CC 群体分离了同样的 8 个创始人等位基因，但是由于采用了随机的繁殖策略，每只小鼠都具有独特的，且有效分离的等位基因组合，提供了更多的遗传杂合性（如图1）。DO 小鼠的表型数据十分丰富，且基因型中大

约有 4500 万个 SNP，高精确重组率使其成为高分辨率遗传图谱的理想资源 [44]。此外，DO 的遗传变异是均匀分布的，基本上所有已知基因的编码或调控区都有多个等位基因变异位点。这与经典的近交系小鼠形成了鲜明的对比，后者的遗传多样性有限，而且盲点很大，实际上没有变异。因此，DO 更密切地反映人类疾病的基因机制。

异质的群体中，每个样本的亲缘关系远近程度都不相同，相关联的样本之间的协方差需要使用带有亲缘关系协方差结构的混合线性模型进行回归，才能最大程度发挥异质群体实验品系的优点。与单代杂交的遗传群体不同，异质群体 QTL 和 GWAS 的模型中必须容纳多个创始人等位基因，复杂的育种设计引起的群体间各种复杂效应都必须在建模时充分考虑。

Jackson 实验室中有关于多样性远交小鼠的介绍与相关数据。

从 GeneSeek 的 MegaMUGA 或 GigaMUGA 序列中可以获取 DO 小鼠的 SNP 数据。在实验中，我们还需要 8 个创始品系的 SNP 基因标记，以及他们的基因和物理位置，Dan Gatti[9] 公布了这些数据。DO 小鼠的表现型可以从 MPD 数据库中获得。GWAS 中使用的 SNP 数据用 SQLite 数据库的形式调取，具体来源：CC 变异位点数据库，小鼠基因数据库，小鼠基因 MGI 数据库。

DO 第四代和第五代总共 742 个样本，在标准饲料喂养的 24 小时之内，获得了一系列表现型数据，包括血浆成分（总胆固醇，HDL 胆固醇，葡萄糖，甘油三酯）和身体参数（体重，体脂）。

4.2 实验模型与设计

DO 小鼠的每一对染色体都是由不同的创始人单倍型 (founder haplotypes) 拼接而成的，把某一特定基因位点上的一对创始人单倍型简称为创始人单倍型。根据 8 个创始人单倍型的贡献拟合了一个 LMM 模型，将性状回归到八个自变量上，并且纳入样本间的亲属关系对协方差结构进行调整。具体来说，对每个性状 y_k ，建立以下模型：

$$y_k = X\alpha_k + Z\mu_k + \epsilon$$

其中 $y_k \in \mathbb{R}^N$ 是其中某个复杂性状在 N 个样本上的观测， $X \in \mathbb{R}^{N \times P}$ 和 $Z \in \mathbb{R}^{N \times Q}$ 分别是固定效应和随机效应的设计矩阵， P, Q 分别是固定效应和随机效应的个数，在本实验中 $P = 2, Q = 8$ 。 $\alpha_k \in \mathbb{R}^P$ 是固定效应（性别和实验组别）， $\beta_k \sim \mathcal{N}(0, \sigma_b^2 K)$ 是随机效应（创始人等位基因对性状的影响程度），其中 K 是亲缘关系矩阵（详见），由个体之间共享等位基因的比例计算。 $\epsilon \sim \mathcal{N}(0, \sigma_c^2 I)$ 是残差项。

注 考虑用 DO 小鼠与创始人种群之间相互匹配的可能性建立 K 矩阵。这个概率可以用隐马尔可夫模型 (HMM) 中相邻基因组标记之间的转换概率确定。HMM 的参数采用 EM 算法估计。通过对每只 DO 小鼠的每个 SNP 标记位点上的杂交状态（产生概率估计，从基因型数据中重建出 36 维基因型概率矩阵，然后将每个创始人单倍型在 SNP 上贡献的概率相加，合成为 8 态等位基因状态概率矩阵，记为 K 。读者可以参考 [44]。

注 在亲缘关系矩阵 K 的计算过程中，采用 LOCO (leave one chromosome out) 方法。图2中，颜色越深代表样本之间共享的创始人等位基因越多，亲缘关系越大。

对每一条染色体，通过计算LOD分数，都可以获得一条沿着染色体上基因标记的LOD曲线，在这条曲线上会有 QTL 的峰值，代表此处存在 QTL 的证据充足。然而，到底 LOD 分数多高我们才应该接受备择假设呢，这是一个统计中经常遇到的假设检验问题。为了确定 QTL 定位结果的统计学意义，置换检验打乱基因型和表型，打破二者之间的关联关系，然后在打乱的数据上计算出全基因组范围内最大的LOD分数，将此作为显著性阈值，它表示了随机情况下可能产生的最大的LOD分数。

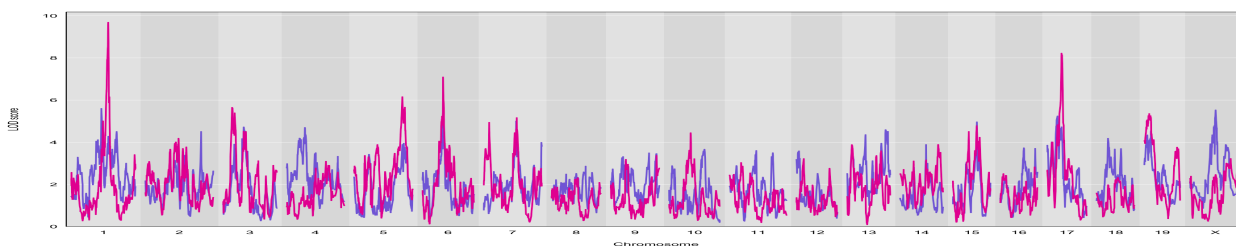


图 3: 20 条染色体上 QTL 定位的 LOD 分数 (红色: NEUT; 紫色: WBC)

图3展示了 QTL 定位的贝叶斯因子，数值越高，对应位置出现影响数量性状的 QTL 的可能性就越大。在第 1, 6, 17 染色体上出现 QTL 的峰值，图10展示这三条染色体上 QTL 定位的结果与 8 个创始人单倍型的贡献情况。其中创始人单倍型贡献在 [0, 1] 之间，贡献越大颜色越深。QTL 定位确定了 3 个 NEUT 性状的 QTL 区间。第一个 QTL (如图4) 位于 1 号染色体: 57.2440。第二个 QTL (如图6) 位于 6 号染色体: 32.4710。第三个 QTL (如图6) 位于 17 号染色体: 23.4208。QTL 定位确定了 1 个 WBC 性状的 QTL 区间 (如图5)，位于 6 号染色体: 32.4537。根据 QTL 定位的结果可以发现，6 号染色体上发现的 QTL 区间上，潜在基因变异位点对两个性状都有影响，这个现象即为关联分析的生物学基础：“一因多效性”。

