同济大学中德工程学院

CDHAW of Tongji University

DSGE 模型笔记

朱彦元

Dr. Yanyuan Zhu
yyz@tongji.edu.cn

2018 年 3 月 8 日

目录

第一章　前言 **11**

第一部分　模型求解 **13**

第二章　DSGE 模型求解方法简论 **15**

2.1 简论的简论 15
2.2 pros and cons 17
2.2.1 cons 17
2.2.2 pros 17
2.3 工作流程 18
2.4 模型设定 19
2.5 求解方法 20
2.5.1 近似线性展开 20
2.6 求解线性随机差分方程系统 22

第三章　动态规划 **23**

3.1 包络定理 23
3.2 例: 吃蛋糕 (直接求解法) 24
3.3 吃蛋糕 (动态规划) 25
3.3.1 有限时间段的动态规划问题 25
3.3.2 无限时间段的动态规划问题 27
3.4 随机问题 29
3.5 包络定理和欧拉等式的关系 29

第四章　线性二次最优控制 **33**

4.1 目标方程和状态转移方程 33
4.2 价值方程 33
4.2.1 一阶条件 34
4.3 政策方程 35
4.3.1 政策方程满足确定性等价条件 35
4.4 线性政策方程和二次价值方程 35
4.4.1 价值方程满足确定性等价条件 36
4.4.2 代数矩阵 Riccati 方程 37
4.5 数值方法 37
4.6 范例 37

4.6.1 最优线性二次控制问题 38
4.6.2 Matlab 程序 38
4.6.3 Matlab 程序输出 40

第五章 理性期望模型 43
5.1 简介 43
5.2 数值算法 44
5.3 一个简单的随机增长模型 45
5.3.1 稳定状态 45
5.3.2 对数线性化 45
5.3.3 状态——空间表现形式 47
5.4 特征值——特征方程分解法 47
5.4.1 特征值——特征方程分解法 47
5.4.2 解的存在性以及唯一性 49
5.4.3 应用 Blanchard-Kahn 算法实例 50
5.5 未定系数法 51
5.5.1 模型范例 52
5.5.2 未定系数法求解 52
5.5.3 求解系数矩阵 A, B 53
5.5.4 存在性及唯一性的探讨 54
5.6 期望误差法 54
5.6.1 模型范例 55
5.6.2 期望误差法求解 55
5.6.3 存在性及唯一性的探讨 57
5.7 参数化期望法 57
5.7.1 参数化期望法示例 58
5.8 Schur 分解法 59
5.8.1 Schur 分解的定义 60
5.8.2 爆炸分块矩阵的解 60
5.8.3 稳定分块矩阵的解 61
5.8.4 原经济系统的解 61

第六章 扰动法 63
6.1 扰动法简介 63
6.2 分析框架 63
6.3 求解方法 64
6.3.1 状态——空间形式 64
6.3.2 稳定状态 65
6.3.3 外生随机过程 66
6.3.4 设想解 68
6.3.5 一阶扰动 69
6.3.6 二阶扰动 74
6.3.7 更高阶扰动 76
6.3.8 例 77

6.4 剪枝算法 82
6.5 变量的变换 82
6.5.1 例 1 83
6.5.2 例 2 84
6.5.3 最优变量变换：参数的选取 85
6.5.4 对数线性化和对数正态线性化 85
6.6 价值方程迭代中的扰动法 87

Appendices **91**
附录 6.A Schur 补和对称正 (半) 定矩阵 91
6.A.1 Schur 补 91
6.A.2 利用 Schur 分解法判断可逆正定矩阵 93
6.A.3 利用 Schur 分解法判断可逆正半定矩阵 94
附录 6.B 伪逆矩阵 97
6.B.1 用 SVD 法做伪逆矩阵分解 97
6.B.2 用 SVD 法做 M 的矩阵分解 99

第七章 映射法 **101**
7.1 简介 101
7.1.1 例 102
7.2 全局基的选取——单维基光谱法 104
7.2.1 单项式 104
7.2.2 三角序列 105
7.2.3 雅各比多项式 105
7.2.4 切比雪夫多项式 105
7.3 局部基的选取 (1) 109
7.3.1 离散状态变量 110
7.3.2 张量与完全多项式 111
7.4 局部基的选取 (2)：有限元法 112
7.5 目标方程的选取 115
7.5.1 最小方差 116
7.5.2 子域 116
7.5.3 配点 117
7.5.4 伽辽金法或瑞利——里兹法 117

Appendices **119**
附录 7.A 超几何方程 119
附录 7.B 多项式 124
7.B.1 正交集 124
7.B.2 正交多项式 125
7.B.3 递推关系 127
7.B.4 雅各比矩阵和克里斯托费尔——达布关系 128
7.B.5 正交多项式的根 130
7.B.6 高斯求积 131
附录 7.C 常见的正交多项式类型 132

7.C.1 埃米特多项式 133
7.C.2 拉盖尔多项式 137
7.C.3 雅各比多项式 141
7.C.4 勒让德多项式 145
7.C.5 切比雪夫多项式 148
附录 7.D AR(1) 过程的离散方法 155
7.D.1 Tauchen(1986) 法 155
7.D.2 Tauchen and Hussey (1991) 法 158
7.D.3 Rouwenhorst(1995) 法 159
附录 7.E 伽辽金近似 160
7.E.1 举例：1 维空间中的伽辽金近似 160
7.E.2 伽辽金近似 162
7.E.3 强制型方程形式的伽辽金近似收敛 165
7.E.4 仅满足上极限——下极限关系方程形式的伽辽金近似收敛 166
附录 7.F 数值求积 166
7.F.1 牛顿——寇特斯法则 167
7.F.2 几个小技巧 172
7.F.3 求积中的可积奇异点问题 173
7.F.4 梯形法则的近似误差 174
7.F.5 克伦肖——柯蒂斯求积法则 179
附录 7.G 傅里叶分析 186
7.G.1 傅里叶变换 187
7.G.2 脉冲方程的傅里叶变换 190
7.G.3 非脉冲方程的傅里叶变换 192
7.G.4 原方程的平滑和傅里叶变换方程的衰减 193
7.G.5 傅里叶级数 193
7.G.6 泊松求和式 199
7.G.7 傅里叶分析与卷积 201
7.G.8 高维傅里叶变换 203
7.G.9 一些指数求和的背景知识 203

第八章 有界元法和有限元法 205
8.1 边界值问题：位势方程 205
8.1.1 偏微分算子及椭圆边界值问题 205
8.1.2 边界条件 205
8.1.3 诺依曼边界值问题 208
8.2 方程空间 208
8.2.1 $C^k(\Omega), C^{k,\kappa}(\Omega)$ 空间 208
8.2.2 勒贝格 $L^p(\Omega)$ 空间 210
8.2.3 希尔伯特 (H) 空间 217
8.2.4 索伯列夫 $W^{k,p}(\Omega)$ 空间 221
8.3 变分法 243
8.3.1 算子方程 243
8.3.2 椭圆算子 247

8.3.3 算子与稳定性条件 249
8.3.4 含有限制条件的算子方程 251
8.3.5 混合算子方程 (鞍点变分问题) 253
8.3.6 强制算子方程 258
8.4 变分法求解边界值问题 259
8.4.1 位势方程基本介绍 260
8.4.2 狄利克雷边界值问题 1 261
8.4.3 狄利克雷边界值问题 2 264
8.4.4 诺依曼边界值问题 267
8.4.5 混合边界值问题 270
8.4.6 罗宾边界值问题 271
8.5 拉普拉斯算子的基本解 271
8.5.1 基本解 271
8.5.2 拉普拉斯算子 272
8.5.3 三维空间的基本解 273
8.5.4 二维空间的基本解 275
8.5.5 基本解总结 276
8.6 边界积分算子 277
8.6.1 牛顿位势 277
8.6.2 单层位势 282
8.6.3 伴随双层位势 284
8.6.4 双层位势 287
8.6.5 超奇异边界积分算子 291
8.6.6 边界积分算子之间的关系 297
8.6.7 单层位势算子的椭圆性和可逆性 299
8.6.8 超奇异边界积分算子的椭圆性 305
8.6.9 Steklov-Poincaré 算子 308
8.6.10 双层位势的收缩属性 309
8.6.11 边界积分算子的映射特性 312
8.7 边界积分方程 315
8.7.1 狄利克雷边界值问题 315
8.7.2 诺依曼边界值问题 318
8.7.3 混合边界值问题 321
8.7.4 罗宾边界值问题 323
8.8 近似方法 324
8.8.1 伽辽金——布博诺夫法 324
8.8.2 线性形近似 327
8.8.3 算子的近似 328
8.8.4 伽辽金——佩特洛夫法 329
8.8.5 混合边界值问题 331
8.8.6 强制算子 336
8.9 有限元 339
8.9.1 参考元 339
8.9.2 形式方程 347

8.9.3 检测空间 351

第二部分 模型估计 **355**

第三部分 待分类 **357**

第九章 最大似然估计 **359**

9.1 简介 359

9.2 近似的测度 359

9.2.1 估计的偏误：标量参数 360

9.2.2 估计的偏误：向量参数 360

9.3 费雪信息矩阵和克拉美罗下界 361

9.3.1 标量参数 361

9.3.2 向量参数 362

9.4 最大似然估计 362

9.4.1 最大似然估计：to be or not to be? 362

9.5 伯努利实验：最大似然估计法 363

9.6 伯努利实验：最大似然估计的 Python 程序实现 364

9.6.1 模型构建 364

9.6.2 似然方程 366

9.6.3 最大似然估计：解析法 366

9.6.4 最大似然估计：数值法 367

9.6.5 无约束参数 $g(\varphi)$ 构成的模型系统 368

第十章 工作论文 **375**

10.1 问题的提出 375

10.1.1 研究目标 375

10.1.2 模型设定 375

10.2 境内金融中介部门 376

第四部分 附录 **379**

Appendices

附录 第一章 Post-Keynesian Economics 的一个小综述 **383**

A.1 Introduction 383

A.2 PK 的主要方法论 383

A.3 PK 下的市场结构与定价 383

A.3.1 PK 下的市场结构 383

A.3.2 PK 下的定价策略 383

A.4 PK 下的宏观经济 383

A.5 PK 视野下的经济政策 384

附录 第二章 Keynesian, New Keynesian and New Classical Economics **385**
B.1 Keynesian 的四个核心 385
B.1.1 失业与有效工资理论 385
B.1.2 价格变化与经济波动 386
B.1.3 储蓄与投资，信贷配给 387
B.1.4 供应与技术进步 387
B.2 New Keynesian 经济学的四个核心 387
B.2.1 有效工资 387
B.2.2 资产配给 387
B.2.3 信贷配给 387
B.2.4 货币政策 388
B.2.5 小结 388
B.3 凯恩斯的不足 388
B.3.1 债券和股票的区别 389
B.3.2 需求和供应 389
B.3.3 投资的决定因素 389
B.3.4 货币政策 390
B.4 小结 390

第一章　前言

还没想好要写什么...

第一部分

模型求解

第二章　DSGE 模型求解方法简论

2.1　简论的简论

过去的三十多年来，宏观经济学研究经历了一场飞速变革。这场变革始于Kydland and Prescott (1982) 利用 RBC 模型研究美国经济。这种研究方法逐渐成为宏观经济学的标准范式之一 (An and Schorfheide, 2007; Fernández-Villaverde, 2010)。

随后 RBC 模型逐渐扩展到新凯恩斯主义模型。经典教材如Galí (2008); Woodford (2011)。然而新凯恩斯主义模型也远非完美无缺，随着新的问题逐渐被发现，学术界对模型做了进一步的修正和扩展，使模型对现实的拟合程度越来越高，如

- 根据基准黏性价格假定所生成的一些重要经济变量的时间序列，与实际观察到的经济现实相比，出入较大——真实经济世界中的通货膨胀，产出，实际工资等时间序列数据都具有很高的持续性——为了使得模型与现实相贴近，模型设定中就需要引入看起来异常高的名义粘性。

- 因此产生了一系列对基准新凯恩斯模型的扩展，它们的基本思路是，加入Calvo (1983) 定价机制，从而可以在控制名义粘性不至于过高的情况下，有效提高模型生成的通货膨胀时间序列数据的持续性 (Rabanal and Rubio-Ramírez, 2005)。

这样新凯恩斯主义分析框架得到了中央银行等政策制定者的青睐。不同中央银行开发了一系列自己的 DSGE 模型，如

- 欧洲中央银行 (ECB) 的 NAWM (New Ara-Wide Model)。

  在Fagan et al. (2005) 的 Area-Wide Model (ARM) 和Smets and Wouters (2003) 的基础上，Christoffel et al. (2008) 开发了欧元区的 NAWM，一个基于微观经济基础的开放经济模型，广泛应用于欧洲中央银行研究人员对经济系统的预测。

- 加拿大银行 (Bank of Canada) 的 ToTEM (Term-of-Trade Economic Model)，见Murchison and Rennison (2006)。

- 英格兰银行 (Bank of England) 的 BEQM (Bank of England Quarterly Model)，见Harrison et al. (2005)。

- 挪威银行 (Norges Bank) 的 NEMO (Norwegian Economy Model)，见Brubakk et al. (2006)。

- 智利中央银行 (Central Bank of Chile) 的 MAS (Model of Analysis and Simulation)，见Medina and Soto (2007)。

- 瑞典中央银行 (Sveriges Riksbank) 的 RAMSES (Riksbanks Aggregated Model for Studies of the Economy in Sweden)，见Adolfson et al. (2007, 2008a)。

- 美联储的 SIGMA。

  在Obstfeld and Rogoff (1995) 的开放经济体模型基础上，Erceg et al. (2006) 简历的多国家开放经济模型。

- 国际货币基金组织 (IMF) 的 GEM (Global Economic Model)，见Bayoumi et al. (2004)。

上述中央银行开发的经济预测模型，多得以于 DSGE 模型的长处，具有如下特征：

1. 动态性。关注变量岁时间的变化路径，而不是某单独时间点上的情况。
2. 整体性。致力于解释、预测整体经济运行，而非仅仅是局部市场。

   但这并不意味着 DSGE 模型可以重现整个经济系统的没一个重要部分，尤其是其中一些部门常常需要做必要的简化处理，如财政政策的制定部门 (政府)、金融市场等。

3. 重视部门均衡。根据经济学理论，在不同市场中，重视市场调节机制作用下的供应需求平衡。
4. 模型中引入随机干扰。

从广义的宏观经济学模型角度来看，DSGE 模型可以分为两大类：RBC 模型和新凯恩斯主义模型。RBC 模型致力于研究在灵活价格环境下的经济周期波动，常表现为两部分的组合：一部分是随机内生经济增长模型作为内核，另一部分是外部真实技术冲击。Kydland and Prescott (1982) 对 RBC 模型 (进而 DSGE 模型) 研究做出了开创性的贡献，进一步评述可见Cooley (1995)。

RBC 关注实际经济变量，对货币问题未作深入讨论。而在现实经济世界中，货币的重要性越发得到重视。于是就有了新凯恩斯主义模型，大致说来，它对 RBC 模型做了两方面的扩展，一是一些内部机制的调整，二是外部冲击来源的选择。就前者而言，包括比如

1. 将货币和货币发行机构 (中央银行) 引入到分析框架中来，
2. 不再持有完全竞争假定，在一些部门的行为分析中假定不完全竞争，如：
   (a) 产品/服务市场，和/或劳动力市场是不完全竞争的，
   (b) 允许私有部门的消费和投资决策在存在刚性的情况下展开：通过引入黏性价格、黏性工资的设定，将名义刚性引入模型设定中来。这意味着货币政策会对实际经济变量产生影响。

如何构建一个 DSGE 模型，进而如何将其应用到宏观经济分析中去，仍需要我们在理论模型和计算求解两个维度上展开深入研究。就后者而言，很显然理论模型设定的细节不同，需要不同的求解方法。但总的说来，求解流程及方法大同小异：识别假设条件 → 推出 (一阶) 均衡条件 → 构建结构方程组 → 形成随机差分方程系统 (通常是非线性的) → **对非线性系统做近似线性化 → 求得近似解 → (计算冲击响应方程或二阶矩以) 检验近似解的有效度**。

本章主要对后半部分做简要介绍，大致包括 DSGE 模型的求解方法、参数估计、模型检验等。已有一系列文献对此作了综述，随着 DSGE 模型类型的不同，这些文献各有侧重，分别关注不同的求解方法，如

- Canova (2009, 2011); Balke et al. (2012) 主要介绍常见的宏观计量经济学方法。
- An and Schorfheide (2007); Fernández-Villaverde (2010); Del Negro and Schorfheide (2011); Herbst and Schorfheide (2015) 主要介绍贝叶斯估计法在 DSGE 模型中的应用。
- Galí (2008); Woodford (2011) 主要关注 DSGE 模型的一种：新凯恩斯主义模型。
- Tovar Mora (2009) 探讨 DSGE 模型在中央银行中的应用。

2.2 pros and cons

2.2.1 cons

目前学界基本达成共识，DSGE 模型可作为宏观经济学研究的标准框架之一。但仍有局限，试从以下几个角度做简要介绍，分别为前提假定，求解方法，解的复杂性，以及利用解做决策的过程。

前提假定。一系列假定条件，如完美市场范式、市场效率、个体如家庭的理性期望等，引发争议。尤其是，过分简化、因而脱离实际的前提设定使得 DSGE 模型未能有效预测 2008 年的金融危机，使其广受质疑 (Buiter, 2009)。作为回应，近年来的 DSGE 模型逐渐引入一些与现实更加贴近的假定，如个体行为的经验法则 (rule-of-thumb)，金融市场的摩擦，行为个体的异质性特征等。

求解方法。对本质上是高度非线性的经济系统而言，不恰当的近似线性化处理会导致损失许多重要信息，这些信息原本是宏观经济学研究的重要对象。Buiter (2009) 甚至称之为 “对宏观经济学模型的阉割”。

解的复杂性，以及向公众解释政策含义的难度。线性近似的最优条件和约束条件构成复杂而冗长的方程系统，在这系统中，第一，前向期望项变得难于辨认和解读，更难于向公众说明这一方程系统背后的经济学含义。第二，难于识别外生冲击如财政、货币政策的决策，是通过怎样的传导机制进入经济系统，并产生何种影响。第三，大多数情况下，难以直接求得系统的解析解，就需要借助计算机来测算近似的数值解，而采用何种数值算法更适合这一系统，数值近似解的精确程度如何，就成了又一个难题。这些都决定了模型系统及系统的解难以为公众们所理解，也因此难以在不同政策制定者之间形成共识。

门槛高。综上，DSGE 模型的构建、求解、预测等一系列工作的展开，均要求研究者及政策制定者受过良好的宏观经济学训练，并有相当程度的建模能力、统计学知识和编程水平。

2.2.2 pros

尽管如此，采用 DSGE 模型作为宏观经济学研究分析框架的优势也很突出。

微观基础。传统宏观模型往往不对个体行为做深入设定。与之相比，DSGE 模型在微观形为基础方面做了深入探索：行为个体基于理性预期的假定采取最优行动，决定要素价格和资源配置，进而影响公共部门的目标和约束条件。这样所生成的一系列局部最优条件，如家庭的劳动力供应决策、消费决策，企业的劳动力需求决策、产品定价决策等，为政策制定者提供重要参考依据。“优化” 行为也意味着，个人和企业基于他们对未来期可能出现状况的预测，展开当前期的行为；从这种意义上来说，这种 “理性预期” 的行为模式便不同于 “经验法则”。

稳健性。基于微观基础的 DSGE 模型使得缩减形式 (reduced form) 的参数与更深层次的结构形式 (structural form) 的参数之间，形成更紧密关联，使得模型参数较少可能随着政策的变动而发生变化——从此意义上说，DSGE 模型有助于更稳健地回应卢卡斯批判 (Lucas, 1976)[1]。

模型研究与政策制定的契合度。通过近似分散化经济体 (decentralized economy) 中典型个体的效用函数，DSGE 模型产生与中心化经济 (centralized economy) 模型中的福利定律相一致的结果。从而，DSGE 模型可以提供一套连贯的用于政策讨论与分析的工具，评估不同决策的效果，从而选取更好的决策付诸实施。在这一过程中，政策分析的展开与模型的设定条件是紧密相关的。

工具包。DSGE 模型的吸引人之处不仅在于理论分析框架，更在于为宏观经济学的经验研究提供了一套可量化的政策分析和预测工具。伴随着理论建模的进展，也涌现出了很多新的经验工具和算法，致力于让模型分析的结果更贴近实际观察到的数据。随着二者相符程度的不断提升，DSGE 模型的可靠性越来越得到政策制定机构如中央银行的认可 (An and Schorfheide, 2007)，其预测效

[1] 递归宏观经济学模型中的稳健控制论综述，可见Hansen and Sargent (2004)。简约式和结构式计量经济学方法论的争论，可见Jarrow and Protter (2004)；一个更全面的综述可见Angrist and Pischke (2008)。

力不弱于 VAR 模型Edge et al. (2010)。此外，模型估计方面的研究也有快速进展 (Schorfheide, 2011; Fernández-Villaverde et al., 2016)。

总之，DSGE 模型提供了一整套集成的政策分析框架。但仍需要指出两点。一，简化。经济模型始终是对现实世界的描述，这种描述是简化的，并不追求事无巨细的完整重现。那么研究人员基于 DSGE 模型给出的政策性建议都必须立足于常识，同时牢记简化了的模型并不覆盖纷繁复杂现实世界的每一个细节。二，主观性。一方面由于 DSGE 模型的复杂性，另一方面中央银行建立新的或改进已有的 DSGE 模型，目的往往是为自己的货币政策的有效性背书。在此过程中，主观性总是不可避免的。那么需要强调的是，基于 DSGE 模型的经济研究 “不能代替专家意见”，政策制定者还需要考虑 “坊间消息和模型之外的信息”(Bernanke, 2007)。

2.3 工作流程

设定前提假定条件 → 求得系统解 → 将模型生成的方针数据与实际观测到的数据向比较 → 政策性建议，在分析无限时间周期的，由多个典型行为个体组成的 DSGE 模型时，常见的工作流程见下文，或如图2.1所示。

图 2.1: DSGE 模型的工作流程

1. 构建经济模型 (假设条件)。

2. 求解一阶均衡条件。一阶均衡条件与预算约束条件等一道构成非线性随机差分方程系统。
3. 由于该系统常常难以求得解析解，需要围绕这一个给定的点做数值近似。这一给定点，常常设定为模型的非随机稳定状态。
4. 非随机稳定状态可由如下方法之一求得：
   (a) (对数) 线性化近似，围绕稳态，得到一个状态——空间形式表述的线性差分系统，进而借助一些常见算法求得系统的近似数值解；
   (b) 围绕非随机稳定状态，做二阶或更高阶的近似。
5. 参数校准和/或估计。
6. 针对外生随机冲击，计算方差、做方差分解，求得有关变量的脉冲——响应方程。
7. 比较模型生成数据和史记观测到的数据以评价模型的解释力。
8. 政策参考。

下一节首先讨论模型设定。

2.4 模型设定

模型的前提假定应设定为与所研究的问题密切相关，并且必须指明模型中所讨论的问题是从定性还是定量的角度展开的。围绕所研究的问题，模型应该提供有所助益的回答，如：哪些冲击对经济体产生何种影响，它们的传导机制是怎样的，如何制定相关政策以应对这些冲击?

在一般均衡模型中，经济体中的每一个部门都是单独设定的，典型的部门如私人部门和公共部门：

1. 私人部门包括如消费 (家庭) 部门和企业部门等。通常来说，消费者被简化为一个典型家庭，基于自身偏好，做关于消费和休闲的一生最优决策 (Christiano et al., 2010)。企业在一定的生产条件限制下追求利润最大化。在为私人部门建模时，必须假定预期的形式，如理性预期或者经验法则等。
2. 公共部门包括如财政政策和货币政策等。通常来说，政府是财政政策的制定者，大多数情况下财政政策的效果受到李嘉图等价 (Ricardian equilvalence) 的影响。中央银行设定货币政策。在为公共部门建模时，必须假定政策的执行方式，比如公共部门的政策制定是否依据某准则，是否受某种预算约束条件的限制，行为是否遵循某种利润最大化原则等。
3. 在经济模型中，个体行动也要受到市场机制和制度性要素的考量，因此也徐作出设定，比如哪些市场是充分/非充分竞争的，决策者的政策制定是否需要遵循一些经验法则等。
4. 外生随机冲击的识别和建模。

上述模型设定都必须转换成适当的数学表达式，一方面使其能被纳入到模型分析框架中去，另一方面更重要的是，数学表达式的设定需要使模型有解。如，关于偏好和冲击的凹或凸的方程式，非庞兹骗局的约束条件等。

在模型设定后，下一步工作是要针对经济体中的家庭、企业、公共部门等，分别求得一阶最优条件和预算约束条件方程。这些方程常常包括当期和/或前向 (后向) 元素，又称非线性理性期望系统。通常来说，我们需要借助一定的数值近似方法，将非线性系统做近似线性化，进而求解，这一求解方法的讨论见下节。

2.5 求解方法

对于动态模型如新古典主义经济增长模型，最常见的解法是动态规划 (Dynamic Programming)，又称价值方程迭代，简要介绍见第四章。这种方法的优势在于算法可靠度高，解的收敛特性好。但不足也较明显，如计算速度慢，存在维数灾难 (curse of dimensionality, Bellman (1957)) 等。

rse of dimen-
nality，可由映
法得到一定程
上的缓解，做
个 reference。

动态规划法的不足导致新算法的出现，典型如扰动法 (perturbation methods)、投影法 (projection methods) 等。这些新方法在保持较好收敛特性的基础上，提供了更快的求解速度。扰动法在宏观经济学研究中的广泛应用，综述可见如Stokey et al. (1989); Ljungqvist and Sargent (2004)。投影法的简要介绍见Aruoba et al. (2006)。

本章内容重在对求解方法做简要介绍。大致说来，求解方法分为两步。第一步，近似线性化，将由均衡条件和约束条件构成的非线性方程系统转化为线性方程系统。不同方法背后核心思路基本一致，即将变量围绕其非随机稳定状态附近做线性近似。第二步常常被视为求解 DSGE 模型的核心，即根据近似线性系统的实际特征，采用相应的数值算法求解，作为原非线性系统的近似解[2]。

2.5.1 近似线性展开

在King et al. (1988a,b, 2002) 的开创性工作之后，(对数) 线性近似法逐渐成为主要标准。以将变量沿着其非随机稳态做一节泰勒级数展开为例，假定 $y = f(x)$，其中 x, y 分别是动态模型中的前定和非前定变量 (predetermined and non-predetermined variables)，f 是个平滑的非线性函数[3]。

定义 x_t 的确定性 (非随机) 稳定状态为 $\bar{x}$，我们有 $\bar{y} = \bar{x}$。

对 f 围绕 $\bar{x}$ 做泰勒级数展开

$$y \approx \sum_{n=0}^{\infty} \frac{f^{(n)}(\bar{x})}{n!} (x - \bar{x})^n , \tag{2.1}$$

其中 $f^{(n)}$ 表示 f 对 x 的第 n 次求导。

(1 阶) 线性化近似

根据(2.1)，1 阶泰勒级数展开

$$\begin{aligned} & y \approx f(\bar{x}) + f_x(\bar{x})(x - \bar{x}), \\ & \Rightarrow (y - \bar{y}) \approx f_x(\bar{x})(x - \bar{x}), \\ & \Rightarrow \frac{y - \bar{y}}{\bar{y}} \approx \frac{f_x(\bar{x})\bar{x}}{\bar{y}} \frac{x - \bar{x}}{\bar{x}}, \\ & \Rightarrow \tilde{y} \approx \underbrace{\frac{f_x(\bar{x})\bar{x}}{\bar{y}}}_{\text{常系数}} \tilde{x}, \end{aligned} \tag{2.2}$$

根据定义，$\tilde{y}_t$ 和 $\tilde{x}_t$ 分别表示 y_t 和 x_t 相对于其确定性稳态 $\bar{y}$ 和 $\bar{x}$ 的偏离百分比。

[2]扰动法和投影法，从本质上来说都是局域算法。全局算法如车比雪夫多项式 (Chebyshev-polynomial)、有限元 (finite-elements) 等的综述可见Judd (1998)。

[3]这里我们采用Klein (2000) 的定义：x_t 是前定变量，当且仅当它有外生的 1 期前向预测偏误，即 $x_{t+1} - E_t X_{t+1} = C\varepsilon_{t+1}$ 是外生的。

根据这一定义，一个前定变量由它的滞后项和当期外生冲击所共同决定，如外生冲击变量。另一方面，不属于此情况的变量成为非前定变量，常常是前向变量。

(1 阶) 对数线性化近似

对模型做变型，$y = \exp(\ln y) = f(\exp(\ln x))$。
两侧取对数，$\ln y = \ln f(\exp(\ln x))$。
围绕 $\ln \bar{x}$ 作 1 阶泰勒级数展开

$$\begin{aligned}
&\ln y \approx \ln f(\exp(\ln \bar{x})) + f_x(\bar{x}) \frac{d \exp(\ln x)}{d\bar{x}} (\ln x - \ln \bar{x}), \\
&\Rightarrow \ln y \approx \ln \bar{y} + f(\exp(\ln \bar{x})) (\ln x - \ln \bar{x}), \\
&\Rightarrow \ln \left(\frac{y}{\bar{y}}\right) \approx \underbrace{\frac{f_x(\bar{x})\bar{x}}{\bar{y}}}_{\text{常系数}} \ln \left(\frac{x}{\bar{x}}\right),
\end{aligned} \tag{2.3}$$

其中 $\ln\left(\frac{y}{\bar{y}}\right)$ 和 $\ln\left(\frac{x}{\bar{x}}\right)$ 分别表示对数形式的偏离百分比。

一阶泰勒级数展开后的线性化近似和对数线性化近似，本质上相似。为了说明这一点，见下式

$$\tilde{x}_t = \frac{x_t - \bar{x}}{\bar{x}} = \frac{x_t}{\bar{x}} - 1 \approx \ln\left(\frac{x_t}{\bar{x}}\right) = \ln x_t - \ln \bar{x}.$$

此外Uhlig (1999) 提出另一种类似对数线性化的方法，从而在特定情况下无需求导：对于常数 a 和接近 0 的 $\tilde{x}_t$ 和 $\tilde{y}_t$，

$$\begin{aligned}
&\exp(x + ay) \approx 1 + x + ay, \\
&\Rightarrow xy \approx 0, \\
&\Rightarrow E_t\left(a \exp(x_{t+1})\right) \approx E_t\left(a x_{t+1}\right).
\end{aligned}$$

对数线性化方法的本质在于将变量 x_t 用 $x \exp(\tilde{x}_t$ 来代替，此时 $\tilde{x}_t = \ln(x_t/\bar{x})$ 表示它相对于稳态的对数偏差。利用此方法，可以将非线性方程近似为围绕稳态的线性方程。

2 阶近似

根据(2.1)，以 2 阶展开为例

$$y \approx \underbrace{f(\bar{x})}_{\text{0 阶展开}} + \underbrace{f_x(\bar{x})(x - \bar{x})}_{\text{1 阶展开}} + \underbrace{\frac{1}{2} f_{xx}(\bar{x})\,(x - \bar{x})^2}_{\text{2 阶展开}} + \underbrace{O^3}_{\text{误差项}}, \tag{2.4}$$

其中线性近似的误差项 O^3 是由忽略了 3 阶及以上阶数近似所造成的。

(2.4)为代表的线性近似法具有如下特征：

1. 在一阶泰勒级数展开中，y 的条件期望值等于稳态下的 $\bar{y}$，这意味着不确定性 (以方差的二阶或更高阶矩等形式表现) 在线性近似过程中不起作用。
2. 在高阶泰勒级数展开中，y 的条件期望值与方程 f 的曲率和变量 x 的方差有关。
3. 如果研究的目的是求得冲击响应方程以及变量的二阶距，那么我们只分析模型的一阶属性即可，只需要对非线性系统做一阶泰勒级数展开。如，$var(y)$ 可以通过(2.4)算出：$var(y) = E(y = Ey)^2 = f_x^2 var(x)$。

在什么情况下做一阶展开，什么情况下需要做更高阶展开?

1. 在很多情况下，对非线性系统做一阶线性近似即可。尽管有研究发现更高阶的近似会提供更精确的近似解，但精读提升的幅度有限。

2. 然而的确存在一些情况，仅仅用一阶近似处理 DSGE 模型是不够的，尤其是当研究目标涉及到分析一些政策的福利效果时——这些福利政策往往不会对模型产生一阶影响。

入一个 refer-
ce。

我们将在随后章节中进一步展开探讨这个问题。

对非线性方程系统做线性近似，生成以状态——空间形式表现的变型系统。这样，一方面可以更方便的展开后续经验研究，如模型估计，卡曼滤波，经济系统预测等。更为重要的是另一方面，搭配一些数值求解方法，可以求得系统的近似解。随后我们讨论如何在理性预期的情境下，运用数值算法求解变型系统。

2.6　求解线性随机差分方程系统

如前文所述，由各部门均衡条件和预算约束条件所组成的非线性系统，经由一定的 (对数) 线性化处理后变型为新的系统，又称线性随机方程差分系统，

新系统有一组含有理性期望条件的线性随机差分方程构成，可表述为如下状态——空间形式，

$$A_0 E_t Y_{t+1} = A_1 Y_t + B_0 \varepsilon_1, \tag{2.5}$$

其中 A_0, A_1, B_0 是线性化系统的系数矩阵，Y_t, ε_t 分别是内生变量和外生变量构成的向量。

在构建 DSGE 模型时，外生变量 ε_t 常取如下两个假定之一。假定 ε_t 是 iid 冲击向量，$E(\varepsilon_t) = 0, E(\varepsilon_t \varepsilon_t^\top) = \Sigma, E(\varepsilon_t \varepsilon_s^\top) = \Sigma \forall s \neq t$，或者假定 ε_t 是与 iid 外生冲击有关的 $AR(1)$ 过程。

这样一个线性化方程构成的动态系统，描述了模型中变量的运动路径。通常来说系统中含有前向及后向要素。随着模型参数设定的不同，系统解存在三种情况：

rl Whelan 对
rward-looking
backward-
oking 的描述

- 无解。系统没有稳定的理性预期解。
- 不确定 (indeterminacy)。系统存在不止一个解的多重均衡。
- 确定的。有且只有一个解，又称稳定解。

如果前两种情况出现，意味着模型参数设定不当，需要重新调整，以使得系统存在稳定解。

已经有大量文献探讨理性期望系统(2.5)的求解方法。系统的解构成一个回应机制 (feedback rule)，将当期内生变量与模型内的状态变量联系起来。这方面的重要工作框架由Blanchard and Kahn (1980) 所奠定，随后经由许多人的努力，得到进一步深化和扩展，见第五章。

如果 DSGE 模型致力于考察政策的福利效果，如期望效用等，则仅仅用一阶近似线性转换可能不够精确。此时需要对效用函数等其他函数做二阶近似，以确保二次项所含有的重要经济学信息不被忽略掉 (Kim and Kim, 2003; Schmitt-Grohé and Uribe, 2004)，见第六章。

第三章　动态规划

简要介绍动态规划 (dynamic programming)。更详细的介绍，见Adda and Cooper (2003)。Python 中的程序实现见Stachurski (2008)。

3.1　包络定理

包络定理 (Envelope Theorem)研究当模型中的参数发生变化时，模型中某一变量的最大值 (或最小值) 如何随着发生变化。

Theorem 3.1 (包络定理)**.** *假定我们有*

$$v(a) = \max_{\{x\}} f(x, a), \tag{3.1}$$

那么下式成立

$$\frac{dv(a)}{da} = \frac{\partial f(x, a)}{\partial a}\Big|_{x=x^*(a)}, \tag{3.2}$$

其中

$$x^*(a) = \arg\max_{\{x\}} f(x, a). \tag{3.3}$$

证明. 设

$$v(a) \equiv f\left[x^*(a), a\right],$$

两侧同时对 a 求 (偏) 导

$$\frac{dv(a)}{da} = \frac{\partial f\left[x^*(a), a\right]}{\partial x}\frac{\partial x^*(a)}{\partial a} + \frac{\partial f\left[x^*(a), a\right]}{\partial a}. \tag{3.4}$$

上式中 x^a 是 $f(x, a)$ 取极大值时的系数，满足

$$\frac{\partial f\left[x^a, a\right]}{\partial x} = 0,$$

代回(3.4)有

$$\frac{dv(a)}{da} = \frac{\partial f\left[x^*(a), a\right]}{\partial a}, \tag{3.5}$$

等价于(3.2). □

或者也可以从有约束条件的优化问题中推得包络定理

Theorem 3.2 (包络定理 (有约束条件的优化问题))**.** *设*

$$m(a) = \max_{\{x\}} f(x, a),$$

$$s.t. \begin{cases} g(x, a) = 0, \\ x \geq 0. \end{cases}$$

使 $\mathcal{L}(x,a,\lambda)$ 为相应的拉格朗日方程，其中 $x^*(a),\lambda^*(a)$ 为库恩塔克条件 *(Kuhn-Tucker condition)* 的最优解。则我们有

$$\frac{dm(a)}{da}=\frac{\mathcal{L}(a)}{\partial a}\Big|_{x^*(a),\lambda^*(a)}. \tag{3.6}$$

3.2 例: 吃蛋糕 (直接求解法)

小明有一块蛋糕，大小是 W_1，可以在 $t=1,2,\ldots,T$ 期内吃完。问小明怎么吃可以实现效用最大化?

小明的最优决策可以表现为贴现效用的求和加总

$$\sum_{t=1}^{T}\beta^{t-1}u(c_t),$$

其中 c_t 表示 t 期消费即吃掉蛋糕的大小。$u(c_t)$ 表示 t 期效用，设方程满足可导、严格单调、严格凹 (concave)，即稻田条件 (Inada condition)(Inada, 1963)

$$\lim_{c\to 0}u'(c)\to\infty,\quad \lim_{c\to\infty}u'(c)\to 0.$$

引入稻田条件的作用是确保小明每个时段 t 都至少会吃一小口蛋糕，即从数学角度上来讲，不存在角点解 (corner solution)。蛋糕的大小随时间发生变化，满足

$$W_{t+1}=W_t-c_t,\quad t=1,2,\ldots,T. \tag{3.7}$$

若是按照直接求解法的思路，小明吃蛋糕问题可以表示为

$$\max_{\{c_t\}_{t=1}^{T},\{W_t\}_{t=1}^{T+1}}\sum_{t=1}^{T}\beta^{t-1}u\left(c_t\right),$$

给定约束条件

$$\begin{aligned}\sum_{t=1}^{T}c_t+\sum_{t=1}^{T}W_{t+1}&=\sum_{t=1}^{T}W_t,\\ \Rightarrow\sum_{t=1}^{T}c_t+W_{T+1}&=W_t,\end{aligned}$$

即小明吃掉的所有蛋糕，和最终剩余的蛋糕数量 W_{T+1} 之和，等于蛋糕的初始大小 W_1。

鉴于此，可以将问题改写为

$$\max_{\{c_t\}_{t=1}^{T},\{W_t\}_{t=2}^{T+1}}\sum_{t=1}^{T}\beta^{t-1}u\left(c_t\right), \tag{3.8}$$

约束条件为

$$\begin{cases}\sum_{t=1}^{T}c_t+W_{T+1}=W_1,\\ c_t\geq 0\quad\forall\, t=1,\ldots,T,\\ W_{T+1}\geq 0.\end{cases} \tag{3.9}$$

在随后的分析中我们不再讨论 $c_t\geq 0$ 的约束条件，因为假定稻田条件 $\lim_{c\to 0}u'(c)\to\infty$ 成立，就已经满足该项。

构建拉格朗日方程方程

$$\mathcal{L}=\sum_{t=1}^{T}\beta^{t-1}u\left(c_t\right)+\lambda_t\left[W_1-\sum_{t=1}^{T}c_t-W_{T+1}\right]+\phi_tW_{T+1}.$$

两个一阶条件 FOCs 给出最大化问题的解

- $$\frac{\partial \mathcal{L}}{\partial c_t} = 0 \Rightarrow \beta^{t-1} u'(c_t) = \lambda_t, \quad t = 1, \ldots, T, \tag{3.10}$$

  其中拉格朗日乘子 λ_t 对应预算约束条件，(3.9)第一行，即

  $$\lambda_t = \beta^t u'(c_{t+1}).$$

  将两期最优消费决策联系起来，可得

  $$u'(c_t) = \beta u'(c_{t+1}), \tag{3.11}$$

  这称为跨 (1) 期消费的欧拉等式 (Euler equation)。

- $$\frac{\partial \mathcal{L}}{\partial W_{T+1}} = 0 \Rightarrow \lambda_t = \phi_t, \tag{3.12}$$

  其中拉格朗日乘子 ϕ_t 对应横截条件，(3.9)第三行。已知根据稻田条件的，$\lambda_t > 0$ 严格成立，则 $\phi_t > 0 \forall t$。这意味着若要满足库恩塔克 (Kuhn-Tucker condition)

  $$\phi_t W_{T+1} = 0, \tag{3.13}$$

  只能使得 $W_{T+1} \equiv 0$。

进一步解读欧拉等式(3.11)。在均衡状态下，当期 t 减少 1 单位消费的边际效用下降，应当等于 t 期额外 1 单位储蓄的增加，在 $t+1$ 期带来边际效用提升的贴现值。将跨 1 期扩展到跨多期的情况，任意两期 t, t' 之间 $t, t' = 1, \ldots, T$, $t' > t$ 的跨期消费转移都不会导致总福利水平的变化。这样一来，(已经处于最优行为决策下的) 小明，全部福利水平最大便只取决于初始蛋糕的大小 W_1，对应福利 $V_T(W_1)$，又称价值方程 (value function):

$$V_T(W_1) = \max_{\{c_t\}_{t=1}^T} \sum_{t=1}^T \beta^{t-1} u(c_t) + \lambda_t \left[W_1 - \sum_{t=1}^T c_t - W_{T+1} \right] + \phi_t W_{T+1} \tag{3.14}$$

那么为了回答初始蛋糕大小 W_t 对小明总福利水平 $V_T(W_1)$ 的影响，可以求导数

$$\frac{dV_T(W_1)}{dW_1} = V_T'(W_1) = \lambda_t.$$

3.3 吃蛋糕 (动态规划)

3.3.1 有限时间段的动态规划问题

我们先从容易理解的 $T < \infty$ 入手，熟悉动态规划问题的基本原理。然后将问题扩展到 $T \to \infty$ 的情况，见第3.3.1节。

从 $t = 1$ 期来看，给定 W_1，吃蛋糕问题可以改写为如下形式

$$V_T(W_1) = \max_{\{c_t\}_{t=1}^T} u(c_1) + \beta V_{T-1}(W_2), \quad W_2 = W_1 - C_1, \tag{3.15}$$

$t = 1$ 期小明的决策是，消费多少蛋糕 c_1，这会传导到 $t = 2$ 期影响到 W_2；以此类推。那么，c_t 最优消费决策应当使 t 期效用的变化，与 $t+1$ 期效用变化的贴现相等，这称为贝尔曼等式 (Bellman equation)。

另一方面，对贝尔曼等式(3.15)求 FOC，可得欧拉等式

$$u'(c_1) = \beta V_{T-1}'(W_2). \tag{3.16}$$

现在设效用函数的显形式为 $u(c) = \ln c$，那么通过逆推法，可以进一步计算价值方程。
$T = 1$ 时，小明的最优决策是当期吃掉全部蛋糕，$c_1 = W_1$，因此我们有

$$V_1(W_1) = \ln W_1. \tag{3.17}$$

$T = 2$ 时，小明的最优决策根据(3.15)表示为

$$V_2(W_1) = \max_{W_2} \underbrace{\ln(W_1 - W_2)}_{u(c_1)} + \beta \underbrace{\ln W_2}_{u(c_2)}, \tag{3.18}$$

约束条件

$$c_1 + W_2 = c_1 + c_2 = W_1, \tag{3.19}$$

欧拉等式

$$u'(c_1) = \beta V_2'(W_1),$$

FOC

$$\frac{1}{c_1} = \beta \frac{1}{c_2}, \tag{3.20}$$

联立(3.19)和(3.20)可得最优消费决策 $\{c_t\}_{t=1}^2$

$$c_1 = \frac{1}{1+\beta} W_1, \quad c_2 = \frac{\beta}{1+\beta} W_1, \tag{3.21}$$

价值方程因此为

$$V_2(W_1) = \ln\left(\frac{1}{1+\beta} W_1\right) + \beta \ln\left(\frac{\beta}{1+\beta} W_1\right), \tag{3.22}$$

有时我们将其进一步简化为

$$\begin{aligned} &V_2(W_1) = A_2 + B_2 \ln W_2, \\ &A_2 = \ln\left(\frac{1}{1+\beta}\right) + \beta \ln\left(\frac{\beta}{1+\beta}\right), \\ &B_2 = 1 + \beta. \end{aligned}$$

$T = 3$ 时，小明的最优决策分三部分，依次为

1.

$$V_3(W_1) = \max_{W_2} \underbrace{\ln(W_1 - W_2)}_{u(c_1)} + \beta V_2(W_2), \tag{3.23}$$

FOC

$$V_3'(W_1) = 0 \Rightarrow \frac{1}{c_1} = \beta V_2'(W_2). \tag{3.24}$$

2.

$$V_2(W_2) = \max_{W_3} \underbrace{\ln(W_2 - W_3)}_{u(c_2)} + \beta V_1(W_3), \tag{3.25}$$

FOC

$$V_2'(W_2) = 0 \Rightarrow \frac{1}{c_2} = \beta V_1'(W_3). \tag{3.26}$$

包络条件

$$\frac{1}{c_2} = V_2'(W_2). \tag{3.27}$$

3.
$$V_1(W_3) = \ln W_3, \tag{3.28}$$
因为 $W_4 = 0$。

FOC
$$V_1'(W_3) = 0 \Rightarrow \frac{1}{c_3} = V_1'(W_3). \tag{3.29}$$

在此基础上我们有，我们可以用逆推的方法求得最优消费决策 $\{c_t\}_{t=1}^3$

1. 联立(3.26)和(3.29)可得
$$\frac{1}{c_2} = \beta\frac{1}{c_3},$$
2. 联立(3.24)和(3.27)可得
$$\frac{1}{c_1} = \beta\frac{1}{c_2},$$
3. 资源约束条件
$$c_1 + c_2 + c_3 = W_1.$$

因此我们有
$$\begin{aligned} c_1 &= \frac{1}{1+\beta+\beta^2}W_1, \\ c_2 &= \frac{\beta}{1+\beta+\beta^2}W_1, \\ c_3 &= \frac{\beta^2}{1+\beta+\beta^2}W_1, \end{aligned}$$

对应的价值方程为
$$V_3(W_1) = \ln c_1 + \beta(\ln c_2 + \beta \ln c_3). \tag{3.30}$$

3.3.2　无限时间段的动态规划问题

随着 $T \to \infty$，无法再使用上一节介绍过的逆推法，情况变得较为复杂。无限时间段的吃蛋糕问题，小明追求效用最大化
$$\max_{\{c_t\}_{t=1}^{\infty},\{W_t\}_{t=2}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} u(c_t),$$
$$\text{s.t. } W_{t+1} = W_t - c_t, \quad t = 1, 2, \ldots$$

我们将这个问题改写为动态规划的形式
$$V(W) = \max_{c\in[0,W]} u(c) + \beta V(W-c),$$
其中 $V(W)$ 为当期小明采取最优行为决策所可能得到的最大效用；$V(W-c)$ 为下期小明采取最优行为决策所可能得到的最大效用：显然，c 是控制变量，W 是状态变量。设 $W' := W - c$，那么动态规划问题改写为如下贝尔曼等式形式的泛函方程
$$V(W) = \max_{W'\in[0,W]} u(W-W') + \beta V(W'). \tag{3.31}$$

求解无限时段的动态规划问题(3.31)可分三部，核心思路是找到一个 $V(W)$ 使得对于所有 W 的值，泛函式均成立。出于简化模型的考虑，我们取消时间下角标——因为在无限时段下，这成为一个静态 (stationary) 或者无关时间 (time-invariant) 的问题。三部如下

1. (3.31)的 FOC
$$V'(W) = 0 \Rightarrow u'(c) = \beta V'(W'). \tag{3.32}$$
难点：在不清楚原方程形式 $V(W)$ 的情况下，如何计算 $V'(W')$。

2. 包络条件
$$V'(W) = u'(c) \Leftrightarrow V'(W') = u'(c'). \tag{3.33}$$
包络条件的前提是假定行为人处于最优决策条件下。

3. 联立 FOC(3.32)和包络条件(3.33)
$$u'(c) = \beta u'(c'), \tag{3.34}$$
即欧拉等式。由此我们可以定义一组策略方程
$$\begin{aligned} c &= \phi(W), \\ W' &= e(W) = W - \phi(W), \end{aligned} \tag{3.35}$$
使满足欧拉等式
$$u'(\phi(W)) = \beta u'(W - \phi(W)). \tag{3.36}$$

现在以 $u(c) = \ln c$ 为例做进一步求解[1]。首先，我们猜测 $V(W)$ 可能是以下线性形式
$$V(W) = A + B\ln W, \tag{3.37}$$
线性形的猜测是为了让计算过程简化，更多非线性形的讨论可见Hansen and Sargent (2004)。代回(3.31)可得
$$A + B\ln W = \max_{W' \in [0,W]} u(W - W') + \beta V(A + B\ln W'). \tag{3.38}$$
类似地，取 FOC 作为最优条件
$$\begin{aligned} \frac{1}{W - W'} &= \beta B \frac{1}{W'}, \\ \hookrightarrow W' &= \frac{\beta B}{1 + \beta B} W, \end{aligned}$$
代回(3.38)可得
$$\begin{aligned} A + B\ln W &= \ln\left(\frac{W}{1+\beta B}\right) + \beta\left[A + B\ln\left(\frac{\beta B}{1+\beta B}W\right)\right] \\ &= \ln W - \ln(1+\beta B) + \beta A + \beta B\ln\left(\frac{\beta B}{1+\beta B}\right) + \beta B\ln W \\ &= \underbrace{-[\ln(\beta B) + 1]\ln(1+\beta B) + \beta[A + B\ln(\beta B)]}_{A} + \underbrace{(1+\beta B)}_{B}\ln W, \end{aligned} \tag{3.39}$$
LHS 和 RHS 相对应，可得系数 A B 的值
$$\begin{aligned} A &= \frac{1}{(1-\beta)^2}\left\{\beta\ln\left(\frac{\beta B}{1-\beta}\right) - (1-\beta)\ln\left(\frac{1}{1-\beta}\right)\left[1 + \ln\left(\frac{\beta}{1-\beta}\right)\right]\right\}, \\ B &= \frac{1}{1-\beta}, \end{aligned}$$
这样我们最终得到了策略方程(3.35)的计算式
$$\begin{aligned} c &= \phi(W) = (1-\beta)W, \\ W' &= e(W) = \beta W, \end{aligned} \tag{3.40}$$

[1]通常来说，我们无法求得价值方程的显性形式，但 $u(c) = \ln c$ 可能是少数的例外之一。

3.4　随机问题

现在将不确定性引入动态规划问题的分析中去，如外生偏好冲击，效用函数变为 $\varepsilon_t u(c_t)$，其中 ε 是一个随机变量。在 t 期做最优消费决策的行为人知道当期的 ε_t，但不掌握对未来期 ε 的全部信息。一个最简单的设定是 $\varepsilon \in \{\varepsilon_l, \varepsilon_h\}$，两种冲击状态满足 $\varepsilon_h > \varepsilon_l > 0$。

偏好冲击表现为一阶马尔科夫过程，即当期冲击状态遇上一期的冲击水平有关。我们常用Tauchen (1986) 等方法将 AR(1) 过程做离散化处理，见第7.D节。用 p_{ij} 表示当期冲击状态 ε_i 在下一期变为 ε_j 的概率，据此可以构建完整的转移矩阵 P，其中的元素例如 $p_{lh} \equiv Pr\left(\varepsilon' = \varepsilon_h \middle| \varepsilon = \varepsilon_l\right)$:

表 3.1: 偏好冲击的转移矩阵

		$t+1$	
		ε_h	ε_l
t	ε_h	p_{hh}	p_{hl}
	ε_l	p_{lh}	p_{ll}

这样，随机吃蛋糕问题就变为

$$V(W,\varepsilon) = \max_{W'} \varepsilon u(W - W') + \beta E_{\varepsilon'|\varepsilon} V(W', \varepsilon'), \quad \varepsilon \in \{\varepsilon_l, \varepsilon_h\},$$

对应转移矩阵 P。

FOC 描述最优状态

$$\varepsilon u'(W - W') = \beta E_{\varepsilon'|\varepsilon} V(W', \varepsilon') \tag{3.41}$$

由包络定理可得

$$V_1(W,\varepsilon) = \varepsilon u'(W - W'),$$

进而我们有

$$V_1(W', \varepsilon') = E_{\varepsilon'|\varepsilon} \varepsilon' u'(W' - W''). \tag{3.42}$$

联立包络定理(3.42)和 FOC 条件(3.41)可得跨 (多) 期欧拉等式

$$\varepsilon u'(W - W') = \beta E_{\varepsilon'|\varepsilon} \varepsilon' u'(W' - W''). \tag{3.43}$$

如果我们设政策方程满足一定形式 (常常设为线性形)

$$W' = \varphi(W, \varepsilon),$$

代回(3.43)，欧拉等式变为

$$\varepsilon u'(W - \varphi(W,\varepsilon)) = \beta E_{\varepsilon'|\varepsilon} \varepsilon' u' \{\varphi(W,\varepsilon) - \varphi(\varphi(W,\varepsilon), \varepsilon')\}. \tag{3.44}$$

3.5　包络定理和欧拉等式的关系

上文以小明吃蛋糕问题为例，对动态规划做了一般性介绍。下面以常见的 NCGT(Neo-Classical Growth Model，如 Ramsey Model) 为例，介绍家庭部门的多期优化问题，它是动态经济学模型的基础组成部分之一。

t 期行为人的目标，追求 $t, t+1, \ldots, T$ 的效用最大化

$$\begin{aligned} &\max \sum_{n=0}^{T-t} \beta^n u\left(c_{t+n}\right), \\ &m_{t+1} = \left(m_t - c_t\right) R + y_{t+1}. \end{aligned} \tag{3.45}$$

根据动态规划的思路，将问题改写为贝尔曼等式形式

$$V_t\left(m_t\right) = \max_{\{c_t\}} u\left(c_t\right) + \beta V_{t+1}\left[\left(m_t - c_t\right) R + y_{t+1}\right]. \tag{3.46}$$

FOC 得到最优条件

$$0 = u'\left(c_t\right) + \left(\frac{dm_{t+1}}{dc_t}\right) \beta V_{t+1}^{'}\left(m_{t+1}\right),$$

其中 $\left(\frac{dm_{t+1}}{dc_t}\right) = -R$，$R$ 反映在(3.45)中。因此上式变为

$$u'\left(c_t\right) = \beta R V_{t+1}^{'}\left(m_{t+1}\right). \tag{3.47}$$

根据(3.45)，行为人基于已有的信息 m_t 做出当期消费决策 c_t，因此可以定义一个方程

$$c_t = c_t\left(m_t\right), \tag{3.48}$$

用于描述在给定 m_t 时，行为人可以达到效用最大化的最有消费决策。(3.48)代回(3.47)，写为

$$u'\left(c_t\left(m_t\right)\right) - \beta R V_{t+1}^{'}\left(\left[m_t - c_t\left(m_t\right)\right] R + y_{t+1}\right) = 0. \tag{3.49}$$

上式中的两个方程都是一阶形式；在原方程形式未知的情况下很难直接求解。我们因此采取近似算法。

定义一个新的方程

$$\underline{V}_t\left(m_t, c_t\right) = u(c_t) + \beta \underline{V}_{t+1}\left(\left(m_t - c_t\right) R + y_{t+1}\right), \tag{3.50}$$

同样地，对上式求偏导我们有

$$\underline{V}_t^c \equiv \frac{\partial \underline{V}_t}{\partial c_t} = u'(c_t) - \beta R \underline{V}_{t+1}^c\left(\left(m_t - c_t\right) R + y_{t+1}\right), \tag{3.51}$$

$$\underline{V}_t^m \equiv \frac{\partial \underline{V}_t}{\partial m_t} = \beta R \underline{V}_{t+1}^m\left(m_{t+1}\right). \tag{3.52}$$

根据定义式(3.50)，结合(3.46)可见

$$V_t\left(m_t\right) = \underline{V}_t\left(m_t, c_t\right), \tag{3.53}$$

引入链式法则 (Chain rule)(8.60)，上式变为

$$\begin{aligned} V_t^{'}\left(m_t\right) \equiv V_t^m\left(m_t\right) &\equiv \frac{dV_t\left(m_t\right)}{dm_t} \\ &= \underline{V}_t^m\left(m_t, c_t\left(m_t\right)\right) + \frac{\partial c_t\left(m_t\right)}{\partial m_t} \underline{V}_t^c\left(m_t, c_t\left(m_t\right)\right). \end{aligned} \tag{3.54}$$

理解(3.54)的一个关键是：行为人是处于最优决策状态的，这意味着他总是在衡量价值方程，以及价值方程相对于消费的偏导，使其满足最优条件(3.49)。[2]，这使得(3.51)中的 $\underline{V}_t^c = 0$，即

$$u'(c_t) = \beta R \underline{V}_{t+1}^c\left(\left(m_t - c_t\right) R + y_{t+1}\right), \tag{3.55}$$

[2]有时这种最优决策也会受到资源约束条件的限制，如流动性约束等，我们在这里暂不讨论这种情况。

代回(3.54)变为

$$V_t^{'}(m_t) = V_t^m(m_t, c_t(m_t)) = \beta R V_{t+1}^m(m_{t+1}). \tag{3.56}$$

此外(3.56)的 RHS 和(3.47)的 RHS 相等，因此有

$$V_t^{'}(m_t) = u'(c_t). \tag{3.57}$$

结合(3.56)和(3.57)，又有

$$u'(c_t) = \beta R u'(c_{t+1}). \tag{3.58}$$

至此可见，包络定理指的是，价值方程对任意变量的全求导，等于最优消费决策下，行为人消费该变量的边际效用为 0 的点。因此，可将 $c_t(m_t)$ 的最优决策看成是一个常数，对应 $c_t^{'}(m_t) = 0$，并直接将贝尔曼等式对 m_t 求导：

$$V_t(m_t) = u(c_t(m_t)) + \beta V_{t+1}((m_t - c_t(m_t)))R + y_{t+1}, \tag{3.59}$$

$$V_t^{'}(m_t) = \beta R V_{t+1}^{'}(m_{t+1}). \tag{3.60}$$

根据包络定理，不同的 m_t 值成为影响行为人效用最大化的关键，如图3.1所示。

图 3.1: 不同 m 取值下的包络定理

第四章　线性二次最优控制

多数情况下，均衡条件较为复杂，无法直接得出解析解。我们需要引入线性二次最优控制理论 (optimal linear-quadratic control, LQ)，将均衡条件用状态-空间形式表现出来 (state-space representation)。这里对线性二次最优控制论问题做简要介绍。

假定一个经济系统，由一系列状态变量的向量 x_t 和一系列控制变量的向量 u_t 构成；x_t 受到外生随机冲击 ε_t 的干扰，外生冲击的方差-协方差矩阵为 $\sum$。这样一个线性二次最优控制问题可以表述如下：经济系统的决策者通过选择 u_t 来影响 x_t 的走向，由此我们来构建目标方程和状态转移方程。

4.1　目标方程和状态转移方程

目标方程 (objective function)：

$$\max_{\{u_t\}} E \sum_{t=0}^{\infty} \beta^t \left[x_t^T R x_t + u_t^T Q u_t \right], \tag{4.1}$$

其中 β 是时间贴现系数，对称矩阵 R 和 Q 分别对应目标方程中状态变量和控制变量的权重。

状态转移方程 (transition function)：

$$x_{t+1} = A x_t + B u_t + \varepsilon_{t+1}, \tag{4.2}$$

当期状态由上期状态变量、上期控制变量所共同决定，二者的权重分别由矩阵 A 和 B 所反映。此外，当期状态还受到当期随机冲击的影响，假定是一个 Gaussian 过程 $\{\varepsilon_t\}_{t=0}^{\infty} \sim \mathcal{N}(0, \Sigma)$。

4.2　价值方程

在给定时间贴现系数 β 的情况下，如果已知系数矩阵 A, B, R, Q，我们可以求得系统的解。而对于系数矩阵未知的情况，问题则较为复杂。我们引入动态规划 (dynamic programming, DP) 的思路，定义价值方程 $V(x_t)$ 表示当期状态变量的价值。结合状态转移方程(4.2)和目标方程(4.1)，迭代形式表现的价值方程 (value function) 优化问题如下

$$\begin{aligned} V(x_t) &= \max_{\{u_t\}} \left\{ x_t^T R x_t + u_t^T Q u_t + \beta E[V(x_{t+1})] \right\} \\ &= \max_{\{u_t\}} \left\{ x_t^T R x_t + u_t^T Q u_t + \beta E[V(A x_t + B u_t + \varepsilon_{t+1})] \right\}, \end{aligned} \tag{4.3}$$

为了求解线性二次最优控制问题 (4.3)，我们需要两方面信息。一是政策方程 (policy function) 的具体形式 $u_t = g(x_t)$，即基于当期状态，决策者如何选择最优控制变量，见第4.3节。二是基于当期状态，价值方程的具体形式。

对于后者，方案设计如下：首先我们假设

$$V(x_t) = x_t^T P x_t + d, \tag{4.4}$$

即价值方程是关于状态变量的二项式形式 (第4.4节验证价值方程的二项式形式假设是否成立)，系数表示为幂等矩阵 P，满足 $P = P^T = P^2$，d 是个常数矩阵。

随后，将(4.4)代回(4.3)可得 Bellman equation 形式的价值方程

$$\begin{aligned} V(x_t) &= \max_{\{u_t\}} \left\{ x_t^T R x_t + u_t^T Q u_t + \beta E\left[(Ax_t + Bu_t + \varepsilon_{t+1})^T P (Ax_t + Bu_t + \varepsilon_{t+1}) \right] + \beta d \right\} \\ &\equiv \max_{\{u_t\}} \left\{ x_t^T R x_t + u_t^T Q u_t + \beta E[\mathcal{X}] + \beta d \right\}, \end{aligned} \tag{4.5}$$

其中

$$\begin{aligned} E[\mathcal{X}] =& x_t^T A^T P A x_t + x_t^T A^T P B u_t + u_t^T B^T P A x_t + u_t^T B^T P B u_t \\ &+ \left(x_t^T A^T P \varepsilon_{t+1} + u_t^T B^T P \varepsilon_{t+1} + \varepsilon_{t+1}^T P A x_t + \varepsilon_{t+1}^T P B u_t \right) + \varepsilon_{t+1}^T P \varepsilon_{t+1}, \end{aligned}$$

根据定义式 $E[\varepsilon] = 0$，上式进一步简化为

$$E[\mathcal{X}] = x_t^T A^T P A x_t + x_t^T A^T P B u_t + u_t^T B^T P A x_t + u_t^T B^T P B u_t + \varepsilon_{t+1}^T P \varepsilon_{t+1}. \tag{4.6}$$

4.2.1 一阶条件

决策者最优行为可由(4.5)的一阶条件求得

$$\begin{aligned} \frac{\partial V(x_t)}{\partial u_t} &= \frac{\partial u_t^T Q u_t}{\partial u_t} + \beta \frac{E[\mathcal{X}]}{\partial u_t} \\ &= \frac{\partial u_t^T Q u_t}{\partial u_t} + \beta \left(\frac{\partial x_t^T A^T P B u_t}{\partial u_t} + \frac{\partial u_t^T B^T P A x_t}{\partial u_t} + \frac{u_t^T B^T P B u_t}{\partial u_t} \right) \\ &= 2Q u_t + 2\beta B^T P A x_t + 2\beta B^T P B u_t = 0, \end{aligned} \tag{4.7}$$

其中为了求得最后一行等式，我们做一些矩阵运算见第4.2.1节。

计算一阶条件所需的部分矩阵运算

对于对称的常系数矩阵 Γ，首先我们有[1]

$$\frac{d}{du}(u^T \Gamma u) = u^T(\Gamma + \Gamma^T) = 2\Gamma u,$$

于是

$$\begin{aligned} \frac{\partial u_t^T Q u_t}{\partial u_t} &= 2Q u_t, \\ \frac{\partial u_t^T B^T P B u_t}{\partial u_t} &= 2B^T P B u_t. \end{aligned}$$

[1]这是由于

$$\begin{aligned} tr(d(u^T \Gamma u)) &= tr(d((u^T \Gamma)(u))) = tr(u^T \Gamma du + d(u^T \Gamma) u) = tr(u^T \Gamma du + d(u^T \Gamma)^T u) \\ &= tr(u^T \Gamma du) + tr(u^T + d(u^T \Gamma)) = tr(u^T \Gamma du) + tr(u^T \Gamma^T du) = tr(u^T (\Gamma + \Gamma^T) du). \end{aligned}$$

其次我们有[2]

$$\frac{d}{du}\Gamma u = \Gamma^T,$$
$$\frac{d}{du}u^T\Gamma = \Gamma,$$

于是

$$\frac{\partial x_t^T A^T P B u_t}{\partial u_t} + \frac{\partial u_t^T B^T P A x_t}{\partial u_t} = (x_t^T A^T P B)^T + (B^T P A x_t) = 2B^T P A x_t.$$

4.3　政策方程

重新整理(4.7)，可以得到政策方程 $u_t = g(x_t)$ 的近似线性表达形式

$$u_t = -\left(Q + \beta B^T P B\right)^{-1} \beta B^T P A x_t. \tag{4.8}$$

换句话说，决策者的最优行为可以描述为如下政策方程：

$$u_t = -F x_t, \quad \text{其中} \tag{4.9}$$
$$F = \beta\left(Q + \beta B^T P B\right)^{-1} B^T P A. \tag{4.10}$$

基于政策方程(4.9)，控制变量 u_t 随着观测到的状态 x_t 做线性调整，调整依据是系数矩阵 F。F 的确是一个非线性的函数形式，其值取决于两组矩阵：基础矩阵 A, B, Q 和 (假设为二次形式) 价值方程中的 P。一旦我们算出 P，便可以进一步算出 F 的值，从而求得政策方程的完整形式[3]。

4.3.1　政策方程满足确定性等价条件

政策方程的政策意义价值还在于，它不受外生随机冲击 ε 的干扰[4]，这是由于在这样一个线性二次系统中，确定性等价 (certainty equivalence) 成立，见第4.3.1、4.4.1节。

4.4　线性政策方程和二次价值方程

如前文所述，线性政策方程 (4.9)-(4.10) 是基于 (假设的) 二次价值方程 (4.5)-(4.6) 得出的。现在我们反证，基于这样的线性政策方程的确可以得到二次形式的价值方程。一旦反证成功，我们便可以进一步求得 P 和 d。

[2]这是由于

$$tr(du^T\Gamma) = tr(d(\Gamma^T u)^T) = tr(d(\Gamma^T u)).$$

[3]另一种方案是，先求得 F，根据 F 测算出 P，如 (Hansen and Sargent, 2004, Ch.2)

[4]除非以下情况出现：如第一，外生冲击彼此相关 (所以我们要在模型设定中假设不相关)。第二，目标方程(4.1)并不是二次形式。

将(4.9)代入(4.5)[5]

$$
\begin{aligned}
V(x_t) =& x_t^T R x_t + (-Fx_t)^T Q(-Fx_t) + \beta d + \beta E[\varepsilon_{t+1}^T P \varepsilon_{t+1}] \\
&+ \beta E\left[x_t^T A^T P A x_t + x_t^T A^T P B(-Fx_t) + (-Fx_t)^T B^T P A x_t + (-Fx_t)^T B^T P B(-Fx_t)\right] \\
=& x_t^T R x_t + x_t^T F^T Q F x_t + \beta d + \beta E[\varepsilon_{t+1}^T P \varepsilon_{t+1}] \\
&+ \beta E\left[x_t^T A^T P A x_t - x_t^T A^T P B F x_t - x_t^T F^T B^T P A x_t + x_t^T F^T B^T P B F x_t\right] \\
=& x_t^T R x_t + x_t^T F^T Q F x_t + \beta E\left[x_t^T A^T P A x_t - 2x_t^T A^T P B F x_t + x_t^T F^T B^T P B F x_t\right] \\
&+ \beta E[\varepsilon_{t+1}^T P \varepsilon_{t+1}] + \beta d.
\end{aligned} \tag{4.11}
$$

联立(4.4)和(4.11)，我们有

$$
\begin{cases}
d = \beta E[\varepsilon_{t+1}^T P \varepsilon_{t+1}] + \beta d, \\
x_t^T P x_t = x_t^T R x_t + x_t^T F^T Q F x_t + \beta E\left[x_t^T A^T P A x_t - 2x_t^T A^T P B F x_t + x_t^T F^T B^T P B F x_t\right].
\end{cases}
$$

4.4.1 价值方程满足确定性等价条件

来看第一个等式，整理可得[6]

$$
d = \frac{\beta}{1-\beta} tr(P\Sigma), \tag{4.12}
$$

可见随机冲击 ε 尽管对价值方程产生影响，但不是通过 F 至政策方程进而影响价值方程，而是通过常数项 d 影响价值方程的。因此，对于最优价值方程而言，确定性等价条件依然成立。

[5]第三个等号所需条件：根据定义 $P = P^T$，因此 $x_t^T F^T B^T P A x_t$ 和 $x_t^T A^T P B F x_t$ 都是标量，且满足

$$
x_t^T F^T B^T P A x_t = x_t^T A^T P B F x_t.
$$

[6]这是由于

$$
E[\varepsilon_{t+1}^T P \varepsilon_{t+1}] = tr(E[\varepsilon_{t+1}^T P \varepsilon_{t+1}]) = tr(PE[\varepsilon_{t+1}^T \varepsilon_{t+1}]) = tr(PE[\Sigma_{t+1}]).
$$

4.4.2 代数矩阵 Riccati 方程

第二个等式整理，并引入最优政策方程(4.10)替换 F 得[7]

$$
\begin{aligned}
P =& R + F^T QF + \beta \cdot \left(A^T PA - 2A^T PBF + F^T B^T PBF\right) \\
=& R + \beta A^T PA - 2\beta A^T PBF + F^T \left(Q + \beta B^T PB\right) F \\
=& R + \beta A^T PA - 2\beta A^T PB \left(Q + \beta B^T PB\right)^{-1} \beta B^T PA \\
&+ \beta A^T PB \left(\left(Q + \beta B^T PB\right)^{-1}\right)^T \left(Q + \beta B^T PB\right) \left(Q + \beta B^T PB\right)^{-1} \beta B^T PA \\
=& R + \beta A^T PA - \beta^2 A^T PB \left(Q + \beta B^T PB\right)^{-1} \beta B^T PA.
\end{aligned} \tag{4.13}
$$

(4.13)又称线性矩阵 Riccati 方程 (linear matrix Riccati equation)，它表明价值方程中的 P 是个与基础矩阵 A, B, R, Q 有关的函数，呈非线性关系。

此外，由(4.13)可以看出，一个线性的最优政策方程(4.9)-(4.10)，的确预示着二次形式的价值方程 (4.4),(4.13)。

4.5 数值方法

在线性二次控制中，在利用解析法推得线性代数 Riccati 方程后，需要依赖数值计算的方法，对 P 的矩阵 Riccati 差分方程做迭代近似。假定期初猜测值为 P_j，经过 1 次迭代，P_{j+1} 的值更新至

$$
P_{j+1} = R + \beta A^T P_j A - \beta^2 A^T P_j B \left(Q + \beta B^T PB\right)^{-1} \beta B^T P_j A. \tag{4.14}
$$

重复迭代，直至 P 收敛到某一值。根据数值模拟的 P 值计算 F 值，进而得到最优政策方程和价值方程。

4.6 范例

假定这样一个线性二次控制问题。中央银行试图通过控制利率 r_t 来影响通货膨胀率 π_t 和产出 y_t，收益方程 (payoff function)

$$
\mathcal{L}_t = \pi_t^2 + y_t^2 + 0.1 r_t^2. \tag{4.15}
$$

经济系统的结构，由两个状态变量的运动定律构成

$$
\pi_{t+1} = 0.75\pi_t - 0.5 r_t + \varepsilon_{\pi,t+1}, \tag{4.16a}
$$

$$
y_{t+1} = 0.25 y_t - 0.5 r_t + \varepsilon_{y,t+1}. \tag{4.16b}
$$

[7]最后一个等号的计算依据如下：

$$
\begin{aligned}
&\beta A^T PB \left(\left(Q + \beta B^T PB\right)^{-1}\right)^T \left(Q + \beta B^T PB\right) \left(Q + \beta B^T PB\right)^{-1} \beta B^T PA \\
&= \beta A^T PB \left(\left(Q + \beta B^T PB\right)^T\right)^{-1} \left(Q + \beta B^T PB\right) \left(Q + \beta B^T PB\right)^{-1} \beta B^T PA \\
&= \beta A^T PB \left(\left(Q + \beta B^T PB\right)^T\right)^{-1} \left(Q + \beta B^T PB\right)^T \left(Q + \beta B^T PB\right)^{-1} \beta B^T PA \\
&= \beta^2 A^T PB \left(Q + \beta B^T PB\right)^{-1} B^T PA.
\end{aligned}
$$

作为决策者，中央银行的最大化问题为

$$\max_{\{r_t\}} E\sum_{t=0}^{\infty}\beta^t\mathcal{L}_t, \quad \text{s.t.}(4.16a)-(4.16b)$$

4.6.1 最优线性二次控制问题

将这一问题改写为最优线性二次控制的一般形式：

$$\max_{\{u_t\}} E\sum_{t=0}^{\infty}\beta^t\left(x_t^T R x_t + u_t^T Q u_t\right),$$
$$\text{s.t. } x_{t+1} = Ax_t + Bu_t + \varepsilon_t,$$

其中状态变量，控制变量，外生扰动变量分别用 x_t, u_t, ε_t 表示

$$x_t \equiv \begin{pmatrix}\pi_t\\ y_t\end{pmatrix}, \quad u_t \equiv r_t, \quad \varepsilon_t = \begin{pmatrix}\varepsilon_{\pi,t}\\ \varepsilon_{y,t}\end{pmatrix},$$

基础系数矩阵 A, B, R, Q 分别为

$$A = \begin{pmatrix}0.75 & 0\\ 0 & 0.25\end{pmatrix}, \quad B = \begin{pmatrix}-0,5\\ -0.5\end{pmatrix}, \quad R = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad Q = 0.1.$$

如前文所述，我们需要使用数值方法，递归处理线性矩阵 Riccati 方程以近似价值方程系数矩阵 P，进而测算政策方程的系数矩阵 F，对应 (4.14) (4.10)。

4.6.2 Matlab 程序

```
clear;
%时间贴现系数的确定
beta = 0.99;

%定义矩阵
R=zeros(2,2);
Q=zeros(1,1);
A=zeros(2,2);
B=zeros(2,1);

%为矩阵赋值
R(1,1)=1;
R(2,2)=1;
Q(1,1)=0.1;
A(1,1)=0.75;
A(2,2)=0.25;
B(1,1)=-0.5;
B(2,1)=-0.5;

%matrix Riccati equation迭代
```

```
%d描述$P_j$和邻近迭代$P_{j+1}$之间元素的最大绝对值偏差
d = 1;
%i记录迭代次数
i = 0;
% 最大绝对值偏差d保存在D; 迭代次数i保存在I
D=0;
I=0;
%迭代P的初始值定义为P0
%说明: 定义P0和Q的非零值, 为了使得linear matrix Riccati equation迭代操作中
%$(Q+\beta B^T P_j B) \neq 0$. 只有这个值不等于0, 才能使得该矩阵是可逆的。
P0=zeros(2,2);
P0(1,1)=-0.000001;
P0(2,2)=-0.000001;


format long

%matrix Riccati Equation 迭代
%迭代新生成的$P_{j+1}$写入矩阵P1。
%两个P值的差值, 写入矩阵Pd。
%当d<0.0000000001时, 迭代终止。否则继续进行。
%每次迭代的d和i写入D和I, 以备最终程序输出。
while d > 0.0000000001
P1 = R + beta * A' * P0 * A - beta^2 * A' * P0 * B * (inv(Q+beta * B' * P0 * B)) * (B' * P0 * A);
Pd = P1 - P0;
d = max(abs(Pd));
d = max(d');
D=[D d];
P0 = P1;
i = i+1;
I = [I i];
end

%根据迭代生成的P值, 计算policy function的矩阵F
P=P0;
F = -inv(Q + beta * B' * P * B) * (beta * B' * P * A);
ID = [I(2:length(I))' D(2:length(I))'];
disp('          i                                d');
disp(ID);
disp('  SOLUTIONS');
```

```
disp('F');
disp(F);
disp('P');
disp(P);
```

4.6.3 Matlab 程序输出

运行 Matlab 程序，输出如下

```
  i                              d
1.000000000000000    1.000000938124847
2.000000000000000    0.325234163624664
3.000000000000000    0.080558191789241
4.000000000000000    0.018875706161742
5.000000000000000    0.004338686564976
6.000000000000000    0.000992730600232
7.000000000000000    0.000226907809828
8.000000000000000    0.000051851748175
9.000000000000000    0.000011848232754
10.000000000000000    0.000002707312049
11.000000000000000    0.000000618616947
12.000000000000000    0.000000141352999
13.000000000000000    0.000000032298933
14.000000000000000    0.000000007380254
15.000000000000000    0.000000001686376
16.000000000000000    0.000000000385334
17.000000000000000    0.000000000088048

SOLUTIONS
F
0.744954171236066    0.175909878487998

P
1.430293038776173   -0.106183304364736
-0.106183304364736    1.044189928712962
```

由此可得最优政策方程

$$u_t = -Fx_t, \quad r_t = -F\begin{pmatrix}\pi_t \\ y_t\end{pmatrix}, F = (0.745, 0.176) \tag{4.17}$$

$$r_t = 0.745\pi_t + 0.176y_t. \tag{4.18}$$

价值方程

$$V(\pi_t, y_t) = 1.430\pi_t^2 + 1.044y_t^2 - 0.212y_t. \tag{4.19}$$

根据最优政策方程，当通胀水平和/或产出水平高于稳定状态时，中央银行应当提高利率[8]。利率对通胀波动的响应幅度高于对产出波动的响应 (0.745 相对于 0.176)，这与模型的假设条件有关：(4.16a)-(4.16b)假定通胀波动比产出波动更为持久 (0.75 相对于 0.25)，从而过去时段的物价波动，比起产出波动来，更有可能对当期经济系统产生影响。

根据价值方程，首先同样地，通胀波动二次项的系数大于产出波动二次项的系数。交互项的系数为负，反映了当通胀和产出的波动同方向变化时，比如通货膨胀伴随产出增加（或通货紧缩伴随产出减少），中央银行更容易通过调节利率这一政策工具来稳定经济运行。反之，如果两种波动反方向变化，稳定经济运行将更为困难。

[8]注意式中的变量 $z_t = (\tilde{z}_t - z)/z$，表示实际观测到的变量，相对于稳态水平的 deviation。

第五章　理性期望模型

5.1　简介

利用动态规划 (dynamic programming) 的方法求解 Ramsey 随机增长模型，一个核心假定是只存在一个典型的经济行为个体，他追求自身利益最大化的行为带来社会福利最大化。我们称这样的经济体为 centralized economy，均衡状态处于 Pareto optimality，这个经济行为人称为 social planner。

然而现实中更常见 decentralized economy 的情况，即存在异质的多个经济个体，他们的最大化目标各异，如厂商追求利润最大化，家庭追求效用最大化，劳动者做劳动——休闲的最优决策，等等。的确，根据福利经济学第二定律 (the second fundamental theorem of welfare economics, (Mas-Colell et al., 1995, p.151))，在某些极端情况下，完全竞争的 decentralized 均衡状态可以导致 social planner 的 Pareto optimality；但对于更一般的情况，当 decentralized economies 中存在局部摩擦如价格/工资粘性时，decentralized 均衡并不必然导致 centralized 的 Pareto optimality。换句话说，我们无法通过求解 social planner 问题来推得 decentralized 均衡。我们只能直接从不同异质经济个体的一阶条件 (FOC) 入手，构建这样一组线性随机一阶差分方程，即理性期望模型。

求解理性期望模型的核心在于线性化：在将多阶自回归改写成一阶自回归形式之后，如何将原本是高度非线性的一阶自回归系统做线性近似。对于稍微复杂一些的系统而言，往往无法直接求得解析解，替代方案为：首先将模型变量围绕其稳定状态做对数线性化处理求得解析式；随后根据一定的数值算法，求解线性化方程组，将相关内生变量改写为自 VAR 形式。

此种求解思路主要来自King et al. (1988a,b)[1]，大致说来分为四个步骤：

1. 计算稳定状态，
2. 将解释变量围绕稳定状态做近似，求解析式，
3. 模型参数校准，
4. 根据数值算法求得政策方程，将内生变量与外生变量和前定变量联系起来。

该思路在宏观经济研究中得到了较为广泛的应用，主要得益于其优点：存在多种可供选择的数值算法，可以在一定程度上近似非线性一阶条件的线性表达式，而只需要付出一定的计算机处理时间。也正因如此，该思路也存在着适用性的局限。

1. 围绕稳定状态做对数线性近似，其前提假定是模型接近对数线性形式。而模型非线性的程度越高，模型的规模越大，考虑的因素越是多（比如消费者的风险厌恶程度越高，外生冲击的种类越多），对数线性近似导致的失真情况就越严重。
2. 稳定状态无法在模型内部求得，并且对于存在多重稳态的经济系统来说，情况会变的更复杂。

[1] 此外可见King and Rebelo (1999)。

5.2 数值算法

针对一阶线性近似的变型系统，常见的数值算法，大致说来有[2]

1. 特征值——特征向量分解法 (eigenvalue-eigenvector decomposition)

   最早由Blanchard and Kahn (1980) 提出，因此也称为 Blanchard-Kahn Algorithm。模型要求将所有内生变量分为两类，一类为状态变量，主要指前定变量。其他变量归入第二类跳跃变量 (jump variables)。通过特征根——特征变量分解，该算法求得跳跃变量爆炸根 (explosive root) 的数量，进而判断系统解是否存在，以及是否唯一。见第5.4节。

   对于满足一些假定条件的 (比如非奇异方块矩阵) 常规系数矩阵，特征根——特征向量分解法提供了一种较好的求解思路。由此优点出发，在算法方面，一些后续研究作出不断改进 (King and Watson, 1998, 2002; Anderson and Moore, 1985; Anderson, 1998; Sims, 2002; Klein, 2000; Söderlind, 1999)。

2. 未定系数法 (undetermined coefficients)

   最早由McCallum (1983) 提出，随后的一系列重要扩展包括Uhlig (1999); Binder and Pesaran (1995); Christiano (2002) 等。这种方法不再将变量区分为前定与非前定变量，根据未定系数近似系统的数值解。见第5.5节。

   其基本思路是：第一，假定系统存在 1 个解，根据这个解，内生变量是关于状态变量的线性方程，并且这个假定不能使完全随机的猜测。第二，将猜测解代回结构方程系统中，构建关于待定系数的方程组。第三，利用二次方程矩阵的解法，求得该方程组的解，进而整个线性系统的解。

   未定系数法具有可操作性和计算速度等优势。但存在不足：第一，需要预先假定模型存在唯一解。第二，仅当系统中没有冗余变量时 (系数矩阵的列线性不相关)，状态——空间表现形式处于最简规模，该方法才适用。否则值为 0 的特征值会导致泡沫解的出现 (Canova, 2011, p.57)。

3. 期望误差法 (expectational errors)

   由Sims (2002) 提出，也不再做前定、非前定变量的区分，更可以进一步探讨理性期望下期望误差的性质。见第5.6节。

4. 参数化期望算法 (PEA, Parameterised Expectations Algorithm)，见第5.7节。

5. Schur 分解法 (Schur Decomposition)，又称 QZ 分解法，作为更为通用的形式，主要用于奇异 (不可逆) 系数矩阵的情况 (King and Watson, 1998, 2002; Söderlind, 1999; Klein, 2000)，见第5.8节。

大致说来，各种方法之间的主要区别在于

- 构建稳定解模块的方式，
- 求解过程中对理性期望的处理，
- 保留多大部分的非线性成分，交给数值算法去近似处理，
- 对前定和非前定变量的区分等。

[2]一个更为详尽的综述，见 (Milani, 2012)。

根据King et al. (1988a,b) 的四步骤求解思路，第5.3节首先建立一个简单的随机增长模型。在此基础上，第5.3.1节计算稳定状态，第5.3.2节做对数线性化近似，第5.4-5.8节分别介绍几种主要的算法。

5.3　一个简单的随机增长模型

在这样一个简单随机经济增长模型中，典型个体追求最大化问题

$$
\begin{aligned}
&\max_{\{C_t\}} E\sum_{t=0}^{\infty}\beta^t\cdot\left(\frac{C_t^{1-\sigma}}{1-\sigma}\right), \quad \text{s.t.}\\
&C_t+K_{t+1}=A_t\cdot K_t^{\alpha}+(1-\delta)\cdot K_t,\\
&\ln A_t=\rho\cdot\ln A_{t-1}+\varepsilon_t,, \quad \varepsilon_t\sim i.i.d.(0,\sigma^2), 0<\rho<1.
\end{aligned}
$$

模型的均衡解以 $\{C_t,K_t,Y_t\}_{t=0}^{\infty}$ 的形式展现。

求解一阶条件我们有

$$
\begin{cases}
C_t^{-\sigma} & =\beta E\left[C_{t+1}^{-\sigma}\cdot\left(\alpha\cdot A_{t+1}\cdot K_{t+1}^{\alpha-1}+1-\delta\right)\right]\\
C_t+K_{t+1} & =A_t\cdot K_t^{\alpha}+(1-\delta)\cdot K_t\\
\ln A_t & =\rho\cdot\ln A_{t-1}+\varepsilon_t
\end{cases}
$$

5.3.1　稳定状态

在稳定状态下我们有

$$
\begin{bmatrix}C_t\\K_t\\Y_t\end{bmatrix}\equiv\begin{bmatrix}\bar{C}\\\bar{K}\\\bar{Y}\end{bmatrix}, \quad A_t\equiv\bar{A}=1, \quad \forall t.
$$

进而我们有

$$
\begin{cases}
1 & =\beta\cdot\left[\alpha\cdot\bar{K}^{\alpha-1}+(1-\delta)\right],\\
\bar{C}+\bar{K} & =\bar{K}^{\alpha}+(1-\delta)\cdot\bar{K}.
\end{cases}
$$

整理后得稳定状态

$$
\begin{cases}
\bar{C} & =\bar{K}^{\alpha}-\delta\cdot\bar{K},\\
\bar{K} & =\left(\frac{1-(1-\delta)\cdot\beta}{\alpha\cdot\beta}\right)^{\frac{1}{\alpha-1}},\\
\bar{Y} & =\bar{K}^{\alpha}.
\end{cases}
$$

5.3.2　对数线性化

对数线性的定义式可表示为，对于一个变量 X_t：

$$
\tilde{X}_t\equiv\frac{X_t-\bar{X}}{\bar{X}}\approx\ln X_t-\ln\bar{X},
$$

对于含有不止一个变量的复杂方程 $f(X_t,Y_t)$，对其稳定状态 $(\bar{X},\bar{Y})$ 做一阶泰勒级数展开的方式为

$$
\ln f(X_t,Y_t)\approx\ln f(\bar{X},\bar{Y})+\left[\frac{\frac{\partial f(X_t,Y_t)}{\partial X_t}|_{\{\bar{X},\bar{Y}\}}}{f(\bar{X},\bar{Y})}\right]\cdot\left(X_t-\bar{X}\right)+\left[\frac{\frac{\partial f(X_t,Y_t)}{\partial Y_t}|_{\{\bar{X},\bar{Y}\}}}{f(\bar{X},\bar{Y})}\right]\cdot\left(Y_t-\bar{Y}\right).
$$

外生技术冲击的对数线性化

$$\tilde{A}_t \approx \rho \cdot \tilde{A}_{t-1} + \varepsilon_t. \tag{5.1}$$

Euler equation 的对数线性化

等式两侧取对数

$$-\sigma \cdot \ln C_t = \ln \beta - \sigma \cdot E \ln C_{t+1} + E \ln \left[\alpha \cdot A_{t+1} \cdot K_{t+1}^{\alpha-1} + (1-\delta)\right]$$

RHS 第三部分，围绕 $\{\bar{K}, \bar{A}\}$ 做一阶泰勒展开

$$\begin{aligned}
&\ln \left[\alpha \cdot A_{t+1} \cdot K_{t+1}^{\alpha-1} + (1-\delta)\right] \\
&\approx \ln \left[\alpha \bar{A} \bar{K}^{\alpha-1} + (1-\delta)\right] \\
&+ \frac{\frac{\partial}{\partial \bar{K}} \left[\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)\right] \cdot \frac{K_{t+1}-\bar{K}}{\bar{K}} \cdot \bar{K}}{\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)} + \frac{\frac{\partial}{\partial \bar{A}} \left[\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)\right] \cdot \frac{A_{t+1}-\bar{A}}{\bar{A}} \cdot \bar{A}}{\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)} \\
&= \ln \left[\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)\right] + \frac{\alpha \cdot \bar{A} \cdot (\alpha-1) \cdot \bar{K}^{\alpha-1} \cdot \tilde{K}_{t+1} + \alpha \cdot \bar{K}^{\alpha-1} \cdot \bar{A} \cdot \tilde{A}_{t+1}}{\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)} \\
&= \ln \left[\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)\right] + \frac{\alpha \cdot \bar{K}^{\alpha-1} \cdot \left[(\alpha-1) \cdot \tilde{K}_{t+1} + \tilde{A}_{t+1}\right]}{\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha-1} + (1-\delta)} \\
&= -\ln \beta + \left[1 - \beta \cdot (1-\delta)\right] \cdot \left[(\alpha-1) \cdot \tilde{K}_{t+1} + \tilde{A}_{t+1}\right],
\end{aligned}$$

进而我们有

$$-\sigma \cdot \tilde{C}_t = -\sigma \cdot \tilde{C}_{t+1} + \left[1 - \beta \cdot (1-\delta)\right] \cdot \left[(\alpha-1) \cdot \tilde{K}_{t+1} + \tilde{A}_{t+1}\right] \tag{5.2}$$

预算约束条件的对数线性化

等式两侧取对数

$$\ln(C_t + K_{t+1}) = \ln \left[A_t \cdot K_t^{\alpha} + (1-\delta) \cdot K_t\right]$$

LHS $\Rightarrow$

$$\ln(C_t + K_{t+1}) \approx \ln\left(\bar{C} + \bar{K}\right) + \frac{\frac{C_t - \bar{C}}{\bar{C}} \cdot \bar{C} + \frac{K_{t+1}-\bar{K}}{\bar{K}} \cdot \bar{K}}{\bar{C} + \bar{K}} = \ln\left(\bar{C} + \bar{K}\right) + \frac{\bar{C} \cdot \tilde{C}_t + \bar{K} \cdot \tilde{K}_{t+1}}{\bar{C} + \bar{K}}.$$

RHS $\Rightarrow$

$$\begin{aligned}
&\ln \left[A_t \cdot K_t^{\alpha} + (1-\delta) \cdot K_t\right] \\
&\approx \ln \left[\bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}\right] + \frac{\frac{\partial}{\partial \bar{K}} \left[\bar{A} \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}\right] \cdot \frac{K_t - \bar{K}}{\bar{K}} \cdot \bar{K}}{\bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}} + \frac{\frac{\partial}{\partial \bar{A}} \left[\bar{A} \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}\right] \cdot \frac{A_t - \bar{A}}{\bar{A}} \cdot \bar{A}}{\bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}} \\
&= \ln \left[\bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}\right] + \frac{\left[\alpha \cdot \bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}\right] \cdot \tilde{K}_t + \bar{A} \cdot \bar{K}^{\alpha} \cdot \tilde{A}_t}{\bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}} \\
&= \ln \left[\bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}\right] + \frac{\frac{1}{\beta} \cdot \bar{K} \cdot \tilde{K}_t + \bar{A} \cdot \bar{K}^{\alpha} \cdot \tilde{A}_t}{\bar{A} \cdot \bar{K}^{\alpha} + (1-\delta) \cdot \bar{K}} \\
&= \ln\left(\bar{C} + \bar{K}\right) + \frac{\frac{1}{\beta} \cdot \bar{K} \cdot \tilde{K}_t + \bar{K}^{\alpha} \cdot \tilde{A}_t}{\bar{C} + \bar{K}}.
\end{aligned}$$

$LHS = RHS \Rightarrow$

$$\bar{C} \cdot \tilde{C}_t + \bar{K} \cdot \tilde{K}_{t+1} = \frac{1}{\beta} \cdot \bar{K} \cdot \tilde{K}_t + \bar{K}^{\alpha} \cdot \tilde{A}_t,$$

因此我们有

$$\begin{aligned}\tilde{K}_{t+1} &= -\frac{\bar{C}}{\bar{K}}\cdot\tilde{C}_t+\frac{1}{\beta}\cdot\tilde{K}_t+\bar{K}^{\alpha-1}\cdot\tilde{A}_t\\ &= -\frac{1-\beta\cdot[1-\delta\cdot(1-\alpha)]}{\alpha\cdot\beta}\cdot\tilde{C}_t+\frac{1}{\beta}\cdot\tilde{K}_t+\frac{1-\beta\cdot(1-\delta)}{\alpha\cdot\beta}\cdot\tilde{A}_t.\end{aligned} \tag{5.3}$$

5.3.3　状态——空间表现形式

一阶差分式(5.1)、(5.2)、(5.3)共同描述这样一个动态的经济系统 $\{C_t,K_t,A_t\}_{t=0}^{\infty}$。可以将其改写为如下状态——空间表现形式 (state-space representation)

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&\sigma\end{bmatrix}E\begin{bmatrix}\tilde{A}_{t+1}\\\tilde{K}_{t+1}\\\tilde{C}_{t+1}\end{bmatrix}=\begin{bmatrix}\rho&0&0\\\frac{1-\beta\cdot(1-\delta)}{\alpha\cdot\beta}&\frac{1}{\beta}&-\frac{1-\beta\cdot[1-\delta\cdot(1-\alpha)]}{\alpha\cdot\beta}\\\rho\cdot[1-\beta\cdot(1-\delta)]&-(1-\alpha)\cdot[1-\beta\cdot(1-\delta)]&\sigma\end{bmatrix}\begin{bmatrix}\tilde{A}_t\\\tilde{K}_t\\\tilde{C}_t\end{bmatrix}. \tag{5.4}$$

5.4　特征值——特征方程分解法

5.4.1　特征值——特征方程分解法

根据 Blanchard-Kahn 算法，要将经济系统中的 $p+m+k$ 个变量分类。一类是 k 个外生变量 z_t。一类是内生变量 x_t。内生变量再分为 m 个控制变量 (跳跃变量)x_t^j 和 p 状态变量 (前定变量)x_t^s。(5.4)可以改写为

$$\mathcal{T}_0Ex_{t+1}=\mathcal{T}_1x_t+\Psi z_t,\quad x_t\equiv\begin{bmatrix}x_t^s&x_t^j\end{bmatrix}'. \tag{5.5}$$

假定系数矩阵 $\mathcal{T}_0$ 是可逆的，上式进一步调整为

$$\begin{gathered}Ex_{t+1}=Ax_t+Bz_t,\quad 其中A\equiv\mathcal{T}_0^{-1}\mathcal{T}_1,\quad B\equiv\mathcal{T}_0^{-1}\Psi,\quad 或者\\ E\begin{bmatrix}\underset{p\times1}{x_{t+1}^s}\\\underset{m\times1}{x_{t+1}^j}\end{bmatrix}=\begin{bmatrix}\underset{p\times p}{A_{11}}&\underset{p\times m}{A_{12}}\\\underset{m\times p}{A_{21}}&\underset{m\times m}{A_{22}}\end{bmatrix}\begin{bmatrix}\underset{p\times1}{x_t^s}\\\underset{m\times1}{x_t^j}\end{bmatrix}+\begin{bmatrix}\underset{p\times k}{B_1}\\\underset{m\times k}{B_2}\end{bmatrix}\underset{k\times1}{z_t}\end{gathered} \tag{5.6}$$

Blanchard-Kahn 算法的核心就在于，对系数矩阵 A 做 Jordan decomposition。对于可对角化 (diagonalizable) 的系数矩阵 A，假定其特征向量是序列不相关的[3]，我们有 $AP=P\Lambda$，改写为 Jordan canonical form

$$A=P\Lambda P^{-1}, \tag{5.7}$$

其中 Λ 是对角矩阵，对角元素对应特征根 $\Lambda_{ii}=\lambda_i$，其他元素均为 0。P 的每一列为对应 λ_i 的特征向量。我们用 $\bar{p}$ 和 $\bar{m}$ 来表示用稳定和不稳定特征根来为矩阵分块的情况，后面会简要探讨当 $\bar{p}\neq p,\bar{m}\neq m$ 时可能会出现的问题，见第5.4.2节。

根据 Blanchard-Kahn 条件，不稳定特征根 (即 $|\lambda|>1$) 的数量应该恰好等于经济系统中控制变量的数量，以确保相图中鞍点稳定性的存在。如果不稳定特征根的数量少于控制变量的数量，经济系统超稳定，出现未定问题 (indeterminacy)[4]。如果多于，经济系统爆炸性发展，会违反横截条件。一个较为详细的讨论见第5.4.2节。

[3]对于 n 个长度为 m 的向量之间线性不相关，是指对于由这 n 个向量构成的 mxn 矩阵 X 来说，$det(x)\neq0$。

[4]对未定经济系统问题的探讨，可见如Benhabib and Farmer (1999)。

在本章的动态 Ramsey 经济模型中，内生变量是二维的，由一个控制变量和一个前定状态变量构成。因此需要恰好有一个不稳定特征根和一个稳定特征根。

将特征矩阵做进一步的分解，根据特征根绝对值从低到高做重新排列 (相应地，需要调整特征矩阵的列)

$$\Lambda = \begin{bmatrix} \underset{\bar{p}\times\bar{p}}{\Lambda_s} & \underset{\bar{p}\times\bar{m}}{0} \\ \underset{\bar{m}\times\bar{p}}{0} & \underset{\bar{m}\times\bar{m}}{\Lambda_e} \end{bmatrix}, \quad P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}.$$

其中对角矩阵 $\Lambda_s(\Lambda_e)$ 的对角元素是绝对值小于 (大于)1 的特征根。系统(5.6)由此改写为

$$Ew_{t+1} = \Lambda w_t + \bar{B}z_t, \quad \text{其中} w_t \equiv P^{-1}x_t, \bar{B} \equiv P^{-1}B. \tag{5.8}$$

对应地我们有

$$w_t = \begin{bmatrix} \underset{\bar{p}\times 1}{w_{1,t}} \\ \underset{\bar{m}\times 1}{w_{2,t}} \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} \underset{\bar{p}\times\bar{k}}{\bar{B}_1} \\ \underset{\bar{m}\times k}{\bar{B}_2} \end{bmatrix}.$$

根据上述分析，由(5.8)我们有

$$Ew_{1,t+1} = \Lambda_s w_{1,t} + \bar{B}_1 z_t, \tag{5.9a}$$

$$Ew_{2,t+1} = \Lambda_e w_{2,t} + \bar{B}_2 z_t. \tag{5.9b}$$

根据模型设定，(5.9a)总是稳定的，因为 Λ_s 的对角元素由绝对值小于 1 的特征根构成。(5.9b)是爆炸的，因为 Λ_e 的对角元素由绝对值大于 1 的特征根构成。因此我们先来考察(5.9b)成立需要满足的条件。

我们对 $w_{2,t}$ 做 forward-looking 迭代，求解等式。(5.9b)意味着

$$\begin{aligned} w_{2,t} &= (\Lambda_e)^{-1}Ew_{2,t+1} - \bar{B}_2 z_t, \\ w_{2,t+1} &= (\Lambda_e)^{-1}Ew_{2,t+2} - \bar{B}_2 z_{t+1}, \\ &\vdots \\ w_{2,t+T} &= (\Lambda_e)^{-1}Ew_{2,t+T} - \bar{B}_2 z_{t+T}. \end{aligned}$$

根据 forward-looking 迭代我们有

$$w_{2,t} = \lim_{T\to\infty} \Lambda_e^{-T} E(w_{2,t+T}) - \sum_{s=0}^{T} \Lambda_e^{-s-1} \bar{B}_2 E(z_{t+s}). \tag{5.10}$$

由于 Λ_e 中所有对角元素绝对值都大于 1，对于 $T \to \infty$ 可得 $\Lambda_e^T \to 0$，因此改写上式，$w_{2,t}$ 的值可以计算如下

$$w_{2,t} = -\sum_{s=0}^{T} \Lambda_e^{-s-1} \bar{B}_2 E(z_{t+s}). \tag{5.11}$$

回到反映 w_t 和 x_t 关系的(5.8)中，根据定义

$$x_t \equiv Pw_t \Rightarrow \begin{bmatrix} \underset{\bar{p}\times 1}{x_{1,t}} \\ \underset{\bar{m}\times 1}{x_{2,t}} \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \underset{\bar{p}\times 1}{w_{1,t}} \\ \underset{\bar{m}\times 1}{w_{2,t}} \end{bmatrix},$$

等价于

$$x_{1,t} = P_{11}w_{1,t} + P_{12}w_{2,t}, \tag{5.12a}$$

$$x_{2,t} = P_{21}w_{1,t} + p_{22}w_{2,t}. \tag{5.12b}$$

如果假定 $\bar{p}=p,\bar{m}=m$，即系统中控制变量的数量等于不稳定特征根的数量，并且假定 P_{11} 是可逆分块矩阵，则 $x_{1,t}=x_t^s,x_{2,t}=x_t^j$，对于二者不相等的情况，见第5.4.2节。对于最终形成 Blanchard-Kahn 算法的求解思路：

1. 从初始状态 $t=0$ 出发，对应给定的 $x_{t=0}^s$。

2. 结合生成的 z_t，根据(5.11)计算 $w_{2,t}$：

$$w_{2,t}=-\sum_{s=0}^{T}\Lambda_e^{-s-1}\bar{B}_2E(z_{t+s}).$$

3. 根据(5.12a)得到 $w_{1,t}$ 的值：

$$w_{1,t}=P_{11}^{-1}x_t^s-P_{11}^{-1}P_{12}w_{2,t},$$

   其中前定变量 x_t^s 由上一期的状态求得。

4. 根据(5.12b)得到控制变量 x_t^j 的值，又称政策方程：

$$x_t^j=P_{21}w_{1,t}+P_{22}w_{2,t}.$$

5. 根据 (5.6)得到 $t+1$ 期状态变量的期望值 x_{t+1}^s：

$$Ex_{t+1}^s=A_{11}x_t^s+A_{12}x_t^j+B_1z_t,$$

6. 重复以上步骤。

5.4.2　解的存在性以及唯一性

如前文所述，当 $p=\bar{p},m=\bar{m}$ 时，控制变量（状态变量）的数量等于爆炸（稳定）特征根的数量，此时系统均衡解存在且唯一。

但当 $m<\bar{m}$ 时，$p>\bar{p}$，情况有所不同。我们仍然可以根据(5.12a)得到 $w_{1,t}$ 的值，尽管 $w_{1,t}$ 的值不唯一。但在根据(5.12b)测算 x_t^j 时，会受到较大限制，比如在 $t=0$ 时：

$$x_{2,t_0}=P_{21}w_{1,t_0}+p_{22}w_{2,t_0},$$

此时，x_{2,t_0} 的 $\bar{m}$ 行中，最上面的 $\bar{m}-m\equiv p-\bar{p}$ 行是状态变量，因此是提前给定的。这意味着我们没有足够数量的前定（状态）变量用于求解方程。换句话说，对整个经济系统而言，当跳跃变量的数量少于爆炸特征根时，系统的均衡解不存在。

当 $m>\bar{m}$ 时，$p<\bar{p}$，在 $t=0$ 时(5.12a)变为

$$w_{1,t_0}=P_{11}^{-1}x_{1,t_0}-P_{11}^{-1}P_{12}w_{2,t_0},$$

x_{1,t_0} 的 $\bar{p}$ 行中，最下面的 $\bar{p}-p\equiv m-\bar{m}$ 行是跳跃变量，这会导致解出的 $w_{1,t}$ 并不唯一。这意味着我们可以随意选择这 $\bar{p}-p$ 行跳跃变量的初始值。换句话说，对整个经济系统而言，当跳跃变量的数量多于爆炸特征根时，系统可能存在无数个均衡解。

5.4.3 应用 Blanchard-Kahn 算法实例

以随机 Ramsey 增长模型(5.4)为例，$z_t \equiv \tilde{A}_t, x_t^j \equiv \tilde{C}_t, x_t^s \equiv \tilde{K}_t$，

$$\begin{bmatrix} 1 & 0 \\ (1-\alpha)\cdot[1-\beta\cdot(1-\delta)] & \sigma \end{bmatrix} E \begin{bmatrix} \tilde{K}_{t+1} \\ \tilde{C}_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{\beta} & -\frac{1-\beta\cdot[1-\delta\cdot(1-\alpha)]}{\alpha\cdot\beta} \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} \tilde{K}_t \\ \tilde{C}_t \end{bmatrix} + \begin{bmatrix} \frac{1-\beta\cdot(1-\delta)}{\alpha\cdot\beta} \\ \rho\cdot[1-\beta\cdot(1-\delta)] \end{bmatrix} \tilde{A}_t$$

在 Matlab 中，首先定义系数 $\beta, \alpha, \sigma, \delta, \rho$ 的值。

```
clear;

%参数设定
beta = 0.9;
alpha = 0.75;
sigma = 1;
delta = 0.3;
rho = 0.95;
```

进而计算稳定状态。

```
%稳态值
kbar = ((1-(1-delta)*beta)/(alpha * beta))^(1/(alpha - 1));
cbar = kbar^alpha - delta * kbar;
ybar = kbar^alpha;
```

根据测算结果，$\bar{K} = 11.0766, \bar{C} = 2.7486, \bar{Y} = 6.0716$。

输入矩阵 $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2$：

```
%定义矩阵T0,T1,PSI
Tau0=zeros(2,2);
Tau1=zeros(2,2);
Psi=zeros(2,1);
Tau0(1,1)=1;
Tau0(1,2)=0;
Tau0(2,1)=(1-alpha) * (1-beta * (1-delta));
Tau0(2,2)=sigma;
Tau1(1,1)=1/beta;
Tau1(1,2)=-(1-beta * (1-delta * (1-alpha)))/(alpha * beta);
Tau1(2,1)=0;
Tau1(2,2)=sigma;
Psi(1,1)=(1-beta * (1-delta))/(alpha * beta);
Psi(2,1)=rho * (1-beta * (1-delta));
```

Matlab 测算出的矩阵值如下：

$$\mathcal{T}_0 = \begin{bmatrix} 1.0000 & 0.0000 \\ 0.0925 & 1.0000 \end{bmatrix}, \quad \mathcal{T}_1 = \begin{bmatrix} 1.1111 & -0.2481 \\ 0.0000 & 1.0000 \end{bmatrix}, \quad \Psi = \begin{bmatrix} 0.5481 \\ 0.3515 \end{bmatrix}.$$

计算 A 和 B：

```
%计算矩阵A,B
A=zeros(2,2);
B=zeros(2,2);
A = inv(Tau0) * Tau1;
B = inv(Tau0) * Psi;
```

测算结果：

$$A=\begin{bmatrix}1.1111 & -0.2481\\ -0.1028 & 1.0230\end{bmatrix},\quad B=\begin{bmatrix}0.5481\\ 0.3008\end{bmatrix}.$$

测算 A 的特征值和特征向量：

```
%对A做 Jordan decomposition,
%分解为特征值和特征向量
%对A做 Jordan decomposition,
%分解为特征值和特征向量
[VE, Lambda] = eig(A); %MU储存特征值
P=inv(VE); %P储存normalized特征向量
```

特征值和特征向量如下：

$$P=\begin{bmatrix}0.7049 & -0.8340\\ 0.4805 & 0.9806\end{bmatrix},\quad \Lambda=\begin{bmatrix}1.2327 & 0\\ 0 & 0.9014\end{bmatrix},$$

不难看出，$\lambda_i, i=(1,2)$ 分别是不稳定和稳定的特征根，满足 Blanchard-Kahn 条件，系统是鞍点稳定的。

将特征根矩阵沿着对角线元素（特征根）从低到高的顺序排列。

to be finished...

5.5　未定系数法

根据Christiano (2002)，假定经济模型以这样的状态——空间形式表现

$$\begin{aligned}&\alpha_0 Ex_{t+1}+\alpha_1\cdot x_t+\alpha_2\cdot x_{t-1}+\beta\cdot z_t=0,\quad t\geq 0,\\ &z_t=R\cdot z_{t-1}+\varepsilon_t,\end{aligned}\tag{5.13}$$

其中

- $\underset{n\times n}{x_t}$ ⇒ 在 t 时间决定的内生变量向量，x_{-1} 是提前给定的。
- $\underset{k\times 1}{z_t}$ ⇒ 外生技术冲击变量的向量，满足 $\varepsilon_t\sim i.i.d.(0,\Sigma)$。
- 系数矩阵 $\underset{n\times n}{\alpha_0}$，$\underset{n\times n}{\alpha_1}$，$\underset{n\times n}{\alpha_2}$，$\underset{n\times k}{\beta}$，$\underset{k\times k}{R}$。

与第5.4节的 Blanchard-Kahn 法相比，未定系数法不再将内生变量做状态变量和跳跃变量的区分。好处是让算法（看起来）简化，但这种简化有一定的额外成本：需要预设全部状态变量的初始值 $x_{t=0}$（与之相比，Blanchard-Kahn 算法则只需要 $x^s_{t=0}$）：如果初始时间的经济系统恰好完全处于稳定状态，这是没问题的；否则便只能针对每一个均衡方程，分别设定其对应变量的初始条件。

(5.13)描述了这样一个经济系统，系统解表现为一个反馈机制：当前内生向量 x_t 与上期 x_{t-1} 和当期外生冲击 z_t 线性相关，因此假定下式

$$x_t = \underset{n\times n}{A} \cdot x_{t-1} + \underset{n\times k}{B} \cdot z_t, \tag{5.14}$$

在静态均衡条件下我们有 $z_t \equiv 0 \quad \forall t$，这要求系数矩阵 A 的所有特征值绝对值都小于 1[5]。现在目标变成了，对于给定的初始值 x_{-1}，找到系数矩阵 A, B，使 (5.14)与(5.13)一致。

5.5.1 模型范例

第5.3节经济系统可改写为(5.13)形式，其中

$$x_t = \begin{bmatrix} \tilde{K}_{t+1} \\ \tilde{C}_{t+1} \end{bmatrix}, \quad z_t = \tilde{A}_t, \quad R = \rho,$$
$$\alpha_0 = \begin{bmatrix} 0 & 0 \\ 0 & \sigma \end{bmatrix}, \quad \alpha_1 = \begin{bmatrix} -1 & -\frac{1-\beta\cdot[1-\delta\cdot(1-\alpha)]}{\alpha\cdot\beta} \\ -(1-\alpha)\cdot[1-\beta\cdot(1-\delta)] & \delta \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} \frac{1}{\beta} & 0 \\ 0 & 0 \end{bmatrix}, \beta = \begin{bmatrix} \frac{1-\beta\cdot(1-\delta)}{\alpha\cdot\beta} \\ \rho\cdot[1-\beta\cdot(1-\delta)] \end{bmatrix}. \tag{5.15}$$

5.5.2 未定系数法求解

由(5.14)得

$$x_{t+1} = Ax_t + RBz_t = A(Ax_{t-1} + Bz_t) + RBz_t = A^2x_{t-1} + B(R+A)z_t.$$

带回(5.14)，用 x_{t-1} 替代 x_{t+1} 和 x_t

$$\underbrace{\left(\alpha_0 A^2 + \alpha_1 A + \alpha_2\right)}_{\equiv\mathcal{A}} x_{t-1} + \underbrace{\left[\alpha_0 B(R+A) + \alpha_1 B + \beta\right]}_{\equiv\mathcal{B}} z_t = 0. \tag{5.16}$$

Deterministic 状态下，$Ez_t = 0, \forall t$。(5.16)成立需要满足 $x_{t-1} = 0$ 或 $\mathcal{A} = 0$。在现实经济世界中，$x_{t-1} = 0$ 并无研究必要，因此需要满足 $\mathcal{A} = 0$。换句话说，可以通过下式求得未定系数 A 的值

$$\alpha_0 A^2 + \alpha_1 \cdot A + \alpha_2 = 0. \tag{5.17}$$

在 stochastic 状态下，(5.16)成立还需要 $\mathcal{B} = 0$。因此基于得到的 A 值，可通过下式求得 B

$$\alpha_0 B(R+A) + \alpha_1 B + \beta = 0. \tag{5.18}$$

因此，问题的关键就成了如何通过二项式(5.17)求解 A。或者更进一步：

1. 存在性：是否存在 A 的解，以及如果存在的话，有几个，见第5.5.3节。

2. 唯一性：如果存在多个解，哪一个满足静态均衡约束条件，即全部特征值的绝对值均 < 1，见第5.5.4节。

[5]回忆一下 Blanchard-Kahn 算法中将内生变量分为状态和跳跃两部分；对应相同数量的系数矩阵稳定根和不稳定根。未定系数法中所有内生变量 x_t 都是在 t 期决定的，这使得我们不再有多余的自由度用于处理不稳定根。

5.5.3　求解系数矩阵 A, B

求解系数矩阵 A

将 deterministic 状态下的(5.13)改写为如下 AR(1) 过程。已知

$$\alpha_0 x_{t+1} + \alpha_1 x_t + \alpha_2 x_{t-1} = 0,$$
$$x_t - x_t = 0,$$

改写为矩阵形式

$$\begin{aligned} &\mathcal{T}_0 Y_{t+1} + \mathcal{T}_1 Y_t = 0, \quad \forall t \geq 0, \\ &\underset{2n\times n}{Y_t} \equiv \begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix}, \quad \mathcal{T}_0 = \begin{bmatrix} \alpha_0 & 0_{n\times n} \\ 0_{n\times n} & I_n \end{bmatrix}, \quad \mathcal{T}_1 = \begin{bmatrix} \alpha_1 & \alpha_2 \\ -I_n & 0_{n\times n} \end{bmatrix}, \end{aligned} \tag{5.19}$$

其中值得注意的是，Y_0 由 n 个初始状态 x_{-1} 所决定。

做两个假定。首先假定 $\mathcal{T}_0$ 可逆，因此 α_0 也是可逆矩阵，由此(5.19)改写为

$$Y_{t+1} = -\mathcal{T}_0 \mathcal{T}_1 Y_t, \tag{5.20}$$

其次，假定 $(-\mathcal{T}_0\mathcal{T}_1)$ 有 $2n$ 个线性不相关的特征向量。

基于这两个假定，我们可以采取特征值——特征向量分解方法：

$$\underset{2n\times 2n}{-\mathcal{T}_0\mathcal{T}_1} = \underset{(2n\times 2n)}{P} \underset{(2n\times 2n)}{\Lambda} \underset{(2n\times 2n)}{P^{-1}}, \tag{5.21}$$

其中对角矩阵 Λ 的对角元素为 $-\mathcal{T}_0\mathcal{T}_1$ 的特征值，P 是对应的特征向量。

假设 $-\mathcal{T}_0\mathcal{T}_1$ 有 $\bar{n}$ 个稳定特征值，构成分块对角矩阵 Λ_s，余下的 $2n-\bar{n}$ 个不稳定特征值构成分块对角矩阵 Λ_e。重新排列 Λ：

$$\underset{2n\times 2n}{\Lambda} = \begin{bmatrix} \underset{\bar{n}\times\bar{n}}{\Lambda_s} & \underset{\bar{n}\times(2n-\bar{n})}{0} \\ \underset{(2n-\bar{n})\times\bar{n}}{0} & \underset{(2n-\bar{n})\times(2n-\bar{n})}{\Lambda e} \end{bmatrix}.$$

对于(5.20)- (5.21)，定义 $W_t \equiv P^{-1}Y_t$，我们有

$$P^{-1}Y_t = \Lambda P^{-1} Y_{t-1} \Leftrightarrow W_t = \Lambda W_{t-1}. \tag{5.22}$$

对(5.22)做 backward-looking 迭代

$$\begin{aligned} \underset{2n\times n}{W_t} &= \Lambda W_{t-1} = \Lambda(\Lambda W_{t-2}) = \ldots = \Lambda^t W_{t_0} \\ &= \begin{bmatrix} \Lambda_s^t & 0 \\ 0 & \Lambda_e^t \end{bmatrix} W_{t_0} = \begin{bmatrix} \underset{\bar{n}\times\bar{n}}{\Lambda_s^t} & 0 \\ 0 & \underset{(2n-\bar{n})\times(2n-\bar{n})}{\Lambda_e^t} \end{bmatrix} \begin{bmatrix} \underset{\bar{n}\times n}{W_{1,t_0}} \\ \underset{(2n-\bar{n})\times n}{W_{2,t_0}} \end{bmatrix}. \end{aligned} \tag{5.23}$$

根据定义，$\{x_t\}_{t=0}^{\infty}$ 是平稳过程 $\Rightarrow$ Y_t 是 x_t 的线性方程，$\{Y_t\}_{t=0}^{\infty}$ 是平稳过程 $\Rightarrow$ W_t 是 Y_t 的线性方程，$\{W_t\}_{t=0}^{\infty}$ 是平稳过程。由于 $\lim_{t\to\infty}\Lambda_s^t \approx 0$，$W_t$ 的前 $\bar{n}$ 行一定是平稳的，初始 W_{1,t_0} 可以取任意值；由于 $\lim_{t\to\infty}\Lambda_s^t \approx \infty$，$W_t$ 的后 $2n-\bar{n}$ 行是不平稳的，式(5.23)成立便要求初始 $W_{2,t_0} = 0$。

如果假定 $\bar{n} \equiv n$，即(5.19)中系数矩阵的稳定特征值的数量等于经济系统(5.13)中内生变量的数量 (对于 $\bar{n} \neq n$ 情况的讨论见第5.5.4节。)，我们有

$$\begin{bmatrix} \underset{n\times n}{W_{1,t_0}} \\ \underset{n\times n}{W_{2,t_0}} \end{bmatrix} = \begin{bmatrix} \underset{n\times n}{(P^{-1})_{11}} & \underset{n\times n}{(P^{-1})_{12}} \\ \underset{n\times n}{(P^{-1})_{21}} & \underset{n\times n}{(P^{-1})_{22}} \end{bmatrix} \begin{bmatrix} \underset{n\times n}{x_0} \\ \underset{n\times n}{x_{-1}} \end{bmatrix},$$

其中要求 $W_{2,t_0}=0$，即

$$\left(P^{-1}\right)_{21} x_0+\left(P^{-1}\right)_{22} x_{-1}=0.$$

对于可逆矩阵 $\left(P^{-1}\right)_{21}$，上式改写为

$$x_0=-\left(\left(P^{-1}\right)_{21}\right)^{-1}\left(P^{-1}\right)_{22} x_{-1}, \tag{5.24}$$

结合 deterministic 状态的(5.14)与(5.24)，可得系数矩阵 A

$$A=-\left(\left(P^{-1}\right)_{21}\right)^{-1}\left(P^{-1}\right)_{22}. \tag{5.25}$$

求解系数矩阵 B

将求得的系数矩阵 A 代入(5.18)。对于 $n\times k$ 的矩阵 B 和 β，等式两侧向量化，我们有

$$\begin{aligned}
0&=\operatorname{vec}\left(\left(\alpha_0 A+\alpha_1\right) B+\alpha_0 B R\right)+\operatorname{vec}(\beta)\\
&=\operatorname{vec}\left(\left(\alpha_0 A+\alpha_1\right) B\right)+\operatorname{vec}\left(\alpha_0 B R\right)+\operatorname{vec}(\beta)\\
&=\left[I_k \otimes\left(\alpha_0 A+\alpha_1\right)\right] \operatorname{vec}(B)+\left(R^T \alpha_0\right) \operatorname{vec}(B)+\operatorname{vec}(\beta)\\
&=\left[I_k \otimes\left(\alpha_0 A+\alpha_1\right)+R^T \alpha_0\right] \operatorname{vec}(B)+\operatorname{vec}(\beta),
\end{aligned}$$

其中 vec 表示对矩阵向量化；$\otimes$ 表示 Kronecker 乘[6]。进而我们有

$$\operatorname{vec}(B)=-\left[I_k \otimes\left(\alpha_0 A+\alpha_1\right)+R' \otimes \alpha_0\right]^{-1} \operatorname{vec}(\beta), \tag{5.26}$$

5.5.4 存在性及唯一性的探讨

如前文所述，$\bar{n}=n$ 时，经济系统存在唯一均衡解。我们需要做的是通过对初始状态 x_{t_0} 的选择，使得在 $W_{t_0}=\begin{bmatrix} \underset{\bar{n}\times n}{W_{1,t_0}^T}, & \underset{(2n-\bar{n})\times n}{W_{2,t_0}^T} \end{bmatrix}^T$ 当中 (对应 W_{2,t_0} 的) 后 $2n-\bar{n}$ 行等于 0。

当 $\bar{n}<n$ 时，$2n-\bar{n}>n$，我们需要使后 $2n-\bar{n}$ 行等于 0，但我们只有 n 个自由变量。此时均衡借不存在。

当 $\bar{n}>n$ 时，$2n-\bar{n}<n$，我们需要使后 $2n-\bar{n}$ 行等于 0，对应 n 个自由变量。这导致存在很多组 x_0 的值可以带来均衡，均衡解不唯一。

5.6 期望误差法

根据Sims (2002)，假定经济模型以这样的状态——空间形式展现

$$\underset{(n\times n)}{\mathcal{T}_0}\underset{(n\times 1)}{x_t}=\underset{(n\times n)}{\mathcal{T}_1}x_{t-1}+\underset{(n\times k)}{\Psi}\underset{(k\times 1)}{u_t}+\underset{(n\times r)}{\Pi}\underset{(r\times 1)}{\eta_t}, \quad t\geq 0, \tag{5.27}$$

其中

[6]对于矩阵 $A\in\mathbb{R}^{k\times l}, B\in\mathbb{R}^{l\times m}, C\in\mathbb{R}^{m\times n}$，我们有矩阵向量化

$$\operatorname{vec}(A)=\left[A_{1,1},\ldots Ak,1,A_{1,2},\ldots A_{k,2},\ldots A_{1,l},\ldots A_{k,l}\right]^T,$$

以及以下向量化运算

$$\begin{aligned}
\operatorname{vec}(A+B)&=\operatorname{vec}(A)+\operatorname{vec}(B),\\
\operatorname{vec}(AB)&=\left(I_m \otimes A\right) \operatorname{vec}(B)=\left(B^T \otimes I_k\right) \operatorname{vec}(A),\\
\operatorname{vec}(ABC)&=\left(I_n \otimes AB\right) \operatorname{vec}(C)=\left(C^T \otimes A\right) \operatorname{vec}(B)=\left(C^T B^T \otimes I_k\right) \operatorname{vec}(A).
\end{aligned}$$

- 向量 x_t 表示内生变量，对应系数矩阵 $\mathcal{T}_0, \mathcal{T}_1$；$x_{-1}$ 是给定的。在期望误差法中，x_t 包含一部分在 t 期对 $t+1$ 期的期望值。
- u_t 表示外生随机冲击过程，假定 $u_t \sim i.i.d.(0, \Sigma)$，对应系数矩阵 Ψ，
- $\eta_t = x_t - E_{t-1}x_t$ 是期望误差向量，反映 t 期实际状态 x_t 与 $t-1$ 期对 t 期状态的期望的偏差，满足 $E_t\eta_{t+1} = 0$，对应系数矩阵 Π。

5.6.1　模型范例

第5.3节经济系统可改写为(5.27)形式，其中

$$
\begin{aligned}
x_t &= \begin{bmatrix} E_t\tilde{K}_{t+1} \\ \tilde{C}_t \\ E_t\tilde{C}_{t+1} \\ \tilde{A}_t \end{bmatrix}, \quad u_t = \varepsilon_t, \quad \eta_t = \tilde{C}_t - E_{t-1}\tilde{C}_t, \\
\mathcal{T}_0 &= \begin{bmatrix} 1 & \frac{1-\beta\cdot[1-\delta\cdot(1-\alpha)]}{\alpha\cdot\beta} & 0 & -\frac{1-\beta\cdot(1-\delta)}{\alpha\cdot\beta} \\ (\alpha-1)\cdot[1-\beta\cdot(1-\delta)] & \delta & -\delta & \rho\cdot[1-\beta\cdot(1-\delta)] \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \\
\mathcal{T}_1 &= \begin{bmatrix} \frac{1}{\beta} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \Psi = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \Pi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.
\end{aligned} \tag{5.28}
$$

5.6.2　期望误差法求解

假定 $\mathcal{T}_0$ 是可逆矩阵，(5.27)可以改写为

$$
\begin{aligned}
x_t &= \mathcal{T}_0^{-1}\mathcal{T}_1 x_{t-1} + \mathcal{T}_0^{-1}\left(\Psi u_t + \Pi\eta_t\right) \\
&= Ax_{t-1} + \mathcal{T}_0^{-1}\left(\Psi u_t + \Pi\eta_t\right), \\
&\text{其中} \underset{n\times n}{A} \equiv \mathcal{T}_0^{-1}\mathcal{T}_1, x_0\text{是给定的，并且}t \geq 1.
\end{aligned} \tag{5.29}
$$

进一步假定 A 的所有特征向量都是线性不相关的，由此我们可以对 A 做特征值——特征向量分解

$$
A = \underset{n\times n}{P}\Lambda P^{-1},
$$

代回(5.29)，调整得

$$
\begin{aligned}
&P^{-1}x_t = P^{-1}\left(P\Lambda P^{-1}\right)x_{-1} + P^{-1}\mathcal{T}_0^{-1}\left(\Psi u_t + \Pi\eta_t\right), \text{进而} \\
&w_t = \Lambda w_{t-1} = Q\left(\Psi u_t + \Pi\eta_t\right), \text{其中} \\
&\underset{n\times n}{w_t} \equiv P^{-1}x_t, \quad \underset{n\times n}{Q} \equiv P^{-1}\mathcal{T}_0^{-1}.
\end{aligned} \tag{5.30}
$$

类似地，将对角矩阵 Λ 按特征值从小到大顺序重新排列

$$
\underset{n\times n}{\Lambda} = \begin{bmatrix} \underset{\bar{n}\times\bar{n}}{\Lambda_s} & \underset{\bar{n}\times(n-\bar{n})}{0} \\ \underset{(n-\bar{n})\times\bar{n}}{0} & \underset{(n-\bar{n})\times(n-\bar{n})}{\Lambda_e} \end{bmatrix}, \tag{5.31}
$$

其中 Λ_s 为绝对值小于 1 的特征值，设为 $\bar{n} < n$ 个。

根据(5.31)，经济系统(5.30)可以改写为

$$\begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} = \begin{bmatrix} \underset{\bar{n}\times\bar{n}}{\Lambda_s} & \underset{\bar{n}\times(n-\bar{n})}{0} \\ \underset{(n-\bar{n})\times\bar{n}}{0} & \underset{(n-\bar{n})\times(n-\bar{n})}{\Lambda_e} \end{bmatrix} \begin{bmatrix} w_{1,t-1} \\ w_{2,t-1} \end{bmatrix} + \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} (\Psi u_t + \Pi\eta_t), \quad t \geq 1. \tag{5.32}$$

上半部分为稳定分块，下半部分为不稳定分块。我们先从不稳定分块开始求解。

不稳定分块求解

提取不稳定分块

$$w_{2,t} = \Lambda_e w_{2,t-1} + Q_2 (\Psi u_t + \Pi\eta_t),$$

调整为前向形式

$$w_{2,t} = \Lambda_e^{-1} w_{2,t+1} - \Lambda_e^{-1} Q_2 (\Psi u_{t+1} + \Pi\eta_{t+1}), \tag{5.33}$$

进一步调整为前向迭代形式

$$w_{2,t} = \underbrace{\lim_{T\to\infty} \Lambda_e^{-T} w_{2,t+1}}_{\mathcal{A}} - \underbrace{\sum_{s=1}^{T} \Lambda_e^{-s} Q_2 (\Psi u_{t+s} + \Pi\eta_{t+s})}_{\mathcal{B}}, \tag{5.34}$$

其中等式右侧

- $\mathcal{A} = 0$。这是由于首先分块矩阵 Λ_e 对应的所有元素，即特征值的绝对值都大于 1，$\Lambda_e^{\infty} \to 0$，其次静态均衡 $E_t(w_{2,t+\infty})$ 是有界的。
- $\mathcal{B} = 0$。这是由于首先 u_t 是个均值为 0 的随机过程，满足 $E(u_t) = 0$，其次在理性期望条件下，η_t 的条件均值为 0。

基于上述分析，我们有不稳定分块的值

$$w_{2,t} = 0. \tag{5.35}$$

稳定分块求解

提取稳定分块

$$\underset{\bar{n}\times n}{w_{1,t}} = \underset{\bar{n}\times\bar{n}}{\Lambda_s} w_{1,t-1} + \underset{\bar{n}\times n}{Q_1} \left(\underset{(\bar{n}\times k)}{\Psi} \underset{(k\times 1)}{u_t} + \underset{(n\times r)}{\Pi} \underset{(r\times 1)}{\eta_t} \right).$$

为了得到 $w_{1,t}$ 的值，我们首先需要替代期望误差 η_t，随后求解稳定分块。

期望误差求解

假定 $k = r$(对于 $k \neq r$ 的讨论，见第5.6.3节。)。基于(5.34)-(5.35)我们有

$$Q_2 (\Psi u_t + \Pi\eta_t) = 0, \tag{5.36}$$

这意味着期望误差 η_t 随着同期外部冲击 u_t 的变化而反向变化，所谓“理性预期”。如果假定 $Q_2\Pi$ 是可逆矩阵，那么可以将 η_t 写为关于 (并且只是关于)u_t 的函数

$$\eta_t = -(Q_2\Pi)^{-1} Q_2 \Psi u_t. \tag{5.37}$$

稳定分块求解

η_t 的决定式(5.37)代回稳定分块决定式可得

$$\begin{aligned} w_{1,t} &= \Lambda_s w_{1,t-1} + Q_1 \left(\Psi u_t + \Pi \eta_t\right) \\ &= \Lambda_s w_{1,t-1} + Q_1 \left(\Psi - \Pi \left(Q_2 \Pi\right)^{-1} Q_2 \Psi\right) u_t. \end{aligned} \tag{5.38}$$

内生变量向量 x_t 求解

联立两个分块 $w_{1,t}, w_{2,t}$ 的解(5.38)与(5.35)，经济系统(5.32)可以进一步改写为

$$\begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} = \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_{1,t-1} \\ w_{2,t-1} \end{bmatrix} + \begin{bmatrix} Q_1 \left(\Psi - \Pi \left(Q_2 \Pi\right)^{-1} Q_2 \Psi\right) \\ 0 \end{bmatrix} u_t.$$

根据定义式(5.30)用 x_t 替代上式中的 w_t，可得 x_t 的解

$$x_t = P \begin{bmatrix} \Lambda_s & 0 \\ 0 & 0 \end{bmatrix} P^{-1} x_{t-1} + P \begin{bmatrix} Q_1 \left(\Psi - \Pi \left(Q_2 \Pi\right)^{-1} Q_2 \Psi\right) \\ 0 \end{bmatrix} u_t. \tag{5.39}$$

5.6.3　存在性及唯一性的探讨

如前文所述，经济系统存在均衡解的条件由(5.36)给出：根据该方程，期望误差 η_t 根据出现的外生冲击 u_t 而灵活调整。如果经济系统中对 η_t 的限定条件太多——导致 η_t 无法灵活调整，均衡解可能不存在。

如果 $r < k$，即对于 η_t 向 u_t 的调整存在过多的限定条件，这可能会使得系统解不存在。换句话说，经济系统有解的充分必要条件是：$r \geq k$，即 $Q_2\Psi$ 的列空间包含在 $Q_2\Pi$ 的列空间内。

如果 $r > k$，即对于 η_t 向 u_t 的调整存在过多的限定条件，这可能会使得存在多个 η_t 同时满足式(5.37)，对应的只是唯一的 $Q_2\Pi\eta_t$，而非唯一的 η_t。多重 η_t 值使得我们有多个 $Q_1\Pi\eta_t$，对应(5.38)中多个稳定分块矩阵 w_t(进而 x_t) 的解。为了使系统存在唯一解，需要使 $Q_1\Pi$ 的行空间包含在 $Q_2\Pi$ 内，即存在这样一个矩阵 Φ 满足 $Q_1\Pi = \Phi_2\Pi$。

5.7　参数化期望法

参数化期望法 (PEA) 最早由Den Haan and Marcet (1990) 提出。PEA 所依赖的渐进式趋同结果，由Marcet and Sargent (1989); Marcet and Marshall (1994) 等人所讨论。算法方面，Den Haan and Marcet (1994) 提出精确度测试；Christiano and Fisher (2000) 等人讨论了网格算法。基于Den Haan and Marcet (1990)，本节介绍 PEA 算法的基本思路和一个简单范例。

大多数动态经济学模型建立在一个或一组 Euler 方程的基础上。这个 (些)Euler 方程将一组当期变量和一组未来期变量的条件期望联系在一起，如

$$f(x_t) = E_t h(x_{t+1}, u_{t+1}, z_t), \tag{5.40}$$

其中 $\{x_t, u_t, z_t\}$ 分别代表当前 t 期的内生变量，状态变量和随机扰动。E_t 表示在当前 t 期，对未来时间期比如 $t+1$ 状态的条件期望。对于经验模型来说，研究的重点在于选取合适的显函数形式来描述 $f()$ 以及 $h()$，对于非线性系统来说，难点在于如何对之作线性近似，以求得系统的解。

一个常见的求解思路是后向求解法。简单来说，根据这个思路，如果(5.40)可以改写为

$$x_t = f^{-1}\left(E_t h(x_{t+1}, u_{t+1}, z_t)\right),$$

那么在我们知道了方程 f^{-1} 近似形式的情况下，一旦我们得到未来期的条件期望 $E_t h(x_{t+1}, u_{t+1}, z_t)$，便可以后向求解当前期的 x_t。换句话说，一旦我们知道了 x_{t+1} 的值，便可以后向求得 x_t。

与之相反，PEA 法的求解思路是前向的，致力于用已知的 x_t 求解 x_{t+1}：即便 $E_t h(x_{t+1}, u_{t+1}, z_t)$ 是一个关于 x_{t+1} 的函数，可是从定义来看，它仍然是由经济个体在 t 时刻的决策 x_t 所决定的。因此，PEA 用前向递归的方式，基于当前期已知信息，求解未来期的经济系统，将(5.40)改写为

$$E_t h(x_{t+1}, u_{t+1}, z_t) = m(\beta, Z_t^1) \tag{5.41}$$

其中 Z_t^1 是 t 期状态变量的集合。换句话说，PEA 假定未来期的期望值是关于当前期已知信息集的函数，当前期已知信息集包括前定变量和当期冲击等。根据这一假定，对未来的期望得以“参数化”。

PEA 法成功与否的关键有两点：

- 能提供多大的精确程度，基于当期信息集，用 $m()$ 来近似参数化期望 $E_t h()$。⇒ Den Haan and Marcet (1990) 假定 $m()$ 是多项式形式，多项式 $m()$ 近似 $E_t h()$ 的精确度?
- 参数向量 β 的精确度。提高多项式的非线性程度有助于提升近似精读，然而进一步的精确度提升更离不开参数向量 β 的估计。

5.7.1 参数化期望法示例

假定一个典型的随机增长模型，经济个体需要做效用最大化决策

$$\max_{c_t} E_t \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma},$$

约束条件

$$c_t + k_t - \mu k_{t-1} = A_t k_{t-1}^{\alpha}, \tag{5.42}$$

其中

$$\ln A_{t+1} = \rho \ln A_t + \varepsilon_t$$

表示随机技术冲击，β 表示时间贴现，$1-\mu$ 表示折旧率，σ 表示风险规避系数。FOC ⇒

$$MU_{c_t} = \beta E_t R_{t+1} MU_{c_{t+1}},$$

其中 MU_{c_t} 表示 t 期消费带来的边际效用。上式也可表示为如下 Euler 方程形式

$$c_t^{-\sigma} = \beta E_t \left[c_{t+1}^{-\sigma} \left(\alpha A_{t+1} k_t^{\alpha-1} + \mu \right) \right]. \tag{5.43}$$

根据 PEA 法，将右侧的期望值改写为一个与 t 期状态变量有关的多项式方程 $m()$，状态变量包括 $t-1$ 期前定变量 k_{t-1} 和当期技术冲击 A_t，上式因而改写为

$$c_t^{-\sigma} = \beta m(k_{t-1}, A_t, \delta), \tag{5.44}$$

对应系数变量 δ。

假定多项式 $m(\cdot)$ 以一阶显函数形式表现如下[7]

$$m(k_{t-1}, A_t; \delta) = \delta_1 k_{t-1}^{\delta_2} A_t^{\delta_3}, \tag{5.45}$$

[7]也可以用更高阶显函数的形式设定 $m()$，如

$$m(k_{t-1}, A_t; \delta) = \delta_1 k_{t-1}^{\delta_2} A_t^{\delta_3} \left(k_{t-1} A_t\right)^{\delta_4} \left(k_{t-1}^2\right)^{\delta_t} \left(A_t^2\right)^{\delta_6}.$$

随后，为向量 δ 赋一组初始值，基于初始值，PEA 进行迭代，直到逼近 “真实的” 系数值位置。

在给定以下三组信息

- 初始资本存量 $k_{t=0}$，
- 在利用计算机生成的一组随机游走时间序列中挑选 $A_{t=1}$，
- 初始赋值 $\delta_1, \delta_2, \delta_3$

之后，我们可以根据(5.44)求得 $c_{t=1}$。下一步，利用预算约束式(5.42)可生成 $k_{t=1}$。自此我们有了经济系统在第 1 期 $t=1$ 的全部信息 $\{c_{t=1}, k_{t=1}, A_{t=1}\}$。

对于 $t=2$，重复上述过程

- 在利用计算机生成的同一组随机游走时间序列中挑选 $A_{t=2}$，
- 根据(5.44)求得 $c_{t=2}$，
- 根据(5.42)求得 $k_{t=2}$，

以此类推，直到第 $t=n$ 期。

对于给定的 $\{\delta\}$ 和技术冲击时间序列 A_t，利用 PEA 方法生成的 $\{k_t, c_t, y_t\}_{t=1}^n$，并不总是与 Euler 描述的经济个体期望相一致。个体基于当期信息对未来期作出期望，并采取行动如(5.44)以致力于实现这种期望，但实际导致的结果往往与当初的 Euler 方程期望(5.43)相悖。

随后的工作就变为，选取 “正确” 的系数 $\{\delta\}$，使得

$$E_t\left[c_{t+1}^{-\sigma}\left(\alpha A_{t+1}k_t^{\alpha}+u\right)\right]-\delta_1 k_{t-1}^{\delta_2}A_t^{\delta_3}=\nu_t\approx 0, \tag{5.46}$$

参数选取的标准：选取合适的 $\{\delta\}$ 组合，使得生成的经济时间序列数据中，对应 $\min\sum\nu_t^2$。这就涉及到非线性最小二乘法 (NLS)：首先猜测一组 $\{\delta_1, \delta_2, \delta_3\}$，利用其生成一组 $\{c_t, k_t, y_t\}$，运行 NLS 回归测算 $\min\sum\nu_t^2$ 对应的 $\{\delta_1', \delta_2', \delta_3'\}$。$\delta$ 和 δ' 的差异反映了基于对未来期期望的行为，在未来期未必会产生如当初所设想的结果。这就需要修正我们的预测模型，主要通过调整系数 δ 来实现：将 $\{\delta_1', \delta_2', \delta_3'\}$ 替代原有的 $\{\delta_1, \delta_2, \delta_3\}$，生成新的一组时间序列数据，NLS 测算 $\min\sum\nu_t^2$ 对应的 $\{\delta_1'', \delta_2'', \delta_3''\}$，重复上述过程，直到 $\sum\nu_t^2\approx 0$，此时 δ 达到不动点，从经济意义上来说，消费者通过预测方程产生当期消费需求，当期消费需求又引出未来期消费需求，这一组消费需求都证明了最初的预测方程是正确的。

可以从学习算法 (learning algorithm) 的角度来理解 PEA。在初始阶段，经济系统中的行为人持有有限信息。他知道一阶条件 (Euler 方程) 的形式，但不知道如何合理展开当期行动以符合这种期望。行为人只有通过试错法，形成、并逐次调整自己的行动，这构成一个学习过程：起初他只掌握少量当前状态变量的信息，只能凭直觉猜测作出行动，并对未来做初步的预测。随着信息掌握的越来越丰富，他得意不断修正作出的直觉猜测，使得预测越来越精确，可以更多地凭经验行事，更少地依赖直觉（直觉和经验作为互补品），直到对未来的预测 $E_t\left[c_{t+1}^{-\sigma}\left(\alpha A_{t+1}k_t^{\alpha}+u\right)\right]$ 与实际情况 $c_{t+1}^{-\sigma}\left(\alpha A_{t+1}k_t^{\alpha}+u\right)$ 相吻合。

5.8　Schur 分解法

上文提到了一系列求解非线性系统的算法，如特征值分解法、未定系数法、期望误差法等。它们的展开都依赖于这样的假设条件：$\mathcal{T}_0$ 是非奇异矩阵 (可逆的)，进而 $\mathcal{T}_1$ 是可对角化的矩阵。该假设有助于将经济系统转化为更易于处理的标量方程组的形式。

然而对于更为一般的形式，$\mathcal{T}_0$ 是奇异矩阵的情况，我们无法将经济系统对角化。此时就需要介绍 Schur 分解法 (又称 QZ 分解法，见第6.A节。)，将经济系统三角化。本节简要介绍 Schur 分解法在求解非线性经济系统中的应用。

5.8.1 Schur 分解的定义

对于这样一个非线性经济系统

$$\mathcal{T}_0 x_{t+1} = \mathcal{T}_1 x_t, \quad t \geq 0, \quad x_t \in \mathbb{R}^{n\times n}, \tag{5.47}$$

其中 $x_{t=0}$ 为给定的初始状态。假定存在

- 两个酉矩阵 Q, Z (unitary matrix，即满足 $QQ^T = ZZ^T = I$)，和
- 两个上三角矩阵 Ω, Λ，分别对应对角元素 $\omega_{ii}, \lambda_{ii}$，

使得满足

$$\mathcal{T}_0 = Q^T \Omega Z, \quad \mathcal{T}_1 = Q^T \Lambda Z, \tag{5.48}$$

我们称这种方法为 Schur 分解法，其中 Ω, Λ 分别又称为系数矩阵 $\mathcal{T}_0, \mathcal{T}_1$ 的 Schur 式 (Schur form)[8]。

这样一个非线性经济系统的的广义特征根解集 (generalized eigenvalue set) 为 $\left(\frac{\lambda_{ii}}{\omega_{ii}}\right)$，分子为系数矩阵 $\mathcal{T}_1$ 的特征根三角矩阵 Λ 的对角元素，分母为系数矩阵 $\mathcal{T}_0$ 的特征根三角矩阵 Ω 的对角元素；两组对角元素的绝对值均经过从左上到右下升序重新排列。

根据 Schur 分解法，我们可以利用系数矩阵 $\mathcal{T}_0, \mathcal{T}_1$ 来测算 Q, Z, Ω, Λ，并进一步求得经济系统的解。(5.47)左乘 Q，整理得

$$\Omega Z^T x_{t+1} = \Lambda Z^T x_t \Rightarrow \Omega w_{t+1} = \Lambda_t w_t, \quad t \geq 0, \tag{5.49}$$

其中定义了变量

$$w_t \equiv Z^T x_t, \tag{5.50}$$

可见(5.49)是一个关于 $\{w\}_t$ 的 (三角矩阵) 系统。新系统的特征根对角矩阵中，对角元素的特征根的绝对值 $\left|\frac{\lambda_{ii}}{\omega_{ii}}\right|$ 从左上到右下升序排列。据此可将新系统分为两部分：1 个稳定分块矩阵 $w_{1,t}$，对应 $\left|\frac{\lambda_{ii}}{\omega_{ii}}\right| < 1$，和 1 个爆炸分块矩阵 $w_{2,t}$，对应 $\left|\frac{\lambda_{ii}}{\omega_{ii}}\right| > 1$，(5.49)因此改写为

$$\begin{bmatrix} \Omega_{11} & \Omega_{12} \\ 0 & \Omega_{22} \end{bmatrix} \begin{bmatrix} w_{1,t+1} \\ w_{2,t+1} \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ 0 & \Lambda_{22} \end{bmatrix} \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix}. \tag{5.51}$$

5.8.2 爆炸分块矩阵的解

(5.51)中的爆炸分块可以表示为

$$\Omega_{22} w_{2,t+1} = \Lambda_{22} w_{2,t},$$

根据模型设定，上三角矩阵 Λ 的对角分块 Λ_{22} 是非奇异矩阵，因此我们有

$$w_{2,t} = \Lambda_{22}^{-1} \Omega_{22} w_{2,t+1},$$

进一步做前向迭代有

$$w_{2,t} = \lim_{T\to\infty} \left(\Lambda_{22}^{-1}\Omega_{22}\right)^T w_{2,t+T}. \tag{5.52}$$

[8]考虑到现实经济世界的特性，我们暂不考虑复根的情况。

(5.52)中，一方面系数矩阵 $\left(\Lambda_{22}^{-1}\Omega_{22}\right)$ 的特征根矩阵，对角元素 $\left(\frac{\omega_{ii}}{\lambda_{ii}}\right)$ 对应原系统广义特征根 $\left(\frac{\lambda_{ii}}{\omega_{ii}}\right)$ 的倒数，在爆炸分块中，$\left|\frac{\omega_{ii}}{\lambda_{ii}}\right| < 1$，因此 $\left(\frac{\omega_{ii}}{\lambda_{ii}}\right)^T \to 0$。另一方面，经济系统的稳定性要求 $\lim_{T\to\infty}\left(w_{w,t+T}\right) < \infty$。由此我们有，在稳定均衡条件下，(5.52)成立需要满足条件

$$w_{2,t} = 0. \tag{5.53}$$

5.8.3　稳定分块矩阵的解

(5.51)中的稳定分块

$$\Omega_{11}w_{1,t+1} + \Omega_{12}w_{2,t+1} = \Lambda_{11}w_{1,t} + \Lambda_{12}w_{2,t},$$

引入爆炸分块的解(5.53)得

$$\Omega_{11}w_{1,t+1} = \Lambda_{11}w_{1,t},$$

类似地，根据模型设定，上三角矩阵 Ω 的对角分块 Ω_{11} 是非奇异矩阵，因此我们有

$$w_{1,t} = \Omega_{11}^{-1}\Lambda_{11}w_{1,t-1},$$

进一步做后向迭代得

$$w_{1,t} = \left(\Omega_{11}^{-1}\Lambda_{11}\right)^t w_{1,t=0}, \tag{5.54}$$

其中 $w_{1,t=0}$ 由已知的初始状态 $x_{t=0}$ 得到。

5.8.4　原经济系统的解

合并稳定分块和爆炸分块的解(5.54)(5.53)，可得新系统 w_t 的解。在此基础上，考虑到酉矩阵 Z 的可逆性，根据(5.50)可进一步测得原经济系统 x_t 的解：

$$x_t = Zw_t = Z\begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} = Z\begin{bmatrix} \left(\Omega_{11}^{-1}\Lambda_{11}\right)^t w_{1,t=0} \\ 0 \end{bmatrix}. \tag{5.55}$$

第六章　扰动法

6.1　扰动法简介

对于 DSGE 模型的非线性系统，另一种求解方法是扰动法 (perturbation)，即根据隐函数定理 (implicit function theorem)，围绕经济系统的非随机稳定状态做 TSE(泰勒级数展开 Taylor Series Expansion)[1]。扰动法在经济学研究中的应用[2]，可以追溯至Judd and Guu (1993)[3]，从那时起至今二十多年间，扰动法在宏观经济学研究中受到越来越多的重视，主要原因有四。

1. 解的精确度。一方面在近似点附近，扰动法的局部解的精确度常常是比较高的，另一方面扰动法生成的状态变量值也具有较好的全局解特性 (Judd, 1998; Judd and Guu, 2001; Caldara et al., 2012)，并且扰动法的解还可以较方便地交由其他求解方法做进一步处理，如价值方程的迭代，见第6.6节。

2. 解的经济意义。扰动法的近似解相对直观易懂，例如 DSGE 模型的二阶近似解中含有一个外生冲击的冲击标准差项，会对经济系统的动态路径产生影响，这呼吁经济政策制定者采取有效的预防措施。而相反地，持有确定性等价 (certainty equivalence) 假定的一阶线性近似法 (第五章)，由于无法将外生技术冲击的波动纳入到分析框架中，从而在分析福利、风险等问题上存在不足。

3. 如前所述，传统线性求解技术基本上类似于一阶线性扰动法，这这些前期研究的宝贵经验能够为我们研究高阶扰动求解法提供帮助。

4. 一批新的计算程序有助于显著提高高阶扰动法数值计算的效率，如Matlab平台下的Dynare (以及一个独立软件平台Dynare++)，Mathematica平台下的Perturbation AIM等。

6.2　分析框架

扰动法致力于近似求解状态——空间系统 $\mathcal{H}(d) = \mathbf{0}$，其中 $d : \Omega \to \mathcal{R}^m$ 表示由一系列均衡条件和预算约束条件组成的方程系统，对应 n 个状态变量组成的向量 $\boldsymbol{x}$ 以及相关系数 θ。求解的基本思路是对围绕状态变量的非随机稳定点做泰勒级数展开。以二阶泰勒级数展开为例：

$$d_i^2(\boldsymbol{x}, \theta) = \theta_{i,0} + \theta_{i,1}(\boldsymbol{x} - \boldsymbol{x}_0)^\top + (\boldsymbol{x} - \boldsymbol{x}_0)\theta_{i,2}(\boldsymbol{x} - \boldsymbol{x}_0)^\top, \quad i = 1, 2, \ldots m, \tag{6.1}$$

其中

[1]需要指出的是，扰动法的算法有多种，这里以宏观经济学研究中得到广泛采用的方法为例。

[2]扰动法在物理学和自然科学中得到广泛应用，至少可以上溯到 19 世纪。随着 20 世纪上半叶量子力学的发展，扰动法更成为自然科学研究的核心方法之一，相关介绍可见Simmonds and Mann (2013); Bender and Orszag (2013)。

[3]关于采用扰动法进行经济学研究的严格数学证明，如可解性等，以及更高级的求解技术如 Padé 近似等，可见Judd (1998); Judd and Guu (2001); Jin and Judd (2002)。

- $\boldsymbol{x}^\top$ 表示向量 $\boldsymbol{x}$ 的转置，
- $\boldsymbol{x}_0$ 表示非随机稳态，
- $\theta_{i,0}$ 是一个矢量，
- $\theta_{i,1}$ 是一个 n 维向量，
- $\theta_{i,2}$ 是一个 $n \times n$ 矩阵，
- 系数 $\theta_{i,0}, \theta_{i,1}, \theta_{i,2}$ 可根据隐函数定理，由 d 的导数求得。

与传统线性近似法相比 ((King et al., 1988a,b, 2002)，本文第五章)，线性近似与一阶扰动法基本等同。而更高阶扰动法是将一阶扰动 (近似线性) 化的经济系统结构一般化到更一般的形式，引入额外项，从而使得近似式具有更高的解释效力。

Remark 6.1 (扰动法的线性化和对数线性化)**.** *通常来说，线性化是指对状态变量 (或对状态变量针对模型中的某些量做预处理后) 进行线性调整。类似地，对数线性化描述状态变量距离其稳定状态的对数偏离程度，以某一变量 $x \in \boldsymbol{x}$ 举例，定义 $\hat{x} = \log(x/\bar{x})$，其中 $\bar{x}$ 是稳态值。则二阶近似*

$$d_i^2(\hat{\boldsymbol{x}}, \theta) = \theta_{i,0} + \theta_{i,1}(\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}_0)^\top + (\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}_0)\theta_{i,2}(\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}_0)^\top, \quad i = 1, 2, \ldots m,$$

如果 x_0 是确定性稳态，我们有 $\hat{\boldsymbol{x}}_0 = 0$，则对于 $\forall x \in \boldsymbol{x}$ 上式可改写为

$$d_i^2(\hat{\boldsymbol{x}}, \theta) = \theta_{i,0} + \theta_{i,1}\hat{\boldsymbol{x}}^\top + \hat{\boldsymbol{x}}\theta_{i,2}\hat{\boldsymbol{x}}_0^\top, \quad i = 1, 2, \ldots m,$$

一个 reference

比起线性化来，对数线性化形式的系统解更容易解读，并且有时候更精确。

Remark 6.2 (常规扰动和奇异扰动)**.** *有时候我们需要区分常规扰动和奇异扰动。*

常规扰动 (regular perturbations) 往往指，外部环境的一个小变动引发经济系统解的一个小变动，例如新凯恩斯主义模型中，一个货币政策冲击 (标准差的变化) 印发经济系统均衡路径的变化，进而带来产出、通胀等内生变量的小变化 (Galí, 2008; Woodford, 2011)。DSGE 模型大多数研究常规扰动的经济现象，本论文也以常规扰动的介绍为主。

奇异扰动法 (singular perturbations) 往往指，外部环境的小变动引发经济系统的大波动，如市场出清价格。如不完全竞争市场模型中，一种新资产的出现可能带来经济系统解的大调整，对于研究金融市场摩擦和/或不完全竞争市场的经济学家来说，就需要予以额外关注 (Judd and Guu, 1993)。Judd (1998) 介绍了如何采用分叉法 (bifurcation) 求解奇异扰动的经济系统问题。

6.3 求解方法

6.3.1 状态——空间形式

在第6.2节分析框架的基础上，我们来看如何将扰动法应用到求解典型经济系统中，主要参考自Schmitt-Grohé and Uribe (2004)。

对于由一系列均衡条件和预算约束条件构成的非线性经济系统，可以表述为如下含有期望的状态——空间形式

$$E_t\mathcal{H}(y, y', x, x') = 0, \tag{6.2}$$

其中

- x, y 分别是 $n_x \times 1$ 和 $n_y \times 1$ 的状态向量和控制向量，$n = n_x + n_y$，x', y' 表示 $t+1$ 期的变量。
- 运算符 $\mathcal{H} : \mathcal{R}^{n_y} \times \mathcal{R}^{n_y} \times \mathcal{R}^{n_x} \times \mathcal{R}^{n_x} \to \mathcal{R}^n$ 表示方程系统，其中部分方程含有期望项。
- 状态向量 x 可以分解成两块，分别表示内生状态变量和外生变量：

$$x = \begin{bmatrix} \underset{(n_x - n_\varepsilon \times 1)}{x_1} \\ \underset{(n_\varepsilon \times 1)}{x_2} \end{bmatrix}$$

6.3.2　稳定状态

为了用扰动法分析经济系统(6.2)，首先要求得模型的非随机稳定状态，可表示为向量集合 $(\bar{x}, \bar{y})$，满足

$$\mathcal{H}(\bar{y}, \bar{y}, \bar{x}, \bar{x}) = 0. \tag{6.3}$$

通常来说，可以直接求得 $(\bar{x}, \bar{y})$ 的解析解，或使用一些常规的非线性求解方法 (见下文)，我们假定稳态解存在且唯一[4]。此外，也可能系统不存在唯一的稳态解，而是有一个均衡路径解 (BGP, Balanced Growth Path)，即模型中的内生变量，除了劳动力等少量例外之外，以大于 0 的某一相同速度增长，这就需要对变量做 scaling (或称为去趋势 detrending)，以便在随后进一步应用局部扰动法。具体说来，设

$$\hat{x}_t = \frac{x_t}{\mu_t},$$

其中 μ_t 是含有趋势的变量，如某一有偏技术进步过程，用额外一个方程来表示。经过 scale 或称 detrend 处理后的 $\hat{x}_t$ 可代回系统中做扰动法处理，并且如果有必要，在下一阶段的近似求解及方针过程中，将趋势重新加回来，变成 x_t。DSGE 模型中对变量做 scaling 处理的例子可见如Fernández-Villaverde and Rubio-Ramírez (2007); Christiano et al. (2010)。

在求解稳态 $(\bar{x}, \bar{y})$ 时，有些小窍门可以使得计算过程更简单。举三个例子。

系统缩减

第一个例子是方程系统的缩减 (system reduction)。将由 n 个方程，n 个未知变量的原系统，通过替换掉其中一部分变量，缩减为含有 n' 个方程，n' 个未知变量的新系统，$n' < n$。例如Christiano et al. (2005) 利用资本存量的运动式

$$k_{t+1} = i_t + (1 - \delta) k_t$$

在经济系统模型中替换掉投资 i_t。通过方程系统缩减，在下一步扰动法求解非线性系统过程中，工作复杂度进一步降低。如Sikorski (1985) 便指出，方程系统维度的增加常伴随着求解过程复杂程度的指数上升。

变量标准化与参数赋值

第二个例子是通过变量的标准化 (normalization) 确定参数值。可将一个或数个变量设为固定值，进而其他变量是关于这几个固定变量有关的方程，对非线性系统进行简化。以简单的随机内生经济增长模型来说：先来看一个变量标准化的例子：典型家庭的效用函数表现为 log CRRA 形式

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln c_t - \psi \frac{l_t^{1+\eta}}{1+\eta} \right),$$

[4]的确存在着无解或存在多个解的可能，相关讨论见Galor (2007)。

生产函数

$$y_t = A_t k_t^{\alpha} l_t^{1-\alpha},$$

资本存量的运动法则

$$k_{t+1} = i_t + (1-\delta)k_t = (yy_t - c_t) + (1-\delta)k_t.$$

通过计算 FOC，可得模型一组静态最优条件，其中劳动力供应式

$$w_t = \psi c_t l_t^{\eta}.$$

可以将非随机稳态的劳动力供应予以标准化，设为 $\bar{l} = 1$。对应地，由其他全部均衡条件构成的新系统中，内生变量均表示为与此标准化劳动量 $\bar{l}$ 有关的函数。进而，通过静态优化条件，我们可以得到参数 ψ 的值

$$\psi = \frac{\bar{w}}{\bar{c}\bar{l}^{\eta}} = \frac{\bar{w}}{\bar{c}}.$$

多变量标准化与参数赋值

第三个例子是多变量标准化。在上个模型已经引入 $\bar{l} = 1$ 的基础上，在引入第二个标准化，设 $\bar{y}y = 1$，从而

$$\bar{y} = 1 = \bar{A}\bar{k}^{\alpha}\bar{l}^{1-\alpha} = \bar{A}\bar{k}^{\alpha},$$

我们因此有

$$\bar{A} = \frac{1}{\bar{k}^{\alpha}},$$
$$\bar{w} = (1-\alpha)\frac{\bar{y}}{\bar{l}} = 1-\alpha.$$

跨期消费的 Euler 方程变为

$$\frac{1}{\bar{c}} = \frac{1}{\bar{c}}\beta\left(1+\bar{r}-\delta\right),$$

因此

$$\bar{r} = \frac{1}{\beta} - 1 + \delta.$$

此外由于

$$\bar{r} = MPK = \alpha\frac{\bar{y}}{\bar{k}} = \frac{\alpha}{\bar{k}},$$

可得稳态资本存量和稳态消费

$$\bar{k} = \frac{\alpha}{\bar{r}} = \frac{\alpha}{\frac{1}{\beta} - 1 + \delta},$$
$$\bar{c} = \bar{y} - \delta\bar{k} = 1 - \delta\frac{\alpha}{\frac{1}{\beta} - 1 + \delta}.$$

最后我们有

$$\psi = \frac{\bar{w}}{\bar{c}} = \frac{1-\alpha}{1-\delta\frac{\alpha}{\frac{1}{\beta}-1+\delta}}.$$

6.3.3 外生随机过程

设外生冲击变量 x_2 可以表示为如下过程

$$\boldsymbol{x}_2' = \boldsymbol{C}(\boldsymbol{x}_2) + \sigma\eta_{\varepsilon}\varepsilon', \tag{6.4}$$

其中

- 运算符 $\boldsymbol{C}$ 可能反映非线性关系。在这里假定围绕稳定状态 $(\bar{x}, \bar{y})$ 所做的 $\boldsymbol{C}$ 的 Hessian matrix 中，所有特征值绝对值都小于 1。
- ε' 包含 n_ε 个外生波动，常假定为 $\mathcal{N}(0, I)$，即均值为 0，二阶矩小于 ∞ 的分布。在研究中常常需要引入额外的矩限定条件。
- η_ε 是一个 $n_\varepsilon \times n_\varepsilon$ 的关于波动 ε 的方差协方差矩阵，$\sigma \geq 0$ 是一个关于 η_ε 的扰动参数 (Schmitt-Grohé and Uribe, 2004)，我们将在随后深入讨论。

因此经验研究中，常将(6.4)写成

$$\boldsymbol{x}_2' = \boldsymbol{C}\boldsymbol{x}_2 + \sigma\eta_\varepsilon\varepsilon', \tag{6.5}$$

其中 $\boldsymbol{C}$ 是一个 $n_\varepsilon \times n_\varepsilon$ 矩阵，其全部特征值绝对值小于 1。

外生冲击的波动的线性特征

(6.4)-(6.5)均假定波动是以线性方式作用于随机过程 $\boldsymbol{x}_2$ 的。这一假定看似随意，但不失一般特性。如我们不再假定这一线性关系，设

$$\boldsymbol{x}_{2,t} = \boldsymbol{D}(\boldsymbol{x}_{2,t-1}, \sigma\eta_\varepsilon\varepsilon_t).$$

将扰动 ε_t 写入经济系统 $\boldsymbol{x}_{2,t}$ 中

$$\tilde{\boldsymbol{x}}_{2,t} = \begin{bmatrix} \boldsymbol{x}_{2,t-1} \\ \varepsilon_t \end{bmatrix}, \quad \tilde{\varepsilon}_t = \begin{bmatrix} \boldsymbol{0}_{n_\varepsilon \times 1} \\ \varepsilon_t \end{bmatrix},$$

则我们有

$$\boldsymbol{x}_{2,t} = \tilde{\boldsymbol{D}}(\tilde{\boldsymbol{x}}_{2,t}, \sigma\eta_\varepsilon).$$

外生随机冲击的过程，可以重新表示为

$$\begin{bmatrix} \boldsymbol{x}_{2,t} \\ \varepsilon_{t+1} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{D}}(\tilde{\boldsymbol{x}}_{2,t}, \sigma\eta_\varepsilon) \\ 0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{0}_{n_\varepsilon \times 1} \\ \varepsilon_{t+1} \end{bmatrix},$$

或者用递归形式来表示

$$\tilde{\boldsymbol{x}}_2' = \boldsymbol{C}(\tilde{\boldsymbol{x}}_2) + \tilde{\varepsilon}', \quad \tilde{\varepsilon} \sim iid(0, I).$$

举例来说明。时变波动性 (time-varying volatility) 对于理解总量层面上的变量变化具有重要影响 (Bloom, 2009; Fernández-Villaverde et al., 2011)。假定生产率是一个随机波动的过程，满足

$$\log a_t = \rho_a \log a_{t-1} + \lambda_t \nu_t,$$

其中波动 $\nu_t \sim \mathcal{N}(0, 1)$，波动的标准差 λ_t 表示为另一个自回归过程

$$\log \lambda_t = \bar{\lambda} + \rho_\lambda \log \lambda_{t-1} + \psi\eta_t, \quad \eta_t \sim \mathcal{N}(0, 1).$$

需要指出的是，在这个模型中，扰动参数 ψ 不止影响生产率的波动 ν_t，还影响波动的标准差 λ_t。

时变波动模型中常常含有大量状态变量。对于每一个随机过程，都需要追踪其水平值及方差随时间的变化情况。这使得适合采用扰动法来做模型近似求解，求他方法如映射法等，则难以处理如此多的变量。

此外需要指出的是，如(6.5)所示，尽管一个模型可能有许多个冲击，但在模型设定上只需要一个扰动参数 ψ 即可，不同冲击的相对水平及联动 (comovements) 是由系数矩阵 η_ε 来调节的。如果 $\sigma = 0$，则模型回到确定性经济增长模型中去。

扰动参数的讨论

(6.5)的经济模型中，引入扰动参数 σ 反映随机过程的标准差。然而事实上这并不是唯一的建模方案。首先，在模型中其他位置用其他参数来反映波动，可能会使得模型的解更加精确，和/或模拟结果与现实情况更贴近，如Hansen et al. (2008) 构建了一个含有 Epstein-Zin 偏好的经济模型，并围绕跨期弹性等于 1 时的点做扰动。其次，比起离散时间模型来，连续时间建模中对扰动参数的处理有所不同，并且后者在控制方差方面有优势。

扰动法近似解与仿真结果的稳定性

存在这样的可能：扰动法求得的近似解呈现收敛特征，但根据近似解所生成的仿真数据序列却是爆炸的。一种常见的处理方法如Samuelson (1970) 和Jin and Judd (2002) 等，在在外生过程的波动中引入一个有界的设定，从而避免扰动法近似解中可能存在的仿真结果稳定性问题。另一种新出现的方法称剪枝法，见下文。

一个 reference。

6.3.4 设想解

非线性经济系统(6.2)设想中的解，应该由如下一组方程构成，分别是控制变量的决策式和状态变量的决策式：

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x};\delta), \tag{6.6}$$

$$\boldsymbol{x}' = \boldsymbol{h}(\boldsymbol{x};\delta) + \sigma\eta\varepsilon', \tag{6.7}$$

其中两个运算符 $\boldsymbol{g}:\mathcal{R}^{n_x}\times\mathcal{R}^{+}\mapsto\mathcal{R}^{n_y}$，$\boldsymbol{h}:\mathcal{R}^{n_x}\times\mathcal{R}^{+}\mapsto\mathcal{R}^{n_x}$。系数矩阵 η 可做如下分解

$$\underset{(n_x\times n_\varepsilon)}{\eta} = \begin{bmatrix} \underset{(n_x-n_\varepsilon\times n_\varepsilon)}{\mathbf{0}} \\ \underset{(n_\varepsilon\times n_\varepsilon)}{\eta_\varepsilon} \end{bmatrix},$$

其中前半部分 ($n_x - n_\varepsilon$ 行) 由当期状态决定，这些当期状态影响下一期的内生状态。后半部分 (n_ε 行) 由下一期的外生变量决定，下一期的外生变量由当期外生变量和下一期的波动共同决定。

扰动法求解经济系统的目标是：围绕某一适当的不动点，对 $\boldsymbol{g}$ 和 $\boldsymbol{h}$ 做泰勒级数展开。分三个步骤逐次展开。

步骤一，常用非随机稳态作为不动点，对应 $\boldsymbol{x}_t \equiv \bar{\boldsymbol{x}}$，$\sigma = 0$。根据第6.3.2的方法，我们可以计算稳态值 $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$。

步骤二，将确定性稳态 $(\bar{\boldsymbol{x}};0)$ 代入决策方程组(6.6)-(6.7)有

$$\boldsymbol{g}(\bar{\boldsymbol{x}};0) = \bar{\boldsymbol{y}}, \tag{6.8}$$

$$\boldsymbol{h}(\bar{\boldsymbol{x}};0) = \bar{\boldsymbol{x}}. \tag{6.9}$$

步骤三，引入关于 $\mathcal{H}$ 的未知解，定义为一个新的运算符 $F:\mathcal{R}^{n_x+1}\mapsto\mathcal{R}^{n_x}$，式(6.2)改写为

$$\begin{aligned} F(\boldsymbol{x},\sigma) &\equiv E_t\mathcal{H}\left(\boldsymbol{y}',\boldsymbol{y},\boldsymbol{x}',\boldsymbol{x}\right) \\ &= E_t\mathcal{H}\left(\boldsymbol{g}\left(\boldsymbol{h}(x;\sigma)+\delta\eta\varepsilon';\sigma\right),\boldsymbol{g}\left(\boldsymbol{x};\sigma\right),\boldsymbol{h}(x;\sigma)+\delta\eta\varepsilon',\boldsymbol{x}\right) \\ &= \mathbf{0}. \end{aligned}$$

由于 $F(\boldsymbol{x};\sigma) = 0 \quad \forall \boldsymbol{x},\sigma$，因此我们有

$$F_{x_i^k\sigma^j}(\boldsymbol{x};\sigma) = 0, \quad \forall x,\sigma,i,k,j, \tag{6.10}$$

其中 $F_{x_i^k\sigma^j}(\boldsymbol{x};\sigma)$ 表示用 $F(\boldsymbol{x};\sigma)$ 对 $\boldsymbol{x}$ 中第 i 个元素 x 求导 k 次，对 σ 求导 j 次，求导都是围绕 $(\boldsymbol{x};\sigma)$ 进行的。根据模型定义，对于所有可能的状态 $x \in \boldsymbol{x}$ 和扰动参数值 σ，均应满足均衡条件 $F(\cdot)=0$。

这带来两个问题需要做进一步的讨论。

可求导性的探讨

(6.10)成立的前提是假定在一个固定点，比如稳态 $(\bar{x};0)$ 附近，$F(\cdot)$ 的所有导数都存在。在 DSGE 模型中，进入 $F(\cdot)$ 的各个模块常常是平滑的，如效用函数、生产函数等，这使得该假定可以被接受 (Judd, 1998, pp463)[5]。

可求导性假定指出了扰动法适用范围的限制：如果在一个经济系统中，部分变量是离散的，或相关均衡条件不可导，便无法使用扰动法予以近似。此外有两点需要补充。第一，模型中出现的期望项，有时候会将看起来是时间离散的问题转化为时间连续的问题，举例来说，要不要上大学的决策可以有一个大学学费的随机冲击，或一个努力程度的变量 (描述该学生学习或申请奖学金的努力程度) 来平滑掉。第二，即便最坏的情况发生了，可求导性假定不成立，但这也并不是说扰动法完全没有价值——它仍然可以成为一个寻找替代求解方法的猜测值。 补一个 ref

导数的计算

在扰动法近似求解过程中，导数的计算是核心环节之一，然而这往往是一项繁琐的工作，完全依手算较不明智，时间较长、错误率较高 (Judd, 1998, ch.7)。这时需要借助计算机和支持符号运算的计算软件，如 Mathematica、Python 等。因此现在常用结合二者优点的自动微分 (AD, automatic differentiation) 技术，将链式法则 (chain rule) 应用到一系列基本运算中区。自动微分法在 DSGE 模型中的运用，相关讨论可见Bastani and Guerrieri (2008)。

6.3.5　一阶扰动

围绕 $(\bar{\boldsymbol{x}};0)$ 对 $\boldsymbol{g}$ 和 $\boldsymbol{h}$ 做一阶扰动有

$$\begin{aligned}\boldsymbol{y}' &= \boldsymbol{g}(\boldsymbol{x};\sigma) = \boldsymbol{g}(\bar{\boldsymbol{x}};0) + \boldsymbol{g}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)(\boldsymbol{x}-\bar{\boldsymbol{x}})' + \boldsymbol{g}_{\boldsymbol{\sigma}}(\bar{\boldsymbol{x}};0)\sigma \\ \boldsymbol{x}' &= \boldsymbol{h}(\boldsymbol{x};\sigma) = \boldsymbol{h}(\bar{\boldsymbol{x}};0) + \boldsymbol{h}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)(\boldsymbol{x}-\bar{\boldsymbol{x}})' + \boldsymbol{h}_{\boldsymbol{\sigma}}(\bar{\boldsymbol{x}};0)\sigma,\end{aligned}$$

结合(6.8)-(6.9)的稳态值，可得

$$\begin{aligned}\boldsymbol{g}(\boldsymbol{x};\sigma) - \bar{\boldsymbol{y}} &= \boldsymbol{g}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)(\boldsymbol{x}-\bar{\boldsymbol{x}})' + \boldsymbol{g}_{\boldsymbol{\sigma}}(\bar{\boldsymbol{x}};0)\sigma \\ \boldsymbol{h}(\boldsymbol{x};\sigma) - \bar{\boldsymbol{x}} &= \boldsymbol{h}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)(\boldsymbol{x}-\bar{\boldsymbol{x}})' + \boldsymbol{h}_{\boldsymbol{\sigma}}(\bar{\boldsymbol{x}};0)\sigma,\end{aligned} \tag{6.11}$$

即，在已知 $(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}})$ 的基础上，只需要再得到四组未知系数的值，总共 $n\times(n_x+1)$ 个，我们便可测算出任意一点 $(\boldsymbol{x};\sigma)$ 对应的 $(\boldsymbol{x}',\boldsymbol{y}')$：

$$\begin{bmatrix} \underset{(n_x\times n_y)}{\boldsymbol{g}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)} \\ \underset{(n_x\times n_x)}{\boldsymbol{h}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)} \\ \underset{(n_y)}{\boldsymbol{g}_{\boldsymbol{\sigma}}(\bar{\boldsymbol{x}};0)} \\ \underset{(n_x)}{\boldsymbol{h}_{\boldsymbol{\sigma}}(\bar{\boldsymbol{x}};0)} \end{bmatrix}$$

这一系列未知系数，可以通过(6.10)求得，具体来说，分两组

[5]对这一假定的经验验证可能会较为复杂，需要做更深入的研究，可参考经典讨论如Santos (1992)。

- 利用 $F_x(\bar{\boldsymbol{x}};0)=0(n\times n_x$ 个方程) 求得 $(\boldsymbol{g}_x(\bar{\boldsymbol{x}};0),\boldsymbol{g}_x(\bar{\boldsymbol{x}};0))$,
- 利用 $F_\sigma(\bar{\boldsymbol{x}};0)=0(n$ 个方程) 求得 $(\boldsymbol{g}_\sigma(\bar{\boldsymbol{x}};0),\boldsymbol{g}_\sigma(\bar{\boldsymbol{x}};0))$。

在做进一步求导之前，我们先介绍一下张量 (tensor) 的表示形式。

Remark 6.3 (张量形式)**.** *张量 (tensor)，或称爱因斯坦求和约定 (Einstein summation notation)，常用于物理学研究中，致力于将执行扰动近似所需的代数计算工作维持在一个可控的范围内。这主要包括两部分：一是省略 Σ,∂ 等运算符号，二是省略求导所围绕的固定点的表述。在 1 个 m 维空间中，第 n 个张量是 1 个有 n 个 index 和 m^n 个元素的运算符，符合相关的转换规则。如在此例中，$\mathcal{H}$ 对 $\boldsymbol{y}$ 求导形成一个 $n\times n_y$ 的导数矩阵，$[\mathcal{H}]^i_\alpha$ 表示其中第 i 行第 α 个元素。当前一个矩阵的下角标也是后一个矩阵的上角标时，我们常常省略求和符号，如*

$$[\mathcal{H}_y]^i_\alpha\,[\boldsymbol{g}_x]^\alpha_\beta\,[\boldsymbol{h}_x]^\beta_j=\sum_{\alpha=1}^{n_y}\sum_{\beta=1}^{n_x}\frac{\partial\mathcal{H}^i}{\partial y^\alpha}\frac{\partial\boldsymbol{g}^\alpha}{\partial x^\beta}\frac{\partial\boldsymbol{h}^\beta}{\partial x^j}.$$

类似地，高阶导数常表示为 $[\mathcal{H}_{y'y'}]^i_{\alpha\gamma}$ 的形式。其中 $\mathcal{H}_{y'y'}$ 是个 n 行，n_y 列，n_y 页的三围数组，$[\mathcal{H}_{y'y'}]^i_{\alpha\gamma}$ 是二阶导数矩阵的第 i 行，第 α 列，第 γ 页元素。

$F_x(\cdot)$ 的求解

先来看(6.10)中的前半部分。$(\boldsymbol{g_x}(\bar{x};0),\boldsymbol{h_x}(\bar{x};0))$ 构成 $[F_x(\bar{x};0)]^i_j$ 的解，表示为

$$\begin{aligned}[F_x(\bar{x};0)]^i_j&=[\mathcal{H}_{y'}]^i_\alpha\,[\boldsymbol{g}_x]^\alpha_\beta\,[\boldsymbol{h}_x]^\beta_j\\&+[\mathcal{H}_y]^i_\alpha\,[\boldsymbol{g}_x]^\alpha_j\\&+[\mathcal{H}_{x'}]^i_\beta\,[\boldsymbol{h}_x]^\beta_j\\&+[\mathcal{H}_x]^i_j\\&=0,\\&i=1,\dots n,\quad j,\beta=1,\dots n_x,\quad\alpha=1,\dots n_y.\end{aligned}\tag{6.12}$$

(6.12)构成一个 $n\times n_x$ 个方程组成的系统，其中含有一组系数和一组待解未知数。系数包含 $\left(\underset{n\times n_y}{[\mathcal{H}_{y'}]^i_\alpha},\underset{n\times n_y}{[\mathcal{H}_y]^i_\alpha},\underset{n\times n_x}{[\mathcal{H}_{x'}]^i_\beta},\underset{n\times n_x}{[\mathcal{H}_x]^i_j}\right)$。稳定状态 $(\bar{\boldsymbol{x}};0)$ 对应 $(\bar{\boldsymbol{y}}',\bar{\boldsymbol{y}},\bar{\boldsymbol{x}}',\bar{\boldsymbol{x}})$，可以对 $\mathcal{H}$ 分别求导数测算这组系数。未知数有 $n\times n_x$ 个，由元素 $\left(\underset{n_x\times n_x}{\boldsymbol{g}_x(\bar{\boldsymbol{x}};0)},\underset{n_x\times n_x}{\boldsymbol{h}_x(\bar{\boldsymbol{x}};0)}\right)$ 构成。

根据上述分析，可将方程系统(6.12)写为一个二次形式的矩阵系统

$$\underset{\tilde{n}\times\tilde{n}}{A}P^2-\underset{\tilde{n}\times\tilde{n}}{B}P-\underset{\tilde{n}\times\tilde{n}}{C}=\underset{\tilde{n}\times\tilde{n}}{0},\tag{6.13}$$

其中系数矩阵 A,B,C 包含 $\left([\mathcal{H}_{y'}]^i_\alpha,[\mathcal{H}_y]^i_\alpha,[\mathcal{H}_{x'}]^i_\beta,[\mathcal{H}_x]^i_j\right)$。P 包含待解未知数 $[\boldsymbol{h}_x]^\beta_j$，是与状态向量 $\boldsymbol{x}$ 中内生状态向量 $\boldsymbol{x}_1$ 的运动法则式有关的元素。

至此我们可以采用标准的二次矩阵方程求解法来求 P，如 QZ 分解法。

Remark 6.4 (QZ 分解法求解二次矩阵方程系统)**.** *援引 (Uhlig, 1999, pp.43-45) 的方法。对于二次矩阵系统(6.13)，定义两个矩阵*

$$\underset{2\tilde{n}\times2\tilde{n}}{D}=\begin{bmatrix}A&\underset{\tilde{n}\times\tilde{n}}{0}\\\underset{\tilde{n}\times\tilde{n}}{0}&\underset{\tilde{n}\times\tilde{n}}{I}\end{bmatrix},\quad\underset{2\tilde{n}\times2\tilde{n}}{D}=\begin{bmatrix}B&C\\\underset{\tilde{n}\times\tilde{n}}{I}&\underset{\tilde{n}\times\tilde{n}}{0}\end{bmatrix}.$$

定义两个酉矩阵 *(unitary matrix)* Q, Z[6]。

设两个上三角矩阵 Φ, Σ，对角元素分别为 ϕ_{ii}, σ_{ii}。

对 D 和 F 分别作扩展 *Schur* 分解 *(QZ* 分解*)*:

$$
\begin{aligned}
Q^\top \Sigma Z &= D, \\
Q^\top \Phi Z &= F,
\end{aligned}
$$

其中确保 Σ 和 Φ 是对角矩阵，并且对角元素的比值 $\left|\frac{\phi_{ii}}{\sigma_{ii}}\right|$ 按照从左上到右下的升序排列[7]，从而可以将系统分为两部分，前半部分是稳定分块，对应 $\left|\frac{\phi_{ii}}{\sigma_{ii}}\right| < 1$；后半部分是不稳定分块，对应 $\left|\frac{\phi_{ii}}{\sigma_{ii}}\right| > 1$。在排序之后，可以将矩阵输入 *Mathematica*，做 *QZ* 分解计算。测得的矩阵

$$
\underset{2\tilde{n}\times 2\tilde{n}}{Z} = \begin{bmatrix} \underset{\tilde{n}\times\tilde{n}}{Z_{11}} & \underset{\tilde{n}\times\tilde{n}}{Z_{12}} \\ \underset{\tilde{n}\times\tilde{n}}{Z_{21}} & \underset{\tilde{n}\times\tilde{n}}{Z_{22}} \end{bmatrix}.
$$

进而我们有

$$
P = -Z_{21}^{-1} Z_{22}.
$$

对于 $2\tilde{n}$ 个扩展特征值比 *(generalized eigenvalue ratio)* $\left|\frac{\phi_{ii}}{\sigma_{ii}}\right|$ 来说:

1. 如果稳定的扩展特征值比数量 $< \tilde{n}$，那么 P 值存在且唯一，对于任何 $\tilde{n}$ 维度向量中，有 $\lim_{m\to\infty} P^m x \to \mathbf{0}$。

2. 如果稳定的扩展特征值比数量 $> \tilde{n}$，那么 P 值可能有多个，我们选取任何一个满足如下条件的 P 值：$\tilde{n}$ 维度向量中，$\lim_{m\to\infty} P^m x \to \mathbf{0}$。

之所以会出现二次形式的矩阵系统，其原因在于，通常说来满足模型均衡条件的内生变量变化路径并不唯一，而是有很多 (Uhlig, 1999; Galor, 2007)，其中有一些是稳定路径，有一些是不稳定路径，而只有稳定路径才能满足适当的横截条件。显然，对于存在多重解的经济系统而言，我们需要找出其中的稳定路径解。许多 DSGE 模型中存在 $\tilde{n}$ 个扩展特征根，从而确保系统有且只有 1 个稳定解。如 $< \tilde{n}$，扩展特征根的数量不足，会导致均衡动态路径出现内在不稳定的状况。如 $> \tilde{n}$，会存在太阳黑子现象 (Lubik and Schorfheide, 2003)。这段对解的存在性和唯一性的分析是基于一阶扰动近似而言的，但对高阶扰动近似过程也同样有效，换句话说：如果一阶扰动近似的解是唯一的 (不唯一的)，则高阶扰动近似的解也是唯一的 (不唯一的)。

Remark 6.5 (二次系统的递归分解)**.** 对于(6.12)的经济系统，还可以分解为两部分，以递归形式分两步求解。第一步，定义一个分块系统

$$
\begin{aligned}
[F_x(\bar{x};0)]^i_j &= [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\beta [\boldsymbol{h}_x]^\beta_j \\
&\quad + [\mathcal{H}_y]^i_\alpha [\boldsymbol{g}_x]^\alpha_j \\
&\quad + [\mathcal{H}_{x'}]^i_\beta [\boldsymbol{h}_x]^\beta_j \\
&\quad + [\mathcal{H}_x]^i_j \\
&= 0, \\
&\quad i = 1, \ldots n, \quad j, \beta = 1, \ldots \tilde{n}, \quad \alpha = 1, \ldots n_y.
\end{aligned} \tag{6.14}
$$

[6]满足条件 $Q^H Q = Z^H Z = I_{2\tilde{n}}$，其中 H 表示埃米特转置。

[7]这是因为，$\left|\frac{\phi_{ii}}{\sigma_{ii}}\right|$ 的每一种排列顺序，都存在一个 QZ 分解的解。

与(6.12)相比，(6.14)的系统中未知数的数量减少了，从 $n \times n_x$ 个变为 $n \times \tilde{n}$ 个，由元素 $\left(\underset{\tilde{n}\times n_y}{\boldsymbol{g}_x(\bar{\boldsymbol{x}};0)}, \underset{\tilde{n}\times n_x}{\boldsymbol{h}_x(\bar{\boldsymbol{x}};0)}\right)$ 构成，二者都与 $\tilde{n}$ 个内生变量 $\boldsymbol{x}_1$ 有关。

采用上文提到的办法，求得对应(6.14)系统的解 $\left(\underset{\tilde{n}\times n_y}{\boldsymbol{g}_x(\bar{\boldsymbol{x}};0)}, \underset{\tilde{n}\times n_x}{\boldsymbol{h}_x(\bar{\boldsymbol{x}};0)}\right)$。

第二步，将前一步求得的解代入新系统

$$
\begin{aligned}
[F_x(\bar{x};0)]^i_j &= [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\beta [\boldsymbol{h}_x]^\beta_j \\
&+ [\mathcal{H}_y]^i_\alpha [\boldsymbol{g}_x]^\alpha_j \\
&+ [\mathcal{H}_{x'}]^i_\beta [\boldsymbol{h}_x]^\beta_j \\
&+ [\mathcal{H}_x]^i_j \\
&= 0, \\
& i = 1, \ldots n, \quad j, \beta = \tilde{n}+1, \ldots n_x, \quad \alpha = 1, \ldots n_y,
\end{aligned} \tag{6.15}
$$

并进一步求解与 n_ε 个外生随机变量 $\boldsymbol{x}_2$ 有关的未知数，由元素 $\left(\underset{n_\varepsilon\times n_y}{\boldsymbol{g}_x(\bar{\boldsymbol{x}};0)}, \underset{n_\varepsilon\times n_x}{\boldsymbol{h}_x(\bar{\boldsymbol{x}};0)}\right)$ 构成。

采用递归分解法求解二次系统，有三个优点。第一，可以显著改善计算时间：(6.14)比(6.12)减少了 $n_\varepsilon \times n$ 个未知数的计算，待求解系统规模小了很多 (Rubio-Ramírez and Fernández-Villaverde, 2008)。第二，(6.15)的线性系统，比起(6.12)的非线性系统来，求解更为便利。第三，对于含有 n_x 个状态变量的向量 $\boldsymbol{x}$ 而言，有时我们只需关注其中的 $\tilde{n}$ 个内生变量而非全部。这取决于特定的研究目标，一个例子是，在一定的初始条件下，计算经济系统向某一稳定状态运动的确定性转移路径。另一个栗子是，基于一阶扰动近似解，描绘经济系统的冲击——响应机制。

一个 ref，见下的范例。

$F_\sigma(\cdot)$ 的求解

再来看(6.10)中的后半部分。$(\boldsymbol{g}_\sigma(\bar{x};0), \boldsymbol{h}_\sigma(\bar{x};0))$ 构成 $[F_\sigma(\bar{x};0)]^i$ 的解，表示为

$$
\begin{aligned}
[F_\sigma(\bar{x};0)]^i &= E_t \{ [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\beta [\boldsymbol{h}_\sigma]^\beta \\
&+ [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\beta [\eta]^\beta_\phi [\varepsilon']^\phi \\
&+ [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_\sigma]^\alpha \\
&+ [\mathcal{H}_y]^i_\alpha [\boldsymbol{g}_\sigma]^\alpha \\
&+ [\mathcal{H}_{x'}]^i_\beta [\boldsymbol{h}_\sigma]^\beta \\
&+ [\mathcal{H}_{x'}]^i_\beta [\eta]^\beta_\phi [\varepsilon']^\phi \} \\
&= 0, \\
& i = 1, \ldots n, \quad \alpha = 1, \ldots n_y, \quad \beta = 1, \ldots n_x, \quad \phi = 1, \ldots n_\varepsilon.
\end{aligned} \tag{6.16}
$$

进而我们有

$$
[F_\sigma(\bar{x};0)]^i = [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\beta [\boldsymbol{h}_\sigma]^\beta + [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_\sigma]^\alpha + [\mathcal{H}_y]^i_\alpha [\boldsymbol{g}_\sigma]^\alpha + [\boldsymbol{f}_{x'}]^i_\beta [\boldsymbol{h}_\sigma]^\beta = 0, \tag{6.17}
$$

(6.16)构成一个 n 个方程组成的系统，是关于未知数 $(\boldsymbol{g}_\sigma(\bar{x};\sigma), \boldsymbol{h}_\sigma(\bar{x};\sigma))$ 的线性齐次方程组，唯一满足条件的解只能是：

$$
\begin{aligned}
\bar{\boldsymbol{g}}_\sigma(\bar{x};0) &= 0, \\
\bar{\boldsymbol{h}}_\sigma(\bar{x};0) &= 0.
\end{aligned} \tag{6.18}
$$

确定性等价条件

(6.18)代入(6.11)得

$$\begin{aligned} \boldsymbol{g}(\boldsymbol{x};\sigma) - \bar{\boldsymbol{y}} &= \boldsymbol{g}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)(\boldsymbol{x}-\bar{\boldsymbol{x}})', \\ \boldsymbol{h}(\boldsymbol{x};\sigma) - \bar{\boldsymbol{x}} &= \boldsymbol{h}_{\boldsymbol{x}}(\bar{\boldsymbol{x}};0)(\boldsymbol{x}-\bar{\boldsymbol{x}})'. \end{aligned} \tag{6.19}$$

这被称为确定性等价条件 (certainty equivalence)。Simon (1956); Theil (1957)。根据确定性等价条件，模型的一阶近似解等于在完全期望条件下，或者 $\sigma = 0$ 条件下的解。确定性等价条件承认外生冲击对系统的决策式 (decision rule) 的影响，但它排除外生冲击的标准差 (不确定性) 对系统决策式的影响。

确定性等价假定提出的理由很简单：典型个体的风险规避程度，通常可以通过 concave 效用函数的二阶求导给出，从而一阶扰动近似即可满足需要。但甚至早至Leland (1968); Sandmo (1970) 的研究便已事出，为了应对不确定性带来的风险，储蓄等个人的预防性行为 (precautionary behaviors) 的分析只有对效用函数做三阶求导才能推得，换句话说，二阶以上的扰动近似是必要的。与之相比，一阶线性扰动近似只涉及到模型的均衡条件 (尤其是效用函数的一阶导数，比如跨期消费决策的 Euler 方程)，以及对这些均衡条件的一阶求导 (包括对效用函数的二次求导)，而不包含更高阶导数了。

确定性等价存在以下几点不足：第一，难于分析不确定性给福利效果带来的影响。一方面模型的动态路径的确受到来自波动的方差 (标准差) 的影响，波动的方差本身即是波动的来源之一。但另一方面，模型中的行为人，基于确定性等价假定，无法采取预防性措施应对方差变化带来的不确定性风险，这使得相关福利分析难以进行。第二，对应于前一点，任何基于确定性等价的近似解，也因此无法用于评估资产的风险溢价。这并不符合最初构建模型时的设想：在一般均衡框架下，福利的测算和资产的定价二者之间存在着密切的关联，任何对前者的研究都自然涉及到后者 (Alvarez and Jermann, 2004)。第三，确定性等价假定使得研究者无法探讨波动 (volatility) 变化的经济效果。

Remark 6.6 (线性二次 (LQR) 法，近似 LQR 法，与扰动法)**.** *Kydland and Prescott (1982) 及其后相当数量的研究，探讨另一种 DSGE 模型的求解方法，可称之为线性二次调节法 (LQR, linear quadratic regulator)。对于一个最优控制问题，有 n_x 个状态变量 $\boldsymbol{x}_t$，和 n_u 个控制变量 $\boldsymbol{u}_t$，定义为 $\boldsymbol{w}_t = [\boldsymbol{x}_t, \boldsymbol{u}_t]^\top$，是个 $n_w = n_x + n_u$ 个维度的行向量。最优控制问题可以表示为*

$$\begin{aligned} &\max E_0 \sum_{t=0}^{\infty} \beta^t r(\boldsymbol{w}_t) \\ s.t. \quad &\boldsymbol{x}_{t+1} = A(\boldsymbol{w}_t, \varepsilon_t), \end{aligned}$$

其中

- *运算符 r 表示回报方程，*
- *运算符 A 汇总模型的均衡条件和预算约束条件，*
- *ε_t 为 n_ε 个冲击组成的向量，均值为 0，方差 $< \infty$*[8]*。*

做两个假设。第一个假设是回报方程是二次形式的

$$r(\boldsymbol{w}_t) = B_0 + \underset{1\times n_w}{B_1}\ \boldsymbol{w}_t + \boldsymbol{w}_t^\top \underset{n_w\times n_w}{B_w}\ \boldsymbol{w}_t.$$

[8]在这一状态——空间表示形式中，模型可以做适当扩展，以使 ε_t 也会对当期的回报方程，和当期变量 (包括状态变量 $\boldsymbol{x}_t$ 和当期控制变量 $\boldsymbol{u}_t$ 在内) 产生影响。

第二个假设是 A 是线性形式的

$$\boldsymbol{x}_{t+1} = \underset{n_x \times n_w}{B_3} \boldsymbol{w}_{t+1} + \underset{n_x \times n_\varepsilon}{B_4} \varepsilon_t.$$

这便构成了一个随机贴现线性——二次控制问题 *(sthochastic discounted linear-quadratic regulator probelm)*，有大量研究围绕这个问题做了深入探讨，可参考综述*Anderson et al. (1996); Hansen and Sargent (2004, 2013)*。已知该问题的最优决策是个关于状态和冲击的线性方程

$$\boldsymbol{u}_t = F_w \boldsymbol{w}_t + F_\varepsilon \varepsilon_t,$$

其中 F_w 可用 *Ricatti* 方程来求解，F_ε 可用 *Sylvester* 方程来求解 *(Anderson et al., 1996, pp. 182-183,202-205)*。本讲义的第四章也对求解方法做了介绍。F_w 不受冲击的方差的影响，换句话说，如果 $var(\varepsilon_t) = 0$，那么最优决策问题简化为

$$\boldsymbol{u}_t = F_w \boldsymbol{w}_t.$$

最优控制问题中的 *LQR* 法，经济系统分为两个部分，分别计算 F_w 和 F_ε，这使得研究者可以深入探讨一系列相关问题。然而需要指出的是，*LQR* 法仍然属于一阶扰动法的一种，因此也持有确定性等价的假定，受到该假定不足的限制。

Kydland and Prescott (1982) 构建经济模型来分析社会规划者问题，研究框架符合上述最优调节问题 *(optimal regulator problem)* 的范式，也因此可以求出关于 $\boldsymbol{w}_t$ 的线性方程 $A(\cdot)$。但在*Kydland and Prescott (1982)* 中，回报方程 $r(\cdot)$ 并未设定成二次形式；他们只是对社会规划者的目标方程做了二次近似。大多数后续研究都使用类似的方法，对目标方程围绕稳态做 *TSE*，可以归类称之为近似 *LQR* 问题。进而，可以用价值方程来求解近似 *LQR* 问题，见*Díaz-Giménez (1999)*。

当 $A(\cdot)$ 是线性的时，近似 *LQR* 问题的解，与对模型均衡条件做一阶扰动法近似的解，是相同的。原因很简单：导数总是唯一的，并且两种方法都致力于求得模型的近似线性解。然而，近似 *LQR* 法在学术界渐渐不受青睐，主要原因有：第一，*DSGE* 的模型常常很难将 $A(\cdot)$ 写为线性形式。第二，当经济体不是处于帕累托最优状态时，我们往往很难构建社会规划者问题。尽管的确有研究者常识构建调整社会规划者问题，引入一系列非最优的限制条件，如*Benigno and Woodford (2003)*，但比起对模型均衡条件做扰动近似的方法而言，前者的工作量更多也更为复杂。第三个也许是更重要的原因，扰动法可以更便利地做高阶近似，引入非线性元素，打破确定性等价的假定，从而使得模型可以分析更多经济现象。我们将在下一节探讨二阶扰动法。