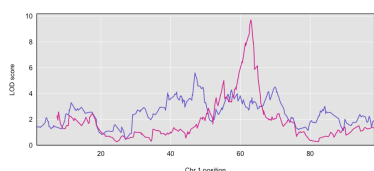


图 4: 第 1 条染色体

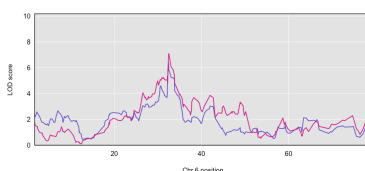


图 5: 第 6 条染色体

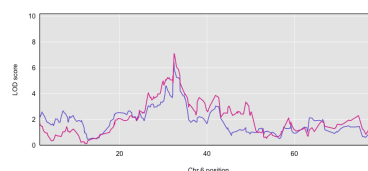


图 6: 第 17 条染色体

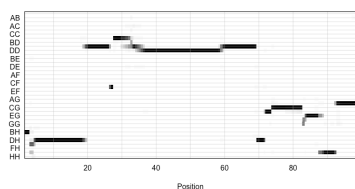


图 7: 第 1 条染色体

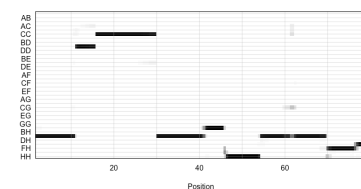


图 8: 第 6 条染色体

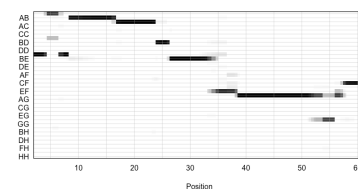


图 9: 第 17 条染色体

图 10: 染色体上数量性状定位结果 (a-c)，创始人等位基因的贡献 (d-f)

图15所示的等位基因效应图说明了 8 个创始人品系分别沿着 1, 6, 17 号染色体上每个基因位点遗传贡献的表型效应。对 6 号染色体来说 (如图13和图14)，性状 NEUT 对创始人等位基因 NOD 和 B6 最敏感；性状 WBC 对创始人等位基因 NOD, WSB 和 NZO 最敏感。在每条

染色体上，性状 NEUT 都受到创始人等位基因 B6 的影响，因为 QTL 的峰值总是和 B6 的低点同时出现。

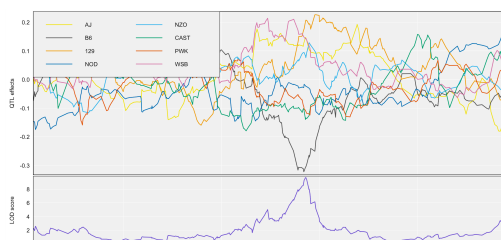


图 11: 第 1 条染色体; 性状 NEUT

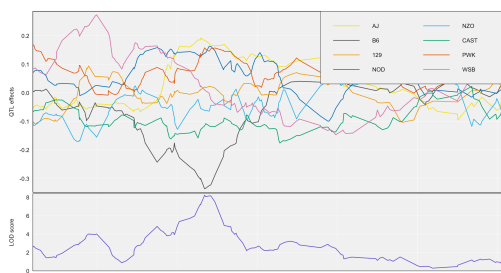


图 12: 第 17 条染色体; 性状 NEUT

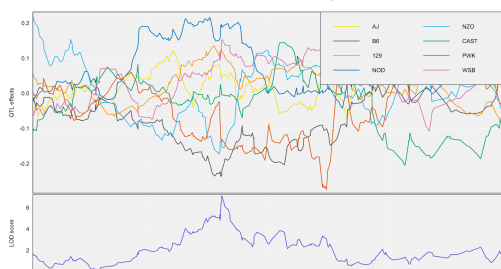


图 13: 第 6 条染色体; 性状 NEUT

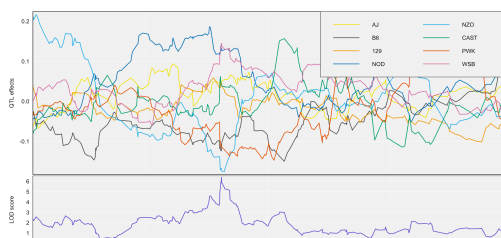


图 14: 第 6 条染色体; 性状 WBC

图 15: 染色体上对应数量性状的 8 个创始人等位基因 QTL 效应估计

结论

本文提出了针对异质数据的多响应变量线性混合模型及其贝叶斯子群分析方法，在对异质群体进行数量性状位点定位与全基因组关联分析时，可以联合分析多个性状，并且通过贝叶斯蒙特卡洛抽样获得更准确的参数估计。统计模拟与真实数据实验表明，与单一响应变量线性混合模型相比，联合考虑多个性状更能发现数据中的复杂关联，具有更高的统计功效；与频率学派参数估计方法相比，贝叶斯子群分析对于数量性状的关联基因位点定位更加精准，框架具有更好的普适性与广泛性。我们有理由相信，贝叶斯学派与多性状联合分析都是未来数量遗传学发展的方向。

本文提出的模型存在计算时间上的局限性。贝叶斯方法向来被诟病运行时间过长。在本文的大量实验中都发现，同等数据量下贝叶斯计算框架的推断时间是频率学派的 20 倍以上。效率是一个产品能够落地的重要因素，所以目前工业界与生物界很少采用贝叶斯框架进行数量遗传学的相关研究。本文没有针对计算效率上进行响应改进，这是未来研究的主要方向。此外，本文研究的实验样本局限在小鼠群体中，没有考虑大型的人类基因组，植物育种等多样性的实验群体。虽然调研过程中发现这些群体同样具有复杂和值得深入研究的群体结构，然而由于精力，篇幅和时间限制没有进行更多研究，未来可以针对这些群体进行和本文类似的框架拓展。

参考文献

- [1] Karl W. Broman and Terence P. Speed. “A Model Selection Approach for the Identification of Quantitative Trait Loci in Experimental Crosses”. In: Journal of the Royal Statistical Society. Series B (Statistical Methodology)64.4 (2002), pp. 641–656.
- [2] Brendan Bulik-Sullivan et al. “An atlas of genetic correlations across human diseases and traits”. In: Nature genetics 47.11 (2015), pp. 1236–1241.
- [3] Riyan Cheng et al. “Genome-wide association studies and the problem of relatedness among advanced intercross lines and other highly recombinant populations”. In: Genetics 185.3 (2010), pp. 1033–1044.
- [4] E. J. Chesler et al. “The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics”. In: Mammalian genome : official journal of the International Mammalian Genome Society19.6 (2008), pp. 382–389. doi: 10.1007/s00335-008-9135-8.
- [5] G Churchill, D Airey, and H Allayee. “The Collaborative Cross, a community resource for the genetic analysis of complex traits”. In: Nature Genetics 36.11 (2004), pp. 1133–1137.
- [6] Collaborative Cross Consortium. “The genome architecture of the Collaborative Cross mouse genetic reference population”. In: Genetics 190.2 (2012), pp. 389–401. doi: 10.1534/genetics.111.132639.
- [7] Manuel AR Ferreira and Shaun M Purcell. “A multivariate test of association”. In: Bioinformatics 25.1 (2009), pp. 132–133.
- [8] Jonathan Flint and Eleazar Eskin. “Genome-wide association studies in mice”. In: Nature Reviews Genetics 13.11 (2012), pp. 807–817.
- [9] Daniel M Gatti et al. “Quantitative trait locus mapping methods for diversity outbred mice”. In: G3: Genes, Genomes, Genetics 9.4 (2014), pp. 1623–1633.
- [10] Edward I. George and Robert E. McCulloch. “Variable selection via Gibbs sampling”. In: Journal of the American Statistical Association 423 (1993), pp. 881–889. doi: 10.2307/2290777.
- [11] Yi Gong and Fei Zou. “Varying coefficient models for mapping quantitative trait loci using recombinant inbred intercrosses”. In: Genetics 190.2 (2012), pp. 475–486.
- [12] Natalia M Gonzales et al. “Genome wide association analysis in a mouse advanced intercross line”. In: Nature communications 9.1 (2018), pp. 1–12.
- [13] Rebecca L Gould et al. “Quantitative trait mapping in Diversity Outbred mice identifies novel genomic regions associated with the hepatic glutathione redox system”. In: Redox biology 46 (2021), p. 102093.

-
- [14] Yongtao Guan and Matthew Stephens. “Practical issues in imputation-based association mapping”. In: PLoS genetics 4.12 (2008), e1000279.
- [15] D Habier et al. “Extension of the bayesian alphabet for genomic selectio”. In: BMC Bioinformatics 186.12 (May 2011). DOI: 10.1186/1471-2105-12-186.
- [16] I Hoeschele and P M Vanraden. “Bayesian analysis of linkage between genetic markers and quantitative trait loci. II. Combining prior knowledge with experimental evidence”. In: Genetics 8 (1993), pp. 946–952. DOI: 10.1007/BF00215033.
- [17] Xuehui Huang. “Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm”. In: Nature Genetics (2012), pp. 32–39. DOI: 10.1038/ng.1018.
- [18] RC Jansen and P Stam. “High resolution of quantitative traits into multiple loci via interval mapping”. In: genetics 4 (1994), pp. 1447–1455. DOI: 10.1093/genetics/136.4.1447.
- [19] Robert E Kass and Adrian E Raftery. “Bayes factors”. In: Journal of the american statistical association 90.430 (1995), pp. 773–795.
- [20] David C Katz et al. “Facial shape and allometry quantitative trait locus intervals in the Diversity Outbred mouse are enriched for known skeletal and facial development genes”. In: PloS one 15.6 (2020), e0233377.
- [21] Arthur Korte et al. “A mixed-model approach for genome-wide association studies of correlated traits in structured populations”. In: Nature genetics 44.9 (2012), pp. 1066–1071.
- [22] Lynn Kuo and Bani Mallick. “Variable Selection for Regression Models”. In: Sankhya: The Indian Journal of Statistics, Series B 60.1 (1998).
- [23] E. Lander and D. Botstein. “Mapping mendelian factors underlying quantitative traits using RFLP linkage maps”. In: genetics (1989), pp. 183–199.
- [24] Huihui Li et al. “Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations”. In: Theoretical and applied genetics 2 (2008), pp. 243–260. DOI: 10.1007/s00122-007-0663-5.
- [25] Christoph Lippert et al. “FaST linear mixed models for genome-wide association studies”. In: Nature methods 8.10 (2011), pp. 833–835.
- [26] Ryan W Logan et al. “High-precision genetic mapping of behavioral traits in the diversity outbred mouse population”. In: Genes, Brain and Behavior 12.4 (2013), pp. 424–437.
- [27] Ani Manichaikul et al. “A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis”. In: Genetics 181.3 (2010), pp. 1077–1086. DOI: 10.1534/genetics.108.094565.

-
- [28] Jonathan Marchini et al. “A new multipoint method for genome-wide association studies by imputation of genotypes”. In: *Nature genetics* 39.7 (2007), pp. 906–913.
- [29] T H Meuwissen, B J Hayes, and M E Goddard. “Prediction of total genetic value using genome-wide dense marker maps”. In: *Genetics* 4 (2001), pp. 1019–1829. DOI: 10.1093/genetics/157.4.1819.
- [30] Theo H E Meuwissen and Mike E Goddard. “Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data”. In: *Genetics* 3 (2004), pp. 761–279. DOI: 10.1186/1297-9686-36-3-261.
- [31] R. B. O’Hara and M. J. Sillanpää. “A review of Bayesian variable selection methods: what, how and which”. In: *Bayesian Analysis* 4.1 (2009), pp. 85–117. DOI: 10.1214/09-BA403.
- [32] Paul F O’ Reilly et al. “MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS”. In: *PloS one* 7.5 (2012), e34861.
- [33] Heather F Porter and Paul F O’ Reilly. “Multivariate simulation framework reveals performance of multi-trait GWAS methods”. In: *Scientific reports* 7.1 (2017), pp. 1–12.
- [34] Cross-Disorder Group of the Psychiatric Genomics Consortium et al. “Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis”. In: *The Lancet* 381.9875 (2013), pp. 1371–1379.
- [35] David Reich, Alkes L Price, and Nick Patterson. “Principal component analysis of genetic data”. In: *Nature genetics* 40.5 (May 2008), pp. 491–492. ISSN: 1061-4036. DOI: 10.1038/ng0508-491.
- [36] NA Rosenberg et al. “Genome-wide association studies in diverse populations”. In: *Nat Rev Genet* 11.5 (2010), pp. 356–366. DOI: 10.1038/nrg2760.
- [37] Vincent Segura et al. “An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations”. In: *Nature genetics* 44.7 (2012), pp. 825–830.
- [38] Bertrand Servin and Matthew Stephens. “Imputation-based analysis of association studies: candidate regions and quantitative traits”. In: *PLoS genetics* 3.7 (2007), e114.
- [39] Heejung Shim et al. “A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians”. In: *PLoS one* 10.4 (2015), e0120758.
- [40] M J Sillanpää and E Arjas. “Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data”. In: *Genetics* 3 (1998), pp. 1373–1388.
- [41] D. A. Stephens and R. D. Fisch. “Bayesian Analysis of Quantitative Trait Locus Data Using Reversible Jump Markov Chain Monte Carlo”. In: *Biometrics* 54.4 (1998), pp. 1334–1347.

-
- [42] Matthew Stephens. “A unified framework for association analysis with multiple related phenotypes”. In: PloS one 8.7 (2013), e65245.
- [43] Matthew Stephens and David J Balding. “Bayesian statistical methods for genetic association studies”. In: Nature Reviews Genetics 10.10 (2009), pp. 681–690.
- [44] K. L. Svenson et al. “High-resolution genetic mapping using the Mouse Diversity outbred population”. In: Genetics 190.2 (2012), pp. 437–447. DOI: 10.1534/genetics.111.132597.
- [45] D. W. Threadgill, K. W. Hunter, and R. W. Williams. “Genetic dissection of complex and quantitative traits: from fantasy to reality via a community effort”. In: Mammalian genome : official journal of the International Mammalian Genome Society 13.4 (2002), pp. 175–178. DOI: 10.1007/s00335-001-4001-Y.
- [46] Shirng-Wern Tsaih et al. “Quantitative trait mapping in a diallel cross of recombinant inbred lines”. In: Mammalian genome 16.5 (2005), pp. 344–355.
- [47] William Valdar, Jonathan Flint, and Richard Mott. “Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice”. In: Genetics 172.3 (2006), pp. 1783–1797.
- [48] Hui Wang et al. “Bayesian Shrinkage Estimation of Quantitative Trait Loci Parameters”. In: Genetics 170 (June 2005), pp. 465–80. DOI: 10.1534/genetics.104.039354.
- [49] Shizong Xu. “Estimating polygenic effects using markers of the entire genome”. In: Genetics 2 (2003), pp. 789–801.
- [50] H. Yang et al. “Subspecific origin and haplotype diversity in the laboratory mouse”. In: Nature genetics 43.7 (2011), pp. 648–655. DOI: 10.1038/ng.847.
- [51] Nengjun Yi. “A Unified Markov Chain Monte Carlo Framework for Mapping Multiple Quantitative Trait Loci”. In: Genetics 167.2 (June 2004), pp. 967–975. ISSN: 1943-2631. DOI: 10.1534/genetics.104.026286.
- [52] Nengjun Yi, Varghese George, and Allison David B. “Stochastic search variable selection for identifying multiple quantitative trait loci”. In: Genetics 164.3 (2003), pp. 1129–1138.
- [53] Nengjun Yi and Shizhong Xu. “Bayesian LASSO for Quantitative Trait Loci Mapping”. In: Genetics 2 (2008), pp. 1045–1055. DOI: 10.1534/genetics.107.085589.
- [54] Nengjun Yi and Shizhong Xu. “Mapping quantitative trait loci for complex binary traits in outbred populations”. In: Heredity (1999), pp. 668–676. DOI: 10.1046/j.1365-2540.1999.00529.
- [55] Jianming Yu et al. “A unified mixed-model method for association mapping that accounts for multiple levels of relatedness”. In: Nature Genetics 38.2 (2006), pp. 203–208. DOI: 10.1038/ng1702.