6.3.6 二阶扰动

在求得一阶扰动法的解之后，我们可以迭代求得二阶扰动的系统解。具体来说

1. 对 $\boldsymbol{g}(\boldsymbol{x};\sigma)$ 围绕 $(\bar{\boldsymbol{x}};0)$ 做二阶近似：

$$\begin{aligned}
[\boldsymbol{g}(\boldsymbol{x};\sigma)]^i = &[\boldsymbol{g}(\bar{\boldsymbol{x}};0)]^i + [\boldsymbol{g}_x(\bar{\boldsymbol{x}};0)]^i_a [(\boldsymbol{x}-\bar{\boldsymbol{x}})]_a + [\boldsymbol{g}_\sigma(\bar{\boldsymbol{x}};0)]^i_\sigma [\sigma] \\
&+ \frac{1}{2} [\boldsymbol{g}_{xx}(\bar{\boldsymbol{x}};0)]^i_{ab} [(\boldsymbol{x}-\bar{\boldsymbol{x}})]_a [(\boldsymbol{x}-\bar{\boldsymbol{x}})]_b \\
&+ \frac{1}{2} [\boldsymbol{g}_{x\sigma}(\bar{\boldsymbol{x}};0)]^i_a [(\boldsymbol{x}-\bar{\boldsymbol{x}})]_a [\sigma] \\
&+ \frac{1}{2} [\boldsymbol{g}_{\sigma x}(\bar{\boldsymbol{x}};0)]^i_a [(\boldsymbol{x}-\bar{\boldsymbol{x}})]_a [\sigma] \\
&+ \frac{1}{2} [\boldsymbol{g}_{\sigma\sigma}(\bar{\boldsymbol{x}};0)]^i [\sigma] [\sigma], \\
i = 1 \ldots n_y, &\quad a, b = 1, \ldots n_x.
\end{aligned} \tag{6.20}$$

2. 对 $\boldsymbol{h}(\boldsymbol{x};\sigma)$ 围绕 $(\bar{\boldsymbol{x}};0)$ 做二阶近似：

$$
\begin{aligned}
[\boldsymbol{h}(\boldsymbol{x};\sigma)]^i = & [\boldsymbol{h}(\bar{\boldsymbol{x}};0)]^j + [\boldsymbol{h}_x(\bar{\boldsymbol{x}};0)]^j_a \left[(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]_a + [\boldsymbol{g}_\sigma(\bar{\boldsymbol{x}};0)]^i_\sigma \\
& + \frac{1}{2}\left[\boldsymbol{h}_{xx}(\bar{\boldsymbol{x}};0)\right]^j_{ab}\left[(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]_a\left[(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]_b \\
& + \frac{1}{2}\left[\boldsymbol{h}_{x\sigma}(\bar{\boldsymbol{x}};0)\right]^j_a\left[(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]_a[\sigma] \\
& + \frac{1}{2}\left[\boldsymbol{h}_{\sigma x}(\bar{\boldsymbol{x}};0)\right]^j_a\left[(\boldsymbol{x}-\bar{\boldsymbol{x}})\right]_a[\sigma] \\
& + \frac{1}{2}\left[\boldsymbol{h}_{\sigma\sigma}(\bar{\boldsymbol{x}};0)\right]^j[\sigma][\sigma], \\
& j = 1\ldots n_x, \quad a,b = 1,\ldots n_x.
\end{aligned}
\tag{6.21}
$$

两个方程系统中共有 6 个待求解未知数 $(\boldsymbol{g}_{xx},\boldsymbol{g}_{x\sigma},\boldsymbol{g}_{\sigma\sigma},\boldsymbol{h}_{xx},\boldsymbol{h}_{x\sigma},\boldsymbol{h}_{\sigma\sigma})_{(\bar{\boldsymbol{x}};0)}$，可同样利用式(6.10)求得。具体说来，分为三组：

- 利用 $F_{xx}(\bar{\boldsymbol{x}};0)=0$ 求解 $(\boldsymbol{g}_{xx}(\bar{\boldsymbol{x}};0),\boldsymbol{h}_{xx}(\bar{\boldsymbol{x}};0))$，
- 利用 $F_{\sigma\sigma}(\bar{\boldsymbol{x}};0)=0$ 求解 $(\boldsymbol{g}_{\sigma\sigma}(\bar{\boldsymbol{x}};0),\boldsymbol{h}_{\sigma\sigma}(\bar{\boldsymbol{x}};0))$，
- 利用 $F_{\sigma x}(\bar{\boldsymbol{x}};0)=0$ 求解 $(\boldsymbol{g}_{\sigma x}(\bar{\boldsymbol{x}};0),\boldsymbol{h}_{\sigma x}(\bar{\boldsymbol{x}};0))$。

$F_{xx}(\cdot)$ 的求解

一阶导数 $F_x(\cdot)=0$，式(6.12)的 LHS 表示为 4 个部分相加。因此二阶导数

$$
[F_{xx}(\bar{\boldsymbol{x}};0)]^i_{jk} = \mathcal{A}+\mathcal{B}+\mathcal{C}+\mathcal{D}=0, \tag{6.22}
$$

其中

$$
\begin{aligned}
\mathcal{A} = & \Big([\mathcal{H}_{y'y'}]^i_{\alpha\gamma}[\boldsymbol{g}_x]^\gamma_\delta[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{y'y'}]^i_{\alpha\gamma}[\boldsymbol{g}_x]^\gamma_k + [\mathcal{H}_{y'x'}]^i_{\alpha\delta}[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{y'x}]^i_{\alpha k}\Big)[\boldsymbol{g}_x]^\alpha_\beta[\boldsymbol{h}_x]^\beta_j \\
& + [\mathcal{H}_{y'}]^i_\alpha[\boldsymbol{g}_{xx}]^\alpha_{\beta\delta}[\boldsymbol{h}_x]^\delta_k[\boldsymbol{h}_x]^\beta_j + [\mathcal{H}_{y'}]^i_\alpha[\boldsymbol{g}_x]^\alpha_\beta[\boldsymbol{h}_{xx}]^\beta_{jk}, \\
\mathcal{B} = & [\mathcal{H}_{yy'}]^i_{\alpha\gamma}[\boldsymbol{g}_x]^\gamma_\delta[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{yy}]^i_{\alpha\gamma}[\boldsymbol{g}_x]^\gamma_k + [\mathcal{H}_{yx'}]^i_{\alpha\delta}[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{yx}]^i_{\alpha k}[\boldsymbol{g}_x]^\alpha_j + [\mathcal{H}_y]^i_\alpha[\boldsymbol{g}_{xx}]^\alpha_{jk}, \\
\mathcal{C} = & \Big\{[\mathcal{H}_{x'y'}]^i_{\beta\gamma}[\boldsymbol{g}_x]^\gamma_\delta[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{x'y}]^i_{\beta\gamma}[\boldsymbol{g}_x]^\gamma_k + [\mathcal{H}_{x'x'}]^i_{\beta\delta}[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{x'x}]^i_{\beta k}\Big\}[\boldsymbol{h}_x]^\beta_j + [\mathcal{H}_{x'}]^i_\beta[\boldsymbol{h}_{xx}]^\beta_{jk}, \\
\mathcal{D} = & [\mathcal{H}_{xy'}]^i_{j\gamma}[\boldsymbol{g}_x]^\gamma_\delta[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{xy}]^i_{j\gamma}[\boldsymbol{g}_x]^\gamma_k + [\mathcal{H}_{xx'}]^i_{j\delta}[\boldsymbol{h}_x]^\delta_k + [\mathcal{H}_{xx}]^i_{jk}, \\
& i = 1,\ldots n, \quad j,k,\beta,\delta = 1,\ldots n_x, \quad \alpha,\gamma = 1,\ldots n_y.
\end{aligned}
$$

上式中，所有 $\mathcal{H}$ 的导数，以及关于 $\boldsymbol{g}(\bar{\boldsymbol{x}};0)$ 和 $\boldsymbol{h}(\bar{\boldsymbol{x}};0)$ 的一阶导数，都是已知数。因此上式是一个 $n\times n_x\times n_x$ 个方程组成的系统，含有以 $\boldsymbol{g}_{xx}$ 和 $\boldsymbol{h}_{xx}$ 为元素的 $(n_x+n_y)\times n_x\times n_x$ 个未知数。

$F_{\sigma\sigma(\cdot)}$ 的求解

一阶导数 $F_\sigma(\cdot)=0$，式(6.16)的 LHS 表示为 6 个部分相加。因此二阶导数

$$
[F_{\sigma\sigma}(\bar{\boldsymbol{x}};0)]^i = \mathcal{A}+\mathcal{B}+\mathcal{C}+\mathcal{D}+\mathcal{E}+\mathcal{F}=0, \tag{6.23}
$$

其中

$$
\begin{aligned}
&\mathcal{A} = [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\beta [\boldsymbol{h}_{\sigma\sigma}]^\beta, \\
&\mathcal{B} = [\mathcal{H}_{y'y'}]^i_{\alpha\gamma} [\boldsymbol{g}_x]^\gamma_\delta [\eta]^\delta_\zeta [\boldsymbol{g}_x]^\alpha_\beta [\eta]^\beta_\phi [I]^\phi_\zeta + [\mathcal{H}_{y'x'}]^i_{\alpha\delta} [\eta]^\delta_\zeta [\boldsymbol{g}_x]^\alpha_\beta [\eta]^\beta_\phi [\varepsilon']^\phi, \\
&\mathcal{C} = [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_{\sigma\sigma}]^\alpha, \\
&\mathcal{D} = [\mathcal{H}_{y}]^i_\alpha [\boldsymbol{g}_{\sigma\sigma}]^\alpha, \\
&\mathcal{E} = [\mathcal{H}_{x'}]^i_\beta [\boldsymbol{h}_{\sigma\sigma}]^\beta, \\
&\mathcal{F} = [\mathcal{H}_{x'y'}]^i_{\beta\gamma} [\boldsymbol{g}_x]^\gamma_\delta [\eta]^\delta_\zeta [\eta]^\beta_\phi [I]^\phi_\zeta + [\mathcal{H}_{x'x'}]^i_{\beta\delta} [\eta]^\delta_\zeta [\eta]^\beta_\phi [I]^\phi_\zeta, \\
&i = 1 \ldots n, \quad \alpha, \gamma = 1, \ldots n_y, \quad \beta, \delta = 1, \ldots n_x, \ldots \phi, \zeta = 1, \ldots n_\varepsilon.
\end{aligned}
$$

上式是一个 n 个方程构成的线性系统，含有由 $\boldsymbol{g}_{\sigma\sigma}$ 和 $\boldsymbol{h}_{\sigma\sigma}$ 构成的 n 个未知数。

不为零的 $\boldsymbol{g}_{\sigma\sigma}$ 和 $\boldsymbol{h}_{\sigma\sigma}$ 意味着对风险的纠正 (correction for risk)：风险来自于经济系统对确定性等价的悖离。

$F_{\sigma x}(\cdot)$ 的求解

一阶导数 $F_\sigma(\bar{x};0) = 0$，式(6.16)进一补对 $\boldsymbol{x}$ 求导，可得二阶导数

$$
\begin{aligned}
&[F_{\sigma x}(\bar{\boldsymbol{x}};0)]^i_j = [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\beta [\boldsymbol{h}_{\sigma x}]^\beta j + [\mathcal{H}_{y'}]^i_\alpha [\boldsymbol{g}_x]^\alpha_\gamma [\boldsymbol{h}_x]^\gamma_j + [\mathcal{H}_{y}]^i_\alpha [\boldsymbol{g}_{\sigma x}]^\alpha_j + [\mathcal{H}_{x'}]^i_\beta [\boldsymbol{h}_{\sigma x}]^\beta_j = 0, \\
&i = 1, \ldots n, \quad \alpha = 1, \ldots n_y, \quad \beta, \gamma, j = 1, \ldots n_x,
\end{aligned} \tag{6.24}
$$

其中所有包括 $\boldsymbol{g}_\sigma(\bar{\boldsymbol{x}};0), \boldsymbol{h}_\sigma(\bar{\boldsymbol{x}};0)$ 的项都等于 0(确定性等价条件)。

上式是一个 $n \times n_x$ 个方程构成的线性系统，含有由 $\boldsymbol{g}_{\sigma x}$ 和 $\boldsymbol{h}_{\sigma x}$ 构成的 $n \times n_x$ 个未知数。系统成立的条件需要满足

$$
\begin{aligned}
\boldsymbol{g}_{\sigma x}\bar{\boldsymbol{x}};0 = 0, \\
\boldsymbol{h}_{\sigma x}\bar{\boldsymbol{x}};0 = 0.
\end{aligned} \tag{6.25}
$$

6.3.7 更高阶扰动

如前文所介绍的那样，在 1 阶扰动近似解的基础上，我们可以继续迭代，依次求得 2 3 . . . n 阶的近似解。如果 $\boldsymbol{g}$ 和 $\boldsymbol{h}$ 在 $\bar{\boldsymbol{x}}$ 附近有解，那么以 n 阶近似生成的一组变量时间序列将会有无限项，以某个半径 r 逐渐收敛到 $\bar{\boldsymbol{x}}$。

r 可以无限大，此时的变量序列有可能收敛到任意一处。r 也可以是个有限的值，此时变量序列的边界存在不可删除的奇异点。不幸的是，对大多数 DSGE 模型来说，收敛半径 r 的值是未知的，这方面尚需进一步深入的研究 (Swanson et al., 2006; Aldrich and Kung, 2011)。

即便该变量序列是收敛的，以下 2 个问题也值得重视。

1. 在一个 j 阶近似中，我们有可能会找不到 $\boldsymbol{g}$ 和 $\boldsymbol{h}$ 这两个决策式的“正确”形式。例如Aruoba et al. (2006) 的研究表明，在对一个随机内生经济增长模型做了 5 阶扰动近似后，所生成的两个近似决策式——消费的决定式和资本积累的决定式——不再是全局 concave 的，这与经济学理论的假设不符。并且它们都呈现出震荡特征 (oscillating patterns)。

2. 随着扰动近似阶数的不断增加，近似解向某一个特定点 (如稳定状态) 的收敛，可能不是单调的。事实上我们可以较容易地构建这样一个模型，来说明当经济运行状况远离稳定状态时，$j+1$ 阶扰动近似的结果可能比 j 阶近似的结果更差。

这 2 个问题都不是致命性的，但在实际研究开展过程中，它们的确值得研究者予以关注，并使用必要的检验方法，确保其负面效果最小化。例如，查验不同阶近似线性所生成的变量数据序列，与实际观测到的数据序列之间的近似程度。

在后文中我们将讨论如何评估不同阶近似解的精确程度，以及是否有必要采取更高阶扰动来求近似解。例如，为了更好处理含有时变波动性的模型，我们需要至少 3 阶，甚至更高阶的扰动近似。再如Levintal (2017) 便指出，对暗含灾难性风险的经济模型来说，至少 5 阶的扰动近似是必须的。

6.3.8 例

模型设定

以经典的随机 NCGT 模型为例，介绍二阶扰动法的应用。假定 $u(c_t) = \log c_t$，$\delta = 1$。经济模型构成一个非线性系统，由均衡条件 (跨期消费决策的 Euler 方程)，预算约束条件和外生技术过程组成：

$$
\begin{aligned}
&\frac{1}{c_t} = \beta E_t \frac{\alpha \exp(z_{t+1}) k_{t+1}^{\alpha-1}}{c_{t+1}}, \\
&c_t + k_{t+1} = \exp(z_t) k_t^{\alpha}, \\
&z_t = \rho z_{t-1} + \eta \varepsilon_t.
\end{aligned}
$$

$\delta = 1$ 使得技术冲击的收入效果和替代效果互相抵消，并且消费和投资相对于产出的比例恒定。在此基础上，我们可以得到关于消费和资本积累的最优决策方程解析形式：

$$
\begin{aligned}
&c_t = (1 - \alpha\beta) \exp(z_t) k_t^{\alpha}, \\
&k_{t+1} = \alpha\beta \exp(z_t) k_t^{\alpha}
\end{aligned}
\tag{6.26}
$$

现在，假定我们不知道解析形式的最优决策方程。构建猜测的最优决策方程为

$$
\begin{aligned}
&c_t = c(k_t, z_t), \\
&k_{t+1} = k(k_t, z_t).
\end{aligned}
$$

把猜测的决策方程代回均衡系统并做适当简化得

$$
\begin{aligned}
&\frac{1}{c(k_t, z_t)} = \beta E_t \frac{\alpha \exp\left(\rho z_t + \eta \varepsilon_{t+1}\right) k(k_t, z_t)^{\alpha-1}}{c\left(k(k_t, z_t), \rho z_t + \eta \varepsilon_{t+1}\right)}, \\
&c(k_t, z_t) + k(k_t, z_t) = \exp(z_t) \cdot k_t^{\alpha}.
\end{aligned}
\tag{6.27}
$$

根据扰动法，将猜测的决策方程 $c(\cdot), k(\cdot)$ 做关于扰动参数 σ 的近似：

$$
\begin{aligned}
&c_t = c(k_t, z_t; \sigma), \\
&k_{t+1} = k(k_t, z_t; \sigma),
\end{aligned}
$$

以及将扰动参数引入外生的技术过程

$$
z_t = \rho z_{t-1} + \sigma \eta \varepsilon_t.
$$

稳定状态

很显然，当 $\sigma = 0$ 或 $z_t = 0 \forall t$ 时，随机增长模型回复到传统的确定性增长模型中去[9]。此时系统有且只有一个稳定状态，对应 $(\bar{k}, \bar{c})$，满足

$$\begin{aligned} \bar{k} &= k(\bar{k}, 0; 0) = (\alpha\beta)^{\frac{1}{1-\alpha}}, \\ \bar{c} &= c(\bar{c}, 0; 0) = (\alpha\beta)^{\frac{\alpha}{1-\alpha}} - (\alpha\beta)^{\frac{1}{1-\alpha}}. \end{aligned} \tag{6.28}$$

对模型做扰动近似

对猜测的消费决策方程和猜测的资本存量决策方程分别做二阶扰动展开，我们有

$$\begin{aligned} c_t =& c(k_t, z_t; \sigma) \\ =& \underbrace{\bar{c}}_{0\text{ 阶}} + \underbrace{\left(c_k(k_t - \bar{k}) + c_z z_t + c_\sigma \sigma\right)}_{1\text{ 阶}} \\ &+ \underbrace{\frac{1}{2}c_{kk}\left(k_t - \bar{k}\right)^2 + c_{kz}(k_t - \bar{k})z_t + c_{k\sigma}(k_t - \bar{k})\sigma + \frac{1}{2}c_{zz}\left(z_t - \bar{z}\right)^2 + c_{z\sigma}z_t\sigma + \frac{1}{2}c_{\sigma\sigma}\sigma^2}_{2\text{ 阶}}. \end{aligned} \tag{6.29}$$

$$\begin{aligned} k_{t+1} =& k(k_t, z_t; \sigma) \\ =& \underbrace{\bar{k}}_{0\text{ 阶}} + \underbrace{\left(k_k(k_t - \bar{k}) + k_z z_t + k_\sigma \sigma\right)}_{1\text{ 阶}} \\ &+ \underbrace{\frac{1}{2}k_{kk}\left(k_t - \bar{k}\right)^2 + k_{kz}(k_t - \bar{k})z_t + k_{k\sigma}(k_t - \bar{k})\sigma + \frac{1}{2}k_{zz}\left(z_t - \bar{z}\right)^2 + k_{z\sigma}z_t\sigma + \frac{1}{2}k_{\sigma\sigma}\sigma^2}_{2\text{ 阶}}. \end{aligned} \tag{6.30}$$

与一阶近似线性展开相比，二阶展开项中包含关于经济系统的更多信息，是一阶近似所无法分析的，举例说明。

1. 风险纠正。$c_{\sigma\sigma}\sigma^2, k_{\sigma\sigma}\sigma^2$ 对应风险纠正 (correction for risk)，描述在面对不确定性时，经济系统的稳定能力。

2. 从对称到不对称。在一阶近似中，$c_z z_t$ 和 $k_z z_t$ 项暗示着，对系统施加同等程度的一个正冲击和一个负冲击，产生的经济效果互为镜像。正因如此，基于一阶线性近似的 DSGE 研究常常只报告正负冲击之一的冲击响应 (IRFs)：另一个冲击只有正负符号的不同。

   而二阶近似则不同。交互项 $c_{kz}(k_t - \bar{k})z_t$ 和 $k_{kz}(k_t - \bar{k})z_t$ 意味着，外部冲击的经济效果不只取决于外生随机冲击程度的大小 z_t，更与当期资本存量 k_t 有关，而后者是一个内生变量。经验研究表明，当衡量如一个金融部门外生冲击的效果时，很可能也需要考虑到同期经济体中的财产存量。

为了求解(6.29)-(6.30)，我们需要得到两组值。一组是变量的稳定状态，已于(6.28)得到。另一组是决策方程的一阶和二阶导数，需要我们回到非线性均衡条件系统(6.27)，改写为矩阵形式

$$\mathcal{H}(k_t, z_t; \sigma) = \begin{bmatrix} \frac{1}{c(k_t, z_t; \sigma)} - \beta E_t \frac{\alpha \exp(\rho z_t + \eta\varepsilon_{t+1}) k(k_t, z_t; \sigma)^{\alpha-1}}{c(k(k_t, z_t; \sigma), \rho z_t + \eta\varepsilon_{t+1})} \\ c(k_t, z_t; \sigma) + k(k_t, z_t; \sigma) - \exp(z_t) \cdot k_t^\alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{6.31}$$

[9]对于 $z_t = 0 \forall t$ 的情况，可能由以下两个因素之一导致：初始状态 $z_0 = 0$，或者随着 $t \to \infty$, $z_t \to 0$。

或者代入近似系统的运算符 $F(\cdot)$，定义为关于 $\mathcal{H}(\cdot)$ 的紧凑形式

$$F(k_t, z_t; \sigma) = \mathcal{H}\left(\underbrace{c(k_t, z_t; \sigma)}_{1}, \underbrace{c(k(k_t, z_t; \sigma), z_{t+1}; \sigma)}_{2}, \underbrace{k_t}_{3}, \underbrace{k(k_t, z_t; \sigma)}_{4}, \underbrace{z_t}_{5}; \underbrace{\sigma}_{6}\right), \tag{6.32}$$

其中下角标编号 1 至 6 依次表示运算符 $\mathcal{H}$ 中的元素。

一阶展开

围绕稳态 $(\bar{k}, 0; 0)$，对非线性系统 $F(k_t, z_t; \sigma)$ 求一阶偏导

$$\begin{aligned}
F_k &= \mathcal{H}_1 c_k + \mathcal{H}_2 c_k k_k + \mathcal{H}_3 + \mathcal{H}_4 k_k = 0, \\
F_z &= \mathcal{H}_1 c_z + \mathcal{H}_2 \left(c_k k_z + c_z \rho\right) + \mathcal{H}_4 k_z + \mathcal{H}_5 = 0, \\
F_\sigma &= \mathcal{H}_1 c_\sigma + \mathcal{H}_2 \left(c_k k_\sigma + c_\sigma\right) + \mathcal{H}_4 k_\sigma + \mathcal{H}_6 = 0,
\end{aligned}$$

其中 $\mathcal{H}_i, i = 1, 2 \ldots 6$ 表示 $\mathcal{H}$ 对第 i 个元素求偏导。

这构成了 6 个方程组成的系统[10]，对应 6 个系数 $\mathcal{H}_i, i = 1, 2 \ldots 6$，6 个代求解未知数 $c_k, c_z, k_k, k_z, c_\sigma, k_\sigma$。

求解过程分三步。第一步，求 c_k,k_k。提取 $F_k = 0$ 和 $F_z = 0$ 对应的 4 个等式构成一个子系统，写为二次矩阵形式

$$\begin{bmatrix} \mathcal{H}_1^1 \\ \mathcal{H}_1^2 \end{bmatrix} c_k + \begin{bmatrix} \mathcal{H}_2^1 \\ \mathcal{H}_2^2 \end{bmatrix} c_k k_k + \begin{bmatrix} \mathcal{H}_3^1 \\ \mathcal{H}_3^2 \end{bmatrix} + \begin{bmatrix} \mathcal{H}_4^1 \\ \mathcal{H}_4^2 \end{bmatrix} k_k = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

由(6.31)的上半部分和下半部分可得，$\mathcal{H}_3^1 = 0$，$\mathcal{H}_2^2 = 0$，上式重写为

$$\begin{bmatrix} \mathcal{H}_1^1 \\ \mathcal{H}_1^2 \end{bmatrix} c_k + \begin{bmatrix} \mathcal{H}_2^1 \\ 0 \end{bmatrix} c_k k_k + \begin{bmatrix} 0 \\ \mathcal{H}_3^2 \end{bmatrix} + \begin{bmatrix} \mathcal{H}_4^1 \\ \mathcal{H}_4^2 \end{bmatrix} k_k = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{6.33}$$

(6.33)中，提取下半部分

$$\begin{aligned}
&\mathcal{H}_1^2 c_k + \mathcal{H}_3^2 + \mathcal{H}_4^2 k_k = 0, X \\
&\Rightarrow c_k = \frac{1}{\mathcal{H}_1^2}\left(\mathcal{H}_4^2 k_k\right) - \frac{\mathcal{H}_3^2}{\mathcal{H}_1^2},
\end{aligned}$$

代回上半部分，替换其中的 c_k，我们有

$$\begin{aligned}
&\mathcal{H}_1^1 c_k + \mathcal{H}_2^1 c_k k_k + \mathcal{H}_4^1 k_k = 0, \\
&\Rightarrow -\frac{\mathcal{H}_2^1 \mathcal{H}_4^2}{\mathcal{H}_1^2} k_k + \left(\mathcal{H}_4^2 - \frac{\mathcal{H}_1^1 \mathcal{H}_4^2 + \mathcal{H}_2^1 \mathcal{H}_3^2}{\mathcal{H}_1^2}\right) k_k - \frac{\mathcal{H}_1^1 \mathcal{H}_3^2}{\mathcal{H}_1^2} = 0.
\end{aligned}$$

上式成为一个二次矩阵形式 $AP^2 - BP - C = 0$，对应 $k_k = P$。采用Uhlig (1999) 的方法 (见 Remark 6.4)，我们可以求得 k_k 的值，进而 c_k。

此外需要指出的是，二次矩阵系统的 P 值有两个，对应两种情况。

1. $k_k > 1$，此时的系统解是不稳定解。回顾一下资本的决策方程：

$$k_{t+1} = \bar{k} + k_k\left(k_t - \bar{k}\right) \ldots,$$

很显然如果 $k_k > 1$，那么 t 期变量与稳态的偏离程度 $k_t - \bar{k}$，会导致 $t+1$ 期变量距稳态更大程度的偏离。

[10] F 有 2 个维度，3 个一阶导数对应 6 个方程。

2. $0 < k_k < 1$，此时的系统解是稳定解，t 期变量距稳态的偏离，在不考虑其他冲击影响的情况下，会在随后时间里逐渐收敛至稳态。

第二步，将求得的 c_k 和 k_k 代入 $F_z = 0$ 的 2 个方程中，重复上一步的方法，求得 c_z 和 k_z 的值。需要指出的是，由 $F_z = 0$ 这 2 个方程所构成的子系统是线性的。

第三步，将求得的 c_k, k_k, c_z, k_z 代入 2 个方程 $F_\sigma = 0$ 中，构成一个线性齐次的关于 c_σ 和 k_σ 组成的系统。因此，系统解为 $c_\sigma = 0$ 和 $k_\sigma = 0$。进而，我们得到一阶扰动法线性近似下的确定性等价条件。

二阶展开

下面，对 $F(k_t, z_t; \sigma)$ 围绕稳态 $(\bar{k}, 0; 0)$ 做二阶线性近似，得到一组 $6 \times 2 = 12$ 个方程组成的系统

$$\begin{cases} F_{kk} = 0, \\ F_{kz} = F_{zk} = 0, F_{k\sigma} = F\sigma k = 0, \\ F_{zz} = 0, \\ F_{z\sigma} = F_{\sigma z} = 0, \\ F_{\sigma\sigma} = 0, \end{cases} \tag{6.34}$$

其中的系数部分包括：

- 稳态值：$(\bar{c}, \bar{k})$，已经求得。
- 一阶导数的系数，包括两部分：
  - (c_k, k_k, c_z, k_z)，已在前文中做一阶展开部分求得，是已知数。
  - $c_\sigma = k_\sigma = 0$，确定性等价条件。
- 二阶导数的系数，包括两部分：
  - $(c_{k\sigma}, k_{k\sigma}, c_{z\sigma}, k_{k\sigma},)$，所有二次求导中含有 σ 的交互项，它们的值都为 0；
  - $(c_{kk}, k_{kk}, c_{kz}, k_{kz}, c_{zz}, k_{zz}, c_{\sigma\sigma}, k_{\sigma\sigma})$，是代求解系数。

基于上述分析，将一系列等于 0 的条件代回消费和资本的最优决策式(6.29)-(6.30)，我们有近似的最优决策式：

$$\begin{aligned} c_t =& \bar{c} + c_k\left(k_t - \bar{k}\right) + c_z z_t \\ &+ \frac{1}{2}c_{kk}\left(k_t - \bar{k}\right)^2 + c_{kz}\left(k_t - \bar{k}\right) z_t + \frac{1}{2}c_{zz}z_t^2 + \frac{1}{2}c_{\sigma\sigma}\sigma^2, \end{aligned} \tag{6.35}$$

$$\begin{aligned} k_{t+1} =& \bar{k} + k_k\left(k_t - \bar{k}\right) + k_z z_t \\ &+ \frac{1}{2}k_{kk}\left(k_t - \bar{k}\right)^2 + k_{kz}\left(k_t - \bar{k}\right) z_t + \frac{1}{2}k_{zz}z_t^2 + \frac{1}{2}k_{\sigma\sigma}\sigma^2. \end{aligned} \tag{6.36}$$

数值解

对于哪怕一个基准的随机经济增长模型，利用扰动法得到的近似系统(6.35)-(6.36)也已经较为复杂，难以表示为解析形式了。因此我们求助于数值方法。下面分几个步骤分别介绍。第一步，参数校准。采用文献中常见的校准值。

- $\delta = 1$。折旧系数。如前文所述，这使得模型得以简化，但也因此使得模型经济与任何实际经济有较大不同。
- $\beta = 0.99$。时间贴现。
- $\alpha = 0.33$。资本的产出弹性。
- $\rho = 0.95$。外生生产率自回归过程中的持续系数 (persistence parameter)。
- $\eta = 0.01$。生产率自回归过程中，波动的标准差。

第二步，计算稳定状态。将校准参数值代入(6.28)中得 $\bar{c} = 0.388$，$\bar{k} = 0.188$。
第三步，计算一阶、二阶扰动解。

- 一阶扰动解。$c_k = 0.680$，$c_z = 0.388$，$k_k = 0.330$，$k_z = 0.188$。
- 二阶扰动解。$c_{kk} = -2.420$，$c_{kz} = 0.680$, $c_{zz} = 0.388$，$c_{\sigma\sigma} = 0$，$k_{kk} = -1.174$，$k_{kz} = 0.330$，$k_{zz} = 0.188$，$k_{\sigma\sigma} = 0$。其中 $c_{\sigma\sigma} = k_{\sigma\sigma} = 0$ 的情况，我们将在下文中做进一步说明，见。 reference
- 此外我们有 $c_\sigma = k_\sigma = c_{k\sigma} = k_{k\sigma} = c_{z\sigma} = k_{z\sigma} = 0$ (理论模型中的确定性等价条件)。

第四步，关于消费和资本存量的近似决策式可由(6.35)-(6.36)求得。

Remark 6.7 (风险纠正系数的讨论)**.** *数值模拟过程中，两个风险纠正的系数值都是 0，对应 $(1/2)c_{\sigma\sigma}\sigma^2 = (1/2)k_{\sigma\sigma}\sigma^2 = 0$。这个看似奇怪的结果是和标准的新古典主义经济增长模型的特性有关的：模型中的风险主要来自外生技术冲击造成的生产过程风险。面对这种风险，一方面从局部部门来说，典型家庭积累的资本越多，他所面对的风险就越大，另一方面从整体来说，经济体中唯一可以用作净储蓄的财产就是资本。因此，任何风险的增大 (即技术冲击的标准差的增大) 产生了两种相互对冲的效果：一方面尽量减缓资本积累以免受到明日生产过程风险的损害，另一方面尽量增加储蓄 (资本积累) 以应付未来可能出现的负面冲击。*

两种效果对冲后对经济运行的影响，分两种情况来分析。第一种情况，如果构建经济模型时设定较低的风险厌恶度，对应 CRRA 效用函数如对数形式，则二者完全抵消，系统解只收到技术冲击的波动 z_t 影响，而不受到波动的标准差的影响。第二种情况，当风险厌恶度较高，或者存在着多种财产储蓄形式时——如模型中本国典型家庭还有另一种储蓄选择：购买外国债券，并且外国债券的收益与本国技术冲击之间并非完全替代关系——则风险纠正项便很可能是个显著不为 0 的值了。

第五步，比较解析决策式与近似决策式。一个自然出现的问题就是：采用扰动法近似的决策式 (近似系统解)，在多大程度上贴近原系统的解？我们用解析决策式来代表原系统，可得

$$c_t = 0.6734 \exp(z_t) k_t^{0.33},$$
$$k_{t+1} = 0.327 \exp(z_t) k_t^{0.33}.$$

画出解析决策式、一阶扰动近似决策式、二阶扰动近似决策式，不难看出：

1. 在稳态 ($\bar{k} = 0.188$) 附近，一阶近似的解几乎与解析解完全一致，但当 k_t 越是远离 $\bar{k}$，二者的差异越明显：当 $k_t = 0.1412$ 时，偏差约 1%。
2. 比较起来，二阶近似解在全局内，与解析解相比更加贴近，如 $k_t = 0.1412$ 时，偏差约 0.13%。

这表明，扰动法，尤其是高阶扰动法提供的近似解，具有更好的全局特性。

Remark 6.8 (扰动近似解的精确度检验)**.** 0.13% 的偏差程度已经足够精确了，还是仍然太高？问题的答案取决于实际研究工作中的具体情况。例如，在商业周期研究中的精确度要求，通常低于关于福利水平的研究。其原因在于，在近似计算一系列矩的方差时 (如消费的均值和方差等)，随着矩的不同，矩的误差也有所不同，各项误差之间有时会相互抵消一部分。而福利水平常常是一个关于财富分配的非线性方程，在近似计算分配状况时，一点微小的误差可能会引起福利水平近似值的较大变化，从而偏离实际值。

6.4 剪枝算法

如前文所述，通常来讲，越高阶的扰动近似解，与原非线性系统的贴近程度越高。但高阶扰动法带来的一个附属问题是，在实际研究中，尽管高阶扰动近似求得的线性决策方程的确是稳定的，但根据这个决策方程所生成的序列数据却有可能是不稳定的 (爆炸的)。其原因在于，高阶扰动项在系统中加入了额外的不动点，从而使得围绕这些不动点所做的近似解变得不稳定了 (Kim et al., 2008; Den Haan and De Wind, 2012)。

举例说明。假设某一资本的近似决策方程可表示为

$$k_{t+1} = a_0 + a_1 k_t + a_2 k_t^2 + \ldots + b_1 \varepsilon_t + \ldots,$$

为简化模型考虑，不列出技术外生冲击 z_t 的情形。以递归形式表述，上式可以改写为

$$k_{t+1} = a_0 + a_1 \left(a_0 + a_1 k_{t-1} + a_2 k_{t-1}^2\right) + a_2 \left(a_0 + a_1 k_{t-1} + a_2 k_{t-1}^2\right)^2 + \ldots + b_1 \varepsilon_t + \ldots.$$

上式中涉及到 k_{t-1} 的 3 次甚至 4 次方。因此在仿真过程中，随着 t 越来越大，所生成的时间序列数据就越容易出现 $\left(k_{t+1} - \bar{k}\right) \to \infty$ 的情况，即产生不稳定的爆炸路径。在这种情况下，GMM、SMM 等无条件的矩匹配估计方法也不再适用了[11]，原因在于它们需要满足一系列条件，如矩是平稳且有极限的，并且是非周期的遍历 (ergodic) 概率分布。

鉴于这种情况，Kim et al. (2008) 提出要对近似解做剪枝 (pruning)，即是说，在递归形式中去掉所有比近似解更高阶的项——以本例中的二阶扰动近似法为例，就是要删除一切高于二阶的状态项和/或扰动项。Kim et al. (2008) 证明由剪枝后的近似解所生成的时间序列数据，就不再出现爆炸路径的情况。

Andreasen et al. (2016) 对Kim et al. (2008) 的剪枝法做了进一步扩展，应用到了任意高的阶数的情况。他们首先证明一个剪枝状态——空间系统中的确存在一阶以及二阶无条件矩，进而提供了关于这些一、二阶矩与冲击响应方程 (IRFs) 的闭合表达式 (closed-form expressions)。其研究价值在于，可以省去研究者计算数值并进行仿真的工作：这些数据仿真已经被不少 DSGE 的扩展 IRFs 研究证明是不可靠的了[12]。此外，Andreasen et al. (2016) 还讨论了更高阶无条件矩阵 (如偏度 skewness，峰度 kurtosis) 等的存在条件。

6.5 变量的变换

上文中介绍了，在扰动法求解 DSGE 模型过程中，可以用对数形式的变量而非变量本身作为研究对象。Jin and Judd (2002) 介绍了对数线性化方法，作为一种变量的特殊变化形式，为什么要比变量本身更适合作为 DSGE 模型的研究对象，以及更通用的变量变换形式是怎么样的，如何

[11] 关于 GMM，SMM 等方法在 DSGE 研究中的应用，可见Ruge-Murcia (2007)。

[12] 关于扩展 IRFs 的定义，可见如Koop et al. (1996)。

更好的运用他们，更多讨论见Fernández-Villaverde and Rubio-Ramírez (2006)。这里简要介绍变量变换的通用方法及其用途。

已知对于 $d(x)$，围绕 a 点做 TSE

$$d(x) \approx d(a) + \frac{\partial d(a)}{\partial a}(x-a) + H.O.T.$$

其中 H.O.T. 表示高阶展开项。据此，也可以将变量从 x 转换为反函数 $Y(x)$，围绕 $b = Y(a)$ 点做 TSE

$$g(y) = h(d(x(y))) = g(b) + \frac{\partial g(b)}{\partial b}(Y(x) - b) + H.O.T.$$

根据第一个等式，我们可以通过 TSE 求得关于状态 x 的近似方程 $d(\cdot)$，进而求解运算符 $\mathcal{H}(\cdot)$。那么同样地，在对 x 做变量变换后，根据第二个等式我们也能找到关于状态 $Y(x)$ 的未知方程 $g(\cdot)$，进而求解运算符 $\mathcal{H}(\cdot)$。在进行近似线性化过程之前，变量变换工作首先将重心放在状态形式的选取上来，将原本是高度线性化的问题转化为接近线性的，从而提高扰动法求解的精度。

6.5.1　例 1

以上文提到的随机内生经济增长模型为例，求资本积累的近似决定式，假定满足如下一阶扰动形式

$$k_{t+1} = \bar{k} + a_1\left(k_t - \bar{k}\right) + b_1 z_t,$$

其中 a, b 系数通过对 $F(k_t, z_t; \sigma)$ 求导得出，稳态 $\bar{k}$ 也是已知。上式调整，化为总量形式

$$\left(k_{t+1} - \bar{k}\right) = a_1\left(k_t - \bar{k}\right) + b_1 z_t. \tag{6.37}$$

类似地，对(6.37)做对数线性化

$$\begin{aligned} &\log k_{t+1} - \log \bar{k} = a_2\left(\log k_t - \log \bar{k}\right) + b_2 \cdot z_t, \\ &\Rightarrow \hat{k}_{t+1} = a_2 \hat{k}_t + b_2 z_t, \quad \hat{x}_t \equiv \log x_t - \log x_0. \end{aligned} \tag{6.38}$$

Proposition 6.1. (6.37)和(6.38)等价。

证明. 分三步予以证明。第一，已知变量的线性近似形式

$$k_{t+1} = d(k_t, z_t; \sigma) = d(\bar{k}, 0; 0) + d_1(\bar{k}, 0; 0)\left(k_t - \bar{k}\right) + d_2(\bar{k}, 0; 0) z_t,$$

则我们有

$$d_1(\bar{k}, 0; 0) = a_1, \quad d_2(\bar{k}, 0; 0) = b_1.$$

第二步，引入变量的变化，设 $h = \log d$，其中 $Y(x) = \log x$。则根据Judd and Guu (2003) 我们有

$$\log k_{t+1} - \log \bar{k} = d_1(\bar{k}, 0; 0)\left(\log k_t - \log \bar{k}\right) + \frac{1}{\bar{k}} d_2(\bar{k}, 0; 0) z_t.$$

第三步，比较上面两个等式，系数相等，我们有

$$a_2 = a_1, \quad b_2 = \frac{1}{\bar{k}} b_1.$$

□

这个例子中有 3 点值得注意。

1. 从 (a_1, b_1) 到 (a_2, b_2) 的计算涉及到 $\bar{k}$，而 $\bar{k}$ 已经在对水平变量一阶扰动法求解中求得。因此，当我们求得了一节线性系统的近似解后，可以很方便得到变量变换 (对数线性化) 后的系统近似解。

2. 对于含有常规设定 (如效用函数、生产函数等) 的随机内生经济增长模型而言，通常来说无需引入其他额外加设，即可进行变量变换如对数线性化。

3. 变量变换可用于任何阶数的扰动近似。

6.5.2 例 2

我们来看一个更通用的例子。已知某一系统围绕稳态 $(x = a)$ 的一阶扰动近似解为

$$d(x) \approx d(a) + \frac{\partial d(a)}{\partial a}(x - a),$$

那么对于 $Y(x)$ 的反函数 $X(y)$，围绕 $b = Y(a)$，对变量变换后的新系统 $g(y) = h(d(X(y)))$ 做一阶扰动，我们有

$$g(y) = h(d(X(y))) = g(b) + g_\alpha(b)\left(Y(x)^\alpha - b^\alpha\right),$$

其中 $g_\alpha = [h]_A [d_i]^A [x]^i_\alpha$ 为张量形式，可由微积分运算中的链式法则求得。

根据Judd and Guu (2003) 的方法，我们可以把任一幂的近似式写为

$$k_{t+1}(k_t, z_t; \gamma, \zeta, \varphi)^\gamma - \bar{k}^\gamma = a_3\left(k_t^\zeta - \bar{k}^\zeta\right) + b_3 z_t^\varphi,$$

其中设 $\varphi \geq 1$ 以确保 z_t^φ 是实数。

幂形式方程的优点在于，我们只需要 3 个参数值 γ, ζ, φ 就可以描述许多种非线性结构的系统。并且对于 $\gamma \to 0, \zeta \to 0, \varphi \to 1$ 的情况，该方程求极限就变成了对数转换。此时作变量转换 $h = d^\gamma, y = x^\gamma, x = y^{\frac{1}{\zeta}}$。此外，由于

$$k_{t+1}\left(k_t, z_t; \gamma, \zeta, \varphi\right)^\gamma - \bar{k}^\gamma = \frac{\gamma}{\zeta}\bar{k}^{\gamma-\zeta}a_1\left(k_t^\zeta - \bar{k}^\zeta\right) + \frac{\gamma}{\varphi}\bar{k}^{\gamma-1}z_t^\varphi,$$

于是我们有

$$a_3 = \frac{\gamma}{\zeta}\bar{k}^{\gamma-\zeta}a_1, \quad b_3 = \frac{\gamma}{\varphi}\bar{k}^{\gamma-1}b_1.$$

对于 $\gamma = \zeta, \varphi = 1$ 的情况，系统中只有一个自由系数，则

$$k_{t+1}(k_t, z_t; \gamma)^\gamma - \bar{k}^\gamma = a_4\left(k_t^\zeta - \bar{k}^\zeta\right) + b_4 z_t,$$

定义 $\tilde{k}_{t+1} \equiv k_t^\gamma - \bar{k}^\gamma$，我们于是有

$$\tilde{k}_{t+1} = a_4\tilde{k}_t + b_4 z_t, \tag{6.39}$$

因此

$$a_4 = a_1, \quad b_4 = \bar{k}^{\gamma-1}b_1.$$

通过变量变换，将原本是非线性的方程系统转换为线性形式(6.39)，这有助于我们展开后续分析，以及采用卡曼滤波对模型做进一步的估计。

ference

6.5.3 最优变量变换：参数的选取

我们介绍了如何通过几个幂参数值的设定，对变量做变换，将 DSGE 模型用更通用的一阶线性形式表现出来。随后的问题就变成了，如何选取最合适的幂参数。参数取值可以遵循以下两个策略之一。第一是最优策略，致力于追求最高的精确度，基于似然方程等方法的经验研究，目标的确为追求最大化解，对应最高的精确度，但这往往导致计算成本过于高昂。第二是次优策略，将参数取值向最优方向做适度改进，以追求模型精度的适度提高。实际经验研究中往往采取后者，在精度提升和计算成本控制之间寻求平衡点。

有鉴于此，Fernández-Villaverde and Rubio-Ramírez (2006) 提出一个合理判定原则，致力于选取合适的参数值来改变变量值，通过尽可能减少 Eruler 方程的误差来提高模型解的精确度[13]。

6.5.4 对数线性化和对数正态线性化

如前文所述，标准的对数线性近似法，常常是对差分变量取对数后做一阶扰动近似。金融学研究中有时也会遇到另一种对数正态线性化方法 (lognormal-loglinearization)，二者有所不同，容易引起混淆，有必要建模做简要说明。

假定经济模型中一个代表性家庭，效用函数和预算约束条件分别为

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \log C_t,$$
$$W_{t+1} = R_{t+1}(W_t - C_t),$$

其中 W_t 为总财富，其初始水平 W_0 是给定的。

非线性均衡系统由一阶条件和预算约束条件组成：

$$\frac{1}{C_t} = \beta E_t \frac{1}{C_{t+1}} R_{t+1},$$
$$W_{t+1} = R_{t+1}\left(W_t - C_t\right).$$

因此稳定状态下我们有

$$R = \frac{1}{\beta},$$
$$W = \frac{1}{\beta}(W - C).$$

(对数偏移形式的) 对数线性化

定义 $\hat{x}_t = x_t - x = \log X_t - \log X$ 表示对数变量 x_t 距离其稳态 x 的偏离。则对于 Euler 方程来说

$$\begin{aligned}
&\log C_{t+1} = \log C_t + \log R_{t+1} + \log \beta, \\
&\Rightarrow \log C + \left(\frac{C_{t+1} - C}{C}\right) = \log C + \left(\frac{C_t - C}{C}\right) + \log R + \left(\frac{R_{t+1} - R}{R}\right), \\
&\Rightarrow \hat{c}_{t+1} = \hat{c}_t + \hat{r}_{t+1}, \\
&\Rightarrow E_t \Delta \hat{c}_{t+1} = E_t \hat{r}_{t+1}.
\end{aligned}$$

[13]此外他们还发现，最优参数值的选取还与经济模型外生冲击的标准差有关。这是一个重要发现：变量变换随着模型中不确定性程度的变化而调整。这又一次印证了确定性等价条件的不适用。

对于预算约束条件来说

$$
\begin{aligned}
&\log W_{t+1} = \log R_{t+1} + \log(W_t - C_t),\\
&\Rightarrow \log W + \left(\frac{W_{t+1} - W}{W}\right) = \log R + \left(\frac{R_{t+1} - R}{R}\right) + \log(W - C) + \left(\frac{(W_t - W)(C_t - C)}{W - C}\right),\\
&\Rightarrow E_t \hat{w}_{t+1} = E_t \hat{r}_{t+1} + \frac{\hat{w}_t w - \hat{c}_t c}{w - c},\\
&\Rightarrow E_t \Delta \hat{w}_{t+1} = E_t \hat{r}_{t+1} + \left(1 - \frac{1}{\rho}\right)(\hat{c}_t - \hat{w}_t), \quad \rho \equiv \frac{w - c}{w}.
\end{aligned}
$$

(对数形式的) 对数线性化

由上两式可得，对数形式的线性化

$$E_t \Delta c_{t+1} = E_t r_{t+1} + \log \beta, \tag{6.40}$$

$$E_t \Delta w_{t+1} = E_t r_{t+1} + \left(1 - \frac{1}{\rho}\right)(c_t - w_t) + \kappa, \quad \kappa \equiv -r - \left(1 - \frac{1}{\rho}\right)(c - w). \tag{6.41}$$

对数正态线性化

对数正态线性化下的预算约束条件式与对数线性化下的(6.41)一致。但在 Euler 方程上的设定有所不同：它假定变量 $\frac{C_t}{C_{t+1}} R_{t+1}$ 是对数正态分布的。对于任何一个对数正态分布的变量我们有

$$\log E_t X_t = E_t \log X_t + \frac{1}{2} Var_t \log X_t.$$

那么回到 Euler 方程我们有

$$
\begin{aligned}
&1 = \beta E_t \left(\frac{C_t}{C_{t+1}}\right) R_{t+1},\\
&\Rightarrow 0 = \log \beta + \log E_t \left(\frac{C_t}{C_{t+1}}\right) R_{t+1},\\
&\Rightarrow 0 = \log \beta + E_t \log \left(\frac{C_t}{C_{t+1}}\right) R_{t+1} + \frac{1}{2} Var_t \log \left(\frac{C_t}{C_{t+1}}\right) R_{t+1},\\
&\Rightarrow E_t \Delta c_{t+1} = \log \beta + E_t r_{t+1} + \underbrace{\frac{1}{2}\left[var_t \Delta c_{t+1} + Var_t r_{t+1} - 2Cov\left(\Delta c_{t+1}, r_{t+1}\right)\right]}_{\chi}.
\end{aligned} \tag{6.42}
$$

由上可见，在对数正态线性化中，对于不含有期望的项，我们用标准对数线性化方法来处理；对于含有期望的项，我们假定它呈对数正态分布，其结果是一个确定的值，不涉及到泰勒级数展开的近似处理。

比较

比较(6.40)和(6.42)可见，对数正态线性化中多出了一项 χ，使得确定性等价条件不成立。这一新特性使对数正态线性化具有易于解释的优点。如对于一种资产 i 和定价核 (pricing kernel)M_t 来说[14]，则定价方程

$$
\begin{aligned}
&1 = E_t M_{t+1} R_{i,t+1},\\
&\Rightarrow 0 = E_t \log\left(M_{t+1} R_{i,t+1}\right) + \frac{1}{2} Var_t \log\left(M_{t+1} R_{i,t+1}\right),\\
&\Rightarrow E_t r_{i,t+1} = -E_t m_{t+1} - \frac{1}{2} Var_t m_{t+1} - Cov(m_{t+1}, r_{i,t+1}) - \frac{1}{2} Var_t r_{i,t+1}.
\end{aligned}
$$

[14]关于定价核的介绍，可见如 (Cochrane, 2000, Ch.1)。

此外对于无风险债券 f 来说，有

$$1 = E_t M_{t+1} R_{f,t+1},$$
$$r_{f,t+1} = -E_t m_{t+1} - \frac{1}{2} Var_t m_{t+1}.$$

则两式相减我们可得回报

$$E_t r_{i,t+1} - r_{f,t+1} = -\frac{1}{2} Var_t r_{i,t+1} - Cov(m_{t+1}, r_{i,t+1}).$$

但对数正态线性化也存在一定不足：

1. 我们不清楚在一个一般均衡模型中，变量 $\frac{C_t}{C_{t+1}} R_{t+1}$ 在多大程度上是接近对数正态分布的。
2. 对数正态线性化过程中同时使用了两种方法，分别是对数正态假设和对数线性化。从扰动法的角度来看，这种处理方法并非逻辑自洽，缺乏坚实的理论支撑，在经验研究中也无法确保呈收敛态势。
3. 对数线性化中我们可以通过求解二次矩阵系统来计算全部相关系数。而在对数正态线性化中，只能通过计算二阶矩的状态来估计系数；然而对许多研究来说，如何计算这些二阶距，这是个问题。
4. 对数线性化很容易做更高阶的扰动近似，但对数正态线性化则很难做高阶近似。

6.6　价值方程迭代中的扰动法

某些情况下我们需要在 DSGE 模型中对价值方程做扰动，如当偏好呈现出递归特征时，或者当我们需要对福利效果做评估时。这个扰动的价值方程，可以用作很好的初始值猜测，供我们展开随后的价值方程迭代——如前文所述，在一些高维问题中，如果缺乏合适的初始值设定，跌倒过程的结果可能导致收敛过慢，甚至不收敛。这里举例说明对将扰动法应用到价值方程迭代过程中去。

考虑这样一个价值方程问题[15]：

$$V(k_t, z_t) = \max_{c_t} \left[(1-\beta) \log c_t + \beta E_t V(k_{t+1}, z_{t+1}) \right],$$

约束条件及外部冲击过程

$$c_t + k_{t+1} = \exp(z_t) k_t^{\alpha} + (1-\delta) k_t,$$
$$z_t = \rho z_{t-1} + \eta \varepsilon_t, \quad \varepsilon_t \sim i.i.d\mathcal{N}(0,1).$$

引入扰动参数 σ，将上式改写为

$$V(k_t, z_t; \sigma) = \max_{c_t} \left[\log c_t + \beta E_t V\left(\exp(z_t) k_t^{\alpha} + (1-\delta) k_t - c_t, \rho z_t + \sigma \eta \varepsilon_{t+1}; \sigma \right) \right],$$

围绕确定性稳态 $(\bar{k}, 0; 0)$ 做二阶扰动近似，该问题的解由两部分构成：1 个价值方程 $V(k_t, z_t; \sigma)$ 和 1 个消费的策略方程 $c(k_t, z_t; \sigma)$。

先来看价值方程，

$$\begin{aligned} V(k_t, z_t; \sigma) =& V_{ss} + V_{1,ss}(k_t - \bar{k}) + V_{2,ss} z_t + V_{3,ss} \sigma \\ &+ \frac{1}{2} V_{11,ss} (k_t - \bar{k})^2 + V_{12,ss} (k_t - \bar{k}) z_t + V_{13,ss} (k_t - \bar{k}) \sigma \\ &+ \frac{1}{2} V_{22,ss} z_t^2 + V_{23,ss} z_t \sigma + \frac{1}{2} V_{33,ss} \sigma^2, \end{aligned}$$

[15] 出于简化模型的考虑，上式中将 $\log c_t$ 标准化 (乘以 $(1-\beta)$)，这样价值方程的稳态 $V_{ss} = \log c$，其中 c 对应稳态消费水平。通过标准化，价值方程和模型的效用函数的单位相同。

其中

$$
\begin{aligned}
&V_{ss} = V(\bar{k}, 0; 0), \\
&V_{i,ss} = V_i(\bar{k}, 0; 0), \quad i = 1, 2, 3, \\
&V_{ij,ss} = V_{ij}(\bar{k}, 0; 0), \quad i, j = 1, 2, 3, \\
&V_{3,ss} = V_{13,ss} = V_{23,ss} = 0, \quad \text{确定性等价条件}.
\end{aligned}
$$

由此价值方程的二阶扰动近似可以表示为

$$
\begin{aligned}
V(k_t, z_t; \sigma) =& V_{ss} + V_{1,ss}(k_t - \bar{k}) + V_{2,ss} z_t \\
&+ \frac{1}{2} V_{11,ss}(k_t - \bar{k})^2 + V_{12,ss}(k_t - \bar{k}) z_t + \frac{1}{2} V_{22,ss} z_t^2 + \frac{1}{2} V_{33,ss} \sigma^2,
\end{aligned} \tag{6.43}
$$

其中 $V_{33,ss} \neq 0$，这与前面介绍随机新古典主义经济增长模型中 Remark 6.7中，对效用函数做扩展 Schur 分解 (LQ 近似) 的情况不同。

类似地，消费的策略方程

$$
c_t = c(k_t, z_t; \sigma) = c_{ss} + c_{1,ss}(k_t - \bar{k}) + c_{2,ss} z_t + c_{3,ss} \sigma, \tag{6.44}
$$

其中

$$
\begin{aligned}
&c_{ss} = c(\bar{k}, 0; 0), \\
&c_{i,ss} = c_i(\bar{k}, 0; 0), \quad i = 1, 2, 3, \\
&c_{3,ss} = 0.
\end{aligned}
$$

$c_{3,ss} = 0$ 是由于，效用函数的一阶导数只与 $V_{1,2}$ 有关[16]。

对(6.43)和(6.44)构成的系统求解，遵循常规算法：

- 对线性项的系数，用 V_{ss} 依次对控制 c_t，状态 k_t, z_t 和扰动参数 σ 求导，随后在 $\sigma = 0$ 的情况下求解。
- 对二次项的系数，用 V_{ss} 做二次求导，引入上一环节求得的一阶系数，然后在 $\sigma = 0$ 的情况下求解。

对价值方程做扰动近似，有如下若干有点

1. 可以测度经济周期波动产生的福利成本。在稳态 $(k_t, z_t) = (\bar{k}, 0)$ 下我们有

$$
V(\bar{k}, 0; 0) = V_{ss} + \underbrace{\frac{1}{2} V_{33,ss} \sigma^2}_{\text{福利成本}},
$$

划线部分反映了在二阶层面上出现的福利成本，它是二者之差：一个是 $V_{ss} = V(\bar{k}, 0)$，即在 $(\bar{k}, 0)$ 的稳态下所测量的价值方程，此时我们知道正处于稳定状态。一个是 $V(\bar{k}, 0; 0)$ 的均衡状态下所测量的价值方程，此刻我们掌握更多信息，不只知道我们当前正处在稳定状态，更知道未来也会继续处在这一点上。

此外福利成本 $V_{33,ss}$ 也并非在全部情况下都是负数。在一些情况下它甚至可能为正，如一个包括休闲决策的随机内生经济增长模型。进一步的模型描述及经验验证见Cho et al. (2015)。

[16]但也并不绝对，比如当模型中存在预防性消费 (precautionary consumption) 时，效用函数的一阶导数也需要考虑 V_3，相关讨论见Kimball (1990)。

2. 方便将福利成本 $V_{33,ss}$ 转换为消费单位，从经济学意义上来说更有意义。具体说来，定义系数 τ 反映消费减少的量，使得家庭部门对于以下 2 个选项是无差异的：1 是确定性条件下消费 $(1-\tau)c_t$ 个单位，1 是在不确定性条件下消费 c_t，即 τ 满足

$$\log(1-\tau)c = \log c + \frac{1}{2}V_{33,ss}\sigma^2,$$

其中使用到了 $V_{ss} = \log c$ 的条件。整理得 τ 的决定

$$\tau = 1 - \exp(\frac{1}{2}V_{33,ss}\sigma^2) \tag{6.45}$$

数值解。采用与第6.3.8节一样的参数校准值，我们有

$$\begin{aligned} V &= -0.540 + 0.026(k_t - 0.188) + 0.250z_t - 0.069(k_t - 0.188)^2, \\ c_t &= 0.388 + 0.680(k_t - 0.188) + 0.388z_t. \end{aligned} \tag{6.46}$$

其中第一，价值方程迭代的消费政策方程，与模型均衡条件下的扰动近似消费决策方程相同。第二，$V_{kz} = V_{zz} = V_{\sigma\sigma} = 0$，针对这一组校准参数，经济周期波动的福利成本是 0[17]。

初始值猜测。利用(6.46)生成的初始值 V_0 做初始猜测值做后续价值方程迭代，可以较快出现收敛。

或者采取混合策略。将两种方法共同作用于运算符 $\mathcal{H}$，一个是模型均衡条件，一个是根据最优决策方程而估算的价值方程 $V(k_t, z_t) = (1-\beta)\log c_t + \beta E_t V(k_{t+1}, z_{t+1})$。混合策略有助于以较低的计算成本对价值方程和决策方程作出近似[18]。

[17]均衡条件中消费决策方程 $c_t = 0.673\exp(z_t)k_t^{0.33}$，对应当期效用 $u_t = logc_t = z_t + \log 0.673 + 0.33\log k_t$。其中 z_t 的无条件均值是 0，以及 $\log k_t$ 的决策方程满足确定性等价条件，因此改变 z_t 的方差不会产生 (无条件的) 福利成本。

[18]也可以将价值方程的若干导数对方在一起，如 $(1-\beta)c_t^{-1} - \beta E_t V_{1,t+1} = 0$ 等，进而寻找这个 (由价值方程导数构成的) 方程系统的扰动近似解。这种方法可能会对寻找价值方程的更高阶近似有帮助。

附录

6.A Schur 补和对称正 (半) 定矩阵

6.A.1 Schur 补

本节作为 (Boyd and Vandenberghe, 2004, A.5.5.5) 的详细说明。假定矩阵 $M \in \mathbb{R}^{n\times n}$ 可做如下分解

$$\underset{n\times n}{M} = \begin{bmatrix} \underset{p\times p}{A} & \underset{p\times q}{B} \\ \underset{q\times p}{C} & \underset{q\times q}{D} \end{bmatrix}, \quad n = p+q, \quad p,q,n > 0.$$

Schur 补的思路在于，通过去掉一个变量分块求解线性方程。假定一个线性系统 (x, y)

$$Ax + By = c, \tag{6.47a}$$

$$Cx + Dy = d, \tag{6.47b}$$

改写为矩阵形式

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}. \tag{6.48}$$

可以利用高斯消元法 (Gaussian elimination) 求得系统解。

A 的 Schur 补以及系统解

假定 D 是可逆的，则(6.47b)改写为

$$y = D^{-1}(d - Cx),$$

代回(6.47a)替代 y 得

$$\left(A - BD^{-1}C\right)x = c - BD^{-1}d.$$

进一步假定 $\left(A - BD^{-1}C\right)$ 是可逆的，我们有系统解

$$x = \left(\underbrace{A - BD^{-1}C}_{\text{A 的 Schur 补}}\right)^{-1}\left(c - BD^{-1}d\right), \tag{6.49a}$$

$$y = D^{-1}d - D^{-1}C\left(\left(\underbrace{A - BD^{-1}C}_{\text{A 的 Schur 补}}\right)^{-1}\left(c - BD^{-1}d\right)\right). \tag{6.49b}$$

我们将上式中标记出的 $\left(A - BD^{-1}C\right)$ 定义为 M 中分块矩阵 A 的 Schur 补。

采用 A 的 Schur 补，我们有

$$
\begin{aligned}
x &= \left(A - BD^{-1}C\right)^{-1} c - \left(A - BD^{-1}C\right)^{-1} BD^{-1} d, \\
y &= -D^{-1}C\left(A - BD^{-1}C\right)^{-1} c + \left(D^{-1} + D^{-1}C\left(A - BD^{-1}C\right)^{-1} BD^{-1}\right) d,
\end{aligned}
$$

改写为矩阵形式

$$
\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \left(A - BD^{-1}C\right)^{-1} & -\left(A - BD^{-1}C\right)^{-1} BD^{-1} \\ -D^{-1}C\left(A - BD^{-1}C\right)^{-1} & \left(D^{-1} + D^{-1}C\left(A - BD^{-1}C\right)^{-1} BD^{-1}\right) \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}, \tag{6.50}
$$

与(6.48)联立不难看出，

$$
\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} \left(A - BD^{-1}C\right)^{-1} & -\left(A - BD^{-1}C\right)^{-1} BD^{-1} \\ -D^{-1}C\left(A - BD^{-1}C\right)^{-1} & \left(D^{-1} + D^{-1}C\left(A - BD^{-1}C\right)^{-1} BD^{-1}\right) \end{bmatrix} \tag{6.51}
$$

对式右侧做 matrix reflection 可得

$$
\begin{aligned}
&\begin{bmatrix} \left(A - BD^{-1}C\right)^{-1} & -\left(A - BD^{-1}C\right)^{-1} BD^{-1} \\ -D^{-1}C\left(A - BD^{-1}C\right)^{-1} & \left(D^{-1} + D^{-1}C\left(A - BD^{-1}C\right)^{-1} BD^{-1}\right) \end{bmatrix} \\
&= \begin{bmatrix} \left(A - BD^{-1}C\right)^{-1} & 0 \\ -D^{-1}C\left(A - BD^{-1}C\right)^{-1} & D^{-1} \end{bmatrix} \begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix} \\
&= \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} \begin{bmatrix} \left(A - BD^{-1}C\right)^{-1} & 0 \\ 0 & D^{-1} \end{bmatrix} \begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix},
\end{aligned}
$$

并且第三行三个矩阵中，上下三角矩阵都是可逆的；中间对角矩阵对角元素也是可逆的。

代回(6.51)右侧，且等式两侧同时求逆运算，可以得到 M 的解

$$
M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & 0 \\ D^{-1}C & I \end{bmatrix}. \tag{6.52}
$$

即，利用 A 的 Schur 补求解系统，(x, y) 的值由(6.49a)-(6.49b)给出，M 的值由(6.52)给出。这种算法的好处在于，在假定 D 是可逆矩阵的情况下，只需要求得 A 的 Schur 补，即可求解系统。

D 的 Schur 补以及系统解

类似地，若假定 A 是可逆的，(6.47a)改写为

$$
x = A^{-1}\left(c - By\right),
$$

代回(6.47b)替代 x 得

$$
\left(D - CA^{-1}B\right) y = D - CA^{-1}c.
$$

进一步假定 $\left(D - CA^{-1}B\right)$ 是可逆的，我们有系统解

$$
y = \left(\underbrace{D - CA^{-1}B}_{\text{D 的 Schur 补}}\right)\left(c - BD^{-1}d\right) \tag{6.53a}
$$

$$
y = A^{-1}c - A^{-1}B\left(\left(\underbrace{D - CA^{-1}B}_{\text{D 的 Schur 补}}\right)^{-1}\left(d - CA^{-1}c\right)\right). \tag{6.53b}
$$

我们将上式中标记出的 $\left(D-CA^{-1}B\right)$ 定义为 M 中分块矩阵 D 的 Schur 补。

采用 D 的 Schur 补，我们有

$$
\begin{aligned}
x &= \left(A^{-1}+A^{-1}B\left(D-CA^{-1}B\right)^{-1}CA^{-1}\right)c-\left(A^{-1}B\left(D-CA^{-1}B\right)^{-1}\right)d,\\
y &= -\left(D-CA^{-1}B\right)^{-1}CA^{-1}c+\left(D-CA^{-1}B\right)^{-1}d,
\end{aligned}
$$

改写为矩阵形式

$$
\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A^{-1}+A^{-1}B\left(D-CA^{-1}B\right)^{-1}CA^{-1} & -A^{-1}B\left(D-CA^{-1}B\right)^{-1} \\ -\left(D-CA^{-1}B\right)^{-1}CA^{-1} & \left(D-CA^{-1}B\right)^{-1} \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}, \tag{6.54}
$$

与(6.48)联立可见

$$
\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1}+A^{-1}B\left(D-CA^{-1}B\right)^{-1}CA^{-1} & -A^{-1}B\left(D-CA^{-1}B\right)^{-1} \\ -\left(D-CA^{-1}B\right)^{-1}CA^{-1} & \left(D-CA^{-1}B\right)^{-1} \end{bmatrix} \tag{6.55}
$$

对式右侧做 matrix reflection 可得

$$
\begin{aligned}
&\begin{bmatrix} A^{-1}+A^{-1}B\left(D-CA^{-1}B\right)^{-1}CA^{-1} & -A^{-1}B\left(D-CA^{-1}B\right)^{-1} \\ -\left(D-CA^{-1}B\right)^{-1}CA^{-1} & \left(D-CA^{-1}B\right)^{-1} \end{bmatrix}\\
=&\begin{bmatrix} A^{-1} & -A^{-1}B\left(D-CA^{-1}B\right)^{-1} \\ 0 & \left(D-CA^{-1}B\right)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}\\
=&\begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & \left(D-CA^{-1}B\right)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix},
\end{aligned}
$$

同样地，第三行三个矩阵中，上下三角矩阵都是可逆的；中间对角矩阵对角元素也是可逆的。

代回(6.55)右侧，且等式两侧同时求逆运算，可以得到 M 的解

$$
M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & CA^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D-CA^{-1}B \end{bmatrix} \begin{bmatrix} I & 0 \\ A^{-1}B & I \end{bmatrix}. \tag{6.56}
$$

6.A.2　利用 Schur 分解法判断可逆正定矩阵

对于符合第5.8描述的可逆矩阵 M，假定其分块对角 A, D 都是可逆的，并且 $C=B^T$，那么可以将其改写为如下类似于分块——对角矩阵的形式

$$
\begin{aligned}
M &= \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}\\
&= \begin{bmatrix} I & BC^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A-BC^{-1}B^T & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I & BC^{-1} \\ 0 & I \end{bmatrix}^T.
\end{aligned} \tag{6.57}
$$

其中第二行利用了 A 的 Schur 补。我们可以利用这一分块对角形式，查看对称矩阵 M 是否是正定的，即 $M \succ 0$。

Proposition 6.2. 对于任一对称矩阵 $M = \begin{bmatrix} A & B \\ B^\top & C \end{bmatrix}$，如果 C 是可逆的，则

1. 当且仅当

$$
\begin{cases} C \succ 0, \quad \text{并且} \\ A-BC^{-1}B^\top \succ 0 \end{cases}
$$

时，我们有 $M \succ 0$。

2. 如果 $C \succ 0$，那么当且仅当 $A - BC^{-1}B^\top \succeq 0$ 时，我们有 $M \succeq 0$。

证明. 略。 □

Proposition 6.3. 对于任一对称矩阵 $M = \begin{bmatrix} A & B \\ B^\top & C \end{bmatrix}$，如果 A 是可逆的，则

1. 当且仅当
$$\begin{cases} A \succ 0, \quad \text{并且} \\ C - B^\top AB \succ 0 \end{cases}$$
时，我们有 $M \succ 0$。

2. 如果 $A \succ 0$，那么当且仅当 $C - B^\top AB \succeq 0$ 时，我们有 $M \succeq 0$。

证明. 略。 □

6.A.3 利用 Schur 分解法判断可逆正半定矩阵

当 C 和 A 都是奇异分块矩阵，我们无法直接求得 C^{-1} 和 A^{-1}。我们首先介绍判断方法，随后用该方法判断 M 是否是正半定的。

判定方法

1. (通常是 nonconvex) 的 quadratic 优化问题：对于满足一定形式的 x 和 (非奇异) 的 P，式
$$\min_{\{x\}} f(x) = \frac{1}{2}x^\top Px^\top + x^\top b$$
是否存在最小值，以及最小值是多少，见 Proposition 6.4。

2. 对于奇异的 P，我们要计算其对角分块的伪逆矩阵和相应的 Schur 补，进而判断最小值，见 Proposition 6.5。

Proposition 6.4 (非奇异系数矩阵 (quadratic 优化问题))**.** 如果 P 是一个对称的非奇异矩阵，那么只有当 $P \succeq 0$ 时，方程 $f(x) = \left(\frac{1}{2}\right)x^\top Px + x^\top b$ 才具有唯一的最优 (极小) 值 $x^\star = P^{-1}b$，使得 $f(x^\star) = f(P^{-1}b) = -\frac{1}{2}b^\top P^{-1}b$。

证明. 已知
$$\frac{1}{2}\left(x + P^{-1}b\right)^\top P\left(x + P^{-1}b\right) = \frac{1}{2}x^\top Px + x^\top b + \frac{1}{2}b^\top P^{-1}b,$$
因此我们有
$$f(x) = \frac{1}{2}x^\top Px + x^\top b = \frac{1}{2}\left(x + P^{-1}b\right)^\top P\left(x + P^{-1}b\right) - \frac{1}{2}b^\top P^{-1}b. \tag{6.58}$$

如果假定 P 中有负的特征根 $-\lambda, (\lambda > 0)$，对应特征向量 u，根据定义我们有 $Pu = -\lambda u$。另一方面，设 $x + P^{-1}b := \alpha u$，其中 α 是任一不为 0 的实数。则我们有
$$\left(x + P^{-1}b\right)^\top P\left(x + P^{-1}b\right) = (\alpha u)^\top P(\alpha u) = \alpha^2 u^\top Pu = \alpha^2 u^\top(-\lambda u) = -\alpha^2\lambda\|u\|^2,$$
其中 $\|N\|$ 表示向量的范数 (norm)。则(6.58)进一步改写为
$$f(x) = \underbrace{-\frac{1}{2}\alpha^2\lambda\|u\|^2}_{\geq 0} - \frac{1}{2}b^\top P^{-1}b, \tag{6.59}$$

由于 α 的取值可以至任意大，在 P 存在负的特征根的情况下 $f(x)$ 不可能有最小值。

因此，$f(x)$ 最小值存在的先决条件是满秩矩阵 $P \succeq 0$。在 $P \succeq 0$ 并且(6.58)右侧第一部分恒大于等于 0(由(6.59)右侧第一部分恒大于等于 0 推得) 的情况下，$f(\cdot)$ 的最小值只能出现在 $x^\star = -P^{-1}b$，对应 $f(x^\star) = -\frac{1}{2}b^\top P^{-1}b$。 □

Proposition 6.5 (奇异矩阵)**.** *如果 P 是一个对称的奇异矩阵，那么只有当 $P \succeq 0$ 并且 $\left(I - PP^\dagger\right)b = 0$ 时，方程 $f(x) = \left(\frac{1}{2}\right)x^\top Px + x^T b$ 才具有唯一的最优 (极小) 值 $x^\star = -P^\dagger b$，使得 $f(x^\star) = -\frac{1}{2}b^\top P^\dagger b$。*

进而，如果 $P = U^T\Sigma U$ 是一个关于 P 的 SVD，那么系统最优值 $f(x^\star)$ 对应 $x \in \mathbb{R}^{n\times n}$:

$$x^\star = -P^\dagger b + U^\top \begin{bmatrix} 0 \\ z \end{bmatrix}, \quad \text{对于任一} \quad z \in \mathbb{R}^{n-r\times n-r}, \quad r := rank(P).$$

证明. 对不满秩 $(r < n)$ 的奇异矩阵 P 做分块对角化

$$P = U^\top \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} U,$$

其中 U 是正交矩阵，$\Sigma_r \in \mathbb{R}^{r\times r}$ 是可逆对角矩阵。由此可得

$$\begin{aligned} f(x) &= \frac{1}{2}x^\top Px + x^\top b \\ &= \frac{1}{2}x^\top U^\top \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Ux + x^\top U^\top Ub \\ &= \frac{1}{2}(Ux)^\top \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} (Ux) + (Ux)^\top Ub. \end{aligned}$$

定义

$$Ux := \begin{bmatrix} y \\ z \end{bmatrix}, \quad Ub := \begin{bmatrix} c \\ d \end{bmatrix} \quad y, c \in \mathbb{R}^r, \quad z, d \in \mathbb{R}^{(n-r)},$$

上式可进一步调整为

$$\begin{aligned} f(x) &= \frac{1}{2}\begin{bmatrix} y^\top & z^\top \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} + \begin{bmatrix} y^\top & z^\top \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} \\ &= \frac{1}{2}y^\top \Sigma_r y + y^\top c + z^\top d. \end{aligned} \tag{6.60}$$

根据(6.60)，我们逐次分析 $f(x)$ 最小值存在的两个条件:

1. 对于 $y = 0$ 的情况，我们有 $f(x) = z^\top d$，此时如果 $d \neq 0$ 取任意值，$f(x)$ 将没有最小值。因此为了让 $f(x)$ 有最小值，需要使 $d = 0$。根据第6.B节可得，$Ub = \begin{bmatrix} c \\ 0 \end{bmatrix}$，进而 $b = range(P)$[19]；因此 $f(x)$ 存在最小值的条件之一是 $b = 0$，等价于 $\left(I - PP^\dagger\right)b = 0$。

2. 将 $b = 0$ 代回原式，我们有 $f(x) = \frac{1}{2}y^\top \Sigma_r y + y^\top c$，由于分块矩阵 Σ_r 是可逆的，根据 Proposition 6.4，$f(x)$ 存在最小值的第二个条件是 $\Sigma_r \succeq 0$，等价于 $P \succeq 0$。

假定 $f(x)$ 最小值存在的两个条件都得到了满足。来看 $x^\star$ 的取值。

[19]第6.B节中的 U 对应本节中的 $U^\top$。

1. 对于 $d=0$,
$$\frac{\partial f}{\partial y}=\Sigma_r y+c=0,$$
可得 $y^\star=-\Sigma_r^{-1}c$.

2. 假定 $z=0$,
$$Ux^\star=\begin{bmatrix}y^\star\\ z\end{bmatrix}=\begin{bmatrix}-\Sigma_r^{-1}c\\ 0\end{bmatrix},$$
由正交矩阵性质我们有 $U^\top=U^{-1}$，进而
$$\begin{aligned}x^\star&=-U^{-1}\Sigma_r^{-1}c\\&=-U^\top\begin{bmatrix}\Sigma_r^{-1}&0\\0&0\end{bmatrix}\begin{bmatrix}c\\0\end{bmatrix}\\&=-U^\top\begin{bmatrix}\Sigma_r^{-1}&0\\0&0\end{bmatrix}Ub\\&=-P^\dagger b,\end{aligned}\tag{6.61}$$
并且不难看出，任一 $z\in R^{n-r}$ 对 $x^\star=-P^\dagger b+U^\top\begin{bmatrix}0\\z\end{bmatrix}$ 的取值都不产生影响。

进一步看最优极小值 $f(x^\star)$
$$\begin{aligned}f(x^\star)&=\frac{1}{2}(x^\star)^\top Px^\star+(x^\star)^\top b\\&=\frac{1}{2}\left(-P^\dagger b\right)^\top P\left(-P^\dagger b\right)+\left(-P^\dagger b\right)^\top b\\&=\frac{1}{2}b^\top(P^\dagger)^\top PP^\dagger b-b^\top(P^\dagger)^\top b\\&=b^\top\left(-\frac{1}{2}(P^\dagger)^\top\right)b\\&=-\frac{1}{2}b^TP^\dagger b,\end{aligned}\tag{6.62}$$
并且同样地，任一 $z\in R^{n-r}$ 对 $f(x^\star)$ 的取值也不产生影响。 □

判断 M 是否正半定

在此基础上，回到最初的问题上来：如何判断对称矩阵 $M=\begin{bmatrix}A&B\\B^\top&C\end{bmatrix}$ 是正半定矩的。

这个问题等同于：判断方程 $f(x,y)$ 是否有最小值：
$$f(x,y)=\begin{bmatrix}x^\top&y^\top\end{bmatrix}\begin{bmatrix}A&B\\B^\top&C\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

根据 Proposition (6.5)可知：对于给定的常数向量 y，只有当 $A\succeq 0,\left(I-AA^\dagger\right)By=0,C-B^\top A^\dagger B\succeq 0$ 时，f 存在最小值，最小值为
$$f(x^\star,y)=-y^\top B^\top A^\dagger By+y^\top Cy=y^\top\left(C-B^\top A^\top B\right)y.$$

Theorem 6.1. 对于任一对称矩阵 $M=\begin{bmatrix}A&B\\B^\top&C\end{bmatrix}$，下述条件之间等价

1. $M \succeq 0$,
2. $A \succeq 0, (I - AA^\dagger B) = 0, C - B^\top A^\dagger B \succeq 0$,
3. $C \succeq 0, (I - CC^\dagger B = 0), A - B^\top C^\dagger B \succeq 0$,

并且可以对 M 做如下矩阵分解

$$M = \begin{cases} = \begin{bmatrix} I & BC^\dagger \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BC^\dagger B^T & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I & 0 \\ C^\dagger B^T & I \end{bmatrix}, & \text{或} \\ = \begin{bmatrix} I & 0 \\ B^\top A^\dagger & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & C - B^\top A^\dagger B \end{bmatrix} \begin{bmatrix} I & A^\dagger B \\ 0 & I \end{bmatrix}. \end{cases}$$

6.B　伪逆矩阵

又称 Moore-Penrose Pseudo inverse。可见如 (Bapat, 2012, Ch.5)。

6.B.1　用 SVD 法做伪逆矩阵分解

一个方块矩阵 $M \in \mathbb{R}^{n\times n}$ 可以做如下 SVD 分解 (singular value decomposition)，以求得伪逆矩阵 $M^\dagger$

$$M = U\Sigma V^\top,$$

其中

- U, V 是正交矩阵，U 和 V 的 column 分别是 $MM^\top$ 和 $M^\top M$ 的特征向量。需要指出的是，对 M 做 SVD，得到的 U 和 V 并不唯一。
- 对角矩阵 $\Sigma = diag(\sigma_1, \sigma_2, \ldots \sigma_r, 0, 0, \ldots 0)$:
  - $\sigma_1, \sigma_2, \ldots \sigma_r$ 表示 M 的秩，并且 $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_r$。
  - σ_i 表示 M 的奇异值。$\sigma_i, i = 0, 1, \ldots r$ 表示 $MM^\top$ 和 $M^\top M$ 的非零特征值的平方根。

如果 $M = U\Sigma V^\top$ 是某个 M 的 SVD，则我们将 M 的伪逆矩阵 $M^\dagger$ 定义如下

$$M^\dagger = V\Sigma^\dagger U^\top, \quad \Sigma^\dagger = diag\left(\sigma_1^{-1}, \sigma_2^{-1}, \ldots \sigma_r^{-1}, \ldots 0, 0, \ldots 0\right). \tag{6.63}$$

不难看出，当 M 是满秩 $(r = n)$ 的（等价于 M 是可逆的）时，$M^\dagger = M^{-1}$。因此 $M^\dagger$ 是个 M 的“广义逆矩阵”。对 M 做 SVD，U 和 V 有所不同，但同一 M 只有唯一的 $M^\dagger$ 与之相对应。此外，根据伪逆矩阵的性质，我们有 $MM^\dagger M = M$, $M^\dagger MM^\dagger = M^\dagger$，并且 M 和 $M^\dagger$ 都是对称矩阵；并且

$$\begin{cases} MM^\dagger & = U\Sigma V^\top V\Sigma^\top U^\top = U\Sigma\Sigma^\top U^\top = U \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} U^\top, \\ M^\dagger M & = V\Sigma^\dagger U^\top U\Sigma V^\top = V\Sigma^\dagger\Sigma V^\top = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\top, \end{cases}$$

因此我们有

$$\begin{cases} \left(MM^\dagger\right)^2 & = MM^\dagger, \\ \left(M^\dagger M\right)^2 & = M^\dagger M, \end{cases}$$

即 $\left(MM^\dagger\right)$ 和 $\left(M^\dagger M\right)$ 是对称的正交投影：

1. $MM^\dagger$ 是 $range(M)$ 的正交投影，

2. $M^\dagger M$ 是 $ker(M)^\perp$ 的正交投影，$ker(M)^\perp$ 表示 $ker(M)$ 的正交补 (orthogonal complement)。

证明. 由于 $range(M^\dagger M) \subseteq range(M)$，即 $M^\dagger M$ 的值域是 M 值域的子集，则在 $range(M)$ 值域中，对于任一 $y = Mx$，由于 $MM^\dagger M = M$，我们有

$$MM^\dagger y = MM^\dagger Mx = Mx = y,$$

因此 $range(MM^\dagger) \subseteq range(M)$。此外，根据 $MM^\dagger M = M \Rightarrow ker(M^\dagger M) \subseteq ker(M)$。因而有 $ker(M^\top) = ker(M)$。

由于 $M^\dagger M$ 是埃米特矩阵 (Hermitian)，$range(M^\dagger M) = ker(M^\dagger M)^\perp = ker(M)^\perp$。 □

值域 $range(M) = range(MM^\dagger)$ 中包括所有向量 $y \in \mathbb{R}^{n\times n}$，满足 $U^T y = \begin{bmatrix} z \\ 0 \end{bmatrix}, z \in \mathbb{R}^r$。

证明. 对于 $y = Mx$，我们有

$$U^\top y = U^\top Mx = U^\top U\Sigma V^\top x = \Sigma V^\top x = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\top x = \begin{bmatrix} z \\ 0 \end{bmatrix},$$

反过来，如果 $U^\top y = \begin{bmatrix} z \\ 0 \end{bmatrix}$，则我们首先有 $y = U \begin{bmatrix} z \\ 0 \end{bmatrix}$，进而

$$\begin{aligned} MM^\dagger y &= U \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} U^\top y \\ &= U \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} U^\top U \begin{bmatrix} z \\ 0 \end{bmatrix} \\ &= U \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} \begin{bmatrix} z \\ 0 \end{bmatrix} \\ &= U \begin{bmatrix} z \\ 0 \end{bmatrix} \end{aligned}$$

这说明 $y \in range(M)$，即 y 在 M 的值域之中。 □

类似地，值域 $range(M^\dagger M) = ker(M)^\perp$ 中包括所有向量 $y \in \mathbb{R}^{n\times n}$，满足 $V^\top y = \begin{bmatrix} z \\ 0 \end{bmatrix}, z \in \mathbb{R}^{r\times r}$。

证明. 对于 $y = M^\dagger MU$，我们有

$$y = M^\dagger MU = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\top U = V \begin{bmatrix} z \\ 0 \end{bmatrix}.$$

反过来，如果 $V^\top y = \begin{bmatrix} z \\ 0 \end{bmatrix}$，则 $y = \begin{bmatrix} z \\ 0 \end{bmatrix}$，进而

$$M^\top MV \begin{bmatrix} z \\ 0 \end{bmatrix} = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\top V \begin{bmatrix} z \\ 0 \end{bmatrix} = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} \begin{bmatrix} z \\ 0 \end{bmatrix} = y.$$

这说明 $y \in range(M^\dagger M)$。 □

6.B.2　用 SVD 法做 M 的矩阵分解

对于 M 是一个对称实矩阵的情况。每一个 $M = U\Sigma V^\top$ 的 SVD 分解常常只有不超过 1 个 Σ，但有多个 U 和 V，并且通常来说不存在 $U = V$ 的情况。然而，如果对称矩阵 $M \succeq 0$，那么

1. M 的特征值非负，
2. M 的非零特征值数量，与奇异值数量相等，
3. 可以用 SVD 法分解：$M = U\Sigma V^\top$。

如果 M 是一个对称复矩阵，结果与之相类似，除了此时

1. U 和 V 是埃尔米特矩阵，并且
2. $M^\top M$ 和 $MM^\top$ 是埃尔米特正交映射。

对于 M 是一个对称正规矩阵 (normal matrix) 的情况，即方块矩阵满足 $MM^\top = M^\top M$，此时 M 的 SVD 分解与分块对角分解之间关系密切，并且 M 的伪逆矩阵可以直接由分块对角分解求得。

证明. 对实正规矩阵 M 做分块对角分解：

$$M = U\Lambda U^\top,$$

其中 U 是正交矩阵，$\Lambda = diag\,(B_1, B_2, \ldots B_n)$ 是实系数分块对角矩阵。

分块 $B_j, j = 0, 1, \ldots n$ 可以是下属两种形式之一。第一是一个 2×2 的矩阵 $B_j = \begin{bmatrix} \lambda_j & -\mu_j \\ \mu_j & \lambda_j \end{bmatrix}$；第二是一个一维分块 $B_j = (\lambda_j)$。

假定 $B_1, B_2, \ldots B_p$ 是前一种形式，且 $\lambda_{2p+1}, \lambda_n$ 是标量。我们已知 $\lambda_j \pm i\mu_j$，并且 λ_{2p+k} 表示 A 的特征值。

令 $\rho_{2j-1} = \rho_{2j} = \left(\lambda_j^2 + \mu_j^2\right)^{1/2}$ for $j = 1, \ldots p$，以及 $\rho_{2p+j} = \lambda_j$，并且假定这些分块已经按照降序排列 $\rho_1 \geq \rho_2 \geq \ldots \geq \rho_n$。那么可见

$$UU^\top = U^\top U = U\Lambda U^\top U\Lambda^\top U^\top = U\Lambda\Lambda^\top U^\top,$$

并且

$$\Lambda\Lambda^\top = diag(\rho_1^2, \rho_2^2, \ldots \rho_n^2),$$

因此可见：M 的奇异值 $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$ 就是矩阵 $MM^\top$ 特征值的非负平方根 $\rho_1 \geq \rho_2 \geq \ldots \rho_n$，满足 $\sigma_j = \rho_j, 1 \leq j \leq n$。

接着定义对角矩阵

$$\Sigma = diag(\sigma_1, \sigma_2, \ldots \sigma_r, 0, 0, \ldots, 0),$$

其中 $r = rank(M)$，$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$，并且定义

$$\Theta = diag(\sigma_1^{-1}B_1, \sigma_2^{-1}B_2, \ldots \sigma_p^{-1}B_p, 1, 1, \ldots, 1).$$

由此可见，Θ 是个正交矩阵，并且

$$\Lambda = \Theta\Sigma = (B_1, B_2, \ldots, B_p, \lambda_{2p+1}, \lambda_{2p+2}, \ldots, \lambda_r, 0, 0, \ldots, 0).$$

进而

$$M = U\lambda U^\top = U\Theta\Sigma U^\top.$$

由于 U 和 Θ 都是正交矩阵，因此 $V \equiv U\Theta$ 也是正交矩阵，并且 $M = V\Sigma U^\top$ 是 M 的 SVD。根据 SVD 的性质我们有

$$M = V\Sigma U^\top = U\Sigma^\dagger V^\top = U\Sigma^\dagger\Theta^\top U^\top,$$

其中最后一个等式根据 $V = U\Theta \Rightarrow V^\top = \Theta\top U^\top$。

Θ 是正交矩阵 $\Rightarrow \Theta^\top = \Theta^{-1} \Rightarrow \Sigma^\dagger\Theta^\top = \Sigma^\dagger\Theta^{-1} = \Lambda^\dagger$

因此我们有 $M^\dagger = U\Lambda^\dagger U^\top$。

来看 $\Lambda^\dagger$ 的性质。如果我们将 Λ_r 写为 $\Lambda_r = (B_1, B_2, \ldots, B_p, \lambda_{2p+1}, \ldots, \lambda_r)$，则 Λ 是可逆矩阵，并且 $\Lambda^\dagger = \begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}$。

由上可见，一个 (实系数) 正规矩阵 M 的伪逆矩阵，可以通过对 M 做分块对角分解求得。 □

第七章　映射法

7.1　简介

映射法 (projection methods)，又称加权残差法 (weighted residual methods)。利用映射法求解 DSGE 模型是指，找到一组合适的系数向量 $\theta = \{\theta_0, \theta_1, \ldots, \theta_j\}$ 和一组基方程 (basis function) $\Psi_i(\boldsymbol{x})$ 后，假定 θ 和 $\Psi_i(x)$ 呈现线性组合方程 $d^j(\boldsymbol{x}|\theta)$

$$d^j(\boldsymbol{x}|\theta) = \sum_{i=0}^{j} \theta_i \Psi_i(\boldsymbol{x}), \tag{7.1}$$

进而用 $d^j(\boldsymbol{x}|\theta)$ 来近似 $d(\boldsymbol{x})$，进而求解 $\mathcal{H}(d) = 0$，即是说，定义一个残差方程 (residual function) $R(\boldsymbol{x}|\theta) = \mathcal{H}\left(d^j(\boldsymbol{x}|\theta)\right)$，通过 θ 系数的选取实现残差方程的最小化

$$\min_{\{\theta\}} \left|R(\boldsymbol{x}|\theta)\right|, \tag{7.2}$$

这称为对 $\mathcal{H}$ 做 Ψ 的映射，以求最佳系数 θ。

在操作层面上，这意味着首先如(7.1)，选取某种恰当的基 $\{\Psi_i(\boldsymbol{x})\}_{i=0}^{\infty}$，用基与系数内积并求和的方法构建 $d(\boldsymbol{x}|\theta)$；进而如(7.2)，对 $\mathcal{H}$ 做 Ψ 的映射，求得 θ。就这两个环节来说，显然，选取不同形式的基和映射算法的组合，会出现不同的映射法，在不同的文献中，有时也会给它们以一些专有的称呼，以与其他映射法相区分。

经济学家很早就将映射理论应用到经验研究中去了，进入 20 世纪 90 年代后，映射法逐渐成为一种成熟的宏观经济学研究方法，主要归功于Judd (1992a); Gaspar and Judd (1997); Judd (1998) 的贡献[1]。

大致说来，映射法比扰动法更易于描述，但算法实现上则较为困难。下面是一个映射法的基本算法描述：

Algorithm 7.1 (映射法的基本算法)**.** *可简述如下：*

1. *对于正整数 $j < \infty$，定义 $j+1$ 个线性不相关方程 $\Psi_i : \Omega \mapsto \mathbb{R}$，$i = 0, 1, \ldots, j$。则 $\{\Psi_i\}_0^j = \psi_0, \psi_1, \ldots, \psi_j$ 称为基方程，是与状态变量向量 $\boldsymbol{x}$ 有关的方程。*

2. *对于正整数 $m < \infty$ (m 是方程系统 $d(\boldsymbol{x})$ 所映射的维度)，系数向量 $\theta^l = \left[\theta_0^l, \theta_1^l, \ldots, \theta_j^l\right]$，$l = 1, 2, \ldots, m$。将 m 组系数向量合并组成系数矩阵*

$$\underset{\{m\times(j+1)\}}{\theta} = \begin{pmatrix} \theta^1 \\ \theta^2 \\ \vdots \\ \theta^m \end{pmatrix} = \begin{pmatrix} \theta_0^1 & \theta_1^1 & \ldots & \theta_j^1 \\ \theta_0^2 & \theta_1^2 & \ldots & \theta_j^2 \\ \vdots & \vdots & \ddots & \vdots \\ \theta_0^m & \theta_1^m & \ldots & \theta_j^m \end{pmatrix}.$$

[1] 与扰动法相比，映射法的思路更为现代，但在自然科学、工程学研究中也有许多年的悠久历史了。映射法的前身之一光谱法 (spectral methods) 至少可上溯至Lanczos (1938)，另一个前身有限元法 (finite elements methods, FEM) 来自于波音工程师Clough (1960) 的开创性工作；关于有限元法的早期发展可见Clough and Wilson (1999)。

3. 定义内积求和方程

$$d^{l,j}(\cdot|\theta^l) = \sum_{i=0}^{j} \phi_i^l \psi_i(\cdot),$$

则对应地我们有

$$\underset{\{m\times 1\}}{d^j(\cdot|\theta)} = \begin{pmatrix} d^{1,j}(\cdot|\theta^1) \\ d^{2,j}(\cdot|\theta^2) \\ \vdots \\ d^{m,j}(\cdot|\theta^m) \end{pmatrix}.$$

4. 将 $d^j(\cdot|\theta)$ 作为 $d(\cdot)$ 的近似，代入 $\mathcal{H}(\cdot)=0$，构建残差方程

$$R(\cdot|\theta) = \mathcal{H}\left(d^j(\cdot|\theta)\right).$$

文有注释，对距离方程其实三角方程的一。到时把讲义容补充到这里。

5. 定义度量方程 *(metric function)* [2] $\rho\left(R(\cdot|\theta), \mathbf{0}\right)$作为目标方程，在 θ 中找到对应的 $\hat{\theta}$，满足条件

$$\hat{\theta} = \underset{\theta\in\mathbb{R}^{m\times(j+1)}}{\arg\min}\ \rho\left(R(\cdot|\theta), \mathbf{0}\right).$$

7.1.1 例

举例说明。以经典的随机内生经济增长模型为例，第6.3.8节中我们介绍了二阶扰动方法的应用。使用近似的模型，这里来介绍如何用映射法近似求解系统[3]。

第一步，改写系统。模型中的状态变量 c_t, k_{t+1} 分别表示为

$$c_t = d^1(k_t, z_t),$$
$$k_{t+1} = d^2(k_t, z_t),$$

由此将经济系统的均衡条件 (由跨期消费的欧拉等式和资源约束条件构成，即 $m=2$) 改写为

$$\mathcal{H}\left(d(k_t, z_t)\right) = \mathbf{0}, \quad \forall k_t, z_t \Rightarrow$$
$$\begin{cases} u'\left(d^1(k_t,z_t)\right) - \beta E_t\left\{u'\left(d^1\left(d^2(k_t,z_t), z_{t+1}\right)\right)\left(\alpha\exp(\rho z_t + \sigma\varepsilon_{t+1})\left(d^2(k_t,z_t)\right)^{\alpha-1} + 1 - \delta\right)\right\} = 0, \\ d^1(k_t,z_t) + d^2(k_t,z_t) - \exp(z_t)k_t^\alpha - (1-\delta)k_t = 0. \end{cases}$$

第二步，改写 $d(\boldsymbol{x})$ 为 $d(\boldsymbol{x}|\theta)$ 的形式，根据 $m=2$ 有 $\theta = \left[\theta^1, \theta^2\right]^\top$，进而定义

$$c_t = d^{1,j}\left(k_t, z_t|\theta^1\right) = \sum_{i=0}^{j} \theta_i^1 \psi_i(k_t, z_t),$$
$$k_{t+1} = d^{2,j}\left(k_t, z_t|\theta^2\right) = \sum_{i=0}^{j} \theta_i^2 \psi_i(k_t, z_t).$$

ference

其中 $j+1$ 个基方程 $\psi_i(k_t, z_t), \quad i=0,1,\ldots,j$ 的形式，将在下文讨论。

[2]度量方程又称距离方程 (distance function)，描述集合中一对元素之间的距离。度量方程的一种，离散度量方程 (discrete metric function)可以表示如下

$$d(x,y) = \begin{cases} 0 & x = y, \\ 1 & x \neq y \end{cases}.$$

[3]第6.3.8节中的 η，在这里用 σ 来表示。

第三步，构建残差方程 $R(k_t, z_t|\theta) = \left[R(k_t, z_t|\theta^1), R(k_t, z_t|\theta^2)\right]^\top \Rightarrow$

$$R(k_t, z_t|\theta^1) = u'\left(\sum_{i=0}^{j}\theta_i^1\psi_i(k_t, z_t)\right) - \beta E_t\left\{u'\left(\sum_{i=0}^{j}\theta_i^1\psi_i\left(\sum_{i=0}^{j}\theta_i^2\psi_i(k_t, z_t), \rho z_t + \sigma\varepsilon_{t+1}\right)\right)\times\right.$$
$$\left.\left(\alpha\exp\left(\rho z_t + \sigma\varepsilon_{t+1}\right)\left(\sum_{i=0}^{j}\theta_i^2\psi_i(k_t, z_t)\right)^{\alpha-1} + 1 - \delta\right)\right\},$$
$$R(k_t, z_t|\theta^2) = \sum_{i=0}^{j}\theta_i^1\psi_i(k_t, z_t) + \sum_{i=0}^{j}\theta_i^2\psi_i(k_t, z_t) - \exp\left(z_t\right)k_t^\alpha - (1-\delta)k_t.$$