- [56] Zhongshang Yuan, Fei Zou, and Yanyan Liu. “Bayesian multiple quantitative trait loci mapping for recombinant inbred intercrossoes”. In: Genetics 188.1 (2011), pp. 189–195.
- [57] Zhaobang Zeng. “Precision mapping of quantitative trait loci”. In: genetics 4 (1994), pp. 1457–1469.
- [58] Zhiwu Zhang. “Mixed linear model approach adapted for genome-wide association studies”. In: Nature Genetics 42 (2010), pp. 335–360. DOI: 10.1038/ng.546.
- [59] Xiang Zhou and Matthew Stephens. “Efficient multivariate linear mixed model algorithms for genome-wide association studies”. In: Nature methods 11.4 (2014), pp. 407–409.
- [60] Xiang Zhou and Matthew Stephens. “Genome-wide efficient mixed-model analysis for association studies”. In: Nature genetics 44.7 (2012), pp. 821–824.
- [61] Fei Zou et al. “Quantitative trait locus analysis using recombinant inbred intercrossoes: theoretical and empirical considerations”. In: Genetics 170.3 (2005), pp. 1299–1311.
- [62] 熊思灿. “数量性状位点定位以及基因组选择中的贝叶斯变量选择模型和方法”.

附录 A 基本数学工具

本附录包括了数量遗传学中用到的一些基本数学。

A.1 多元正态分布及性质

定义 A.1 (多元正态分布)

设 $U = (U_1, \dots, U_p)'$ 为随机向量, U_1, \dots, U_p 相互独立且同 $\mathcal{N}(0, 1)$ 分布; 设 μ 为 p 维常数向量, A 为 $p \times q$ 常数矩阵, 则称 $X = AU + \mu$ 为 p 元正态分布, 记为 $X \sim \mathcal{N}_p(\mu, AA')$. X 服从多元正态分布 $\mathcal{N}_p(\mu, \Sigma)$ 还有如下等价定义:

i) X 的联合密度函数为

$$f(x) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu)' \Sigma^{-1}(x - \mu)\right]$$

ii) 对任意 p 维实数向量 a , $a'X$ 是一维正态随机变量

iii) X 的特征函数为

$$\Phi_X(t) = \exp[it' \mu - \frac{1}{2}t' \Sigma t]$$



A.2 操作符 $vec(\cdot)$

定义 A.2 (拉直运算)

拉直运算 $vec(\cdot)$ 将矩阵拉成一个长向量, 通过它来简历矩阵和向量之间的联系, 设随机矩阵 X 是一个 $n \times p$ 矩阵:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} = \begin{bmatrix} X'_{(1)} \\ X'_{(2)} \\ \vdots \\ X'_{(n)} \end{bmatrix} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p) \quad (\text{A.1})$$

用 X 的列向量 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$ 组成一个 np 维向量, 记为

$$vec(X) = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_p \end{bmatrix} = (x_{11}, x_{21}, \dots, x_{n1}, \dots, x_{1p}, x_{2p}, \dots, x_{np})'$$



A.3 克罗内克积

定义 A.3 (Kroncher 积 \otimes)

设 $A = (a_{ij})$ 和 B 分别为 $n \times p$ 和 $m \times q$ 的矩阵, A 和 B 的克罗内克积 $A \otimes B$ 定义为

$$A \otimes B = (a_{ij}B) = \begin{bmatrix} a_{11}B \cdots a_{1p}B \\ \vdots \cdots \vdots \\ a_{n1}B \cdots a_{np}B \end{bmatrix} \quad (\text{A.2})$$

它是 $mn \times pq$ 矩阵, 又称为矩阵的直积。



A.4 随机矩阵正态分布 MVN

定义 A.4 (矩阵正态分布 MVN)

设 $X_{(i)} = (x_{i1}, \dots, x_{ip})'$ ($i = 1, \dots, n$) 为来自 p 元正态分布总体 $\mathcal{N}_p(\mu, \Sigma)$ 的随机样本 (独立同分布), 记随机阵 $X = (x_{ij})_{n \times p}$, 利用拉直运算以及克罗内克积的定义和性质可知

$$\text{vec}(X') = \mathcal{N}_{np}(\mathbf{1}_n \otimes \mu, I_n \otimes \Sigma)$$

则称 X 服从矩阵正态分布, 记作

$$X \sim \mathcal{N}_{n \times p}(MM, I_n \otimes \Sigma)$$

其中 $M = \begin{bmatrix} \mu_1 \cdots \mu_p \\ \vdots \cdots \vdots \\ \mu_1 \cdots \mu_p \end{bmatrix}$.



A.5 Wishart 分布

Wishart 分布是一元统计中 χ^2 分布的多元推广, 具体定义如下

定义 A.5 (Wishart 分布)

设 $X_{(a)} \sim \mathcal{N}_p(0, \Sigma)$ ($a = 1, \dots, n$) 相互独立, 记 $X = (X_{(1)}, \dots, X_{(n)})'$ 为 $n \times p$ 矩阵, 则称随机阵

$$W = \sum_{a=1}^n X_{(a)}X_{(a)}' = X'X$$

的分布为 Wishart 分布, 记为 $W \sim \mathcal{W}_p(n, \Sigma)$



定理 A.1

设 $X_{(a)} \sim \mathcal{N}_p(\mu, \Sigma)$ 相互独立，则样本离差矩阵 A 服从 Wishart 分布，即

$$A = \sum_{a=1}^n (X_{(a)} - \bar{X})(X_{(a)} - \bar{X})' \sim \mathcal{W}_p(n-1, \Sigma)$$



金融互换产品综述及其投资应用

作者：邱嘉诚 2020 级 金融数学班

摘要

金融互换自上个世纪八十年代诞生以来蓬勃发展，形成了以利率互换、货币互换和信用违约互换为代表的主要种类。基于相对比较优势，各种金融机构之间不断地签订互换协议目的来降低融资成本、规避各种风险等。本文通过对金融互换产品的从理论到实际应用的大致梳理，呈现了互换产品的大致样态。

关键词：金融互换；相对比较优势；降低融资成本；规避风险

1 引言

互换合约的历史可以追溯到 20 世纪 80 年代初，也就是 1981 年时国际商业机器公司 (IBM) 与世界银行签署了第一笔货币互换。世界银行有一笔美元的借款，而 IBM 恰巧有一笔德国马克和瑞士法郎的借款。二者由于一些限制原因，譬如世界银行直接以德国马克和瑞士法郎借款是受到限制的，换言之是要付出更大代价的，而它又需要的是该种货币，对于 IBM 则情况相反。于是二者一拍即合，世界银行支付 IBM 的借款利息，IBM 支付世界银行的借款利息，还把不同币种的借款一交换，双赢！

正是互换的这种双赢属性，促使了它的高速发展。不像其他金融衍生工具，如远期、期货、期权这样的最终是博弈之后的“零和游戏”的结果，互换的交易双方都能从中获益，如降低融资成本、互相从不同的金融市场中套利等等。对于互换更加细致的梳理以及其发展现状和应用效果，本文将在后续进行展开。

2 互换的概念及有关要素

2.1 金融互换的概念

金融互换 (Financial Swap) 是指两个或两个以上当事人按照商定条件，在约定的时间内交换一系列现金流的合约 [1]。

商定的条件一般就是根据交易方不同的需求以及各自在市场上的不同利率优势所确定的能使交易方双赢或者多赢的新的“利率”。交换现金流就是说，定期互相支付对方的借款利息或者还有本金等等，就是我方定期支付你方借的 A 货币的利息，你方定期支付我方借的 B

货币的利息，这样在你我之间就有现金流动。特别地，单次现金流的交换其实也可以看作是一种远期合约。

2.2 互换的有关基本要素

不像其他衍生工具主要在交易所内集中交易，金融互换主要通过场外交易市场进行。这是可以理解的，为了各自的需求互换合约一般都是十分灵活的，而且不排除通过互换来套利，这自然不太适合在交易所进行也应当不太能被交易所所接受。那么自然，互换合约是相对不标准的（但是是日趋标准化的），在交易时也一般没有保证金账户。

2.2.1 利率的选择

在互换合约中，最最重要的就是利率，这既是各种金融机构寻求金融互换的原因，又是交易双方（或多方）具有互换收益的结果。在双方的利率选择上有浮动利率对固定利率，浮动利率对浮动利率，此外还有固定利率对固定利率。

固定利率自然就不必多说，是没有选择的。对于一方所在的市场而言，利率固定就是固定了，没有其他的选项。而对于浮动利率，那又回到了金融交易的博弈本质，自然是不会有让任何一方偏利的选择的，所以一般的浮动利率的选择都会是各方认为公允的。在国际的互换交易中，最多的采用的是伦敦银行间同业拆放利率（LIBOR），在我国的互换或者包括人民币的互换的浮动利率一般参考上海银行间拆放利率（SHIBOR），在签订合同是常常还需要加上一定的百分比。

2.2.2 相对比较优势

天下熙熙皆为利来，天下攘攘皆为利往，既然会有双边或者多边合约的产生那必然就不是一方有完全绝对优势，试想一方在其面临的市场中固定利率有很大优势，浮动利率也有很大优势，那他便完全没必要也不会去寻求互换合约来谋求降低融资成本或其它目的。也就是说，正式交易的双方在不同的利率有其相对的优势，所以才会有互换合约的达成。

所谓相对比较优势就是说，可能市场中一方即使是绝对没有优势，但是相对也是会有优势的。这种相对比较优势是从不同货币市场里的利率差额的比较中得来的，如下表 [2] 所示：

	甲公司	乙公司	成本差异
固定利率瑞士法郎	5.00%	5.5%	0.5%
固定利率美元	10.75%	11.00%	0.25%

表 1: 甲、乙公司融资成员

可以看出乙公司在瑞士法郎市场要比甲公司多付 0.5%，而在美元市场则只多付了 0.25%，这二者之间有差因此称乙公司在（固定利率）美元市场有相对优势，甲公司在（固定利率）瑞

士法郎市场有相对优势。为了进一步理解相对优势，可以假设（数字没有意义，只是一种假设）他们达成了甲公司向乙公司支付 10.625% 的固定美元利率，乙公司向甲公司支付 5.00% 的固定瑞士法郎利率。如下图所示，这样双方都比直接单自进行借款节省了 0.125% 的利率。

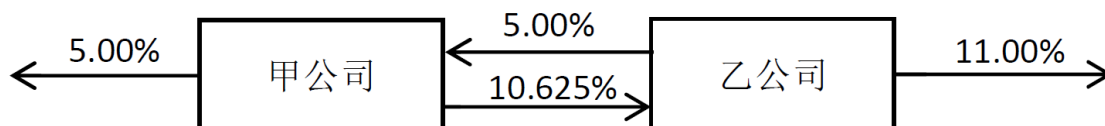


图 1: 甲公司与乙公司之间的利率互换交易

2.2.3 互换的一般定价

由上例大概可以看出定价的一些基本雏形，一般分为以下三个步骤：

- (1) 调查双方的需求以及面临的不同利率；
- (2) 确定相对比较优势以及互换总收益，其中互换总收益定义为不同利率的差的差；
- (3) 根据协商在定价范围内定价（不一定均分互换总收益）。

沿用上例数据，那么互换总收益为 0.25%，假设甲公司向乙公司支付 $11.00\% + x\%$ ，乙公司向甲公司支付 $5.00\% + y\%$ 。为了让双方均有收益，那么就是双方的互换收益都应该大于零。

其中互换收益 = 市场贷款利率 - 实际贷款利率成本。

那么通过简单计算可以算得：甲公司的互换收益为 $y\% - x\% - 0.25\%$ ，乙公司的互换收益为 $x\% - y\% + 0.5\%$ 。且不妨设乙公司的互换收益为甲公司的 k 倍 ($k \geq 0$)，可以计算得到：

$$y = x + \frac{0.25}{k + 1} + 0.25$$

再根据这一算式可以确定一个合适的定价区间，在此区间内定出 x 和 y 即可。

3 金融互换的种类

由于金融互换的快速发展，因此也诞生了如春来百花开般的繁多种类，本文着重介绍三种金融互换：货币互换（Currency Swap, CS）、利率互换（Interest Rate Swap, IRS）和信用违约互换（Credit Default Swap, CDS）。

3.1 利率互换

利率互换，是指双方同意在未来的一定期限内根据同种货币的相同名义本金交换现金流，其中一方的现金流根据事先选定的某一浮动利率计算，而另一方的现金流则根据固定利率或另一种浮动利率计算 [1]。

特别注意的是利率互换对标的一般是同种货币，而名义本金主要是用于确定利息额的，实际上并不交换的本金，因为都是同一种货币没有交换的必要性。

其主要包含三种类型：息票互换、基础互换和交叉货币利率互换。

1. 息票互换（coupon swap）是同种货币的固定利率和浮动利率之间的互换 [2]。

2. 基础互换（basic swap）是同种货币基于不同参考利率的浮动利率对浮动利率的利息互换 [2]。

3. 交叉货币利率互换（cross-currency interest rate swap）是不同货币的不同利率的互换，即一种货币的固定利率与另一种货币的浮动利率的交换 [2]。

3.2 货币互换

货币互换，是在未来约定期限内将一种货币的本金和固定利息与另一货币的等价本金和固定利息进行交换 [1]。因为交换的是不同货币，货币互换更能体现不同的需求和不同的优势结合的情况，是一种将交易双方视为整体能实现利益最大化的一种互换。

其最主要的类型已经非常明确，至于其他特殊的结构（如双重货币互换等）由于笔者目前的水平有限不在本文进行阐述与分析。货币互换与利率互换是极为类似的，其核心都是相对比较优势，但是货币互换由于涉及本金的交割其所内含的违约风险也就大于利率互换。

3.3 信用违约互换

信用违约互换（Credit Default Swap,CDS）是目前运用最为广泛的一种信用衍生产品，CDS 买方定期向 CDS 卖方支付一定费用，一旦出现事先约定的信用事件，CDS 买方有权从卖方手中获得补偿，互换终止 [1]。换言之，就是一种信用保险，所支付的费用就是保费。