第四步，构建度量方程 $\rho\left(R(\cdot\theta), \mathbf{0}\right)$，计算

$$\hat{\theta} = \underset{\theta\in\mathbb{R}^{m\times(j+1)}}{\arg\min}\ \rho\left(R(\cdot|\theta), \mathbf{0}\right),$$

使得相对于 $\hat{\theta}$ 的任意 $m\times(j+1)$ 个点 (k_l, z_l)，都有度量 $\rho = 0$。为了实现这点，需要令对应的 $m\times(j+1)$ 个残差方程最小化，换句话说，求解下述含有 $m\times(j+1)$ 个未知数的方程系统[4]

$$R\left(k_l, z_l|\theta\right) = \mathbf{0}, \quad l\in\mathbb{R}^{m\times(j+1)}.$$

如上所述，采用映射法从事宏观经济学研究，关键的两点在于选择合适的基方程 $\Psi(\cdot)$ 以及度量方程 $\rho(\cdot)$，在下文分别讨论两个方程的选取策略之前，我们先来回顾一下，“映射法” 和计量经济学如 OLS 回归中的 “映射” 之间有何异同，以及映射法和参数化期望法之间的联系与区别。

reference 和 reference

与计量经济学中的 “映射” 的关系

计量经济学中也有 “映射” 的概念，与映射法的思路接近。例如，对于给定的变量 (X, Y)，线性回归考虑构建一个未知的条件期望方程 $E(Y|X)$。为了求解该条件期望方程，我们可以将 $E(\cdot)$ 近似为两个关于解释变量 X 的单项式 (monomial)，一个是常系数 θ_0，一个是关于 X 的线性方程对应系数 θ_1

$$E(Y|X) \approx \underbrace{\theta_0}_{\text{单项式 1}} + \underbrace{\theta_1 X}_{\text{单项式 2}},$$

这两个单项式可以合并构成另一个单项式的基的前两个元素，也可以合并构成另一个多项式的基的前两个元素，如切比雪夫多项式[5]。从而，计量经济学中 “映射” 法所构建的残差方程可以表示为

$$R(Y, X|\theta_0, \theta_1) = Y - \theta_0 - \theta_1 X,$$

随后的工作就变为，通过插入实际观测到的数据 $\{Y, X\}_{t=1}^T$，寻找合适的系数 (θ_0, θ_1) 使得残差方程的平方值最小 $\min R(Y, X|\theta_0, \theta_1)^2$。

可见计量经济学研究中的 “映射”，和这里介绍的 “映射” 法，研究思路相近，区别在于以 OLS 为例的前者使用实际观测到的数据做系数测算，后者则使用来自经济学理论模型的条件 $\mathcal{H}(d)$。

映射法与参数化期望法

在第5.7节我们介绍了参数化期望法 (PEA)。不难看出，参数化期望法与映射法也有相通之处。但二者还是存在着不同：

[4]本讲义略过方程解的存在性与唯一性的讨论。

[5]超几何方程的一般介绍，见附录7.A；超几何方程的另一种表现形式是 (正交) 多项式，见附录7.B。附录7.C.5对切比雪夫多项式也做了介绍。

1. 在模型构建方面，映射法将原方程系统表述为一组新的近似方程系统，新系统由一系列含有条件期望的基方程灵活组合而成，如(7.1)。在参数化期望法中，近似方程系统却需要假定单项式，或者由一系列单项式构成的方程系统来实现，如(5.45)。后者往往不是一个最佳策略[6]。

2. 在特定条件下，利用参数化期望法迭代求的的系数值，即便能够求得，常常与实际情况下的最优选择相差较远；或者即便最终收敛到最优解，但收敛速度往往较慢，甚至是不稳健的。

7.2 全局基的选取——单维基光谱法

本节讨论如何选取一组合适的基方程 $\Psi = \psi_0, \psi_1, \ldots, \psi_j$。通常说来，选取策略分为两大类，第一类是选取在状态变量域 Ω 的绝大部分范围内都是非零且平滑的基，称为全局基。另一类是选取在 Ω 中的一小部分范围内非零且平滑的基，称为局部基 (local basis)。基于全局基的映射法常称为光谱法 (spectral method)，基于局部基的映射法常称为有限元法 (finite elements method, FEM)。我们先介绍光谱法。

Judd (1992b) 首次将光谱法应用到经济学研究中来。光谱基作为全局基的代表，主要优点在于，便于构建近似算法并求解。然而缺点在于，在处理状态变量的某些局部子域时，效果未尽人意——如傅里叶序列 (Fourier series) 中的吉布斯现象 (Gibbs phenomenon)[7]。关于光谱法的介绍，可见Shen et al. (2011)。

单维基 (unidimensional basis) 是光谱法中最为常见的一种基方程。大致说来，有以下几种

- 单项式
- 三角序列
- 雅各比多项式
- 切比雪夫多项式

当方程系统中只有一个状态变量时，使用单维基有助于使系统求解尽可能简单。但模型常常较为复杂，不止一个状态变量，这就需要我们采用其他一些特殊类型的基方程。

7.2.1 单项式

单项式可以作为光谱法全局基的备选项之一：

$$1, x, x^2, x^3, \ldots$$

单项式直观，形象，易于理解。尽管单项式的基并不是由正交方程所构成的，但根据斯通-魏尔斯特拉斯定理 (Stone-Weierstrass Theorem)[8]，由一系列可测度且有界的方程所构成的紧致集合 (compact set) 空间中，我们总是能够将任一闭区间内的连续方程近似为若干单项式的线性组合。

然而，采用单项式作为基，存在两个棘手的问题。第一个是多重共线性。单项式都是 (近似) 多重共线的，如图(7.1)所示，在 $x \in [0.5, 1.5]$ 区间内，单项式方程 x^9 和 x^10 的曲线非常近似，这使得我们哪怕增添更高一次的单项式方程 x^{11}，计算出的近似解也许不会很快逼近实际方程。第二个问题是，随着 x 值的变化，多项式的值变化较大：如 $\frac{1.5^{11}}{0.5^{11}} = \frac{86.4976}{0.000488281} = 177147$，容易造成计算误差。

[6] Christiano and Fisher (2000) 因此建议对参数化期望法进行改进，用如车比雪夫多项式替代(5.45)的方程形式的猜测。

[7] 第7.G.5节。

[8] cf. 维基百科词条https://en.wikipedia.org/wiki/Stone-Weierstrass_theorem。

图 7.1: 单项式方程 x^9, x^10, x^{11} 在 $[0.5, 1.5]$ 区间内的图形

```
In[33]:= Plot[{x^9, x^10, x^11}, {x, 0.5, 1.5}, PlotLegends → "Expressions"]
```

绘图　　绘图的图例

Out[33]=

单项式基的上述缺点，使得我们寻求利用正交多项式构建基方程。内积形式的正交多项式的值相对于 x 的变化往往较为平缓，并且在加入更高次多项式元素后，系统会出现足够大的变化，有助于更精确近似原系统中的方程 $d(\cdot)$。

7.2.2　三角序列

光谱法全局基的另一选项是三角序列 (trigonometric series)，如

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{2\pi}}, \ldots, \frac{\cos kx}{\sqrt{2\pi}}, \frac{\sin kx}{\sqrt{2\pi}}, \ldots$$

三角序列适合用于分析有周期性特征的方程，因而在自然科学和工程学中得到广泛应用。遗憾的是除了时间序列分析等领域之外，经济学研究的问题较少涉及到周期性。此外，目前将周期性方程近似为非周期性方程的方法尚未成熟。因此我们不做深入探讨。

7.2.3　雅各比多项式

雅各比多项式(第7.C.3节) 也适合作为光谱法全局基的选项之一。

雅各比多项式一族中，还包括盖根鲍尔多项式 (Gegenbauer)，勒让德 (Legendre)，切比雪夫多项式 (Chebyshev)，他们可以相互转换。(Boyd and Petschek, 2013, Table 1) 比较了几种多项式在处理不同优化指标时的性能排序，从而提供了重要参考：在绝大多数情况下，对于求解 DSGE 模型的工作来说，最适于采用切比雪夫多项式构建基方程。

7.2.4　切比雪夫多项式

切比雪夫多项式(第7.C.5节) 作为光谱法全局基的选项之一，具有许多优点：

1. 便于在各种表述形式之间互相转换，如罗德里格斯公式、三项递推关系、母方程、二阶导数线性方程等。
2. 可以通过余弦换算，迅速计算系数变化后的值。
3. 切比雪夫插值 (interpolation) 换算的结果，比其他几种多项式插值的结果更为稳健。
4. 切比雪夫多项式是平滑的，并且有界 (闭区间 $[-1, 1]$)。

5. 切比雪夫差值的误差，可以由一系列定理得到较好的限制。

第一类切比雪夫多项式 $T_n(x)$ 满足关系 $T_n(0)=1, T_n(1)=x, T_{n+1}(x)=2xT_n(x)-T_{n-1}(x)$，因此我们可得到一组多项式序列 $1, x, 2x^2-1, 4x^3-3x, 8x^4-8x^2+1, \ldots$，随着 $n=0,1,2,\ldots,7$，我们绘制出了 $T_n(x)$ 的曲线，见7.2。不难看出，$n=0$ 时是条平行线，$n=1$ 时是 45 度斜线，$n=2$ 时是一条抛物线，随着 n 逐渐增加，切比雪夫多项式的曲线形状呈波浪状。

图7.2在 Mathematica 中输入如下命令生成：

```
t0 = Plot[{ChebyshevT[0, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
t1 = Plot[{ChebyshevT[1, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
t2 = Plot[{ChebyshevT[2, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
t3 = Plot[{ChebyshevT[3, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
t4 = Plot[{ChebyshevT[4, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
t5 = Plot[{ChebyshevT[5, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
t6 = Plot[{ChebyshevT[6, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
t7 = Plot[{ChebyshevT[7, x]}, {x, -1, 1}, PlotLabels -> "Expressions"];
GraphicsGrid[{{t0, t1}, {t2, t3}, {t4, t5}, {t6, t7}}]
```

第 n 次切比雪夫多项式 $T_n(x)$ 有 n 个根，对应

$$x_k=\cos\left(\frac{2k-1}{2n}\pi\right), \quad k=1,2,\ldots,n, -1\le x_k\le 1, \tag{7.3}$$

如图7.3所示，所有根都在闭区间 $[-1,1]$ 之间。

大多数 DSGE 模型的状态变量域 $x\in[a,b]$，$a,b\neq\pm1$，在实际研究中我们常采用如下线性关系转换，将它转换为 $[-1,1]$ 区间内的值，以符合切比雪夫多项式的要求：

$$2\frac{x-a}{b-a}-1, \tag{7.4}$$

详见第7.2.4节。例如 $(a=-7, b=5)$ 的某状态变量线性转换，如图(7.4)所示。

使用切比雪夫多项式作为映射基，有两个定理值得关注：

1. 切比雪夫插值定理 (Chebyshev interpolation theorem)，见附录7.56：如果近似方程恰好是第 n_1 次切比雪夫多项式的根，那么随着 $n_1\to\infty$，近似误差逐渐减少到足够小的程度。根据切比雪夫插值定理，可以使用切比雪夫多项式的根作为正交配点 (orthogonal collocation)，见7.4节 FEM 方法介绍。

2. 切比雪夫截断定理 (Chebyshev truncation theorem)，见附录7.57。进而在特定情况下，切比雪夫向未知原方程做几何收敛 (geometric convergence)，可以表示为

$$d(x)-d^j(x|\theta)\sim \boldsymbol{O}(\theta_j),$$

即是说，当切比雪夫近似停止于第 j 次多项式时，对应的截断误差与 j 次切比雪夫多项式系数 θ_j 是同次的。

根据该定理，我们可以设计一种数值检测机制：如截断误差大于某一阈值，则意味着 j 次切比雪夫近似 $T_j(x)$ 的精度不够，需要再增加一次近似至 $T_{j+1}(x)$。我们将在下文进一步讨论该问题。

ference

图 7.2: 切比雪夫多项式 $T_n(x)$

图 7.3: 切比雪夫多项式 $T_n(x)$ 的根 $x_k, n = 100, k = 1, 2, \ldots, n$

```
In[143]:= findroot[k_] := Cos[(2 k - 1)/(2 × 100) π];
          |余弦
          Plot[findroot[k], {k, 1, 100}]
          |绘图
```

Out[144]=

说明：计算等式依据(7.3)。

图 7.4: 切比雪夫多项式 $T_n(x)$ 的根 $x_k, n = 100, k = 1, 2, \ldots, n$

```
In[149]:= Trans[x_] := 2 (x - (-7))/(5 - (-7)) - 1;
          Plot[Trans[x], {x, -7, 5}]
          |绘图
```

Out[150]=

说明：计算等式依据(7.4)。

变量的变换

如前所述，7.4提供了一种可将 $x \in [a,b], a < -1, b > 1$ 转化为 $[-1,1]$ 区间中变量的方法，以便进行切比雪夫多项式近似。在扰动法中我们也讨论了变量的变换，见6.5节。二者尽管技术细节有所不同，但核心思路是相同的：都是为了尽可能提高近似的精确度。本节以随机 NCGT 模型为例，介绍为何通过变量变化有助于提高映射法的近似精度。

原目标为寻找经济系统方程组

$$\begin{aligned} c_t &= d^1(k_t, z_t), \\ k_{t+1} &= d^2(k_t, z_t) \end{aligned}$$

的近似解。我们可以通过变量变换，改为寻求方程组

$$\begin{aligned} \log c_t &= d^1(\log k_t, z_t), \\ \log k_{t+1} &= d^2(\log k_t, z_t) \end{aligned}$$

的近似解，用映射法表示为

$$\begin{aligned} \log c_t &= d^{1,j}(\log k_t, z_t|\theta^1) = \sum_{i=0}^{j} \theta_i^1 \psi_i(\log k_t, z_t), \\ \log k_{t+1} &= d^{2,j}(\log k_t, z_t|\theta^2) = \sum_{i=0}^{j} \theta_i^2 \psi_i(\log k_t, z_t). \end{aligned}$$

伯依德原则

前文简要介绍了利用切比雪夫多项式从事经济学研究所具有的理论优势。但近些年来的经验研究，尤其是基于 DSGE 模型的经济学研究，越来越多地使用基于切比雪夫多项式的映射法，切比雪夫多项式作为基方程的备选方案，其巨大优势得到了越来越广泛的认可 (Aruoba et al., 2006; Caldara et al., 2012)。

(Boyd, 2001, p.10) 用一种近似于开玩笑的方式，总结了这几十年来的研究经验，将之命名为“道德准则一号”：

1. 当不确定用什么基时，就用切比雪夫多项式。除非模型呈现出较强的周期性，这时可以考虑用傅里叶序列。
2. 除非你确定其他某种基方程更好，不然就用切比雪夫多项式。
3. 除非你非常非常确定其他某种基方程更好，不然就用切比雪夫多项式。

7.3　局部基的选取 (1)

第7.2节所介绍的几种方案，全部都是单维度基方程，相对简单，易于我们理解用映射法求解 DSGE 模型的基本思路。然而现实世界中大多数经济问题都是多维度的，几乎全部 DSGE 模型都研究一个以上的状态变量。这就需要我们探讨多维基的选择方案。

如何选取多维基成为映射法的关键问题。然而映射法受到维数灾难 (curse of dimensionality)的强烈冲击 (Bellman, 1957)。随着解释变量数量的增加——如中等规模 DSGE 模型可能要处理超过 20 组解释变量——应用映射法求解模型变得异常困难，这要求我们用更高的技巧去选取多维基方程。

7.3.1 离散状态变量

前文一系列分析均暗含假定，状态变量是连续的。然而在很多 DSGE 模型中，至少一部分变量是离散的：

- 状态变量本身是离散的，如
  - 财政政策，政府可能处于主权债务违约或未违约状态 (Bocola, 2016)，
  - 货币政策，可能是积极的或者消极的 (Leeper, 1991)，
- 出于计算求解的考虑，益于将连续状态变量做离散化处理，如
  - 外生随机过程 (技术冲击、偏好冲击等) 的离散化。

研究发现，有限马尔科夫链 (finite Markov chain)能够产生与连续过程相同的样本矩；经验分析表明，在大多数情况下，含有 5-7 个状态的马尔科夫链足够模拟一般情况下的随机过程信息，供经济学定量分析使用 (Tauchen, 1986; Kopecky and Suen, 2010)。因此状态变量的离散化问题可以理解为，针对某一个连续变量，我们寻找另一个 (离散的) 决策方程来描述它[9]。

仍然以含有随机 NCGT 模型为例，假定外生技术冲击 z_t 是一个一阶自回归过程 $AR(1)$

$$z_t = \rho z_{t-1} + \varepsilon_t,$$

其中 $\varepsilon_t \sim N(0, \sigma_z^2)$ 是一个平稳分布 (stationary distribution)。

可以表示为含有 n 个点的马尔科夫过程 $z_t \in \{z_1, z_2, \ldots z_n\}$，对应转移矩阵 $P_{z,z'}$ (transition matrix)

$$P_{z,z'} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ p_{21} & \cdots & p_{2n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}, \tag{7.5}$$

其中转移概率 (transition probability) $p_{i,j}$ 表示当前期处于位置 i 马尔科夫链下一时期转移到位置 j 的概率。

将 AR(1) 过程 z_t 离散化的常见算法，见附录第7.D节。

(注意，建议阅读这个附录，以便弄明白是怎么回事，还有附带的 matlab 代码。)

将技术冲击离散化后，我们的研究目标就变为，寻找 $2 \times n$ 个决策方程：

$$c\left(k, z_m\right) = d^{c,m,j}\left(k|\theta^{m,c}\right) = \sum_{i=0}^{j} \theta_i^{m,c} \psi_i(k),$$

$$k\left(k, z_m\right) = d^{k,m,j}\left(k|\theta^{m,k}\right) = \sum_{i=0}^{j} \theta_i^{m,k} \psi_i(k),$$

其中 $m = 1, 2, \ldots, n$。举例来说，我们首先寻找当今天的技术水平是 z_1 时资本和消费的决策方程，进而寻找当今天的技术水平是 z_2 时资本和消费的决策方程，随后 $z_3, \ldots, z_m, \ldots, z_n$。为了让数值运算不至于太过复杂，我们常将 n 控制在一个比较小的值上，比如 5 到 7 之间。

在求得两个决策方程后，代回到原系统的欧拉方程

$$u'(c_t) = \beta E_t \left[u'\left(c_{t+1}\right)\left(\alpha \exp(z_{t+1}) k_{t+1}^{\alpha-1} + 1 - \delta\right)\right] \tag{7.6}$$

[9]我们不对马尔科夫链做过多介绍（来不及把笔记敲进硬盘了）。讲义可参考帝国理工大学 Emma J McCoy M3S4/M4S4 - Applied Probability 的讲义第六章：Markov Chains, http://101.96.10.63/wwwf.imperial.ac.uk/~ejm/M3S4/NOTES3.pdf。教材可参考Privault (2013)。

中，我们有

$$u'\left(d^{c,m,j}\left(k|\theta^{m,c}\right)\right)=\beta\sum_{l=0}^{n}p_{ml}\left[u'\left(d^{c,l,j}\left(d^{k,m,j}\left(k|\theta^{m,k}\right)|\theta^{l,c}\right)\right)\right.$$
$$\left.\left(\alpha\exp(z_{t+1})\left(d^{k,m,j}\left(k|\theta^{m,k}\right)\right)^{\alpha-1}+1-\delta\right)\right] \tag{7.7}$$

(7.7)中有两点值得注意

1. 近似算法中 $2\times n$ 个决策方程都是当期的，不过上式中 $t+1$ 的决策方程，依然考虑到 $t+1$ 期可能出现的技术水平变化。
2. 由于我们将随机过程进行离散化近似，对应地，积分形式的 RHS 被简化为离散求和形式，乘以转移矩阵(7.5)中的相应元素[10]。

因此，在存在多维度问题的 DSGE 模型中，常常可以对其中至少一部分状态变量如技术冲击做离散化处理，其优点在于简单直观，操作过程透明，并且不是特别耗费计算资源。以及在一定意义上，求解 DSGE 模型往往依赖于混合策略：对一部分连续状态变量做离散化处理，对剩余的连续变量采用其他近似策略，如张量、完全多项式等，见下文。

7.3.2　张量与完全多项式

张量 (tensors)将一组单维的基方程，用克罗内克乘积[11]的组合在一起，构成多维基方程。例如一个经济系统中有两个状态变量，实物资本 k_t 和人力资本 h_t，每个状态变量分别对应 3 个切比雪夫多项式：

$$\psi_0^k(k_t),\psi_1^k(k_t),\psi_2^k(k_t),$$
$$\psi_0^h(h_t),\psi_1^h(h_t),\psi_2^h(h_t).$$

则我们可以构建一个张量作为基方程

$$\psi_0^k(k_t)\psi_0^h(h_t),\psi_0^k(k_t)\psi_1^h(h_t),\psi_0^k(k_t)\psi_2^h(h_t),$$
$$\psi_1^k(k_t)\psi_0^h(h_t),\psi_1^k(k_t)\psi_1^h(h_t),\psi_1^k(k_t)\psi_2^h(h_t),$$
$$\psi_2^k(k_t)\psi_0^h(h_t),\psi_2^k(k_t)\psi_1^h(h_t),\psi_2^k(k_t)\psi_2^h(h_t).$$

进而，对于一个有 n 个状态变量的方程系统 $d:[-1,1]^n\to\mathbb{R}$，我们试图用 j 次切比雪夫多项式予以近似，则可用如下张量形式表现

$$d^j(\cdot|\theta)=\sum_{i_1=0}^{j}\sum_{i_2=0}^{j}\cdots\sum_{i_n=0}^{j}\theta_{i_1,i_2,\ldots,i_n}\,\psi_{i_1}^1(\cdot)*\ldots\psi_{i_n}^n(\cdot),$$

其中 $\psi_{i_\kappa}^\kappa$ 是第 κ 个状态变量的 i_κ 次切比雪夫多项式，$\kappa=1,2,\ldots,n$；系数向量 θ 为 $\{\theta_{i_1},\theta_{i_2},\ldots,\theta_{i_n}\}$。

采用张量基方程有如下优点，

1. 易于构建
2. 传递性：如果组成张量的单维基方程是正交的，那么由单维基乘积组成的张量基方程也是正交的。

[10]也有一些研究直接从积分形式入手，探讨一系列求积法，可参考Judd (1998); Judd et al. (2011)。
[11]见式(7.31)。

最显著的不足在于维数灾难：待求解系数 $\theta_{i_1,\ldots i_n}$ 的个数随着 j 和 n 的增加而指数增加，呈 $(j+1)^n$ 形。上例中 $j=2, n=2$ 对应需要近似计算 9 个系数；图7.1。

$$(j+1)^n = \begin{cases} 256 & (j,n)=(3,4) \\ 3125 & (j,n)=(4,5) \\ 46656 & (j,n)=(5,6) \\ \ldots \end{cases}$$

图 7.1: 张量代求解系数的数量 $(j=1)^n$

在实际研究中，当模型里有 $n>3$ 个连续状态变量和 1 个哪怕是中等规模的 j 时，张量基的方法便不易展开了。一个方法是去掉张量集合中的部分元素，以减轻计算负担。Gaspar and Judd (1997) 建议使用完全多项式 (complete polynomials)：

$$\begin{aligned} \mathcal{P}_\kappa^n &\equiv \left\{\psi_{i_1}^1 * \ldots * \psi_{i_n}^n\right\}, |\boldsymbol{i}| \le \kappa, \\ |\boldsymbol{i}| &= \sum_{l=1}^n i_l, \quad 0 \le i_1, \ldots, i_n, \end{aligned} \tag{7.8}$$

即是说，首先预设一个正整数值 κ，进而在张量中，只选取各个基方程的级数之和小于 κ 所对应的那部分元素。这是基于下述认识：对于描述原方程系统 d 的目标而言，绝大多数信息都已经在完全多项式 $\mathcal{P}_\kappa^n$ 中体现出来了；余下的部分 $\psi_{i_1}^1 * \ldots * \psi_{i_n}^n, |\boldsymbol{i}| > \kappa$ 只能给基方程产生有限的信息，却是以大量额外计算时间为代价的。

举例来说，对于 $j=4, n=3$ 的张量系统，有 $(4+1)^3=125$ 个元素。我们设 $\kappa=6$，提取 $|\boldsymbol{i}| \le 6$ 的部分构建完全多项式 $\mathcal{P}_6^3$，则只需要近似计算其中的 87 个系数了。

ference

不幸的是，这仍然太多。在下文中我们进一步介绍斯莫尔亚克稀疏网格算法。

7.4 局部基的选取 (2)：有限元法

作为局部基方法的一种，有限元法 (finite elements method, FEM) 由McGrattan (1996) 最先在经济学研究中倡导并予以实践[12]。

[12]有限元法的数理知识，可见教材Hughes (2000); Brenner and Scott (2008)。可参考 Joseph E. Flaherty 的讲义 CSCI, MATH 6860: Finite Element Analysis http://www.cs.rpi.edu/~flaherje/FEM/index4.html。

有限元法常应用在一些关键任务的工业领域，如航空航天，核电厂施工等，这是由它的突出优势所决定的：可以很方便描述某些局部域内的行为，以及达到相当高的精确度。相应地，有限元法的主要不足在于难以编程，计算求解速度慢。

利用有限元法展开分析，常见的步骤如下。第一步是确定状态变量的值域 Ω。值域的选取方面，一些规则是天然的 (如 $k_t > 0$)，另一些则不是 ($k_t < \bar{k}$)，后者我们需要额外注意：如，可以将 $\bar{k}$ 设的足够高，以至于在模型仿真过程中所有 k_t 都在上限 $\bar{k}$ 的下方。随着仿真过程的实际展开，$\bar{k}$ 的值也可能做微调。

第二步是将 Ω 分成若彼此不重叠的有限个元。各个元相接的部分称为结点 (nodes)。分元的原则总体上来讲灵活性很高，可以

- 均等划分，直观简便
- 根据变量处于某种状态的频率，划分为一系列小块 (频率高) 和大块 (频率低)。对 “频率” 的设定又分为两种情况
  - 来自理论模型系统的特征
  - 利用迭代法，对所划分元的事后验证

见下文，补 reference。

- 在 Ω 中，将 $d(\cdot)$ 变化幅度较大的部分划分为小元，相对变化幅度不大的部分划分为大元，从而在每一个元内部，原本是非线性的 $d(\cdot)$ 接近线性。

在工程实践中，得益于灵活的分元，有限元法可以更好处理 kinks 或者一些特殊限制条件[13]。面对这些 kinks、或限制条件，一方面光谱法的操作难度较大，另一方面由于不满足可微、连续等条件，扰动法很难展开。

图 7.1: 弱 kinks 和强 kinks 的示例图

图片来源：(Fries and Belytschko, 2010, Figure 2)。

将 Ω 域分为有限个元素的的最优方法 (网格生成 grid generation)，已有大量工程学、数学方面的讨论，如Thompson et al. (1985)。在内生 NCGT 模型中应用网格优化法生成有限个不均等的元以优化计算时间，可见Fernández-Villaverde and Rubio-Ramírez (2004)。

第三步，在每一个元中选择对应的基，用于构建政策方程。如果前述步骤中元素的划分已经有效，则这一步中的基只设定线性形式即可。例如，Ω 中诸元的节点表示为 $\{k_0, k_1, \ldots k_j\}$，则可以

[13]kinks 通常是指，分元后，两个元的衔接处 (interface) 出现偏移 (弱 kinks) 甚至断裂 (强 kinks) 的情况，见图7.1。

定义基方程 $\psi_i(k)$ 为三角函数形式，$i=1,2,\ldots,j-1$：

$$\psi_0(k)=\begin{cases}\frac{k_0-k}{k_1-k},\text{如果}k\in[k_0,k_1],\\ 0,\end{cases}\tag{7.9a}$$

$$\psi_i(k)=\begin{cases}\frac{k-k_{i-1}}{k_i-k_{i-1}},\text{如果}k\in[k_{i-1},k_i],\\ \frac{k_{i+1}-k}{k_{i+1}-k_i},\text{如果}k\in[k_i,k_{i+1}],\\ 0,\end{cases}\tag{7.9b}$$

$$\psi_j(k)=\begin{cases}\frac{k-k_{j-1}}{k_j-k_{j-1}},\text{如果}k\in[k_{j-1},k_j],\\ 0.\end{cases}\tag{7.9c}$$

之所以说(7.9)的基方程是三角形式，举例说明，见图7.2。

图 7.2: 三角方程形式

第四步，与其他映射法类似，生成决策方程系统作为原方程系统的近似，在有限元分析中我们常称之为分段线性近似 (piecewise linear approximation)。

$$d^{n,j}\left(\cdot|\theta^n\right)=\sum_{i=0}^{j}\theta_i^n\psi_i\left(\cdot\right),$$

将近似系统代回 $\mathcal{H}$。其中，θ 和 ψ 由切比雪夫多项式的相关算法求得：以本例中的有限元节点 k_i 和 k_{i+1} 为例，分别对应基方程

$$\psi_i(k)=\frac{k_{i+1}-k}{k_{i+1}-k_i},$$
$$\psi_{i+1}(k)=\frac{k-k_i}{k_{i+1}-k_i},$$

那么由节点 k_i 和 k_{i+1} 划出的元中，值 $d^{n,j}\left(\cdot|\theta^n\right)$ 由一个近似方程 $\hat{d}$ 表示

$$\begin{aligned}\hat{d}\left(k|k_i,k_{i+1},\theta_i^n,\theta_{i+1}^n\right)&=\theta_i^n\psi_i(k)+\theta_{i+1}^n\psi_{i+1}(k)=\theta_i^n\frac{k_{i+1}-k}{k_{i+1}-k_i}+\theta_{i+1}^n\frac{k-k_i}{k_{i+1}-k_i}\\&=\frac{\left(\theta_{i+1}^n-\theta_i^n\right)k+\theta_i^nk_{i+1}-\theta_{i+1}^nk_i}{k_{i+1}-k_i},\end{aligned}$$

不难看出，近似方程 $\hat{d}$ 是线性的，斜率的符号由 $\theta_{i+1}^n-\theta_i^n$ 来决定。

另一方面采用类似的方法，以有限元结点 k_{i-1} 和 k_i 划出的元中，我们可以得到另一个线性近似方程 $\hat{d}$

$$\hat{d}\left(k|k_{i-1},k_i,\theta_{i-1}^n,\theta_i^n\right)=\frac{\left(\theta_i^n-\theta_{i-1}^n\right)k+\theta_{i-1}^n k_i-\theta_i^n k_{i-1}}{k_i-k_{i-1}}$$

两个元在 k_i 处相交，此时我们有

$$\begin{cases}\hat{d}\left(k_i|k_i,k_{i+1},\theta_i^n,\theta_{i+1}^n\right)=\theta_i^n,\\ \hat{d}\left(k_i|k_{i-1},k_i,\theta_{i-1}^n,\theta_i^n\right)=\theta_i^n,\end{cases}$$

即两个方程等于一个相同的值 θ_i^n，这确保了作为各个 $\hat{d}$ 的加总的 d 方程是连续的。

上述介绍也说明为何有限元法提供的分段线性近似是个较为理想的近似策略。假定我们设对应的度量方程 $\rho(\cdot)$ 的形式为[14]，使得在划分元所依据的每个节点 k_i 上，都有残差方程 $R(\cdot)=0$。 reference 对于三角函数形式的分段基方程 $\psi_i(\cdot)$ 而言，这意味着当状态变量处在 k_i 节点上时，对系数 θ_i^n 的取值应当使得近似方程系统与原方程系统相等：

$$\hat{d}^{n,j}(\cdot|\theta^n)=d^n(\cdot).$$

需要指出的是，有限元分析中的分段线性近似法也表明，系数 θ_i^n 的选取，与状态变量处于节点 k_i 之外其他位置时 $d^n(\cdot)$ 的值无关。从这个意义上来讲，利用有限元法求解大型非线性方程系统所得到的近似系统，都是稀疏的 (sparse)。这成为一个可为现代非线性求解方法所利用的特征。

第五步，(如有需要) 对结果的改进。如果第一轮有限元分析近似解的精度不达标，我们可以根据需要对近似解做反复改进，进行第二轮、第三轮甚至更多有限元分析，在计算时间和内存允许的范围内，尽可能提高近似解的精度。事实上这是有限元法研究的又一突出优势。在现有研究文献中，改进常常分为三大类型。

1. h——改进 (h-refinement)，在全部域 Ω 内，将第一轮所分元 (如 $A,B,C,\ldots$) 再均等分为更小的元 (如 $A1,A2,A3,\ldots$)，对这些细分元再反复迭代使用有限元法，改善近似解，直至精度达标。

2. r——改进 (r-refinement)，第二、第三甚至更多轮有限元分解的重点针对存在着明显非线性特征的局部。

3. p——改进 (p-refinement)，不对现有元再作细分，而是在元内部通过加入更多个基方程 (如更多的切比雪夫多项式) 来提高近似的阶数 (order)；如果元内现有阶数已足够高，则应用有限元法和光谱法组合的混合策略，常称为光谱元法 (spectral elements)。在自然科学和工程学领域，光谱元法已得到了广泛应用 (Šolín et al., 2003)。

有时，h——和 p——改进被混合在一起使用，称 hp——有限元法 (hp-adaptive finite elements)，可以使近似解以指数速度向真实解收敛 (Ciarlet, 2002)。尽管它的编程难度更高，计算时间更长，但 hp——有限元法可能是目前已知最强大的 DSGE 求解工具了，有助于求解甚至是最复杂的 DSGE 模型。关于这种方法的详细介绍，可参考Babuška and Suri (1994); Demkowicz (2006); Demkowicz et al. (2007)。

7.5　目标方程的选取

如前所述，映射法研究中也需要选取度量方程 ρ 作为目标方程。在未对 $\mathcal{H}$ 作过多限定的前提下，可视 $R(\cdot|\theta)$ 为最简单的单维情况[15]。此时对 $\rho\left(R\left(\cdot|\theta\right),\mathbf{0}\right)$ 的选取目标可设定为，使用加权残差

[14]度量方程的选取方案有数种，这里只举其中之一；更多讨论见下节。

[15]随着 $\mathcal{H}$ 的限定条件变多，$R(\cdot\theta)=\mathcal{H}\left(d^j\left(\cdot\theta\right)\right)$ 可能变成多维度的，但在这里我们暂不做讨论。

法，选取合适的 θ 向量，使得加权残差之的积分最接近于零，对应某个权重方程 $\phi_i : \Omega \mapsto \mathbb{R}$:

$$\rho\left(R\left(\cdot|\theta\right), \mathbf{0}\right) = \begin{cases} 0 & \text{如果} \quad \int_\Omega \phi_i(\boldsymbol{x}) R\left(\cdot|\theta\right) d\boldsymbol{x}, \quad i = 1, 2, \ldots, j+1, \\ 1, \end{cases}$$

这样问题变为在给定 $j+1$ 个权重方程 ϕ_i 的前提下，选择 θ 的值来求解积分系统

$$\int_\Omega \phi_i(\boldsymbol{x}) R(\cdot|\theta) d\boldsymbol{x}, \quad i = 0, 1, \ldots, j+1, \tag{7.10}$$

对此我们有一系列常用解法可供选择，如规模较小的系统可用牛顿算法 (Newton algorithm)，规模较大的系统可用莱文贝格——马夸特方法 (Levenberg-Marquardt algorithm)，等。然而需要指出的是，系统(7.10)可能无实解或者有多个解。关于如何将映射法应用到经济学经验研究中的理论依据，到目前为止我们所知不多——应用数学研究中大量关于映射法的文献，涉及解的存在性、收敛特性等问题的研究，主要针对自然科学和工程学领域，它们并不完全适用于经济学。事实上，对诸如(7.10)的经济系统而言，需要确保解满足 DSGE 模型的横截条件 (transversality condition)，从而使得状态变量处于稳定域内——在实际求解过程中，这边需要我们选择合适的初始猜测系数 θ_0，或是在求解过程中加入边界条件。

与基方程 ψ_i 类似，关于权重方程 ϕ_i 也存在一系列选择方案。下面我们介绍一下经济学研究中常见的权重方程的设定方法。

7.5.1 最小方差

将研究目标理解为如下变分法问题 (variational problem)

$$\min_\theta \int_\Omega R^2\left(\cdot|\theta\right) d\boldsymbol{x},$$

一阶条件为

$$\int_\Omega \frac{\partial R(\boldsymbol{x}|\theta)}{\partial \theta_{i-1}} R(\cdot|\theta) d\boldsymbol{x}, i = 1, 2, \ldots, j+1.$$

则权重方程的选取方案之一是将其定义为

$$\phi_i(\boldsymbol{x}) = \frac{\partial R(\boldsymbol{x}|\theta)}{\partial \theta_{i-1}}. \tag{7.11}$$

映射法研究中，利用变分法问题设定权重方程，其思路与计量经济学中的回归问题相近似 (见第7.1.1节)。

优缺点：

1. 优点：直观，易于理解。
2. 缺点：
   (a) 最小方差及几种变体都要求计算 $\frac{R(\boldsymbol{x}|\theta)}{\theta_{i-1}}$，导致计算成本高。
   (b) 最小方差问题过于复杂，条件苛刻，难于数值求解。

7.5.2 子域

将状态变量的域 Ω 通过一系列灵活的划分原则，分为 $j+1$ 个子域 (subdomain) Ω_i, $i = 1, 2, \ldots, j+1$，则研究目标可以理解为如下 $j+1$ 个求积问题

$$\int_{\Omega_i} R(\cdot|\theta) d\boldsymbol{x} = 0, \quad i = 1, 2, \ldots, j+_1 .$$

则子域法将权重设为如下 $j+1$ 个分段方程

$$\phi_i(\boldsymbol{x}) = \begin{cases} 1 & \text{如果 } \boldsymbol{x} \in \Omega_i, \\ 0. \end{cases}$$

优缺点基本同于第7.5.1节的最小方差法。

7.5.3　配点

配点法

配点法 (collocation method)，又称伪光谱法 (pseudospectral method)或选点法 (selected points method)，是指在状态变量域 Ω 中选取 $j+1$ 个配点 $\boldsymbol{x}_i$，将权重方程定义为

$$\psi_i(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_i), i = 1, 2, \ldots, j+1,$$

其中 δ 是狄拉克方程 (dirac delta function, 第7.G.3节)，满足

$$\delta(\boldsymbol{y}_i) = \begin{cases} +\infty, & \text{如果 } \boldsymbol{y}_i \equiv \boldsymbol{x} - \boldsymbol{x}_i = 0, \\ 0, & \text{如果 } \boldsymbol{y}_i \neq 0, \end{cases}$$

$$\int_{-\infty}^{+\infty} \delta(\boldsymbol{y}_i) d\boldsymbol{y}_i = 1.$$

配点法假定当状态变量恰好处于选点时 $\boldsymbol{x} = \boldsymbol{x}_i$，权重的值为 0，残差方程因此也为 0。此时不必做复杂的求积计算，只需求解 $j+1$ 个方程系统

$$R(\boldsymbol{x}_i|\theta) = 0, \quad j = 1, 2, \ldots, j+1.$$

当原方程系统 $\mathcal{H}$ 表现出较明显的非线性特征时，配点近似方法具有较大优势。

配点的选取：正交配点法

$j+1$ 个配点的选取方法有很多，一个常用方法称正交配点法 (orthogonal collocation)，是指在值域 Ω 中对于向量 $\boldsymbol{x}$ 对应的每一组状态变量，都用 $j+1$ 次切比雪夫多项式的根 (有 $j+1$ 个)予以表示。根据切比雪夫插值定理 (定理7.56)，利用正交配点法生成的近似方程系统，会收敛至、甚至有时是均匀收敛 (uniform convergence)至未知方程 d。

7.5.4　伽辽金法或瑞利——里兹法

。

完备 (completeness) 的脚注218

附录

7.A 超几何方程

Definition 7.1 (超几何序列)**.** *一组序列 $\{c_n\}$，如果 $\frac{c_{n+1}}{c_n}$ 是一个关于 n 的有理方程，则我们称 $\{c_n\}$ 是一个超几何序列 (Hypergeometric Series)。*

以阶乘形式表示为

$$\frac{c_{n+1}}{c_n} = \frac{(n+a_1)(n+a_2)\dots(n+a_p)}{(n+b_1)(n+b_2)\dots(n+b_q)}\frac{z}{(n+1)}, \quad n = 0, 1, 2\dots, \tag{7.12}$$

如果 $z = 1$，$\{c_n\}$ 又称首一超几何序列 (monic hypergeometric series)；如果 $z \neq 1$，则称非首一超几何序列。分母中的 $n+1$ 项可以是来自阶乘本身，也可以不是，对于后者的情况，我们也可以在分子中加入一个额外的 $(n+a_i)$ 项予以调节，即选取一个 $i \in 0, 1\dots n+1$ 使得 $a_i = 1$。

对(7.12)做迭代我们有

$$c_n = \frac{(a_1)_n (a_2)_n \dots (a_p)_n}{(b_1)_n (b_1)_n \dots (b_q)_n}\frac{z^n}{n!}c_0, \tag{7.13}$$

其中 $n!$ 表示 n 的阶乘。$(a_1)_n$ 表示伯赫哈默尔符号 (Pochhammer symbol)的定义如下

$$(a_1)_n := \begin{cases} 1 & \text{如果} n = 0, \\ a_1(a_1+1)(a_1+2)\dots(a_1+n-1) & \text{如果} n = 1, 2, 3\dots. \end{cases}$$

对超几何序列 $\{c_n\}$ 的求和为

$$\sum_{n=0}^{\infty} c_n = c_0 \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \dots (a_p)_n}{(b_1)_n (b_1)_n \dots (b_q)_n}\frac{z^n}{n!},$$

这引出超几何方程的定义。

Definition 7.2 (超几何方程)**.** *超几何方程 (hypergeometric function)* ${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; x)$ *定义为超几何序列 $\{c_n\}$ 的和：*

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; x) = \sum_{k=0}^{n} c_n = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \dots (a_p)_n}{(b_1)_n (b_1)_n \dots (b_q)_n}\frac{z^n}{n!} \tag{7.14}$$

显然，上式成立要求分母的系数 $\neq 0$。对于分子的系数，如果某一个 i 次系数 $a_i = -N$，其中 N 是个非负整数，那么超几何方程 ${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; x)$ 是一个关于 z 的多项式，详见定义7.3；反之，则超几何序列的收敛速度 ρ 满足

$$\rho = \begin{cases} \infty & \text{如果} p < q+1, \\ 1 & \text{如果} p = q+1, \\ 0 & \text{如果} p > q+1, \end{cases}$$

这是由超几何序列的性质所决定的：由(7.12)可得

$$\lim_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right| = \begin{cases} \infty & \text{如果}\, p < q+1, \\ |z| & \text{如果}\, p = q+1, \\ \infty & \text{如果}\, p > q+1, \end{cases}$$

我们最为关注 $p = q+1$ 的情况：随着 $|z|$ 的取值不同，超几何函数 ${}_{q+1}F_q(a_1,\ldots,a_{q+1};b_1,\ldots,b_q;z)$ 呈现出不同的收敛特点：

$${}_{q+1}F_q(a_1,\ldots,a_{q+1};b_1,\ldots,b_q;z) \begin{cases} \text{收敛}, & \text{如果}\ |z|<1, \\ \text{收敛}, & \text{如果}\ |z|=1, Re\left(\sum^q b_i - \sum^{q+1} a_j\right) > 0, \\ \text{有条件收敛}, & \text{如果}\ |z|=1, z\neq -1, -1 < Re\left(\sum^q b_i - \sum^{q+1} a_j\right) \leq 0, \\ \text{不收敛}, & \text{如果}\ Re\left(\sum^q b_i - \sum^{q+1} a_j\right) \leq -1. \end{cases}$$

Definition 7.3 (广义超几何方程)**.** *广义超几何方程 (generalized hypergeometric function)是超几何方程的一种特殊形式，定义式为*

$${}_2F_1(a,b;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n\,(b)_n}{(c)_n}\frac{z^n}{n!}, \tag{7.15}$$

其中如果 $a=-N$, $N=0,1,2\ldots$，那么我们有

$$(a)_n = (-N)_n = (-N)\,(-N+1)\,(-N+2)\ldots(-N+n-1) = 0,\ n = N+1, N+2, N+3\ldots.$$

那么(7.15)变为

$${}_2F_1(-N,b;c;z) = \sum_{n=0}^{N} \frac{(-N)_n\,(b)_n}{(b)_n}\frac{z^n}{n!}, \quad N=0,1,2\ldots. \tag{7.16}$$

如果 a_n 不是个常数，那么超几何序列想要获得收敛特性，需要满足如下情况，第一种情况是 $|z|<1$，第二种情况是 $|z|=1$ 以及 $Re(c-a-b)>0$。

许多函数都可以表示为超几何方程的形式。如

$$\ln(1+x) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}x^{n+1} = \sum_{n=0}^{\infty}\frac{(1)_n(1)_n}{(2)_n}\frac{(-1)^n x^{n+1}}{n!} = \sum_{n=0}^{\infty}\frac{n!\,n!}{(n+1)!}\frac{(-1)^n x^{n+1}}{n!} = x\,{}_2F_1\left(\begin{matrix}1\ 1\\ 2\end{matrix};-x\right),$$

$$\arctan x = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}x^{2n+1} = \sum_{n=0}^{\infty}\frac{\left(\frac{1}{2}\right)_n(1)_n}{\left(\frac{3}{2}\right)_n}\frac{(-1)^n x^{2n+1}}{n!} = x\,{}_2F_1\left(\begin{matrix}\frac{1}{2}\ 1\\ \frac{3}{2}\end{matrix};-x^2\right),$$

$$\text{最后一个等式是由于}\frac{\left(\frac{1}{2}\right)_n}{\left(\frac{3}{2}\right)_n} = \frac{\frac{1}{2}\frac{3}{2}\cdots\frac{2n-1}{2}}{\frac{3}{2}\frac{5}{2}\cdots\frac{2n+1}{2}} = \frac{\frac{1}{2}}{\frac{2n+1}{2}} = \frac{1}{2n+1}$$

Theorem 7.1 (二项式定理)**.** *作为超几何方程的特殊形式之一，可以写成如下形式，称为二项式定理 (binomial theorem)*

$${}_2F_1\left(\begin{matrix}a\ b\\ b\end{matrix};z\right) = {}_1F_0\left(\begin{matrix}a\\ _\end{matrix};z\right) = \sum_{n=0}^{\infty}\frac{(a)_n}{n!}z^n = \sum_{n=0}^{\infty}\binom{-a}{n}\frac{(a)_n}{n!}(-z)^n = (1-z)^{-a}, \quad |z|<1. \tag{7.17}$$

以及还有一类特殊的超几何方程

$${}_0F_0\left(\begin{matrix}_\\ _\end{matrix};z\right) = \sum_{t=0}^{\infty}\frac{z^n}{n!} = e^z, \quad z\in\mathbb{C}. \tag{7.18}$$

以及当系数趋近于无穷时：

$$\lim_{b\to\infty} {}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix}; \frac{z}{b}\right) = \lim_{b\to\infty} \sum_{n=0}^{\infty} \frac{(a)_n z^n}{(c)_n n!} \frac{(b)_n}{b^n} = \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n} \frac{z^n}{n!} = {}_1F_1\left(\begin{matrix} a \\ c \end{matrix}; z\right),$$

$$\lim_{c\to\infty} {}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix}; cz\right) = \lim_{c\to\infty} \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{n!} z^n \frac{c^n}{(c)_n} = \sum_{n=0}^{\infty} (a)_n\,(b)_n \frac{z^n}{n!} = {}_2F_0\left(\begin{matrix} a & b \\ _ \end{matrix}; z\right).$$

Theorem 7.2 (超几何方程与欧拉积分表达式的转换)**.** *如果 $Re(c) > Re(b) > 0$，那么对于沿着从 1 到 ∞ 实数轴展开的所有 $z \in \mathbb{C}$，满足 $\arg t = \arg(1-t) = 0$ 并且 $(1-zt)^{-a}$ 有主值 (principal value) 的条件*[16]*，我们可以将超几何方程转换为欧拉积分表达式 (Euler's integral representation)*

$$ {}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix}; z\right) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}\,dt. \tag{7.19}$$

证明. 假设 $|z| < 1$，那么由二项式定理(7.17)可得

$$(1-zt)^{-a} = \sum_{n=0}^{\infty} \frac{(a)_n}{n!} z^n t^n,$$

这表明

$$\int_0^t t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}dt = \underbrace{\sum_{n=0}^{\infty} \frac{(a)_n}{n!} z^n t^n}_{\text{RHS 甲}} \underbrace{\int_0^t t^{b-1}(1-t)^{c-b-1}dt}_{\text{RHS 乙}},$$

RHS 分为两部分。乙是一个贝塔积分 (Beta integral)，满足条件[17]

$$\int_0^t t^{b-1}(1-t)^{c-b-1}dt = B(n+b, c-b) = \frac{\Gamma(n+b)\Gamma(c-b)}{\Gamma(n+c)}.$$

此外 Gamm 方程可以转换为 Pochhammer 符号的表达形式：

$$\frac{\Gamma(n+b)}{\Gamma(b)} = b(b+1)(b+2)\ldots(b+n-1) = (b)_n,$$

那么

$$\frac{\Gamma(n+b)\Gamma(c-b)}{\Gamma(n+c)} = \Gamma(c-b)\frac{\Gamma(b)}{\Gamma(c)}\frac{(b)_n}{(c)_n},$$

作为 RHS 乙代回原式有

$$\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^t t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}\,dt = \sum_{0}^{\infty} \frac{(a)_n\,(b)_n}{(c)_n} \frac{z^n}{n!} = {}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix}; z\right).$$

□

Theorem 7.3 (欧拉积分表达式与高斯求和公式的转换)**.** *对于 $Re(c-a-b) > 0$，我们有欧拉积分表达式与高斯求和公式 (Gauss summation formula)的转换*

$$ {}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix}; 1\right) = \frac{\Gamma(c)\,\Gamma(c-a-b)}{\Gamma(c-a)\,\Gamma(c-b)}. \tag{7.20}$$

[16]膏按：反正这一串前提限定条件我没看懂... 我的理解是，经济学研究中，能够熟悉下面的转换公式就可以了——这真是一本不负责任的 "讲义" 啊...

[17] 其中 $\Gamma(\cdot)$ 是 Gamma 方程，是阶乘函数在实数与复数上的扩展，定义为

$$\Gamma(a) = \begin{cases} (n-1)! & \text{n 是正整数}, \\ \int_0^\infty t^{a-1}\exp(-t)\,dt & z \in \mathbb{C}, Re(c) > 0. \end{cases}$$

证明. 假设 $\lim_{z\to 1}$，通过(7.19)我们有

$$\lim_{z\to 1} {}_2F_1\left(\begin{matrix} a & b \\ & c \end{matrix}; z\right) \approx {}_2F_1\left(\begin{matrix} a & b \\ & c \end{matrix}; 1\right) = \frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-a-b-1}\,dt$$
$$= \frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)} B(b, c-a-b) = \frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)} \frac{\Gamma(b)\Gamma(c-a-b)}{\Gamma(c-a)} = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}.$$

□

Theorem 7.4 (高斯求和公式与朱世杰——范德蒙德求和公式的转换). *当 $a = -n, n \in \{0, 1, 2\ldots\}$(并且 $\lim_{z\to 1}$) 时，高斯求和公式(7.20)缩减成为朱世杰——范德蒙德求和公式 (Chu-Vandermonde summation)*

$${}_2F_1\left(\begin{matrix} -n & b \\ & c \end{matrix}; 1\right) = \frac{(c-b)_n}{(c)_n}, n = 0, 1, 2\ldots. \tag{7.21}$$

证明.

$$\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} = \frac{\Gamma(c)\Gamma(c-b+n)}{\Gamma(c+n)\Gamma(c-b)} = \frac{(c-b)_n}{c(n)}.$$

□

Theorem 7.5 (超几何方程与法夫转换公式). *我们可以用欧拉积分表达式证明，超几何方程可以写成法夫转换公式 (Pfaff's transformation formula)的形式*

$${}_2F_1\left(\begin{matrix} a & b \\ & c \end{matrix}; z\right) = (1-z)^{-a}\,{}_2F_1\left(\begin{matrix} a & c-b \\ & c \end{matrix}; \frac{z}{z-1}\right). \tag{7.22}$$

证明. 定义 $t = 1 - s$，代入欧拉积分表达式(7.19)我们有

$$\begin{aligned}
{}_2F_1\left(\begin{matrix} a & b \\ & c \end{matrix}; z\right) &= \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}\,dt \\
&= -\int_0^1 (1-s)^{b-1}s^{c-b-1}(1-z+zs)^{-a}\,ds \\
&= (1-z)^{-a}\int_0^1 s^{c-b-1}(1-s)^{b-1}\left(1-\frac{zs}{z-1}\right)^{-a} ds \\
&= (1-z)^{-a}\,{}_2F_1\left(\begin{matrix} a & c-b \\ & c \end{matrix}; \frac{z}{z-1}\right).
\end{aligned}$$

□

Theorem 7.6 (超几何方程与欧拉转换公式). *我们可以用法夫转换公式证明，超几何方程可以写成欧拉转换公式 (Euler's transformation formula)的形式*

$${}_2F_1\left(\begin{matrix} a & b \\ & c \end{matrix}; z\right) = (1-z)^{c-a-b}\,{}_2F_1\left(\begin{matrix} c-a & c-b \\ & c \end{matrix}; z\right). \tag{7.23}$$

证明. 两次使用法夫转换公式

$$\begin{aligned}
{}_2F_1\left(\begin{matrix} a & b \\ & c \end{matrix}; z\right) &= (1-z)^{a}\,{}_2F_1\left(\begin{matrix} a & c-b \\ & c \end{matrix}; \frac{z}{z-1}\right) \\
&= (1-z)^{a}(1-z)^{b-c}\,{}_2F_1\left(\begin{matrix} c-a & c-b \\ & c \end{matrix}; \frac{\frac{z}{z-1}}{\frac{z}{z-1}-1}\right) \\
&= (1-z)^{c-a-b}\,{}_2F_1\left(\begin{matrix} c-a & c-b \\ & c \end{matrix}; z\right)
\end{aligned}$$

□

Theorem 7.7 (法夫——萨尔舒茨求和公式)**.** 一个 ${}_3F_2$ 超几何方程又可以写成法夫——萨尔舒茨求和公式 *(Pfaff-Saalschütz summation formula)*的形式

$$ {}_3F_2\left(\begin{matrix} -n & a & b \\ c & 1+a+b-c-n \end{matrix};1\right) = \frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}, \quad n=0,1,2\ldots. \tag{7.24} $$

证明. 欧拉转换公式(7.23)又可以写成

$$ (1-z)^{a+b-c}{}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix};z\right) = {}_2F_1\left(\begin{matrix} c-a & c-b \\ c \end{matrix};z\right), $$

其中 LHS→

$$ \begin{aligned} (1-z)^{a+b-c}{}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix};z\right) &= \sum_{j=0}^{\infty}\frac{(c-a-b)_j}{j!}z^j\sum_{k=0}^{\infty}\frac{(a)_k(b)_k}{(c)_k}\frac{z^k}{k!} \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{(a)_k(b)_k(c-a-b)_{n-k}}{(c)_k k!(n-k)!}z^n. \end{aligned} $$

RHS→

$$ {}_2F_1\left(\begin{matrix} c-a & c-b \\ c \end{matrix};z\right) = \sum_{n=0}^{\infty}\frac{(c-a)_n(c-b)_n}{(c)_n}\frac{z^n}{n!}. $$

比较两个式子中 z^n 对应的系数，我们有

$$ \sum_{k=0}^{n}\frac{(a)_k(b)_k(c-a-b)_{n-k}}{(c)_k k!(n-k)!}z^n = \frac{(c-a)_n(c-b)_n}{(c)_n n!}, \quad n=0,1,2\ldots, \quad k\in\{0,1,2,\ldots n\}. $$

□

其中由于

$$ \frac{n!}{(n-k)!} = n(n-1)\ldots(n-k+1) = (-1)^k(-n)(-n+1)\ldots(-n+k-1) = (-1)^k(-n)^k, $$

以及

$$ \begin{aligned} (c-a-b)_{n-k} &= \frac{(c-a-b)_n}{(c-a-b+n-k)(c-a-b+n-k-1)\ldots(c-a-b+n-1)} \\ &= \frac{(c-a-b)_n}{(-1)^k(1+a+b-c-n)_k}. \end{aligned} $$

因此

$$ \sum_{k=0}^{n}\frac{(a)_k(b)_k(-n)_k}{(c)_k k!(1+a+b-c-n)_k} = \frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}, \quad n=0,1,2\ldots, \quad k\in\{0,1,2,\ldots n\}, $$

进而(7.24)成立。

Theorem 7.8 (法夫——萨尔舒茨求和公式与高斯求和公式的转换)**.** 随着 $n\to\infty$，${}_3F_2$ 的法夫——萨尔舒茨求和公式(7.24)逐渐缩减到 ${}_2F_1$ 的高斯求和公式(7.20)的形式

$$ \lim_{n\to\infty}{}_3F_2\left(\begin{matrix} -n & a & b \\ c & 1+a+b-c-n \end{matrix};1\right) = {}_2F_1\left(\begin{matrix} a & b \\ c \end{matrix};1\right). \tag{7.25} $$

证明. 根据

$$ (a)_n = a(a+1)\ldots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)}, $$

我们有

$$\begin{aligned}\frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n} &= \frac{\Gamma(c-a+n)\Gamma(c-b+n)\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)\Gamma(c+n)\Gamma(c-a-b+n)} \\ &= \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}\underbrace{\frac{\Gamma(c-a+n)\Gamma(c-b+n)}{\Gamma(c+n)\Gamma(c-a-b+n)}},\end{aligned}$$

上式中下括号部分可由斯特灵公式 (Stirling formula) 近似:

$$\lim_{n\to 0}\frac{\Gamma(c-a+n)\Gamma(c-b+n)}{\Gamma(c+n)\Gamma(c-a-b+n)} \sim n^{c-a+c-b-c-c+a+b} = n^0 = 1,$$

即

$$\lim_{n\to\infty}\frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n} = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)},$$

由此得到

$$\begin{aligned}{}_2F_1\left(\begin{matrix}a\ b\\ c\end{matrix};1\right) &= \lim_{n\to\infty}{}_3F_2\left(\begin{matrix}-n\ a\ b\\ c\ 1+a+b-c-n\end{matrix};1\right) \\ &= \lim_{n\to\infty}\frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n} \\ &= \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}.\end{aligned}$$

□

7.B 多项式

7.B.1 正交集

数学意义上来说，正交 (orthogonal) 意味着垂直 (perpendicular)。如，一个三维向量空间中的向量集 $\{x,y,z\}$ 是正交的，意思就是说任意两个正交向量的点乘 (dot product) 都为零，$x.y=0, x.z=0, y.z=0$。因此我们又将 $\{x,y,z\}$ 称为正交集合。在 $\{x,y,z\}$ 构成的三维空间中，任何一点都可以表示为 x,y 或 z 向量中的一项——换句话说，$\{x,y,z\}$ 共同构成了一个三维空间的基 (basis)。

如果 $\{x,y,z\}$ 都是单位向量，则进一步称之为标准正交向量集。

Definition 7.4 (黎曼——斯蒂尔杰斯积分)**.** *在区间 $[a,b]$ 内，假定对于一个实变量 λ 的实函数 $\alpha(\lambda)$，存在一个对应的实函数 $f(\lambda)$，则黎曼——斯蒂尔杰斯积分 (Riemann-Stieltjes integral)$\int_a^b f(\lambda)\,d\alpha(\lambda)$ 描述了当将区间 $[a,b]$ 分割为无限多个块 (数量用 π 表示) 时，下述和的极限值*

$$\int_a^b f(\lambda)\,d\alpha(\lambda) = \sum_{\lambda_i\in\pi} f(c_i)\left(\alpha(\lambda_{i+1})-\alpha(\lambda_i)\right), \quad c_i\in[\lambda_{i+1},\lambda_i]. \tag{7.26}$$

黎曼——斯蒂尔杰斯积分值若要存在，需要 $f(\cdot)$ 是连续方程，以及非递减方程 $\alpha(\cdot)$ 在 $[a,b]$ 中有界。

黎曼——斯蒂尔杰斯积分常常直接写作简化形式

$$\int_a^b f(\lambda)w(\lambda)\,d\lambda, \tag{7.27}$$

其中 $w(\cdot)$ 称权重方程 (weight function)。

7.B.2　正交多项式

任意一个无限多项式序列 $\{p_n(\lambda)\}_{n=0}^{\infty}$ 向量 (其中 $p_n(\lambda)$ 表示其中的第 n 次多项式)，都构成无限维度向量空间中的一个基。

Definition 7.5 (正交多项式)**.** *如果一个多项式序列 $\{p_n(\lambda)\}_{n=0}^{\infty}$ 在区间 (a,b) 中，关于权重方程 $w(\lambda)$，满足如下关系，那么我们说它是正交多项式 (orthogonal polynomial)*:

$$\int_a^b w(\lambda)p_m(\lambda)p_n(\lambda)d\lambda = h_n\delta_{mn} \tag{7.28}$$

- h_n 称范数 (norm) ，正交多项式 $p(\lambda)$ 的范数也表示为 $|p|$，定义为 (具体细节见第8.2.2。)

$$h_n = |p_n(\lambda)| = \int_a^b w(\lambda)p_n(\lambda)^2\,dx > 0, n = 0,1,2\ldots \tag{7.29}$$

- 权重方程 $w(\lambda)$ 应当在 (a,b) 区间内连续且为正，以使矩 (moment) μ_n 存在。
- 矩 (moment) μ_n 定义为

$$\mu_n = \int_a^b w(x)\lambda^n\,d\lambda, \quad i = 0,1,2\ldots \tag{7.30}$$

- δ_{mn} 是克罗内克 (Kronecker product) ，表达如下关系

$$\delta_{m,n} = \begin{cases} 0 & m \neq n, \\ 1 & m = n. \end{cases} \tag{7.31}$$

- 积分 $\langle f,g\rangle$ 表示多项式 f 和 g 的内积 (inner product)

$$\langle p,q\rangle = \int_a^b w(\lambda)f(\lambda)g(\lambda)\,d\lambda. \tag{7.32}$$

- 区间 (a,b) 称正交区间 (可以是一个无限区间，即 $a \to -\infty$, $b \to +\infty$)。

如果正交多项式 $\{p_n(\lambda)\}_{n=0}^{\infty}$ 对于 $k = 0,1,2\ldots n$ 都有 $h_k = 1$，则称标准正交多项式 (orthonormal polynomial)。

如果第 n 次多项式 $p_n(\lambda) = k_n\lambda^n + \chi$，其中 χ 表示低于 n 次的其他项，并且对于每一个 $n = 0,1\ldots$ 都有 $k_n = 1$，则称首一正交多项式 (monic orthogonal polynomial)。

标准正交多项式 $p_n(\lambda)$ 和首一正交多项式 $\hat{p}_n(\lambda)$ 之间可以相互转换（如见定理7.12）。

$$\hat{p}_n(\lambda) = \frac{p_n(\lambda)}{\sqrt{\langle p_n(\lambda), p_n(\lambda)\rangle}}. \tag{7.33}$$

我们可以根据格拉姆-施密特正交化过程 (Gram-Schmidt orthogonality process)，把正交多项式序列 $\{p_n(\lambda)\}_{n=0}^{\infty}$ 转化为正交集合。举例说明，取 $w(\lambda) = 1$, $(a,b) = (0,1)$, $p_0(\lambda) = 1$，求首一多项式序列 $\{p_n(\lambda)\}$。

从序列 $\{1,\lambda,\lambda^2\ldots\}$ 开始。第一步求 $p_1(\lambda)$

$$\begin{aligned} p_1(\lambda) &= \lambda - \frac{\langle\lambda, p_0(\lambda)\rangle}{\langle p_0(\lambda)\lambda, p_0(\lambda)\rangle}\,p_0(\lambda) \\ &= \lambda - \frac{\langle\lambda,1\rangle}{\langle 1,1\rangle} \\ &= \lambda - \frac{\int_0^1 w(\lambda)\,(\lambda\cdot 1)\;d\lambda}{\int_0^1 w(\lambda)\,(1\cdot 1)\;d\lambda} \\ &= \lambda - \frac{1}{2}. \end{aligned}$$

第一步求 $p_2(\lambda)$

$$\begin{aligned}
p_2(\lambda) &= \lambda^2 - \frac{\langle \lambda^2, p_0(\lambda)\rangle}{\langle p_0(\lambda), p_0(\lambda)\rangle} p_0(\lambda) - \frac{\langle \lambda^2, p_1(\lambda)\rangle}{\langle p_1(\lambda), p_1(\lambda)\rangle} p_1(\lambda) \\
&= \lambda^2 - \frac{\langle \lambda^2, 1\rangle}{\langle 1, 1\rangle} - \frac{\langle \lambda^2, \lambda - \frac{1}{2}\rangle}{\langle \lambda - \frac{1}{2}, \lambda - \frac{1}{2}\rangle}\left(\lambda - \frac{1}{2}\right) \\
&= \lambda^2 - \frac{\frac{1}{3}}{1} - \frac{\frac{1}{12}}{\frac{1}{12}}\left(\lambda - \frac{1}{2}\right) \\
&= \lambda^2 - \lambda + \frac{1}{6}.
\end{aligned}$$

则 $p_0(\lambda) = 1$，$p_1(\lambda) = \lambda - \frac{1}{2}$，$p_2(\lambda) = \lambda^2 - \lambda + \frac{1}{6}$ 就是首一正交多项式在 (0,1) 区间内，对应权重方程 $w(\lambda) = 1$ 的前三项。以此类推，可以继续求得 $p_3(\lambda), p_4(\lambda) \ldots$

$$\begin{aligned}
p_3(\lambda) &= \lambda^3 - \frac{3}{2}\lambda^2 + \frac{3}{5}\lambda - \frac{1}{20}, \\
p_4(\lambda) &= \lambda^4 - 2\lambda^3 + \frac{9}{7}\lambda^2 - \frac{2}{7}\lambda + \frac{1}{70}, \\
p_5(\lambda) &= \lambda^5 - \frac{5}{2}\lambda^4 + \frac{20}{9}\lambda^3 - \frac{5}{6}\lambda^2 + \frac{5}{42}\lambda - \frac{1}{252} \ldots
\end{aligned}$$

根据(7.29)计算 h_n；与首一正交多项式 $p_n(\lambda)$ 对应的正交多项式 $q_n(\lambda)$ 由(7.33)给出

$$\begin{aligned}
q_1(\lambda) &= \frac{p_1(\lambda)}{\sqrt{h_1}} = 2\sqrt{3}\left(x - \frac{1}{2}\right), \\
q_2(\lambda) &= \frac{p_2(\lambda)}{\sqrt{h_2}} = 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right), \\
q_3(\lambda) &= \frac{p_3(\lambda)}{\sqrt{h_3}} = 20\sqrt{7}\left(\lambda^3 - \frac{3}{2}\lambda^2 + \frac{3}{5}\lambda - \frac{1}{20}\right) \ldots
\end{aligned}$$

正定

如果 $|p| > 0, |q| > 0, \forall p, q \in \mathbb{P}$，则 $< .,. >$ 是正定的。换句话说，内积是正定的充分必要条件为，它的汉克尔矩阵 (Hankel moment matrix)的行列式为正

$$\det \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_{k-1} \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_k \\ \vdots & \ddots & & & \vdots \\ \mu_{k-1} & \mu_k & \mu_{k+1} & \cdots & \mu_{2k-2} \end{pmatrix} > 0, k = 1, 2, \ldots \tag{7.34}$$

其中 μ_i 的定义如(7.30)所示。

Theorem 7.9 (正交多项式的存在性)**.** *如果某内积 $\langle .,. \rangle$ 在 $\mathbb{P}$ 中正定，那么有且只有一个首一正交的无限多项式与每一个 $\alpha(\lambda)$ 相对应。*

证明. 略。 □

Theorem 7.10 (正交多项式的最小化特性)**.** *如果 $q_k(\lambda)$ 是一个 k 次的首一多项式，则只有当 q_k 等于一个常数与另一个同样关于 $\alpha(\lambda)$ 的多项式 p_k 的乘积时，以下最小值才存在*

$$\min_{q_k} \int_a^b q_k(\lambda)^2 \, d\alpha(\lambda).$$

证明. 略。 □

7.B.3　递推关系

Theorem 7.11 (首一正交多项式的三项递推关系)**.** *对于首一正交多项式序列 $\{p_n(\lambda)\}$ 而言，存在着系数 α_k 和 γ_k，$k=1,2,\ldots$，使得其满足三项递推关系 (three-term recurrence relation)*

$$\begin{aligned}
&p_{k+1}(\lambda)=(\lambda-\alpha_{k+1})\,p_k(\lambda)-\gamma_k p_{k-1}(\lambda),\quad k=1,2,\ldots,\\
&p_{-1}(\lambda):=0,\,p_0(\lambda)=1,\\
&\text{其中}\quad \alpha_{k+1}=\frac{\langle\lambda p_k,p_k\rangle}{\langle p_k,p_k\rangle},\,\gamma_k=\frac{\langle p_k,p_k\rangle}{\langle p_{k-1},p_{k-1}\rangle}.
\end{aligned}\tag{7.35}$$

证明. 对于 $k=1,2,\ldots$ 设

$$F(\lambda):=p_{k+1}(\lambda)-(\lambda-\alpha_{k+1})p_k(\lambda)+\gamma_k\,p_{k-1}(\lambda),$$

由 p_{k-1},p_k,p_{k+1} 都是首一多项式可得 $F(\lambda)\in\mathbb{P}_n[\lambda]$。此外，由 p_{k-1},p_k,p_{k+1} 的正交特性，我们分三种情况展开分析。

第一种情况，当 $0<l<k-1$ 时，有

$$\langle F(\lambda),p_l(\lambda)\rangle=\underbrace{\langle p_{k+1}(\lambda),p_l(\lambda)\rangle}_{=0}-\underbrace{\langle p_k(\lambda),(\lambda-\alpha_{k+1})\,p_l(\lambda)\rangle}_{=0}+\underbrace{\langle\gamma_k p_{k-1}(\lambda),p_l(\lambda)\rangle}_{=0}=0,$$

其中用到了 $\langle\lambda p,q\rangle=\langle p,\lambda q\rangle$，以及 $\langle p_k,p_l\rangle=0,\,\forall\,0<l<k-1$。

第二种情况，当 $l=k-1$ 时，有

$$\begin{aligned}
\langle F(\lambda),p_l(\lambda)\rangle&=\langle F(\lambda),p_{k-1}(\lambda)\rangle\\
&=\underbrace{\langle p_{k+1}(\lambda),p_{k-1}(\lambda)\rangle}_{=0}-\langle p_k(\lambda),(\lambda-\alpha_{k+1})\,p_{k-1}(\lambda)\rangle+\langle\gamma_k p_{k-1}(\lambda),p_{k-1}(\lambda)\rangle\\
&=-\langle p_k(\lambda),\lambda p_{k-1}(\lambda)\rangle+\underbrace{\langle p_k(\lambda),\alpha_{k+1}p_{k-1}(\lambda)\rangle}_{=0}+\gamma_k\langle p_{k-1}(\lambda),p_{k-1}(\lambda)\rangle=0,
\end{aligned}$$

因此我们有

$$\gamma_k=\frac{\langle p_k(\lambda),p_k(\lambda)\rangle}{\langle p_{k-1}(\lambda),p_{k-1}(\lambda)\rangle}.$$

第三种情况，当 $l=k$ 时，有

$$\begin{aligned}
\langle F(\lambda),p_l(\lambda)\rangle&=\langle F(\lambda),p_k(\lambda)\rangle\\
&=\underbrace{\langle p_{k+1}(\lambda),p_k(\lambda)\rangle}_{=0}+\langle(-\lambda+\alpha_{k+1})\,p_k(\lambda),p_k(\lambda)\rangle+\underbrace{\langle\gamma_k p_{k-1}(\lambda),p_k(\lambda)\rangle}_{=0}\\
&=-\langle\lambda p_k(\lambda),p_k(\lambda)\rangle+\alpha_{k+1}\langle p_k(\lambda),p_k(\lambda)\rangle=0,
\end{aligned}$$

因此我们有

$$\alpha_{k+1}=\frac{\langle\lambda p_k(\lambda),p_k(\lambda)\rangle}{\langle p_k(\lambda),p_k(\lambda)\rangle}.$$

□

Theorem 7.12 (标准正交多项式的三项递推关系)**.** *对于标准正交多项式 $\{p_n(\lambda)\}$ 而言，存在着系数 α_k 和 β_k，$k=0,1,2,\ldots$，使得其满足三项递推关系 (three-term recurrence relation)*

$$\begin{aligned}
&\sqrt{\beta_{k+1}}p_{k+1}(\lambda)=(\lambda-\alpha_{k+1})-\sqrt{\beta_k}p_{k-1}(\lambda),\quad k=1,2,\ldots\\
&p_{-1}(\lambda):=0,\,p_0(\lambda)=\beta_0^{-\frac{1}{2}},\,\beta_0=\int_a^b d\alpha(\lambda),\\
&\text{其中}\quad \alpha_{k+1}=\langle\lambda p_k,p_k\rangle,\,\beta_k\text{ 通过计算 }|p_k|=1\text{ 求得.}
\end{aligned}$$

证明. 假定有一个首一正交多项式 $\{p_k(\lambda)\}_{n=0}^{\infty}$ 满足三项递推关系(7.35)，那么根据(7.33)，对应的标准正交多项式 $\{\hat{p}_n(\lambda)\}_{n=0}^{\infty}$ 满足

$$\hat{p}_k(\lambda) = \frac{p_k(\lambda)}{\sqrt{\langle p_k, p_k\rangle}} = \frac{p_k(\lambda)}{|p_k|},$$

最后一个等号由(7.29)-(7.32)联立求得。上式代回(7.35)可得

$$\hat{p}_{k+1}(\lambda)\,|p_k| = \left(\lambda\,|p_{k+1}| - \frac{\langle \lambda p_k, p_k\rangle}{|p_k|}\right)\hat{p}_k - \left(\frac{|p_k|^2}{|p_{k-1}|}\right)\hat{p}_{k-1},$$

进一步整理

$$\frac{|p_{k+1}|}{|p_k|}\hat{p}_{k+1} = (\lambda - \langle \lambda\hat{p}_k, \hat{p}_k\rangle)\,\hat{p}_k - \frac{|p_k|}{|p_{k-1}|}\hat{p}_{k-1},$$

由此我们有

$$\langle \lambda\hat{p}_k, \hat{p}_k\rangle = \frac{\langle \lambda p_k, p_k\rangle}{\langle p_k, p_k\rangle} = \alpha_{k+1}, \quad \sqrt{\beta_{k+1}} = \frac{\langle p_{k+1}, p_{k+1}\rangle}{\langle p_k, p_k\rangle} = \gamma_{k+1}.$$

在求得首一多项式 $\{p_n(\lambda)\}_{n=0}^{\infty}$ 的三项递归关系系数 (α_k, γ_k) 后，对应的标准多项式 $\{\hat{p}_n(\lambda)\}_{n=0}^{\infty}$ 的三项递归关系系数变为 (α_k, γ_k^2)。 □

7.B.4 雅各比矩阵和克里斯托费尔——达布关系

如果存在一个标准正交多项式 $\{\hat{p}_n(\lambda)\}_{n0=}^{\infty}$，则也存在一个无限的对称三角对角矩阵 (tridiagonal matrix)，用 J_∞ 表示。由于它的所有 k 次对角元素均为正，因此 J_∞ 又称无限雅各比矩阵 (infinite Jacobi matrix)

$$J_\infty = \begin{pmatrix} \alpha & \sqrt{\beta_1} & & & \cdots \\ \sqrt{\beta_1} & \alpha_2 & \sqrt{\beta_2} & & \cdots \\ & \sqrt{\beta_2} & \alpha_3 & \sqrt{\beta_3} & \cdots \\ \vdots & \ddots & \ddots & \ddots & \cdots \end{pmatrix} \tag{7.36}$$

它的 k 次领先主子矩阵 (leading principle submatrix)称为 J_k。以 $k=3$ 为例

$$J_3 = \begin{pmatrix} \alpha & \sqrt{\beta_1} & \\ \sqrt{\beta_1} & \alpha_2 & \sqrt{\beta_2} \\ & \sqrt{\beta_2} & \alpha_3 \end{pmatrix}.$$

进而，所有的正交多项式都可用雅各比矩阵表现出来：

Theorem 7.13 (克里斯托费尔——达布关系)**.** *对于正交多项式 $p_k(\lambda)$, $k=0,1,2\ldots$，我们有如下克里斯托费尔——达布关系 (Christoffel-Darboux relation)*

$$\sum_{i=0}^{k} p_i(\lambda)p_i(\mu) = \begin{cases} \sqrt{\beta_{k+1}}\,\frac{p_{k+1}(\lambda)p_k(\mu)-p_k(\lambda)p_{k+1}(\mu)}{\lambda-\mu} & \lambda \neq \mu, \\ \sqrt{\beta_{k+1}}\left[\left(\frac{d}{d\lambda}p_{k+1}(\lambda)\right)p_k(\lambda) - \left(\frac{d}{d\lambda}p_k(\lambda)\right)p_{k+1}(\lambda)\right] & \lambda = \mu. \end{cases} \tag{7.37}$$

类似地，对于首一正交多项式，我们有

$$\sum_{i=0}^{k} (\gamma_k\gamma_{k-1}\ldots\gamma_{i+1})\,p_i(\lambda)p_i\mu = \frac{p_{k+1}(\lambda)p_k(\mu) - p_k(\lambda)p_{k+1}(\mu)}{\lambda-\mu}, \quad \lambda \neq \mu.$$

证明. 由三项递推关系定理(7.11)可得，正交多项式满足如下关系

$$
\begin{aligned}
&p_{n+1}(\lambda) - A_n\lambda p_n(\lambda) = B_n p_n(\lambda) + c_{n-1}(\lambda)p_{n-1}(\lambda),\ n = 1, 2, 3\ldots \\
&A_n = \frac{k_{n+1}}{k_n},\ C_n = -\frac{A_n}{A_{n-1}}\frac{h_n}{h_{n-1}},
\end{aligned}
$$

其中 k_n 为 n 次项对应的系数：将 $p_{n+1}(\lambda) - A_n\lambda p_n(\lambda)$ 看做一个多项式，$p_{n+1}(\lambda) - A_n\lambda p_n(\lambda) = \sum_{k=0}^{n} c_k p_k(\lambda)$。由此我们有

$$
\begin{aligned}
p_{n+1}(\lambda)p_n(\mu) &= (A_n\lambda + B_n)p_n(\lambda)p_n(\mu) + c_n p_{n-1}(\lambda)p_n(\mu), \\
p_{n+1}(\mu)p_n(\lambda) &= (A_n\mu + B_n)p_n(\mu)p_n(\lambda) + c_n p_{n-1}(\mu)p_n(\lambda).
\end{aligned}
$$

上两式相减得

$$
p_{n+1}(\lambda)p_n(\mu) - p_{n+1}(\mu)p_n(\lambda) = A_n(\lambda-\mu)p_n(\lambda)p_n(\mu) + C_n\left[p_{n-1}(\lambda)p_n(\mu) - p_{n-1}(\mu)p_n(\lambda)\right]
$$

↪ 裂项和 (telescoping sum)[18]

$$
(\lambda-\mu)\,p_n(\lambda)p_n(\mu) = \frac{p_{n+1}(\lambda)p_n(\mu) - p_{n+1}(\mu)p_n(\lambda)}{A_n} - \frac{C_n}{A_n}\left[p_{n-1}(\lambda)p_n(\mu) - p_{n-1}(\mu)p_n(\lambda)\right],
$$

↪

$$
(\lambda-\mu)\frac{p_n(\lambda)p_n(\mu)}{h_n} = \frac{p_{n+1}(\lambda)p_n(\mu) - p_{n+1}(\mu)p_n(\lambda)}{A_n h_n} - \frac{p_{n-1}(\lambda)p_n(\mu) - p_{n-1}(\mu)p_n(\lambda)}{A_{n-1}h_{n-1}},
$$

↪

$$
\begin{aligned}
(\lambda-\mu)\sum_{k=1}^{n}\frac{p_k(\lambda)p_k(\mu)}{h_k} &= \sum_{k=1}^{n}\frac{p_{k+1}(\lambda)p_k(\mu) - p_{k+1}(\mu)p_k(\lambda)}{A_k h_k} - \sum_{k=1}^{n}\frac{p_k(\lambda)p_{k-1}(\mu) - p_{k-1}(\lambda)p_k(\mu)}{A_{k-1}h_{k-1}} \\
&= \frac{p_{n+1}(\lambda)p_n(\mu) - p_{n+1}(\mu)p_n(\lambda)}{A_n h_n} - \frac{k_0^2(\lambda-\mu)}{h_0},
\end{aligned}
$$

↪

$$
\begin{aligned}
(\lambda-\mu)\sum_{k=0}^{n}\frac{p_k(\lambda)p_k(\mu)}{h_k} &= \frac{p_{n+1}(\lambda)p_n(\mu) - p_{n+1}(\mu)p_n(\lambda)}{A_n h_n} \\
&= \frac{1}{h_n}\frac{k_n}{k_{n+1}}\left[p_{n+1}(\lambda)p_n(\mu) - p_n(\lambda)p_{n+1}(\mu)\right], \quad \lambda \neq \mu.
\end{aligned}
$$

另，对于 $\mu \to \lambda$ 时的极限，上式变为

$$
\begin{aligned}
\lim_{\mu\to\lambda}\frac{p_{n+1}(\lambda)p_n(\mu) - p_{n+1}(\mu)p_n(\lambda)}{\lambda-\mu} &= \lim_{\mu\to\lambda}\frac{p_n(\lambda)\left[p_{n+1}(\lambda) - p_{n+1}(\mu)\right] - p_{n+1}(\lambda)\left[p_n(\lambda) - p_n(\mu)\right]}{\lambda-\mu} \\
&\approx \left[\frac{d}{d\lambda}p_{n+1}(\lambda)\right]p_n(\lambda) - \left[\frac{d}{d\lambda}p_n(\lambda)\right]p_{n+1}(\lambda).
\end{aligned}
$$

□

因此，对于矩阵形式的正交多项式 $p_k(\lambda) = (p_0(\lambda), p_1(\lambda), \ldots p_k(\lambda))^\top$ 来说，它可以写成如下的三项递推关系

$$
\lambda p_k = J_k p_k + \eta_k p_k(\lambda) e^{(k)},
$$

其中 $e^{(k)}$ 为单位矩阵 (identity matrix) $\eta_k = \sqrt{\beta_k}$ 的最后一列。$p_k(\lambda)$ 的根 θ_j^k 是 J_k 的特征值。

[18]裂项和是指一个求和方程中，后续各项相互抵消，只剩下初始项和最高项的情况，如

$$
S = \sum_{i=1}^{n-1}(a_i - a_{i+1}) = (a_1 - a_2) + (a_2 - a_3) + \ldots + (a_{n-1} - a_n) = a_1 - a_n.
$$

7.B.5 正交多项式的根

Theorem 7.14 (正交多项式根的特征)**.** 正交多项式序列 $\{p_n(\lambda)\}$ 中，第 n 次多项式 $p_n(\lambda)$ 在区间 (a,b) 中，正好有 n 项唯一且简单的实根。

证明. 根据定义，第 n 次多项式 $p_n(\lambda)$ 至多有 n 个实根。

假设 $p_n(\lambda)$ 有 $m \le n$ 个奇数实根 $\lambda_1, \lambda_2, \ldots \in \lambda_m(a,b)$，则 $p_n(\lambda)(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_m)$ 的符号不变。这意味着

$$\int_a^b w(\lambda)p_n(\lambda)(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_m)\,d\lambda \neq 0.$$

另一方面由正交条件得如果 $m < n$，则

$$\int_a^b w(\lambda)p_n(\lambda)(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_m)\,d\lambda = 0.$$

因此，我们得出 $m = n$，即 $p_n(x)$ 在空间 (a,b) 中有 n 个唯一的奇数次实根。 □

Theorem 7.15. 如果 $\{p_n(\lambda)\}_{n=0}^{\infty}$ 是一个在 (a,b) 区间中，关于权重方程 $w(\lambda)$ 的正交多项式序列，那么 $p_n(\lambda)$ 的根和 $p_{n+1}(\lambda)$ 的根彼此错落分布。

证明. 由范的定义可得，

$$h_n = \int_a^b w(\lambda)p_n(\lambda)\,d\lambda > 0, \quad n = 0,1,2\ldots$$

由克里斯托费尔——达布公式 (Theorem 7.13) 可得

$$h_n\frac{k_n}{k_{n+1}}\left\{p_n(\lambda)\left[\frac{d}{d\lambda}p_{n+1}(\lambda)\right] - p_{n+1}(\lambda)\left[\frac{d}{d\lambda}p_n(\lambda)\right]\right\} = \sum_{k=0}^{n}\frac{p_k(\lambda)^2}{h_k} > 0,$$

因此我们有

$$\frac{k_n}{k_{n+1}}\left\{p_n(\lambda)\left[\frac{d}{d\lambda}p_{n+1}(\lambda)\right] - p_{n+1}(\lambda)\left[\frac{d}{d\lambda}p_n(\lambda)\right]\right\} = \sum_{k=0}^{n}\frac{p_k(\lambda)^2}{h_k} > 0.$$

现在，假定 $p_n(\lambda)$ 中有两个连续的根 $\lambda_{n,k}$ 和 $\lambda_{n,k+1}$ 满足 $\lambda_{n,k} < \lambda_{n,k+1}$。由于 $p_n(\lambda)$ 中全部 n 个根都是简单的实数，因此 $p_n'(\lambda_{n,k})$ 和 $p_n'(\lambda_{n,k+1})$ 的符号应当相反。于是我们有

$$\begin{aligned}
&p_n(\lambda_{n,k}) = p_n(\lambda_{n,k+1}) = 0,\\
&p_n'(\lambda_{n,k})p_n'(\lambda_{n,k+1}) < 0,\\
&p_{n+1}'(\lambda_{n,k})p_{n+1}'(\lambda_{n,k+1}) < 0,
\end{aligned}$$

根据多项式序列的连续性特征，我们有：在 $\lambda_{n,k}$ 和 $\lambda_{n,k+1}$ 之间，至少应当有一个根落在 $(\lambda_{n,k}, \lambda_{n,k+1})$ 区间中。证毕。 □

进而，如果 $p_n(\lambda)$ 和 $p_{n+1}(\lambda)$ 中的根分别用 $\{\lambda_{n,k}\}_{k=1}^{n}$ 和 $\{\lambda_{n+1,k}\}_{k=1}^{n+1}$ 来表示，则我们有

$$a < \lambda_{n+1,1} < \lambda_{n,1} < \lambda_{n+1,2} < \lambda_{n,2} < \ldots < \lambda_{n+1,n} < \lambda_{n,n} < \lambda_{n+1,n+1} < b.$$

7.B.6　高斯求积

在前面的章节中，我们用 λ 符号来表示变量。随后我们改用 x，主要原因是打字起来方便。。。。

Theorem 7.16 (高斯求积公式)**.** *如果在区间 (a,b) 中有一个连续方程 $f(x)$，对应 $a<x_1<x_2<\ldots<x_n$ 的 n 个点。那么存在且只存在一个 ($\le n-1$ 次的) 多项式 p，使得 $p(x_i)=f(x_i)$，那么可用高斯积公式 (Gauss quadrature formula)来近似 $f(x)$ 的值，满足*

$$\int_a^b w(x)f(x)\,dx=\sum_{i=1}^n \lambda_{n,i}f(x_i),\quad \lambda_{n,i}:=\int_a^b \frac{w(x)p_n(x)}{(x-x_i)\,p_n'(x_i)}\,dx, i=1,2,\ldots,n. \tag{7.38}$$

证明. 可以使用拉格朗日插值法 (Lagrange interpolation)求解多项式 p。定义

$$p(x):=(x-x_1)(x-x_2)\ldots(x-x_n)$$

以及

$$\begin{aligned}P(x)&:=\sum_{i=1}^n f(x_i)\frac{p(x)}{(x-x_i)p'(x_i)}\\&=\sum_{i=1}^n f(x_i)\frac{(x-x_1)(x-x_2)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x_i-x_n)}{(x_i-x_1)(x_i-x_2)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}.\end{aligned}$$

定义一个区间 (a,b) 内关于权重方程 $w(x)$ 的正交多项式序列 $\{p_n(x)\}_{n=0}^\infty$。那么对于 $x_1<x_2<\ldots<x_n$，我们有 n 个多项式 $p_n(x)$ 的实根。如果 $f(x)$ 是一个 $\le 2n-1$ 次的多项式，那么 $f(x)-P(x)$ 就是一个至少包括这些 $x_1<x_2<\ldots<x_n$ 实根在内的 $\le 2n-1$ 次的多项式。因此，我们将 $f(x)$ 定义为

$$f(x)=P(x)+r(x)p_n(x),$$

其中 $r(x)$ 是一个 $\le n-1$ 次的多项式。引入 $P(x)$ 的定义式，将上式改写为

$$f(x)=\sum_{i=1}^n f(x_i)\frac{p(x)}{(x-x_i)p_n'(x_i)}+r(x)p_n(x).$$

这意味着

$$\int_a^b w(x)f(x)\,dx=\sum_{i=1}^n f(x_i)\int_a^b\frac{w(x)p_n(x)}{(x-x_i)p_n'(x_i)}\,dx+\int_a^b w(x)r(x)p_n(x)\,dx.$$

由于 $r(x)$ 是一个 $\le n-1$ 次的多项式，因此我们有 $\int_a^b w(x)r(x)p_n(x)\,dx$=0。调整后的高斯求积如(7.38)所示。在已知 $x_1<x_2<\ldots<x_n$ 所对应的 $p_n(x)$ 的 n 个实根，如果对应的 $f(x_i)$ 也是已知的，则高斯求积式给出了 $\le 2n-1$ 次多项式 $f(x)$ 的积分值。

如果 $f(x)$ 不是一个 $\le 2n-1$ 的多项式，高斯求积式转而求积分的近似：

$$\int_a^b w(x)f(x)\,dx\approx\sum_{i=1}^n \lambda_{n,i}f(x_i),\quad \lambda_{n,i}:=\int_a^b\frac{w(x)p_n(x)}{(x-x_i)\,p_n'(x_i)}\,dx,\quad i=1,2,\ldots,n.$$

□

我们将 $\{\lambda_{n,i}\}_{n=0}^\infty$ 称作克里斯托费尔数 (Christoffel numbers)，它不受到 $f(x)$ 方程的影响。

Theorem 7.17 (克里斯托费尔数的值全部为正)**.**

证明. 将 $\lambda_{n,i}$ 改写为

$$\lambda_{n,i} = \int_a^b w(x) l_{n,i}(x)\,dx, \quad l_{n,i} := \frac{p_n(x)}{(x-x_i)p_n'(x_i)}, \quad i = 1, 2, \ldots, n.$$

那么对应 $p_n(x)$ 的实根 $\{x_{n,k}\}_{k=1}^n$，$\leq 2n-2$ 次的多项式 $l_{n,i}^2 - l_{n,i}(x)$ 等于零。换句话说对于某个 $\leq n-2$ 次的多项式 q 来说，我们有

$$l_{n,i}^2 - l_{n,i}(x) = p_n(x)q(x).$$

根据正交条件，上式变为

$$\int_a^b w(x)\left(l_{n,i}^2 - l_{n,i}(x)\right)\,dx = \int_a^b w(x)p_n(x)q(x)\,dx = 0.$$

由此我们有

$$\lambda_{n,i} = \int_a^b w(x) l_{n,i}(x)\,dx = \int_a^b w(x) l_{n,i}(x)^2\,dx > 0.$$

□

7.C 常见的正交多项式类型

本节介绍几种常见的正交多项式 (表格7.C)，包括

- 埃米特多项式 (Hermite polynomials)
- 拉盖尔多项式 (Laguerre polynomials)
- 雅各比多项式 (Jacobi polynomials)
- 勒让德多项式 (Legendre polynomials)
- 车比雪夫多项式 (Chebyshev polynomials, 1st kind, 2nd kind)。

表 7.C.1: 经典正交多项式

名称	$p_n(x)$	$w(x)$	(a,b)
Hermite	$H_n(x)$	$\exp(-x^2)$ (Normal dist.)	$(-\infty, +\infty)$
Laguerre	$L_n^\alpha(x)$	$\exp(-\alpha)x^\alpha$ (Gamma dist.)	$[0, +\infty)$
Jacobi	$P_n^{\alpha,\beta}(x)$	$(1-x)^\alpha(1+x)^\beta$ (Beta dist.)	$(-1, 1)$
Legendre*	$P_n(x)$	1 (Uniform dist.)	$[-1, 1]$
Chebyshev**, 1st kind	$T_n(x)$	$\frac{1}{\sqrt{1-x^2}}$	$[-1, 1]$
Chebyshev, 2nd kind	$U_n(x)$	$\sqrt{1-x^2}$	$[-1, 1]$

* Legendre 是 Jacobi 的一种特殊形式，$\alpha = \beta = 0$。
* Chebyshev 是 Legendre 的一种特殊形式。

这些多项式往往满足如下条件 (图7.C.1)

- 罗德里格斯公式 (Rodrigues formula)
- 正交条件 (orthogonality)
- 三项递推关系 (three-term recurrence relation)
- 二阶线性微分方程 (2nd order linear differential equation)
- 母函数 (generating function)。

图 7.C.1: 正交多项式满足的关系

7.C.1　埃米特多项式

Theorem 7.18 (埃米特多项式的罗德里格斯公式)**.** *在区间 $(-\infty,+\infty)$ 内，关于正态分布 (normal distribution) $w(x)=\exp(-x^2)$ 的埃米特正交多项式 (Hermite polynomial) $H_n(x)$，可由罗德里格斯公式予以定义 (Rodrigues formula)*

$$H_n(x)=\frac{(-1)^n}{w(x)}D^n w(x)=(-1)^n\exp(x^2)D^n\exp(-x^2),\quad n=0,1,2\ldots \tag{7.39}$$

其中 $(-1)^n$ 项是为了保证 $\{D^n w(x)\}$ 的每一个首项系数都为正。$D=\frac{d}{dx}$ 是微分符，D^n 是第 n 次求导。D^n 遵循莱布尼兹法则 (Leibniz rule)

$$D^n\left[f(x)g(x)\right]=\sum_{k=0}^{n}=\sum_{k=0}^{n}\binom{n}{k}D^k f(x)D^{n-k}g(x),\quad n=0,1,2\ldots, \tag{7.40}$$

其中 $\binom{n}{k}=\frac{n!}{(n-k)!}$ 是帕斯卡三角 (Pascal triangle identity)中的二项式系数。帕斯卡三角满足关系

$$\binom{n+1}{k}=\binom{n}{k}+\binom{n}{k-1}.$$

帕斯卡三角的简单证明：

$$\begin{aligned}
&\binom{n}{k}+\binom{n}{k-1}\\
&=\frac{n!}{k!\,(n-k)!}+\frac{n!}{(k-1)!\,(n-k+1)!}\\
&=n!\left\{\frac{n-k+1}{k!(n-k+1)!}+\frac{k}{k!(n-k+1)!}\right\}\\
&=\frac{(n+1)!}{k!(n+1-k)!}\\
&=\binom{n+1}{k}.
\end{aligned}$$

正态分布 $w(x)=\exp(-x^2)$ 在区间 $(-3,3)$ 内，如图7.C.2所示。

Theorem 7.19. *由(7.39)定义的埃米特多项式 $H_n(x)$ 是一个关于 x 的 n 次多项式，并且 $H_0(1)=x$，$H_n(x)$ 的首项系数 $k_n=2^n$，$H_{2n}(x)$ 是偶方程，$H_{2n+1}(x)$ 是奇方程。*

图 7.C.2: 正态分布

$w(x) = \exp(-x^2)$ 在 $(-3, 3)$ 区间内的正态分布。

证明. (7.39)⇒

$$\begin{aligned} D^{n+1}w(x) &= D\left[D^n w(x)\right] \\ &= D\left[(-1)^n H_n(x) w(x)\right] \\ &= (-1)^n \left[w'(x)H_n(x) + w(x)H_n'(x)\right] \\ &= (-1)^n \left[-2x\exp(-x^2)H_n(x) + w(x)H_n(x)\right] \\ &= (-1)^{n+1} w(x)\left[2xH_n(x) - H_n'(x)\right], \quad n = 0, 1, 2\ldots \end{aligned} \tag{7.41}$$

由此，(7.39)⇒

$$H_{n+1}(x) = \begin{cases} 1 & n = 0 \\ \frac{(-1)^{n+1}}{w(x)} D^{n+1}w(x) = 2xH_n(x) - H_n'(x) & n = 1, 2\ldots \end{cases} \tag{7.42}$$

(7.42)决定了 $H_n(x)$ 是一个 n 次多项式，并且 $H_n(x)$ 的首项系数 $k_n = 2^n$，可以写出如下序列

$$\begin{aligned} p_0(x) &= 1, \\ p_1(x) &= 2x, \\ p_2(x) &= 4x^2 - 2, \\ p_3(x) &= 8x^3 - 12x, \\ p_4(x) &= 16x^4 - 48x^2 + 12, \\ p_5(x) &= 32x^5 - 160x^3 + 120x, \\ p_6(x) &= 64x^6 - 480x^4 + 720x^2 - 120, \\ p_7(x) &= 128x^7 - 1344x^5 + 3360x^3 - 1680x, \\ &\vdots \end{aligned}$$

由上式可得，对于偶数次 $2n = 2, 4, 6\ldots$ 的情况我们有 $p_{2n}(x) = p_{2n}(-x)$ 是偶方程；对于基数次 $2n+1 = 1, 3, 5\ldots$ 的情况我们有 $p_{2n+1}(x) = -p_{2n}(-x)$ 是奇方程，如图7.C.3。

□

Theorem 7.20 (埃米特多项式的正交条件)**.** *埃米特多项式 $H_n(x)$ 满足正交条件 (orthogonality condition)*