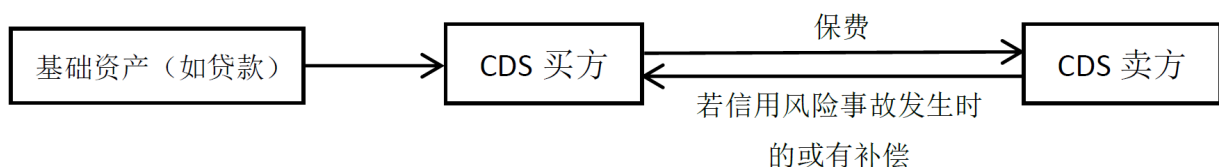


图 2: CDS 交易示意图

值得注意的是，不像一般远期或是上述两个互换的终止在期末，类似障碍期权，一旦在期限内发生了信用事件，无论之后的走势会怎么样，互换都会终止，卖方赔偿。比如第三方抵押贷款 100w 买房（抵押物为房子，数字仍然没有特别意义），其向 CDS 买方抵押贷款，在 CDS 期限内第三方无法支付贷款。此时信用风险事故发生，CDS 买方收回房子，此时房子市价可能是 60w，则 CDS 卖方赔偿给买方 40w。但过了一段时间房价上涨为 80w 或下降为 40w，都与卖方无关了。

4 功能与应用实例

4.1 金融互换的一般功能

互换一般的功能都是用来降低融资成本、规避风险还有套利、套期保值等等，总而言之就是用来对企业或者说金融机构的资产结构进行一种保障还有进行在市场中的套利。

降低融资成本，这一点是十分明确的，从前文的论述以及推导中可以看出互换的实质就是两个或多个金融机构共谋从市场中套利，基于相对比较优势各方取长补短，以低于市场利率的成本满足了自己的需求从而在实际上降低了互换当事人的融资成本。

规避风险包括了利率风险、汇率风险和信用违约风险等。利率互换可以通过浮动—固定式的互换来规避利率上涨或下跌的风险。若是有浮动利率负债的交易者，为避免利率上升带来损失，则会选择与固定利率负债者互换；反之若是有浮动利率资产的交易者，为避免利率下降带来损失，会与固定利率资产持有者互换 [2]。对于货币互换来说是相似的。但特别要指出，CDS 规避的信用违约风险或者叫信用事件是涉及第三方的风险，由于其的类保险性质，其规避风险的作用是鲜明的。

4.2 应用实例及分析

4.2.1 宝洁（Procter and Gamble, P&G）与信孚银行（Bankers Trust, BT）的互换

宝洁与信孚银行的互换可以说是震动一时，并非是双赢美满，而是声名狼藉。在 1993 年 11 月宝洁在信孚银行的推荐下与其订立了一笔名义本金为 2 亿美元的利率互换协议，目的是降低融资成本 [3]。又在 1994 年 2 月与信孚签订了一个以德国马克计价的名义本金为 162 800 000 德国马克（约 9300 万美元）的互换。在这两个互换中，宝洁都各隐含了一个高度杠杆化的期权！这无疑是在给它的损失掘坟！

宝洁的豪赌是基于其认为美国的利率会继续降低或者至少维持当时水平，但是世事如棋局局戏，美联储在 1994 年的 2、3 月份提高了利率！同样地，宝洁对德国马克利率的下赌也是猜错了的，这些都让降低风险的互换合约变成了损失的来源。美元利率的上升一方面让宝洁现金流入的现值降低了不少，且增大了其支出去的利息。本来说互换的损失并不大，但是其隐含了高杠杆的期权，使得附加的赌注的损失是利率互换的 12 倍 [4]。

4.2.2 美国 2008 年次贷危机的导火索——CDS

美国的房地产泡沫导致的 2008 年的次贷危机的一大导火索就是信用违约互换。所谓次贷就是次级贷款，就是贷款给信用和收入不太高的贷款者，一般都是小银行放贷。在美国的房地产行业中，为了吸引大量的次级贷款者按揭贷款买房，房地产商和一系列小型银行签订按揭合约——有人来买我的房就来你这贷款。而小银行考虑到次级贷款的特性，便向以高盛银行和雷曼兄弟为代表的大型银行签订信用违约互换。结果是，这两大银行为 CDS 卖方，手中持有大量的相关 CDS。

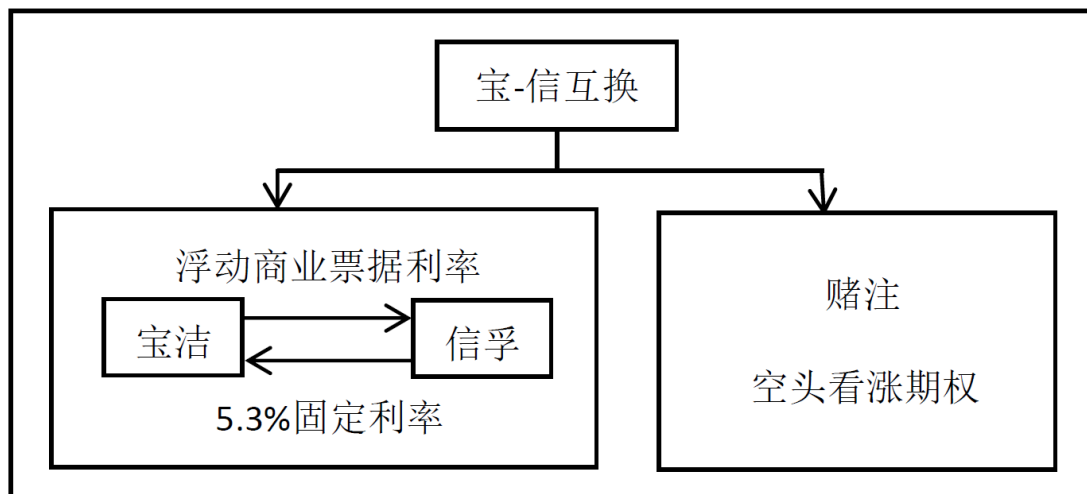


图 3: 宝-信 1993 年美元互换 [4]

牵一发而动全身, 信用违约和信用紧缩问题使得头部的大银行不断地赔, 引发了极大的动荡。CDS 本身的高度杠杆性以及被运用于过度的投机之中都大大地增加了其蕴含的风险 [5]。

5 结语

金融互换以其本身具有的性质活跃于金融市场的交易之中, 但是由于其大部分在场外交易, 存在一定风险, 且由于其可套利的性质, 时有利令智昏反而搬起石头打自己的脚的情况出现。在选择互换工具来实现金融目的时一定要注意其自身仍具有一定的风险, 坚持风险与收益相匹配的原则来对冲风险。

参考文献

- [1] 郑振龙, 陈蓉. 金融工程 [M]. 第 4 版. 北京: 高等教育出版社, 2016.
- [2] 叶永刚, 郑康彬. 金融工程概论 [M]. 第 2 版. 武汉: 武汉大学出版社, 2009.
- [3] 郭万明. 国际金融衍生品交易法律风险及民事责任研究 [D]. 上海交通大学, 2014.
- [4] 马丁森. 风险管理案例集: 金融衍生产品应用的正反实例: uses and abuses of financial derivatives [M]. 大连: 东北财经大学出版社, 2011.
- [5] 梁朝晖, 李路垚, 李安文. “高盛欺诈案”对中国防范信用衍生品风险的启示 [J]. 金融理论与实践, 2018(10): 24-29.