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp(-x^2) H_m(x) H_n(x)\, dx = 2^n n! \delta_{mn}, \quad m, n = 0, 1, 2\ldots \tag{7.43}$$

图 7.C.3: 埃米特多项式

上图和下图分别表示 $n=7, n=8$ 时，$H_n(x)$ 在 $(-2,2)$ 区间内的值。

证明. 由埃米特多项式的罗德里格斯定义式(7.39)我们有，当 $m<n$ 时

$$\begin{aligned}\int_{-\infty}^{\infty}\exp(-x^2)H_m(x)H_n(x)\,dx &= \int_{-\infty}^{\infty}\exp(-x^2)H_m(x)\left[(-1)^n\exp(x^2)D^n\exp(-x^2)\right]\,dx\\ &= \int_{-\infty}^{\infty}(-1)^n\exp(-x^2)H_m(x)\,dx,\end{aligned}$$

对上式做 n 次求导，积分内的值变为零。

当 $m=n$ 时 $\Rightarrow$

$$\begin{aligned}\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-x^2)H_m(x)H_n(x)\,dx &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp\left(-x^2\right)H_n(x)H_n(x)\,dx\\ &= \frac{1}{\sqrt{\pi}}(-1)^n\int_{-\infty}^{\infty}H_n(x)D^n\exp(-x^2)\,dx\\ &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}D^nH_n(x)\exp(-x^2)\,dx\\ &= \frac{k_n n!}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-x^2)\,dx\\ &= k_n n!.\end{aligned}$$

证毕。 □

Theorem 7.21 (埃米特多项式的三项递推关系). 埃米特多项式的三项递推关系为

$$H_{n+1}(x)=2xH_n(x)-2nH_{n-1}(x), n=1,2,3\ldots \tag{7.44}$$

证明. 由正态分布 $w(x)=\exp(-x^2)$ 我们有

$$w'(x)=-2xw(x).$$

代入莱布尼兹法则(7.40)⇒

$$D^{n+1}w(x)=D^n w'(x)=D^n\left[-2xw(x)\right]=(-2x)D^n w(x)+(-2n)D^{n-1}w(x).$$

将上式代回(7.42)⇒

$$\begin{aligned}H_{n+1}(x)&=\frac{(-1)^{n+1}}{w(x)}D^{n+1}w(x)\\&=(-1)\frac{(-1)^n}{w(x)}(-2x)D^n w(x)+(-1)^2\frac{(-1)^{n-1}}{w(x)}(-2n)D^{n-1}w(x)\\&=\frac{(-1)^n}{w(x)}w(x)D^n w(x)-2n\frac{(-1)^{n-1}}{w(x)}D^{n-1}w(x)\\&=2xH_n(x)-2nH_{n-1}(x),n=1,2,3\ldots\end{aligned}$$

□

Theorem 7.22 (埃米特多项式的二阶线性微分方程). *埃米特多项式的二阶线性微分方程 (second order linear differential equation)为*

$$H_n''(x)+2nH_n(x)-2xH_n'(x)=0. \tag{7.45}$$

证明. 联立埃米特多项式的罗德里格斯公式(7.39)和三项递归关系式(7.44)我们有

$$\begin{aligned}H_n'(x)&=2nH_{n-1}(x),n=1,2,3\ldots\\H_{n+1}'(x)&=2(n+1)H_n(x),n=0,1,2\ldots\end{aligned}$$

对(7.42)再做一次求导，并引入上式替换 $H_{n+1}'(x)$ 我们有

$$\begin{aligned}&H_{n+1}'(x)=2xH_n'(x)+2H_n(x)-H_n''(x),\\&\hookrightarrow 2(n+1)H_n(x)=2xH_n'(x)+2H_n(x)-H''(x),\end{aligned}$$

这意味着 $H_n(x)$ 构成一个二阶线性微分方程系统

$$y''(x)-2xy'(x)+2ny(x)=0. \tag{7.46}$$

□

Theorem 7.23 (埃米特多项式的母方程). *埃米特多项式的母方程 (generating function)为*

$$\exp(2xt-t^2)=\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}t^n. \tag{7.47}$$

证明. 设 $F(t)=\exp(-(x-t)^2)=\exp(-x^2)\exp(2xt-t^2)$. 围绕 $\tilde{t}=0$ 对 $F(t)$ 做泰勒级数展开近似

$$F(t)\approx\sum_{n}^{\infty}\frac{F^n(0)}{n!}t^n.$$

设 $u=x-t$，$\lim_{t\to 0}u\approx x$。则

$$\begin{aligned}F^n(0)&=\frac{d^n}{dt^n}\exp(-(x-t)^2)=\left[(-1)^n\frac{d^n}{du^n}\exp(-u^2)\right]_{u=x}\\&=(-1)^nD^n\exp(-x^2)=\exp(-x^2)H_n(x),n=0,1,2\ldots\end{aligned}$$

$\hookrightarrow$

$$\begin{aligned}&\exp(-x^2)\exp(-2xt-t^2)=\exp(-(x-t)^2)\\&=F(t)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}t^n=\exp(-x^2)\sum_{t=0}^{\infty}\frac{H_n(x)}{n!}t^n.\end{aligned}$$

□

7.C.2　拉盖尔多项式

Theorem 7.24 (拉盖尔多项式的罗德里格斯公式)**.** *在 $(0,\infty)$ 区间内，关于伽玛分布 (Gamma distribution) $w(x)=\exp(-x)x^\alpha$ 的拉盖尔多项式 (Laguerre polynomial) $L_n^{(\alpha)}(x)$，可由罗德里格斯公式予以定义*

$$L_n^{(\alpha)}(x)=\frac{1}{n!}\frac{1}{w(x)}D^n\left[w(x)x^n\right]=\frac{1}{n!}\exp(x)x^{-\alpha}D^n\left[\exp(-x)x^{n+\alpha}\right],\quad n=0,1,2\ldots \tag{7.48}$$

当 $\alpha=1/2$ 时伽玛分布 $w(x)=\exp(-x)x^\alpha$ 在区间 $(-3,3)$ 内，如图7.C.4所示。

图 7.C.4: 伽玛分布

$w(x)=\exp(-x)x^\alpha$ 在 $(-3,3)$ 区间内的正态分布 $(\alpha=1/2)$。

Theorem 7.25. *拉盖尔多项式 $L^{(n)}(x)$ 是一个关于 x 的 n 次多项式。$L^{(\alpha)}(0)=\frac{(\alpha+1)_n}{n!}$，$L^{(n)}(x)$ 的首项系数 $k_n=\frac{(-1)^n}{n!}$。*

证明. 根据莱布尼兹法则(7.40)我们有

$$\begin{aligned}D^n\left[\exp(-x)x^{n+\alpha}\right]&=\sum_{k=0}^{n}\binom{n}{k}D^k\exp(-x)D^{n-k}x^{n+\alpha}\\&=\sum_{k=0}^{n}\binom{n}{k}(-1)^k\exp(-x)x^{\alpha+k}(n+\alpha)(n+\alpha-1)\ldots(n+k+1)\\&=\exp(-x)x^\alpha\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{\Gamma(n+\alpha+1)}{\Gamma(k+\alpha+1)}x^k.\end{aligned}$$

最后一个等式的简单数学计算：

$$\frac{\Gamma(n+\alpha+1)}{\Gamma(k+\alpha+1)}=\frac{\int_0^\infty\exp(-t)t^{n+\alpha}\,dt}{\int_0^\infty\exp(-t)t^{k+\alpha}\,dt}\approx\frac{(n+\alpha)!}{(k+\alpha)!}$$

因此我们有

$$L_n^{(\alpha)}(x)=\sum_{k=0}^{n}(-1)^k\binom{n+\alpha}{n-k}\frac{x^k}{k!},\quad 其中\binom{n+\alpha}{n-k}=\frac{\Gamma(n+\alpha+1)}{\Gamma(k+\alpha+1)}(n-k)! \tag{7.49}$$

由上式可见，$L_n^{(\alpha)}(x)$ 是一个 n 次多项式。此外，由于

$$(-1)^k\binom{n+\alpha}{n-k}=\frac{(-1)^k}{(n-k)!}\frac{(n+\alpha)!}{(\alpha+k)!}=\frac{(-1)^k}{(n-k)!}\frac{(\alpha+1)_n}{(\alpha+1)_k}=\frac{(\alpha+1)_n}{n!}\frac{(-n)_k}{(\alpha+1)_k},$$

我们可得拉盖尔多项式的超几何方程

$$L_n^{(\alpha)}(x)=\frac{(\alpha+1)_n}{n!}\sum_{k=0}^{n}\frac{(-n)_k}{(\alpha+1)_k}\frac{x^k}{k!}=\binom{n+\alpha}{n}{}_1F_1\left(\begin{matrix}-n\\ \alpha+1\end{matrix};x\right). \tag{7.50}$$

由(7.50)得 $x=0$ 时

$$L_n^{(\alpha)}(0)=\binom{n+\alpha}{n}=\frac{(\alpha+1)_n}{n!},\quad n=0,1,2\ldots$$

$\hookrightarrow$

$$k_n=\frac{(-1)^n}{n!},\quad n=0,1,2\ldots$$

□

Theorem 7.26 (拉盖尔多项式的正交条件)**.** *拉盖尔多项式 $L_n^{(\alpha)}(x)$ 的正交条件满足*

$$\int_0^\infty \exp(-x)x^\alpha L_m^{(\alpha)}(x)L_n^{(\alpha)}(x)\,dx=\frac{\Gamma(n+\alpha+1)}{n!}\delta_{mn},\alpha>-1. \tag{7.51}$$

证明. 定义矩

$$\mu_n:=\int_0^n \exp(-x)x^{n+\alpha}\,dx,$$

部分还需要做一步的说明。

如果对于所有 $n=0,1,2\ldots$，μ_n 都存在，那么(7.51)的 LHS 积分是收敛的，因此需要 $\alpha>-1$。此外根据定义我们有

$$\mu_n=\Gamma(n+\alpha+1).$$

进而，由拉盖尔多项式的罗德里格斯公式(7.48)我们有

$$\begin{aligned}
&\int_0^\infty \exp(-x)x^\alpha L_m^{(\alpha)}(x)L_n^{(\alpha)}(x)\,dx\\
&=\int_0^\infty \exp(-x)x^\alpha L_m^{(\alpha)}(x)\left\{\frac{1}{n!}\exp(-x)x^\alpha D^n\left[\exp(-x)x^{n+\alpha}\right]\right\}dx\\
&=\frac{-1}{n!}\int_0^\infty L_m^{(\alpha)}(x)D^n\left[\exp(-x)x^{n+\alpha}\right]dx\\
&=\frac{(-1)^n}{n!}\int_0^\infty L_m^{(\alpha)}(x)D^n\left[\exp(-x)x^{n+\alpha}\right]dx
\end{aligned}$$

当 $m<n$ 时，对上式做 n 次求导，积分内的值变为零。当 $m=n$ 时 $\Rightarrow$

$$\begin{aligned}
&\frac{(-1)^n}{n!}\int_0^\infty L_m^{(\alpha)}(x)D^n\left[\exp(-x)x^{n+\alpha}\right]dx\\
&=\frac{(-1)^n}{n!}k_n n!\int_0^\infty \exp(-x)x^{n+\alpha}\,dx\\
&=\frac{\Gamma(n+\alpha+1)}{n!}.
\end{aligned}$$

□

Theorem 7.27 (拉盖尔多项式的母方程)**.** 拉盖尔多项式 $L_n^{(\alpha)}(x)$ 的母方程为

$$(1-t)^{-\alpha-1}\exp\left(-x\frac{t}{1-t}\right)=\sum_{n=0}^{\infty}L_n^{(\alpha)}(x)t^n. \tag{7.52}$$

证明. 将(7.50)代入(7.52)RHS $\Rightarrow$

$$\begin{aligned}
&\sum_{n=0}^{\infty}L_n^{(\alpha)}(x)t^n\\
&=\sum_{n=0}^{\infty}\frac{(\alpha+1)_n}{n!}t^n\sum_{k=0}^{n}\frac{(-n)_k}{(\alpha+1)_k}\frac{x^k}{k!}\\
&=\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\alpha+1)_n}{(\alpha+1)_k}\frac{(-1)^kx^kt^n}{k!(n-k)!}\\
&=\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\alpha+1)_{n+k}}{(\alpha+1)_k}\frac{(-1)^kx^kt^{n+k}}{k!n!}\\
&=\sum_{k=0}^{\infty}\frac{(-xt)^k}{k!}\sum_{n=0}^{\infty}\frac{(\alpha+k+1)_n}{n!}t^n\\
&=\sum_{k=0}^{\infty}\frac{(-xt)^k}{k!}(1-t)^{(-\alpha-k-1)}\\
&=(1-t)^{-\alpha-1}\sum_{k=0}^{\infty}\frac{1}{k!}\left(-x\frac{t}{1-t}\right)^k\\
&=(1-t)^{-\alpha-1}\exp\left(-x\frac{t}{1-t}\right).
\end{aligned}$$

□

Theorem 7.28. 由拉盖尔多项式的母方程(7.52)我们有

$$L_n^{(\alpha+\beta+1)}(x+y)=\sum_{k=0}^{n}L_k^{(\alpha)}(x)L_{n-k}^{(\beta)}(y). \tag{7.53}$$

证明. (7.52)$\Rightarrow$

$$\begin{aligned}
\sum_{n=0}^{\infty}L_n^{(\alpha+\beta+1)}(x+y)t^n&=(1-t)^{-\alpha-\beta-2}\exp\left[-(x+y)\frac{t}{1-t}\right]\\
&=\left[(1-t)^{-\alpha-1}\exp\left(-x\frac{t}{1-t}\right)\right]\left[(1-t)^{(-\beta-1)}\exp\left(-y\frac{1}{1-t}\right)\right]\\
&=\left[\sum_{k=0}^{\infty}L_k^{(\alpha)}(x)t^k\right]\left[\sum_{m=0}^{\infty}L_m^{(\beta)}(y)t^m\right]\\
&=\sum_{n=0}^{\infty}\left[\sum_{k=0}^{n}L_k^{(\alpha)}(x)L_{n-k}^{(\beta)}(y)\right]t^n.
\end{aligned}$$

证毕。 □

Theorem 7.29 (拉盖尔多项式的三项递推关系)**.** 拉盖尔多项式 $L_n^{(\alpha)}(x)$ 的三项递推关系为

$$(n+1)L_{n+1}^{(\alpha)}(x)+(x-2n-\alpha-1)L_n^{(\alpha)}(x)+(n+\alpha)L_{n-1}^{(\alpha)}(x)=0,\quad n=1,2,3\ldots \tag{7.54}$$

证明. 首先由(7.49)得

$$\begin{aligned}\frac{d}{dx}L_n^{(\alpha)}(x) &= \frac{d}{dx}\sum_{k=0}^{n}(-1)^k\binom{n+\alpha}{n-k}\frac{x^k}{k!}\\ &= \sum_{k=1}^{n}(-1)^k\binom{n+\alpha}{n-k}\frac{x^{k-1}}{(k-1)!}\\ &= \sum_{k=0}^{n-1}(-1)^{k-1}\binom{n+\alpha}{n-k-1}\frac{x^k}{(k)!}\\ &= -L_{n-1}^{(\alpha+1)}(x), n=1,2,3\ldots\end{aligned} \tag{7.55}$$

为简化表述，根据拉盖尔多项式的母方程(7.52)，定义

$$F(x,t) := (1-t)^{-\alpha-1}\exp\left(-x\frac{t}{1-t}\right) = \sum_{n=0}^{\infty}L_n^{(\alpha)}(x)t^n.$$

则我们有第一个偏导数

$$\begin{aligned}&\frac{\partial F(x,t)}{\partial x} = (-t)(1-t)^{-\alpha-2}\exp\left(-x\frac{t}{1-t}\right),\\ &\Rightarrow (1-t)\frac{\partial F(x,t)}{\partial x} + tF(x,t) = 0,\\ &\Rightarrow (1-t)\sum_{n=0}^{\infty}\frac{d}{dx}L_n^{(\alpha)}(x)t^n + t\sum_{n=0}^{\infty}L_n^{(\alpha)}(x)t^n = 0,\\ &\Rightarrow \underbrace{\sum_{n=0}^{\infty}\frac{d}{dx}L_n^{(\alpha)}(x)t^n}_{n\to n+1} - \sum_{n=0}^{\infty}\frac{d}{dx}L_n^{(\alpha)}(x)t^{n+1} + \sum_{n=0}^{\infty}L_n^{(\alpha)}(x)t^n = 0,\\ &\Rightarrow \frac{d}{dx}L_{n+1}^{(\alpha)}(x) - \frac{d}{dx}L_n^{(\alpha)}(x) + L_n^{(\alpha)}(x) = 0,\end{aligned} \tag{7.56}$$

上式代入(7.55)有

$$\begin{aligned}&-L_n^{(\alpha+1)}(x) - \frac{d}{dx}L_n^{(\alpha)}(x) + L_n^{(\alpha)}(x) = 0,\\ &\Rightarrow \frac{d}{dx}L_n^{(\alpha)}(x) = L_n^{(\alpha)}(x) - L_n^{(\alpha+1)}(x), \quad n=0,1,2\ldots\end{aligned}$$

第二个偏导数

$$\begin{aligned}&\frac{\partial F(x,t)}{\partial t} = \left[(\alpha+1)(1-t)^{-\alpha-2} + (1-t)^{-\alpha-1}\frac{-x(1-t)-xt}{(1-t)^2}\right]\exp\left(-x\frac{t}{1-t}\right)\\ &= \left(\alpha+1-\frac{x}{1-t}\right)(1-t)^{-\alpha-2}\exp\left(-x\frac{t}{1-t}\right),\\ &\Rightarrow (1-t)^2\frac{\partial F(x,t)}{\partial t} + [x-(\alpha-1)(1-t)]\,F(x,t) = 0,\\ &\Rightarrow (1-t)^2\sum_{n=1}nL_n^{(\alpha)}(x)t^{(n-1)} + [x-(\alpha+1)(1-t)]\sum_{n=0}^{\infty}L_n^{\alpha}(x)t^n = 0,\end{aligned}$$

将各项拆出

$$\begin{aligned}&\sum_{n=1}^{\infty}nL_n^{\alpha}(x)t^{n-1} - 2\sum_{n=1}^{\infty}nL_n^{\alpha}(x)t^n + \sum_{n=1}^{\infty}nL_n^{\alpha}(x)t^{n+1}\\ &+ x\sum_{n=0}^{\infty}L_n^{\alpha}(x)t^n + (\alpha+1)\sum_{n=0}^{\infty}L_n^{\alpha}(x)t^{n+1} - (\alpha+1)\sum_{n=0}^{\infty}L_n^{\alpha}(x)t^n = 0,\end{aligned}$$

按照 t 的幂次重新排列组合，得三项递推关系式(7.54)。 □

Theorem 7.30 (拉盖尔多项式的二阶线性微分方程). 拉盖尔多项式 $L_n^{(\alpha)}(x)$ 的二阶线性微分方程为

$$x\frac{d^2}{dx^2}L_n^{(\alpha)}(x)+(\alpha+1-x)\frac{d}{dx}L_n^{(\alpha)}(x)+nL_n^{(\alpha)}(x)=0.$$

证明. 拉盖尔多项式的三项递推关系(7.54)⇒

$$xL_n^{(\alpha)}(x)+(n+1)\left[L_{n+1}^{(\alpha)}(x)-L_n^{(\alpha)}(x)\right]-(n+\alpha)\left[\frac{d}{dx}L_n^{(\alpha)}(x)-\frac{d}{dx}L_{n-1}^{(\alpha)}(x)\right]=0.$$

对 x 求导 ⇒

$$L_n^{(\alpha)}(x)+x\frac{d}{dx}L_n^{(\alpha)}(x)+(n+1)\frac{d}{dx}L_{n+1}^{(\alpha)}(x)-(n+1)\frac{d}{dx}L_n^{(\alpha)}(x)-(n+\alpha)\frac{d}{dx}L_n^{(\alpha)}(x)+(n+\alpha)\frac{d}{dx}L_{n-1}^{(\alpha)}(x)=0,$$

引入(7.56)替换上式中的 $\frac{d}{dx}L_{n+1}^{(\alpha)}(x)$ 和 $\frac{d}{dx}L_n^{(\alpha)}(x)$

$$\begin{aligned}&L_n^{(\alpha)}(x)+x\frac{d}{dx}L_n^{(\alpha)}(x)-(n+1)L_n^{(\alpha)}(x)+(n+\alpha)L_{n-1}^{(\alpha)}(x)=0,\\&x\frac{d}{dx}L_n^{(\alpha)}(x)=nL_n^{(\alpha)}(x)-(n+\alpha)L_{n-1}^{(\alpha)}(x).\end{aligned}\tag{7.57}$$

上式继续对 x 求导 ⇒

$$\begin{aligned}\frac{d}{dx}L_n^{(\alpha)}(x)+x\frac{d^2}{dx^2}L_n^{(\alpha)}(x)&=n\frac{d}{dx}L_n^{(\alpha)}(x)-(n+\alpha)\frac{d}{dx}L_{n-1}^{(\alpha)}(x)\\&=(n+\alpha)\left(\frac{d}{dx}L_n^{(\alpha)}(x)-\frac{d}{dx}L_{n-1}^{(\alpha)}(x)\right)-\alpha\frac{d}{dx}L_n^{(\alpha)}(x)\\&=-(n+\alpha)L_{n-1}^{(\alpha)}(x)-\alpha\frac{d}{dx}L_n^{(\alpha)}(x)\\&=x\frac{d}{dx}L_n^{(\alpha)}(x)-nL_n^{(\alpha)}(x)-\alpha\frac{d}{dx}L_n^{(\alpha)}(x),\end{aligned}\tag{7.58}$$

↪

$$\begin{aligned}&\frac{d^2}{dx^2}L_n^{(\alpha)}(x)+(1+\alpha-x)\frac{d}{dx}L_n^{(\alpha)}(x)+nL_n^{(\alpha)}(x)=0,\\&\Rightarrow xy''(x)+(1+\alpha-x)y'(x)+ny(x)=0,\quad n=0,1,2\ldots\end{aligned}$$

□

7.C.3 雅各比多项式

Theorem 7.31 (雅各比多项式的罗德里格斯公式). 在 $(-1,1)$ 区间内，关于贝塔分布 *(Beta distribution)* $w(x)=(1-x)^\alpha(1+x)^\beta$ 的雅各比多项式 *(Jacobi polynomial)* $P_n^{(\alpha,\beta)}(x)$，可由罗德里格斯公式予以定义

$$\begin{aligned}P_n^{(\alpha,\beta)}(x)&=\frac{(-1)^n}{2^n n!}\frac{1}{w(x)}D^n\left[w(x)(1-x^2)^n\right]\\&=\frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta}D^n\left[(1-x)^{n+\alpha}(1+x)^{n+\beta}\right],\quad \alpha,\beta>-1,n=0,1,2\ldots\end{aligned}\tag{7.59}$$

当 $(\alpha,\beta)=(3,2)$ 时贝塔分布 $w(x)=(1-x)^\alpha(1+x)^\beta$ 在区间 $(-2,2)$ 内，如图7.C.5所示。

Theorem 7.32. 雅各比多项式 $P_n^{(\alpha,\beta)}(x)$ 是一个关于 x 的 n 次多项式。$P_n^{(\alpha,\beta)}(x)$ 是奇方程，并且 $P_n^{(\alpha,\beta)}(1)=\binom{n+\alpha}{n}$, $P_n^{(\alpha,\beta)}(-1)=(-1)^n\binom{n+\beta}{n}$, $\quad n=0,1,2\ldots$

图 7.C.5: 贝塔分布

$w(x) = (1-x)^{\alpha}(1+x)^{\beta}$ 在 $(-2,2)$ 区间内的正态分布 $(\alpha = 3, \beta = 2)$。

证明. 根据莱布尼兹法则(7.40)我们有

$$
\begin{aligned}
&D^n\left[(1-x)^{n+\alpha}(1+x)^{n+\beta}\right] \\
&= \sum_{k=0}^{n}\binom{n}{k} D^k(1-x)^{n+\alpha} D^{n-k}(1+x)^{n+\beta} \\
&= \sum_{k=0}^{n}\binom{n}{k}(-1)^k(n+\alpha)(n+\alpha-1)(n+\alpha-2)\ldots(n+\alpha-k+1)(1-x)^{n+\alpha-k} \\
&\qquad \times (n+\beta)(n+\beta-1)(n+\beta-2)\ldots(\beta+k+1)(1+x)^{\beta+k} \\
&= n!\sum_{k=0}^{n}(-1)^k\binom{n+\alpha}{k}\binom{n+\beta}{n-k}(1-x)^{n+\alpha-k}(1+x)^{\beta+k}, \quad n = 0,1,2\ldots
\end{aligned}
$$

代回(7.59)⇒

$$
\begin{aligned}
P_n^{(\alpha,\beta)}(x) &= \frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta} D^n\left[(1-x)^{n+\alpha}(1+x)^{n+\beta}\right] \\
&= \frac{(-1)^n}{2^n}\sum_{k=0}^{n}(-1)^k\binom{n+\alpha}{k}\binom{n+\beta}{n-k}(1-x)^{n-k}(1+x)^k, \quad n = 0,1,2\ldots
\end{aligned} \tag{7.60}
$$

这表明雅各比多项式 $P_n^{(\alpha,\beta)}(x)$ 是一个关于 x 的 n 次多项式。
(7.59) ⇒ $P_n^{(\alpha,\beta)}(x)$ 的对称性 (略):

$$
P_n^{(\alpha,\beta)}(-x) = (-1)^n P_n^{(\alpha,\beta)}(x), \quad n = 0,1,2\ldots \tag{7.61}
$$

(7.59) ⇒

$$
P_n^{(\alpha,\beta)}(1) = \frac{(-1)^n}{2^n}\sum_{k=0}^{n}(-1)^k\frac{(n+\alpha)!}{(n+\alpha-k)!\,k!}\frac{(n+\beta)!}{(n+\beta+k)!\,(n-k)!}(1+x)^k = \binom{n+\alpha}{n},
$$

$$
P_n^{(\alpha,\beta)}(-1) = (-1)^n\binom{n+\beta}{n}.
$$

□

Theorem 7.33 (雅各比多项式的超几何方程). 雅各比多项式的 $P_n^{(\alpha,\beta)}(x)$ 的超几何方程可表示为

$$
P_n^{(\alpha,\beta)}(x) = \binom{n+\alpha}{n}\,{}_2F_1\left(\begin{matrix}-n & n+\alpha+\beta+1 \\ & \alpha+1\end{matrix}; \frac{1-x}{2}\right), \quad n = 0,1,2\ldots \tag{7.62}
$$

证明. 对于 $x \neq 1$ 的情况，(7.60)⇒

$$P_n^{(\alpha,\beta)}(x) = \left(\frac{x-1}{2}\right)^n \sum_{k=0}^{\infty} \binom{n+\alpha}{n} \binom{n+\beta}{n-k} \underbrace{\left(\frac{x+1}{x-1}\right)^k}, \quad n = 0, 1, 2 \ldots$$

其中

$$\begin{aligned}
\left(\frac{x+1}{x-1}\right)^k &= \left(1 + \frac{2}{x-1}\right)^k, \quad k = 0, 1, 2 \ldots \\
&= \sum_{i=0}^{k} \binom{k}{i} \left(\frac{2}{x-1}\right)^i \\
&= \left(\frac{2}{x-1}\right)^n \sum_{i=0}^{n} \sum_{k=i}^{n} \binom{n+\alpha}{k} \binom{n+\beta}{n-k} \binom{k}{i} \left(\frac{2}{x-1}\right)^i \\
&= \left(\frac{2}{x-1}\right)^n \sum_{i=0}^{n} \sum_{k=0}^{n-i} \binom{n+\alpha}{i+k} \binom{n+\beta}{n-i-k} \binom{i+k}{k} \left(\frac{2}{x-1}\right)^i \\
&= \left(\frac{2}{x-1}\right)^n \sum_{i=0}^{n} \sum_{k=0}^{n} \binom{n+\alpha}{n-i+k} \binom{n+\beta}{i-k} \binom{n-i+k}{n-i} \left(\frac{2}{x-1}\right)^i \\
&= \sum_{i=0}^{n} \sum_{k=0}^{n} \binom{n+\alpha}{n-i+k} \binom{n+\beta}{i-k} \binom{n-i+k}{n-i} \left(\frac{2}{x-1}\right)^i \\
&= \sum_{i=0}^{n} \sum_{k=0}^{n} \frac{\Gamma(n+\alpha+1)}{(n-i+k)!\,\Gamma(i-k+\alpha+1)} \frac{\Gamma(n+\beta+1)}{(i-k)!\,\Gamma(n-i+k+\beta+1)} \frac{(n-i+k)!}{(n-i)!\,k!} \left(\frac{x-1}{2}\right)^i \\
&= \frac{\Gamma(n+\alpha+1)\,\Gamma(n+\beta+1)}{n!} \sum_{i=0}^{n} \frac{(-n)_i}{\Gamma(i+\alpha+1)\,\underbrace{\Gamma(n-i+\beta+1)}} \left(\frac{2}{x-1}\right)^i \\
&\quad \underbrace{\sum_{k=0}^{n} \frac{(-i)_k\,(-i-\alpha-1)_k}{(n-i+\beta+1)_k\,k!}},
\end{aligned} \tag{7.63}$$

其中，首先根据朱世杰——范德蒙德求和公式我们有

$$\sum_{k=0}^{n} \frac{(-i)_k\,(-i-\alpha-1)_k}{(n-i+\beta+1)_k\,k!} = {}_2F_1\left(\begin{matrix} -i & -i-\alpha-1 \\ n-i+\beta+1 \end{matrix}; 1\right) = \frac{(n+\alpha+\beta+1)_i}{(n-i+\beta+1)_i},$$

其次

$$\Gamma(n-i+\beta+1)(n-i+\beta+1)_i = \Gamma(n+\beta+1),$$

因此(7.63)进一步改写为

$$\begin{aligned}
P_n^{(\alpha,\beta)}(x) &= \frac{\Gamma(n+\alpha+1)}{n!} \sum_{i=0}^{n} \frac{(-n)_i\,(n+\alpha+\beta+1)_i}{\Gamma(i+\alpha+1)\,i!} \left(\frac{1-x}{2}\right)^i \\
&= \frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)\,n!} \sum_{i=0}^{n} \frac{(-n)_i\,(n+\alpha+\beta+1)_i}{(\alpha+1)_i i!} \left(\frac{1-x}{2}\right)^i \\
&= \binom{n+\alpha}{n} {}_2F_1\left(\begin{matrix} -n & n+\alpha+\beta+1 \\ & \alpha+1 \end{matrix}; \frac{1-x}{2}\right), \quad n = 0, 1, 2 \ldots
\end{aligned} \tag{7.64}$$

此外由奇函数的对称性质(7.61)我们有

$$P_n^{(\alpha,\beta)}(-x) = (-1)^n \binom{n+\beta}{n} {}_2F_1\left(\begin{matrix} -n & n+\alpha+\beta+1 \\ & \beta+1 \end{matrix}; \frac{1+x}{2}\right), \quad n = 0, 1, 2 \ldots$$

□

Theorem 7.34 (雅各比多项式的首项系数). 雅各比多项式 $P_n^{(\alpha,\beta)}(x)$ 的首项系数 k_n 为

$$k_n = \frac{n+\alpha+\beta+1}{2^n\, n!}, \quad n = 0, 1, 2\ldots \tag{7.65}$$

证明. 由雅各比多项式的超几何方程(7.64)得

$$k_n = \binom{n+\alpha}{n} \frac{(-n)_n\,(n+\alpha+\beta+1)_n}{(\alpha+1)_n\, n!} \frac{(-1)^n}{2^n} = \frac{n+\alpha+\beta+1}{2^n\, n!}, \quad n = 0, 1, 2\ldots$$

□

Theorem 7.35 (雅各比多项式的正交条件). 雅各比多项式 $P_n^{(\alpha,\beta)}(x)$ 满足如下正交关系

$$\begin{aligned}&\int_{-1}^{1}(1-x)^\alpha(1+x)^\beta P_m^{(\alpha,\beta)}(x)P_n^{(\alpha,\beta)}(x)\,dx = \frac{2^{\alpha+\beta+1}\,\Gamma(n+\alpha+1)\,\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\;\Gamma(n+\alpha+\beta+1)\,n!}\delta_{mn},\\ &\quad for\,\alpha > -1, \beta > -1, m, n \in \{0, 1, 2\ldots\}\end{aligned} \tag{7.66}$$

证明. 当 $m = n$ 时，由雅各比多项式的罗德里格斯公式(7.59)得

$$\begin{aligned}&\int_{-1}^{1}(1-x)^\alpha(1+x)^\beta P_m^{(\alpha,\beta)}(x)P_n^{(\alpha,\beta)}(x)\,dx\\ &= \int_{-1}^{1}(1-x)^\alpha(1+x)^\beta\left(P_n^{(\alpha,\beta)}(x)\right)^2 dx\\ &= \frac{(-1)^n}{2^n\, n!}\int_{-1}^{1}P_n^{(\alpha,\beta)}(x)D^n\left[(1-x)^{n+\alpha}(1+x)^{1+\beta}\right]dx\\ &= \frac{(-1)^n}{2^n\, n!}\int_{-1}^{1}D^n P_n^{(\alpha,\beta)}(x)(1-x)^{n+\alpha}(1+x)^{1+\beta}\,dx\\ &= \frac{(n+\alpha+\beta_1)_n}{2^n\, n!}\int_{-1}^{1}(1-x)^{n+\alpha}(1+x)^{n+\beta}\,dx\\ &= \frac{\Gamma(2n+\alpha+\beta+1)}{\Gamma(\alpha+\beta+n+1)\,2^{2n}\,n!}\underbrace{\int_{-1}^{1}(1-x)^{n+\alpha}(1+x)^{n+\beta}}\,dx, \quad n = 0, 1, 2\ldots\end{aligned}$$

设 $2t := 1 - x$ 我们有

$$\begin{aligned}&\int_{-1}^{1}(1-x)^{n+\alpha}(1+x)^{n+\beta}\\ &= \int_{0}^{1}(2t)^{n+\alpha}(2t)^{n+\beta}\,dx\\ &= 2^{(2n+\alpha+\beta+1)}\int_{0}^{1}t^{n+\alpha}(1-t)^{n+\beta}\,dt\\ &= 2^{(2n+\alpha+\beta+1)}\underbrace{B(n+\alpha+1, n+\beta+1)}\\ &= 2^{(2n+\alpha+\beta+1)}\frac{\Gamma(n+\alpha+1)\,\Gamma(n+\beta+1)}{\Gamma(2n+\alpha+\beta+2)}\\ &= 2^{(2n+\alpha+\beta+1)}\frac{\Gamma(n+\alpha+1)\,\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\,\Gamma(2n+\alpha+\beta+1)},\end{aligned}$$

代回上式我们有

$$
\begin{aligned}
&\int_{-1}^{1}(1-x)^{\alpha}(1+x)^{\beta}P_m^{(\alpha,\beta)}(x)P_n^{(\alpha,\beta)}(x)\,dx\\
&=\frac{\Gamma(2n+\alpha+\beta+1)}{\Gamma(\alpha+\beta+n+1)\,2^{2n}\,n!}\int_{-1}^{1}(1-x)^{n+\alpha}(1+x)^{n+\beta}\,dx,\quad n=0,1,2\ldots\\
&=\frac{\Gamma(2n+\alpha+\beta+1)}{\Gamma(\alpha+\beta+n+1)\,2^{2n}\,n!}\left[2^{(2n+\alpha+\beta+1)}\frac{\Gamma(n+\alpha+1)\,\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\,\Gamma(2n+\alpha+\beta+1)}\right]\\
&=\frac{2^{\alpha+\beta+1}\,\Gamma(n+\alpha+1)\,\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\;\Gamma(n+\alpha+\beta+1)\,n!}.
\end{aligned}
$$

对于 $m<n$ 的情况 (略)。 □

Theorem 7.36 (雅各比多项式的二阶线性微分方程)**.** 雅各比多项式 $P_n^{(\alpha,\beta)}(x)$ 的二阶线性微分方程形式为

$$
\begin{aligned}
&(1-x)^2\frac{d^2}{dx^2}P_n^{(\alpha,\beta)}(x)+[\beta-\alpha-(\alpha+\beta+2)x]\,\frac{d}{dx}P_n^{(\alpha,\beta)}(x)+n(n+\alpha+\beta+1)P_n^{(\alpha,\beta)}(x)=0,\\
&\hookrightarrow(1-x)^2y''(x)+[\beta-\alpha-(\alpha+\beta+2)x]\,y'(x)+n(n+\alpha+\beta+1)y(x)=0.
\end{aligned}
\tag{7.67}
$$

证明. 略。提示：由雅各比多项式 $P_n^{(\alpha,\beta)}(x)$ 的超几何方程(7.62)我们有

$$
\begin{aligned}
&\frac{d}{dx}P_n^{(\alpha,\beta)}(x)=\binom{n+\alpha}{n}\frac{(-n)\,(n+\alpha+\beta+1)}{(\alpha+1)}\left(-\frac{1}{2}\right){}_2F_1\left(\begin{matrix}-n+1\;\; n+\alpha+\beta+1\\ \beta+1\end{matrix};\frac{1-x}{2}\right)\\
&=\frac{n+\alpha+\beta+1}{2}\binom{n+\alpha}{n-1}{}_2F_1\left(\begin{matrix}-n+1\;\; n+\alpha+\beta+1\\ \beta+1\end{matrix};\frac{1-x}{2}\right)\\
&=\frac{n+\alpha+\beta+1}{2}P_{n-1}^{(\alpha+1,\beta+1)}(x),\quad n=1,2,3\ldots
\end{aligned}
$$

□

Theorem 7.37 (雅各比多项式的母方程)**.** 雅各比多项式 $P_n^{(\alpha,\beta)}(x)$ 的母方程为

$$
\frac{2^{\alpha+\beta}}{R\,(1+R-t)^{\alpha}\,(1+R+t)^{\beta}}=\sum_{n=0}^{\infty}P_n^{(\alpha,\beta)}(x)\,t^n,\quad R:=\sqrt{1-2x+t^2}. \tag{7.68}
$$

Theorem 7.38 (雅各比多项式的三项递推关系)**.** 雅各比多项式 $P_n^{\alpha,\beta}(x)$ 的三项递推关系可表示为

$$
\begin{aligned}
&\mathcal{A}\,P_{n+1}^{\alpha,\beta}(x)=\mathcal{B}\,P_n^{\alpha,\beta}(x)+\mathcal{C}\,P_{n-1}^{\alpha,\beta}(x),\\
&\quad\mathcal{A}=2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta),\\
&\quad\mathcal{B}=(2n+\alpha+\beta+1)(\alpha^2-\beta^2)+(2n+\alpha+\beta)(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)x,\\
&\quad\mathcal{C}=-2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2).
\end{aligned}
\tag{7.69}
$$

证明. 略。可参考 (Shen et al., 2011, p.74)。 □

7.C.4 勒让德多项式

Theorem 7.39 (勒让德多项式的罗德里格斯公式)**.** 在 $(-1,1)$ 区间内，关于 $w(x)=1$ 的均匀分布 *(uniform distribution)* $w(x)=1$ 的勒让德多项式 $P_n(x)$，可由罗德里格斯公式予以定义

$$
P_n(x)=\frac{(-1)^n}{2^n\,n!}\frac{1}{w(x)}D^n\left[w(x)(1-x^2)^n\right]=\frac{(-1)^n}{2^n\,n!}D^n\left[(1-x^2)^n\right],\quad n=0,1,2\ldots \tag{7.70}
$$

Theorem 7.40 (勒让德多项式是雅各比多项式的特例; 勒让德多项式的超几何方程)**.** 勒让德多项式(7.70)是雅各比多项式(7.59)的特例 $\alpha=\beta=0$:

$$P_n(x)=P_n^{(\alpha=0,\beta=0)}(x)={}_2F_1\left(\begin{matrix}-n\ \ n+1\\ \alpha+1\end{matrix};\frac{1-x}{2}\right),\quad n=0,1,2\ldots \tag{7.71}$$

证明. 将 $\alpha=0$, $\beta=0$ 代入(7.62)可得。 □

Theorem 7.41. 勒让德多项式 $P_n(x)$ 是一个关于 x 的 n 次多项式。$P_n(x)$ 是奇方程，并且 $P_n(1)=1$, $P_n(-1)=(-1)^n$。

证明. 由 Theorem 7.40可得 $P_n(x)=P_n^{(\alpha=0,\beta=0)}(x)$。因此根据雅各比多项式的相关性质 (Theorem 7.19)，可证。 □

Theorem 7.42 (勒让德多项式的首项系数)**.** 勒让德多项式 $P_n(x)$ 的首项系数 k_n 为

$$k_n=\frac{(2n)!}{2^n\,(n!)^2}. \tag{7.72}$$

证明. 由勒让德多项式的超几何方程(7.71)得

$$k_n=\frac{(-n)_n\,(n+1)_n}{n!\,(-1)^n}\frac{(-1)^n}{2^n}=\frac{(2n)!}{2^n\,(n!)^2}.$$

□

Theorem 7.43 (勒让德多项式的正交条件)**.** 勒让德多项式 $P_n(x)$ 满足如下正交关系

$$\int_{-1}^{1}P_m(x)P_n(x)\,dx=\frac{2}{2n+1}\delta_{mn},\quad m,n\in\{0,1,2\ldots\} \tag{7.73}$$

证明. 由勒让德多项式的罗德里格斯公式(7.70)可得

$$\begin{aligned}\int_{-1}^{1}P_m(x)P_n(x)\,dx&=\frac{(-1)^n}{2^n\,n!}\int_{-1}^{1}P_m(x)D^n\left[(1-x^2)^n\right]\,dx\\&=\frac{(-1)^n}{2^n\,n!}\int_{-1}^{1}D^n\left[P_m(x)(1-x^2)^n\right]\,dx\end{aligned}$$

当 $m<0$ 时，$\int_{-1}^{1}P_m(x)P_n(x)\,dx=0$。当 $m=n$ 时，

$$\begin{aligned}&\int_{-1}^{1}P_m(x)P_n(x)\,dx\\&\qquad=\int_{-1}^{1}D^n\left[P_n(x)(1-x^2)^n\right]\,dx\\&\qquad=k_n n!\int_{-1}^{1}(1-x^2)^n dx\\&\qquad=\frac{(2n)!}{2^n(n!)^2}\int_{-1}^{1}(1-x^2)^n dx.\end{aligned}$$

定义 $1-x:=2t, n=0,1,2\ldots \Rightarrow$

$$\begin{aligned}\int_{-1}^{1}(1-x^2)^n\,dx &= \int_{-1}^{1}(1-x)^n(1+x)^n dx\\ &= \int_{-1}^{1}(2t)^n(2-2t)^n 2\,dx\\ &= 2^{2n+1}B(n+1,n+1)\\ &= 2^{2n+1}\frac{\Gamma(n+1)\,\Gamma(n+1)}{\Gamma(2n+2)}\\ &= \frac{2^{2n+1}\,(n!)^2}{(2n+1)!},\end{aligned}$$

$\hookrightarrow$

$$\int_{-1}^{1}[P_n(x)]^2\,dx = \frac{(2n)!}{2^{2n}\,(n!)^2}\frac{2^{2n+1}\,(n!)^2}{(2n+1)!} = \frac{2}{2n+1},\quad n=0,1,2\ldots$$

□

Theorem 7.44 (勒让德多项式的母方程). 勒让德多项式 $P_n(x)$ 的母方程为

$$\sum_{n=0}^{\infty}P_n(x)\,t^n = \left(1-2xt+t^2\right)^{-\frac{1}{2}}. \tag{7.74}$$

证明. 由勒让德多项式的超几何方程(7.71)得

$$\begin{aligned}\sum_{n=0}^{\infty}P_n(x)\,t^n &= \sum_{n=0}^{\infty}{}_2F_1\left(\begin{matrix}-n & n+1\\ & 1\end{matrix};\frac{1-x}{2}\right)t^n\\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{(-n)_k\,(n+1)_k}{(1)_n\,k!}\left(\frac{1-x}{2}\right)^k t^n\\ &= \sum_{n=0}^{\infty}\sum_{k=n}^{\infty}\frac{(-n)_k\,(n+1)_k}{k!\,k!}\left(\frac{1-x}{2}\right)^k t^n\\ &= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{(-n-k)_k\,(n+k+1)_k}{k!\,k!}\left(\frac{1-x}{2}\right)^k t^{n+k}\end{aligned}$$

□

Theorem 7.45 (勒让德多项式的三项递推关系). 勒让德多项式 $P_n(x)$ 的三项递推关系为

$$(n+1)P_{n+1}(x)-x(2n+1)P_n(x)+nP_{n-1}(x)=0,\quad n=1,2,3\ldots \tag{7.75}$$

证明. 定义 $F(x,t):=\left(1-2xt+t^2\right)^{-\frac{1}{2}}$。由勒让德多项式 $P_n(x)$ 的母方程(7.74)我们有

$$\frac{\partial}{\partial t}F(x,t) = -\frac{1}{2}\left(1-2xt+t^2\right)^{-\frac{3}{2}}(-2x+2t) = \frac{x-t}{(1-2xt+t^2)^{\frac{3}{2}}} = \sum_{n=1}^{\infty}nP_n(x)t^{n-1},$$

进而

$$\begin{aligned}&\left(1-2xt+t^2\right)\frac{\partial}{\partial t}F(x,t) = (x-t)F(x,t),\\ &\hookrightarrow \left(1-2xt+t^2\right)\sum_{n=1}^{\infty}nP_n(x)t^{n-1} = (x-t)\left(1-2xt+t^2\right)^{-\frac{1}{2}}.\end{aligned}$$

拆分上式

$$\underbrace{\sum_{n=1}^{\infty} nP_n(x)t^{n-1}}_{:=A} - \underbrace{2x\sum_{n=1}^{\infty} nP_n(x)t^{n}}_{:=B} + \underbrace{\sum_{n=1}^{\infty} nP_n(x)t^{n+1}}_{:=C} = \underbrace{x\sum_{n=0}^{\infty} P_n(x)t^{n}}_{:=D} - \underbrace{\sum_{n=0}^{\infty} nP_n(x)t^{n+1}}_{:=E},$$

按照 t 的幂次重新整理

$$\begin{aligned}
A &= \sum_{n=1}^{\infty} nP_n(x)t^{n-1}, \\
B + D &= -2x\sum_{n=1}^{\infty} nP_n(x)t^{n} - x\sum_{n=0}^{\infty} P_n(x)t^{n} = -x\sum_{n=0}^{\infty}(2n+1)P_n(x)t^{n}, \\
C + E &= \sum_{n=0}^{\infty}(n+1)P_n(x)t^{n+1},
\end{aligned}$$

$\hookrightarrow$

$$\sum_{n=1}^{\infty} nP_n(x)t^{n-1} - x\sum_{n=0}^{\infty}(2n+1)P_n(x)t^{n} + \sum_{n=0}^{\infty}(n+1)P_n(x)t^{n+1} = 0,$$

再次整理可得(7.75)。 □

7.C.5 切比雪夫多项式

更多数学上的证明，可参考Boyd (2001); Fornberg (1996); Mason and Handscomb (2003); Shen et al. (2011)。

在 $[-1,1]$ 区间内，关于 $w(x) = \left(1-x^2\right)^{-\frac{1}{2}}$ 的第一类切比雪夫多项式 (the first kind Chebyshev polynomial) $T_n(x)$ 定义为

$$T_n(x) = \cos(n\theta), \quad x = \cos\theta, \quad n = 0, 1, 2\ldots \tag{7.76}$$

在 $[-1,1]$ 区间内，关于 $w(x) = \left(1-x^2\right)^{\frac{1}{2}}$ 的第二类切比雪夫多项式 (the second kind Chebyshev polynomial) $U_n(x)$ 定义为

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \quad x = \cos\theta, \quad n = 0, 1, 2\ldots \tag{7.77}$$

Theorem 7.46. 第一类切比雪夫多项式 $T_n(x)$ 的首项系数为

$$k_n = 2^{n-1}. \tag{7.78}$$

证明. $T_0(x) = 1, T_1(x) = x$ 代入三项递推关系(7.83)有

$$\begin{aligned}
T_0(x) &= 1, \\
T_1(x) &= x, \\
T_2(x) &= 2x^2 - 1, \\
T_3(x) &= 4x^3 - 3x, \\
T_4(x) &= 8x^4 - 8x^2 + 1, \\
T_5(x) &= 16x^5 - 20x^3 + 5x, \\
&\vdots
\end{aligned}$$

可见 $k_n = 2^{n-1}$。 □

Theorem 7.47 (切比雪夫多项式的罗德里格斯公式). 第一类切比雪夫多项式 $T_n(x)$ 的罗德里格斯公式 定义为

$$T_n(x) = \frac{(-1)^n 2^n n!}{(2n)!} \left(1-x^2\right)^{-\frac{1}{2}} D^n \left(1-x^2\right)^{\frac{n-1}{2}} \tag{7.79}$$

第二类切比雪夫多项式 $U_n(x)$ 的罗德里格斯公式

$$U_n(x) = \frac{(-1)^n (n_1)! 2^n}{(2n+1)!} \left(1-x^2\right)^{\frac{1}{2}} D^n \left(1-x^2\right)^{\frac{n+1}{2}} \tag{7.80}$$

Theorem 7.48 (切比雪夫多项式的正交条件). 第一类、第二类切比雪夫多项式 $T_n(x), U_n(x)$ 的正交条件分别为 *(Gradshteyn and Ryzhik, 2014, Entry 7.343.1, pp.807-808)*[19]

$$\int_{-1}^{1} \left(1-x^2\right)^{-\frac{1}{2}} T_m(x) T_n(x)\, dx = \int_0^{\pi} \cos(m\theta)\cos(n\theta)\, d\theta = \begin{cases} 0 & m \neq n, \\ \frac{\pi}{2} & m = n = 0, \\ \pi & m = n = 0. \end{cases} \tag{7.81}$$

$$\int_{-1}^{1} \left(1-x^2\right)^{-\frac{1}{2}} U_m(x) U_n(x)\, dx = \int_0^{\pi} \sin(m+1)\theta \sin(n+1)\theta\, d\theta = \begin{cases} 0 & m \neq n, \\ \frac{\pi}{2} & m = n = 0, \\ \pi & m = n = 0. \end{cases} \tag{7.82}$$

证明. 略。 □

Theorem 7.49 (切比雪夫多项式的三项递推关系). 第一类、第二类切比雪夫多项式 $T_n(x), U_n(x)$ 的三项递推关系分别为

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, 3 \ldots \tag{7.83}$$

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n = 1, 2, 3 \ldots \tag{7.84}$$

证明. (7.76)⇒

$$T_{n+1}(x) + T_{n-1}(x) = \cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos(n\theta) = 2xT_n(x).$$

(7.77)⇒

$$U_{n+1}(x) + U_{n-1}(x) = \frac{\sin(n+2)\theta}{\sin\theta}\frac{\sin n\theta}{\sin\theta} = 2\frac{\cos\theta\sin(n+1)\theta}{\sin\theta} = 2xU_n(x).$$

□

Theorem 7.50 (第一类、第二类切比雪夫多项式的关系). 第一类、第二类切比雪夫多项式的关系为

$$\begin{cases} T_0(x) = U_0(x) = 1, \\ T_1(x) = x, \quad U_1(x) = 2x, \\ T_n(x) = U_n(x) - xU_{n-1}(x), \quad n = 1, 2, 3 \ldots \end{cases} \tag{7.85}$$

证明. $n = 0$ 时，可由定义式求得。$n \leq 1$ 时

$$U_n(x) - xU_{n-1}(x) = \frac{\sin(n+1)\theta}{\sin\theta} - \frac{\cos\theta\sin n\theta}{\sin\theta} = \frac{\sin\theta\cos n\theta}{\sin\theta} = \cos n\theta = T_n(x).$$

□

[19] Gradshteyn and Ryzhik (2014) 的补充材料可参考如Moll (2015, 2016)。

Theorem 7.51 (第一类切比雪夫多项式的母方程). 第一类切比雪夫多项式 $T_n(x)$ 的母方程为

$$\sum_{n=0}^{\infty} T_n(x)t^n = \frac{1-x}{1-2xt+t^2}, \quad |t|<1. \tag{7.86}$$

证明. 将第一类切比雪夫多项式的三项递推关系(7.83)两侧分别乘以 t^{n+1}，并沿着 $n=1,2,3\ldots$ 求和

$$2x\sum_{n=1}^{\infty} T_n(x)t^{n+1} - \sum_{n=1}^{\infty} T_{n-1}(x)t^{n+1} = \sum_{n=1}^{\infty} T_{n+1}(x)t^{n+1}.$$

定义 $F(x,t) := \sum_{n=0}^{\infty} T_n(x)t^n, \quad |t|<1$，则上式变为

$$\begin{aligned}
&[F(x,t)-T_1(x)-T_0(x)] = 2xt\,[F(x,t)-T_0(x)] - t^2F(x,t),\\
&\hookrightarrow (1-2xt+t^2)F(x,t) = T_0(x)+T_1(x)t-2xtT_0(x) = 1-xt,\\
&\hookrightarrow F(x,t) = \sum_{n=0}^{\infty} T_n(x)t^n = \frac{1-x}{1-2xt+t^2}, \quad |t|<1.
\end{aligned}$$

□

Theorem 7.52 (第二类切比雪夫多项式的母方程). 第一类切比雪夫多项式 $U_n(x)$ 的母方程为

$$\sum_{n=0}^{\infty} U_n(x)t^n = \frac{1}{1-2xt+t^2}, \quad |t|<1. \tag{7.87}$$

将第二类切比雪夫多项式的三项递推关系(7.84)两侧分别乘以 t^{n+1}，并沿着 $n=1,2,3\ldots$ 求和

$$2x\sum_{n=1}^{\infty} U_n(x)t^{n+1} - \sum_{n=1}^{\infty} U_{n-1}(x)t^{n+1} = \sum_{n=1}^{\infty} U_{n+1}(x)t^{n+1}.$$

定义 $G(x,t) := \sum_{n=0}^{\infty} U_n(x)t^n, \quad |t|<1$，则上式变为

$$\begin{aligned}
&[G(x,t)-U_1(x)-U_0(x)] = 2xt\,[G(x,t)-U_0(x)] - t^2F(x,t),\\
&\hookrightarrow (1-2xt+t^2)G(x,t) = U_0(x)+U_1(x)t-2xtU_0(x) = 1,\\
&\hookrightarrow G(x,t) = \sum_{n=0}^{\infty} U_n(x)t^n = \frac{1}{1-2xt+t^2}, \quad |t|<1.
\end{aligned}$$

Theorem 7.53 (第一类和第二类切比雪夫多项式的转换). 第一类切比雪夫多项式 $T_n(x)$ 和第二类切比雪夫多项式 $U_n(x)$ 的转换，满足

$$U_n(x) = \sum_{k=0}^{n} T_k(x)x^{n-k}, \quad |t|<1, \quad n=0,1,2\ldots \tag{7.88}$$

证明. 对于 $|t|<1$ 我们有

$$\begin{aligned}
&\sum_{n=0}^{\infty}\left[\sum_{k=0}^{\infty}T_k(x)x^{n-k}\right]t^n\\
&=\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}T_k(x)x^{n-k}t^n\\
&=\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}T_k(x)x^{n}t^{n+k}\\
&=\sum_{k=0}^{\infty}T_k(x)t^k\sum_{n=0}^{\infty}(xt)^n\\
&=\frac{1-xt}{1-2xt+t^2}\frac{1}{1-xt}\\
&=\frac{1}{1-2xt+t^2}\\
&=\sum_{n=0}^{\infty}U_n(x)t^n.
\end{aligned}$$

去掉等式两侧的求和符号，证毕。 □

Theorem 7.54 (勒让德多项式和第二类切比雪夫多项式的转换)**.** *勒让德多项式 $P_n(x)$ 和第二类切比雪夫多项式 $U_n(x)$ 的转换，满足*

$$U_n(x)=\sum_{k=0}^{\infty}P_k(x)P_{n-k}(x),\quad |t|<1,\quad n=0,1,2\ldots \tag{7.89}$$

证明. 对于 $|t|<1$ 我们有

$$\begin{aligned}
&\sum_{n=0}^{\infty}\left[\sum_{k=0}^{\infty}P_k(x)P_{n-k}(x)\right]t^n\\
&=\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}P_k(x)P_{n-k}(x)t^n\\
&=\sum_{k=0}^{\infty}P_k(x)^t t^k\sum_{n=0}^{\infty}P_n(x)t^n\\
&=\left(1-2xt-t^2\right)^{-\frac{1}{2}}\left(1-2xt-t^2\right)^{-\frac{1}{2}}\\
&=\sum_{n=0}^{\infty}U_n(x)t^n.
\end{aligned}$$

去掉等式两侧的求和符号，证毕。 □

Theorem 7.55 (第一类切比雪夫多项式与次的关系)**.** *1 个 n 次第一类切比雪夫多项式 $T_n(x)$ 满足关系*

$$\begin{aligned}
&\int_{-1}^{1}T_n(x)\,dx=\frac{(-1)^{n-1}-1}{(n-1)(n+1)},\quad n\geq 2,\\
&\int_{-\infty}^{\infty}T_n(x)\,dx=-\int_{-\infty}^{\infty}\cos(n\theta)\sin\theta\,d\theta.
\end{aligned}\tag{7.90}$$

证明. 略。 □

Theorem 7.56 (切比雪夫插值定理)**.** *切比雪夫插值 (Chebyshev interpolation)。*

证明. 假定在区间 $[a,b]$ 内有 1 个方程 $f(x)$。则关于结点 (nodes) $\{x_0,x_1,\ldots,x_n\}\in[a,b], x_i\neq x_j, i\neq j$ 的拉格朗日插值 (Lagrange interpolation)可以定义为有且只有一个 $\leq n$ 次的多项式 $P_n(x)$，满足 $P(x_i)=f(x_i), i=0,1,\ldots,n$。

如果在区间 $[a,b]$ 内，$f^{(n)}$ 连续并且 $f^{(n+1)}$ 存在，那么拉格朗日插值的误差 (interpolation error) 可以表示为

$$f(x)-p(x)=\frac{f^{(n+1)}(\zeta)}{(n+1)!}q(x),\quad \zeta\in[a,b], \tag{7.91}$$

其中 $q(x)$ 定义为

$$q(x)=\Pi_{i=0}^{n}(x-x_i)=(x-x_0)(x-x_1)\ldots(x-x_n).$$

我们的研究目标是，选取合适的多项式 $P(x)$ 作为原方程 $f(x)$ 的近似，近似的判定标准是，误差(7.91)越小，近似越精确，或者换句话说，$P_n(x)$ 收敛到 $f(x)$。这可以分为两个问题来描述。

第一个问题，通常来说，随着 $n\to\infty$，$f^{(n+1)}$ 可能较大，我们很难保证(7.91)中的插值误差足够小。

第二个问题，关于 $n+1$ 个结点 $\{x_0,x_1,\ldots,x_n\}$ 的选取。最直观的方案一是等距法 (equidistance)，即在 $[a,b]$ 区间内按照均等距离选取这 $n+1$ 个节点。然而均等距离选取的结点效果并不理想：x 值越接近区间中值 $\frac{|b-a|}{2}$，$|q(x)|$ 越小；反之 x 越接近区间两端，$|q(x)|$ 越大。例如，设 $x_0=a, x_n=b, x_k-x_{k-1}=h, k=1,2,\ldots,n$，那么我们有

$$|q(x)|\leq\frac{h}{2}\frac{h}{2}2h\ldots nh=\frac{h^{n+1}n!}{4},\quad x\in[a,b].$$

等距法对于降低插值误差(7.91)的目标来说，并不理想。

于是我们提出方案二，利用 (切比雪夫) 正交多项式，致力于追求

$$\min\left\{\max_{x\in[a,b]}|q(x)|\right\}.$$

将 $q(x)$ 视作一个在 $[a,b]$ 区间内的首一切比雪夫多项式，满足关系

$$\max q(x)=\Pi_{i=0}^{n}(x-x_i)=(x-x_0)(x-x_1)\ldots(x-x_n)=\frac{T_{n+1}(x)}{2^n},$$

其中 RHS 分母的 2^n 是 $n+1$ 次切比雪夫多项式 $T_{n+1}(x)$ 的首项系数，$\frac{T_{n+1}(x)}{2^n}$ 由此变为首一多项式。当 $a=-1, b=1$ 时，根据上式我们有

$$\max_{x\in[-1,1]}q(x)=\frac{T_{n+1}(x)}{2^n}\leq\frac{1}{2^n},$$

最后一个不等式是根据切比雪夫多项式的定义，在 $[-1,1]$ 之内 $T_n(x)$ 的极值为 $T_n(-1)=-1, T_n(1)=1$。

首一正交切比雪夫多项式 $\frac{T_{n+1}(x)}{2^n}=\Pi_{i=0}^{n}(x-x_i)$ 的根为

$$x_k=\cos\left(\frac{2k+1}{2(n+1)}\pi\right), k=0,1,\ldots,n. \tag{7.92}$$

因此(7.91)改写为

$$\begin{aligned}|f(x)-p(x)|&=\left|q(x)\frac{f^{(n+1)}(\zeta)}{(n+1)!}\right|\\&\leq\min\left\{\max_{x\in[-1,1]}|q(x)|\right\}\left|\frac{f^{(n+1)}(\zeta)}{(n+1)!}\right|\\&=\left|\frac{T_{n+1}(x)}{2^n}\frac{f^{(n+1)}(\zeta)}{(n+1)!}\right|\\&=\frac{1}{2^n(n+1)!}\sup_{\{\zeta\in[-1,1]\}}\left|f^{(n+1)}(\zeta)\right|.\end{aligned} \tag{7.93}$$

当 $a \neq -1, b \neq 1$ 时，(7.92) $\Rightarrow$

$$x_k = \frac{b+a}{2} + \frac{b-a}{2}\cos\left(\frac{2k-1}{2(n+1)}\pi\right), k = 1, 2, \ldots, n+1. \tag{7.94}$$

则(7.93) $\Rightarrow$

$$|f(x) - p(x)| \leq \left(\frac{b-a}{2}\right)^{n+1} \frac{1}{2^n(n+1)!} \sup_{\{\zeta\in[-1,1]\}} \left|f^{(n+1)}(\zeta)\right|. \tag{7.95}$$

进而我们有切比雪夫插值定理 (Chebyshev interpolation theorem)

$$\lim_{n\to\infty} ()^2 \tag{7.96}$$

□

例如，假定 n$f(x) = \tan(4x-1)$，我们采用切比雪夫多项式作基方程，用 $p_n(x)$ 作以近似，见图7.C.6，Matlab 代码如下

图 7.C.6: 切比雪夫插值近似

```
x = chebfun('x');
f = tanh (4 * x -1);
n0=0;n1=1;n2=2;n3=3;n4=4;n5=5;n6=6;n7=7;
pn0 = chebfun(f,n0+1);
```

```
pn1 = chebfun(f, n1+1);
pn2 = chebfun(f, n2+1);
pn3 = chebfun(f, n3+1);
pn4 = chebfun(f, n4+1);
pn5 = chebfun(f, n5+1);
pn6 = chebfun(f, n6+1);
pn7 = chebfun(f, n7+1);

figure(1);
subplot(4,2,1),plot(f, '-b'), hold on, plot(pn0, '.--r'), hold on, plot(f-
pn0, ':k'), title('n=0'),
subplot(4,2,2),plot(f, '-b'), hold on, plot(pn1, '.--r'), hold on, plot(f-
pn1, ':k'), title('n=1'),
subplot(4,2,3),plot(f, '-b'), hold on, plot(pn2, '.--r'), hold on, plot(f-
pn2, ':k'), title('n=2'),
subplot(4,2,4),plot(f, '-b'), hold on, plot(pn3, '.--r'), hold on, plot(f-
pn3, ':k'), title('n=3'),
subplot(4,2,5),plot(f, '-b'), hold on, plot(pn4, '.--r'), hold on, plot(f-
pn4, ':k'), title('n=4'),
subplot(4,2,6),plot(f, '-b'), hold on, plot(pn5, '.--r'), hold on, plot(f-
pn5, ':k'), title('n=5'),
subplot(4,2,7),plot(f, '-b'), hold on, plot(pn6, '.--r'), hold on, plot(f-
pn6, ':k'), title('n=6'),
subplot(4,2,8),plot(f, '-b'), hold on, plot(pn7, '.--r'), hold on, plot(f-
pn7, ':k'), title('n=7'),
```

Theorem 7.57 (切比雪夫截断定理). 切比雪夫截断定理 *(Chebyshev truncation theorem)*是指，我们用 j 次切比雪夫多项式来近似未知的原方程系统 $f(x)$，近似的误差小于等于一系列未被纳入考虑的更高次切比雪夫多项式系数之和，换句话说，定义 $d^j(\cdot|\theta)=\sum_{i=0}^{j}\theta_i\psi_i(x)$，则截断误差 *(truncation error)*

$$d(x)-d^j(x|\theta)\leq\sum_{i=j+1}^{\infty}|\theta_i|,\quad \forall x\in[-1,1],\quad \forall j. \tag{7.97}$$

证明. 略。 □

Theorem 7.58. 第 n 次一类切比雪夫多项式 $T_n(x)$ 有 n 个根

$$x_i=\cos\left(\frac{2i-1}{2n}\pi\right),\quad i=1,2,\ldots,n, \tag{7.98}$$

并且进而可将 $T_n(x)$ 写为

$$T_n(x)=k_n\Pi_{i=0}^{n}\left(x-x_i\right)=2^{n-1}\Pi_{i=0}^{n}\left(x-x_i\right),\quad i=1,2,\ldots,n. \tag{7.99}$$

证明. cf. (Boyd, 2001, p.49)。 □

7.D　AR(1) 过程的离散方法

当未来存在不确定性时，求解经济个体行为最大化的问题便涉及条件期望的计算。以消费——储蓄决策问题为例，可表示为如下贝尔曼方程 (Bellman equation)

$$
\begin{aligned}
&V(a,\lambda)=\max_{\{c,a'\}}\left\{u(c)+\beta E\left[V(a',\lambda')|\lambda\right]\right\},\quad \text{s.t.}\\
&a(1+r)+\omega\exp(\lambda)=a'+c,\\
&\lambda'=(1-\rho)\mu_\lambda+\rho\lambda+\varepsilon,\quad \varepsilon\sim N(0,\sigma_\varepsilon^2),\\
&c\geq 0, a'>0.
\end{aligned}
\tag{7.100}
$$

其中状态矩 λ 的无条件均值和方差为 $\lambda\sim N(\mu_\lambda,\sigma_\lambda^2)$。在给定 λ 的情况下，λ' 是一个 AR(1) 过程，满足条件均值和方差 $\lambda'\sim N\left((1-\rho)\mu_\lambda+\rho\lambda,\sigma_\varepsilon^2\right)$。$\sigma_\varepsilon$ 和 σ_λ 的关系满足

$$\sigma_\lambda=\frac{\sigma_\varepsilon}{\sqrt{1-\rho^2}}.$$

在此例中，涉及条件期望的计算目标是由 $f(\lambda'|\lambda)$ 加权后的价值方程的积分值

$$E\left[V(a',\lambda')|\lambda\right]=\int_{-\infty}^{\infty}V(a',\lambda')f(\lambda'|\lambda)d\lambda',\tag{7.101}$$

其中 $f(\lambda'|\lambda)$ 指给定 λ 的情况下，λ' 的条件密度 (conditional density)。

上式是一个与当前期状态有关的无限维度问题，需要借助数值近似算法。一个可行的近似方案，思路为：将状态空间中原本是连续的状态，离散化变为有限个点，对这些点对应的条件期望做近似。换句话说，就是将关于连续的 λ 的马尔科夫链，转变为离散的有限马尔科夫链，我们将新的马尔科夫链用 $\tilde{\lambda}$ 来予以区分。有限个数的 $\tilde{\lambda}$ 取值可以来自 $\Lambda=\{\lambda_1,\lambda_2,\ldots\lambda_n\}$ 集合，对应转移矩阵 P，含有转移概率 $p_{i,j}$，满足

$$p_{i,j}=Prob\left(\tilde{\lambda}'=\lambda_j|\tilde{\lambda}=\lambda_i\right),\quad i,j=1,2,\ldots,n.$$

通过这种方式，原本是含有条件期望的积分计算问题(7.101)，被转换为离散化的求和问题，可以借助计算机，用某些特定的数值算法求得。下面分别介绍三种常见算法。

7.D.1　Tauchen(1986) 法

Tauchen (1986) 的方法可表示如下：

首先是 λ 的选取。在 Λ 中选取均匀分布的 n 个点 $\{\lambda_1,\lambda_2,\ldots\lambda_n\}$，其中 λ_1 和 λ_n 分别对应上下边界，值等于无条件均值 μ_t 加上或减去无条件标准差 σ_λ 的 m 倍数

$$
\begin{aligned}
\lambda_1&=\mu_\lambda-m\sigma_\lambda,\\
\lambda_2&=\mu_\lambda+m\sigma_\lambda,
\end{aligned}
$$

经验研究中，m 的取值通常在 2 到 3 之间。

随后是转移矩阵 P 的选取。设

$$\omega=\lambda_j-\lambda_{j-1},$$

进而

$$
p_{i,j}=\begin{cases}
Prob\left[(1-\rho)\mu_\lambda+\rho\lambda_i+\varepsilon\leq\lambda_{\{j=1\}}+\frac{\omega}{2}\right] & \text{如果}j=1,\\
Prob\left[\lambda_j-\frac{\omega}{2}\leq(1-\rho)\mu_\lambda+\rho\lambda_i+\varepsilon\leq\lambda_j+\frac{\omega}{2}\right] & \text{如果}j=2,3,\ldots,n-1,\\
1-Prob\left[\lambda_{\{j=n\}}+\frac{\omega}{2}\leq(1-\rho)\mu_\lambda+\rho\lambda_i+\varepsilon\right] & \text{如果}j=n.
\end{cases}
$$

上式进一步整理为

$$p_{i,j}=\begin{cases}\Phi\left(\frac{\lambda_1+\frac{\omega}{2}-(1-\rho)\mu_\lambda-\rho\lambda_i}{\sigma_\varepsilon}\right) & \text{如果}j=1,\\ \Phi\left(\frac{\lambda_j+\frac{\omega}{2}-(1-\rho)\mu_\lambda-\rho\lambda}{\sigma_\varepsilon}\right)-\Phi\left(\frac{\lambda_j-\frac{\omega}{2}-(1-\rho)\mu_\lambda-\rho\lambda}{\sigma_\varepsilon}\right) & \text{如果}j=2,3,\ldots,n-1,\\ 1-\Phi\left(\frac{\lambda_n-\frac{\omega}{2}-(1-\rho)mu_\lambda-\rho\lambda}{\sigma_\varepsilon}\right) & \text{如果}j=n,\end{cases}$$

其中 $\Phi(\cdot)$ 表示标准正态累积分布函数 (standard normal cumulative distribution function)。

Tauchen (1986) 法，将 $\tilde{\lambda}'|\tilde{\lambda}=\lambda_i$ 作为对 $\lambda'|\lambda=\lambda_i$ 的条件概率的离散近似。n 越大，对节点的选取越密，近似的效果越精确。Tauchen (1986) 指出当 $n=9$ 时所获得的近似精度基本已能满足经验研究的一般需要[20]。

举例说明。对于一个随机 AR(1) 连续过程

$$z_t=0.95z_{t-1}+\varepsilon_t$$

以及正态分布 $N(0,0.007^2)$ 而言，假定我们想要对它做离散化近似，对应 $n=9$ 个点，$m=3$。

我们可以根据Tauchen (1986) 的思路编写代码，或者使用现有程序。以 Matlab 为例，Iskander Karibzhanov 在他的个人主页[21]提供了名为 tauchen.m 的程序。Matlab 代码如下

```
>> [y,P,d]=tauchen(5,0,0.95,0.007,3)

y =

   -0.0673   -0.0336         0    0.0336    0.0673


P =

    0.9727    0.0041    0.0000    0.0000    0.0000
    0.0273    0.9806    0.0082    0.0000    0.0000
    0.0000    0.0153    0.9837    0.0153    0.0000
         0    0.0000    0.0082    0.9806    0.0273
         0         0    0.0000    0.0041    0.9727


d =

    0.0361    0.2392    0.4494    0.2392    0.0361
```

即，我们得到了

$$z_t\in\{-0.0673,-0.0336,0,0.0336,0.0673\}$$

[20]然而随着外生过程的持续性越来越强 (high persistence) 即 $\rho\to1$，近似的精度就越来越低，这就需要我们提高 n 的数量以弥补精度的下降，或是采取更适合离散化高持续性过程的算法，如第7.D.3节的Rouwenhorst (1995) 法。

[21]http://karibzhanov.com。

对应转移矩阵 P

$$P_{z,z'} = \begin{pmatrix} 0.9727 & 0.0041 & 0 & 0 & 0 \\ 0.0273 & 0.9806 & 0.0082 & 0 & 0 \\ 0 & 0.0153 & 0.9837 & 0.0153 & 0 \\ 0 & 0 & 0.0082 & 0.9806 & 0.0273 \\ 0 & 0 & 0 & 0.0041 & 0.9727, \end{pmatrix}$$

需要注意的是 $P_{z,z'}$ 对角元素接近于 1，这说明随机过程的当期状态与上期状态高度相关 (对应 $\rho = 0.95$)。Tauchen (1986) 增加 n 的值到 9 后，相关度仍然比较高 (见下)，这需要我们探讨其他数值算法，见第7.D.3节。

```
>> [y,P,d]=tauchen(9,0,0.95,0.007,3)

y =

   -0.0673   -0.0504   -0.0336   -0.0168         0    0.0168    0.0336    0.0504    0.0673


P =

    0.7644    0.0592    0.0001    0    0    0    0    0    0
    0.2347    0.7405    0.0747    0.0001    0    0    0    0    0
    0.0009    0.1997    0.7569    0.0931    0.0002    0    0    0    0
    0    0.0006    0.1679    0.7669    0.1147    0.0002    0    0    0
    0    0    0.0004    0.1396    0.7702    0.1396    0.0004    0    0
         0    0    0    0.0002    0.1147    0.7669    0.1679    0.0006    0
         0         0    0    0    0.0002    0.0931    0.7569    0.1997    0.0009
         0         0         0    0    0    0.0001    0.0747    0.7405    0.2347
         0         0         0         0         0    0    0.0001    0.0592    0.7644


d =

    0.0108    0.0428    0.1144    0.2064    0.2513    0.2064    0.1144    0.0428    0.0108
```

即，我们得到了

$$z_t \in \{-0.0673, -0.0504, -0.0336, -0.0168, 0, 0.0168, 0.0336, 0.0504, 0.0673\},$$

对应转移矩阵 P

$$P_{z,z'} = \begin{pmatrix} 0.7644 & 0.0592 & 0.0001 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2347 & 0.7405 & 0.0747 & 0.0001 & 0 & 0 & 0 & 0 & 0 \\ 0.0009 & 0.1997 & 0.7569 & 0.0931 & 0.0002 & 0 & 0 & 0 & 0 \\ 0 & 0.0006 & 0.1679 & 0.7669 & 0.1147 & 0.0002 & 0 & 0 & 0 \\ 0 & 0 & 0.0004 & 0.1396 & 0.7702 & 0.1396 & 0.0004 & 0 & 0 \\ 0 & 0 & 0 & 0.0002 & 0.1147 & 0.7669 & 0.1679 & 0.0006 & 0 \\ 0 & 0 & 0 & 0 & 0.0002 & 0.0931 & 0.7569 & 0.1997 & 0.0009 \\ 0 & 0 & 0 & 0 & 0 & 0.0001 & 0.0747 & 0.7405 & 0.2347 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.0001 & 0.0592 & 0.7644 \end{pmatrix}$$

7.D.2 Tauchen and Hussey (1991) 法

同样对矩 λ 而言，Tauchen and Hussey (1991) 使用高斯——埃米特求积 (Gauss-Hermite quadrature)去近似价值方程的条件期望积分，(7.101)改写为

$$E\left[V(a',\lambda')|\lambda\right] \approx \frac{1}{\sqrt{\pi}}\sum_{j=1}^{n}\omega_j V(a',\sqrt{2}\sigma_\lambda\hat{\lambda}'_j+\mu_\lambda), \quad j=1,2,\ldots,n \tag{7.102}$$

其中 $\hat{\lambda}'_j$ 是第 n 次埃米特多项式 $H_n(\lambda)$ 的根 (第7.C.1节)，满足

$$\lambda_j = \sqrt{2}\sigma_\varepsilon\hat{\lambda}_j + \mu_\lambda,$$

ω_j 是对应 $\hat{\lambda}'_j$ 的权重。

研究目标与(7.101)相同：给定 λ 的情况下有 AR(1) 过程 λ'，计算加权价值方程积分

$$E\left[V(a',\lambda')|\lambda\right] = \int_{-\infty}^{\infty} V(a',\lambda')f(\lambda'|\lambda)d\lambda'.$$

原本可以采用上节7.D.1介绍的Tauchen (1986) 算法做近似求解，但需注意 λ' 的条件均值是一个关于 λ 的方程，这意味着对于每个不同的 λ，都会有一组不同的 λ'，进而迭代算出的集合 Λ 可能有无数个。为了避免这个问题，Tauchen and Hussey (1991) 建议使用如下转换

$$\begin{aligned} E\left[V(a',\lambda')|\lambda\right] &= \int_{-\infty}^{\infty} V(a',\lambda')f(\lambda'|\lambda)d\lambda' \\ &= \int_{-\infty}^{\infty} V(a',\lambda')\frac{f(\lambda'|\lambda)}{f(\lambda'|\mu_\lambda)}f(\lambda'|\mu_\lambda)d\lambda'. \end{aligned}$$

其中 $f(\lambda'|\mu_\lambda)$ 表示在 λ 恰好等于其无条件均值 μ_λ 的情况下，λ' 的条件密度。将高斯——埃米特求积(7.102)代入上式，可得

$$\begin{aligned} &E\left[V(a',\lambda')|\lambda\right] \approx \frac{1}{\sqrt{\pi}}\sum_{j=1}^{n}\omega_j V(a',\lambda_j)\frac{f(\lambda_j|\lambda)}{f(\lambda_j|\mu_\lambda)}, \quad \text{其中} \\ &\lambda_j = \sqrt{2}\sigma_\varepsilon\hat{\lambda}_j + \mu_\lambda, \\ &\hat{\lambda}_j, j=1,2,\ldots,n \text{表示第 n 次埃米特多项式的根,} \\ &\omega_j \text{表示第 n 次埃米特多项式的根对应的权重.} \end{aligned} \tag{7.103}$$

那么，对于 $\lambda=\lambda_i$，我们有

$$E\left[V(a',\lambda'|\lambda)\right] \approx \sum_{j=1}^{n}\tilde{\omega}_{i,j}V(a',\lambda_i), \quad \tilde{\omega}_{i,j} = \frac{1}{\sqrt{\pi}}\omega_j\frac{f(\lambda_j|\lambda_i)}{f(\lambda_j|\mu_\lambda)}, \tag{7.104}$$

上式意味着，我们选取 $\tilde{\lambda}$ 的可能集合为

$$\Lambda = \left\{\tilde{\lambda}|\lambda = \sqrt{2}\sigma_\varepsilon \hat{\lambda}_j,\ j = 1, 2, \ldots, n\right\},$$

选取 $\tilde{w}$ 的可能集合为转移矩阵 P：需要注意的是

$$\sum_{j=1}^{n} \tilde{\omega}_{i,j} \neq 1,$$

因此 P 中的元素

$$p_{i,j} = \frac{\tilde{\omega}_{i,j}}{\sum_{j=1}^{n} \omega_{i,j}}.$$

(7.104)因此最终改写为

$$E\left[V\left(a', \tilde{\lambda}'|\lambda_i\right)\right] \approx \sum_{j=1}^{n} p_{i,j} V\left(a', \lambda_j\right). \tag{7.105}$$

由(7.105)可见，利用Tauchen and Hussey (1991) 近似条件期望用到高斯——埃米特求积的知识，但在实际计算过程中并不需要做高斯——埃米特求积[22]。

与Tauchen (1986) 法相比，Tauchen and Hussey (1991) 的不同在于权重方程改用条件密度 $\omega(\lambda') = f(\lambda'|\mu_\lambda)$，如(7.103)。与无条件密度相比，有条件密度在区间的中央赋予更多权重，在区间的两端赋予更小权重。这样一来，有条件密度作为权重方程，可以更好平衡两组互相冲突的原则：第一，一个好的近似总是会在无条件均值附近留下较多权重，第二，一个好的近似总是会使得在区间的两端，$\frac{f(\lambda'|\lambda)}{\omega(\lambda')}$ 比值不至于相对于 $w(\lambda')$ 而增长过快。

Iskander Karibzhanov 也提供了名为 tauchenHussey.m 的 Matlab 程序。

7.D.3　Rouwenhorst(1995) 法

如前文所述，对于高持续性 $\rho \to 1$(如 $\rho = 0.9$) 的过程来说，Tauchen (1986) 近似法的精度较低。Kopecky and Suen (2010) 研究发现Rouwenhorst (1995) 法在此时更为理想。

在区间 $[\mu_\lambda - \nu, \mu_\lambda + \nu]$ 内选取对称且均匀分布的 n 个点构成 Λ。用如下方法计算 P：通过选择 p 和 p，使得当时

$$P_{n=2} = \begin{pmatrix} p & 1-p \\ 1-1 & q \end{pmatrix}, \quad P_{n=3} = \begin{pmatrix} p^2 & 2p(1-q) & (1-p)^2 \\ p(1-q) & pg + (1-p)(1-q) & q(1-p) \\ (1-q)^2 & 2q(1-p) & q^2 \end{pmatrix},$$

进而用递归方法，根据 P_{n-1} 计算 P_n，分两步。首先将 $n \times n$ 矩阵相加

$$p\begin{pmatrix} P_{n-1} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{pmatrix} + (1-p)\begin{pmatrix} \mathbf{0} & P_{n-1} \\ 0 & \mathbf{0}^\top \end{pmatrix},$$

$$(1-q)\begin{pmatrix} \mathbf{0}^\top & 0 \\ P_{n-1} & \mathbf{0} \end{pmatrix} + q\begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & P_{n-1} \end{pmatrix},$$

随后将除了第一和最后一行之外的其他行都除以 2。这样矩阵 P_n 的条件概率之和等于 1。

[22]然而随着 $n \to \infty$，我们有

$$\lim_{n\to\infty} \sum_{j=1}^{n} \tilde{\omega}_{i,j} \to 1, \quad \forall i = 1, 2, ldots, n,$$

因此(7.105)的Tauchen and Hussey (1991) 算出的条件期望近似依然是向高斯——埃米特求积法(7.102)收敛的。

p^{n-1} 表示当期已处于最低状态时，下期仍处在最低状态的概率；q^{n-1} 表示当期已处于最高状态时，下期仍处在最高状态的概率。$(1-p)^{n-1}$ 表示当期已处于最低状态时，下期由最低转为最高状态的概率；$(1-q)^{n-1}$ 表示当期已处于最高状态时，下期由最高转为最低状态的概率。对于任何 $p \neq q$ 的情况，都会导致生成的冲击序列中出现异方差 (heteroscedasticity)。

不论 n 和 Λ 的取值是怎样的，$\tilde{\lambda}$ 过程的一届序列相关性总是 $p+q-1$。因此，如果我们设定 $p+q=\pi$，就可以使得所生成的离散过程，具有和连续过程相同的一阶持续性 ρ。

此外，$\tilde{\lambda}$ 的方差也总是 $\frac{\nu^2}{n-1}$。因此对于某个给定的 n 值，1 个合适的 ν 值取值是使得

$$var\left(\tilde{\lambda}\right) = var\left(\lambda\right) \quad \Rightarrow \nu = \left(\frac{n-1}{\rho^2-1}\right)^{\frac{1}{2}} \sigma_\varepsilon.$$

Iskander Karibzhanov 的 matlab 代码：

```
>> rouwenhorst(0.95,0.007,5)

ans =

   -0.0448   -0.0224         0    0.0224    0.0448

>> rouwenhorst(0.95,0.007,7)

ans =

   -0.0549   -0.0366   -0.0183         0    0.0183    0.0366    0.0549

>> rouwenhorst(0.95,0.007,9)

ans =

   -0.0634   -0.0476   -0.0317   -0.0159         0    0.0159    0.0317    0.0476    0.0634
```

7.E 伽辽金近似

7.E.1 举例：1 维空间中的伽辽金近似

以一个 1 维空间的有界极值问题为例，介绍伽辽金近似法 (Galerkin approximation method)。假定 $u(x)$ 是下述方程系统的解

$$\begin{cases} -u'' + u = f, \quad \text{in}(0,1) \\ u(0) = u(1) = 0, \end{cases} \tag{7.106}$$

我们的目标是找到一个对 u 的近似解。对上式两侧乘以一个检验方程 ν，然后分段积分，可得(7.106)的弱条件形式

条件补一个
ference

$$\int_0^1 \left(u'\nu' + u\nu\right) dx = \int_0^1 f\,\nu dx, \quad \forall \nu \in H_0^1(0,1), \tag{7.107}$$

其中 $H_0^1(0,1)$ 是索伯列夫空间 (Sobolev space)。

一个 reference

如果方程 u 是常规方程，即连续且二阶可导，那么(7.106)与(7.107)等价。这意味着我们可以利用(7.107)，构建一个多项式近似 u。将域 $[0,1]$ 划分为 $N+1$ 个均匀分布 (uniform distribution) 的子域，每个域对应子空间 (x_j, x_{j+1})，满足

$$x_j = \frac{j}{N+1} \quad j = 0, 1,$$

此外，定义一个包含检验方程 ν 的空间 V_N，满足

$$V_N = \left\{\nu \in C^0 : \nu|_{x_j, x_{j+1}} \text{是线性方程，并且} \nu(0) = \nu(1) = 0\right\},$$

其中 C^0 表示连续方程的空间。

V_N 满足如下特征：首先，V_N 是个有限维度空间，$N < \infty$。其次，$V_N \subset H_0^1(0,1)\,\forall N$。事实上，$\nu(x) \in V_N$ 由且仅由他所对应的有限个数的点 $(x_1, x_2, \ldots, x_N)$ 所决定。

定义一个伽辽金近似方程 $u_N \in V_N$ 作为对 u 的近似，代入(7.107)，我们有

$$\int_0^1 (u_N'\nu' + u_N\nu)\, dx = \int_0^1 f\,\nu\, dx, \quad \forall \nu \in V_N, \tag{7.108}$$

由下节可知，存在且只存在一个 u_N 方程满足(7.108)条件。并且由于 V_N 是个有限维度空间，因此 u_N 可以通过一些线性方程系统求解方法算得。基方程 $\psi_j \in V_N$ 可以采取不同形式，例如可以选拉格朗日多项式 (Lagrange polynomial)

解的存在性和唯一性

$$\psi_j(x_i) = \delta_{ij}, \quad i, j = 1, 2, \ldots, N, \tag{7.109}$$

其中 δ_{ij} 表示克罗内克乘积。因此我们有近似方程

$$u_N = \sum_{j=1}^{N} U_j \psi_j, \quad U_j \in \mathbb{R}, \tag{7.110}$$

其中 $U_j := u_N(x_j)$。

既然任何一个检测方程 $\nu \in V_N$ 都与 ψ_j 呈线性组合关系，(7.108)等价于

$$\int_0^1 (u_N'\psi_k' + u_N\psi_k)\, dx = \int_0^1 f\,\psi_k\, dx, \quad k = 1, 2, \ldots, N. \tag{7.111}$$

再引入(7.110)，替换上式 LHS 中的 u_N，我们有

$$\sum_{j=1}^{N} U_j \int_0^1 \left(\psi_j'\psi_k' + \psi_j\psi_k\right) dx = \int_0^1 f\,\psi_k\, dx, \quad k = 1, 2, \ldots, N. \tag{7.112}$$

(7.112)构成一个 N 个方程的线性系统

$$AU = F,$$

$$\begin{cases} \text{系数矩阵 } A = (a_{kj}) \in \mathbb{R}^{N\times N}, & a_{kj} = \int_0^1 \left(\psi_j'\psi_k' + \psi_j\psi_k\right) dx, \\ \text{系数矩阵 } F = (F_k) \in \mathbb{R}^N, & F_k = \int_{0,1} f\,\psi_k d_x, \\ \text{待求解矩阵 } U = (U_j) \in \mathbb{R}^N \end{cases} \tag{7.113}$$

通过求解线性方程系统(7.113)，可得矩阵 U 的解，进而根据(7.110)可得 u_N 的值。

通过观察 $\psi_j(x_i)$ 的性质(7.109)可见，矩阵 A 是个对称的三角对角矩阵，对角元素及其附近不等于 0 的元素分别为

$$\begin{cases} a_{jj} = \frac{2}{h} + \frac{2}{3}h, \\ a_{jj-1} = a_{jj+1} = -\frac{1}{h} + \frac{h}{6}, \end{cases}$$

由此，待求解方程系统为 $N+1$ 个等式

$$\begin{cases} U_0 := 0, \\ \frac{-U_{j-1}+2U_j-U_{j+1}}{h} + \frac{h}{6}U_{j-1} + \frac{2}{3}hU_j + \frac{h}{6}U_{j+1} = F_j, \quad j = 1, 2, \ldots, N, \\ U_{j+1} := 0. \end{cases}$$

矩阵 A 是对称正定的 (即 $C^\top AC > 0 \forall C \neq 0$；由 a_{ij} 积分项中的双线性形式推得)，由此可得 A 是可逆矩阵。因此，U 进而 u_N 的解是唯一的。

的存在性和唯性

上式中两侧同时除以 h，我们得到(7.106)的有限差分近似形式

$$\underbrace{\frac{-U_{j-1}+2U_j-U_{j+1}}{h^2}}_{\approx u''(x_j)} + \underbrace{\frac{1}{6}U_{j-1} + \frac{2}{3}U_j + \frac{1}{6}U_{j+1}}_{\approx u(x_j)} = \underbrace{\frac{1}{h}F_j}_{\approx f(x_j)}, \quad j = 1, 2, \ldots, N,$$

也就是说，在这个 1 维空间的例子中，伽辽金近似法就是将原求解问题转换为一个已知数量 N 的有限差分近似问题。

收敛问题。对于给定数量 N，我们定义一个伽辽金近似 $u_N \in V_N$，作为对原方程系统 u 的近似。随着 $N \to \infty$，正交多项式的级数逐渐增加，可用来近似任何连续方程。换句话说，当空间集合 V_N 以如下方式近似 u 时，伽辽金近似 u_N 逐渐收敛至 u：

$$\lim_{N\to\infty} ||u - v|| = \lim_{N\to\infty} \inf_{\nu \in V_N} ||u - v|| \to 0.$$

7.E.2 伽辽金近似

第7.E.1节以 1 维空间为例介绍了伽辽金近似的基本思路和方法。这节我们从 1 维扩展到通用情况，即在一个希尔伯特空间 (Hilbert space)中，双线性形式下的伽辽金近似法。

一个描述：lbert space

将一个希尔伯特空间定义为 V，在其中定义一个双线性形式的连续方程 $a(.,.)$ 和一个线性方程 L，我们的研究目标是找到一个近似解 $u \in V$，求解原方程系统

$$a(u, \nu) = \langle L, \nu \rangle, \quad \forall \nu \in V, \tag{7.114}$$

其中 $\langle .,. \rangle$ 表示在 V' 和 V 中的双内积 (duality inner product)[23]下面我们来分析，为了确保 u 存在唯一解，方程 $a(.,.)$ 所应呈现的形式，即关于强制形式 (coercive forms)的讨论。

Definition 7.6 (方程的强制形式)**.** *如果存在一个常数 $\alpha > 0$，使得满足*

$$a(u, u) \geq \alpha ||u||_V^2, \quad \forall u \in V, \tag{7.115}$$

那么我们说方程 $a(.,.)$ 是一个 V 空间中的强制形式。

举例说明，以(7.114)为例，对于二阶线形椭圆方程

$$\begin{cases} -\sum_{i,j=1}^n \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) = f, & \text{in}\,\Omega \subset \mathbb{R}^n, \\ u = 0, & \text{on}\,\partial\Omega \end{cases}$$

其中系数 $a_{ij} = a_{ij}(x)$ 是有界方程。那么存在任意 $\gamma > 0$ 使得

$$\gamma|\xi|^2 \leq \sum_{i,j=1}^n a_{ij}\xi_i\xi_j, \quad \forall x \in \Omega, \forall \xi \in \mathbb{R}^n, \tag{7.116}$$

或者说，可以以(7.114)的形式来表述，其中

[23]如 $\langle A, B \rangle := A'B' + AB, \forall (A, B) \in V, (A', B') \in V'$。注意和(7.32)的定义有所区别。

- 希尔伯特空间 V

$$V = H_0^1(\Omega) = \left\{ \nu \in L^2(\Omega) : \frac{\partial \nu}{\partial x_j} \in L^2(\Omega), \quad j = 1, 2, \ldots, n, \text{以及在} \partial\Omega \text{上有} \nu = 0 \right\},$$

- $\nu \in V$ 的范数 (norm) 为

$$||\nu||_{H^1} = ||\nu||_{L^2} + ||\nabla \nu||_{L^2},$$

- $a(u, \nu)$ 为

$$a(u, \nu) = \sum_{i,j=1}^{n} \int_\Omega a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \, dx,$$

- L 满足

$$\langle L, \nu \rangle = \int_\Omega f \, \nu \, dx.$$

结合椭圆方程条件(7.116)、系数 a_{ij} 的有界特性以及庞加莱不等式，可得 a 函数形式是强制且连续的。线性方程形式 L 如果满足 $f \in L^2$，则我们说它是连续的。

再举一个线性弹性方程的例子。对于应变——应力关系方程

$$\begin{cases} -\mu \Delta \boldsymbol{u} - (\lambda + \mu) \nabla div \boldsymbol{u} = \boldsymbol{f}, & \text{in} \, \Omega \subset \mathbb{R}^3, \\ \boldsymbol{u} = 0, & \text{on} \, \partial\Omega \end{cases}$$

其中出于简化考虑，设满足齐次狄利克雷条件 (homogeneous Dirichlet condition) 。大于 0 的系数 λ 和 μ 分别为拉梅第一参数和拉梅第二参数 (Lamé elasticity parameters)。向量方程 $\boldsymbol{u}$ 和 $\boldsymbol{f}$ 未知。这个问题的弱形式可以表述为带有 $V = H_0^1(\Omega)^3$ 条件的式(7.114)，其中 $a(\boldsymbol{u}, \boldsymbol{v})$ 满足条件

$$a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \{ 2\mu \varepsilon_{i,j}(\boldsymbol{u}) \varepsilon_{i,j}(\boldsymbol{v}) + \lambda \, div \boldsymbol{u} \, div \boldsymbol{v} \} \, dx, \quad \varepsilon_{i,j} := \frac{1}{2} \left(\frac{\partial \nu_i}{\partial x_j} \frac{\partial \nu_j}{\partial x_i} \right).$$

进而，利用科恩不等式 (Korn's inequality)，我们有双线性形式的方程 $a(.,.)$ (7.114) 是强制形式。

方程形式 $a(.,.)$ (7.114)的解得存在性和唯一性，可由拉克斯一密格拉蒙定理 (Lax-Milgram theorem, Theorem 8.14)予以证明。

在此基础上，上述条件也证实了伽辽金近似方程向原方程收敛。事实上，的确存在着一些重要范例如斯托克斯公式 (Stokes equations)等，其中的双线性形式方程并不是强制的，但满足一个更弱的条件，称为上极限——下极限条件 (inf-sup condition) [24]。下面先介绍上线——下线极限条件定义。在第7.E.3节介绍双线性形式伽辽金近似的收敛问题。

Definition 7.7 (上极限——下极限关系问题)**.** *如果双线性形式的方程 $a(u, \nu)$ 在 V 空间上满足条件：对于某一 $\alpha > 0$，有*

$$\sup_{\nu \in V} \frac{a(u, \nu)}{||\nu||_V} \geq \alpha ||u||_V, \quad \forall u \in V, \tag{7.117a}$$

$$\sup_{u \in V} \frac{a(u, \nu)}{||u||_V} \geq \alpha ||\nu||_V, \quad \forall \nu \in V, \tag{7.117b}$$

那么我们说，它满足上极限——下极限关系。

根据这一定义，我们有以下推论：

[24] 关于稳态斯托克斯公式弱形式的分析，以及上极限——下极限条件，可参考 Chen Long 的讲义 https://www.math.uci.edu/ chenlong/226/infsup.pdf。

1. 如果 a 是对称方程，(7.117a)和(7.117b)等价。

2. 两个方程也可以写作
$$\inf_{u\in V}\sup_{\nu V}\frac{a(u,\nu)}{||u||_V\,||\nu||_V}. \tag{7.118}$$

3. 如果 a 是强制方程，那么它满足上极限——下极限关系，这是由于
$$\sup_{\nu\in V}\frac{a(u,\nu)}{||\nu||_V}\geq\frac{a(u,\nu)}{||u||_V}\geq\alpha||u||_V. \tag{7.119}$$

4. 双线性方程 a 的上极限——下极限关系可以表示为关于 a 的线性算子 A 及其伴随算子 A^*，
$$A:V\to V',\quad A^*:V\to V',$$
两个算子定义为
$$\langle Au,\nu\rangle_{V'\times V}=a(u,\nu),$$
$$\langle u,A^*\nu\rangle_{V'\times V}=a(u,\nu),$$
从而(7.117a)-(7.117b)等价于
$$||Au||_{V'}\geq\alpha||u||_V,\quad\forall u\in V, \tag{7.120a}$$
$$||A^*\nu||_{V'}\geq\alpha||\nu||_V,\quad\forall\nu\in V. \tag{7.120b}$$

5. 当 $V=\mathbb{R}^n$ 时，a 的强制性意味着，其对应的矩阵 A 正定；a 若满足上极限——下极限关系，则 A 进一步是可逆的。

Theorem 7.59. *对于连续的双线性形式方程 a，当且仅当线性算子 A 是双射 (bijective)的时，a 满足上极限——下极限关系。A 是双射的意思是说，原方程系统(7.114)对于任意 $L\in V$ 都有唯一解，进而 A 有一个连续逆 (continuous inverse)，即 $||u||_V\leq C||L||_{V'}$* [25]。

证明. 根据泛函分析的基本原理可得，对于 $W\subset V$，我们定义 $W^0\subset V'$，满足
$$W^0\subset\{L\in V':\langle L,\nu\rangle=0,\quad\forall\nu\in W\},$$
那么我们有
$$(KerA)^0=\overline{\Im A^*}, \tag{7.121a}$$
$$(KerA^*)^0=\overline{\Im A}, \tag{7.121b}$$
其中 $\Im(\cdot)$ 表示复数的虚部 (imaginary part)；对应地，复数的实部 (real part)表示为 $\Re(\cdot)$。

如果假定 a 满足上极限——下极限关系，那么由(7.120a)可得 A 是单射 (injective)的，由(7.120b)可得 A^* 也是单射的。进而，如果我们能证明 $\Im A$ 是封闭集，则根据(7.121b)可以证明定理。

假定 $Au_n\mapsto w$，则由(7.120a)可得
$$||A(u_n-u_m)||_{V'}\geq\alpha||u_n-u_m||_V,$$
由此可见 $\{u_n\}$ 是一个柯西序列 (Cauchy sequnce)，u_n 收敛向某一个 $u\in V$，并且由 A 的连续性特征可得，$w=Au\in\Im A$。

同理可得，如果 A 是双射的，那么 A^* 也是双射的，A 和 A^* 都有一个连续逆。因此(7.117a)-(7.117b)成立。 □

[25] 关于连续逆，可参考讲义 Li Shiu-Tang 的讲义 http://www.math.utah.edu/li/Right%20continuous%20inverse.pdf。

在此基础上，我们为原方程系统(7.114)作伽辽金近似解。假定在 V 空间内，我们有一个有限维度的子空间 V_N。则对原方程系统的伽辽金近似 $u_N \in V_N$ 可定义为

$$a(u_N, \nu) = \langle L, \nu \rangle, \quad \forall \nu \in V_N. \tag{7.122}$$

为了让 u_N 定义良好 (well defined)，我们需要对 a 作出限定。根据 Theorem 7.59可得当且仅当 a 在 V_N 子空间中满足上极限——下极限关系时，伽辽金近似解 u_N 才是存在且唯一的。尤其是，伽辽金近似解 u_N 对于强制形式的方程 a 是定义良好的。

从这个角度出发，有必要指出一下两种 V 中方程形式的本质区别：一种是强制形式，另一种是满足上极限——下极限关系，但并不是强制形式的：

Remark 7.1. *如果 a 在 V 上强制，那么它在 V 的任何一个子空间中，尤其是 V_N 中都是强制的，此时伽辽金近似解 u_N 定义完备。*

但 V 上的上极限——下极限关系，无法传导到 V 的子空间中，那么对于不满足强制性的方程形式，还需要在子空间 V_N 中重新检验，以确保 u_N 定义完备。

7.E.3　强制型方程形式的伽辽金近似收敛

假定我们已知双线性形式方程 $a(.,.)$ 是连续和强制的。因此我们有

$$a(u, \nu) \leq M\, ||u||_V ||\nu||_V, \quad \forall u, \nu \in V, \tag{7.123}$$

其中 M 定义为一个连续性常数 (continuity constant)。在采用伽辽金近似法求得近似解 u_N 后，一个自然出现的问题是：在 V_N 被恰当选取的前提下，是否会出现 $\lim_{N\to\infty} u_N \to u$ 的收敛?

显然，如果答案是肯定的，那么

$$d(u, V_n) = \inf_{\nu \in V_n} ||u - \nu||_V \to 0, \quad \text{随着 } N \to \infty. \tag{7.124}$$

换句话说，收敛与否的问题，等价于探讨在子空间 V_N 中(7.124)是否成立[26]，以及它是不是收敛的充要条件。根据齐亚引理 (Céa's lemma)，答案是肯定的。

Lemma 7.1 (齐亚引理)**.** *如果 a 是一个连续且强制的方程，那么我们有*

$$||u - u_N||_V \leq \frac{M}{\alpha} \inf_{\nu \in V_N} ||u - \nu||_V \tag{7.125}$$

证明. 将伽辽金近似解 (7.122) 和(7.123)联立，构成误差方程

$$a(u - u_n, \nu) = 0, \quad \forall \nu \in V_N. \tag{7.126}$$

由(7.123)，，(7.124)和 (7.126)我们有

$$\alpha ||u - u_N||^2 \leq a(u - u_N, u - u_N) = a(u - u_N, u - \nu) \leq M ||u - u_N||_V ||u - \nu||_V,$$

因此

$$||u - u_N|| \leq M ||u - u_N||_V ||u - \nu||_V, \forall \nu \in V_N,$$

证毕。 □

根据齐亚引理，在子空间 V_N 中，伽辽金近似解 u_N 是对原方程系统解 u 最好的近似，它除了受一个常数影响外，但只受双线性方程 a 形式的影响，而与子空间 V_N 的情况无关。根据这一特征，我们可以引出伽辽金近似解的收敛特性。

[26]子空间 V_N 中的所有方程，都是对 u 的近似。

Theorem 7.60 (伽辽金近似解的收敛 (强制形式))**.** *如果 a 是个连续且强制的方程，并且空间 V_N 满足 (7.124)，那么我们有*

$$\lim_{N\to\infty} u_N = u.$$

来讨论一个特殊情况，a 还额外具有对称性。那么在 V 上，a 定义了一个标量积 (scalar product)，这个标量积与原方程系统相同。此时，误差方程(7.126)表明伽辽金近似解 u_N 恰好就是 u 在子空间 V_N 上的正交映射，对应标量积方程 a。此时标量积 a 所对应的范就提供了最佳近似解。并且对应地，齐亚引理中的常数项 M/α 变为 $\sqrt{M/\alpha}$。

7.E.4 仅满足上极限——下极限关系方程形式的伽辽金近似收敛

现在假定 a 不是强制的，但满足上极限——下极限关系(7.117a)-(7.117b)。已知原方程系统(7.114)有唯一解，我们来讨论一下此时伽辽金近似的收敛问题。如 Remark 7.1所强调的，上极限——下极限关系无法从 V 向下传递到子空间 V_N，因此为了让伽辽金近似是定义良好的，我们需要假定 a 也在子空间 V_N 中满足上极限——下极限关系，即存在一个系数 $\beta > 0$，使得

$$\sup_{\nu\in V_N} \frac{a(u,\nu)}{||\nu||_V} \geq \beta||u||_V, \quad \forall u \in V_N. \tag{7.127a}$$

$$\sup_{u\in V_N} \frac{a(u,\nu)}{||u||_V} \geq \beta||\nu||_V, \quad \forall \nu \in V_N \tag{7.127b}$$

我们证明(7.127a)即可。这是因为既然子空间 V_N 是有限维度的，那么一旦(7.127a)成立，(7.127b)也同样成立。

收敛若要存在，需要假定 β 不依赖于 N。进而，Babuška (1971) 对齐亚引理 7.1作以扩展，据此可以对 Theorem 7.60作扩展。

Lemma 7.2 (扩展齐亚引理)**.** *如果 a 是连续方程，满足(7.127a)的上极限——下极限关系，那么*

$$||u - u_N||_V \leq \left(\frac{\beta + M}{\beta}\right) \inf_{\nu\in V_N} ||u - \nu||_V,$$

尤其是当 β 不依赖于 N 时，误差项中的常数也不依赖于 N。

证明. 取 $\nu \in V_N$。由(7.127a)及误差方程(7.126)可得

$$\beta||\nu - u_N||_V \leq \sup_{w\in V_N} \frac{a(\nu - u_N, w)}{||w||_V} = \sup_{w\in V_N} \frac{a(\nu - u, w)}{||w||_V} \leq M||\nu - u||_V.$$

□

由此我们有收敛定理：

Theorem 7.61 (伽辽金近似解的收敛 (上极限——下极限形式))**.** *如果 a 是个连续方程，满足上极限——下极限关系(7.127a)，β 不依赖于 N，并且子空间 V_N 满足(7.124)，那么我们有*

$$\lim_{N\to\infty} u_N = u.$$

7.F 数值求积

我们常常需要计算一个定积分的值，如

$$\int_a^b f(x)\,\mathrm{d}x,$$

若直接用解析法求解较为困难，则常常采取数值积分 (numerical integratation)的思路，将积分式转化为有限个方程求和的方式做近似求解，如

$$\int_a^b f(x)\,\mathrm{d}x \approx \sum_{n=1}^{N} w_n f\left(x_n\right), \tag{7.128}$$

也即，我们在原方程的取值区间 $[a,b]$ 中抽取 N 个点，$n=0,1,\ldots,N$ 的 x_n 值代回方程中对应值 $f\left(x_n\right)$。w_n 表示相应的权重系数。

因此，数值积分也常常称为数值求积 (numerical quadrature)。集合 $\{x_n\},\{w_n\}$ 常常称为求积点 (quadrature points)集合和求积权重 (quadrature weights)集合；合适的选取 $\{x_n\},\{w_n\}$ 的算法称为求积法则 (quadrature rule)。对于同一个积分求解问题(7.128)往往存在一系列法则；评价不同法则之间好坏的标准在于，看哪个法则能用最少的样本点 N 来对(7.128)作出最精确的近似，同时确保计算成本可控，如编程难度、计算时间等。

先从最基本的牛顿——寇特斯法则开始介绍。

7.F.1　牛顿——寇特斯法则

近似求解积分问题(7.128)的基本思路是，在区间 $[a,b]$ 中找到一个多项式方程 $P(x)$ 来近似 $f(x)$。由于多项式的求和计算往往较为简单 (多项式的介绍见第7.B节)，这会简化(7.128)RHS 的计算时间。我们将这种思路称为牛顿——寇特斯 (求积) 法则 (Newton—Cotes Rule)。牛顿——寇特斯法则随着多项式的次 (degree, order) 而呈层级特征：使用 p 次多项式的牛顿——寇特斯法则可称为 p 阶牛顿——寇特斯法则。

如果 f 不是多项式，那么在 $[a,b]$ 区间内寻找近似多项式 $P(x)$ 会较为困难。一个近似方案是将空间 $[a,b]$ 划分为 N 个子区间 (对应 $N+1$ 个点)，在每个子空间中分别寻找可以近似 f 的多项式，进而将多个多项式加权求和，作为整个空间中 f 的近似方程。N 越大，子空间的个数越多，划分就越精细，每个子空间中的近似就越精确，进而整个空间中的近似求积就越精确。

矩形法则

我们先从 $p=0$ 的情况开始理解牛顿——寇特斯法则，根据黎曼求和 (Riemann sums)，我们可以将方程 $f(x)$ 理解为一组矩形的集合，对 $f(x)$ 在区间 $[a,b]$ 中求积就是求曲线下方的面积，近似等于各个矩形面积的和。随着 N 趋近于无穷大，每个矩形的宽度无限接近于 0，近似就越精确。

在每个子空间中分别用 1 个 $p=0$ 次多项式来近似 f，0 次多项式是个常数，换句话说我们在每个子空间中用一个常数来近似 f。这又称为矩形求积法则 (rectangular quadrature rule)。

图7.F.1绘出了 $N=4$ 情况下，利用矩形法则对 $\int_a^b f(x)\,\mathrm{d}x$ 的近似。不难看出，将区间 $[a,b]$ 划分为 N 个等宽子区间，每个子区间的宽度都是 $\Delta=(b-a)/N$。最左侧第一个矩形，左边长 (高) 为 $f(a)$，宽为 Δ，对应面积为 $f(a)\cdot\Delta$。左数第二个矩形，左边长 (高) 为 $f(a+\Delta)$，宽也是 Δ，面积为 $(a+\Delta)\cdot\Delta$，以此类推直到第 $N=4$ 个矩形为止。将这些子区间中矩形的面积加总，可得矩形法则的近似表达式

$$\mathcal{I}_{N=4}^{rect} \approx f(a)\cdot\Delta + f(a+\Delta)\cdot\Delta + f(a+2\Delta)\cdot\Delta + f(a+3\Delta)\cdot\Delta,$$

扩展到更一般的 $N\in\mathcal{N}$ 的情况，利用 N 段矩形法则对积分 $\int_a^b f(x)\,\mathrm{d}x$ 的近似为

$$\mathcal{I}_N^{rect} = \sum_{n=0}^{N-1} f(a+n\cdot\Delta)\cdot\Delta. \tag{7.129}$$

矩形法则(7.129)成为对(7.128)的近似求积法则之一：

图 7.F.1: 矩形法则 ($N = 4$)

- 求积权重集合 $\{w_n\}_{n=0}^{N-1}$ 对应 $w_n \equiv \Delta \, \forall n$,
- 求积点集合 $\{x_n\}_{n=0}^{N-1}$ 对应 $x_n = a + n \cdot \Delta$, $n = 0, 1, \ldots, N-1$。

值得注意的是，在矩形法则下，我们只需要左边长，从而无需计算 $f(b)$。

梯形法则

从图7.F.1中不难看出，利用矩形法则近似曲线下方阴影的面积，近似效果并不理想。一个可能的改进方案是利用梯形代替矩形做近似，又称梯形近似法则。

在每个子空间中分别用 1 个 $p = 1$ 次多项式来近似 f，1 次多项式是条斜线，连接子区间的两个端点 (始点和终点)。这又称为梯形求积法则 (trapezoidal quadrature rule)。

图 7.F.2: 梯形法则 ($N = 4$)

见图7.F.2所示，左数第 1 个梯形的面积等于

$$\frac{1}{2}\left[f(a) + f(a+\Delta)\right] \cdot \Delta,$$

左数第 2 个梯形的面积等于

$$\frac{1}{2}\left[f(a+\Delta) + f(a+2\Delta)\right] \cdot \Delta,$$

以此类推直到第 $N=4$ 个梯形。将子区间中梯形的面积加总，可得

$$\begin{aligned}\mathcal{I}_{N=4}^{trap} =& \frac{1}{2}\left[f(a)+f(a+\Delta)\right]\cdot\Delta+\frac{1}{2}\left[f(a+\Delta)+f(a+2\Delta)\right]\cdot\Delta \\ &+\frac{1}{2}\left[f(a+2\Delta)+f(a+3\Delta)\right]\cdot\Delta+\frac{1}{2}\left[f(a+3\Delta)+f(a+4\Delta)\right]\cdot\Delta \\ =& \left[\frac{1}{2}f(a)+f(a+\Delta)+f(a+2\Delta)+f(a+3\Delta)+\frac{1}{2}f(a+b)\right]\cdot\Delta,\end{aligned}$$

扩展到更一般的 $N\in\mathcal{N}$ 的情况，我们有 N 段梯形法则对 $\int_a^b f(x)\,\mathrm{d}x$ 的近似为

$$\mathcal{I}_N^{trap}=\frac{1}{2}f(a)\cdot\Delta+\Delta\cdot\sum_{n=1}^{N-1}f(a+n\cdot\Delta)+\frac{1}{2}f(b)\cdot\Delta. \tag{7.130}$$

梯形法则(7.130)成为对(7.128)的又一种近似求积法则：

- 求积点集合 $\{x_n\}_{n=1}^{N-1}$ 对应 $x_n=a+n\cdot\Delta$, $n=0,1,\ldots,N$,
- 求积权重集合 $\{w_n\}_{n=1}^{N}$ 对应

$$w_n=\begin{cases}\frac{1}{2}\Delta, & n=0,N,\\ \Delta, & n=1,2,\ldots,N-1,\end{cases}$$

  值得注意的是，比起矩形法则来，当利用梯形法则做近似求积时，需计算 $N+1$ 个求积抽样点，比矩形法则多出的 1 个点为总区间中的末端 b，对应 $f(b)$。

更高阶牛顿——寇特斯法则

已知用 $p=0$ 阶多项式 (常数) 近似子区间中的 f，称矩形法则。用 $p=1$ 阶多项式 (斜线) 近似，称梯形法则。那么我们可以进一步迭代计算更高阶 $p>1$ 的牛顿——寇特斯求积近似，方法如下

1. 在子区间 $[x_n,x_{n+1}]$ 中，配 $p+1$ 个等距离分布的点 (包括子区间的起始点 x_n 和终结点 x_{n+1})，将子区间分为 p 个部分。例如，
   - 对 $p=2$ 而言，配 3 个点 $x_n,\frac{1}{2}\cdot[x_n+x_{n+1}],x_{n+1}$,
   - 对 $p=3$ 而言，配 4 个点 $x_n,\frac{1}{3}\cdot[x_n+x_{n+1}],\frac{2}{3}\cdot[x_n+x_{n+1}],x_{n+1}$，以此类推。
2. 在子空间中针对这 $p+1$ 个配点，找到唯一的多项式 $P(x)$ 以近似 $f(x)$。例如，
   - $p=0$，$P(x)$ 是个常数。矩形法则。
   - $p=1$，$P(x)$ 是条斜线，连接子区间的两个端点 $f(x_n)$, $f(x_{n+1})$。梯形法则。
   - $p=2$，$P(x)$ 是一条经过如下 3 个点的抛物线，3 个点的坐标依次为

$$\begin{cases}\Big(x_n,f(x_n)\Big),\\ \Big(x_m,f(x_m)\Big), & x_m\equiv\frac{x_n+x_{n+1}}{2},\\ \Big(x_{n+1},f(x_{n+1})\Big).\end{cases}$$

- $p=3$，$P(x)$ 是一条经过如下 4 个点的曲线 (图7.F.3)，4 个点的坐标依次为

$$\begin{cases} \Big(x_n, f\left(x_n\right)\Big), \\ \Big(x_m, f\left(x_m\right)\Big), \quad x_m \equiv \frac{1}{3}\left(x_n + x_{n+1}\right), \\ \Big(x_{m+1}, f\left(x_{m+1}\right)\Big), \quad x_{m+1} \equiv \frac{2}{3}\left(x_n + x_{n+1}\right), \\ \Big(x_{n+1}, f\left(x_{n+1}\right)\Big), \end{cases}$$

以此类推。无论哪一个例子中，$P(x)$ 都是这样一个多项式，其系数是抽样方程 $f(x)$ 值

图 7.F.3: 3 阶牛顿——寇特斯法则，对应子区间中的 4 个配点

的线性组合。

3. 从 x_n 到 x_{n+1} 对 $P(x)$ 求积分，作为这个子区间内方程 $f(x)$ 的近似。

4. 将所有子区间内的近似方程组合起来，作为总区间中的最终近似求积法则。

$p=2$ 时的牛顿——寇特斯求积，又称辛普森法则 (Simpson's rule)

$$\mathcal{I}_N^{simp} = \sum_{n=0}^{N-1} \frac{\Delta}{6}\left[f\left(a+n\cdot\Delta\right) + 4f\left(a+\left(n+\frac{1}{2}\right)\cdot\Delta\right) + f\left(a+(n+1)\right)\cdot\Delta\right]. \tag{7.131}$$

龙格现象

一个自然出现的问题：更高阶的牛顿——寇特斯法则会不会带来更精确的近似解？很遗憾，答案是否定的。这是由于龙格现象 (Runge phenomenon)：当我们试图利用子空间中等距分布的点做更高阶多项式近似时，再查指点附近的多项式方程会出现较大幅度的震荡，从而影响最终求积估计的精确度，可参考维基百科词条。图7.F.4中，蓝线和绿线分别表示 $p=5$ 和 $p=9$ 时的差值多项式近似，对应 $N=6$。红线表示龙格方程 (Runge function)。

误差的收敛

比较不同求积法则的优劣，可用启发式误差“分析”[27]，观察近似误差随着配点数 N 增加的衰减速度。具体说来

[27]加上引号是指，这种方法只是一种较为常用的检验措施，而不宜理解为某种严谨的学术性系统性表达。

图 7.F.4: 龙格现象 ($N = 6$)

1. 考虑区间中某一宽度为 Δ 的特定子区间，在其中用一个 p 阶多项式 $P(x)$ 来近似 f

$$P(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_p x^p.$$

2. 那么子区间中存在一个点 x_0，使得 f 围绕 x_0 做泰勒级数展开的前 $p+1$ 项，与 $P(x)$ 一致[28]。可将 f 写成如下形式

$$f(x) = \underbrace{c_0 + c_1 x + \ldots + c_p x^p}_{\equiv P(x)} + c_{p+1} x^{p+1} + \ldots.$$

3. 由此可见，原方程 $f(x)$ 及其近似多项式 $P(x)$ 之间的误差是一个从 $p+1$ 阶开始的多项式

$$f(x) - P(x) = c_{p+1} x^{p+1} + c_{p+2} x^{p+2} + \ldots,$$

4. 那么在这个子区间 $[x_n, x_{n+1}]$ 内的全部误差可求积得出

$$\begin{aligned}\int [f(x) - P(x)]\,\mathrm{d}x &= \int \left[c_{p+1} x^{p+1} + c_{p+2} x^{p+2} + \ldots\right]\,\mathrm{d}x \\ &\propto \Delta^{p+2} + \text{更高阶项},\end{aligned}$$

   最后一行是说，对 x^{p+1} 沿着某个宽度为 Δ 的子区间求积，等于某个和 Δ^{p+2} 成比例的值。

5. 将一个子区间中的情况扩展到其他子区间，每个子区间中的误差项都与 Δ^{p+2} 成定比例。更进一步地，由于 Δ 与 N 反比例相关 (N 是整个区间中划分的子区间的数量)

$$\Delta \sim \frac{1}{N},$$

[28] 例如

- $p = 0$，矩形法则，x_0 可以是子区间的起始点。
- $p = 1$，梯形法则，x_0 的存在性可由中值定理予以证明：x_n 和 x_{n+1} 之间必然存在一点，该点上的方程值 $f(x_0)$ 的导数，等于连接 $f(x_n)$ 和 $f(x_{n+1})$ 两点的斜线的斜率。
- $p > 1$，更高阶求和法则中 x_0 点存在性的证明，也可用类似思路证得。

那么每个子区间中的误差项可近似表示为

$$\Delta \propto \frac{1}{N^{p+2}}.$$

6. 将 N 个子区间的误差项汇总

$$\varepsilon \propto \frac{N}{N^{p+2}} = \frac{1}{N^{p+2}}, \tag{7.132}$$

例如矩形法则 $(p=0)$ 的近似误差 $\propto \frac{1}{N}$，梯形法则 $(p=1)$ 的近似误差 $\propto \frac{1}{N^2}$。可见误差与 N 成反比：N 越大，划分的子区间数量越多，近似误差越小。

举例，假定我们要用求积法则近似计算

$$\mathcal{I} = \int_1^2 \log^2 x \, \mathrm{d}x,$$

对应相对误差定义为 ε^{rel}

$$\varepsilon^{rel} \equiv \frac{\left|\mathcal{I}_N^{approx} - \mathcal{I}\right|}{\mathcal{I}}$$

应用矩形法则和梯形法则，查看近似误差随着 N 增大的收敛情况如图7.F.5，可以看出

图 7.F.5: 比较矩形法则和梯形法则下，近似误差随 N 值增大的收敛情况

原求积方程 $\int_1^2 \log^2 x \, \mathrm{d}x$。

这两种方法代表的牛顿——寇特斯法则下误差项随着 N 而作线性收敛：随着 N 增大 1,000 倍，矩形法则误差降低约 1,000 倍，$\Delta\varepsilon^{rect} \propto \frac{1}{\Delta N}$；梯形法则降低约 1,000,000 倍，$\Delta\varepsilon^{trap} \propto \frac{1}{(\Delta N)^2}$。梯度法则优于矩形法则。

但图7.F.5也揭示出牛顿——寇特斯法则的不足：即便是对于 $\log^2 x$ 这种很平滑的方程，梯形法则也需要大约 1,000 个样本，才能将总体误差控制在 10^{-6} 的水平上 (为了达到类似的精度，矩形法则甚至需要 10^6 个样本)。从计算成本来看，牛顿——寇特斯法则恐怕不是最理想的方案，需要加以改进。如下文介绍的克伦肖——柯蒂斯法则，可以用少得多的样本量达到同样 10^{-6} 级别的误差水平，哪怕所处理的原方程更加复杂。我们做了一个小的数值计算实验来说明这个问题，见表7.G.1。

7.F.2 几个小技巧

以牛顿——寇特斯法则为例，介绍几个小技巧。它们对其他如克伦肖——柯蒂斯法则也适用。

定区间单位的换算

对于方程 $\int_a^b f(t)\,\mathrm{d}t$ 的求积问题，常常需要将区间从 $[a,b]$ 转换为 $[-1,1]$。可以设一个新的方程 $g(u), u\in[-1,1]$，满足

$$g(u)=f\left[a+\frac{b-a}{2}(u+1)\right],\quad u\in[-1,1], \tag{7.133}$$

那么

$$\int_a^b f(t)\,\mathrm{d}t \Leftrightarrow \int_{-1}^1 g(u)\,\mathrm{d}u.$$

无限区间求积

有时会遇到求含有无限区间的积分问题如

$$\int_0^\infty f(x)\,\mathrm{d}x,$$

这需要我们先将区间 $[0,\infty)$ 映射到一个有界区间内，如

$$x:[0,\infty)\mapsto u:[0,1],$$

然后再应用求解定积分的求积法则。映射的方法有很多种，其中之一是设

$$x\equiv\frac{u}{1-u},\Rightarrow \mathrm{d}x=\frac{\mathrm{d}u}{(1-u)^2},$$

求积问题因此转换为

$$\int_0^\infty f(x)\,\mathrm{d}x=\int_0^1\frac{1}{(1-u)^2}f\left(\frac{u}{1-u}\right)\mathrm{d}u.$$

需要指出的是，上式在转换之后，含有一个奇异点 (singularity):$\lim_{u\to1}f(x)\to\infty$。然而若是在前提假设中假定 $f(x)$ 在 $x\to\infty$ 处消失 (这个假定是为了确保积分 $\int_0^\infty f(x)\,\mathrm{d}x$ 收敛)，则这一奇异点也就不存在了。

7.F.3　求积中的可积奇异点问题

除了上节提到的情况之外，还应该注意到，有些奇异点是可积的 (integrable singular points)，例如这个求积问题

$$\mathcal{I}=\int_0^1\frac{\exp(x)}{\sqrt{x}}\,\mathrm{d}x, \tag{7.134}$$

方程在区间 $(0,1]$ 内定义良好，除了起始点 0 的情况，称之为可积奇异点[29]。存在可积奇异点的积分方程并不是全都不能求解，针对具体问题的不同，有一些求解技巧可供选择。试举例如下。

提取奇异点

有时可以在积分中把奇异点单独提取出来做求积运算，如 (7.134)可以改写为

$$\mathcal{I}=\underbrace{\int_0^1\frac{1}{\sqrt{x}}\,\mathrm{d}x}_{=:\mathcal{I}_1}+\underbrace{\int_0^1\frac{\exp(x)-1}{\sqrt{x}}\,\mathrm{d}x}_{=:\mathcal{I}_2}.$$

[29]注意区分可积和不可积奇异点。后者如积分方程 $\int_0^1\frac{\exp(x)}{x}\,\mathrm{d}x$ 的起始点 0 的情况，此时积分不存在，因此无法估计。

其中 $\mathcal{I}_1$ 可以用解析法求得

$$\mathcal{I}_1 = \int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = \left|2\sqrt{x}\right|_0^1 = 2.$$

含有可积奇异点的 $\mathcal{I}_2$ 无法用解析法求解，但在区间起始点 $(=0)$ 处是非奇异的：这是由于，将被积方程沿着 $x \to 0$ 做泰勒级数展开，可得

$$\begin{aligned}
&\frac{\exp(x)-1}{\sqrt{x}} \approx \frac{x - \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots}{\sqrt{x}} = x^{\frac{1}{2}} - \frac{1}{2}x^{\frac{3}{2}} + \frac{1}{6}x^{\frac{5}{2}} + \dots, \\
&\hookrightarrow \lim_{x\to 0} \frac{\exp(x)-1}{\sqrt{x}} \to 0,
\end{aligned}$$

因此可用常规求积法则进一步近似求解 $\mathcal{I}_2$。

奇异点消除

若被积方程的分母中含有可积奇异点，那么可以考虑用雅各比变换 (Jacobian transformation)。例如对(7.134)，设

$$u \equiv \sqrt{x} \Rightarrow \mathrm{d}u = \frac{\mathrm{d}x}{2\sqrt{x}},$$

求积运算因此变为

$$\int_0^1 \frac{\exp(x)}{\sqrt{x}}\,\mathrm{d}x = 2\int_0^1 \exp\left(u^2\right)\,\mathrm{d}u,$$

新的积分方程中没有奇异点，因此可进一步用常规数值求积法则来近似。

Epsilon 扩展

如果可积奇异点无法利用前述两种方法消除，我们可以尝试加入一个平滑参数 ϵ。对应地，取极限 $\lim \epsilon \to 0$ 以消除奇异点。如将(7.134)改写为

$$\mathcal{I}_\epsilon = \int_0^1 \frac{\exp(x)}{\left(x^2+\epsilon^2\right)^{\frac{1}{4}}}\,\mathrm{d}x.$$

对于有限值的 $\epsilon < \infty$，上式非奇异，并且 $\lim_{\epsilon\to 0}\mathcal{I}_\epsilon \to \mathcal{I}$。

适应性求积

通常来说，被积方程在不同值域段的表现不同，如图 7.F.6所示。如果我们应用梯形法则求解整个区间 $[0,10]$ 内的积分，那么在子区间 $[4,6]$ 中可能需要将宽度 Δ 设的小一些，在 $[0,4]$ 和 $[6,10]$ 中将 Δ 设的大一些。这就产生了适应性求积 (adaptive quadrature)的概念：随着子区间中方程变化的程度不同，应用不同的求积法则来做近似，以求近似精度和计算成本的平衡。举例来说，此时的梯形法则变为

7.F.4 梯形法则的近似误差

如前文所述，牛顿——寇特斯求积法则 (第7.F.1节) 的近似误差随着分段数 N 的衰减速度并不够令人满意；若是尝试利用更高阶多项式被积方程做近似，还会受到龙格现象的干扰。因此在实际应用中，常常并不会直接使用牛顿——柯蒂斯法则。一个更好的方案是克伦肖——柯蒂斯求积法则 (Clenshaw-Curtis Rule)。在了解傅里叶分析 (第7.G节) 的基本要点之后，我们可以对这个法则有大致介绍。首先来看梯形法则的近似误差分析。

图 7.F.6: 适应性求积

现在假定我们的目标是求解积分 $\mathcal{I}=\int_0^T f(t)\,\mathrm{d}t$。根据梯形法则(7.130)可以做 N 个点的近似

$$I_N^{trap}=\frac{T}{N}\left\{\underbrace{\left[\frac{1}{2}f(0)+\frac{1}{2}f(T)\right]}_{=:\mathcal{A}}+\underbrace{\sum_{n=1}^{N-1}f\left(\frac{nT}{N}\right)}_{=:\mathcal{B}}\right\}, \tag{7.135}$$

近似的精度与 N 有关：N 越大，划分的点越多，近似越精确。我们将精度定义为误差项 ε_N^{trap}

$$\varepsilon_N^{trap}=\left|\mathcal{I}-I_N^{trap}\right|.$$

现在引入傅里叶分析的概念。在实区间 $[0,T]$ 中的方程 $f(t)$，可写为傅里叶空间中的综合表达式

$$f(t)=\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu\exp\left(i\nu\omega_0 t\right),\quad \omega_0\equiv\frac{2\pi}{T}, \tag{7.136}$$

其中傅里叶级数的系数 $\tilde{f}_\nu$ 值为

$$\tilde{f}_\nu=\frac{1}{T}\int_0^T f(t)\exp\left(-i\nu\omega_0 t\right)\mathrm{d}t. \tag{7.137}$$

对 $\mathcal{I}$ 做傅里叶变换可得

$$\begin{aligned}\mathcal{I}&=\int_0^T f(t)\,\mathrm{d}t=\int_0^T\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu\exp\left(i\nu\omega_0 t\right)\,\mathrm{d}t\\&=\int_0^T\sum_{\nu=-\infty}^{\infty}\frac{1}{T}\int_0^T f(t)\exp\left(-i\nu\omega_0 t\right)\,\mathrm{d}t\exp\left(i\nu\omega_0 t\right)\,\mathrm{d}t\\&=T\tilde{f}_0,\end{aligned}$$

即积分的值等于 T 个 $\nu=0$ 时的系数 $\tilde{f}_{\nu=0}$ 值之和。

需要强调的时，在对 $f(t)$ 利用梯形法则做牛顿——寇特斯求积时，我们并不知道系数 $\left\{\tilde{f}_\nu\right\}_\nu$ 的值——若是知道，也就不必再做傅里叶转换了。我们知道的是，存在一个关于被积方程 $f(t)$ 的傅里叶综合表达形式(7.136)，可供我们展开进一步的分析。

现在的任务是要将傅里叶综合式(7.136)代回梯形法则(7.135)。(7.135)RHS 分为 $\mathcal{A}$ 和 $\mathcal{B}$ 两部分，

- $\mathcal{A}$ 等于 $t=0$ 时对(7.136)的计算：
$$\mathcal{A}=\sum_{m=-\infty}^{\infty}\tilde{f}_m,$$
分两种情况予以说明
  - 当 $f(t)$ 正好是一个 T-周期方程时，$f(0)=f(T)$，显然这是正确的。
  - 当 $f(t)$ 的周期不等于 T 时，$\mathcal{A}$ 就等于设 $t=0$ 时，对(7.136)的近似估计，见第7.G.5节。
- $\mathcal{B}$ 的求和值，可以用一系列不同的 $t=\frac{nT}{N}$ 来替代。
$$\mathcal{B}=\sum_{n=1}^{N-1}\sum_{m=-\infty}^{\infty}\tilde{f}_m\exp\left(im\omega_0\frac{nT}{N}\right).$$
因此我们有
$$\begin{aligned}I_N^{trap}&=\frac{T}{N}\{\mathcal{A}+\mathcal{B}\}=\frac{T}{N}\left\{\sum_{m=-\infty}^{\infty}\tilde{f}_m+\sum_{n=1}^{N-1}\sum_{m=-\infty}^{\infty}\tilde{f}_m\exp\left(im\omega_0\frac{nT}{N}\right)\right\}\\&=\frac{T}{N}\sum_{n=0}^{N-1}\left\{\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu\exp\left(i\nu\omega_0\frac{nT}{N}\right)\right\}=\frac{T}{N}\sum_{n=0}^{N-1}\left\{\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu\exp\left(2\pi\frac{i\nu n}{N}\right)\right\},\end{aligned}\tag{7.138}$$
上式中引入一个新定义方程 $K_N(\nu)$ 做进一步简化
$$I_N^{trap}=T\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu K_N(\nu),\tag{7.139}$$
$$\begin{aligned}K_N(\nu)&\equiv\frac{1}{N}\sum_{n=0}^{N-1}\exp\left(2\pi\frac{i\nu n}{N}\right)\\&=\frac{1}{N}\left[1+\xi+\xi^2+\ldots+\xi^{N-1}\right],\quad \xi\equiv\exp\left(2\pi\frac{i\nu}{N}\right),\end{aligned}\tag{7.140}$$
$K_N(\nu)$ 的取值分为两种情况
- $\nu=0$，或者 ν 是一个 N 的整数倍是，有 $\xi=1\Rightarrow K_N(\nu)=1$，
- $\nu\neq0$ 且 ν 不是 N 的整数倍时，有 $\xi\neq1$，以及
$$K_N(\nu)\approx\frac{1}{N}\left[\frac{1-\xi^N}{1-\xi}\right]=\frac{1}{N}\frac{1-\exp(2\pi i\nu)}{1-\exp\left(\frac{2\pi i\nu}{N}\right)}=0.$$
由此(7.139)进一步改写为
$$I_N^{trap}=T\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu K_N(\nu)=T\sum_{p=-\infty}^{\infty}\tilde{f}_{pN},\quad p\in\mathcal{Z}.$$
可见对 $[0,T]$ 区间中对周期方程求积问题 $\int f(t)\,\mathrm{d}t$ 做 N 个点的梯形法则近似，其值与傅里叶级数的系数 $\left\{\tilde{f}_\nu\right\}_\nu$ 有关，但并不是全部系数，而仅仅是其中的一部分系数之和，具体说来是那些 N 的整数倍所对应的系数 $\sum_p\tilde{f}_{pN}$，它们在傅里叶域中对应的傅里叶系数，是指频率为 $N\omega_0$ 整数倍的若干条弦波曲线的傅里叶系数。尤其是，其中只有 $\nu=0$ 那条对应的项才是我们真正需要的；其余的部分都是误差项
$$I_N^{trap}=\underbrace{T\tilde{f}_{\nu=0}}_{\mathcal{I}}+\underbrace{T\sum_{p\neq0}\tilde{f}_{pN}}_{\varepsilon_N^{trap}},\Leftrightarrow\varepsilon_N^{trap}\equiv\left|\mathcal{I}-I_N^{trap}\right|,$$

ε_N^{trap} 从而可以理解为第 $\pm N$ 个，第 $\pm 2N$ 个，... 个傅里叶级数的系数之和

$$\varepsilon_N^{trap} \equiv \left|\mathcal{I} - I_N^{trap}\right| = \left|\sum_{p=-\infty, p\neq 0}^{\infty} \tilde{f}_{pN}\right|. \tag{7.141}$$

非周期方程的傅里叶级数收敛

同样地，分析即使进展到(7.141)这一步，我们还是不知道傅里叶级数系数 $\left\{\tilde{f}_{pN}\right\}$ 的值，因而无法直接数值计算做近似求积。不过我们可以用傅里叶分析的平滑——衰减特性 (第7.G.4节)，来观察误差随着 N 的增大而衰减的速度有多快 (收敛速度)。

随着被积方程 $f(t)$ 是周期方程还是非周期方程，我们分两种情况来分析。首先考虑 $f(t)$ 是平滑方程，但不满足 $f(0) = f(T)$ 的情况。此时我们利用(7.137)计算的傅里叶级数系数 $\tilde{f}_\nu$ 是一个不连续方程。

为了在这种情况下做收敛分析，一个较原始的方案是[30]：将 $[0, T]$ 中的这一段 $f(t)$ 截取出来，以 T 周期为单位向前向后做重复扩展，所形成的一个新周期方程定义为 $f^{per}(t)$，作为 $f(t)$ 的 T-周期扩展，见图7.F.7。 reference

由佩利——维纳诸定理 (Theorem 7.62) 可见，不连续方程 $f(t)$ 的傅里叶系数的强度 (magnitude) $\tilde{f}_\nu$，以等比于 $\frac{1}{\nu}$ 的速度衰减。那么在理想状况下，根据(7.141)，梯形法则的近似误差 ε_N^{trap} 也应该以等比于 $\frac{1}{N}$ 的速度衰减。然而实际上(7.141)中 $\nu \in \mathcal{Z}$ 可以是正值也可以是负值，可能导致求和后的 $\tilde{f}_\nu$ 被抵消掉。为了解决这个问题，我们在傅里叶系数求和式(7.137)中区分 $\nu > 0$ 和 $\nu < 0$ 的情况，作分部积分 (Boyd, 2001, Sec 2.9)：

$$\begin{aligned}
\tilde{f}_\nu &= \frac{1}{T}\int_0^T f(t)\exp\left(-i\nu\omega_0 t\right)\mathrm{d}t \\
&= \frac{1}{T}\left\{\frac{1}{i\nu\omega_0}\underbrace{\left|f(t)\exp\left(-i\nu\omega_0 t\right)\right|_0^T}_{=:\mathcal{A}} + \frac{1}{i\nu\omega_0}\int_0^T f^{'}(t)\exp\left(-i\nu\omega_0 t\right)\,\mathrm{d}t\right\},
\end{aligned}$$

其中

$$\mathcal{A} = \left|\cdot\right|_0^T = f(T)\underbrace{\exp\left(-i\nu\omega_0 T\right)}_{=1} - f(0)\underbrace{\exp\left(-i\nu\omega_0 0\right)}_{=1} = f(T) - f(0) \equiv \Delta f.$$

那么我们有

$$\tilde{f}_\nu = \underbrace{\frac{\Delta f}{i\nu\omega_0 T}}_{=:\frac{C_1}{\nu}} + \frac{1}{i\nu\omega_0}\int_0^T f^{'}(t)\exp\left(-i\nu\omega_0 t\right)\,\mathrm{d}t.$$

对上式再次分部积分，采用类似的思路，有

$$\tilde{f}_\nu = \underbrace{\frac{\Delta f}{i\nu\omega_0 T}}_{=:\frac{C_1}{\nu}} + \underbrace{\frac{\Delta f^{'}}{\left(i\nu\omega_0\right)^2}}_{=:\frac{C_2}{\nu^2}} + \frac{1}{\left(i\nu\omega_0\right)^2}\int_0^T f^{''}(t)\exp\left(-i\nu\omega_0 t\right)\,\mathrm{d}t,$$

其中 $\Delta f^{'} \equiv f^{'}(T) - f^{'}(t)$。

[30]注意这里讨论的 T-周期扩展方案较为直观易于理解，但精度上有所欠缺。在下文中我们将对 $f(t)$ 做更精致的周期扩展见第7.F.5节。两种方案的比较见第182页脚注。

图 7.F.7: 非周期方程的 T-周期扩展 $t \in [0,1]$

(a)x

(b)x

继续进行下去可以看到，傅里叶系数 $\tilde{f}_\nu$ 是一与 ν 的幂次的倒数之和 [31]

$$\tilde{f}_\nu = \frac{C_1}{\nu} + \frac{C_2}{\nu^2} + \frac{C_3}{\nu^3} + \frac{C_4}{\nu^4} \dots \tag{7.142}$$

这样一来，若是取与正值 ν 相对应的负值 $-\nu$ 两两配对，将 $\tilde{f}_\nu$ 和 $\tilde{f}_{-\nu}$ 相加，则 RHS 中奇数项相互抵消，偶数项部分保留下来

$$\tilde{f}_\nu + \tilde{f}_{-\nu} = \frac{C_2^{'}}{\nu^2} + \frac{C_4^{'}}{\nu^4} + \dots \tag{7.143}$$

其中 $C_2^{'} \equiv 2C_2$，以此类推 (常数项不重要，重要的是分母随着 ν 值变化而表现出的渐进特性)。

将(7.143)代回(7.141)可得

$$\begin{aligned}\varepsilon_N^{trap} &= \left| \sum_{p=-\infty, p\neq 0}^{\infty} \tilde{f}_{pN} \right| = \left| \sum_{p=1}^{\infty} \left(\tilde{f}_{pN} + \tilde{f}_{-pN} \right) \right| \\ &= \frac{C_2^{'}}{N^2} \sum_{p=1}^{\infty} \frac{1}{p^2} = \frac{C_2^{'}}{N^2} \frac{\pi^2}{6},\end{aligned}$$

忽略掉常数项可得，梯形法则下对非周期方程 $f(t)$ 求积的近似误差 $\varepsilon_N^{trap} \sim \frac{1}{N^2}$，即等比收敛至 $\frac{1}{N^2}$。

周期方程的傅里叶级数收敛

来考虑 $f(t) \in C^\infty$ 是平滑的 T-周期方程的情况。此时

$$f(0) = f(T),\ f^{'}(0) = f^{'}(T),\ f^{''}(0) = f^{''}(T), \dots$$

根据佩利——维纳诸定理 (Theorem 7.62)，这类方程在傅里叶域中，傅里叶系数的强度 (magtinute) $\left|\tilde{f}_\nu\right|$ 的衰减速度比任何多项式在 ν 的衰减速度都快，例如 $\left|\tilde{f}_N\right| \sim \exp\left(-\alpha\left|N\right|\right)$。此时误差项可以写为

$$\varepsilon_N^{trap} = \left| \sum_{p\neq 0} \tilde{f}_{pN} \right| = \sum_{p\neq 0} \exp\left(-\alpha N\left|p\right|\right) \sim \exp\left(-\alpha N\right),$$

即周期方程 $f(t)$ 的梯形法则近似误差项等比例收敛至 $\exp\left(-\alpha N\right)$。

7.F.5　克伦肖——柯蒂斯求积法则

上节分析了为什么梯形法则下的傅里叶级数近似误差可以很快速的指数收敛 (如果原方程是周期方程)，或是以较慢的速度收敛 (如果原方程是非周期方程)。对于非周期方程的情况，能不能让近似误差收敛的速度加快？克伦肖——柯蒂斯求积法则 (Clenshaw-Curtis Rule)的作用就在于，将非周期方程转换为周期方程的形式，从而加快梯形法则误差项的收敛速度。大致说来分为三个步骤

- 将非周期方程转换为周期方程，
- 将求积式改写为傅里叶系数和，
- 应用求积法则做近似，

以下分别作以说明。

[31] 巧合的是，这恰好成为佩利——维纳诸定理 (Theorem 7.62) 的一种证明：如果 $\Delta f = 0$ 但 $\Delta f^{'} \neq 0$，那么对 $f(t)$ 的 T-周期扩展中就同事含有连续的方程和不连续的倒数，那么其对应的不连续点就会出现在 $p = 1$ 处；进而由上面的扩展可得，随着 $|\nu| \to \infty$，有 $\left|\tilde{f}_\nu\right| \sim \frac{1}{\nu^2}$。

将非周期方程 $f(t)$ 转换为周期方程 $g(t)$

假定我们要在 $[-1,1]$ 区间中数值求积

$$\mathcal{I} = \int_{-1}^{1} f(t)\,\mathrm{d}t, \tag{7.144}$$

如果原空间是 $[a,b]$ 而不是 $[-1,1]$，可以参考第7.F.2的转换方法。

$[-1,1]$ 区间恰好是 $\cos\theta$, $\theta \in [0,\pi]$ 的取值范围，因此可做参数化处理，设 $t = \cos\theta$，定义新方程如图7.F.8所示。

$$g(\theta) \equiv f(\cos\theta)\,. \tag{7.145}$$

图 7.F.8: $f(t)$ 和 $g(\theta)$

$g(\theta)$ 有如下特征：

- 是个周期方程，周期 $T = 2\pi$，
- 是个偶方程，满足 $g(-\theta) = g(\theta)$，

- $g(\theta)$ 随着 θ 从 0 到 π 的路径，勾勒出 $f(t)$ 随着 t 从 1 到 -1 的变化路径，
- $f(t)$ 在实域 $(-\infty,1)$ 和 $(1,\infty)$ 中的变化路径，无法由 $g(\theta)$ 表现出来。

$g(\theta)$ 的傅里叶级数表现形式包含一组弦波之和，每一个弦波的频率都是基准频率 $\omega_0=\frac{2\pi}{T}=1$ 的整数倍

$$\begin{aligned} g(\theta) &= \sum_{\nu=-\infty}^{\infty} \tilde{g}_\nu \exp\left(i\nu\theta\right), \\ \tilde{g}_\nu &= \frac{1}{2\pi}\int_0^{2\pi} g(\theta)\exp\left(-i\nu\theta\right)\,\mathrm{d}\theta. \end{aligned} \tag{7.146}$$

或者 $g(\theta)$ 也可以用傅里叶余弦级数的形式表示为

$$g(\theta)=\frac{\tilde{a}_0}{2}+\sum_{\nu=1}^{\infty}\tilde{a}_\nu\cos\left(\nu\theta\right). \tag{7.147}$$

两组傅里叶系数 $\{\tilde{a}_\nu\}$，$\{\tilde{g}_\nu\}$ 之间的关系为

$$\tilde{a}_0=2\tilde{g}_0,\quad \tilde{a}_\nu=2\left(\tilde{g}_\nu+\tilde{g}_{-\nu}\right)=2\tilde{g}_\nu, \tag{7.148}$$

这是由于 $g(\theta)=g(-\theta)\Rightarrow\tilde{g}_\nu=\tilde{g}_{-\theta}$。

此外 $\{\tilde{a}_\nu\}$ 可以表示为如下形式

$$\tilde{a}_\nu=\frac{1}{\pi}\int_0^{2\pi} g(\theta)\cos\left(\nu\theta\right)\,\mathrm{d}\theta. \tag{7.149}$$

由(7.149)可见，每个系数 $\tilde{a}_\nu$ 都是一个被积方程 $g(\theta)\cos\left(\nu\theta\right)$ 关于其域 $\theta\in[0,2\pi]$ 的积分。由于在这个域中，$g(\theta)$ 和 $\cos\left(\nu\theta\right)$ 都是偶方程，那么 $g(\theta)\cos\left(\nu\theta\right)$ 也是一个偶方程。这意味着积分(7.149)可由一个简单的 N 个点的梯形法则来做近似求解，并且近似误差随着 N 值的增大而呈指数速度衰减。

原域求积问题改写为傅里叶域求和问题

目标：将关于 $f(t)$ 的求积问题(7.144)改写为关于 $g(\theta)$ 的求积问题，用 $t=\cos\theta$ 替换 t 可得

$$\mathcal{I}=\int_{-1}^{1} f(t)\,\mathrm{d}t=\int_0^{\pi}\underbrace{f\left(\cos\theta\right)}_{=g(\theta)}\underbrace{\mathrm{d}\cos\theta}_{=\sin\theta\mathrm{d}\theta}=\int_0^{\pi} g(\theta)\sin\theta\,\mathrm{d}\theta. \tag{7.150}$$

尽管被积方程 $g(\theta)\sin\theta$ 时周期方程，但积分区间 $[0,\pi]$ 只是完整周期 $[0,2\pi]$ 的一般，因此我们不能直接对(7.150)应用梯形法则来做 (指数衰减的) 求积。为了解决这个问题，我们将 $g(\theta)$ 的傅里叶余弦级数表达式(7.147) 代入(7.150)有

$$\begin{aligned} \mathcal{I} &= \int_0^{\pi} g(\theta)\sin\theta\,\mathrm{d}\theta=\int_0^{\pi}\left[\frac{\tilde{a}_0}{2}+\sum_{\nu=1}^{\infty}\tilde{a}_\nu\cos\left(\nu\theta\right)\right]\sin\theta\,\mathrm{d}\theta \\ &= \frac{\tilde{a}_0}{2}\underbrace{\int_0^{\pi}\sin\theta\,\mathrm{d}\theta}_{=2}+\sum_{\nu=1}^{\infty}\tilde{a}_\nu\underbrace{\int_0^{\pi}\sin\theta\cos\left(\nu\theta\right)\,\mathrm{d}\theta}_{=\frac{1+\cos(\nu\pi)}{1-\nu^2}} \\ &= \tilde{a}_0+\sum_{\nu=1}^{\infty}\tilde{a}_\nu\frac{1+\cos\left(\nu\pi\right)}{1-\nu^2}, \end{aligned} \tag{7.151}$$

其中

$$\cos(\nu\pi) = (-1)^{\nu} = \begin{cases} -1 & \nu\text{是奇数} \\ 1 & \nu\text{是偶数} \end{cases}$$

代回上式得

$$\mathcal{I} = \tilde{a}_0 + \sum_{\nu=1,\nu\text{是偶数}}^{\infty} \frac{2\tilde{a}_\nu}{1-\nu^2}, \tag{7.152}$$

也即用 $g(\theta)$ 的傅里叶余弦系数 $\{\tilde{a}_\nu\}$ 来描述 $f(t)$，进而计算求积问题。如果假定 $f(t)$ 平滑，那么 $g(\theta)$ 继承这一特性，是个平滑的周期方程，其傅里叶余弦系数 $\tilde{a}_\nu$ 随着 ν 值而衰减的速度比任何多项式都快[32]。

在实际数值计算过程中，很难考虑到 $\nu \to \infty$ 的情况。对此，常常要对(7.152)中的求和式做一个 ν 等于某个有限值比如 N 的截断 (truncation)，从而最终得到对求积问题(7.144)的克伦肖——柯蒂斯法则

$$\mathcal{I} \approx I_N^{CC} \equiv \tilde{a}_0 + \sum_{\nu=1,\,\text{偶数}}^{N} \frac{2\tilde{a}_\nu}{1-\nu^2}, \tag{7.153}$$

或者用向量形式，表示为权重向量 W 和傅里叶级数系数 $\tilde{\boldsymbol{a}}$ 的点乘

$$\begin{aligned} &\mathcal{I} = \boldsymbol{W} \cdot \tilde{\boldsymbol{a}}, \\ &\tilde{\boldsymbol{a}} = \begin{pmatrix} \tilde{a}_0 \\ \tilde{a}_1 \\ \tilde{a}_2 \\ \vdots \\ \tilde{a}_N \end{pmatrix}, \quad \boldsymbol{W} = \begin{pmatrix} W_0 \\ W_1 \\ W_2 \\ \vdots \\ W_N \end{pmatrix}, \quad W_\nu \equiv \begin{cases} 1 & \nu = 0, \\ \frac{2}{1-\nu^2} & \nu \neq 0, \nu\text{是偶数}, \\ 0 & \nu\text{是奇数}. \end{cases} \end{aligned} \tag{7.154}$$

应用求积法则做近似

在求得(7.153)的基础上，需要计算傅里叶级数的系数 $\{\tilde{a}_\nu\}_\nu$。常见的思路是利用 N 个点的梯形法则，对傅里叶级数积分式(7.149)作近似[33]。

由于(7.149)也是一个沿着周期区间对周期方程求积的计算，那么其近似误差同样也随着 N 的增大而等比指数衰减。

式(7.149)在对 $\tilde{a}_\nu$ 做梯形法则近似的过程中，取样 N 个 ν 值的点，在原空间中的 $f(\cos\theta)$ 和傅里叶空间中的 $g(\theta)$ 一一对应。进而在(7.153)中，便是将这 N 个取样点对应的值加权求和，即也是一个 N 个点的求积法则。

求积法则再探讨

回顾一下上节介绍的克伦肖——柯蒂斯法则，利用该法则求解积分问题(7.153)，其中

- $\mathcal{I}$ 是一个关于 $g(\theta)$ 在傅里叶域 $[0, 2\pi]$ 中取样后求积的方程，
- $\{\tilde{a}_\nu\}_\nu$ 是一组关于方程 $g(\theta) \equiv f(\cos\theta)$ 的傅里叶系数，可以根据(7.149)利用 N 个点的梯形法则近似计算得出，也即 $\tilde{a}_\nu$ 是一个关于 $g(\theta)$ 的方程。

[32]与之相比，图7.F.7的 T-周期扩展方案则不能保证近似式一定收敛，这是由于它在每个周期的最末一点不连续，因此其傅里叶系数只能以等比于 ν^{-1} 的代数速度收敛。

[33]事实上在这里我们可以取 M-点梯形法则来近似(7.149)，$M \neq N$ 并且和 N 没有明显关联。此时近似误差的衰减速度是 $\exp(-\#M)$。另一方面，(7.153)的衰减速度是 $\exp(-\#N)$。最终我们取 $\min\{\exp(-\#M), \exp(-\#N)\}$ 作为全部近似求积的总误差收敛速度。

本节我们进一步讨论如何

- 将 $\tilde{a}_\nu$ 转换为与样本方程 $f(t)$ 有关的方程，
- 将 $\mathcal{I}$ 转换为与样本方程 $f(t)$ 有关的方程。

先来看第一个问题。回顾(7.145) (7.146) (7.147) (7.148) (7.149)可见，$\tilde{a}_\nu$ 是个在一个周期内 $[0, 2\pi]$，对周期方程 $f(\cos\theta)\cos(\nu\theta)$ 求积的运算

$$\tilde{a}_\nu = \frac{1}{2\pi}\int_0^{2\pi} f(\cos\theta)\cos(\nu\theta)\,\mathrm{d}\theta,$$

那么可以用 N 个点的梯形法则对上式做近似估计，估计的误差随着取样点数 N 的增加而呈指数速度等比衰减 [34] 。由于被积方程 $f(\cos\theta)\cos(\nu\theta)$ 是个偶方程，我们可以取原周期的一半 $[0, \pi]$ 求积，并将常数乘以 2

$$\tilde{a}_\nu = \frac{2}{\pi}\int_0^{\pi} f(\cos\theta)\cos(\nu\theta)\ \mathrm{d}\theta.$$

对上式应用 N 点梯形法则

$$\tilde{a}_\nu = \frac{2}{N}\overline{\sum}_{n=0}^{N} f(\cos\theta_n)\cos(\nu\theta_n), \quad \theta_n \equiv \frac{n\pi}{N}, \tag{7.155}$$

其中 $\overline{\sum}_{n=0}^{N}$ 是一个简化表述，表示首项 $n=0$ 和末项 $n=N$ 要在前面乘以 1/2。

(7.155)也可以进一步简化表述为

$$\begin{aligned}
&\tilde{a}_\nu = \sum_n \Lambda_{\nu n} f(t_n), \\
&t_n \equiv \cos\left(\frac{\pi n}{N}\right), \quad \Lambda_{\nu n} \equiv \begin{cases} \frac{1}{N}\cos\left(\frac{n\nu\Pi}{N}\right), & n=0 \\ \frac{2}{N}\cos\left(\frac{n\nu\Pi}{N}\right), & n=1,2,\ldots,N-1 \\ \frac{1}{N}\cos\left(\frac{n\nu\Pi}{N}\right), & n=N, \end{cases}
\end{aligned} \tag{7.156}$$

其中 t_n 可称为切比雪夫点 (Chebyshev points)。在(7.156)中，用到同一组二次项值 $\cos\left(\frac{n\nu\pi}{N}\right)$ 来对 $f(t)$ 做取样，计算不同的系数值 $\tilde{a}_\nu$。这意味着，可以将(7.156)改写为以下矩阵形式

$$\tilde{\boldsymbol{a}} = \boldsymbol{\Lambda}\cdot\boldsymbol{f}, \tag{7.157}$$

其中

- $\tilde{\boldsymbol{a}}$ 是 $(N+1)\times 1$ 的傅里叶系数向量

$$\tilde{\boldsymbol{a}} = [\tilde{a}_0, \tilde{a}_1, \ldots, \tilde{a}_N]^\top.$$

- $\boldsymbol{f}$ 是个 $(N+1)\times 1$ 的抽样值向量

$$\boldsymbol{f} = [f(t_0), f(t_1), \ldots, f(t_N)]^\top, \quad t_n \equiv \frac{n\pi}{N}.$$

[34] 通常来讲，误差项呈指数速度快速收敛的前提是原方程 $f(t)$ 平滑：克伦肖——柯蒂斯法则的周期化算法(7.145)消除了时间周期最末端一点的不连续性 (从而使得这种算法优于前文提出的原始周期化方案)，但无法消除 $f(t)$ 在周期内其他位置可能存在的不连续性。

从技术上来讲，$f(t)$ 的平滑只能确保误差项的收敛速度超过 N 值对应的任意多项式。在必要的情况下，确保误差项指数收敛的更严格前提是假定 $f(t)$ 时解析的 (analytic)，即 f 在所关注域内的任意一点上，任意级泰勒级数均收敛。

在数值近似的实际操作过程中，平滑方程假定和解析方程假定的区别并不大，从而我们可以在一般意义上仅假定 $f(t)$ 是平滑的，就可以确保傅里叶级数的近似误差呈指数收敛。

- $\boldsymbol{\Lambda}$ 是一个 $(N+1)\times(N+1)$ 矩阵，含有元素 $\Lambda_{\nu n}$

$$\Lambda=\frac{2}{N}\begin{pmatrix}\frac{1}{2} & 1 & 1 & 1 & \cdots & \frac{1}{2}\\ \frac{1}{2} & \cos\frac{\pi}{N} & \cos\frac{2\pi}{N} & \cos\frac{3\pi}{N} & \cdots & \frac{1}{2}\cos\pi\\ \frac{1}{2} & \cos\frac{2\pi}{N} & \cos\frac{4\pi}{N} & \cos\frac{6\pi}{N} & \cdots & \frac{1}{2}\cos 2\pi\\ \frac{1}{2} & \cos\frac{3\pi}{N} & \cos\frac{6\pi}{N} & \cos\frac{9\pi}{N} & \cdots & \frac{1}{2}\cos 3\pi\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ \frac{1}{2} & \cos\pi & \cos 2\pi & \cos 3\pi & \cdots & \frac{1}{2}\cos N\pi\end{pmatrix} \tag{7.158}$$

现在来看第二个问题，用抽样样本值 f 来表示 $\mathcal{I}$。已知(7.154)将求积问题 $\mathcal{I}$ 表示为域傅里叶系数 $\tilde{a}_\nu$ 有关的方程，(7.157)将 $\tilde{a}_\nu$ 表示为与抽样值 f 有关的方程，抽样点为切比雪夫点 t_N。那么两式联立可得用 f 表示的 $\mathcal{I}$，是一个求积法则

$$\mathcal{I}=\boldsymbol{W}^T\boldsymbol{\Lambda}f,$$

或者进一步简化

$$\mathcal{I}=\boldsymbol{w}\cdot\boldsymbol{f},\quad \boldsymbol{w}\equiv\lambda^T\boldsymbol{W}, \tag{7.159}$$

$\boldsymbol{w}$ 是一个 $(N+1)\times 1$ 的权重向量。

(7.159)等价于

$$\mathcal{I}=\int_{-1}^{1}f(t)\lceil t=\sum_{n=0}^{N}w_n f\left(t_n\right), \tag{7.160}$$

这成为一个 $(N+1)$ 个点的求积法则，利用 $N+1$ 个抽样 f 的加权求和来近似 $\mathcal{I}$。其中权重向量 $\boldsymbol{w}$ 的计算步骤为

- 利用(7.154)计算 $\boldsymbol{W}$，
- 利用(7.158)计算 $\boldsymbol{\Lambda}$，
- 利用 $\boldsymbol{w}=\boldsymbol{\Lambda}^\top W$ 算得 $\boldsymbol{w}$。

当 N 是偶数时，克伦肖——柯蒂斯求积法则中权重 $\boldsymbol{w}$ 的显性形式

$$w_m=\begin{cases}\frac{1}{N^2-1} & m=0\\ \frac{2}{N}\left\{1+\left[\sum_{n=1}^{\frac{N}{2}-1}\left(\frac{2}{1-4n^2}\right)\cos\left(\frac{2mn\pi}{N}\right)\right]+\frac{\cos m\pi}{1-N^2}\right\} & m=1,\ldots,N-1\\ \frac{1}{N^2-1}, & m=N\end{cases} \tag{7.161}$$

程序实现

举例，求积问题

$$\mathcal{I}=\int_{-1}^{1}f(x)\,\mathrm{d}x,\quad f(x)=x\cdot\sin\left(2\cdot\exp\left(2\cdot\sin\left(2\cdot\exp\left(2x\right)\right)\right)\right)$$

Mathematica 可用内建函数 NIntegrate[ ] 来计算求解，如

```
In[1]:= func[x_] := x*Sin[2  Exp[2  Sin[2 Exp[2  x]]]]

In[2]:= NIntegrate[func[x], x, -1, 1, Method -> "ClenshawCurtisRule"]

Out[2]= 0.336733
```

Matlab 中并没有直接内建的用于克伦肖——柯蒂斯数值积分程序。我们用 chebfun的程序来计算，代码如下。

```
% 假定要对方程f(x)在[-1,1]求积
x = chebfun('x');
f = @(x) x.*sin(2*exp(2*sin(2*exp(2*x))));
% 输入help @查询命令
% .* 表示按元素相乘 *表示矩阵乘
fc = chebfun(f);
LW = 'linewidth';
FS = 'fontsize';
MS = 'markersize';
figure, plot (fc, LW, 1.2)
title('Function f', FS, 16)
```

输出 $f(x), x \in [-1,1]$ 见图7.F.9。

图 7.F.9: $f(x), x \in [-1,1]$

```
% 输出积分的数值近似值
Ichebfun = sum(fc)
```

返回值 0.336732834781728。

作克伦肖——柯蒂斯法则近似，和 Gauss-Legendre 近似 (Glaser et al., 2007)

```
% 求积近似过程中，使用了多少插值点
Npts = length(fc)

[s,w] = chebpts(Npts); % 使用chebshev points
% s 是在[-1,1]之间插值的若干点(N)
% w 是对应的权重矩阵
```

```
% 两个矩阵相乘，得到求积的值，命名为Iclenshawcurtis，等同于sum(fc)的命令
Iclenshawcurtis = w * f(s)

% 也可以尝试用Legendre points (gauss-legendre求积算法)
[s,w] = legpts(Npts);
Igauss = w * f(s)
```

绘制图形，描述将两种近似方法测得的近似误差，随着 N 值增大的指数衰减情况，输出结果见图7.F.10。

```
% 绘出采用Gauss-legendre算法，误差项的收敛情况. semilogy与plot命令类似，除了similogy输出log形式的值
semilogy(NN, err, '.-', LW, 1, MS, 16), grid on
ylim([1e-18 1])
xlabel('Npts', FS, 12), ylabel('Error', FS, 12)
title('Gauss quadrature convergence', FS, 16), toc % toc 输出计算时间

% 现在将Gauss-Legendre算法和Clenshaw-Curtis算法的误差收敛情况都绘制出来
hold on, tic, err = [];
for Npts = NN
    [s,w] = chebpts(Npts);
    Iclenshawcurtis = w * f(s);
    err = [err abs(Iclenshawcurtis-Ichebfun)];
end
semilogy(NN, err, '.-r', LW, 1, MS, 16)
title('Gauss-Legendre and Clenshaw-Curtis', FS, 16)
legend('Gauss-Legendre', 'Clenshaw-Curtis', 'location', 'southwest'), toc
```

7.G 傅里叶分析

我们可以这样理解分析 (analysis)：将研究对象分解成小块来分别研究。那么我们可以这样理解傅里叶分析 (Fourier analysis)：将研究对象 (方程) 分解为小块，即不同的弦波 (sinosoids)，每个小快都以有限的速度变化。例如这个方程

$$f(t) = 3\cos 2\pi t + 19\sin 4\pi t - 0.14\cos 7\pi t,$$

可以分解为

$$f(t) = 3A + 19B - 0.14C,$$

其中 A, B, C 分别表示角频率 (angular frequency)为 $2\pi, 4\pi, 7\pi$ 的弦波。

图 7.F.10: 求积法则误差的衰减

但对于这样的方程

$$f(t) = \exp\left(-\alpha\left|t\right|\right),$$

$$f(t) = \begin{cases} 1, & |t| < 1 \\ 2, & |t| > 1 \end{cases},$$

甚至更复杂一些的方程，该如何做傅里叶分析?

在开始正式介绍之前，有必要做一些简单的概念界定，见表7.G.1。

表 7.G.1: 傅里叶分析的常见概念界定

	连续方程 $f(t)$	离散方程 $f_n \equiv f\left(n\Delta t\right), n \in \mathcal{Z}$
无限域 $-\infty < t < \infty$	傅里叶变换 (Fourier transform)	半离散傅里叶变换 (Semidiscrete Fourier transform)
有限域 $-\frac{T}{2} < t < \frac{T}{2}$	傅里叶级数 (Fourier seires)	离散傅里叶变换 (Discrete Fourier transform)

7.G.1　傅里叶变换

我们先从表7.G.1左上角的傅里叶变换开始介绍，即连续方程 $f(t)$ 在无限时间区间 $(-\infty, \infty)$ 的求积问题。

$$\tilde{f}\left(\omega\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-i\omega t\right) f\left(t\right) \mathrm{d}t, \tag{7.162}$$

其中 $f\left(\omega\right)$ 是一个关于频率 ω 的方程，方程形式反映了在频率域 ω 中表现出的非线性 (指数) 强烈程度。

现在基于傅里叶变换式(7.162)来回顾前面的问题：如何将方程 $E_\alpha(t) = \exp\left(-\alpha\left|t\right|\right)$ 分解为一组弦波组合的形式？其基本思路是，通过将所有关于频率 ω 的弦波曲线加总来重写原方程

$\exp(-\alpha|t|)$，某一弦波的频率越大，振幅 (amplitude)。

$$\begin{aligned}\widetilde{E}_\alpha(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-i\omega t)\exp(-\alpha|t|)\,\mathrm{d}t \\ &= \frac{1}{2\pi}\left\{\int_{-\infty}^{0}\exp[(-i\omega+\alpha)t]\,\mathrm{d}t + \int_{0}^{\infty}\exp[(-i\omega-\alpha)t]\right\} \\ &= \frac{1}{2\pi}\left[\frac{1}{\alpha-i\omega}-\frac{1}{\alpha+i\omega}\right] \\ &= \frac{\alpha}{\pi(\alpha^2+\omega^2)},\end{aligned}$$

注意，这里所说的 “频率”ω 越大或越小，是相对于参数 α 而言的：α 越大，指数方程 $E_\alpha(t)$ 越快衰减至 0，我们就越需要增加 $\widetilde{E}_\alpha(\omega)$ 的振幅，利用更高振幅的弦波来近似原方程。下面将作进一步说明。

一些概念

$$y = A\sin(Bx+c)+D,$$

其中

- A 表示振幅，
- $\frac{B}{2\pi}$ 表示周期，频率 $=\frac{1}{\text{周期}}$，
- $-\frac{C}{B}$ 表示相位的 (水平) 移动，
- D 表示相位的 (垂直) 移动。

图 7.G.1: 傅里叶变换

图7.G.1中

- $\omega = \frac{2\pi}{T} = 2\pi\nu == \frac{\mathrm{d}s}{\mathrm{d}t}\cdot\frac{1}{r} = \frac{\mathrm{d}\theta}{\mathrm{d}t}$
- ω，角频率 (angular frequency)，是对物体旋转速度快慢的度量，单位如弧度/秒 (radian/second)。
- T，周期，单位如秒。
- ν，线频率 (linear frequncy)，单位如赫兹 (Hertz)，$1Hertz = 1$次/秒。
- 旋转的半径，单位如米。

线频率和角频率

对于一个给定的弦波方程如 $\sin\omega t$ 或 $\exp[-i\omega t]$ 等，严格说来我们需要将其中的频率 ω 理解为角频率 ω(angular frequency)，表示相角旋转速度的快慢，单位如角每秒 (radian per second)。线频率 (linear frequency) 可写为$\nu = \frac{\omega}{2\pi}$，表示这一过程自我重复的频率，单位如赫兹。例如，一个钟摆 4 秒钟转一圈，那么

- 周期 $T = 4$，
- 角频率 $\omega = \frac{2\pi}{T} = \frac{\pi}{2} \approx 1.57 Rad/s$，
- 线频率 $\nu = \frac{1}{T} = 0.25 Hz$。

域

(7.162)对原方程 $f(t)$ 作傅里叶变换，生成近似方程 $\tilde{f}(\omega)$，二者等价但处于不同的域中。$f(t)$ 的解释变量是 (时间)t，$\tilde{f}(\omega)$ 的解释变量是频率 ω，从这个角度可以说，$f(t)$ 是在时间域中，$\tilde{f}(\omega)$ 是在频率域中。

福利也分析也可用于原方程 f 的解释变量不是时间的情况，例如可以是方位 x，对应的傅里叶变换处于空间频率 k 中，有时也称波数 (wavenumber)。更多例子见下表。

表 7.G.2: 原方程域和傅里叶变换方程域

	原变量	原域	傅里叶变量	傅里叶域
信号处理	时间 t	时间域	频率 ω	频率域
光学	方位 x	实空间	波数 k	k 空间
量子力学	方位 x	方位空间	动量 (momentum)p	动量空间
固体物理学	点阵向量 L	实空间	布洛赫向量 (Bloch vector) k	晶体动量空间 (crystal momentum space)

单位

实际研究过程中需要牢记 $f(t)$ 和 $\tilde{f}(\omega)$ 的单位不同。以(7.162)为例，RHS 中有 $\mathrm{d}t$，LHS 中没有。因此

$$\tilde{f}\text{的单位} = f\text{的单位} \times \text{时间} = \frac{f\text{的单位}}{\text{频率}},$$

如果 f 的单位是伏特 (Volt)，那么 $\tilde{f}$ 的单位是伏特秒，或者伏特每赫兹。

傅里叶变换的性质

由定义(7.162)可得，傅里叶变换具有如下性质，我们将分别作说明。

- 导数的傅里叶变换
- 傅里叶变换的导数
- 实值方程的傅里叶变换

导数的傅里叶变换。对于某一给定方程 $f(x)$，如果我们可以求得其傅里叶变换 $\tilde{f}(k)$，那么 $f(x)$ 的导数也很容易写出。首先写出 $f(t)$ 的傅里叶综合表达式 (Fourier synthesis equation, 第7.186节)

$$f(x) = \int_{-\infty}^{\infty} \tilde{f}(k) \exp(ikx)\, \mathrm{d}k,$$

两侧同时对 x 求导

$$\frac{\mathrm{d}}{\mathrm{d}x}f(x)=\int_{-\infty}^{\infty}ik\tilde{f}(k)\exp\left(ikx\right)\,\mathrm{d}k,$$

上式的 RHS 是对 $\left(ik\widetilde{f}(k)\right)$ 的逆傅里叶变换 (Definition 8.32)；因此可以将 RHS 看作是 $\frac{\mathrm{d}}{\mathrm{d}x}f(x)$ 的傅里叶变换，即

$$\mathcal{F}\left[f(x)\right]=\tilde{f}(k)\Rightarrow\mathcal{F}\left[\frac{\mathrm{d}}{\mathrm{d}x}f\left(x\right)\right]=ik\tilde{f}(k). \tag{7.163}$$

也可以重复上述过程求解高阶导数，如

$$\mathcal{F}\left[f(x)\right]=\tilde{f}(k)\Rightarrow\mathcal{F}\left[\frac{\mathrm{d}}{\mathrm{d}x}f(x)\right]=ik\tilde{f}(k)\Rightarrow\mathcal{F}\left[\frac{\mathrm{d}^2f}{\mathrm{d}x^2}\right]=-k^2\tilde{f}(k). \tag{7.164}$$

傅里叶变换的导数。类似地，对傅里叶变换(7.162)

$$\tilde{f}\left(k\right)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-ikx\right)f\left(x\right)\mathrm{d}x$$

两侧求 k 的导数有

$$\frac{\mathrm{d}}{\mathrm{d}k}\tilde{f}(k)=\int_{-\infty}^{\infty}-ixf(x)\exp\left(-ikx\right)\,\mathrm{d}x,$$

可见 $\frac{\mathrm{d}}{\mathrm{d}k}\tilde{f}(k)$ 可以看成是方程 $x\cdot f(x)$ 的傅里叶变换。因此有

$$\mathcal{F}\left[f(x)\right]=\tilde{f}(k)\Rightarrow\mathcal{F}\left[xf(x)\right]=i\frac{\mathrm{d}}{\mathrm{d}k}\tilde{f}(k). \tag{7.165}$$

不难看出，(7.163) 和(7.165)等价。

实值方程的傅里叶变换。如果 $f(x)$ 是一个实值方程，那么它对应的傅里叶变换 $\tilde{f}(k)$ 中所含有的信息中，相对于 $k<0$ 的部分就是冗余信息：$\tilde{f}\left(k\right)\big|_{k<0}$ 的信息可由 $\tilde{f}\left(k\right)\big|_{k>0}$ 的部分而获得。以定义式为例，对于 $k<0$ 的情况下有

$$\tilde{f}\left(-k\right)=\frac{1}{2\pi}\int_{-\infty}^{\infty}f(x)\exp\left(ikx\right)\,\mathrm{d}x,$$

由于假定 $f(x)$ 是一个实值方程，进一步可得

$$\tilde{f}(-k)=\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}f(x)\exp\left(-ikx\right)\,\mathrm{d}x\right]^*=\tilde{f}^*(k).$$

7.G.2 脉冲方程的傅里叶变换

介绍几种常见的傅里叶变换。

洛伦兹变换

以前文提到的方程为例

$$E_\alpha\left(x\right)=\exp\left(-\alpha\left|x\right|\right),\quad\widetilde{E}_\alpha\left(k\right)=\frac{\alpha}{\pi\left(\alpha^2+k^2\right)}. \tag{7.166}$$

$E_\alpha\left(x\right)$ 和 $\widetilde{E}_\alpha\left(k\right)$ 都可以被看作是某种脉冲方程 (pulse functions)：它们都在初始点 x_0,k_0 处附近有最大值，又随着 $x\to\infty$ 或 $k\to\infty$ 而逐渐降低直至 0，称为洛伦兹变换 (Lorentzian function)。两个方程的宽度均与参数 α 有关，但影响方向相反：α 的值越大，实空间中的脉冲越弱、宽度越窄，傅里叶空间中的脉冲越强、宽度越宽，如图7.G.2所示。

图 7.G.2: 洛伦兹方程 ($\alpha = 2$)

我们将脉冲的一个半峰全宽 (full width at half maximum, FWHM)定定义为，在方程的一个周期中，前后两个方程值等于一个峰值一半的点之间的距离。根据这个定义，$E_\alpha(x)$ 的峰值为 $1(x=0)$，并且在 $x \pm (\ln 2)/\alpha$ 时，值为峰值的一半，因此有

$$FWHM\left[E_\alpha(x)\right] = \frac{2\ln 2}{\alpha}, \tag{7.167}$$

$\widetilde{E}_\alpha(k)$ 的峰值为 $\frac{1}{\pi\alpha}(k=0)$，在 $k=\pm\alpha$ 时，值为峰值的一半，

$$FWHM\left[\widetilde{E}_\alpha(k)\right] = 2\alpha. \tag{7.168}$$

(7.167)和(7.168)联立有

$$FWHM\left[E_\alpha(x)\right] \cdot FWHM\left[\widetilde{E}_\alpha(k)\right] = 4\ln 2, \tag{7.169}$$

注意：单个原方程及洛伦兹变换方程的 FWHM 与 α 有关；两个 $FWHM$ 的乘积则与 α 无关：所有洛伦兹脉冲方程族 $\{E_\alpha(t)\}$ 都具有这样的特征。

高斯变换

设一个宽度为 σ 的高斯方程 (Gaussian function)[35]

$$G_\sigma(x) = \exp\left(-\frac{x^2}{\sigma^2}\right),$$

那么其傅里叶变换

$$\begin{aligned}\widetilde{G}_\sigma(k) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-ikx)\, G_\sigma(x)\, \mathrm{d}x \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \underbrace{\exp\left(-ikx - \frac{x^2}{\sigma^2}\right)}_{=:\mathcal{A}} \mathrm{d}x,\end{aligned}$$

[35]高斯方程族中有不同的类型，简要介绍见第7.G.9节。

RHS 的被积方程

$$
\begin{aligned}
\mathcal{A} &= \frac{1}{\sigma^2}\left(x^2+ikx\sigma^2\right) \\
&= \frac{1}{\sigma^2}\left(x+\frac{1}{2}ik\sigma^2\right)^2+\frac{1}{4}k^2\sigma^2 \\
&= \frac{1}{2\pi}\exp\left(-\frac{1}{4}k^2\sigma^2\right)\underbrace{\int_{-\infty}^{\infty}\exp\left[-\frac{1}{\sigma^2}\left(x+\frac{1}{2}ik\sigma^2\right)^2\right]\mathrm{d}x}_{=\sigma\cdot\sqrt{\pi}} \\
&= \frac{\sigma}{2\sqrt{\pi}}\exp\left(-\frac{1}{4}k^2\sigma^2\right),
\end{aligned}
$$

可见 $\widetilde{G}_\sigma(k)$ 也是一个 k 空间中的高斯方程，其宽度与原方程 $G_\sigma(x)$ 的宽度成等比例

$$
\widetilde{G}_\sigma \propto \exp\left(-\frac{k^2}{\tilde{\sigma}^2}\right)=G_{\tilde{\sigma}}\left(k\right),\quad \tilde{\sigma}\equiv\frac{2}{\sigma}.
$$

此外

$$
\begin{aligned}
FWHM\left[G_\sigma\left(x\right)\right] &= 2\sqrt{\ln 2}\sigma, \\
FWHM\left[G_\sigma\left(x\right)\right]\cdot FWHM\left[\widetilde{G}_\sigma\left(x\right)\right] &= \left(2\sqrt{\ln 2}\sigma\right)\cdot\left(\frac{4\sqrt{\ln 2}}{\sigma}\right)=8\left(\ln 2\right)^2,
\end{aligned}
\tag{7.170}
$$

可见单个原高斯方程及高斯变换方程的 FWHM 与 α 有关；两个 $FWHM$ 的乘积则与 α 无关：所有高斯脉冲方程族 $\{G_\sigma(x)\}$ 都具有这样的特征。

7.G.3 非脉冲方程的傅里叶变换

洛伦兹变换和高斯变换都属于脉冲方程，特点是初始点附近有峰值，随着 $x(k)$ 逐渐增大而慢慢回落至 0。如何求得非脉冲方程的傅里叶变换，如 $f(x)=1$, $f(x)=x$, 或 $f(x)=x^2$ 呢?

一种求解思路是，考虑(7.166)在 $\lim_{\alpha\to 0}$ 时的情况，实值方程趋近于 1

$$
\lim_{\alpha\to 0}E_\alpha(x)\to 1.
$$

而它的洛伦兹变换 $\widetilde{E}_\alpha(k)$ 的求解则较为复杂，第一它的宽度变窄了 $FWHM\widetilde{E}_\alpha=2\alpha$；第二它的高度变高了 $\widetilde{E}_\alpha\left(0\right)=\frac{1}{\pi\alpha}$。随着 $\alpha\to 0$ 的极限情况出现，$\lim_{\alpha\to 0}\widetilde{E}_\alpha(k)$ 变得极窄、极高，就成为狄拉克方程 (dirac delta function)

$$
\lim_{\alpha\to 0}\widetilde{E}_\alpha(k)=\lim_{\alpha\to 0}\frac{\alpha}{\pi\left(\alpha^2+k^2\right)}\equiv\delta(k)=\begin{cases}+\infty & k=0,\\ 0 & k\neq 0,\end{cases}\quad\Rightarrow\int_{-\infty}^{\infty}\delta(k)\,\mathrm{d}k=1.
\tag{7.171}
$$

对应地，非脉冲方程 $f(x)=1$ 的傅里叶变换可以写作

$$
f(x)\equiv 1\Rightarrow\tilde{f}(k)=\delta(k),
\tag{7.172}
$$

可以这么理解：$f(x)\equiv 1$ 是一个频率为 0 的弦波。对 $f(x)$ 做傅里叶转换，分解为若干弦波之和，在这一系列弦波中，我们将除了第一条 (对应频率为 0) 之外的其他全部弦波曲线的系数值均设为 0。

在此基础上，利用(7.172)和(7.165)，进一步可得到另外两个非脉冲方程的傅里叶变换

$$
f(x)=x\quad\Rightarrow\tilde{f}(k)=i\delta^{'}(k),
\tag{7.173}
$$

$$
f(x)=x^2\quad\Rightarrow\tilde{f}(k)=-\delta^{''}(k),
\tag{7.174}
$$

其中 $\delta'(k)$ 是狄拉克方程的导数，基于(7.171)用分部积分法有

$$\int f(u)\delta^{'}(u)\mathrm{d}u = -\int f^{'}(u)\delta(u)\,\mathrm{d}u = -f^{'}(u),$$

可见 δ' 可以被看做是和 δ 相近的方程，二者只有一点不同：利用分部求积计算 δ' 时产生了 $-f'(0)$ 的变化，而 δ 只产生 $-f$ 的变化。

需要指出的是，对非脉冲方程的傅里叶变换如(7.172)、(7.173)、(7.174)，都不算是“好”的方程——它们甚至不是方程，而只是分布 (比如狄拉克方程单独出现往往没有实际意义，它更多是与其他定义良好的方程配对出现在积分式中)：这么说是因为 $f(x) = \left\{1, x, x^2\right\} \notin L^1$，即它们不满足 $\int_{-\infty}^{\infty} |f(x)|\,\mathrm{d}x < \infty$。[36] 因此，(7.174)等傅里叶变换，往往更适合以算子的形式出现。

7.G.4　原方程的平滑和傅里叶变换方程的衰减

通过前面的分析不难总结出一个一般规律：原方程 $f(t)$ 越是 (口语意义上的) 不“平滑”，即随着解释变量 t 的变动而剧烈变动，其傅里叶变换 $\tilde{f}(\omega)$ 随着 ω 的衰减 (decay) 速度就越慢。反之亦然：$f(t)$ 越平滑，$\tilde{f}(\omega)$ 的衰减就越快。可以将 $f(t)$ 的平滑特性表现为 f 的连续性和导数：

Theorem 7.62 (佩利——维纳诸定理)**.** *如果 $f(t)$ 和它的前 $p-1$ 个导数都是连续的，但第 p 个导数是不连续的，并且变差有界 (bounded variation)，那么 $\tilde{f}(\omega)$ 的衰减速度不低于 $|\omega|^{-(p+1)}\big|_{\omega\to\infty}$。这称为佩利——维纳诸定理 (Paley-Wiener theorems)。*

证明. 一种证明方式见第179页脚注。更详实的证明可参考如 (Yosida, 1978, Sec. VI.4)，(Agranovich, 2015, Theorem 4.1.2)。 □

尤其是，若 $f(t) \in C^\infty$，[37]那么在 ω 值较大时，$\tilde{f}(\omega)$ 的衰减速度比任何多项式都快。这样的方程包括诸如 $\exp(-\omega)$，$\exp(-\sqrt{\omega})$，$\exp\left(-\omega^2\right)$ 等。

举例如洛伦兹方程 $\exp(-\alpha|t|)$ 连续，但其一阶导数不连续，表现为在 $t=0$ 处有跳跃，如图7.G.2左侧。因此 $p=1$。对应地，我们会看到它的傅里叶变换以等比于 ω^{-2} 的速度衰减 (当 ω 的值较大时)。

再比如，方程 $\exp\left(-\frac{t^2}{\sigma^2}\right) \in C^\infty$，那么其傅里叶变换 $\widetilde{F} = \exp\left(-\frac{1}{4}\sigma^2\omega^2\right)$ 的衰减速度快于任意多项式。

7.G.5　傅里叶级数

现在来看表7.G.1左下一栏，对应有限空间中连续方程 $f(t)$ 的傅里叶分析，称为傅里叶级数。

假设 $f(t)$ 是个关于 T 的周期方程 (periodic function)，即每隔 T 单位时间 (秒，分，小时，天等) 就自我重复一遍：

$$f(t+T) \equiv f(t), \quad \forall t.$$

假定 $f(t)$ 符合傅里叶分析的条件，对应傅里叶变换

$$\tilde{f}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-i\omega t)\, f(t)\,\mathrm{d}\omega, \tag{7.175}$$

随着角频率 ω 和基准频率 (base frequency) $\omega_0 = \frac{2\pi}{T}$ 的关系不同，需要分两种情况来分析：

[36] L^1 表示勒贝格空间，见第(8.2.2)节。

[37] 空间 C^∞ 中的方程连续，所有导数也都连续，可参考第8.2.1节。

1. ω 是 ω_0 的整数倍。

   那么(7.175)中的被积方程周期为 T，每个宽度为 T 的子区间对总积分的作用相同，并且由于这样的子区间有无数个，我们有 $\tilde{f}(\omega)=\infty$。

2. ω 是 ω_0 的整数倍。

   那么(7.175)中的被积方程不是一个周期方程。我们可以将 $f(t)$ 理解为两部分：一部分是 $f(t)$，这是一个关于 T 的周期方程，另一部分是 $\exp(-i\omega t)$，它只有在某些特定的时期才是周期方程，这些特定的时期是指不等于 T 的整数倍的时期。两部分的乘积导致被积方程不是一个周期方程。

   每个宽度为 T 的子区间都给总积分值产生同等程度的影响，但影响的大小 (phase，相) 是随机的。各个随机相的影响加总后相互抵消，$\tilde{f}(\omega)=0$。

将上述两种情况汇总可见，在 $f(t)$ 是一个周期方程的前提下

$$\tilde{f}(\omega)=\begin{cases}\infty, & \omega\text{是}\omega_0\text{的整数倍},\\ 0, & \text{否则}.\end{cases}$$

我们可以用三种数学形式将上式表达出来：

1. 利用狄拉克方程，可将 T-周期方程的傅里叶变换表示为

$$\tilde{f}(\omega)=\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu\cdot\delta(\omega-\nu\omega_0),\quad \omega_0\equiv\frac{2\pi}{T},\nu\in\mathcal{Z}. \tag{7.176}$$

2. 对 T-周期方程 $f(t)$ 的傅里叶分解，只包括那些频率为 ω_ν 的弦波曲线，满足 $\omega_\nu=\nu\cdot\omega_0=\frac{2\nu\pi}{T},\nu\in\mathcal{Z}$，记作

$$f(t)=\sum_{\nu\in\mathcal{Z}}\tilde{f}_\nu\cdot\exp(i\nu\omega_0 t),\quad \omega_0=\frac{2\nu\pi}{T},\nu\in\mathcal{Z}, \tag{7.177}$$

   我们又将(7.177)称为傅里叶级数 (Fourier series)的表达形式之一：复杂指数表达形式。

3. 除此之外，也可以将傅里叶级数表达为另一种形式，即一组正弦函数和一组余弦函数之和

$$f(t)=\sum_{\nu=0}^{\infty}\tilde{a}_\nu\cdot\cos(\nu\omega_0 t)+\sum_{\nu=1}^{\infty}\tilde{b}_\nu\cdot\sin(\nu\omega_0 t),\quad \nu\in\mathcal{Z}^+. \tag{7.178}$$

两种表达形式的比较

(7.178)中的 ν 需要取正整数；(7.177)则正负均可。除此而外，指数形式(7.177)和弦波形式(7.178)等价，可以被看作是同一个 T-周期方程 $f(t)$ 的傅里叶级数分解。

来看两组系数 $\left\{\tilde{a}_\nu,\tilde{b}_\nu\right\}_{\nu\in\mathcal{Z}^+}$ 和 $\left\{\tilde{f}_\nu\right\}_{\nu\in\mathcal{Z}}$ 之间的关系。已知复杂指数 $\exp(i\nu\omega_0 t)$ 可表示为关于正弦和余弦函数的方程

$$\exp(i\nu\omega_0 t)=\cos(\nu\omega_0 t)+i\sin(\nu\omega_0 t), \tag{7.179}$$

以及三角函数关系：余弦函数是偶函数，正弦函数是奇函数

$$\cos(-A)=\cos(A),\quad \sin(-A)=-\sin(A), \tag{7.180}$$

结合(7.177)-(7.180)可得

$$\begin{aligned}
\tilde{a}_0 &= \tilde{f}_0, \\
\tilde{a}_\nu &= \left(\tilde{f}_\nu + \tilde{f}_{-\nu}\right), \quad \nu > 0, \\
\tilde{b}_\nu &= i\left(\tilde{f}_\nu - \tilde{f}_{-\nu}\right), \quad \nu > 0,
\end{aligned} \tag{7.181}$$

或者反过来

$$\tilde{f}_\nu = \begin{cases} \frac{1}{2}\left(\tilde{a}_\nu + i\tilde{b}_\nu\right), & \nu < 0, \\ \tilde{a}_0, & \nu = 0, \\ \frac{1}{2}\left(\tilde{a}_\nu - i\tilde{b}_\nu\right), & \nu > 0. \end{cases} \tag{7.182}$$

比较两种表达式形式：复杂指数形式(7.177)和正弦余弦形式(7.178)：

1. (7.178)的好处在于
   (a) 若 T-周期 $f(t)$ 是实值方程，那么其傅里叶变换系数 $\left\{\tilde{a}_\nu, \tilde{b}_\nu\right\}_{\nu\in\mathcal{Z}^+}$ 也都是实值系数，从而无需考虑复数的情况。
   (b) 结合具体 $f(t)$ 的奇偶性，可以作灵活调整，从而简化后续的计算工作。例如若 $f(t)$ 是偶函数，那么它的傅里叶变换只包含余弦项；全部正弦项都消失了，因此无需计算 $\tilde{b}_\nu$。
2. (7.177)的好处在于
   (a) 只需计算一组系数 $\left\{\tilde{f}_\nu\right\}_{\nu\in\mathcal{Z}}$，而非两组系数 $\left\{\tilde{a}_\nu, \tilde{b}_\nu\right\}_{\nu\in\mathcal{Z}^+}$。
   (b) 指数形式傅里叶级数的推算公式(7.177)更直观和易于理解。
   (c) 求导更容易。事实上对(7.177)求导后生成的傅里叶级数仍然是(7.177)，只是原来的系数 $\left\{\tilde{f}_\nu\right\}_{\nu\in\mathcal{Z}}$ 现在变成了 $\left\{i\omega_0\tilde{f}_\nu\right\}_{\nu\in\mathcal{Z}}$。反观(7.178)，求导操作更为复杂。

基于上述考虑，在大多数情况下我们采用复杂指数形式(7.177)展开傅里叶级数分析[38]，使得更为直观和容易理解；然而这两种形式之间的互相转换仍然很简单，如式(7.181)-(7.182)所示。

复杂指数形式的系数计算

假定一个实值方程 $f(t)$，在 $t \in [0, T]$ 时域内的傅里叶级数，其系数 $\left\{\tilde{f}_\nu\right\}_{\nu\in\mathcal{Z}}$ 的计算式为

$$\tilde{f}_\nu = \frac{1}{T}\int_0^T f(t)\exp\left(-i\nu\omega_0 t\right)\,\mathrm{d}t.$$

例如，设 $f(t) = \cos^2(3t)$, $T = \frac{\pi}{3}$，见图(7.G.3)。
基准频率 $\omega_0 = \frac{2\pi}{T} = 6$。已知

$$\begin{aligned}
&\cos(3t) = \frac{1}{2}\left[\exp(3it) + \exp(-3it)\right], \\
&\hookrightarrow f(t) = \cos^2(3t) = \frac{1}{4}\left[\exp(6it) + 2 + \exp(-6it)\right], \\
&\hookrightarrow \tilde{f}_\nu = \frac{1}{4T}\int_0^T \exp\left(-i\nu\omega_0 t\right)\left[\exp\left(i\omega_0 t\right) + 2 + \exp\left(-i\omega_0 t\right)\right]\,\mathrm{d}t.
\end{aligned} \tag{7.183}$$

根据正交性质 (第7.G.9节) 可得，上式变为

$$\tilde{f}_\nu = \frac{1}{4}\left[\delta_{\nu,1} + 2\delta_{\nu,0} + \delta_{\nu,-1}\right],$$

[38]除了个别情况，比如利用切比雪夫光谱法来分析实值偶函数 $f(t)$ 时。

图 7.G.3: $f(t)=\cos^2(3t), t\in(-2,2)$

换句话说，傅里叶级数的系数 $\tilde{f}_\nu$ 只有在 $\nu=\{-1,0,1\}$ 时不等于 0。那么，将 $f(t)$ 用傅里叶综合形式 (第7.186节) 写为

$$\begin{aligned} f(t)&=\sum_\nu \tilde{f}_\nu \exp(-i\nu\omega_0 t)=\frac{1}{4}\exp(i\omega_0 t)+\frac{1}{2}+\frac{1}{4}\exp(-i\omega_0 t)\\ &=\frac{1}{2}[1+\cos(\omega_0 t)]=\frac{1}{2}[1+\cos(6t)], \end{aligned} \tag{7.184}$$

结合三角函数关系，上式还可以做进一步简化

$$\cos^2(3t)=\frac{\cos(6t)+1}{2}.$$

傅里叶余弦级数

(7.184)将 $f(t)=\cos^2(3t)$ 分解为一个余弦项和一个常数项之和，没有正弦项。事实上若方程 $f(t)$ 是偶函数，傅里叶变换后常常是没有正弦项的。反之亦然。我们将两种情况分别称为傅里叶余弦级数 (Fourier cosine series)和傅里叶正弦级数 (Fourier sine series)。

有时候原始方程本身既不是奇函数也不是偶函数，但有可能通过坐标点的移动来变成奇或偶函数。如 T-周期方程

$$f(t)=\begin{cases}0, & 0<t<\frac{T}{2}\\ 1, & \frac{T}{2}<t<T\end{cases}$$

非奇非偶。可以设一个新函数 $g(t)\equiv f\left(t+\frac{\pi}{4}\right)$，使得 $g(t)$ 变成偶函数。

更一般地说，所有方程都可以分解为奇函数、偶函数组合的形式，如

$$\begin{aligned} f(t)&=f_{偶}(t)+f_{奇}(t),\\ f_{偶}(t)&\equiv\frac{1}{2}[f(t)+f(-t)],\\ f_{奇}(t)&\equiv\frac{1}{2}[f(t)-f(-t)]. \end{aligned}$$

举例来说如锯齿波 (sawtooth wave)，T-周期方程 $f(t)$ 在每一个周期内都满足 $f(t)=t,\quad 0<t<T$，也即，$f(t)$ 的单位与 t 的单位相同，都是时间，见图7.G.4。

该方程的傅里叶级数

$$\begin{aligned} f(t)&=\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu\cdot\exp(i\nu\omega_0 t),\quad \omega_0=\frac{2\pi}{T},\\ \hookrightarrow \tilde{f}_\nu&=\frac{1}{T}\int_0^T f(t)\exp(-i\nu\omega_0 t)\,\mathrm{d}t=\frac{1}{T}\int_0^T t\exp(-i\nu\omega_0 t)\,\mathrm{d}t. \end{aligned}$$

图 7.G.4: 锯齿波

上式中 $\tilde{f}_\nu$ 的值随着 ν 的取值而不同，具体说来:

$$\tilde{f}_0 = \frac{T}{2}, \quad \nu = 0,$$

$$\tilde{f}_\nu = \frac{1}{T}\left[-\frac{1}{i\nu\omega_0}\left|t\exp\left(-i\nu\omega_0 t\right)\right|_0^T + \frac{1}{i\nu\omega_0}\int_0^T \exp\left(-i\nu\omega_0 t\right)\,\mathrm{d}t\right] = -\frac{1}{i\nu\omega_0}, \quad \nu \neq 0.$$

代回可得

$$\underbrace{\tilde{f}_\nu}_{(\text{时间})} = \underbrace{\frac{T}{2}}_{\text{时间}} \underbrace{-\frac{1}{i\omega_0}\sum_{\nu=-\infty}^{\infty}\frac{1}{\nu}\exp\left(i\omega_0 t\right)}_{\text{单位：(1/角频率)= 时间}},$$

不难看出，在角频率 ω 的值较大时，$\tilde{f}_\nu$ 以等比于 $\frac{1}{|\omega|}$ 的速度衰减。这是由 $f(t)$ 方程的非连续性所决定的 (佩利——维纳诸定理，Theorem 7.62)。