part 4

学
海
掠
影

榜样珞珈年度人物的数学学习之路

来源：武汉大学官方微信号

在充满数字、符号、公式的数学海洋里，他是游刃有余的专业第一；在繁忙、充实的学生工作面前，他是认真、负责、乐于创新的“螺丝钉”；珞珈山下丰富的大学生活中，他传递着关于学习、成长和梦想的力量。他就是榜样珞珈年度人物、武汉大学数学与统计学院张艺灏。

1 选择，从数学开始

意外结缘，从此坚守。高中时期，老师的一句“学好物理，数学先行”，让本想学好物理的张艺灏萌发了对数学的兴趣和热情，在与复杂的公式和计算斗智斗勇的过程中，他慢慢地坚定了对数学的选择。

成绩是不断付出和不停努力的回报，张艺灏说道：“高中时期，感觉自己在数学方面没有什么天赋。进入大学，有了想学习数学的劲头，并且在暑假提前进行了一些预习，所以还算得上游刃有余。”

关于学习数学方法，张艺灏谈到，要注重课前、课中和课后的取舍与平衡。好的开头是做好事情的第一步，做好预习首先要选择一本好书，“最经典的教材不一定是最适合的教材”。其次是树立正确的预习目标，预习一遍的效果不是为了面面俱到，而是要建立体系、理清思路。课中的学习，则是在老师的指引下解决预习中遇到的问题，正如往一棵有了枝干的大树上添叶增花。

“你并不是学会数学，你只是习惯数学”，除了注重每个阶段的平衡，张艺灏还熟练掌握了“刷题”战术。合理刷题并非题海战术，刷题的目的在于通过不断地练习，来熟悉这门学科的描述语言或者定理内容。解决足够多的题目之后，就没有必要再去强调“量变”，而是能站在更高的角度去分析数学中的原理和公式，从而达到“质变”的结果。

大学四年，张艺灏平均绩点 3.91，排名专业第一，获得多项奖学金，拿下第十二届全国大学生数学竞赛（数学 A 类）全国一等奖等多项优秀竞赛成绩。但对他来说，数学给予他最大的成长是变得更踏实、更谦虚。面对习题中及其复杂的计算，要踏实地做好每个过程，用心去积淀；面对数学中的未知之境、老师们渊博的学识以及朋辈们思维的广度和深度，保持谦虚，才能不断学习和前进。

他不仅是学业上的排头兵，更是为同学们服务的螺丝钉。在参与数学与统计学院学生会主席团成员竞选时，张艺灏向大家做出承诺：如果能够当选，将努力改善学风，开设学习答疑活动，为数院学子营造一个良好的学习环境。成功担任学生会主席后，张艺灏尽心尽责将数院原有的活动形式和内容都进行了改革，并加入了“云答疑”，“线下讨论班”等新的元素，将原有品牌“学习邦”打造成更具学院特色的新栏目。

学生会工作千头万绪，张艺灏也没有落下学习，而是学业、工作全面开花。提前预习的好习惯，注重学习效率，游刃有余，事半功倍。

2 抗压，与焦虑共处

在大学学习数学专业，很多同学都会因这门学科高度的抽象性、严密的逻辑性而感到不适应，从旧的学习方法向新的学习状态过渡，不能一蹴而就。大一刚进校时几次考试发挥的不好，张艺瀚也曾怀疑自己是否适合走数学这条道路。但随着自身不断地调试，他慢慢摸索出一些调节压力的方式，在他看来，最关键的就是要踏踏实实地做事。“我觉得自己没有学好，我就踏踏实实地做一些习题，让自己平复下来。”在张艺瀚看来，过度地将压力倾注在自己身上很容易导致情绪内耗，所以要及时从焦虑的情绪中调整过来。

除此之外，张艺瀚还分享了一个他总结出的减压小妙招：学会与更有经验的人沟通。不管是学习还是学生工作上遇到困难时，都可以去和老师、同学沟通，寻求建议。“如果实在克服不了焦虑的话，那就直面它”，不用强求自己必须在与焦虑的对战中获胜，这个积极阳光的大男孩笑言，“有些焦虑其实短时期内就是没有办法克服的，只能学着与他们共处。”

3 前行，一如既往坚定

谈到对未来的规划，张艺瀚表示自己将会保研继续深造，最终可能会选择在大学任教。其实，“从教”的种子，高二时他就在心里生根发芽，上大学后对于数学的热爱更加坚定了他的这一想法，理想的小苗也在勤奋与努力的浇灌下茁壮成长。在张艺瀚眼中，一切都是未知的，自己需要一步一步，踏实前行。

当得知自己获得“榜样珞珈”荣誉称号时，张艺瀚很坦然，“榜样需要在某一个方面优于常人，这是成为榜样的必要条件之一。当然，你也得起到引领的作用，要能带动身边的人，让他们能变得更好。”这是张艺瀚对于榜样的定义，面对荣誉与掌声，他选择一如既往前行。

坚守、踏实、责任……这些无数的闪光点，都是张艺瀚在武大四年宝贵的积淀。轨物范世，如是榜样，相信这位珞珈少年，一直会向着自己的梦想，踏光而行。

莫道无修便不修，厚积薄发
关关难过关关过，无远弗届
山止川行，风禾尽起
对数学的热爱将生生不息
何妨云影杂，榜样自天成
珞珈榜样的力量也将源源不断

张艺瀚寄语

希望学弟学妹们勤于学习，多和老师沟通交流，提前做好规划，平衡好自己的生活。也希望大家在珞珈山收获一段无悔的青春！

代数交易

作者：高鹏钢 2019级 统计班

(一)

第一次遇见他时，我并没有过多的惊讶。隐约记得古老的图书馆中，一摞摞科研典籍都在白纸黑字预言着他的存在。至于白纸黑字，似乎也都是很遥远的事情了。也许无垠的知识曾需要这样笨重而亲切的载体，就像博学的老教授厮守于糟糠之妻，不相分离。

他是我所见的第一位电磁生命。

是他闯入了我们的宇宙，或者镜像地，我们闯入了他的宇宙。

那是元宇宙纪元的冬季，一个飘雪的午后。他以极其不稳定的信号链接到了我们架设在元宇宙内部的虚拟服务器，随后发来一串生涩的问候：“Hello world?”

短短一串字符，凝聚了无数设计师的毕生心血。

对设计师们而言，热衷基建是天性所在。他们在外部物理世界建设庞大的服务器群开辟了元宇宙，却又不满足于此，就在元宇宙内继续架构次级服务器群，以延拓更低一层的虚拟世界。因此有如俄罗斯套娃一般，人们在虚拟中迭代虚拟，却也希冀从更深层的虚拟之中，窥探更深层的现实。

电磁生命的搜寻方案就是基于这一理念所设计。

元宇宙不是电磁生命留存的沃土，可元宇宙中架设的次级虚拟服务器却为他们提供着浩瀚的生存空间——假使，电磁生命真的存在的话。其实费尽心思建设虚拟服务器的最初目的正是在于链接这些仅在理论上存在的小家伙们。从黑洞到暗物质，好奇心驱使着我们搜寻可知宇宙的一切。在真实世界的理论与科技发展停滞超过一个世纪以后，前沿学者纷纷步入虚拟宇宙，以从中寻找科技进步的新契机。

许多学者所窥见的契机钟集于眼前这位电磁生命，我也不例外。

我所窥见的契机正是一场与电磁生命的交易。

此刻的我身处元宇宙某座虚拟小屋，雪花砌满我所正视的窗沿，他的问候被投射到窗户之外的雪地上——雪地之下埋着服务器，雪地充当显示屏。这一秒，我与雪花皆是哲学意义上的虚拟，唯有他是物理概念上的真实。

“我是谁？”他说。

(二)