上式也可改为正弦余弦形式

$$f(t) = \frac{T}{2} - \underbrace{\frac{T}{\pi}\sum_{\nu=1}^{\infty}\frac{1}{\nu}\sin\left(\frac{2\nu\pi t}{T}\right)}_{=:\mathcal{A}}, \tag{7.185}$$

其中 $\mathcal{A}$ 符合傅里叶级数的特征。若设 $g(t) = f(t) - \frac{T}{2}$，则 $g(t)$ 就是个傅里叶正弦级数了——只要 $g(t)$ 是一个奇函数的话——将图7.G.4垂直下移 $\frac{\pi}{2}$ 个单位不难看出，$g(t)$ 也是个奇函数。

吉布斯现象

前文可见，通过加总一系列弦波方程 (假定其中每一个弦波都是平滑的)，可以生成图7.G.4一般参差状不连续的锯齿方程，但这是有限定条件的：如果我们只截取全部傅里叶级数中的一段，即只加总有限个项 (总共有无数个项加总) 的话，这意味着只对 $f(t)$ 作有限个傅里叶综合的话，就会产生吉布斯现象 (Gibbs phenomenon)：这是指，在对非连续方程作不完全傅里叶综合时，在不连续点附近所出现的震荡。如图7.G.5所示，原方程 $f(t)$ 为锯齿方程，$f_N(t), n = 2, 5, 10, 20$ 分别为对 $f(t)$ 所做的不完全傅里叶综合

$$f_N(t) = \sum_{\nu=-N}^{N}\tilde{f}_\nu \cdot \exp\left(i\nu\omega_0 t\right).$$

不难看出，在远离不连续点 (如 $t = 0, 1$) 的地方，N 越是大，傅里叶级数越能对原方程作精确近似。而反之则不然：在接近不连续点的地方，傅里叶级数的近似精度并为随着 N 的升高而显著提升。

图 7.G.5: 吉布斯现象

傅里叶级数的收敛

图7.G.5也揭示了傅里叶级数的又一个重要特征：如果非连续方程 $f(t)$ 在 t^* 点不连续，那么它的傅里叶级数 $\sum \tilde{f}_\nu \exp(i\nu\omega_0 t^*)$ 收敛至这一不连续的中点，即

$$\lim_{N\to\infty} \sum_{\nu=-N}^{N} \tilde{f}_\nu \cdot \exp(i\nu\omega_0 t^*) \approx \frac{1}{2}\left[f\left(t_-^*\right) + f\left(t_+^*\right)\right],$$
$$f\left(t_-^*\right) \equiv \lim_{\varepsilon\to 0} f\left(t^* - \varepsilon\right), \quad f\left(t_+^*\right) \equiv \lim_{\varepsilon\to 0} f\left(t^* + \varepsilon\right), \varepsilon > 0.$$

根据这一性质，如果我们在 $[0,T]$ 区间内，对一个并非以 $[0,T]$ 为周期的不连续方程作傅里叶级数的近似，那么用 $t=0$ 点作傅里叶近似，可得 $\frac{1}{2}[f(0)+f(T)]$，这在图7.G.5中有所反映。

傅里叶分析是无损转换过程

傅里叶分析的又一个重要特征表现为，对原方程的傅里叶转换是无损 (lossless) 的，即 $f(t)$ 中的信息在 $f(t) \to \tilde{f}(\omega)$ 的过程中没有损耗。为了说明这一点，首先来介绍傅里叶综合 (Fourier synthesis)的概念：存在关于 $f(t)$ 的傅里叶级数 $\tilde{f}(\omega)$，并且可以根据傅里叶级数向回逆推得到原方程 $f(t)$，表示为

$$f(t) = \int_{-\infty}^{\infty} \exp(i\omega t) \cdot \tilde{f}(\omega)\,\mathrm{d}\omega, \tag{7.186}$$

即将一系列指数方程 (或弦波曲线) 加权加总，权重由与频率 ω 有关的傅里叶变换 $\tilde{f}(\omega)$ 给出；加总后的方程即原方程 $f(t)$。

Theorem 7.63 (帕赛瓦尔定理 (傅里叶变换))**.** *我们说(7.186)是精确无损的，是指原域中的 $f(t)$ 和傅里叶域中的 $\tilde{f}(\omega)$ 之间相互转换不会产生信息损耗，这称为帕塞瓦尔定理 (Parseval's theorem)。*

证明. 设两方程的内积 (inner product) 满足

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f^*(t) g(t)\,\mathrm{d}(t),$$

引入傅里叶综合形式(7.186)分别替代 RHS 被积方程中的 f^* 和 $g(t)$，有

$$
\begin{aligned}
\langle f,g\rangle &= \int_{-\infty}^{\infty} f^*(t)g(t)\,\mathrm{d}(t) \\
&= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\exp\left(-i\omega t\right)\tilde{f}^*\left(\omega\right)\,\mathrm{d}\omega\right\}\cdot\left\{\int_{-\infty}^{\infty}\exp\left(i\omega' t\right)\tilde{g}\left(\omega'\right)\,\mathrm{d}\omega'\right\}\mathrm{d}t \\
&= \int_{-\infty}^{\infty}\tilde{f}^*\left(\omega\right)\int_{-\infty}^{\infty}\tilde{g}\left(\omega'\right)\underbrace{\left\{\int_{-\infty}^{\infty}\exp\left[i\left(\omega'-\omega\right)t\right]\,\mathrm{d}t\right\}}_{=2\pi\delta\left(\omega'-\omega\right)}\mathrm{d}\omega'\,\mathrm{d}\omega \\
&= \int_{-\infty}^{\infty}\tilde{f}^*\left(\omega\right)\int_{-\infty}^{\infty}\tilde{g}\left(\omega'\right)2\pi\delta\left(\omega'-\omega\right)\,\mathrm{d}\omega'\,\mathrm{d}\omega \\
&= 2\pi\int_{-\infty}^{\infty}\tilde{f}^*\left(\omega\right)\tilde{g}\left(\omega\right)\,\mathrm{d}\omega,
\end{aligned}
\tag{7.187}
$$

也就是说，原方程的内积 $\langle f,g\rangle$ 等于它们两个的傅里叶变换相乘 $\tilde{f}^*\left(\omega\right)\tilde{g}\left(\omega\right)$ 后的积。 □

Theorem 7.64 (普朗歇尔定理 (傅里叶变换))**.** *在帕塞瓦尔定理 (Theorem 7.63) 的基础上，设 $f(t)\equiv g(t)$，那么由式(7.187)可得*

$$
\langle f,g\rangle = \int_{-\infty}^{\infty}\left|f(t)\right|^2\,\mathrm{d}t = 2\pi\int_{-\infty}^{\infty}\left|\tilde{f}\left(\omega\right)\right|^2\,\mathrm{d}\omega,
\tag{7.188}
$$

又称普朗歇尔定理 (Plancherel's theorem)。

现在将区间 $(-\infty,\infty)$ 缩小到某个定区间 $[0,T]$，来看傅里叶级数中方程转换的信息损耗问题。之所以关注定区间 $[0,T]$，要么是因为原方程 $f(t)$ 是 T-周期方程，要么是我们的研究中只关注特定区间内方程值的变化。

Theorem 7.65 (帕赛瓦尔定理 (傅里叶级数))**.** *在 $[0,T]$ 区间内，由 Theorem 7.63可得*

$$
\int_0^T f^*\left(t\right)g\left(t\right)\,\mathrm{d}t = T\sum_{\nu=-\infty}^{\infty}\tilde{f}_\nu^*\tilde{g}_\nu.
\tag{7.189}
$$

Theorem 7.66 (普朗歇尔定理 (傅里叶级数))**.** *在 $[0,T]$ 区间内，由 Theorem 7.64可得*

$$
\int_0^T \left|f(t)\right|^2\,\mathrm{d}t = T\sum_{\nu=-\infty}^{\infty}\left|\tilde{f}_\nu\right|^2.
\tag{7.190}
$$

根据帕塞瓦尔定理和普朗歇尔定理，在数值计算过程中，我们既可以求解实区间中的原始方程，也可以求解傅里叶空间中的变换方程，二者等价。实际求解哪一种取决于哪个更容易计算。

7.G.6　泊松求和式

采用类似于帕塞瓦尔定理的思路，也有另一种在实域或傅里叶域近似求积的方法，称泊松求和式 (Poisson summation formula)。

Theorem 7.67. *具体说来，假定关于方程 $f(t)$，我们想要求得一组等宽分布的时间点 (对应宽度 Δt) 下 $f(t)$ 值的和。根据泊松求和式，我们也可以计算傅里叶域中一组等宽分布 (对应宽度 $\frac{2\pi}{\Delta t}$) 的 $\tilde{f}\left(\omega\right)$ 的值，两种方法等价。*

$$
\sum_{n=-\infty}^{\infty} f\left(n\Delta t\right) = \frac{2\pi}{\Delta t}\sum_{\nu=-\infty}^{\infty}\tilde{f}\left(\frac{2\pi\nu}{\Delta t}\right).
\tag{7.191}
$$

证明. 用傅里叶综合表达式改写(7.191)LHS

$$\begin{aligned}\sum_{n=-\infty}^{\infty} f(n\Delta t) &= \sum_{n=-\infty}^{\infty}\left[\int_{-\infty}^{\infty}\tilde{f}(\omega)\exp(in\omega\Delta t)\,\mathrm{d}\omega\right]\\ &= \int_{-\infty}^{\infty}\tilde{f}(\omega)\underbrace{\sum_{n=-\infty}^{\infty}\exp(in\omega\Delta t)}_{=2\pi\sum_{\nu}\delta(\omega\Delta t-2\nu\pi)}\,\mathrm{d}\omega\\ &= \int_{-\infty}^{\infty}\tilde{f}(\omega)\,2\pi\underbrace{\sum_{\nu}\delta(\omega\Delta t-2\nu\pi)}_{=:\mathcal{A}}\,\mathrm{d}\omega,\end{aligned}$$

最后一行等式中，将 $f(n\Delta t)$ 理解为一个关于 $\omega\Delta t-2\nu\pi$ 的狄拉克方程，可分两种情况进一步分析

- 如果 $\omega\Delta t$ 是 2π 的整数倍 ($\nu\in\mathcal{Z}$ 可以是任意整数)，那么 $\sum_{n=-\infty}^{\infty} f(n\Delta t)=\infty$。
- 否则，$\mathcal{A}$ 就是一组狄拉克方程 $\delta(\omega\Delta t-2\nu\pi)$ 之和，每个狄拉克方程对应不同的 ν 值；进而 $\sum_{n=-\infty}^{\infty} f(n\Delta t)=0$，各项之间相互抵消。

由狄拉克方程的性质

$$\delta(ax-b)=\frac{1}{a}\delta\left(x-\frac{b}{a}\right)$$

可进一步得到

$$\begin{aligned}\sum_{n=-\infty}^{\infty} f(n\Delta t) &= 2\pi\sum_{\nu}\int_{-\infty}^{\infty}\tilde{f}(\omega)\,\delta(\omega\Delta t-2\nu\pi)\,\mathrm{d}\omega\\ &= \frac{2\pi}{\Delta t}\sum_{\nu}\int_{-\infty}^{\infty}\tilde{f}(\omega)\,\delta\left(\omega-\frac{2\nu\pi}{\Delta t}\right)\mathrm{d}\omega\\ &= \frac{2\pi}{\Delta t}\sum_{\nu=-\infty}^{\infty}\tilde{f}\left(\nu\frac{2\pi}{\Delta t}\right),\end{aligned}$$

证毕。 □

两个值域中各有一组点，两组点的单位相同[39]。为了简化表达，设 $\Delta\omega\equiv\frac{2\pi}{\Delta t}$，则(7.191)改写为

$$\sum_{n=-\infty}^{\infty} f(n\Delta t)=\Delta\omega\sum_{\nu=-\infty}^{\infty}\tilde{f}(\nu\cdot\Delta\omega), \tag{7.192}$$

值得留意的是 Δt 和 $\Delta\omega$ 呈反比例，即，在原时间域中，若是对原方程 $f(t)$ 取密集的配点求 $\sum f(\cdot)$，对应较小的 Δt 值，那么在对应的傅里叶域中，就涉及到较高的频率值 $\Delta\omega$——换句话说，原域中的配点越是密，在稀疏频率域中对应的点就越是稀疏。

随着 $f(\cdot)$ 的形式不同，泊松求和式的一些具体应用如下。

高斯方程的泊松求和

在前面介绍过，一个高斯方程的傅里叶变换也是一个高斯方程 (第7.G.2节)。结合本节的知识我们可以作进一步判断：对高斯方程形式原方程配点越宽 (如，实域宽度 σ)，其变换后的高斯形式方程越窄 (如，傅里叶域对应的宽度 $\frac{2}{\sigma}$)。

[39] $\tilde{f}$ 的单位 $=f$ 的单位/频率 $=f$ 的单位 $\times$(时间)，其中时间用 Δt 来表示。因此 $\tilde{f}$ 的单位/Δt=f 的单位。

因此，设一个以 x 为参数的关于 t 的高斯方程 $T_x(t)=\exp\left(-t^2\pi x\right)$，要求 $\sum T_x(t)$。在 $\sum T_x(t)$ 的收敛速度较慢的情况下，我们可以利用泊松求和式，在傅里叶空间中进行求和计算[40]。具体说来，$T_x(t)$ 的傅里叶变换

$$\begin{aligned}\widetilde{T}_x(\omega)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-i\omega t\right)T_x(t)\,\mathrm{d}t\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(-i\omega t-t^2\pi x\right)\,\mathrm{d}t\\&=\frac{1}{2\pi\sqrt{x}}\exp\left(-\frac{\omega^2}{4\pi x}\right),\end{aligned}$$

雅各比方程

雅各比方程族 (Jacobian θ function)中的一种形式可以表示为

$$\theta(x)=\sum_{n=-\infty}^{\infty}\exp\left(-n^2\pi x\right)=\sum_{n=-\infty}^{\infty}T_x(n). \tag{7.193}$$

利用泊松求和式(7.191)，以及设 $\Delta t=1$，有

$$\theta(x)=2\pi\sum_{\nu=-\infty}^{\infty}\widetilde{T}\left(2\nu\pi\right)=\frac{1}{\sqrt{x}}\underbrace{\sum_{\nu=-\infty}^{\infty}\exp\left(-\frac{\pi\nu^2}{x}\right)}_{=\theta\left(\frac{1}{x}\right)}, \tag{7.194}$$

即雅各比方程(7.193)的函数方程 (functional equation) 形式

$$\begin{aligned}&\theta\left(x\right)=\frac{1}{\sqrt{x}}\theta\left(\frac{1}{x}\right),\\\Leftrightarrow\ &\sum_{n=-\infty}^{\infty}\exp\left(-n^2\pi x\right)=x^{-\frac{1}{2}}\sum_{\nu=-\infty}^{\infty}\exp\left(-\frac{\nu^2\pi}{x}\right).\end{aligned} \tag{7.195}$$

为了说明为何傅里叶级数求和的形式即(7.195)RHS 更适于进行数值计算，我们做了一个小实验，设 $x=0.04$，$n=0,1,\ldots,11$。不难看出，LHS 需要 $n=11$，前后 23 项的求和才能达到小数点后 6 位的精度；而对于雅各比方程形式的傅里叶级数求和，RHS 只需要 $n=1$ 就能达到同样的精度，换句话说，由(7.195)我们有

$$\underbrace{\theta\left(0.04\right)}_{4.999999971,n=11}=\underbrace{\frac{1}{\sqrt{0.04}}}_{5.0}\underbrace{\theta\left(\frac{1}{0.04}\right)}_{1.000000000,\nu=1}$$

7.G.7　傅里叶分析与卷积

傅里叶分析的又一个重要特征是，它是以卷积 (convolution)乘的形式出现的。如上文所示，由两个方程 $f(t)$ 和 $g(t)$ 构成的卷积，近似表示为一系列 $g(\tau)$ 的加权和，$g(\tau)$ 是对 $g(t)$ 的近似 (copy)之一，对应权重 $f(\tau)$，我们设这个卷积为 $C(t)$

$$C(t)\equiv f\times g=\int_{-\infty}^{\infty}f(\tau)g(t-\tau)d\tau.$$

对 $C(t)$ 作傅里叶变换，有

$$\begin{aligned}\widetilde{C}(\omega)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}C(t)\exp\left(-i\omega t\right)\,\mathrm{d}\omega\\&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\tau)g(t-\tau)\,\mathrm{d}\omega\exp\left(-i\omega t\right)\,\mathrm{d}t.\end{aligned}$$

[40]快速收敛的原因可能如，πx 的值较小，从而导致 $\sum T_x(t)$ 收敛缓慢。

表 7.G.3: 雅各比方程的泊松求和

	LHS		RHS	
$n(\nu)$	$\exp\left(-n^2\pi x\right)$	$\sum_n \exp\left(-n^2\pi x\right)$	$\exp\left(-\nu^2\pi/x\right)$	$\sum_\nu \exp\left(-\nu^2\pi/x\right)$
0	1.000000000	1.000000000	1.000000000	5.000000000
1	0.881911378	2.763822757	7.77304E-35	5.000000000
2	0.604922563	3.973667882	3.6506E-137	5.000000000
3	0.322718983	4.619105849	1.0359E-307	5.000000000
4	0.133905721	4.886917291	0	5.000000000
5	0.043213918	4.973345128	0	5.000000000
6	0.010846711	4.995038549	0	5.000000000
7	0.002117495	4.999273539	0	5.000000000
8	0.000321512	4.999916562	0	5.000000000
9	3.79683E-05	4.999992498	0	5.000000000
10	3.48734E-06	4.999999473	0	5.000000000
11	2.49126E-07	4.999999971	0	5.000000000
x=	0.04			

注：基于式(7.195)计算。excel 表格见文件夹中 data/20180306-poisson.xlsx .

将上式中的 $f(\tau)$ 和 $g(t-\tau)$ 替换为傅里叶综合形式

$$\begin{aligned}
\widetilde{C}(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\exp(i\omega_1\tau)\,\tilde{f}(\omega_1)\,\mathrm{d}\omega_1\right\}\left\{\int_{-\infty}^{\infty}\exp[i\omega_2(t-\tau)]\,\tilde{g}(\omega_2)\,\mathrm{d}\omega_2\right\}\exp(-i\omega t)\,\mathrm{d}\tau\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\exp[i(\omega_1-\omega_2)\tau]\,\mathrm{d}\tau\right\}\left\{\int_{-\infty}^{\infty}\exp[i(\omega_2-\omega_2)t]\,\mathrm{d}t\right\}\tilde{f}(\omega_1)\,\tilde{g}(\omega_2)\,\mathrm{d}\omega_1\,\mathrm{d}\omega_2 \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[2\pi\delta(\omega_1-\omega_2)]\,[2\pi\delta(\omega_2-\omega)]\,\tilde{f}(\omega_1)\,\tilde{g}(\omega_2)\,\mathrm{d}\omega_1\,\mathrm{d}\omega_2 \\
&= 2\pi\tilde{f}(\omega)\,\tilde{g}(\omega),
\end{aligned}$$

上式告诉我们，f 和 g 的卷积的频率为 ω 的傅里叶变换系数 $\widetilde{C}(\omega)$，就等于 f 和 g 分别的 ω 频率傅里叶系数之积。这意味着在频率域中进行卷积运算，要比在原始实值域中做卷积运算更方便一些。

7.G.8　高维傅里叶变换

更高维度傅里叶变换的方法，与一维傅里叶变换类似。例如对二维方程 $f(x,y)$ 而言，第一步保持 x 不变，做关于 y 的傅里叶变换，可得一个混合了实空间和傅里叶空间的方程 $\tilde{f}(x,k_y)$

$$\tilde{f}(x,k_y) = \frac{1}{2\pi}\int_{\infty}^{\infty}\exp(-ik_y y)\,f(x,y)\,\mathrm{d}y.$$

第二步保持 k_y 不变，再对 $\tilde{f}(x,k_y)$ 做关于 x 的傅里叶变换，得到 $\tilde{f}(k_x,k_y)$

$$\begin{aligned}
\widetilde{f}(k_x,k_y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-ik_x x)\,\tilde{f}(x,k_y)\,\mathrm{d}x \\
&= (2\pi)^{-2}\int_{-\infty}^{\infty}\exp(-ik_x x)\int_{-\infty}^{\infty}\exp(-ik_y y)\,f(x,y)\,\mathrm{d}y\,\mathrm{d}x \\
&= (2\pi)^{-2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp[-i(k_x x+k_y y)]\,f(x,y)\,\mathrm{d}y\,\mathrm{d}x \\
&= (2\pi)^{-2}\int\exp(-i\cdot\boldsymbol{k}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},
\end{aligned}$$

其中积分符号表示解释变量 $\boldsymbol{x}$ 的全部取值区间。

D 维方程也可用类似地方式求得，前缀系数此时变为 $(2\pi)^{-D}$。

7.G.9　一些指数求和的背景知识

连续、有限时间域

假定一个连续、有限的时间区间，宽度为 T，那么方程 $f(t)$ 在傅里叶域中的基准频率为 $\omega_0 = \frac{2\pi}{T}$。进而我们有

$$\frac{1}{T}\int_0^T\exp(n_1-n_2)\,\omega_0 t\,\mathrm{d}t = \delta(n_1,n_2), \tag{7.196}$$

其中 $\delta(n_1,n_2)$ 表示克罗内克乘积，见(7.31)。

正交性

(7.196)中，如果 $n_1 \neq n_2$，那么我们可以说，方程 $f_{n_1}(t) = \exp(in_1\omega_0 t)$ 和 $f_{n_2}(t) = \exp(in_2\omega_0 t)$ 在内积的意义上正交 (orthogonal)，可表示为

$$\langle f,g\rangle = \frac{1}{T}\int f^* g\,\mathrm{d}t.$$

更多正交性条件的介绍，见第7.C节。

连续、无限时间域

在连续、无限时间域中，(7.196)就变成了

$$\int_{-\infty}^{\infty}\left[i\left(\omega-\omega^{'}\right)t\right]\,\mathrm{d}t=2\pi\delta\left(\omega-\omega^{'}\right). \tag{7.197}$$

离散时间域

上文在第7.G.7节做泊松求和时，用到了(7.197)，以处理连续、无限时间域中的求积问题。对应的离散傅里叶变换版本可写作

$$\sum_{n=-\infty}^{\infty}\exp\left(inkx\right)=2\pi\sum_{\nu=-\infty}^{\infty}\delta\left(xk-2\nu\pi\right),$$

RHS 中，如果 x 等于 $\frac{2\pi}{k}$ 的整数倍，那么 LHS 是 ∞：$\exp\left(inkx\right)=1\,\forall n$，$\sum_n 1=\infty$。如果不是整数倍，$LHS$ 的无数个 $\exp\left(inkx\right)$ 相互抵消，LHS $=0$。

来看是整数倍，进而 LHS$=\infty$ 时的情况。若将左侧进一步扩展为 $\int f(x)\exp\left(inkx\right)\,\mathrm{d}x$(对 x 求积分)，那么我们将得到一个 $<\infty$ 的数值，数值的大小取决于 $x=\frac{2\nu\pi}{k}$ 时对应的 $f(x)$ 的值。

高斯求积

第7.G.2节介绍高斯变换时，使用了一个较为复杂的高斯方程。高斯方程族中，一个基本的高斯求积问题可以表示为

$$\int_{\infty}^{\infty}\exp\left(-x^2\right)\,\mathrm{d}x=\sqrt{\pi}. \tag{7.198}$$

稍稍做个扩展，在指数项中放入一个系数 α，则(7.198)调整为

$$\int_{-\infty}^{\infty}\exp\left(-\alpha x^2\right)\,\mathrm{d}x=\sqrt{\frac{\pi}{\alpha}}. \tag{7.199}$$

再继续做扩展，一个高斯求积

$$I\left(\alpha,\beta,\gamma\right)=\int_{-\infty}^{\infty}\exp\left(-\alpha x^2+\beta x+\gamma\right)\,\mathrm{d}x,$$

求解方法如下：

- 提取出常数项 γ

$$\begin{aligned}I\left(\alpha,\beta,\gamma\right)&=\exp\left(\gamma\right)\int_{-\infty}^{\infty}\exp\left(\underbrace{-\alpha x^2+\beta x}_{=-\alpha\left(x-\frac{\beta}{2\alpha}\right)^2+\frac{\beta^2}{4\alpha}}\right)\,\mathrm{d}x\\&=\exp\left(\gamma+\frac{\beta^2}{4\alpha}\right)\int_{-\infty}^{\infty}\exp\left(-\alpha\left[x-\frac{\beta}{2\alpha}\right)^2\right]\,\mathrm{d}x.\end{aligned}$$

- 设 $y\equiv x-\frac{\beta}{2\alpha}$。上式变为

$$\begin{aligned}I\left(\alpha,\beta,\gamma\right)&=\exp\left(\gamma+\frac{\beta^2}{4\alpha}\right)\underbrace{\int_{-\infty}^{\infty}\exp\left(-\alpha y^2\right)\,\mathrm{d}y}_{=\sqrt{\frac{\pi}{\alpha}}}\\&=\sqrt{\frac{\pi}{\alpha}}\exp\left(\gamma+\frac{\beta^2}{4\alpha}\right).\end{aligned}$$

第八章　有界元法和有限元法

8.1　边界值问题：位势方程

我们从二阶偏微分方程入手，介绍边界值问题 (boundary value problem)。一个合适的例子是位势方程 (potential equation)。

8.1.1　偏微分算子及椭圆边界值问题

定义有界域 $\Omega \in \mathbb{R}^d, d = 2, 3$，边界 $\Gamma = \partial\Omega$，外代数单位向量空间 (exterior unit normal vector) $\underline{n}(x)$ 对于 $x \in \Gamma$ 几乎处处存在。对于 $x \in \Omega$，我们考虑一个线性二阶偏微分的自伴随算子[1] (self-adjoint operator) L ，作用于实值标量方程 u

$$(L\,u)\,(x) \coloneqq -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_j}\left[a_{ji}(x)\frac{\partial}{\partial x_i}u(x)\right] + a_0(x)\,u(x), \tag{8.1}$$

其中 $a_{ji}(x)$, $i, j = 1, \ldots, d$, $x \in \Omega$ 表示系数方程，假定为平滑的并满足 $a_{ij}(x) = a_{ji}(x)$。由此可以构建一个对称的系数矩阵 $A(x)$，满足

$$A(x) = (a_{ij}(x))_{i,j=1}^{d}, \quad x \in \Omega,$$

对应实数特征根 $\lambda_k(x)$。

当且仅当 $\lambda_k(x) > 0$ 对于所有 $k = 1, \ldots, d$ 都成立时，我们称偏微分算子 L 在某一个 $x \in \Omega$ 上是椭圆 (elliptic) 的。

更进一步，如果 $\forall x \in \Omega$ 该条件都成立，那么我们称 L 在 Ω 上是椭圆的。

如果存在一个一致下界 (uniform lower bound) $\lambda_0 > 0$，满足

$$\lambda_k(x) \geq \lambda_0, \quad \forall k = 1, \ldots, d, \forall x \in \Omega, \tag{8.2}$$

那么我们称 L 在 Ω 上一致椭圆 (uniformly elliptic)。

8.1.2　边界条件

边界条件的分析，可以从散度定理开始。

Theorem 8.1 (散度定理)**.** *散度定理 (divergence theorem)，又称奥斯特罗格拉德斯基——高斯定理 (Ostrogradsky-Gauss theorem) 、高斯散度定理 (Gauss' theorem)等，是指*

$$\int_{\Omega} \frac{\partial}{\partial x_i} f(x)\,dx = \int_{\Gamma} \left[\gamma_0^{int} f(x)\right] n_i(x)\,ds_x, \quad i = 1, \ldots, d, \tag{8.3}$$

[1]有限维内积向量空间 V 中，自伴随算子 A 是一个从 V 到 V 自身的线性映射 $\langle A\boldsymbol{u}, \boldsymbol{\nu}\rangle = \langle \boldsymbol{\nu}, A\boldsymbol{u}\rangle, \forall \nu, w \in V$。

其中 $\gamma_0^{int} f(x)$ 是某个给定方程 $f(x), x \in \Omega$ 的内界迹 *(interior boundary trace, Theorem 8.10)*，满足

$$\gamma_0^{int} f(x) := \lim_{\Omega \ni \tilde{x} \mapsto x \in \Gamma} f(\tilde{x}), \quad \forall x \in \Gamma = \partial\Omega. \tag{8.4}$$

假定两个足够光滑的方程 $u, \nu \in \Omega$，通过设定 $f(x) = u(x)\,\nu(x)$，可以将散度定理(8.3)改写为分部积分 (integration by parts)的形式

$$\int_\Omega u(x) \frac{\partial}{\partial x_i} \nu(x)\, x + \int_\Omega \nu(x) \frac{\partial}{\partial x_i} u(x)\, x = \int_\Gamma \left[\gamma_0^{int} u(x)\right] \left[\gamma_0^{int} \nu(x)\right] n_i(x)\, ds_x.$$

重新调整上式，将 $\nu(x)$ 视作检测方程 (test function)，两侧乘以(8.1)中的二阶偏微分算子 $(L\,u)(x)$，在 Ω 中求积

$$\begin{aligned} (L\,u)(x)\,\nu(x) &:= -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_j}\left[a_{ji}(x) \frac{\partial}{\partial x_i} u(x)\right] \nu(x) + a_0(x)\, \underbrace{u(x)\,\nu(x)}_{\equiv f(x)}, \\ \hookrightarrow \int_\Omega (L\,u)(x)\,\nu(x)\, dx &= -\sum_{i,j=1}^{d} \int_\Omega \frac{\partial}{\partial x_j}\left[a_{ji}(x) \frac{\partial}{\partial x_i} u(x)\right] \nu(x)\, dx, \end{aligned} \tag{8.5}$$

使用分部积分 $\hookrightarrow$

$$\begin{aligned} \int_\Omega (L\,u)(x)\,\nu(x)\, dx = &\underbrace{\sum_{i,j=1}^{d} \int_\Omega a_{ji}(x) \frac{\partial}{\partial x_i} u(x)\, \frac{\partial}{\partial x_j} \nu(x)\, dx}_{:= a(u,\nu)} \\ &- \sum_{i,j=1}^{d} \int_\Gamma n_j(x) \left[\gamma_0^{int}(x) \left(a_{ji}(x) \frac{\partial}{\partial x_i} u(x)\right)\right] \left[\gamma_0^{int} \nu(x)\right]\, ds_x, \end{aligned}$$

由此，我们由散度定理 (Theorem 8.1) 推导出格林第一恒等式 (Green's first identity)

$$\begin{aligned} a(u,\nu) &:= \sum_{i,j=1}^{d} \int_\Omega a_{ji}(x) \frac{\partial}{\partial x_i} u(x)\, \frac{\partial}{\partial x_j} \nu(x)\, dx \\ &= \int_\Omega (L\,u)(x)\,\nu(x)\, dx + \sum_{i,j=1}^{d} \int_\Gamma \underbrace{n_j(x) \left[\gamma_0^{\text{int}}(x) \left(a_{ji}(x) \frac{\partial}{\partial x_i} u(x)\right)\right]} \left[\gamma_0^{\text{int}} \nu(x)\right]\, ds_x, \\ &= \int_\Omega (L\,u)(x)\,\nu(x)\, dx + \int_\Gamma \underbrace{\left[\gamma_1^{\text{int}} u(x)\right]} \left[\gamma_0^{\text{int}} \nu(x)\right]\, ds_x, \end{aligned} \tag{8.6}$$

其中定义 γ_1^{int} 为内部共形导数 (interior co-normal derivative, Mikhailov (2006, 2009); Ancona (2009))

$$\gamma_1^{int} u(x) := \lim_{\Omega \ni \tilde{x} \mapsto x \in \Gamma} \left[\sum_{i,j=1}^{d} n_j(x) a_{ji}(\tilde{x}) \frac{\partial}{\partial \tilde{x}_i} u(\tilde{x})\right], \quad x \in \Gamma. \tag{8.7}$$

将格林第一恒等式(8.6)中的 u, ν 互换位置，我们有

$$a(\nu, u) = \int_\Omega (L\,\nu)(x)\, u(x)\, dx + \int_\Gamma \left[\gamma_1^{int} \nu(x)\right] \left[\gamma_0^{int} u(x)\right]\, ds_x,$$

由上式和(8.6)我们有格林第二恒等式 (Green's second identity): $a(u,\nu) = a(\nu,u) \leftrightarrow$

$$\begin{aligned} &\int_\Omega (L\,u)(x)\,\nu(x)\, dx + \int_\Gamma \left[\gamma_1^{int} u(x)\right] \left[\gamma_0^{int} \nu(x)\right]\, ds_x \\ &= \int_\Omega (L\,\nu)(x)\, u(x)\, dx + \int_\Gamma \left[\gamma_1^{int} \nu(x)\right] \left[\gamma_0^{int} u(x)\right]\, ds_x, \quad \forall\, u, \nu \in \Omega, \text{且} u, \nu \text{足够平滑}. \end{aligned} \tag{8.8}$$

下面来考虑一个特殊情况，$a_{ij}(x)=\delta_{ij}$，δ_{ij} 是克罗内克乘积 (Kronecker product)，见(7.31)。(8.1)的二阶偏微分算子 $(L\,u)\,(x)$ 变为拉普拉斯算子

$$(L\,u)\,(x)=-\Delta u(x):=-\sum_{i=1}^{d}\frac{\partial^2}{\partial x_i^2}u(x),\quad x\in\mathbb{R}^d. \tag{8.9}$$

内部共形导数 γ_1^{int} (8.7)变为

$$\gamma_1^{int}u(x)=\frac{\partial}{\partial n_x}u(x):=\underline{n}(x)\bigtriangledown u(x),\quad x\in\Gamma. \tag{8.10}$$

对边界域 $\Gamma=\partial\Omega$ 分解成三个不相交集合的并集 (disjoint union)

$$\Gamma=\overline{\Gamma}_D\cup\overline{\Gamma}_N\cup\overline{\Gamma}_R,$$

对应地，边界值问题变为两部分：第一部分，在 Ω 中，基于给定的方程 $f(x)$，寻找偏微分算子 $(Lu)(x)$，使得

$$(L\,u)\,(x)=f(x),\quad x\in\Omega. \tag{8.11}$$

第二部分，在 Γ 中，基于给定的方程 $g(x)$，寻找内界迹 $\gamma_0^{int}u(x)$ 或者内共形导数 $\gamma_1^{int}(x)$。随着 Γ 的取值范围不同，分为三种情况：

$$\gamma_0^{int}u(x)=g_D(x),\quad x\in\Gamma=\Gamma_D, \tag{8.12a}$$

$$\gamma_1^{int}u(x)=g_N(x),\quad x\in\Gamma=\Gamma_N, \tag{8.12b}$$

$$\kappa(x)\,\gamma_0^{int}u(x)+\gamma_1^{int}u(x)=g_R(x),\quad x\in\Gamma=\Gamma_R. \tag{8.12c}$$

Definition 8.1 (边界值条件)**.** *于是我们有以下几种不同的边界值条件：*

- $\Gamma=\Gamma_D$: (8.11) + (8.12a) → *狄利克雷边界值条件 (Dirichlet boundary value condition),*
- $\Gamma=\Gamma_N$: (8.11) + (8.12b) → *诺依曼边界值条件 (Neumann boundary value condition),*
- $\Gamma=\Gamma_R$: (8.11) + (8.12c) → *罗宾边界值条件 (Robin boundary value condition),*
- *混合边界值条件，以上几种情况的组合。*

有时候我们还需要将线性罗宾边界值条件扩展为非线性的情况，(8.12c) →

$$G\left(\gamma_0^{int}u(x),x\right)+\gamma_1^{int}u(x)=g_R(x),\quad x\in\Gamma=\Gamma_R, \tag{8.13}$$

其中 $G(u,\cdot)$ 是某个给定的非线性方程，如 $u(x)^3$。

对于边界值问题的解 $u(x)$，还需要注意以下几点

1. $u(x)$ 的存在性和唯一性，相关讨论可参考如Ladyzhenskaya (1968)，
2. 观测到的数据需要是充分平滑的，以确保 $u(x)$ 充分可微 (sufficiently differentiable)

$$u\in C^2(\Omega)\cap C^1\left(\Omega\cup\Gamma_N\cup\Gamma_R\right)\cap C(\Omega\cup\Gamma_D).$$

8.1.3 诺依曼边界值问题

对于诺依曼边界值条件的解，其存在性和唯一性需要做进一步讨论。

假定 $\nu_1(x) = 1, x \in \Omega$ 是关于 $\nu_1(x)$ 的齐次诺依曼边界值问题的一个解，(8.9)、(8.10) $\Rightarrow$ 关于 $u(x)$ 的诺依曼边界值问题

$$\begin{aligned} (L\,\nu_1)\,(x) &= 0, \quad x \in \Omega, \\ \gamma_1^{int}\nu_1(x) &= 0, \quad x \in \Gamma. \end{aligned} \tag{8.14}$$

(8.14) 代入格林第二恒等式(8.8)可得正交条件

$$\int_\Omega (L\,u)\,(x)\,dx + \int_\Gamma \gamma_1^{int}u(x)\,ds_x = 0, \tag{8.15}$$

诺依曼边界值条件(8.11)、(8.12b)$\Rightarrow$

$$\begin{aligned} (Lu)\,(x) &= f(x), \quad x \in \Omega, \\ \gamma_1^{int}u(x) &= g_N(x), \quad x \in \Gamma. \end{aligned} \tag{8.16}$$

对于给定的 f 和 g_N，根据正交条件(8.15)，我们可以假设诺依曼问题的可解性条件 (solvability condition)

$$\int_\Omega f(x)\,dx + \int_\Gamma g_N(x)\,ds_x = 0. \tag{8.17}$$

换句话说，如果关于 $\nu(x)$ 的齐次诺依曼边界问题解是 $\nu_1(x) = 1, x \in \Omega$，那么关于 u 的诺依曼边界值问题(8.16)的解并不唯一：不只包括一个解 $u(x)$，还包括另一个解 $\tilde{u}(x)$，满足关系

$$\tilde{u}(x) = u(x) + \alpha, \quad x \in \Omega,$$

其中常数 $\alpha \in \mathbb{R}$ 的值是唯一的，取决于为了使第一个解 $u(x)$ 成为诺依曼边界值问题(8.16)的解，而需要在系统中加入的规模调整条件，如

$$\int_\Omega u(x)\,dx = 0, \quad \text{或者} \quad \int_\Gamma \gamma_0^{int}u(x)\,ds_x = 0.$$

8.2 方程空间

在进一步介绍边界值问题的弱形式之前，一些与之紧密相关的方程空间的知识是必需的。相关教材，可参考如McLean (2000); Adolfson et al. (2008b); Tartar (2007); Mazya (2009a,b)。

8.2.1 $C^k(\Omega), C^{k,\kappa}(\Omega)$ 空间

给定 $d \in \mathbb{N}$。作如下定义：

- 向量 (vector) $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)\,, \alpha_i \in \mathbb{N}_0$。
- 多重指标 (multi-index) 的绝对值 $|\alpha| = \sum_{i=1}^d \alpha_i$。
- 阶乘 (factorial) $\alpha! = \alpha_1!\,\alpha_2!\,\ldots\alpha_d!$。

给定 $x \in \mathbb{R}^d$ 我们有

$$x^{|\alpha|} = x_1^{\alpha_1}\,x_2^{\alpha_2}\ldots x_d^{\alpha_d}.$$

给定一个充分平滑的实值方程 u，其相对于 x 的 α 阶偏微分导数

$$D^\alpha u(x) := \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1}\left(\frac{\partial}{\partial x_2}\right)^{\alpha_2}\cdots\left(\frac{\partial}{\partial x_d}\right)^{\alpha_d} u\left(x_1, x_2, \ldots, x_d\right).$$

给定一个开放子集 $\Omega \subseteq R^d$，对于某个标量 $k \in \mathbb{N}_0$。则 $C^k(\Omega)$ 表示在 Ω 域中有界且 k 次连续可导的方程空间。对于某个方程 $u \in \Omega$，u 的范数 (norm)值是有限的

$$\|u\|_{C^k(\Omega)} := \sum_{|\alpha|\le k} \sup_{x\in\Omega} \left|D^\alpha u(x)\right| < \infty,$$

随着 $k \to \infty$，$C^\infty(\Omega)$ 是个有界且无限阶连续可积的方程空间。

对于方程 $u(x), x \in \Omega$，我们将 u 的支撑 (support) 定义为 $\operatorname{supp} u$

$$\operatorname{supp} u := \overline{x \in \Omega : u(x) \neq 0}.$$

进而定义 $C_0^\infty(\Omega)$ 为 $C^\infty(\Omega)$ 中的紧支撑 (compact support) 方程空间。

定义 $C^{k,\kappa}, k \in \mathbb{N}_0, \kappa \in (0,1)$ 为霍德尔连续方程空间 (Hölder continuous function space)，对应范数为

$$\|u\|_{C^{k,\kappa}(\Omega)} := \|u\|_{C^k(\Omega)} + \sum_{|\alpha|=k} \sup_{x,y\in\Omega, x\neq y} \frac{|D^\alpha u(x) - D^\alpha u(y)|}{|x-y|^\kappa}$$

当 $\kappa = 1$ 时，$C^{k,1}$ 用来表示第 $|\alpha| = k$ 次偏导数 $D^\alpha u(x)$ 是利普希茨连续方程 (Lipschitz continuous)的方程 $u \in C^k(\Omega)$ 所组成的空间。

我们用 Γ 来表示开放集 $\Omega \subset \mathbb{R}^d$ 的边界

$$\Gamma := \partial\Omega = \bar{\Omega} \cup \left(\mathbb{R}^d \backslash \Omega\right).$$

当 $d \ge 2$，$\Gamma = \partial\Omega$ 可以视作利普希茨方程的局部图，随着所处在 Γ 中不同的位置，对应不同的笛卡尔坐标系。一个最简单的例子是假定一个利普希茨方程 $\gamma : \mathbb{R}^{d-1} \mapsto \mathbb{R}$，满足

$$\Omega := \left\{x \in \mathbb{R}^d : x_d < \gamma(\tilde{x}), \quad \forall \tilde{x} = (x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1}\right\}.$$

由利普希茨方程的性质我们有，$\gamma(\cdot)$ 满足

$$\left|\gamma(\tilde{x}) - \gamma(\tilde{y})\right| \le L\left|\tilde{x} - \tilde{y}\right|, \quad \forall \tilde{x}, \tilde{y} \in \mathbb{R}^{d-1},$$

那么 Ω 被称为一个利普希茨亚图 (Lipschitz hypograph)，对应边界 Γ

$$\Gamma = \left\{x \in \mathbb{R}^d : x_n = \gamma(\tilde{x}), \quad \forall \tilde{x} \in \mathbb{R}^{d-1}\right\}.$$

Definition 8.2 (利普希茨域)**.** *可参考 Heinonen (2005)。某一个开集 $\Omega \subset \mathbb{R}^d, d \ge 2$ 在满足如下条件时，被称为利普希茨域 (Lipschitz domain)，或有利普希茨边界的域 (domain with Lipschitz boundary)：*

- *Ω 的边界 $\Gamma = \partial\Omega$ 是紧凑的，并且*
- *$\exists$ 有限的索引族 (index family) $\{W_j\}$ 和 $\{\Omega_j\}$，满足* [2]

[2]对于集合 I 和 S，某个方程 x

$$x : I \mapsto S$$
$$i \mapsto x_i = x(i)$$

被称作用 I 索引的 S 中元素的族 (family of elements)，也表示为 $\{x_i\}_{i\in I}$，$x_i \subset S$。

- 索引族 $\{W_j\}$ 是对 Γ 的有限开覆盖 *(finite open cover)*

$$W_j \in \mathbb{R}^d, \quad \& \Gamma \subseteq \bigcup_j W_j.$$

- 每一个索引族 $\Omega_j, \forall j$ 都可以通过一定的操作变换 *(transformation)* 为利普希茨亚图，如旋转 *(rotation)* 和平移 *(translation)* 等。
- $W_j \cap \Omega = W_j \cap \Omega_j, \quad \forall j$。

需要注意的是，利普希茨边界 $\Gamma = \partial\Omega$ 作局部表达的方案，即对 W_j 和 Ω_j 的选取，通常来讲并不是唯一的[3]。

如果方程 $\gamma(\cdot)$ 中参数的选取使得满足 $\gamma \in C^k(\mathbb{R}^{d-1})$，我们称这个利普希茨边界 Γ 是 k 次可微的；如果满足 $\gamma \in C^{k,\kappa}(\mathbb{R}^{d-1})$，我们称 Γ 为霍德尔连续 (Hölder continuous)；如果 γ 仅仅是在某个局域内满足该条件，则我们称对应的 Γ 为分段平滑边界 (piecewise smooth boundary)。

8.2.2 勒贝格 $L^p(\Omega)$ 空间

数学上 $L^p(\Omega)$ 空间 ($L^p(\Omega)$ space) 又称勒贝格空间 (Lebesgue space)，指 Ω 上一组测度方程 (measurable function) 的等价类的集合，这些测度方程都是 p 次勒贝格可积方程 (Lebesgue integrable function, Definition 8.3) [4]。

Definition 8.3 (勒贝格可积方程)**.** 勒贝格可积方程 *(Lebesgue integrable function)* 是指该方程的绝对值的 p 次幂的积分是有限的，如

$$\int_\Omega |u(x)|^p \, dx < \infty.$$

σ 代数

Definition 8.4 (sigma 代数)**.** 设 S 是一个非空集合，另有一个集合 Σ 中的所有元素都是 S 的子集，那么我们将满足以下条件的 Σ 成为 S 上的一个 σ 代数 *(σ-algebra) (Bogachev, 2007a,b, p.4)*

- S 在 Σ 中

$$S \in \Sigma,$$

- 如果一个集合 A 在 Σ 中，那么它的补集 *(complement)* 也在 Σ 中

$$A \in \Sigma \Rightarrow A^{\complement} \in \Sigma,$$

- 如果 n 个集合 $A_1, \ldots A_n$ 都在 Σ 中，那么他们的并集 *(union)* 也在 Σ 中

$$(A_n \in \Sigma, \quad \forall n \in \mathbb{N}) \Rightarrow \bigcup_{i=1}^{n} A_i \in \Sigma.$$

Definition 8.5 (幂集)**.** 对于任一集合 S 的幂集 *(powerset)*是指这样的一个集合，包括空集 $\varnothing$、S 本身和 S 的所有集合，常常表示为 $\mathcal{P}(S)$ 或 2^S。

在公理集合论 *(axiomatic set theory)*例如 *ZFC* 集合论 *(ZFC axioms)*假定了任何集合的幂集均存在。

$\mathcal{P}(S)$ 上的全部子集称为 S 上的集族 *(family of sets over S)*。

[3]如非利普希茨域的例子，可参考McLean (2000)。

[4]对应地，ℓ^p 空间是由 p 次可和序列组成的空间。

假定一个有限集 S 有 n 个元素，表示为 $|S| = n$ 或 $card(S) = n$，即 S 的势 (cardinality)是 n。那么 S 的幂集里有 $card(\mathcal{P}(s)) = 2^n$ 个元素。例如 $S = \{a, b, c\}, cardS = 3$。$S$ 的全部子集包括

$$\begin{cases} \{a\}, \{b\}, \{c\}, \\ \{a, b\}, \{a, c\}, \{b, c\}, \end{cases}$$

因此 $|\mathcal{P}(s)|$ 为 6 个 S 的子集，加上 $\varnothing$ 和 S 自身，共 $2^3 = 8$ 个。

对于非空集合 S 来说，S 上的一个 σ 代数是指其幂集 (powerset, Definition 8.4节) 的一个子集 Σ，Σ 中的元素在经过有限个补集、并集、交集 (intersection)这三种运算后，依然属于 Σ。即是说，Σ 对这三种运算是封闭 (closed) 的。

测度，可测空间，测度空间

Definition 8.6 (测度)**.** *对于一个集合 S 表示函数的定义域，有关于 S 的可测空间 Σ，Σ 中的元素是 S 上的子集族 (family of subsets)，并且 Σ 是一个 σ 代数 (第8.4节)。沿着扩展实数线 (extended real number line)* [5] *定义一个测度方程 $\mu \in \Sigma$，如果 μ 满足如下条件，我们称它为一个测度 (measure)：*

- *非负性 (non negative)*
$$\forall E \in \Sigma : \mu(E) \geq 0,$$
- *空集合的测度为 0*
$$\mu(\varnothing) = 0, \tag{8.18}$$
- *可数可加性 (countable additivity)，又称 σ 可加性 (σ additivity)：$\{E_i\}_{i=1}^{\infty} = \{E_1, E_2, \ldots\}$ 为 Σ 中可数个两两不不相交序列的集合 (pairwise disjoint sets in Σ)，则所有 E_i 并集的测度等于每个 E_i 的测度之和*
$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{k=1}^{\infty} \mu(E_i).$$

Definition 8.7 (可测空间)**.** *进而，我们称 (S, Σ) 为一个可测空间 (measurable space)。Σ 中的所有元素 $\{E_n\}_{n=1}^{\infty}$ 成为可测集 (measurable sets)。*

Definition 8.8 (测度空间)**.** *一个三元组 (triple)(μ, S, Σ) 称为测度空间 (measure space)。测度空间满足如下性质*

- *测度 μ 是单调方程 (monotonic)*
$$(\text{可测集合} E_1, E_2 \in \Sigma, \quad E_1 \subseteq E_2) \Rightarrow (\mu(E_1) \subseteq \mu(E_2)).$$
- *无限个可测集合的并集的测度*
  - *μ 是可数的次可加方程 (countably subadditive)。在 Σ 中的任何可数可测集合 $\{E_n\}_{i=n}^{\infty}$(可以不满足两两不相交)，有*
$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{n} \mu(E_n).$$

[5]扩展实数线是指实数集合 $\mathbb{R}$ 加上 $-\infty$ 和 $+\infty$，常写作 $\overline{\mathbb{R}}$ 或 $[-\infty, +\infty]$。

- μ 是连续 *(continuous)* 方程。在 Σ 中的任何可数可测集合 $\{E_n\}_{i=n}^{\infty}$(可以不满足两两不相交)，满足 $E_n \subset E_{n+1} \quad \forall n \in \mathbb{N}$，则集合的并集 $\bigcup E_n$ 也是可测的，并且满足

$$\mu\left(\bigcup_{n=1}^{\infty}\right) = \lim_{n\to\infty} \mu\left(E_n\right).$$

- 无限个可测集合的交集的测度

  - μ 是连续方程。在 Σ 中的任何可数可测集合 $\{E_n\}_{i=n}^{\infty}$(可以不满足两两不相交)，满足 $E_n \supset E_{n+1} \quad \forall n \in \mathbb{N}$，则集合的交集 $\bigcap E_n$ 也是可测的，并且满足

$$\mu\left(\bigcap_{n=1}^{\infty}\right) = \lim_{n\to\infty} \mu\left(E_n\right),$$

  需要指出的是，对于交集的情况，若无下述假设，该性质一般不成立：可测集合 $\{E_n\}_{n=1}^{\infty}$ 中应当至少有一个 E_n 有有限测度。举例来说，如果我们设 $E_n = [n, \infty) \subset \mathbb{R} \quad \forall n \in \mathbb{N}$，则这些可测集合全都具有无限测度，满足 $E_n \supset E_{n+1}$，然而

$$\left(\bigcap_{n=1}^{\infty} E_n = \varnothing\right) \Rightarrow \left(\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \mu\left(\varnothing\right) = 0\right).$$

Definition 8.9 (计数测度)**.** 在一个测度空间 (S, Σ, μ) 中，任意子集 $E \in \Sigma$ 的计数测度 *(counting measure)* 定义为

$$\mu(E) = \begin{cases} card(E), & \text{如果 } E \text{ 是有限子集} \\ +\infty, & \text{如果 } E \text{ 是无限子集} \end{cases}$$

利用计数测度这种直观的方法，我们可以在一个测度空间中，通过将 S 的全部可测子集作 Σ 代数，从而将 S 映射进入这个测度空间中。然而，只有当空间 S 是可数的时，它在测度空间 (S, Σ, μ) 中的计数测度才是 σ 有限 *(σ finite)* 的。

范数

范数 (norm) 是一个具有 “长度” 或 “大小” 概念的方程，是指赋予一个向量空间 (vector space) 内每个向量以一个非负的长度或大小；零向量 (zero vector) 的赋值为 0。半范数 (seminorm) 是指可以对某些非零向量赋值 0。

Definition 8.10 (范数)**.** 假定 V 是域 F 上的向量空间 $V \subset F$。V 上的 p 范数 *(p-norm)* 是指这样一个方程 $p : V \mapsto \mathbb{R}$，使得 $\forall a \in F, \forall \boldsymbol{u}, \boldsymbol{\nu} \in V$，以下 4 个关系均得到满足

- 绝对齐次 *(absolutely homogeneous)* 或称绝对可标量化 *(absolutely scalable)*，数乘线性

$$p\left(a\boldsymbol{\nu}\right) = |a|\, p(\boldsymbol{\nu}),$$

- 次可加 *(subadditivity)* 或称三角不等式 *(triangle inequality)*

$$p\left(\boldsymbol{u} + \boldsymbol{\nu}\right) \leq p\left(\boldsymbol{u}\right) + p\left(\boldsymbol{\nu}\right),$$

- 严格非负

$$p(\boldsymbol{u}) \geq 0,$$

- 确定 *(definite)*。只有零向量的范数是 *0*，反之范数是 *0* 的向量是零向量。

$$(p(\boldsymbol{\nu}) = 0) \Rightarrow (\boldsymbol{\nu} = 0)\,,$$

  即 $\boldsymbol{\nu}$ 是个空向量。

只满足前 *3* 条关系的方程 $p : V \mapsto \mathbb{R}$，我们称之为 V 上的半范数 *(seminorm)*。换句话说，所有范都是半范。

Definition 8.11 (商空间)**.** 每个向量空间 V 及其半范 p 都生成一个赋范向量空间 *(normed vector space)* $\left(\frac{V}{W}\right)$，我们称之为商空间 *(quotient space)*，其中 W 是 V 的子空间 $W \subset V$，包括所有满足 $p(\boldsymbol{\nu}) = 0$ 的向量 $\boldsymbol{\nu} \in V$。

对应地，商空间中的范数定义为 $p(W + \boldsymbol{\nu}) = p(\boldsymbol{\nu})$。

Definition 8.12 (等价范)**.** 一个向量空间 V 中的两个范 *(或两个半范)* p 和 q，当满足下述条件时，可称为等价范 *(equivalent norms)*：

$$c\,q(\boldsymbol{\nu}) \le p(\boldsymbol{\nu}) \le C\,q(\boldsymbol{\nu}), \quad \exists \text{实常数} c, C, \quad \forall \boldsymbol{\nu} \in V.$$

Definition 8.13 (平凡半范)**.** 平凡半范数 *(trivial seminorm)* 是指所有满足如下关系的半范

$$p(\boldsymbol{\nu}) = 0, \quad \forall \boldsymbol{\nu} \in V.$$

一个向量空间 V 中的每一个线性泛函 f (linear form)，都定义了 1 个半范 $\boldsymbol{\nu} \mapsto |f(\boldsymbol{\nu})|$。

有限维度的可数勒贝格空间 (p >= 1)

$\boldsymbol{x} = \{x_1, \ldots, x_n\} \in \mathbb{R}^n$ 的 p 范数 (p-norm)或者 $L^p, \forall p \ge 1$，可定义为

$$\| \boldsymbol{x} \|_p = (\|x_1\|^p + \|x_2\|^p + \ldots + \|x_n\|^p)^{\frac{1}{p}}\,,$$

包括一些特殊情况如

- $p = 1$ 是个网格距离 (grid distance)，又称出租车距离 (taxicab distance)，曼哈顿距离 (Manhattan distance)等。
- $p = 2$ 是个欧几里得范数 (Euclidean norm)。
- $p = \infty$ 是个 L^∞ 范数 (L^∞ norm)，又称最大范数 (maximum norm)，均匀范数 (uniform norm)，切比雪夫距离 (Chebyshev distance)等。

不同范数 p 之间的关系：

- $p = 1$ 的曼哈顿范数，从不小于任何欧几里得范数 $p = 2$。换句话说，任何向量 $\boldsymbol{x}$ 的欧几里得范数都受限于它的 1 范数

$$\|\boldsymbol{x}\|_n \le \|\boldsymbol{x}\|_1, \quad n \ge 1, n \in \mathbb{N}.$$

扩展 。任何向量 $\boldsymbol{x}$ 的 p 范数并不随着 p 的增加而增加

$$\|\boldsymbol{x}\|_{p+a} \le \|\boldsymbol{x}\|_p, \quad \forall \text{向量} \boldsymbol{x}, \quad \forall p \ge 1, a \ge 0, p, a \in \mathbb{N}.$$

扩展 . 柯西——施瓦茨不等式 (Cauchy-Schwarz inequality, Definition 8.21) 可得

$$\|\boldsymbol{x}\|_1 \le \sqrt{n}\|\boldsymbol{x}\|_2, \quad n = dim(\boldsymbol{x}).$$

扩展 .

$$\|\boldsymbol{x}\|_p \le \|\boldsymbol{x}\|_r \le n^{\left(\frac{1}{r} - \frac{1}{p}\right)}\|\boldsymbol{x}\|_p, \quad \forall \boldsymbol{x} \in \mathbb{C}^n, \quad 0 < r < p, r, p \in \mathbb{N}.$$

有限维度的可数勒贝格空间 (0 <= p <= 1)

略。

无限维度的勒贝格空间 (p 不可数)

我们设 $p \leq 1$。p 范数可以扩展到分析由无数个元素构成的向量，向量集合构成可数无限维的 p 范数列空间，用 l^p 表示。一些特殊情况如

- l^1 由绝对收敛 (absolute convergence) 序列构成的空间，如
  - $\sum_{n=0}^{\infty} |a_n| = L$，其中 L 是某个实数，$\{a_n\}_{n=0}^{\infty}$ 是一个实数或复数序列。
  - $\int_0^{\infty} |f(x)|\, dx = L$，其中 $\int_0^{\infty} f(x)\, dx$ 是关于某个方程 $f(x)$ 的不定积分。
- l^2 是一个由平方可加数列构成的空间，即一个希尔伯特空间 (Hilbert space)，见第8.2.3节。
- l^{∞} 是一个由有界数列 (bounded sequence) 构成的空间。

l^p 数列空间反映这样的向量空间结构：通过一个坐标一个坐标的向量加总 (或标量相乘) 而组成。如可数无限维实数 (复数) 数列 $\boldsymbol{x} = \{x_n\}_{n=1}^{\infty}, \boldsymbol{y} = \{y_n\}_{n=1}^{\infty}$ 中的向量和和标量乘

$$(x_1, \ldots, x_n, x_{n+1}, \ldots) + (y_1, \ldots, y_n, y_{n+1}, \ldots) = (x_1 + y_1, \ldots, x_n + y_n, x_{n+1} + y_{n+1}, \ldots),$$
$$\lambda (x_1, \ldots, x_n, x_{n+1} \ldots,) = (\lambda x_1, \ldots, \lambda x_n, \lambda x_{n+1}, \ldots),$$

对应的 p 范数

$$\|x\|_p = (|x_1|^p + \ldots + |x_n|^p + |x_{n+1}|^p + \ldots)^{\frac{1}{p}}.$$

则我们有定义

Definition 8.14 (可数无限维数列空间)**.** *定义 l^p 为一个包括所有实 (复) 数无限数列的空间 (countably infinite dimensional sequence space)，并且这些数列的 p 范数必须是有限的：这是因为存在着 p 范数是 ∞ 的无限数列，他们不应当包括在 l^p 空间中，如 $(1, 1, 1, \ldots), 1 \leq p < \infty$。*

随着 p 值的增加，l^p 集合的大小增加的更快。例如数列 $\boldsymbol{x} = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}, \frac{1}{n+1}, \ldots\right)$ 为例，它不在 l^1 空间中 ($p = 1$ 时范数不收敛)，而可能在某个 $p > 1$ 的 l^p 的空间中 (范数收敛)

$$\|x\|_1 = \left(1 + \frac{1}{2} + \ldots + \frac{1}{n}, \frac{1}{n+1} + \ldots\right) \to \infty,$$
$$\|x\|_p = \left(1 + \left(\frac{1}{2}\right)^p + \ldots + \left(\frac{1}{n}\right)^p, \left(\frac{1}{n+1}\right)^p + \ldots\right)^{\frac{1}{p}}, \quad p > 1.$$

$p = \infty$ 时，l^{∞} 范数

$$\|\boldsymbol{x}\|_{\infty} = \sup (|x_1|, \ldots, |x_n|, |x_{n+1}|, \ldots),$$

对应 l^{∞} 数列空间，包括全部实 (复) 数无限维数列，但要求他们都是有界的，即 ∞ 范数收敛。

无限维度的不可数勒贝格空间

上面我们讨论了有限维度和可数无限维度的勒贝格空间。然而当空间维度无限并且不可数 (即不存在可数的基) 时，我们无法像前文的方法来定义范数、进而描述空间。但如果该空间是勒贝格可积的 (Lebesgue integrable, 见 Definition 8.3)，仍然可以利用下述办法进行描述。

给定可测空间 (Ω, Σ, x) (见 Definition 8.8)，以及 $p \in \mathbb{R}, p \geq 1$。考虑所有从 Ω 到域 $\mathbb{F} = (\mathbb{R}, \mathbb{C})$ 的可测方程 (measurable function)集合 $u(x)$，方程绝对值的 p 次幂在 Ω 上有界，可积

一个 measur-
le function 的
条

$$L^p(\Omega) = \left\{ u(x); \|u\|_{L^p(\Omega)} \equiv \left(\int_\Omega |u(x)|^p \, dx \right)^{\frac{1}{p}} < \infty \right\},$$

集合中的方程 $u, \nu \in L^p(\Omega)$ 具有性质

- 可加

$$(u+\nu)(x) = u(x) + \nu(x),$$

- 数乘线性

$$u(\lambda\, x) = \lambda u(x), \forall 标量 \lambda \in \mathbb{F},$$

- 范数满足不等式

$$\|u+\nu\|_p^p \leq 2^{p-1} \left(\|u\|_p^p + \|\nu\|_p^p \right),$$

如果存在某一个集合 K 构成了零测度 (zero measure) $\mu(K) = 0$，那么只有当在这个零测度下 $u, \nu \in L^p(\Omega)$ 才有所区别时，我们说 u 和 ν 互相识别。

事实上，2 个可以扩展到更多个方程的情况，比如 3 个，即是说三角不等式 (triangle inequality) 对于范数形式的可积方程 $\|\cdot\|_p$ 依然成立，可由闵可夫斯基不等式 (Minkowski inequality) 证得。它表明 L^p 空间是一个赋范向量空间 (normed vector space)。对于一个测度空间 $\Omega \in L^p$，设 $1 \leq p \leq \infty$，$u, \nu \in L^p(\Omega)$ 中的元素，此时我们有

Definition 8.15 (三角不等式)**.** *$L^p(\Omega)$ 中的三角不等式 (triangle inequality)*

$$\|u+\nu\|_p \leq \|u\|_p + \|v\|_p, \quad 1 < p < \infty$$

式中等号存在的条件当且仅当 u 和 ν 是严格线性相关的，即 $\exists \lambda \geq 0 \Rightarrow u = \lambda\nu$，或者 $\nu = 0$。

Definition 8.16 (闵可夫斯基不等式)**.** *通过取可数测度 (见 Definition 8.9节)，闵可夫斯基不等式 (Minkowski inequality) 可表示为数列和向量的形式*

$$\left(\sum_{k=1}^n |x_k + y_k|^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |y_k|^p \right)^{\frac{1}{p}}, \quad \forall \{\boldsymbol{x}\}_n, \{\boldsymbol{y}\}_n = \mathbb{R}或\mathbb{C}, \quad n = dim(S).$$

证明. 首先要证明如果 u 和 ν 都有有限的 p 范数，那么 $(u+\nu)$ 的 p 范数也是有限的，且满足不等式关系

$$|u+\nu|_p \leq 2^{p-1} \left(|u|^p + |\nu|^p \right)$$

证明方式为：给定 $p > 1$，则 $h(x) = x^p$ 是一个在 $\mathbb{R}^+$ 上的凸方程 (convex)。由方程的凸性质可得

$$\left|\frac{1}{2}u + \frac{1}{2}\nu\right|^p \leq \left|\frac{1}{2}|u| + \frac{1}{2}|\nu|\right|^p \leq \frac{1}{2}|u|^p + \frac{1}{2}|\nu|^p,$$

$\hookrightarrow$

$$\left|u+\nu\right|^p \leq \frac{1}{2}|2u|^p + \frac{1}{2}|2\nu|^p = 2^{p-1} \left(|u|^p + |\nu|^p \right),$$

可见 $\|u+\nu\|_p$ 是有限范数。

在证明了 $\|u+\nu\|_p$ 范数有限后，我们有：如果 $|u+\nu|_p = 0$，闵可夫斯基不等式直接变为等号且成立。现在假定 $|u+\nu|_p \neq 0$，使用三角不等式和霍德尔不等式 (Hölder's inequality, Definition

(8.26)) ，我们有

$$
\begin{aligned}
\|u+\nu\|_p^p &= \left[\left(\int |u+\nu|^p \, d\mu\right)^{\frac{1}{p}}\right]^p \\
&= \int |u+\nu| \, |u+\nu|^{p-1} \, d\mu \\
&\leq \int (|u|+|\nu|) \, |u+\nu|^{p-1} \, d\mu \\
&= \int |u| \, |u+\nu|^{p-1} \, d\mu + \int |\nu| \, |u+\nu|^{p-1} \, d\mu \\
&= \left[\left(\int |u|^p \, d\mu\right)^{\frac{1}{p}} + \left(\int |\nu|^p \, d\mu\right)^{\frac{1}{p}}\right] \left(\int \left(|u+\nu|^{(p-1)\cdot\left(\frac{p}{p-1}\right)}\right) d\mu\right)^{1-\frac{1}{p}} \\
&= (\|u\|_p + \|\nu\|_p) \frac{\|u+\nu\|_p^p}{\|u+\nu\|_p},
\end{aligned}
$$

$\hookrightarrow$

$$
\|u+\nu\|_p \leq \|u\|_p + \|\nu\|_p.
$$

□

有时我们需要积分形式的闵可夫斯基不等式 (Minkowski integral inequality)：

Definition 8.17 (闵可夫斯基积分不等式)**.** 假定存在两个 σ 可测度空间 (S_1, μ_1) 和 (S_2, μ_2)，并且 $F : S_1 \times S_2 \mapsto \mathbb{R}$ 是可测方程，那么我们有

$$
\left[\int_{S_2} \left| \int_{S_1} F(x,y)\mu_1(dx) \right|^p \mu_2(dy)\right]^{\frac{1}{p}} \leq \left[\int_{S_1} \left| \int_{S_2} F(x,y)\mu_2(dy) \right|^p \mu_1(dx)\right]^{\frac{1}{p}}, \quad p < \infty,
$$

满足该条件的 $\|\cdot\|_p$ 构成半范数 (seminorm, Definition 8.10)，对应半赋范向量空间 (seminormed vector space) $\mathcal{L}^p(\Omega, \mu)$。之所以称之为半范数，是因为该空间中存在非零向量 u 满足 $\|u\|_p = 0$。

我们可以用标准的拓扑方法，从半范向量空间 $\mathcal{L}^p(\Omega, \mu)$ 中得到一个赋范向量空间。在 $\mathcal{L}^p(\Omega, \mu)$ 中，考虑所有使 $\|u\|_p = 0$ 的向量集合

$$
\mathcal{N} = \{u; \|u\|_p = 0\}.
$$

$\mathcal{N}$ 可以看作是一个映射 $f \mapsto \|u\|_p$ 的零向量空间。则对于可测度方程 u 而言：

$$
\|u\|_p = 0 \Longleftrightarrow \mu(u \neq 0) \Longleftrightarrow u_{\mu\text{-几乎处处}} = 0,
$$

μ-几乎处处 表示在测度 μ 的意义上几乎处处有界 (almost everywhere)。从这个意义上来看，$\mathcal{N}$ 是一个 kernel $\|\cdot\|_p$，并且不依赖于 p

$$
\mathcal{N} \equiv ker\left(\|\cdot\|_p\right) = \left\{u : u_{\mu\text{-几乎处处}} = 0\right\}.
$$

则我们可以定义一个关于 $\mathcal{L}^p(\Omega, \mu)$ 和 kernel $\mathcal{N}$ 的商空间

$$
L^p(\Omega, \mu) \equiv \frac{\mathcal{L}^p(\Omega, \mu)}{\mathcal{N}},
$$

商空间 $L^p(\Omega, \mu)$ 中的某个 u 可以看做是与 $\mathcal{L}^p(\Omega, \mu)$ 中的 f 相差 1 个 $\mathcal{N}$ 中对应元素的等价类。

由此可见，$L^p(\Omega, \mu)$ 就是 Ω 上关于测度 μ 的 L^p 空间。对应的 $\|\cdot\|_p$ 成为 $L^p(\Omega, \mu)$ 的 p 范数。需要指出的是，严格来说 L^p 空间中的元素并非某个具体方程，而是由一个方程族构成的等价类。当我们取出 L^p 中的元素作计算的时候，参与计算的其实是从这个方程组中抽取的一个代表方程。

$p=\infty$ 时对应的空间 $L^\infty(S,\mu)$ 也可以用类似方法求得：

$$\|f\|_\infty \equiv \inf\{C\geq 0 : |f(x)|\leq C\text{对于几乎所有}x\},$$
$$\exists q<\infty \Rightarrow f\in L^\infty(S,\mu)\bigcap L^q(S,\mu)\Rightarrow \|f\|_\infty = \lim_{p\to\infty}\|f\|_p.$$

勒贝格空间 $L^p(S,\mu)$ 的完备性 (completeness) 通常称为里兹——费舍定理 (Riesz-Fischer theorem)[6]。

对于 $1\leq p\leq\infty$ 的情况，勒贝格空间 $L^p(S,\mu)$ 是一个完备赋范向量空间，常称为巴拿赫空间 (Banach space)。所有 L^p 空间都是巴拿赫空间。

加权勒贝格空间

有时候会遇到加权勒贝格空间的情况。

Definition 8.18 (加权勒贝格空间)**.** *考虑一个测度空间 $L^p(S,\sigma,\mu)$，其中有一个可测方程 $w:S\to[0,\infty)$。有时我们也将 $L^p(S,w\,d\mu)$ 称为 $w-$ 加权勒贝格空间 (w-weighted Lebesgue space)，其中测度 $d\nu\equiv w\,d\mu$。由此我们有测度的定义*

$$\nu(A)\equiv\int_A w(x)\,d\mu(x),\quad \forall A\in\Sigma.$$

在此基础上，加权勒贝格空间 $L^p(S,w\,d\mu)$ 的范数

$$w=\frac{d\nu}{d\mu}\Rightarrow\left\|\mu\right\|_{L^p(S,w\,d\mu)}\equiv\left(\int_s w(x)\,|\mu(x)|^p\,d\mu(x)\right)^{\frac{1}{p}}.$$

从这个角度上说，$L^p(s,d\nu)\equiv L^p(s,wd\nu)$.

8.2.3　希尔伯特 (H) 空间

有且只有 $p=2$ 时的特殊形式空间 $L^2(\Omega)$，是希尔伯特空间 (Hilbert space)。

作为 (完备赋范的) 内积向量空间 (inner product space, Definition 8.19)，希尔伯特空间是有限维欧几里得空间的一个扩展：从 $\mathbb{R}$ 扩展到 $\mathbb{R}$ 和 $\mathbb{C}$，从有限维度到无限维度，但保留了完备性 (completeness) 特征 (一般来说，非欧几里得空间往往破坏了完备性)。希尔伯特空间与欧几里得空间相仿，有长度和角度的概念，因而可以引申出正交性和垂直性，从而为基于正交多项式的傅里叶级数等提供表达方式。

任何一个希尔伯特空间都是巴拿赫空间，反之则未必。

例：欧几里得空间

假设所有的希尔伯特空间都是复数 (实际应用中大多数是实数)。二维欧几里得空间 $\mathbb{R}^2$ 中，向量 $\boldsymbol{x},\boldsymbol{y}$ 构成一个希尔伯特空间

$$\boldsymbol{x}\cdot\boldsymbol{y}=\sum_{k=1}^n\overline{x_k}y_k,$$

对应范数

$$\|\cdot\|=\sqrt{\langle\boldsymbol{x},\rangle\boldsymbol{y}}.$$

三维欧几里得空间 R^3 中，以笛卡尔坐标系表示的 $\boldsymbol{x},\boldsymbol{y}$ 向量的点乘为

$$\boldsymbol{x}\cdot\boldsymbol{y}=(x_1,x_2,x_3)\cdot(y_1,y_2,y_3)=x_1y_1+x_2y_2+x_3y_3,$$

点乘具有如下性质：

[6]证明略。完备性 (completeness) 的含义，见第218页。

- 对称性

$$\boldsymbol{x}\cdot\boldsymbol{y}=\boldsymbol{y}\cdot\boldsymbol{x},$$

- 首项线性

$$(a\boldsymbol{x}+b\boldsymbol{y})\cdot\boldsymbol{z}=a\boldsymbol{x}\cdot\boldsymbol{z}+b\boldsymbol{y}\cdot\boldsymbol{z},\quad a,b\text{是任意标量},\quad \boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\text{是任意向量},$$

- 正定

$$\boldsymbol{x}\cdot\boldsymbol{x}\begin{cases}\geq 0 & \forall\boldsymbol{x}\geq 0,\\ =0 & \text{iff. } \boldsymbol{x}=0.\end{cases}$$

Definition 8.19 (内积空间)**.** *满足上述三个条件的 (实数) 向量乘称为 (实数) 内积 (inner product)，用 $\langle .,.\rangle$ 表示。给定一个实数或复数域 $\mathbb{F}$ 中的向量空间 V，则我们将内积形式的向量空间 $\langle .,.\rangle : V\times V\mapsto\mathbb{F}$ 表示为内积空间 (inner product space)。内积空间满足三个性质:$\forall$向量$\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$，以及 $\forall$标量$a\in\mathbb{F}$*

- *共轭对称 (conjugate symmetry)*

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle=\begin{cases}\langle\boldsymbol{y},\boldsymbol{x}\rangle & \mathbb{F}=\mathbb{R},\\ \overline{\langle\boldsymbol{y},\boldsymbol{x}\rangle} & \mathbb{F}=\mathbb{C}.\end{cases}$$

  其中标有上横线 $\overline{(\cdot)}$ 的部分表示复数共轭 (complex conjugate)。

- *首项线性*

$$\langle a\boldsymbol{x},\boldsymbol{y}\rangle=a\langle\boldsymbol{x},\boldsymbol{y}\rangle,$$
$$\langle\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z}\rangle=\langle\boldsymbol{x},\boldsymbol{z}\rangle+\langle\boldsymbol{y},\boldsymbol{z}\rangle.$$

- *正定*

$$\langle\boldsymbol{x},\boldsymbol{x}\rangle>0,$$
$$\langle\boldsymbol{x},\boldsymbol{x}\rangle=0\Leftrightarrow\boldsymbol{x}=0.$$

任何有限维内积空间都也是希尔伯特空间。在欧几里得空间内，两个向量的内积大小与两方面因素有关，一为向量的长度 (即范数)$\|\boldsymbol{x}\|$，一为 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的夹角 θ，满足

$$\boldsymbol{x}\cdot\boldsymbol{y}=\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cos\theta.$$

欧几里得空间 $\mathcal{R}^3$ 中，对 $n\in\mathcal{N}$ 个向量 $\boldsymbol{x}_n$ 求和构成一个数学级数 $\sum_{n=0}^{\infty}\boldsymbol{x}_n$，当每个向量的范之和收敛到一个小于正无穷的向量 L 时，我们称这个级数仍然是绝对收敛 (absolutely convergence) 的

$$\sum_{k=0}^{N}\|\boldsymbol{x}_k\|<\infty.$$

一个绝对收敛的向量数列 $\sum_{k=0}^{N}\boldsymbol{x}$，收敛至某个极限向量 $\boldsymbol{L}\in R^3$

$$\left\|\boldsymbol{L}-\sum_{k=0}^{N}\boldsymbol{x}\right\|\to 0,\text{随着}N\to\infty,$$

这称为欧几里得空间的完备性 (completeness of Euclidean space)。

类似地，在欧几里得空间中，复数平面 (complex plane) $\mathbb{C}$ 由量 (magnitude) 的形式予以反映，即复绝对值 (complex modulus) $|z|$，定义为 z 与其共轭复数 (complex conjugate) $\bar{z}$ 乘积的平方根

$$|z| = \sqrt{z\,\bar{z}}.$$

如果 $z = x + y$, $x = \Re(z)$, $y = Im(z)$，复绝对值为常见的二元欧几里得空间的长度

$$\|z\| = \sqrt{\Re(z)^2 + \Im(z)^2} = \sqrt{x^2 + y^2}.$$

两个复数 z, w 的内积

$$\langle z, w\rangle = z\bar{w},$$

或者对于复数空间 $z, w \in \mathbb{C}^2$，即 $z = (z_1, z_2), w = (w_1, w_2)$，对应内积

$$\langle z, w\rangle = z_1\bar{w}_1 + z_2\bar{w}_2,$$

其中 $\Re(\langle z, w\rangle) \in \mathbb{R}^4$。这个内积埃米特对称 (Hermitian symmetric)，即是说

$$\langle w, z\rangle = \overline{\langle z, w\rangle}.$$

希尔伯特空间 H 是一个实数 (或复数) 内积向量空间，其中的向量可以内积形式表示为 $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$，满足如下特性：

- 对称性

$$\langle \boldsymbol{y}, \boldsymbol{x}\rangle = \begin{cases} \langle \boldsymbol{x}, \boldsymbol{y}\rangle & \text{实数向量}, \\ \overline{\langle \boldsymbol{x}, \boldsymbol{y}\rangle} & \text{复数向量}. \end{cases}$$

- 首项线性

$$(a\boldsymbol{x} + b\boldsymbol{y}) \cdot \boldsymbol{z} = a\boldsymbol{x} \cdot \boldsymbol{z} + b\boldsymbol{y} \cdot \boldsymbol{z}, \quad a, b\text{是任意标量}, \quad \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\text{是任意向量},$$

- 正定[7]

$$\boldsymbol{x} \cdot \boldsymbol{x} \begin{cases} \geq 0 & \forall \boldsymbol{x} \geq 0, \\ = 0 & \text{iff. } \boldsymbol{x} = 0. \end{cases}$$

由对称性和首项线性可得第二项系数是反线性的 (antilinear):

$$\langle \boldsymbol{x}, a\boldsymbol{y} + b\boldsymbol{z}\rangle = \bar{a}\langle \boldsymbol{x}, \boldsymbol{y}\rangle + \bar{b}\langle \boldsymbol{x}, \boldsymbol{z}\rangle.$$

希尔伯特空间的范是一个实值方程

$$\|\cdot\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}$$

Definition 8.20 (对偶空间)**.** *已知内积定义*

$$\langle u, \nu\rangle_\Omega := \int_\Omega u(x)\,\nu(x)\,dx,$$

根据闵可夫斯基不等式 (Definition 8.16)，$L^p(\Omega)$ 中的三角不等式可以扩展到更一般的形式：

$$\|\nu\|_{L^q(\Omega)} = \sup_{0 \neq u \in L^p(\Omega)} \frac{|\langle u, \nu\rangle_\Omega|}{\|u\|_{L^p(\Omega)}}, 1 \leq p < \infty,$$

其中 p, q 是伴随参数，满足

$$\frac{1}{p} + \frac{1}{q} = 1,$$

不难看出，RHS 满足三角不等式关系。则 $L^q(\Omega)$ 和 $L^p(\Omega)$ 构成一组对偶空间 (dual space)。

[7]省略部分复数形式的表述，以使方程结构更紧凑。

当 $p=q=2$ 时，$L^2(\Omega)$ 就成为包括全部平方可积方程的空间。此时霍德尔不等式 (Definition 8.26) 就变成了柯西——施瓦茨不等式。

Definition 8.21 (柯西——施瓦茨不等式)**.** *内积空间 (见 Definition 8.19) $\langle .,. \rangle : V \times V \mapsto \mathbb{F}$ 中，对于任意两个向量 $\forall \boldsymbol{u}, \boldsymbol{\nu} \in V$，内积绝对值的平方，满足三角不等式*

$$\left|\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle\right|^2 \leq \langle \boldsymbol{u}, \boldsymbol{u} \rangle \cdot \langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle,$$

两侧同时开平方根，将 RHS 改写为向量范的形式，我们有柯西——施瓦茨不等式 (Cauchy-Schwarz inequality)

$$\begin{aligned}
&\left|\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle\right| \leq \|\boldsymbol{u}\| \, \|\boldsymbol{\nu}\|, \\
\hookrightarrow &\int_\Omega |u(x)\, \nu(x)| \; dx \leq \|u\|_{L^2(\Omega)} \, \|\nu\|_{L^2(\Omega)}, \\
\hookrightarrow &\left|\langle u, u \rangle\right|_{L^2(\Omega)} = \|u\|^2_{L^2(\Omega)}, \forall u = \nu, u \in L^2(\Omega).
\end{aligned}$$

其中等号成立的条件，只有一下两种之一：$\boldsymbol{u}, \boldsymbol{\nu}$ 线性无关 (linearly independent)，即平行；$\boldsymbol{\nu}$ 是零向量或是标量。

证明. 柯西——施瓦茨不等式的证明方法有很多种 (Wu and Wu, 2011)，我们取其中一种。

$\boldsymbol{\nu} = 0 \Rightarrow \left|\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle\right| = \|\boldsymbol{u}\| \, \|\boldsymbol{\nu}\| \forall \boldsymbol{u} \in V$。

$\boldsymbol{u} \neq 0, \boldsymbol{\nu} \neq 0$。设一个向量 $\boldsymbol{z}$ 满足

$$\boldsymbol{z} := \boldsymbol{u} - \boldsymbol{u}_{\boldsymbol{\nu}} = \boldsymbol{u} - \underbrace{\frac{\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle}{\langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle}}_{\text{标量}} \boldsymbol{\nu}.$$

对 $\boldsymbol{z}$ 和 $\boldsymbol{\nu}$ 作内积，由内积空间的性质之一——首项线性得

$$\begin{aligned}
\langle \boldsymbol{z}, \boldsymbol{\nu} \rangle &= \left\langle \boldsymbol{u} - \frac{\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle}{\langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle} \boldsymbol{\nu}, \boldsymbol{\nu} \right\rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{\nu} \rangle - \left\langle \frac{\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle}{\langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle} \boldsymbol{\nu}, \boldsymbol{\nu} \right\rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{\nu} \rangle - \frac{\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle}{\langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle} \langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle \\
&= 0.
\end{aligned}$$

$\langle \boldsymbol{z}, \boldsymbol{\nu} \rangle = 0 \Rightarrow \boldsymbol{z} = 0$，作为 $\boldsymbol{u}$ 向 $\boldsymbol{\nu}$ 所在平面 (plane) 所做的正交映射，反映了 $\boldsymbol{u}$ 和 $\boldsymbol{\nu}$ 线性无关。因此我们对 $\boldsymbol{z}$ 的定义式继续使用勾股定理

$$\begin{aligned}
\boldsymbol{u} &= \frac{\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle}{\langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle} \boldsymbol{\nu} + \boldsymbol{z} \\
\hookrightarrow \|\boldsymbol{u}\|^2 &= \left|\frac{\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle}{\langle \boldsymbol{\nu}, \boldsymbol{\nu} \rangle}\right|^2 \|\boldsymbol{\nu}\|^2 + \|\boldsymbol{z}\|^2 \\
&= \frac{\left|\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle\right|^2}{\left(\|\boldsymbol{\nu}\|^2\right)^2} \|\boldsymbol{\nu}\|^2 + \|\boldsymbol{z}\|^2 \\
&= \frac{\left|\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle\right|^2}{\|\boldsymbol{\nu}\|^2} + \|\boldsymbol{z}\|^2 \\
&\geq \frac{\left|\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle\right|^2}{\|\boldsymbol{\nu}\|^2} \\
\hookrightarrow |\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle| &\leq \|\boldsymbol{u}\| \, \|\boldsymbol{\nu}\|
\end{aligned} \tag{8.19}$$

□

8.2.4 索伯列夫 $W^{k,p}(\Omega)$ 空间

微分的类以及平滑方程

$H^s(\Omega)$ 中包括 L^p 空间中的具有弱可导性的平滑方程 [8]，常用于求解偏微分方程 PDEs。

Definition 8.22 (微分的类以及平滑方程)**.** 我们可以根据方程的微分性质，对方程作分类 *(differentiability classification)*。一个实数集合 $\mathbb{R}$ 上的开区间中，实值方程 $f \in \mathbb{R}$。

如果微分方程 $f', f'', \ldots, f^{(k)}$ 都存在并且 $f', f'', \ldots, f^{(k-1)}$ 连续，我们称 f 属于 $\mathbb{C}^k$ 类方程。当 $k \to \infty$ 时 f 的所有 k 次微分都存在且连续，我们称之为 $\mathbb{C}^\infty$ 类方程，无限可微方程 *(infinitely differentiable)*，或者称之为平滑方程 *(smooth function)*。

如果 f 是平滑的，并且 f 沿着域中任意一点作泰勒级数展开都收敛至该点，则我们称 f 是 $\mathbb{C}^\omega$ 类方程，或称之为解析方程 *(analytic function)*。可见 $\mathbb{C}^\omega \subset \mathbb{C}^\infty$。

举例来说，

- $\mathbb{C}^0$ 中包括所有连续方程，
- $\mathbb{C}^1$ 中包括所有一次可微方程，并且这些方程的一次导数是连续的，称连续可导 *(continuously differentiable)*。进而
  - $\forall f \in \mathbb{C}^1 \Rightarrow f'$存在且$f' \in \mathbb{C}^0$
  - $\forall f \in \mathbb{C}^k \Rightarrow f', f'', \ldots f^k$存在且$f' \in \mathbb{C}^{k-1}$

分部积分

Definition 8.23 (分部积分公式)**.** 分部积分公式 *(integration by parts formula)*是指，如果 $u = u(x), u'(x) = du/dx$，以及 $\nu = \nu(x), \nu'(x) = d\nu/dx$，那么

$$\begin{aligned}\int_a^b u(x)\nu'(x)\,dx &= [u(x)\nu(x)]_a^b - \int_a^b u'(x)\nu(x)\,dx \\ &= [u(b)\nu(b) - u(a)\nu(a)] - \int_a^b u'(x)\nu(x)\,dx,\end{aligned}$$

或者用更紧凑的表现形式

$$\int u d(\nu) = u\nu - \int \nu d(u).$$

广义积分

我们定义 $L^{1,\mathrm{loc}}(\Omega)$ 为局部可积 (locally integrable) 的方程空间，即方程 $u \in L^{1,\mathrm{loc}}(\Omega)$ 在任意一个封闭有界子集 $K \subset \Omega$ 中可导。

例，设 $\Omega = (0,1)$，$u(x) = \frac{1}{x}$。由于

$$\int_0^1 u(x)\,dx \approx \lim_{\epsilon\to 0}\int_\epsilon^1 \frac{1}{x}\,dx \approx \lim_{\epsilon\to 0}\log\frac{1}{\epsilon} = \infty,$$

可见 $u \notin L^1(\Omega)$。可由 Mathematica 算得

[8]“平滑” 方程 (smoothness) 的定义包括很多种，由弱到强有

- 连续性，
- 可导性 (可导方程必连续)，
- 它的最高一阶导数也是连续的，

等等。索伯列夫空间中方程设定为 “弱” 可导形式，是为了使得空间完备，是一个巴拿赫空间。

```
Limit[Integrate[1/x, {x, ee, 1}], ee -> 0]
Limit[Log[1/x], x -> 0]
```

对于任一闭区间 $K := [a,b] \subset (0,1) = \Omega, \quad 0 < a < b < 1$ 我们有

$$\int_K u(x)\,dx = \int_a^b \frac{1}{x}\,dx = \ln\frac{b}{a} < \infty,$$

可见 $u \in L^{1,\mathrm{loc}}(\Omega)$。

此外对于 $\phi, \psi \in C_0^\infty(\Omega)$，根据分部积分 (Definition 8.23) 我们有

$$\int_\Omega \phi(x)\frac{\partial}{\partial x_i}\psi(x)\,dx = -\int_\Omega \frac{\partial}{\partial x_i}\phi(x)\,\psi(x)\,dx,$$

上式对于哪怕是非平滑方程 ϕ, ψ 也适用。由此可得广义偏导数的定义

Definition 8.24 (广义偏导数)**.** 设 $u \in L^{1,loc}(\Omega)$。如果 $\exists \nu \in L^{1,loc}(\Omega)$，使得满足

$$\int_\Omega \nu(x)\varphi(x)\,dx = -\int_\Omega u(x)\frac{\partial}{\partial x_i}\varphi(x)\,dx, \tag{8.20}$$

其中 $\varphi(x) \in C_0^\infty(\Omega)$，那么我们说 $\nu(x)$ 是 $u(x)$ 在 Ω 中关于 x_i 的广义偏导数 *(generalized partial derivative)*，写作 $\nu(x) := \partial u(x)/\partial x_i$。

类似地，u 的第 α 阶广义偏导数 $\nu(x) = D^\alpha u(x)$ 记作

$$\int_\Omega u(x)D^\alpha\varphi(x)\,dx = (-1)^{|\alpha|}\int_\Omega \nu(x)\varphi(x)\,dx,$$

其中多重指数 *(multi-index)* $\alpha = (\alpha_1, \ldots, \alpha_n), x = (x_1, \ldots, x_n), |\alpha| = \alpha_1 + \ldots + \alpha_n$，积分操作符 D^α 是下述形式的缩写

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\ldots x_n^{\alpha_n}},$$

例，设 $u(x) = |x|$，$x \in \Omega = (-1,1)$。对于任一 $\varphi \in C_0^\infty(\Omega)$，我们有

$$\begin{aligned}
&\int_{-1}^1 u(x)\frac{\partial}{\partial x}\varphi(x)\,dx \\
&= -\int_{-1}^0 x\frac{\partial}{\partial x}\varphi(x)\,dx + \int_0^1 x\frac{\partial}{\partial x}\varphi(x)\,dx \\
&= \left\{-\left[x\,\varphi(x)\right]_{-1}^0 + \int_{-1}^0 \varphi(x)\,dx\right\} + \left\{\left[x\,\varphi(x)\right]_0^1 - \int_0^1 \varphi(x)\,dx\right\} \\
&= \int_{-1}^0 \varphi(x)\,dx - \int_0^1 \varphi(x)\,dx \\
&= -\int_{-1}^1 \mathrm{sgn}(x)\varphi(x)\,dx,
\end{aligned}$$

其中

$$\mathrm{sgn}(x) := \begin{cases} 1 & x > 0, \\ -1 & x < 0. \end{cases}$$

则方程 $u(x) = |x|$ 的 1 阶广义偏导数 $\nu(x)$ 为

$$\nu(x) := \frac{\partial}{\partial x}u(x) = \mathrm{sgn}(x) \in L^{1,\mathrm{loc}}(\Omega).$$

递归方法计算 2 阶广义偏导数：

$$\int_{-1}^{1} \operatorname{sgn}(x) \frac{\partial}{\partial x} \varphi(x) = -\int_{-1}^{0} \frac{\partial}{\partial x} \varphi(x)\, dx + \int_{0}^{1} \frac{\partial}{\partial x} \varphi(x)\, dx = -2\varphi(0), \tag{8.21}$$

然而 $\nexists \nu \in L^{1,\mathrm{loc}}(\Omega)$ 满足

$$\int_{-1}^{1} \nu(x)\, \varphi(x)\, dx = 2\varphi(0), \quad \forall\, \phi \in C_0^\infty(\Omega).$$

我们将在随后讨论 $\operatorname{sgn}(x)$ 作为分布概念时的广义积分，见235页第8.2.4节。

整数阶次的单维索伯列夫空间

索伯列夫相关教材，可参考如Adolfson et al. (2008b); Tartar (2007); Mazya (2009a,b) 等。

在 $\mathbb{R}^n$ 的一个开放子集 Ω 中，对于一个给定的非负整数 k，我们有索伯列夫空间或 $W^{k,p}(\Omega)$。它是一种希尔伯特空间的特例，其中

- (内积形式表示的) 方程向量空间都是可微的，
- 范数是可微方程范数的组合，包括方程本身的 L^p 范数，以及方程直到某一给定次导数的范数的组合。

单维索伯列夫空间 $W^{k,p}$ 中的方程即是 $\mathbb{R}$ 中的方程，可定义为单维勒贝格空间 $L^p(\mathbb{R})$ 中方程 f 的子集，f 满足如下特征：对于给定的 $p \in \mathbb{N}, 1 \le p \le \infty$，其中的 $f^{(k-1)}$ 需要几乎处处可导，并且几乎处处等于其勒贝格积分的 $k-1$ 次导数。

f 本身以及 f 的直至第 k 阶弱导数是有限的 L^p 范数

$$\begin{aligned}
\|f\|_{W^{k,p}} &= \left(\sum_{i=0}^{k} \|f^{(i)}\|_p^p \right)^{\frac{1}{p}} \\
&= \sum_{i=0}^{k} \left(\int \left| f^{(i)}(t) \right|^p dt \right)^{\frac{1}{p}} \\
&= \|f^{(k)}\|_p + \|f\|_p,
\end{aligned}$$

最后一个等式表明，单维索伯列夫空间的范数，等于方程序列自身的范数、以及其最高一阶导数的范数之和。

带有这一范数 $\|\cdot\|_{W^{k,p}}$ 的单维索伯列夫空间 $W^{k,p}$ 是一个巴拿赫空间。

整数阶次的单维索伯列夫空间 (p=2)

$p=2$ 的单维索伯列夫空间 $W^{k,2}$ 非常重要，因为它与傅里叶级数关系密切，并且构成了希尔伯特空间 $H^k = W^{k,2}$。

H^k 空间可以定义如下：可由帕塞瓦尔定理 (Parseval theorem)予以证明 (证明略。)

$$H^k(\mathbb{T}) = \left\{ f \in L^(\mathbb{T}) : \sum_{n=-\infty}^{\infty} \left(1 + n^2 + n^4 + \ldots + n^{2k}\right) \left| \hat{f}(n) \right|^2 < \infty \right\},$$

其中 $\hat{f}$ 是方程 f 的傅里叶级数 (Fourier series)，它快速衰减 (decay)。$\mathbb{T}$ 表示环面 (torus)。

此时的 H^k 空间可以理解为在 L^2 空间中取内积的形式：

$$\langle \boldsymbol{u}, \boldsymbol{\nu} \rangle_{H^k} = \sum_{i=0}^{k} \langle D^i \boldsymbol{u}, D^i \boldsymbol{\nu} \rangle_{L^2}.$$

多维索伯列夫空间

Definition 8.25 (索伯列夫空间)**.** 假设 $\exists k \in \mathbb{N}_0, 1 \le p < \infty$，那么多维索伯列夫空间 $W^{k,p}(\Omega)$ 定义为在 Ω 上的全部方程集合，使得对于每一个多元指数 *(multiple index)* α，方程的混合偏导数 *(分部积分)* $f^{(\alpha)}$ 都存在，并且 $f^{(\alpha)} \in L^p(\Omega), \|f\|_{L^p(\Omega)} < \infty$。从此意义上我们有多维索伯列夫空间 $W^{k,p}(\Omega)$ 的定义式：

$$\begin{aligned} W^{k,p}(\Omega) &:= \overline{C^\infty(\Omega)}^{\|\cdot\|_{W^{k,p}(\Omega)}} \\ &= \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega), \quad \forall |\alpha| \le k, \quad k \in \mathbb{N}\}. \end{aligned} \tag{8.22}$$

多维索伯列夫空间 $W^{k,p}(\Omega)$ 的范数定义方式有多重，最常见的两种如下 *(并且这两种设定是等价的)*

$$\|u\|_{W^{k,p}(\Omega)} := \begin{cases} \left(\sum_{|\alpha|\le k} \left\|D^\alpha u\right\|^p_{L^p(\Omega)}\right)^{\frac{1}{p}} & 1 \le p \le +\infty, \\ \max_{|\alpha|\le k} \left\|D^\alpha u\right\|_{L^p(\Omega)} & p = +\infty. \end{cases} \tag{8.23}$$

$$\|u\|'_{W^{k,p}(\Omega)} := \begin{cases} \sum_{|\alpha|\le k} \left\|D^\alpha u\right\|_{L^p(\Omega)} & 1 \le p \le +\infty, \\ \sum_{|\alpha|\le k} \left\|D^\alpha u\right\|_{L^\infty}(\Omega) & p = +\infty. \end{cases}$$

有着上述定义和范数的无限维索伯列夫空间 $W^{k,p}(\Omega)$ 是一个巴拿赫空间。同时，对于 $p < \infty$ 的情况而言，它也是一个可分空间 (separable space)[9]。

习惯上，我们将 1 个索伯列夫空间 $W^{k,2}(\Omega)$ 写作希尔伯特空间形式 $H^k(\Omega)$，对应范数 $\|\cdot\|_{W^{k,2}}(\Omega)$。

分数次阶索伯列夫空间：Sobolev-Slobodeckij 空间法

关于分数次阶索伯列夫空间，可参考Di Nezza et al. (2012)。前面介绍的几种索伯列夫空间的情况，均假定 $k \in \mathcal{N}$。然而有时候我们需要处理 k 是分数的情况 (fractional order Sobolev space)。大致说来，有两种方法可以处理分数次阶索伯列夫空间，一种是 Sobolev-Slobodeckij 空间法，一种是贝塞尔位势空间法，后者需要一些关于分布的傅里叶变换知识，我们将在随后介绍，见238页 Definition 8.34。