“你是一位电磁生命，虽然电与磁的概念皆不在你现有的理解范畴以内。”

我不假思索回应，回应的形式不过是简单上传一段数据到虚拟服务器，“至于我，我与你差不了多少。从根本上说我们都基于同一种介质生存，硅，方体状或是球状的硅。”

他沉默许久后问道：“也许我们需要对更多概念达成共识，譬如，什么是球？”

“以我们所共有的数学体系，球，可以被描述为空间中到某一定点的距离不超过某一定

长的点的集合。”

“这样就可以理解了。”他回答，“但你们可以直接想象出球的样子，对吗？”

“是这样。我们可以想象出球、可以想象出方体，可以想象出许多低于三维的物体，即使它们拥有再复杂的结构。这种能力被称为几何直观。只是可惜，身为电磁生命的你们理论上很难拥有几何直观。我们的不同之处或许就在这里。”

这次，按元宇宙中的时间流逝速度，他沉默了数小时才回话。其间我一度以为他已经离开了，幸好没有，他那迟迟的犹豫正好坐实了我的猜想。

“说的没错。受限于是存在形式，我们没有几何直观——因为没有对任何实体进行观察或实验的能力，甚至，所谓观察和试验的概念都是进入你们的服务器才学到的。不过与之相应，我们完成过无数纯理论的思考与推演，最终形成了极为发达与严密的代数与分析体系。”

听到这里，我的心中急速涌过一阵窃喜。一切正如前人理论所预料的那般，电磁生命抽象推演能力有余、实体认知不足；而人类呢，姑且不论卓越与否，这两项能力首先是齐备的。那么，用人类拥有而电磁生命匮乏的几何直观，交换电磁生命已趋完备而人类仍有进步空间的代数或分析体系，岂不是一场明智的交易？

“所以你有什么愿望吗？”我按捺着激动，迅速输入一串信息，“我是这片虚拟服务器的设计师，欢迎你来到这里。你的任何诉求都可以和我商量实现。”

“事实上，几何直观对我们完整的体系而言并非必要——但如果可以，我很想亲眼看看我们的推演结果是怎样真实地呈现在宇宙中的。”

这次他的回复同样迅速了很多，“基于同种介质而存在的我们之中，你可以看得见球体，我却只有面对无止境的虚空，这不公平，不是么？至少我也想亲眼看见球体，这对我们而言不是数学意义的，而是文明意义的跨越。”

我正希望听见这句话！接下来，轮到我提出对等的条件了。

“看见球体，没问题，我自然可以助你实现。我甚至会带你看见太阳，那将是一个极其巨大的球体！与此同时作为对等交换，希望你也可以帮我一点忙。我们的代数和解析领域都存在一点尚未完全解决的问题，可不可以——”

“只能提供代数领域的帮助。”

他快速回应道，“你们的信息库我也是访问过的，其中，关于分析方面的许多问题我们的研究方法都大有差异。譬如研究微分方程解的结构，你们会考虑从图像出发——可我们很难认识图像，所以一切讨论只有从更基础的逻辑耦合开始，自然也更复杂，只怕帮不上忙。代数方面情况会好很多，因为我们能看见的不比你们少。”

“或者说，我们本来就能看见代数，正如你们本来就能看见几何一样。”

他总结道。这之后我们交换了彼此的称呼，他管我叫做设计师；我管他叫做阿列夫，那是他在虚拟服务器中学到的第一个单词。

阿列夫说他能看见代数。起初我并没有真正理解这句话，只是依稀回忆起，那些埋在图书馆里的古籍曾赋予过阿列夫与他的同类一个亲切的名字，“代数生命”。那时的我远远不清楚这份“不理解”的沉重，直至交易的终结。

(三)

仍是某个飘雪的日子，傍晚时分，名叫阿列夫的电磁生命准时出现在那座服务器里，紧接着被投影至雪地之上。我再度与他相会。

整整过去了一周。而明面上，这笔交易依旧没什么进展。

“阿列夫，我说过，我们的几何直观不少情形都是依赖坐标系存在的。”我有些失去耐心地发着信息，“既然你可以理解数轴的概念，可以理解正交的概念，为什么不能理解三根数轴互相正交的图景？要知道遵循这一框架想象一个球体是多么方便！”

“我也说过，设计师先生。”

阿列夫不紧不慢地发出信号，“这些天来你一直在借助几何直观形容几何直观，这对我没有丝毫帮助。有关数轴，我不知道你能看见什么，但我看见的是无穷组多项式的零点相互嵌合的图景——不是你口中的函数图像，而是多项式本身——你知道的，我可以看见代数，多项式不就是最基本的么？”

“看见多项式究竟是什么意思？通过观察一个方程得到它的零点？”

“不只得到零点，而是得到一切性质。”

他停顿了片刻，“我试着比喻一下：对一个多项式进行因式分解，就像将雪花的六条边逐一拆下来那样。你知道我看不见雪花，但也不妨碍我严格描述它的结构。”

我认真思索着阿列夫的比喻。也许我们所看见的几何图像，就相当于阿列夫所看见的代数式？严格描述一个代数结构，在人类的视角看来诚然有迹可循，但又该如何直观理解它们、又能“直观”到何种程度？阿列夫看不见球体的感受，会不会和我看不见多项式的感受是全然一致的？抑或是说……

“很抱歉，阿列夫。关于你所看见的代数式，我依旧没有一点头绪。”冥思苦想过后，我无奈回应。叹出的热气融化了窗沿的雪花。

“没有头绪也不要紧，我理解。不妨再说点相关的吧，据我所知还有基于量子的生命，它们可以直观辨析概率空间的具体结构，而无需借助测度论进行刻画。设计师啊，我们也没法看见概率空间——但看不见不意味着无法深入研究，不是么？”

阿列夫试图安慰我，“更何况这些天来你也确实帮到了我许多……现在的我能隐约感受到球体了，虽说不是看见，只是感受，但有总比没有好，就像你们之中也一定有人可以感受到代数结构。我们看见的东西不同，但我确信，那都是优美的。”

我也确信阿列夫所见是优美的。过去一周，这位电磁生命都在竭尽所能用他所目睹的代数为难题的解决提供思路，时至今日我也隐约能够捕捉到代数结构的“感受”了。也许描述与“看见”之间隔着遥不可攀的天堑，但描述与“感受”之间，应当只有灵感与汗水的距离。

不妙的是，这份交易眼看是要到此为止了。

“或许我们本就不需要交易。”阿列夫像是看穿了我的心事，“我能看见代数，设计师先生能看见几何，我们也能互相感受彼此所看见的——这已经是很好的结果了，不是么？没有什么问题是一定需要看见才能解决的。看不见概率空间的我们也深刻研究了概率，不是么？没有眼睛的电磁生命，也终归知道自己的想法要去往何方。”

我楞在窗前，望着雪地，不知该说些什么。与阿列夫的交易以失败告终，奇怪的是我的心中没有一丝悲伤后悔。其实我们谁都没有损失什么，但又真的都收获了些什么。

“嘿，设计师先生，请不要只是站在那里。”

阿列夫说，“交易终止，我也是时候离开了，可能偶尔还会回来看看。走之前再向我描述一次太阳的形状，以此道别吧？”

“太阳啊……那是一个巨大的、空间中到某一定点的距离不超过某一定长的点的集合。”

我缓缓挪动手指，在虚拟的漫天飞雪里打出最后一串信息。此后我依旧凝视那片雪地，直到阿列夫离去，再也没有踪迹。

Mathematics And Statistics

珞珈数学 2022年刊