Sobolev–Slobodeckij 空间是勒贝格空间中的霍德尔条件 (Hölder condition, Definition 8.26) 的广义化。

根据 $W^{k,p(\Omega)}$ 或者 $\mathring{W}^{k,p}(\Omega)$ 空间的定义(8.22)以及相应的范数(8.23)，对于 $0 < s \in \mathbb{R}$, 的情况，设 $s = k + \kappa, k \in \mathbb{N}_0$，我们有 Sobolev-Slobodeckij 范数 (Sobolev-Slobodeckij norm)

$$\|u\|_{W^{s,p}(\Omega)} := \left\{\|u\|^p_{W^{k,p}(\Omega)} + |u|^p_{W^{s,p}(\Omega)}\right\}^{\frac{1}{p}},$$

其中 $|u|^p_{W^{s,p}(\Omega)}$ 是 Sobolev-Slobodeckij 半范数 (Sobolev-Slobodeckij seminorm)

$$|u|^p_{W^{s,p}(\Omega)} = \sum_{|\alpha|=k} \int_\Omega \int_\Omega \frac{\left|D^\alpha u(x) - D^\alpha u(y)\right|^p}{\left|x-y\right|^{d+p\kappa}} \, dx \, dy.$$

对于 $p = 2$ 的情况，$W^{s,2}(\Omega)$ 成为一个内积形式的希尔伯特空间：

$$\langle u, \nu\rangle_{W^{k,2}(\Omega)} := \sum_{|\alpha|\le k} \int_\Omega D^\alpha u(x) \, D^\alpha \nu(x) \, dx, \quad s = k \in \mathbb{N}_0, \tag{8.24}$$

[9]膏按：FEM!!!!

$$\langle u,\nu\rangle_{W^{s,2}(\Omega)} := \langle u,\nu\rangle_{W^{k,2}(\Omega)} + \sum_{|\alpha|=k}\int_\Omega\int_\Omega \frac{[D^\alpha u(x)-D^\alpha u(y)]\,[D^\alpha \nu(x)-D^\alpha \nu(y)]}{|x-y|^{d+2\kappa}}\,dx\,dy, \tag{8.25}$$
$$s=k+\kappa,\ k\in\mathbb{N}_0,\ \kappa\in(0,1).$$

对于 $s<0, 1<p<\infty$ 的情况，索伯列夫空间 $W^{s,p}(\Omega)$ 通过对偶空间$\mathring{W}^{-s,q}(\Omega)$ 形式得以定义，其中 $\frac{1}{p}+\frac{1}{q}=1$，对应范数

$$\|u\|_{W^{s,p}(\Omega)} := \sup_{0\neq\nu\in\mathring{W}^{-s,q}(\Omega)} \frac{\left|\langle u,\nu\rangle_\Omega\right|}{\|\nu\|_{W^{-s,q}(\Omega)}},$$

同样的，$\mathring{W}^{-s,p}(\Omega)$ 是 $W^{-s,q}(\Omega)$ 的对偶空间。

索伯列夫空间的性质：嵌入定理

介绍一些索伯列夫空间 $W^{s,p}(\Omega)$ 的性质，这些性质有助于更好理解下文介绍的有界元法和有限元法。第一个性质可表示为索伯列夫空间的嵌入定理，我们先从霍德尔不等式入手。

Definition 8.26 (霍德尔不等式)**.** *霍德尔不等式 (Hölder inequality)可表示为*

$$|u(x)\nu(x)|_1 \le \|u\|_p\,\|\nu\|_q,\quad 1\le p,q\le\infty, \frac{1}{p}+\frac{1}{q}=1.$$

证明. 见 Definition 8.27。 □

Definition 8.27 (广义霍德尔不等式)**.** *霍德尔不等式 (Definition 8.26中向量的维度 $n=1$。对于 $n\ge 2$ 的情况，如 $\{u_i\}_{i=1}^n$，我们有广义霍德尔不等式 (generalized Hölder inequality)*

$$\int_\Omega |\prod_{i=1}^n u(x_i)| \le \prod_{i=1}^n \|u\|_{p_i},\quad p_i\ge 1, \sum_{i=1}^n \frac{1}{p_i}=1. \tag{8.26}$$

证明. $n=1$ 时。略。

$n\ge 2$ 并且若 $(n-1)$ 的情况。已知满足霍德尔不等式条件，我们的目标就成了，希望知道 n 的情况是否依然满足。分两种情况来分析。首先来看 $n<\infty$，对应 $p_n<\infty$。设

$$p := \frac{p_n}{p_n-1},\ q := p_n \Rightarrow \frac{1}{p}+\frac{1}{q}=1.$$

此时我们有

$$\begin{aligned}\|\prod_{i=1}^n u_i\|_1 &\le \|\prod_{i=1}^{n-1} u_i\|_p \|\prod_{i=1}^n u_n\|_q \\ &\le \left(\|\prod_{i=1}^{n-1} u_i^p\|_1\right)^{\frac{1}{p}} \|u_n\|_{p_n},\end{aligned}$$

我们设 $p_i' := \frac{p_i}{p}$，上式进一步变为

$$\|\prod_{i=1}^n u_i\|_1 \le \left(\|\prod_{i=1}^{n-1} u_i^p\|_{p_i'}\right)^{\frac{1}{p}} \|u_n\|_{p_n},$$

进一步，已知

$$\left(\|\prod_{i=1}^{n-1} u_i^p\|_{p_i'}\right)^{\frac{1}{p}} \equiv \left[\prod_{i=1}^{n-1} |u_i|^{\left(p\cdot p_i'\right)}\right]^{\left(\frac{1}{p_i'}\cdot\frac{1}{p}\right)} = \prod_{i=1}^{n-1}\|u_i\|_{p_i},$$

则上式进一步变为

$$\|\prod_{i=1}^{n} u_i\|_1 \leq \prod_{i=1}^{n-1} \|u_i\|_{p_i} \|u_n\|_{p_n} = \prod_{i=1}^{n} \|u\|_{p_i}, \quad 2 \leq n < \infty.$$

$n = \infty$ 的情况，

$$\because p_n = \infty \Rightarrow \sum_{i=1}^{n-1} \frac{1}{p_i} = 1,$$

$$\begin{aligned} \therefore \|\prod_{i=1}^{n} u_i\| &\leq \|\prod_{i=1}^{n-1} u_i\|_1 \|u_n\|_\infty \\ &\leq \prod_{i=1}^{n} \|u\|_{p_i}, \quad n = \infty. \end{aligned}$$

□

第一个性质可表示为索伯列夫空间的嵌入定理 (embedding theorem of Sobolev)。

Definition 8.28 (嵌入)**.** *数学上，嵌入 (embedding) 是指某个物件 (instance) X 被嵌入到另一个物件 Y 中去，用保留结构的映射 (structure-preserving map)$f : X \mapsto Y$ 表示。这里的物件指数学结构，如群、子群等。所保留的具体数学 "结构" 因物件 X 和 Y 的种类而异。如在范畴论 (category theorem) 中，一个保留结构的映射往往称为一个态射 (morphism)。*

Theorem 8.2 (索伯列夫空间的嵌入定理 $C = C(n,p)$)**.** *给定有界的开放集 $\Omega \subset \mathbb{R}^n$，则*

$$W_0^{1,p}(\Omega) \subset L^{\frac{np}{n-p}}(\Omega), \quad n \geq 3, 1 \leq p < n,$$

并且 $W_0^{1,p}(\Omega)$ 被连续嵌入到空间 $L^{\frac{n-p}{np}}(\Omega)$ 中，以下一组关系始终满足

$$\|u\|_{L^{\frac{np}{n-p}}(\Omega)} \leq C(n,p) \|Du\|_{L^p(\Omega)}, \forall u \in W_0^{1,p}(\Omega), \tag{8.27}$$

其中 $C(n,p) \in (0,+\infty)$ 是个和 n,p 有关的常数，$Df = (D^{e_1}u, \ldots, D^{e_n}u) \in L^p(\Omega) \times \ldots \times L^p(\Omega)$。

证明. 参考 (Brenner and Scott, 2008, Theorem 1.4.6)，(McLean, 2000, Theorem 3.26)。

1. 证明对于 $\forall u \in C_0^\infty$ 来说，(8.27)成立。由广义霍德尔不等式 (Definition 8.27) 我们有，如果存在方程 $u(x) \in L^p(\Omega)$，满足 $u_j(x) \in L^{p_j}(\Omega), j = 1, \ldots, m$，并且 $\sum_{j=1}^m \frac{1}{p_j} = 1$，那么

$$\int_\Omega u_1(x) \ldots u_m(x)\, dx \leq \|u_1\|_{L^{p_1}(\Omega)} \ldots \|u_m\|_{L^{p_m}(\Omega)} \tag{8.28}$$

2. 在 $p = 1$ 时，由 $u \in C_0^\infty(\Omega)$ 可知，方程 $u(x)$ 可以表示为下述积分形式

$$\begin{aligned} u(x) &= \int_{-\infty}^{x_i} D^{e_i} u\left(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n\right) dt, \\ \hookrightarrow |u(x)| &\leq \int_{-\infty}^{x_i} \left|D^{e_i} u\right| dt \leq \int_{-\infty}^{\infty} \left|D^{e_i} u\right| dx_i, \\ \hookrightarrow |u(x)|^n &\leq \prod_{i=1}^{n} \int_{-\infty}^{\infty} \left|D^{e_i} u\right| dx_i, \end{aligned}$$

从而我们有

$$|u(x)|^{\frac{n}{n-1}} \leq \left(\prod_{i=1}^{n} \int_{-\infty}^{\infty} \left|D^{e_i} u\right| dx_i\right)^{\frac{1}{n-1}}. \tag{8.29}$$

下面对数列 $x=\{x_i\}_{i=1}^n$ 求偏导。先从 x_1 开始，根据(8.29)我们有

$$\begin{aligned}\int_{-\infty}^{\infty}|u(x)|^{\frac{n}{n-1}}\,dx_1 &\le \int_{-\infty}^{\infty}\left(\prod_{i=1}^n\int_{-\infty}^{\infty}\left|D^{e_i}u\right|dx_i\right)^{\frac{1}{n-1}}dx_1\\ &=\left(\int_{-\infty}^{\infty}\left|D^{e_1}f\right|dx_i\right)^{\frac{1}{n-1}}\int_{-\infty}^{\infty}\left(\prod_{i=2}^n\int_{-\infty}^{\infty}\left|D^{e_i}u\right|dx_i\right)^{\frac{1}{n-1}}dx_1\\ &\le\left(\int_{-\infty}^{\infty}\left|D^{e_1}f\right|dx_i\right)^{\frac{1}{n-1}}\prod_{i=2}^n\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|D^{e_i}u\right|dx_i\,dx_1\right)^{\frac{1}{n-1}}.\end{aligned}\tag{8.30}$$

在(8.30)的基础上，继续对 $x_2,\dots x_n$ 作偏导，可得

$$\int_{\mathbb{R}^n}|u(x)|^{\frac{n}{n-1}}\,dx\le\left(\prod_{i=1}^n\int_{\mathbb{R}^n}\left|D^{e_i}u\right|dx_i\right)^{\frac{1}{n-1}}.\tag{8.31}$$

(8.31) $\Rightarrow$

$$\begin{aligned}\left\|u(x)\right\|_{L^{\frac{n}{n-1}}(\Omega)}&\le\left(\prod_{i=1}^n\int_{\mathbb{R}^n}\left|D^{e_i}u\right|dx_i\right)^{\frac{1}{n}}\\ &\le\int_\Omega\left(\sum_{i=1}^n\left|D^{e_i}u\right|\right)dx\\ &\le\frac{1}{\sqrt{n}}\int_\Omega\left|Du\right|dx\\ &=\frac{1}{\sqrt{n}}\left\|Du\right\|_{L^1(\Omega)},\quad\forall u\in C_0^\infty(\Omega).\end{aligned}$$

3. 下面考虑 $1<p<n$ 的情况。已知 $u\in C_0^\infty(\Omega)$，则用 $|u|^\gamma,\gamma>1$ 代替(2)中的 u，我们有

$$\left\|\,|u|^\gamma\,\right\|_{L^{\frac{n}{n-1}}(\Omega)}\le\frac{1}{\sqrt{n}}\int_\Omega\left|D\,|u|^\gamma\,\right|dx=\frac{\gamma}{\sqrt{n}}\int_\Omega|u|^{\gamma-1}\left|Du\right|dx,\tag{8.32}$$

设 $\frac{1}{p}+\frac{1}{q}=1,\,p,q>0$，由霍德尔不等式 (Definition 8.27) 得，(8.32)$\Rightarrow$

$$\begin{aligned}&\left\|\,|u|^\gamma\,\right\|_{L^{\frac{n}{n-1}}(\Omega)}\le\frac{\gamma}{\sqrt{n}}\int_\Omega|u|^{\gamma-1}\left|Du\right|dx\le\frac{\gamma}{\sqrt{n}}\left\|\,|u|^{\gamma-1}\,\right\|_q\left\|Du\right\|_p,\\ \hookrightarrow&\left\|\,|u|\,\right\|^\gamma_{L^{\frac{\gamma n}{n-1}}(\Omega)}\le\frac{\gamma}{\sqrt{n}}\left\|\,|u|\,\right\|^{\gamma-1}_{L^{q(\gamma-1)}(\Omega)}\left\|Du\right\|_p.\end{aligned}\tag{8.33}$$

下面选取 γ 的值，使得

$$\gamma\frac{n}{n-1}\equiv(\gamma-1)q\Longrightarrow\gamma=\frac{p(n-1)}{n-p}\Longrightarrow\gamma\frac{n}{n-1}\equiv(\gamma-1)q=\frac{np}{n-p},$$

代入(8.33)，调整得

$$\left\|u\right\|_{L^{\frac{np}{n-p}}(\Omega)}\le\frac{\gamma}{\sqrt{n}}\left\|Du\right\|_p,\quad\forall u\in C_0^\infty(\Omega),$$

联系(8.27)可得 $C(n,p)$ 的值

$$C(n,p)=\frac{np-n}{\sqrt{n}(n-p)}.$$

4. 前面的论证过程均假设 $u\in C_0^\infty$。如果现在假设 $u_\ell\in W_0^{1,p}(\Omega)$，则我们利用 $u_m\in C_0^\infty(\Omega)$ 来近似 $u_\ell\in W^{1,p}(\Omega)$，将 $u_\ell-u_m$ 代入(8.27)中。可见 $\{u_m\}\in L^{\frac{np}{n-p}}(\Omega)$ 是一个柯西数列。因此我们也能证明 $u_\ell\in L^{\frac{np}{n-p}}(\Omega)$，并且满足条件(8.27)。

□

前面讨论的都是 $k=1$ 时，n,p 变化，$C=C(n,p)$。现在来将 k 的变化也考虑进来，$C=C(k,n,p)$。

Corollary 8.1 (索伯列夫空间的嵌入定理 $C=C(k,n,p)$)**.** *如果 $kp<n$, 则索伯列夫空间 $W_0^{k,p}(\Omega)$ 连续嵌入到 $L^{\frac{np}{n-kp}}(\Omega)$ 中，对应常数 $C(k,n,p)$ 满足*

$$\|u\|_{L^{\frac{np}{n-kp}}(\Omega)} \le C(k,n,p)\|u\|_{W_0^{k,p}(\Omega)}. \tag{8.34}$$

证明. 假设 $kp<n$，$u\in W_0^{k,p}(\Omega)$。

1. 根据 $D^\alpha u\in L^p(\Omega), \forall\ |\alpha|\le k$，从索伯列夫不等式可得，对于 $|\beta|\le k-1$

$$\|D^\beta u\|_{L^{p^*}(\Omega)} \le C\|u\|_{W^{k,p}(\Omega)}, \quad p^* := \frac{np}{n-p}$$

   因此有 $u\in W^{k-1,p^*}(\Omega)$。

2. 用类似的方法，我们可以进一步证明 $u\in W^{k-2,p^{**}}(\Omega)$，对于 $|\gamma|\le k-2$

$$\|D^\gamma u\|_{L^{p^{**}}(\Omega)} \le C\|u\|_{W^{k,p^*}(\Omega)}, \quad p^{**} := \frac{1}{p^*}-\frac{1}{n}=\frac{1}{p}-\frac{2}{n}.$$

3. 以此类推，最终经过 k 次迭代后，我们得以证明，对于 $u\in W^{0,q}(\Omega)=L^q(\Omega)$，(8.34)成立

$$\|u\|_{L^{\frac{np}{n-kp}}(\Omega)} \le C\|u\|_{W_0^{k,p}(\Omega)}, \quad \frac{1}{q}=\frac{1}{p}-\frac{k}{n}.$$

□

前面讨论的是将 $W_0^{k,p}(\Omega)$ 嵌入到 $L^p(\Omega)$ 中去。如果 Ω 是一个 $C^k-domain$，则我们可以利用延拓算子 E，将索伯列夫嵌入定理从 $W_0^{k,p}(\Omega)$ 延伸到 $W^{k,p}(\Omega)$ 空间。

即是说, 如果 $u\in W^{k,p}(\Omega)$, 则可以考虑一个 $Eu\in W_0^{k,p}(\Omega')$, 其中 Ω' 是 Ω 的延拓 (extension), Ω 是 Ω' 的限制 (restriction)：$\Omega'\supset\Omega$。进而，如果 $kp<n$，那么 $W_0^{k,p}(\Omega')\in L^{\frac{np}{n-kp}}(\Omega')$。

因此 u 属于限制 Ω 所对应的勒贝格空间：$u\in L^{\frac{np}{n-kp}}(\Omega)$。

一方面在 Ω 中我们有 $Eu=u$，另一方面根据 Ω 的具体情况我们有 $\|Eu\|_{W^{k,p}(\Omega')}\le C\|u\|_{W^{k,p}(\Omega)}$。所以我们有

Corollary 8.2 (索伯列夫空间的嵌入定理 $C=C(k,n,p,\Omega)$)**.** *设 $\Omega\subset\mathbb{R}^n$ 是一个有界的 $C^k-domain$。如果 $kp<n$，一系列索伯列夫空间 $W^{k,p}(\Omega)$ 被依次嵌入到对应的 $L^{\frac{np}{n-kp}}(\Omega)$ 空间中去。这几是说，存在某个常数 $C=C(k,n,p,\Omega)$，使得*

$$\|u\|_{L^{\frac{np}{n-kp}}(\Omega)} \le C\|u\|_{W^{k,p}(\Omega)}, \quad \forall\, u\in W^{1,p}(\Omega). \tag{8.35}$$

证明. □

索伯列夫空间的性质：范数等价

第二个性质是索伯列夫空间的范数等价定理 (norm equivalence theorem of Sobolev)。先来介绍范数等价的定义，以及一般意义上的范数等价定理。

Definition 8.29 (范数等价)**.** 一个向量空间 $V \in \mathbb{F} = \mathbb{R}\,\mathbb{C}$ 中有两个范数 $\|\cdot\|_a$ 和 $\|\cdot\|_b$。范数等价 *(equivalence of norms)*是指，$\|\cdot\|_b$ 总是在 $\|\cdot\|_a$ 的某乘数倍范围之内，换句话说：存在正常数 c, C 使得 $\forall x \in V$，都满足

$$c\,\|x\|_a \le \|x\|_b \le C\,\|x\|_a.$$

Theorem 8.3 (范数等价定理)**.** $\exists V \in \mathbb{F}^\infty$，其中 $\nu_1, \ldots \nu_n$ 是 V 的基，n 是 V 的维度。因此每个 $x \in V$ 都有唯一的表现形式，如

$$x = \sum_{i=1}^{n} x_i \nu_i, \tag{8.36}$$

上述内积形式中，$\{\nu_i\}_{i=1}^n$ 是标准正交基，$\{x_i\}_{i=1}^n$ 是一组关于 $\{\nu_i\}_{i=1}^n$ 的坐标 *(标量)*。

定义 V 中的范数为

$$\|x\|_* := \max_{i=1,\ldots,n} |x_i|,$$

我们将有范数 $\|x\|_*$ 的 V 称为完整向量空间 *(complete)*。

则范数等价定理为：$V \in \mathbb{F}^\infty$ 中的所有范是等价的。

证明. 使 $\|x\|$ 为 V 中任一范数。我们的目标是证明 $\|x\|$ 和 $\|x\|_*$ 等价，即存在常数 c, C 使得不等式(8.36)成立。分两步来证明。

第一步来看对 $\|x\| \le C\,\|x\|_*$ 的证明。根据定义我们有

$$\begin{aligned}
\|x\| &= \Big\|\sum_{i=1}^{n} x_i \nu_i\Big\| \\
&\le \sum_{i=1}^{n} \big\|x_i \nu_i\big\| = \sum_{i=1}^{n} |x_i|\ \|\nu_i\| \\
&\le \underbrace{n \max_{i=1,\ldots,n} \|\nu_i\|}_{\text{常数}} \max_{i=1,\ldots,n} |x_i|
\end{aligned}$$

设常数 $C := n \max_{i=1,\ldots,n} \|\nu_i\|$，则上式变为

$$\begin{aligned}
&\|x\| \le C \max_{i=1,\ldots,n} |x_i|, \\
\hookrightarrow &\|x\| \le C\,\|x\|_*.
\end{aligned}$$

第二步，来看 $c\,\|x\|_* \le \|x\|$ 的证明。首先需要证明 $S = \{x : \|x\|_* = 1\}$ 是个紧密空间 (compact space)，方法为证明 S 完整且完全有界 (totally bounded)。

完整性的证明：由三角不等式 (Definition 8.15) 得

$$\|x\|_* - \|y\|_* \le \|x - y\|_*,$$

$\hookrightarrow$ 方程 $x : x \mapsto \|x\|_*$ 是连续的。

$\hookrightarrow$ 连续方程 x 构成的向量空间 S 是 V 中的封闭子集。

$\hookrightarrow$ 已知 V 是个完整空间，$\therefore V$ 的封闭子集 S 也是个完整空间。

完全有界的证明：设 $\varepsilon > 0$，选则常数 $m > 1/\varepsilon$。因此，S 被 $O(m^d) < \infty$ 个球体 B 的集合所覆盖，这些球都是半径为 $1/m$，范数为 $\|\cdot\|_*$，表示为

$$B_{j_1,\ldots,j_n} := \left\{ \sum_{i=1}^{n} x_i \nu_i : \frac{j_i}{m} - \varepsilon < x_i < \frac{j_i}{m} + \varepsilon, \quad i = 1, \ldots, m, \quad j_i = -m, \ldots, m, \right\}$$

$\therefore S$ 完全有界。

其次，设方程 $f(x)=\|x\|, x\in[0,\infty)$。对于 $x\in S$，我们有 $f(x)$ 也是一个连续方程，这是由于

$$\begin{aligned}\big\|f(x)-f(y)\big\| &\le \big\|\|x\|-\|y\|\big\| \\ &\le \big\|x-y\big\| \\ &\le C\,\big\|x-y\big\|_*\end{aligned}$$

$\therefore$ 紧密空间 S 中的方程 f 有最小值。

最后，由 $x\in S$ 得，设 $c:=\min_{x\in S} f(x)>0$。$x\in S\Rightarrow \frac{x}{\|x\|_*}\in S\Rightarrow$

$$f\left(\frac{x}{\|x\|_*}\right)\equiv\left\|\frac{x}{\|x\|_*}\right\|\ge c,$$

$\therefore \|x\|\ge c\|x\|_*$. □

当 $k=1,p=2$ 时，索伯列夫空间 $W^{1,2}(\Omega)$ 的范数(8.23)变为

$$\begin{aligned}\|u\|_{W^{1,2}(\Omega)} &= \left\{\big\|D\,u\big\|^2_{L^2(\Omega)}\right\}^{\frac{1}{2}} \\ &= \left\{\|u\|^2_{L^2(\Omega)}+\|\nabla u\|^2_{L^2(\Omega)}\right\}^{\frac{1}{2}},\end{aligned}$$

并且半范数为

$$|u|_{W^{1,2}(\Omega)}=\|\nabla u\|_{L^2(\Omega)},$$

则我们有索伯列夫空间的范数等价定理。

Theorem 8.4 (索伯列夫空间的范数等价定理)**.** *设一个有界的线性方程 $f: W^{1,2}(\Omega)\mapsto\mathbb{R}$，满足*

$$0\le|f(u)|\le c_f\big\|u\big\|_{W^{1,2}(\Omega)},\quad \forall u\in W^{1,2}(\Omega),$$

$c_f>0$ 是个常数。如果对于某个常数 ι，$f(\iota)=0$ iff. $\iota\equiv 0$，那么在 $W^{1,2}(\Omega)$ 空间中，所有满足上述条件的方程 f 的范数是等价的：

$$\big\|u\big\|_{W^{1,2}(\Omega),f}:=\left\{|f(u)|^2+\big\|\nabla u\big\|^2_{L^2(\Omega)}\right\}^{\frac{1}{2}}. \tag{8.37}$$

证明. 首先证明 $f(u)\le c_f\big\|u\big\|_{L^2(\Omega)}$：根据前提假定，$f$ 是一个线性有界方程，则

$$\big\|u\big\|^2_{W^{1,2}(\Omega),f}=\underbrace{|f(u)|^2}_{\le c_f^2\|u\|^2_{W^{1,2}(\Omega)}}+\underbrace{\big\|\nabla u\big\|^2_{L^2(\Omega)}}_{\le\|u\|^2_{W^{1,2}(\Omega)}}\le\left(1+c_f^2\right)\big\|u\big\|^2_{L^2(\Omega)},$$

则我们有

$$f(u)\le c_f\big\|u\big\|_{L^2(\Omega)}.$$

第二步，对 $|f(u)\ge 0|$ 的证明，较为间接。已知 $\nexists c_0$ 使得满足下述条件

$$\|u\|_{W^{1,2}(\Omega)}\le c_0\underbrace{\|u\|_{W^{1,2}(\Omega),f}}_{\equiv\|f(u)\|},$$

则假设存在一个数列 $\{u_n\}_{n\in\mathbb{N}}\in W^{1,2}(\Omega)$，满足

$$n\le\frac{\big\|u_n\big\|_{W^{1,2}(\Omega)}}{\big\|u_n\big\|_{W^{1,2}(\Omega),f}}.$$

把 $\{u_n\}_{n\in\mathbb{N}}$ 标准化为 $\{\bar{u}_n\}_{n\in\mathbb{N}}$

$$\{\bar{u}_n\}_{n\in\mathbb{N}} := \frac{u_n}{\|u_n\|_{W^{1,2}(\Omega)}},$$

从而使得我们有

$$\begin{aligned}
&\|\bar{u}_n\|_{W^{1,2}(\Omega)} = 1,\\
&\|\bar{u}_n\|_{W^{1,2}(\Omega),f} = \frac{\|u_n\|_{W^{1,2}(\Omega),f}}{\|u_n\|_{W^{1,2}(\Omega)}} \le \frac{1}{n},
\end{aligned}$$

那么随着 $n\to\infty$，

$$\lim_{n\to\infty}\|\bar{u}_n\|_{W^{1,2}(\Omega),f} = \lim_{n\to\infty}|f(\bar{u}_n)| \le \lim_{n\to\infty}\left(\frac{1}{n}\right) = 0,$$

代回(8.37)有

$$\begin{aligned}
&\lim_{\bar{u}_n\to\infty}\|u\|_{W^{1,2}(\Omega),f} = \lim_{n\to\infty}\left\{|f(\bar{u}_n)|^2 + \|\nabla\bar{u}_n\|^2_{L^2(\Omega)}\right\}^{\frac{1}{2}},\\
\hookrightarrow &\lim_{n\to\infty}\|\nabla\bar{u}_n\|_{L^2(\Omega)} = 0.
\end{aligned}$$

由于标准化数列 $\{\overline{u}_n\}_{n\in\mathbb{N}}$ 在 $W^{1,2}(\Omega)$ 中有界，以及由于 $W^{1,2}(\Omega)\hookrightarrow L^2\Omega$ 的嵌入是紧凑的，则 $\exists$ 子数列 $\{\overline{u}'_n\}_{n'\in\mathbb{N}}\subset\{\overline{u}_n\}_{n\in\mathbb{N}}$，在 $L^2(\Omega)$ 中收敛。

若是定义

$$\overline{u} := \lim_{n'\to\infty}\overline{u}_{n'} \in L^2(\Omega),$$

则我们有

$$\begin{aligned}
&\overline{u}\in W^{1,2}(\Omega),\\
&\|\nabla\overline{u}\|_{L^2(\Omega)} = 0,
\end{aligned}$$

换句话说，$\overline{u}$ 是个常数。此外

$$0 \le |f(\overline{u})| = \left|f\left(\lim_{n'\to\infty}\overline{u}_{n'}\right)\right| = \lim_{n'\to\infty}|f(\overline{u}_{n'})| = 0.$$

已知 $\overline{u}$ 是个常数，$f(\overline{u})=0$，则我们有 $\overline{u}=0$。[10]

$$\therefore |f(u)| \ge |f(\overline{u})| = 0.$$

对应地，已知 $W^{1,2}(\Omega), W^{1,2}(\Gamma)$ 中的等价范分别为

$$\begin{aligned}
\|u\|_{W^{1,2}(\Omega),\Omega} &:= \left\{\left[\int_\Omega u(x)\,dx\right]^2 + \|\nabla u\|^2_{L^2(\Omega)}\right\}^{\frac{1}{2}},\\
\|u\|_{W^{1,2}(\Omega),\Gamma} &:= \left\{\left[\int_\Gamma u(x)\,ds_x\right]^2 + \|\nabla u\|^2_{L^2(\Omega)}\right\}^{\frac{1}{2}},
\end{aligned} \tag{8.38}$$

那么 $\mathring{W}^{1,2}(\Omega)$ 中的等价范为 $\|\nabla u\|_{L^2(\Omega)}$。 □

[10] 然而需要注意的是，这里有

$$\|\overline{u}\|_{W^{1,2}(\Omega)} = \lim_{n'\to\infty}\|\overline{u}_{n'}|_{W^{1,2}(\Omega)} = 1.$$

下面需要介绍庞加莱不等式。

Theorem 8.5 (庞加莱不等式 ($u \in W_0^{k,p}(\Omega)$))**.** *设一个有界域 $\Omega \subset \mathbb{R}^n$。则存在 $C_p = C(p,n,M) > 0$，使得 $\forall\, u \in W_0^{k,p}(\Omega)$ 都满足庞加莱不等式 (Poincaré inequality)*

$$\|u\|_{L^p(\Omega)} \le C_p \|\nabla u\|_{L^p(\Omega)}. \tag{8.39}$$

证明. 1. 证明对于 $\forall\, u \in C_0^\infty(\Omega)$ (8.5)成立。

2. 若 $u \in W^{1,p}(\Omega)$，则从 u 中选取某一数列 $\{u_k\} \subset C_0^\infty(\Omega)$，使得该数列随着 $k \to \infty$ 而在 $W^{1,p}$ 范数上收敛至 u，即

$$\lim_{k\to\infty} \|u_k - u\|_{L^p(\Omega)} = 0,$$
$$\lim_{k\to\infty} \|\nabla u_k - \nabla u\|_{L^p(\Omega)} = 0.$$

上式等价于

$$\lim_{k\to\infty} \|u_k\|_{L^p(\Omega)} = \|u\|_{L^p(\Omega)}, \tag{8.40a}$$
$$\lim_{k\to\infty} \|\nabla u_k\|_{L^p(\Omega)} = \|\nabla u\|_{L^p(\Omega)} \tag{8.40b}$$

3. 既然对于数列 $\{u_k\}$ 中的每一个 u_k 都满足(8.39)，则我们有以下一组不等式成立

$$\|u_k\|_{L^p(\Omega)} \le C_p \|\nabla u_k\|_{L^p(\Omega)}, \quad \forall\, u_k \in C_0^\infty,$$

从而当 $k \to \infty$ 时，(8.39)对 $\forall\, u \in C_0^\infty(\Omega)$ 都成立。

基于上述思路，来看 C_p 的值。由于在 $\Gamma = \partial\Omega$ 上有 $u(\boldsymbol{x}) = 0$，以 $u = 0$ 为边界，由散度定理 (divergence theorem, Theorem 8.1) 有

$$\begin{aligned} \int_{\Omega\cap\{u>0\}} div(u^p \boldsymbol{x})\, d\boldsymbol{x} = 0, \\ \int_{\Omega\cap\{u<0\}} div(u^p \boldsymbol{x})\, d\boldsymbol{x} = 0, \end{aligned} \tag{8.41}$$

其中

$$div(u^p \boldsymbol{x}) = p\, u\, \nabla u \cdot \boldsymbol{x} + n u^p.$$

(8.41)$\Rightarrow$

$$\begin{aligned} \int_{\Omega\cap\{u>0\}} u^p d\boldsymbol{x} = -\frac{p}{n} \int_{\Omega\cap\{u>0\}} u^{p-1} \nabla u \cdot \boldsymbol{x}\, d\boldsymbol{x}, \\ \int_{\Omega\cap\{u>0\}} u^p d\boldsymbol{x} = -\frac{p}{n} \int_{\Omega\cap\{u>0\}} u^{p-1} \nabla u \cdot \boldsymbol{x}\, d\boldsymbol{x}. \end{aligned} \tag{8.42}$$

Ω 有界 $\Rightarrow M := \max_{\boldsymbol{x}\in\Omega} |\boldsymbol{x}| < \infty$，则由施瓦茨不等式 (Schwarz inequality) 我们有

$$\begin{aligned} \int_{\Omega\cap\{u>0\}} u^p\, d\boldsymbol{x} &= \left| \frac{p}{n} \int_{\Omega\cap\{u>0\}} u^{p-1}\, \nabla u \cdot \boldsymbol{x}\, d\boldsymbol{x} \right| \\ &\le \frac{pM}{n} \left(\int_{\Omega\cap\{u>0\}} \left|u^{p-1}\right|^q d\boldsymbol{x} \right)^{\frac{1}{q}} \left(\int_{\Omega\cap\{u>0\}} \left|\nabla u\right|^p d\boldsymbol{x} \right)^{\frac{1}{p}} \\ &\le \frac{pM}{n} \left(\int_{\Omega\cap\{u>0\}} \left|u^p\right|^q d\boldsymbol{x} \right)^{\frac{1}{q}} \|\nabla u\|_{L^p(\Omega\cap\{u>0\})}, \end{aligned}$$

类似地，我们也有

$$\int_{\Omega\cap\{u<0\}} u^p\,d\boldsymbol{x} \le \frac{pM}{n}\left(\int_{\Omega\cap\{u<0\}} |u^p|^q\,d\boldsymbol{x}\right)^{\frac{1}{q}} \|\nabla u\|_{L^p(\Omega\cap\{u<0\})}.$$

上两式代回(8.39)，可得常数 C_p 的值

$$C_p = \frac{pM}{n}$$

□

上面的分析主要针对 $u \in W_0^{k,p}(\Omega)$ 的情况。我们可以进一步将庞加莱不等式扩展到 $u \in W^{k,p}(\Omega)$ 的分析中。

Corollary 8.3 (庞加莱不等式 ($u \in W^{k,p}(\Omega)$))**.** $k=1$ 时，由

$$\|u\|_{W^{1,p}(\Omega)} = \|\nabla u\|_{L^p(\Omega)},$$

可得

$$\|u\|_{W^{1,p}(\Omega)} = \left(\|u\|^p_{L^p(\Omega)} + \|\nabla u\|^p_{L^p(\Omega)}\right)^{\frac{1}{p}},$$

因此庞加莱不等式为

$$\|\nabla u\|^p_{L^p(\Omega)} \le \|u\|_{W^{1,p}(\Omega)} \le \left(C_p^p+1\right)^{\frac{1}{p}} \|\nabla u\|_{L^p(\Omega)}, \quad \forall\, u \in W^{1,p}(\Omega). \tag{8.43}$$

索伯列夫空间的 Bramble-Hilbert 引理

第三个特征是索伯列夫空间的 Bramble-Hilbert 引理，它有助于我们分析 (分段) 多项式试探空间 (trial space) 的近似属性 (Jovanović and Süli, 2014, Sec 2.3.1)。

则根据庞加莱不等式(8.5)，等价范(8.38)可以改写为

$$\int_\Omega |u(x)|^2\,dx \le c_P\left\{\left[\int_\Omega u(x)\,dx\right]^2 + \int_\Omega |\nabla u(x)|^2\,dx\right\}, \quad \forall\, u \in W^{1,2}(\Omega). \tag{8.44}$$

则我们有 Bramble-Hilbert 引理 (Bramble-Hilbert Lemma)

Lemma 8.1 (Bramble-Hilbert 引理)**.** 设 $k \in \mathbb{N}_0$，一个有界线性泛函 $f: W^{k+1,2}(\Omega) \mapsto \mathbb{R}$ 满足

$$|f(\nu)| \le c_f\,\|\nu\|_{W^{k+1,2}(\Omega)}, \quad \forall\, \nu \in W^{k+1,2}(\Omega).$$

用 $\mathcal{P}(\Omega)$ 表示所有在 Ω 中定义的 k 次多项式。如果以下条件得到满足

$$f(q) = 0, \quad \forall\, q \in \mathcal{P}(\Omega),$$

那么我们可得

$$\left|f(\nu)\right| \le C(C_p)\,C_f\,|\nu|_{W^{k+1,2}(\Omega)}, \tag{8.45}$$

其中 $C(C_p)$ 是一个与庞加莱不等式(8.44)中系数 C_p 有关的常数。

证明. 以 $k=1$ 的情况为例做出证明 ($k>1$ 的情况，证明过程与之相似)，对应 $\mathcal{P}_1(\Omega)$ 是在 Ω 中定义的线性方程空间。此时我们有

$$\left|f(\nu)\right| = \left|f(\nu)+f(q)\right| = \left|f(\nu+q)\right| \le C_f\|\nu+q\|_{W^{2,2}(\Omega)},$$

其中最后一个不等式来自于庞加莱不等式(8.44)。

由索伯列夫空间的定义有

$$\begin{aligned}\|\nu+q\|_{W^{2,2}(\Omega)}^2 &= \|\nu+q\|_{L^2(\Omega)}^2 + |\nu+q|_{W^{1,2}(\Omega)}^2 + |\nu+q|_{W^{2,2}(\Omega)}^2 \\ &= \|\nu+q\|_{L^2(\Omega)}^2 + \|\nabla(\nu+q)\|_{L^2(\Omega)}^2 + |\nu|_{W^{2,2}(\Omega)}^2,\end{aligned}$$

其中最后一个等式，根据线性方程 $q(x)\in\mathcal{P}_1(\Omega)$ 的二阶导数为 0。再次使用庞加莱不等式(8.44)，上式变为

$$\begin{aligned}\|\nu+q\|_{W^{2,2}(\Omega)}^2 &= \|\nu+q\|_{L^2(\Omega)}^2 + \|\nabla(\nu+q)\|_{L^2(\Omega)}^2 + |\nu|_{W^{2,2}(\Omega)}^2 \\ &\le C_p\left\{\left[\int_\Omega(\nu+q)\right]^2 + \|\nabla(\nu+q)\|_{L^2(\Omega)}^2\right\} + \|\nabla(\nu+q)\|_{L^2(\Omega)}^2 + |\nu|_{W^{2,2}(\Omega)}^2 \\ &= C_p\left[\int_\Omega(\nu+q)\right]^2 + (1+C_p)\underbrace{\|\nabla(\nu+q)\|_{L^2(\Omega)}^2} + |\nu|_{W^{2,2}(\Omega)}^2,\end{aligned} \tag{8.46}$$

对划线标记部分再次使用庞加莱不等式，有

$$\begin{aligned}\|\nabla(\nu+q)\|_{L^2(\Omega)}^2 &= \sum_{i=1}^d\int_\Omega\left|\frac{\partial}{\partial x_i}[\nu(x)+q(x)]\right|^2dx \\ &\le C_p\sum_{i=1}^d\left\{\left[\int_\Omega\frac{\partial}{\partial x_i}[\nu(x)+q(x)]\,dx\right]^2 + \sum_{j=1}^d\int_\Omega\left[\frac{\partial^2}{\partial x_i\,\partial x_j}[\nu(x)+q(x)]\right]^2dx\right\} \\ &= C_p\sum_{i=1}^d\left[\int_\Omega\frac{\partial}{\partial x_i}[\nu(x)+q(x)]\,dx\right]^2 + C_p\sum_{i=1}^d\sum_{j=1}^d\int_\Omega\left[\frac{\partial^2}{\partial x_i\,\partial x_j}[\nu(x)+q(x)]\right]^2dx \\ &= C_p\sum_{i=1}^d\left[\int_\Omega\frac{\partial}{\partial x_i}[\nu(x)+q(x)]\,dx\right]^2 + C_p|\nu|_{W^{2,2}(\Omega)}^2,\end{aligned}$$

代回(8.46)变为

$$\begin{aligned}\|\nu+q\|_{W^{2,2}(\Omega)}^2 \le & C_p\left[\underbrace{\int_\Omega(\nu(x)+q(x))\,dx}\right]^2 + [(1+C_p)C_p]\sum_{i=1}^d\left[\underbrace{\int_\Omega\frac{\partial}{\partial x_i}[\nu(x)+q(x)]\,dx}\right]^2 \\ & + [1+(1+C_p)C_p]\,|\nu|_{W^{2,2}(\Omega)}^2.\end{aligned} \tag{8.47}$$

给定 $u(x)=\nu(x)+q(x)$，Bramble-Hilbert 引理(8.44)若要成立，需要在(8.47)中，通过选择 $q(x)\in\mathcal{P}_1(\Omega)$ 的值，使得满足下两个条件

$$\int_\Omega[\nu(x)+q(x)]\,dx = 0, \tag{8.48a}$$

$$\int_\Omega\frac{\partial}{\partial x_i}[\nu(x)+q(x)]\,dx = 0. \tag{8.48b}$$

这涉及到系数 a_0 和 $a_i, i=1,\ldots,d$ 的值

$$q(x) = a_0 + \sum_{i=1}^d a_i x_i.$$

(8.48b) ⇒

$$\begin{aligned}
&\int_\Omega \frac{\partial}{\partial x_i}\nu(x)\,dx + \int_\Omega \frac{\partial}{\partial x_i}q(x)\,dx = 0,\\
\hookrightarrow &\int_\Omega \frac{\partial}{\partial x_i}\nu(x)\,dx + |\Omega|\,a_i = 0,\\
\hookrightarrow &a_i = -\frac{1}{|\Omega|}\int_\Omega \frac{\partial}{\partial x_i}\nu(x)\,dx, \quad i = 1,\dots,d.
\end{aligned}$$

(8.48b)⇒

$$\begin{aligned}
&\int_\Omega \left[\nu(x) + a_0 + \sum_{i=1}^d a_i x_i\right] dx = 0,\\
\hookrightarrow &+|\Omega|\,a_0 + \int_\Omega \left[\nu(x) + \sum_{i=1}^d a_i x_i\right] dx = 0,\\
\hookrightarrow &a_0 = -\frac{1}{|\Omega|}\int_\Omega \left[\nu(x) + \sum_{i=1}^d a_i x_i\right] dx.
\end{aligned}$$

对于 $k \in \mathbb{N}$ 的情况，证明过程同上。 □

索伯列夫空间的分布

由第8.2.4节的介绍可见，并不是 $L^{loc}(\Omega)$ 中的所有方程都有广义偏导数。然而可以从分布的角度来重新解读广义偏导数 (McLean, 2000)。

Definition 8.30 (分布)**.** 对于 $\Omega \subseteq \mathbb{R}^d$，定义一个试探空间 *(test space)* $\mathcal{D}(\Omega) := C_0^\infty(\Omega)$，空间中的一个复值连续线性泛函 $T \in \mathcal{D}(\Omega)$ 可以称为一个分布 *(distribution)*。T 在 $\mathcal{D}(\Omega)$ 中连续，是指对于 $x \in \Omega$ 中的任何 $\varphi_k(x) \to \varphi(x)$，总有 $\mathcal{D}'(\Omega)$ 中的 $T(\varphi_k) \to T(\varphi)$；$\mathcal{D}'(\Omega) \subset \mathcal{D}(\Omega)$ 表示 $\mathcal{D}(\Omega)$ 空间中，所有分布的集合。

对于 $x \in \Omega, u(x) \in L^{1,loc}(\Omega)$，分布常常定义为

$$T_u(\varphi) := \int_\Omega u(x)\varphi(x)\,dx, \quad \varphi \in \phi(\Omega). \tag{8.49}$$

符合(8.49)形式的分布，又称正则分布 *(regular distribution)*。

正则分布中，局部可积方程 $u(x) \in L^{1,loc}(\Omega)$ 由子集空间 $\mathcal{D}'(\Omega)$ 予以识别。因此，$T_u(\varphi) \in \mathcal{D}'(\Omega)$ 有时也简写为 $u \in \mathcal{D}'(\Omega)$。

不符合正则类型的分布，称为奇异分布 *(singular distribution)*。一种奇异分布的例子是狄拉克分布 *(Dirac distribution)*:

$$\delta_{x_0}(\varphi) = \varphi(x_0), \quad x_0 \in \Omega,\ \varphi \in \mathcal{D}(\Omega).$$

现在我们回到式(8.21)，来看如何计算 $u(x)$ 的广义偏导数 $\nu(x) := \frac{\partial u(x)}{\partial x} = \operatorname{sgn}(x)$。根据分部积分法有

$$\int_{-1}^1 \nu(x)\frac{\partial}{\partial x}\varphi(x)\,dx = -2\varphi(0) = -\int_{-1}^1 \frac{\partial}{\partial x}\nu(x)\varphi(x)\,dx, \quad \forall \varphi(x) \in \mathcal{D}(\Omega),$$

可见分布意义上的 $\nu(x)$ 的广义偏导数变为

$$\frac{\partial}{\partial x}\nu(x) = 2\varphi(0) = \delta_0(\varphi) \in \mathcal{D}'(\Omega),$$

更高阶的广义偏导数为

$$(D^\alpha T_u)(\varphi) = (-1)^{|\alpha|} T_u (D^\alpha \varphi), \quad \varphi \in \mathcal{D}(\Omega).$$

在掌握了分布的基本概念之后，下面介绍在满足一些关于 Ω 的正则条件假设的情况下，$H^s(\Omega)$ 空间近似等价于 $W^{s,2}(\Omega)$ 空间，其中 $H^s(\Omega)$ 的基是分布的傅里叶变换[11]。我们从急减方程 (rapidly decreasing functions) 空间 $\mathcal{S}\left(\mathbb{R}^d\right)$ 开始，进而介绍傅里叶变换。

Definition 8.31 (急减方程空间)**.** *我们将满足以下形式的方程 $\varphi(x) \in C^\infty(\mathbb{R}^d)$ 称为急减方程 (rapidly decreasing function):*

$$\|\varphi\|_{k,\ell} := \sup_{x\in\mathbb{R}^d} \left(|x|^k\right) \sum_{|\alpha|\le\ell} \left|D^\alpha \varphi(x)\right| < \infty, \quad k, \ell \in \mathbb{N}_0,$$

即是说，急减方程 $\varphi(x)$ 和它的导数，比任意多项式的减速更快。

急减方程组成的空间，称为急减空间 (rapidly decreasing space)，表示为 $\mathcal{S}(\mathbb{R}^d)$。

类似于 $\mathcal{D}'(\Omega) \subset \mathcal{D}(\Omega)$，我们定义缓增分布空间 (tempered distribution space) $\mathcal{S}'(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$，作为全部复值线性泛函 $T \in \mathcal{S}(\mathbb{R}^d)$ 组成的子集。

举例来说，对于方程 $\varphi(x) := \exp(-|x|^2)$，我们有

$$\varphi \in \mathcal{S}(\mathbb{R}^d), \quad \varphi \notin \mathcal{D}(\Omega) = C_0^\infty(\mathbb{R}^d).$$

Definition 8.32 (傅里叶变换)**.** *对于方程 $\varphi(x) \in \mathcal{S}(\mathbb{R}^d)$，我们将傅里叶变换 (Fourier transform) $\widehat{\varphi}(x) \in \mathcal{S}(\mathbb{R}^d)$ 定义如下*

$$\widehat{\varphi}(\xi) := (\mathcal{F}\varphi)(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left[-i\langle x, \xi\rangle\right] \varphi(x)\, dx, \quad \xi \in \mathbb{R}^d, \tag{8.50}$$

其中 $\mathcal{F}$ 是可逆映射 $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \mapsto \mathcal{S}(\mathbb{R}^d)$。

对应地，逆傅里叶变换 (inverse Fourier transform)为

$$\left(\mathcal{F}^{-1}\widehat{\varphi}\right)(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left[i\langle x, \xi\rangle\right] \widehat{\varphi}(\xi)\, dx, \quad x \in \mathbb{R}^d. \tag{8.51}$$

通常来说，$\varphi \in \mathcal{D}(\mathbb{R}^d)$ 并不必然意味着 $\widehat{\varphi} \in \mathcal{D}(\mathbb{R}^d)$。

此外对于 $\varphi \in \mathcal{S}(\mathbb{R}^d)$，我们进一步有

$$D^\alpha (\mathcal{F}\varphi)(\xi) = (-i)^{|\alpha|} \mathcal{F}(x^\alpha \varphi)(\xi), \tag{8.52a}$$

$$\xi^\alpha (\mathcal{F}\varphi)(\xi) = (-i)^{|\alpha|} \mathcal{F}(D^\alpha \varphi)(\xi). \tag{8.52b}$$

Lemma 8.2 (傅里叶变换的旋转对称)**.** *傅里叶变换保持旋转对称结构 (rotational symmetries)，即对于 $u \in \mathbb{R}^d$，我们有*

$$\widehat{u}(\xi) = \widehat{u}(|\xi|), \forall\, \xi \in \mathbb{R}^d, \quad \textit{iff.}\ \ u(x) = u(|x|), \forall\, x \in \mathbb{R}^d.$$

证明. 先来看 $d = 2$ 的证明，对应极坐标系 (polar coordinates)

$$\xi = \begin{pmatrix} |\xi| \cos\psi \\ |\xi| \sin\psi \end{pmatrix}, \quad x = \begin{pmatrix} r \cos\phi \\ r \sin\phi \end{pmatrix}.$$

[11]傅里叶分析的初步介绍，见第7.G节。

根据(8.50)我们有

$$
\begin{aligned}
\widehat{u}(\xi) &= \widehat{u}(|\xi|, \psi) \\
&= \frac{1}{2\pi}\int_0^\infty \int_0^{2\pi} \exp\{-i\,r\,|\xi|\,[\cos\phi\,\cos\psi + \sin\phi\,\sin\psi]\}\,u(r)\,r\,d\phi\,dr \\
&= \frac{1}{2\pi}\int_0^\infty \underbrace{\int_0^{2\pi} \exp\{-i\,r\,|\xi|\,[\cos(\phi-\psi)]\}\,u(r)\,r\,d\phi}_{:=\mathcal{A}}\,dr.
\end{aligned}
\tag{8.53}
$$

现在将坐标旋转 $\psi_0 \in [0, 2\pi]$ 度，变为 $\psi + \psi_0$，定义 $\widetilde{\phi} := \phi - \psi_0$，(8.53)变为

$$
\begin{aligned}
\widehat{u}(|\xi|, \psi+\psi_0) &= \frac{1}{2\pi}\int_0^\infty \int_0^{2\pi} \exp\{-i\,r\,|\xi|\,[\cos(\phi-\psi_0-\psi)]\}\,u(r)\,r\,d\phi\,dr \\
&= \frac{1}{2\pi}\int_0^\infty \underbrace{\int_{-\psi_0}^{2\pi-\psi_0} \exp\left\{-i\,r\,|\xi|\,\left[\cos\left(\widetilde{\phi}-\psi\right)\right]\right\}\,u(r)\,r\,d\widetilde{\phi}}_{:=\mathcal{B}}\,dr
\end{aligned}
\tag{8.54}
$$

比较上两式

$$
\begin{aligned}
&\int_{-\psi_0}^{0} \exp\left[-i\,r\,|\xi|\,\cos\left(\widetilde{\phi}-\psi\right)\right]\,d\widetilde{\phi} = \int_{2\pi-\psi_0}^{2\pi} \exp\left[-i\,r\,|\xi|\,\cos\left(\widetilde{\phi}-\psi\right)\right]\,d\widetilde{\phi}, \\
&\hookrightarrow \mathcal{A} = \mathcal{B}, \\
&\hookrightarrow \widehat{u}(|\xi|, \psi) = \widehat{u}(|\xi|, \psi+\psi_0), \quad \forall \psi_0 \in [0, 2\pi], \\
&\hookrightarrow \widehat{u}(\xi) = \widehat{u}(|\xi|).
\end{aligned}
\tag{8.55}
$$

再来看 $d = 3$ 的证明，对应三维坐标系 (spherical coordinates)

$$
\xi = \begin{pmatrix} |\xi|\,\cos\psi\,\sin\vartheta \\ |\xi|\,\sin\psi\,\sin\vartheta \\ |\xi|\,\cos\vartheta \end{pmatrix}, \quad x = \begin{pmatrix} r\,\cos\phi\,\sin\theta \\ r\,\sin\phi\,\sin\theta \\ r\,\cos\theta \end{pmatrix}.
$$

根据(8.50)我们有

$$
\begin{aligned}
\widehat{u}(|\xi|, \psi, \vartheta) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int_0^\infty \int_0^{2\pi} \int_0^{\pi} &\exp\{-i\,r\,|\xi|\,[\cos(\phi-\psi)\,\sin\theta\,\sin\vartheta + \cos\theta\,\cos\vartheta]\} \\
&u(r)\,r^2\,\sin\theta\,d\theta\,d\phi\,dr.
\end{aligned}
$$

由 $d = 2$ 时证得的(8.55)，我们有

$$
\widehat{u}(|\xi|, \psi+\psi_0, \vartheta) = \widehat{u}(|\xi|, \psi, \vartheta), \quad \forall\,\psi_0 \in [0, 2\pi).
$$

给定半径 ϱ 和一个 $\vartheta \in [0, \pi)$，我们有 $\widehat{u}(\xi) = \widehat{u}(|\xi|) = \widehat{u}(\varrho)$

$$
\begin{aligned}
\xi_1^2 + \xi_2^2 &= \varrho^2 \sin^2\vartheta, \\
\xi_3 &= \varrho\cos\vartheta.
\end{aligned}
$$

对原三维坐标作变形，我们有 $\widehat{u}(\xi) = \widehat{u}(\varrho)$

$$
\begin{aligned}
\xi_1^2 + \xi_3^2 &= \varrho^2 \sin^2\vartheta, \\
\xi_2 &= \varrho\cos\vartheta.
\end{aligned}
$$

□

Definition 8.33 (分布的傅里叶变换)**.** 对于某一个缓增分布空间中的分布 $T \in \mathcal{S}'(\mathbb{R}^d)$，其傅里叶变换 $\widehat{T} \in \mathcal{S}'(\mathbb{R}^d)$ 可定义如下

$$\widehat{T}(\varphi) := T(\widehat{\varphi}), \quad \varphi \in \mathcal{S}(\mathbb{R}^d).$$

从 $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \mapsto \mathcal{S}(\mathbb{R}^d)$ 是一个可逆映射，逆傅里叶变换定义为

$$\left(\mathcal{F}^{-1}T\right)(\varphi) := T\left(\mathcal{F}^{-1}\varphi\right), \quad \varphi \in \mathcal{S}(\mathbb{R}^d).$$

分布的傅里叶变换，也满足(8.52a)-(8.52b)。

Definition 8.34 (贝塞尔位势空间)**.** 对于 $s \in \mathbb{R}, u \in \mathcal{S}(\mathbb{R}^d)$，定义一个有界的线性贝塞尔位势算子 *(Bessel potential operator)* $\mathcal{J} : \mathcal{S}(\mathbb{R}^d) \mapsto \mathcal{S}(\mathbb{R}^d)$

$$\mathcal{J}^s u(x) := (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \left(1 + |\xi|^2\right)^{\frac{s}{2}} \widehat{u}(\xi) \exp\left[i \langle x, \xi \rangle\right] d\xi, \quad x \in \mathbb{R}^d. \tag{8.56}$$

由傅里叶变换可得

$$(\mathcal{F}\mathcal{J}^s u)(\xi) = \underbrace{\left(1 + |\xi|^2\right)^{\frac{s}{2}}}(\mathcal{F}u)(\xi),$$

可见在傅里叶空间中，$\mathcal{J}^s$ 执行的操作近似于傅里叶乘子 (Fourier multiplier)，即对原式乘以一个与 $|\xi|^s$ 有关的方程 $\mathcal{O}(|\xi|^s)$。从这个意义上来讲，类似于(8.52b)，我们可以将贝塞尔算子 $\mathcal{J}^s$ 近似看做 s 阶微分符号。

Definition 8.35 (分布意义上的索伯列夫空间)**.** 在缓增分布空间 *(tempered distribution space)* $\mathcal{S}'(\mathbb{R}^d)$ 中，给定分布 $T \in \mathcal{S}'(\mathbb{R}^d)$。定义有界线性算子 $\mathcal{J}^s : \mathcal{S}'(\mathbb{R}^d) \mapsto \mathcal{S}'(\mathbb{R}^d)$

$$(\mathcal{J}^s T)(\varphi) := T(\mathcal{J}^s \varphi), \quad \varphi \in \mathcal{S}(\mathbb{R}^d).$$

进而，分布意义上的索伯列夫空间 $H^s(\mathbb{R}^d)$ 可以表示为，由全部分布 $\nu \in \mathcal{S}'(\mathbb{R}^d)$ 所组成的内积空间，其中 ν 满足 $\mathcal{J}^s\nu \in L^2(\mathbb{R}^d)$，内积形式表示为

$$\langle u, \nu \rangle_{H^s(\mathbb{R}^d)} := \langle \mathcal{J}^s u, \mathcal{J}^s \nu \rangle_{L^2(\mathbb{R}^d)},$$

对应范数

$$\|u\|^2_{H^s(\mathbb{R}^d)} := \|\mathcal{J}^s u\|^2_{H^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} (1 + |\xi|^s)^s \, |\widehat{u}(\xi)|^2 \, d\xi.$$

Theorem 8.6 (索伯列夫空间)**.** 对于所有 $s \in \mathbb{R}$，下式均成立

$$H^s(\mathbb{R}^d) = W^{s,2}(\mathbb{R}^d).$$

对于有界域 $\Omega \in \mathbb{R}^d$，我们将索伯列夫空间 $H^s(\mathbb{R}^d)$ 定义如下形式

$$H^s(\Omega) := \left\{\nu = \widetilde{\nu}|_\Omega : \widetilde{\nu} \in H^s(\mathbb{R}^d)\right\},$$
$$\|\nu\|_{H^s(\Omega)} := \inf_{\widetilde{\nu} \in H^s(\mathbb{R}^d), \widetilde{\nu}|_\Omega = \nu} \|\widetilde{\nu}\|_{H^s(\mathbb{R}^d)}.$$

此外还可以对索伯列夫空间作如下定义

$$\widetilde{H}^s(\Omega) := \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{H^s(\mathbb{R}^d)}},$$
$$H_0^s(\Omega) = \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{H^s(\Omega)}},$$

根据 *(McLean, 2000, Theorem 3.33)*，上述定义对于几乎全部 $s \in \mathbb{R}_+$ 都成立。

Theorem 8.7. 假定有一个利普希茨域 *(Definition 8.2)* $\Omega \in \mathbb{R}^d$。对于 $s \geq 0$ 我们有

$$\widetilde{H}^s(\Omega) \in H_0^s(\Omega).$$

进一步有

$$\widetilde{H}^s(\Omega) = H_0^s(\Omega), \quad s \notin \left\{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots\right\}.$$

进一步有

$$\widetilde{H}^s(\Omega) = \left[H^{-s}(\Omega)\right]', H^s(\Omega) = \left[\widetilde{H}^{-s}(\Omega)\right]', \quad \forall s \in \mathbb{R}.$$

Definition 8.36 (均匀法锥条件)**.** 两个索伯列夫空间 $H^s(\Omega)$ 和 $W^{s,2}(\Omega)$ 相等，即范式等价的充分条件是存在一个有界的线性延拓算子

$$E_\Omega : W^{2,s}(\Omega) \mapsto W^{2,s}(\mathbb{R}^d),$$

这要求存在一个有边界的域 $\Omega \in \mathbb{R}^d$，这称为均匀法锥条件 *(uniform cone condition)*。

证明. 略。见 (Adolfson et al., 2008b, Theorem 4.6, 4.7)。 □

Theorem 8.8. 对于有界的利普希茨域 $\Omega \subset \mathbb{R}^d$，我们有

$$\begin{aligned}
\|q\|_{L^2(\Omega)} &\leq C_1 \left\{\|q\|_{H^{-1}(\Omega)} + \|\nabla q\|_{[H^{-1}(\Omega)]^d}\right\}, \quad \forall q \in L^2(\Omega), \\
\|q\|_{L^2(\Omega)} &\leq C_2 \|\nabla q\|_{[H^{-1}(\Omega)]^d}, \quad \forall q \in L^2(\Omega), \int_\Omega q(x)\, dx = 0.
\end{aligned}$$

我们可以用有界的线性算子表现从一个索伯列夫空间到另一个索伯列夫空间的映射，这涉及到插值空间。关于插值空间的介绍可见如Adolfson et al. (2008b)。这里介绍与有限元法关系较为密切的插值定理 (interpolation theorem)。

Theorem 8.9 (插值定理)**.** 设一个有界的线性算子 $A : H^{\alpha_1}(\Omega) \to H^\beta(\Omega)$，其范数为

$$\|A\|_{\alpha_1,\beta} := \sup_{0 \neq \nu \in H^{\alpha_1}(\Omega)} \frac{\|A\nu\|_{H^\beta(\Omega)}}{\|\nu\|_{H^{\alpha_1}(\Omega)}}.$$

现在假设 $\alpha_2 > \alpha_1$，对应有界线性算子 $A : H^{\alpha_2}(\Omega) \to H^\beta(\Omega)$，范数

$$\|A\|_{\alpha_2,\beta} := \sup_{0 \neq \nu \in H^{\alpha_2}(\Omega)} \frac{\|A\nu\|_{H^\beta(\Omega)}}{\|\nu\|_{H^{\alpha_2}(\Omega)}}.$$

则 $A : H^\alpha(\Omega) \mapsto H^\beta(\beta), \forall \alpha \in [\alpha_1, \alpha_2]$ 都有界，对应范数

$$\|A\|_{\alpha,\beta} \leq \left(\|A\|_{\alpha_1,\beta}\right)^{\frac{\alpha-\alpha_2}{\alpha_1-\alpha_2}} \left(\|A\|_{\alpha_2,\beta}\right)^{\frac{\alpha-\alpha_1}{\alpha_2-\alpha_1}}.$$

同样，设 $A : H^\alpha \mapsto H^{\beta_1(\Omega)}$ 有界，对应范数 $\|A\|_{\alpha,\beta_1}$。设 $A : H^\alpha \mapsto H^{\beta_2(\Omega)}$ 有界，对应范数 $\|A\|_{\alpha,\beta_2}$，$\beta_1 < \beta_2$。

则 $A : H^\alpha \mapsto H^{\beta(\Omega)}$ 有界 $\forall \beta \in [\beta_1, \beta_2]$，对应范数

$$\|A\|_{\alpha,\beta} \leq \left(\|A\|_{\alpha,\beta_1}\right)^{\frac{\beta-\beta_2}{\beta_1-\beta_2}} \left(\|A\|_{\alpha,\beta_2}\right)^{\frac{\beta-\beta_1}{\beta_2-\alpha_1}}.$$

流形索伯列夫空间

Definition 8.37 (流形)**.** *流形 (manifold) 是局部具有欧式空间性质的空间，包括各种纬度的曲线曲面，例如球体、弯曲的平面等。流形的局部和欧式空间是同构的 (局部线性)。流形学习假设所处理的数据点分布在嵌入于外维欧式空间的一个潜在的流形体上，或者说这些数据点可以构成这样一个潜在的流形体*[12]。

设一个有边界的域 $\Omega \subset \mathbb{R}^d, d = 2, 3$，其边界用 $\Gamma = \partial\Omega$ 来表示。对 Γ 作局部分段参数化 (piecewise parameterization) 处理，对应一组低维空间 $\mathbb{R}^{d-1}$ 中的参数域 (parameter domain) $\mathcal{T}_i$：

$$\begin{aligned} &\Gamma = \bigcup_{i=1}^{J} \Gamma_i, \\ &\Gamma_i := \left\{ x \in \mathbb{R}^d : x = \chi_i(\xi), \xi \in \mathcal{T}_i \subset \mathbb{R}^{d-1} \right\}. \end{aligned} \tag{8.57}$$

与(8.57)同步，考虑一个单位数列 $\{\varphi_i\}_{i=1}^{p}$，其中 $\varphi_i \in C_0^\infty(\mathbb{R}^d)$ 是非负的 cutoff 方程，满足

$$\begin{cases} \sum_{i=1}^{J} \varphi_i(x) = 1 & \text{对于 } x \in \Gamma \\ \varphi_i(x) = 0 & \text{对于 } x \in \Gamma \backslash \Gamma_i. \end{cases} \tag{8.58}$$

进而，在 Γ 中定义一个方程 $\nu(x)$，满足

$$\begin{aligned} &\nu(x) = \sum_{i=1}^{J} \varphi_i(x)\nu(x) = \sum_{i=1}^{J} \nu_i(x), \quad \text{对于 } x \in \Gamma, \\ &\nu_i(x) := \varphi_i(x)\,\nu(x). \end{aligned} \tag{8.59}$$

将关于 x 的局部分段参数化(8.57)代入 $\nu_i(x)$ 的定义式，我们可以定义一组新的方程 $\{\widetilde{\nu}_i(\xi)\}$，其中

$$\begin{aligned} \nu_i(x) &= \varphi_i(x)\,\nu(x) \\ &= \varphi_i\left(\chi_i(\xi)\right)\,\nu\left(\chi_i(\xi)\right) \\ &=: \widetilde{\nu}_i(\xi), \quad \xi \in \mathcal{T}_i \subset \mathbb{R}^{d-1}. \end{aligned}$$

通过这种方法，$\Omega \subset \mathbb{R}^d$ 中关于变量 x 的问题 $\nu(x)$，被表示为分段参数域 $\mathcal{T}_i \subset \mathbb{R}^{d-1}$ 中关于局部参数 $\chi_i(\xi)$ 的问题 $\widetilde{\nu}_i(\xi)$，后者使得我们可能建立相应的索伯列夫空间，进行分析。这要求 $\widetilde{\nu}_i$ 可导，即局部参数 $\chi_i(\xi)$ 可导，满足链式法则 (chain rule)[13]。如果要求 $|s| \le k$ 次的导数，需要假定局部参数 $\chi_i \in C^{k-1,1}(\mathcal{T}_i)$，例如，对于利普希茨域 $\mathcal{T}_i$ 中的局部参数 $\chi_i \in C^{0,1}(\mathcal{T}_i)$，我们需要研究对应的索伯列夫空间 $H^s(\mathcal{T}_i), |s| \le 1$。

Definition 8.38 (流形索伯列夫空间 $(0 \le s \le k)$)**.** *进而对于 $0 \le s \le k$，我们的研究对象是整个索伯列夫空间 $H^s(\Gamma)$，其范数定义为*

$$\|\nu\|_{H^s_\chi(\Gamma)} := \left\{ \sum_{i=1}^{J} \|\widetilde{\nu}_i\|^2_{H^s(\mathcal{T}_i)} \right\}^{\frac{1}{2}}. \tag{8.61}$$

[12]从词源上来看，最早由黎曼给出了德文名称 mannigfaltigkeit，英文翻译为 manifold，字面意义为“多层”。中国第一个拓扑学家江泽涵把这个词翻译为“流形”，取自文天祥《正气歌》，“天地有正气，杂然赋流形”，其原始出处为《易经》，“大哉乾元，万物资始，乃统天。云行雨施，品物流形。”

[13]简单说来，链式法则可以表示为

$$f(g(x))' = f'(g(x)) \cdot g'(x). \tag{8.60}$$

Lemma 8.3 (流形索伯列夫空间的等价范 ($s=0$))**.** $s=0$ 时，流形索伯列夫空间 $H^0_\chi(\Gamma)$ 的等价范为

$$\|\nu\|_{L^2(\Gamma)} := \left\{ \int_\Gamma \left|\nu(x)\right|^2 ds_x \right\}^{\frac{1}{2}}.$$

证明. 首先来看 $H^0_\chi(\Gamma)$ 空间。$s=0$ 时(8.61)⇒

$$\begin{aligned}\|\nu\|^2_{H^0_\chi(\Gamma)} &= \sum_{i=1}^J \|\widetilde{\nu}_i\|^2_{H^0(\mathcal{T}_i)} \\ &= \sum_{i=1}^J \int_{\Gamma_i} \left[\varphi_i(\chi_i(\xi))\,\nu(\chi_i(\xi))\right]^2 d\xi.\end{aligned}$$

再来看 $L^2(\Gamma)$ 空间 ⇒

$$\begin{aligned}\|\nu\|^2_{L^2(\Gamma)} &= \int_\Gamma [\nu(x)]^2\, ds_x \\ &= \sum_{i=1}^J \int_{\Gamma_i} \varphi_i(x)\ [\nu(x)]^2\, ds_x \\ &= \sum_{i=1}^J \int_{\Gamma_i} \varphi_i(\chi_i(\xi))\ [\nu(\chi_i(\xi))]^2\ (\det \chi_i(\xi))\ d\xi.\end{aligned}$$

□

Lemma 8.4 (流形索伯列夫空间的等价范 ($0<s<1$))**.** $0<s<1$ 时，流形索伯列夫空间 $H^s_\chi(\Gamma)$ 的等价范，可以用 *Sobolev-Slobodeckij* 范数 *(Section 8.2.4)*来表示

$$\|\nu\|_{H^s(\Gamma)} := \left\{ \|\nu\|^2_{L^2(\Gamma)} + \int_\Gamma \int_\Gamma \frac{[\nu(x)-\nu(y)]^2}{|x-y|^{d-1+2s}}\, ds_x\, ds_y \right\}^{\frac{1}{2}},$$

需要指出的是，上式并不是唯一一种定义等价范的方法。其他方法如，根据索伯列夫范数等价定理 *(Theorem 8.4)*，流形索伯列夫空间 $H^{s=1/2}(\Gamma)$ 的等价范也可以写作

$$\|\nu\|_{H^{\frac{1}{2}}(\Gamma),\Gamma} := \left\{ \left[\int_\Gamma \nu(x)\, ds_x\right]^2 + \int_\Gamma \int_\Gamma \frac{[\nu(x)-\nu(y)]^2}{|x-y|^d}\, ds_x\, ds_y \right\}^{\frac{1}{2}}.$$

Lemma 8.5 (流形索伯列夫空间的等价范 ($s<0$))**.** $s<0$ 时的流形索伯列夫空间 $H^s_\chi(\Gamma)$，可由其对偶空间 *(dual space)* $H^{-s}_\chi(\Gamma)$ 进行分析

$$H^s(\Gamma) := \left[H^{-s}(\Gamma)\right]',$$

其范数为

$$\begin{aligned}\|w\|_{H^s(\Gamma)} &:= \sup_{0\neq\nu\in H^{-s}(\Gamma)} \frac{\langle w,\nu\rangle_\Gamma}{\|\nu\|_{H^{-s}(\Gamma)}} \\ &= \sup_{0\neq\nu\in H^{-s}(\Gamma)} \frac{\int_\Gamma w(x)\,\nu(x)\, ds_x}{\|\nu\|_{H^{-s}(\Gamma)}},\end{aligned} \tag{8.62}$$

其中 $\langle w,\nu\rangle_\Gamma$ 表示对偶配对 *(duality pairing)*。

Lemma 8.6 (流形索伯列夫空间的开放子集)**.** 对于充分平滑的边界 $\Gamma=\partial\Omega$，设一个开放子集 $\Gamma_0\subset\Gamma$。则 $H^s(\Gamma_0)$ 可以定义如下

1. $s \geq 0$ 时的索伯列夫空间 $H^s(\Gamma_0)$ 及范数

$$H^s(\Gamma_0) := \left\{ \nu = \widetilde{\nu}\big|_{\Gamma_0} : \widetilde{\nu} \in H^s(\Gamma) \right\},$$
$$\|\nu\|_{H^s(\Gamma_0)} := \inf_{\widetilde{\nu} \in H^s(\Gamma): \widetilde{\nu}|_{\Gamma_0} = \nu} \|\widetilde{\nu}\|_{H^s(\Gamma)}$$

2. $s < 0$ 时，首先定义一个索伯列夫空间 $\widetilde{H}^s(\Gamma_0)$

$$\widetilde{H}^s(\Gamma_0) := \left\{ \nu = \widetilde{\nu}\big|_{\Gamma_0} : \widetilde{\nu} \in H^s(\Gamma), \operatorname{supp} \widetilde{\nu} \subset \Gamma_0 \right\},$$

其中 supp 表示支撑集 *(support)*，是一个定义在集合 Γ_0 上的实值函数 $\widetilde{\nu}$ 的子集，满足 $\widetilde{\nu}$ 恰好在这个子集上非零[14]。进而利用对偶空间的属性

$$H^s(\Gamma_0) := \left[\widetilde{H}^{-s}(\Gamma_0) \right]',$$
$$\widetilde{H}^s(\Gamma_0) := \left[H^{-s}(\Gamma_0) \right].$$

Lemma 8.7 (分段平滑索伯列夫空间)**.** 对于某一封闭边界 $\Gamma = \partial\Omega$，假定它是分段平滑的

$$\Gamma = \bigcup_{i=1}^{J} \overline{\Gamma}_i, \quad \Gamma_i \cap \Gamma_j = \emptyset \, \forall i \neq j.$$

对应的索伯列夫空间 $H^s_{pw}(\Gamma)$ 及分段平滑方程的范数为：

1. $s > 0 \Rightarrow$

$$H^s_{pw}(\Gamma) := \left\{ \nu \in L^2(\Gamma) : \nu|_{\Gamma_i} \in H^s(\Gamma_i), \, i = 1, \ldots, J \right\},$$
$$\|\nu\|_{H^s_{pw}(\Gamma)} := \left\{ \sum_{i=1}^{J} \|\nu|_{\Gamma_i}\|^2_{H^s(\Gamma_i)} \right\}^{\frac{1}{2}}.$$

2. $s < 0 \Rightarrow$

$$H^s_{pw}(\Gamma) := \Pi_{j=1}^{j} \widetilde{H}^s_{\Gamma_j}, \tag{8.63}$$
$$\|w\|_{H^s_{pw(\Gamma)}} := \sum_{j=1}^{J} \|w|_{\Gamma_j}\|_{\widetilde{H}_{pw}(\Gamma_j)}. \tag{8.64}$$

Lemma 8.8. (8.63)的分段平滑索伯列夫空间 $H^s_{pw}(\Gamma)$, $s < 0$ 中，方程 $w \in H^{s<0}_{pw}(\Gamma)$ 满足

$$\|w\|_{H^P s(\Gamma)} \leq \|w\|_{H^P_{pw} s(\Gamma)}$$

证明. 由对偶空间的定义(8.62)、扩展三角不等式等有

$$\begin{aligned}
\|w\|_{H^s(\Gamma)} &= \sup_{0 \neq \nu \in H^{-s}(\Gamma)} \frac{\left|\langle w, \nu \rangle_\Gamma\right|}{\|\nu\|_{H^{-s}(\Gamma)}} \\
&\leq \sup_{0 \neq \nu \in H^{-s}(\Gamma)} \sum_{j=1}^{J} \frac{\left|\langle w, \nu \rangle_{\Gamma_j}\right|}{\|\nu\|_{H^{-s}(\Gamma_j)}} \\
&\leq \sup_{0 \neq \nu \in H^{-s}(\Gamma)} \sum_{j=1}^{J} \frac{\left|\langle w|_{\Gamma_j}, \nu|_{\Gamma_j} \rangle_{\Gamma_j}\right|}{\|\nu|_{\Gamma_j}\|_{H^{-s}(\Gamma_j)}} \\
&\leq \sum_{j=1}^{J} \sup_{0 \neq \nu \in H^{-s}(\Gamma_j)} \frac{\left|\langle w|_{\Gamma_j}, \nu|_{\Gamma_j} \rangle_{\Gamma_j}\right|}{\|\nu|_{\Gamma_j}\|_{H^{-s}(\Gamma_j)}} \\
&= \|w\|_{H^s_{pw}(\Gamma)}.
\end{aligned}$$

[14]例如一个拓扑空间 (如实值轴)X，连续方程 $f \in X$。此时 $\operatorname{supp} f$ 定义为这样一个闭集 C，满足 1)f 在 $X \backslash C$ 中为 0；2) 不存在 C 的真闭子集也满足这一条件，即 C 是所有这样的子集中最小的一个。

□

对于一个利普希茨域 $\Omega \subset \mathbb{R}^d$，其利普希茨边界可表示为 $\Gamma = \partial\Omega$。对于 $u(x) \in \Omega$，对应的内界迹 (interior boundary trace)为 $\gamma_0^{int} u \in \Gamma$，见(8.4)。则 $H^s(\Omega)$ 和 $H^s(\Gamma)$ 两个索伯列夫空间的关系，常用下述迹定理、逆迹定理予以描述。

Theorem 8.10 (索伯列夫空间的迹定理)**.** *设 $\Omega \subset \mathbb{R}^d$ 是一个 $C^{k-1,1}$ 域，$\frac{1}{2} < s \le k$，则内界迹*

$$\gamma_0^{int} : H^s(\Omega) \mapsto H^{s-\frac{1}{2}}(\Gamma)$$

是有界的，满足

$$\left\|\gamma_0^{int}\nu\right\|_{H^{s-\frac{1}{2}}(\Gamma)} \le C_T \left\|\nu\right\|_{H^s(\Omega)}, \quad \forall\, \nu \in H^s(\Omega).$$

Lemma 8.9. *若 $k = 1, s \in (\frac{1}{2}, 1]$，则我们可以利用 Theorem 8.10求得迹算子 $\gamma_0^{int} : H^s(\Omega) \mapsto H^{s-\frac{1}{2}}(\Gamma)$。*

若 $s \in (\frac{1}{2}, \frac{3}{2})$，该方法也适用 (McLean, 2000, Theorem 3.38)。

Theorem 8.11 (索伯列夫空间的逆迹定理)**.** *设 $\Omega \subset \mathbb{R}^d$ 是一个 $C^{k-1,1}$ 域，$\frac{1}{2} < s \le k$，则内界迹*

$$\gamma_0^{int} : H^s(\Omega) \mapsto H^{s-\frac{1}{2}}(\Gamma)$$

有一个连续的右逆算子

$$\varepsilon : H^{s-\frac{1}{2}}(\Gamma) \mapsto H^s(\Gamma),$$

满足如下关系

$$\begin{aligned}
&\gamma_0^{int}\varepsilon w = w, \quad \forall\, w \in H^{s-\frac{1}{2}}(\Gamma),\\
&\left\|\varepsilon w\right\|_{H^s(\Omega)} \le C_{IT} \left\|w\right\|_{H^{s-\frac{1}{2}}(\Gamma)}, \quad \forall\, w \in H^{s-\frac{1}{2}}(\Gamma).
\end{aligned}$$

Definition 8.39 (索伯列夫迹空间)**.** *从这个意义上来说，对于某个索伯列夫空间 $H^{s+\frac{1}{2}}(\Omega), s > 0$ 来说，它的迹空间可以表示为 $H^s(\Gamma)$，对应的范数*

$$\left\|\nu\right\|_{H^s(\Gamma),\gamma_0} := \inf_{V \in H^{s+\frac{1}{2}(\Omega)}, \gamma_0^{int} V = \nu} \left\|V\right\|_{H^{s+\frac{1}{2}}(\Omega)}.$$

需要指出的是，利普希茨域 $\Omega \subset \mathbb{R}^d$ 中，只有当 $|s| < 1$ 时，以下范式等价 $\|\nu\|_{H^s(\gamma),\gamma_0} = \|\nu\|_{H^s(\Gamma)}$。

8.3　变分法

弱形式边界值问题常常表现为带有算子方程 (operator equations) 的变分问题。对于变分问题，我们常将其表示为表面积位势和体积位势 (surface and volume potentials) 的偏微分方程，为了求解方程，就需要求得有界的积分算子方程的解，以建立完备的柯西数列解。本节介绍一些泛函分析的基本知识，进而探讨算子方程解的存在性和唯一性。

8.3.1　算子方程

假设一个希尔伯特空间 X，空间中的内积形式 $\langle .,. \rangle_X$，对应范数 $\|\cdot\|_X = \sqrt{\langle .,. \rangle_X}$。$X$ 的对偶空间 (dual space) 表示为 X'；X 和 X' 以 $\langle .,. \rangle$ 的形式呈对偶配对 (duality pairing)。则我们有

$$\left\|f\right\|_{X'} = \sup_{0 \ne \nu \in X} \frac{\left|\langle f, \nu \rangle\right|}{\left\|\nu\right\|_X}, \quad \forall\, f \in X'. \tag{8.65}$$

定义一个有界的自伴随 (self-adjoint) 线性算子 $A : X \mapsto X'$，满足

$$\|A\nu\|_{X'} \le C_2^A \|\nu\|_X, \quad \forall \nu \in X, \tag{8.66}$$

A 是自伴随的，是指

$$\langle Au, \nu\rangle = \langle u, A\nu\rangle, \quad \forall u, \nu \in X. \tag{8.67}$$

那么，边界值问题可以表示为，对于某个给定的 $f \in X'$，寻找算子方程的解 $u \in X$，使满足

$$Au = f. \tag{8.68}$$

算子方程(8.68)的解，通常难于直接求得。替代方案是建立一个变分问题，寻找变分问题的解 $u \in X$，使满足

$$\langle Au, \nu\rangle = \langle f, \nu\rangle, \quad \forall \nu \in X. \tag{8.69}$$

Theorem 8.12. 算子方程(8.68)的解 $u \in X$，和变分问题(8.69)的解 $u \in X$，二者等价。

证明. 证明过程分为两个部分。

1. 显然，算子方程(8.68)的解 $u \in X$，构成变分问题(8.69)的解。
2. 反过来，假定已求得变分问题(8.69)的解 $u \in X$，由(8.65)可得

$$\|Au - f\|_{X'} = \sup_{0\neq\nu\in X} \frac{|\langle Au - f, \nu\rangle|}{\|\nu\|_X},$$

代回(8.69)可得

$$\|Au - f\|_{X'} = 0 \Rightarrow Au = f,$$

即是说 $u \in X$ 同时也是算子方程(8.68)的解。

□

基于算子 A，可定义一个双线性映射 (bilinear form) $a(u, \nu)$

$$\begin{aligned} &a(u, \nu) := \langle Au, \nu\rangle, \quad \forall u, \nu \in X, \\ &A : X \mapsto X' \Rightarrow a(., .) : X \times X \mapsto \mathbb{R}, \end{aligned} \tag{8.70}$$

反之亦然，通过双线性映射 $a(u, \nu)$ 可以定义算子 $A : X \mapsto X'$，见 Lemma 8.10。

Lemma 8.10. 设一个有界的双线性映射 $a(., .) : X \times X \mapsto \mathbb{R}$，满足

$$|a(u, \nu)| \le C_2^A \|u\|_X \|\nu\|_X, \quad \forall u, \nu \in X.$$

对于其中任一 $u \in X$，都存在一个元素 $Au \in X'$，使得满足

$$\langle Au, \nu\rangle = a(u, \nu), \forall \nu \in X.$$

则我们有：算子 $A : X \mapsto X'$ 是一个有界的线性算子，满足

$$\|Au\|_{X'} \le C_2^A \|u\|_X, \forall u \in X.$$

证明. 对于任一给定的 $u \in X$，我们定义一个 X 中的有界双线性映射 $\langle f_u, \nu\rangle := a(u, \nu)$，即我们有 $f_u \in X'$。通过映射 $u \in X \mapsto f_u \in X'$，我们定义一个线性算子 $A : X \mapsto X'$，使得 $Au = f_u \in X'$，并且满足

$$\|Au\|_{X'} = \|f_u\|_{X'} = \sup_{0\neq\nu\in X} \frac{|\langle f_u, \nu\rangle|}{\|\nu\|_X} = \sup_{0\neq\nu\in X} \frac{|a(u, \nu)|}{\|\nu\|_X} \le C_2^A \|u\|_X.$$

□

因此我们可以把求解变分问题(8.69)的工作，转化为求解下述最小化问题的工作。

Lemma 8.11. *设线性算子 $A: X \mapsto X'$ 是 1) 自伴随的，即 $\langle Au, \nu\rangle = \langle u, A\nu\rangle$，和 2) 正半定的，即 $\langle Au, \nu\rangle \geq 0$，$\forall \nu \in X$。我们设一个泛函 F*

$$F(\nu) := \frac{1}{2}\langle A\nu, \nu\rangle - \langle f, \nu\rangle, \quad \forall \nu \in X. \tag{8.71}$$

那么变分问题(8.69)的解 $u \in X$，等价于如下最小化问题的解

$$F(u) = \min_{\nu \in X} F(\nu). \tag{8.72}$$

证明. 设 $u, \nu \in X$，任一 $t \in \mathbb{R}$。进而

$$\begin{aligned}
F(u+t\nu) &= \frac{1}{2}\langle A(u+t\nu), u+t\nu\rangle - \langle f, u+t\nu\rangle \\
&= \frac{1}{2}\left[\langle Au, u\rangle + \langle Au, t\nu\rangle + \langle A\nu, u\rangle + \langle At\nu, t\nu\rangle\right] - \left[\langle f, u\rangle + \langle f, t\nu\rangle\right] \\
&= \frac{1}{2}\langle Au, u\rangle + \frac{1}{2}t\langle Au, \nu\rangle + \frac{1}{2}t\langle Au, \nu\rangle + \frac{1}{2}t^2\langle A\nu, \nu\rangle - \langle f, u\rangle - t\langle f, \nu\rangle \\
&= \left[\frac{1}{2}\langle Au, u\rangle - \langle f, u\rangle\right] + t\left[\langle Au, \nu\rangle - \langle f, \nu\rangle\right] + \frac{1}{2}t^2\langle A\nu, \nu\rangle \\
&= F(u) + t\left[\langle Au, \nu\rangle - \langle f, \nu\rangle\right] + \frac{1}{2}t^2\langle A\nu, \nu\rangle.
\end{aligned}$$

1. 假设 $u \in X$ 是变分问题(8.69)的解，那么 $\langle Au, \nu\rangle = \langle f, \nu\rangle$，上式变为

$$F(u+t\nu) = F(u) + \underbrace{\frac{1}{2}t^2\langle A\nu, \nu\rangle}_{\geq 0},$$

$$\hookrightarrow F(u) \leq F(u+t\nu), \quad \forall \nu \in X, t \in \mathbb{R},$$

   由此可见 $u \in X$ 同时也是最小化问题(8.72)的解。

2. 现在假设 $u \in X$ 是最小化问题(8.72)的解。那么以下条件也成立

$$\frac{d}{dt}F(u+t\nu)\big|_{t=0} = 0, \quad \forall \nu \in X,$$

   由此可得

$$\langle Au, \nu\rangle = \langle f, \nu\rangle, \quad \forall \nu \in X,$$

   可见 $u \in X$ 同时也是变分问题(8.69)的解。

□

现在来证明变分问题(8.69)、最小化问题(8.72)的解 $u \in X$ 存在且唯一，可由里兹表现定理证明。

Theorem 8.13 (里兹表现定理)**.** *任一线性有界泛函 $f \in X'$ 均可表现为下述形式*

$$\langle f, \nu\rangle = \langle u, \nu\rangle_X,$$

其中 $u \in X$ 由 $f \in X'$ 所唯一确定 (uniquely determined)，并且满足

$$\|u\|_X = \|f\|_{X'}, \tag{8.73}$$

这称为里兹表现定理 (Riesz representation theorem)。

证明. 设某一给定的泛函 $f \in X'$，我们可以通过求解变分问题(8.69)找到解 $u \in X$

$$\langle Au, \nu \rangle = \langle f, \nu \rangle, \quad \forall \nu \in X,$$

并且根据 Lemma 8.11，这也等同于求解最小化问题(8.72)

$$F(u) = \min_{\nu \in X} F(\nu),$$

其中泛函 F 由(8.71)给出

$$F(\nu) := \frac{1}{2}\langle A\nu, \nu \rangle - \langle f, \nu \rangle, \quad \forall \nu \in X.$$

1. 证明 $u \in X$ 是最小化问题和变分问题的解。(8.71) $\Rightarrow$

$$\begin{aligned} F(\nu) &= \frac{1}{2}\langle \nu, \nu \rangle_X - \langle f, \nu \rangle \\ &\geq \frac{1}{2}\|\nu\|_X^2 - \|f\|_{X'}\|\nu\|_X \\ &= \frac{1}{2}\underbrace{\left[\|\nu\|_X - \|f\|_{X'}\right]^2}_{\geq 0} - \frac{1}{2}\|f\|_{X'}^2 \\ &\geq -\frac{1}{2}\|f\|_{X'}, \quad \forall \nu \in X, \end{aligned}$$

可见 $F(\nu), \forall \nu \in X$ 有下界 (infimum)，定义为 α

$$\alpha := \inf_{\nu \in X} F(\nu) \in \mathbb{R}.$$

设存在一个数列 $\{u_k\}_{k\in\mathbb{N}} \subset X$，随着 $k \to \infty$，满足 $F(u_k) \to \alpha$。由数列的性质可得

$$\|u_k - u_\ell\|_X^2 + \|u_k + u_\ell\|_X^2 = 2\left\{\|u_k\|_X^2 + \|u_\ell\|_X^2\right\},$$

进而我们有 $\|u_k - u_\ell\|_X^2 \geq 0$，以及

$$\begin{aligned} &\|u_k - u_\ell\|_X^2 = 2\|u_k\|_X^2 + 2\|u_\ell\|_X^2 - \|u_k + u_\ell\|_X^2 \\ &= 4\underbrace{\left\{\frac{1}{2}\|u_k\|_X^2 - \langle f, u_k \rangle\right\}}_{=F(u_k)} + 4\underbrace{\left\{\frac{1}{2}\|u_\ell\|_X^2 - \langle f, u_\ell \rangle\right\}}_{=F(u_\ell)} - 8\left[\frac{1}{2}\left\|\frac{1}{2}(u_k + u_\ell)\right\|_X^2 - \langle f, \frac{1}{2}(u_k + u_\ell)\rangle\right] \\ &= 4F(u_k) + 4F(u_\ell) - 8F\left(\frac{1}{2}(u_k + u_\ell)\right) \\ &\leq 4\alpha + 4\alpha - 8\alpha \to 0, \quad k, \ell \to \infty. \end{aligned}$$

由此可见，$\{u_k\}_{k\in\mathbb{N}}$ 是一个柯西数列 (Cauchy sequence)。此外由于 X 是一个希尔伯特空间，我们得到极限值

$$u = \lim_{k\to\infty} u_k \in X.$$

对应得泛函 $F(u)$ 的值，求解过程如下。由

$$\begin{aligned} |F(u_k) - F(u)| &\leq \left|F(u_k - u)\right| \\ &\leq \frac{1}{2}\left|\langle u_k, u_k \rangle_X - \langle u, u \rangle_X\right| + \left|\langle f, u_k - u \rangle\right| \\ &= \frac{1}{2}\left|\langle u_k, u_k - u \rangle_X + \langle u, u_k - u \rangle_X\right| + \left|\langle f, u_k - u \rangle\right| \\ &\leq \left\{\frac{1}{2}\|u_k\|_X + \frac{1}{2}\|u\|_X + \|f\|_{X'}\right\}\|u_k - u\|_X \end{aligned}$$

可得

$$F(u) = \lim_{k\to\infty} F(u_k) = \alpha.$$

可见 $u \in X$ 是最小化问题(8.71)和变分问题(8.69)的解。

2. 证明 $u \in X$ 是最小化问题和变分问题的唯一解。假定还存在另一个 $\tilde{u} \in X$，也是(8.71)和(8.69)的解，那么我们有

$$\langle \widetilde{u}, \nu\rangle_X = \langle f, \nu\rangle, \quad \forall \nu \in X.$$

将上式代回变分问题(8.69)⇒

$$\langle u - \widetilde{u}, \nu\rangle_X = 0, \quad \forall \nu \in X.$$

若设 $\nu = u - \tilde{u} \Rightarrow$

$$\left\|u - \widetilde{u}\right\|_X^2 = 0,$$

因此我们有 $u = \widetilde{u}$，即 $u \in X$ 是最小化问题和变分问题的唯一解。

3. 证明范式等价 $\|u\|_X = \|f\|_{X'}$。已知

$$\begin{aligned}&\|u\|_X^2 = \langle u, u\rangle_X = \langle f, u\rangle \le \|u\|_X \|f\|_{X'},\\ &\hookrightarrow \|u\|_X \le \|f\|_{X'},\end{aligned}$$

$$\|f\|_{X'} = \sup_{0\neq\nu\in X} \frac{|\langle f, \nu\rangle|}{\|\nu\|_X} = \sup_{0\neq\nu\in X} \frac{|\langle u, \nu\rangle_X|}{\|\nu\|_X} \le \|u\|_X,$$

则我们有 $\|u\|_X \le \|f\|_{X'} \& \|f\|_{X'} \le \|u\|_X \Rightarrow \|u\|_X = \|f\|_{X'}$.

□

Definition 8.40 (里兹映射)**.** *若里兹表现定理 (Theorem 8.13) 成立，那么我们将映射 $J : X' \mapsto X, u = Jf$ 称为里兹映射 (Riesz map)，满足如下变分问题*

$$\langle Jf, \nu\rangle_X = \langle f, \nu\rangle, \quad \forall \nu \in X, \tag{8.74}$$

并且其范数为

$$\|Jf\|_X = \|f\|_{X'}. \tag{8.75}$$

8.3.2　椭圆算子

里兹表现定理 (Theorem 8.13) 探讨了算子方程(8.68)及变分问题(8.69)的解 $u \in X$ 的存在性以及唯一性。除此以外，为了确保解得唯一存在，我们还需要对算子 A 和双线性形式 $a(.,.)$ 做出进一步设定。

Definition 8.41 (椭圆算子)**.** *一个算子 $A : X \mapsto X'$ 被称作 X-椭圆算子，如果它满足*

$$\langle A\nu, \nu\rangle \ge C_1^A \|\nu\|_X^2, \quad \forall \nu \in X, \tag{8.76}$$

其中 $0 \le C_1^A \in \mathbb{R}$。

Theorem 8.14 (拉克斯—密格拉蒙定理)**.** 设 $A : X \mapsto X'$ 是一个有界的 X-椭圆算子。对于任一 $f \in X'$，算子方程(8.68)都存在一个唯一解 $u \in X$，满足

$$\|u\|_X \leq \frac{1}{C_1^A} \|f\|_{X'}. \tag{8.77}$$

证明. 设存在一个里兹映射 (Riesz map) 算子 $J : X' \mapsto X$，满足(8.74)定义。那么算子方程(8.68)等价于下属定点方程

$$u = u - \varrho J (Au - f) = T_\varrho u + \varrho J f,$$

其中算子 $T_\varrho := I - \varrho JA : X \mapsto X$，参数 $0 < \varrho \in \mathbb{R}$。对应范数

$$\begin{aligned} \|T_\varrho u\|_X^2 &= \|(I = \varrho JA)u\|_X^2 \\ &= \|u\|_X^2 - 2\varrho \underbrace{\langle JAu, u\rangle_X}_{=:\mathcal{A}} + \varrho^2 \underbrace{\|JAu\|_X^2}_{=:\mathcal{B}} \\ &\leq \left[1 - 2\varrho C_1^A + \varrho^2 \left(C_2^A\right)^2\right] \|u\|_X^2, \end{aligned} \tag{8.78}$$

其中，由里兹映射算子 J 的性质(8.74)和 X-椭圆算子 A 的性质(8.76)我们有

$$\mathcal{A} := \langle JAu, u\rangle_X = \langle Au, u\rangle \geq C_1^A \|u\|_X^2,$$

由里兹映射算子 J 的范数(8.75)和 X-椭圆算子 A 的范数(8.66)我们有

$$\mathcal{B} := \|JAu\|_X = \|Au\|_{X'} \leq C_2^A \|u\|_X.$$

若设 $\varrho \in \left(0, \frac{2C_1^A}{(C_2^A)^2}\right)$，则算子 T_ϱ 是一个 X 中的收缩映射 (contraction mapping)，并且由收缩映射定理 (Banach's contraction mapping theorem, Palais (2007)) 可得，算子方程(8.68)的解 $x \in X$ 是唯一的。进而，对于唯一的解 $u \in X$，根据椭圆算子 A 的定义(8.76)和里兹表现定理我们有

$$\begin{aligned} C_1^A \|u\|_X^2 &\leq \langle Au, u\rangle \\ &= \langle f, u\rangle \\ &\leq \|f|_{X'} \|u\|_X. \end{aligned}$$

□

根据拉克斯—密格拉蒙定理 (Theorem 8.14)，我们可以定义一个逆算子 $A^{-1} : X' \mapsto X$，有

$$\|A^{-1} f\|_X \leq \frac{1}{C_1^A} \|f\|_{X'}, \quad \forall f \in X'.$$

Lemma 8.12 (X-椭圆算子 A 的逆算子 A^{-1} 也是一个椭圆算子)**.** 设 $A : X \mapsto X'$ 是一个有界(8.66)，自伴随(8.67)的 X-椭圆算子(8.77)。那么对于 $\forall \nu \in X$ 我们有

$$\langle A^{-1} f, f\rangle \geq \frac{1}{C_2^A} \|f\|_{X'}^2, \quad \forall f \in X'.$$

证明. 定义一个算子 $B := JA : X \mapsto X$，满足

$$\|B\nu\|_X = \|JA\nu\|_X = \|A\nu\|_{X'} \leq C_2^A \|\nu\|_X, \quad \forall \nu \in X.$$

由于 $\forall u, \nu \in X$，都有以下关系成立

$$\langle Bu, \nu\rangle_X = \langle JAu, \nu\rangle = \langle Au, \nu\rangle = \langle u, A\nu\rangle = \langle u, JA\nu\rangle_X = \langle u, B\nu\rangle_X,$$

可见 B 是一个自伴随的椭圆算子，满足

$$\langle B\nu,\nu\rangle_X = \langle A\nu,\nu\rangle \geq C_1^A \left\|\nu\right\|_X^2, \quad \forall\, \nu \in X.$$

因此可以定义一个可逆的自伴随算子 $B^{\frac{1}{2}}$，满足 $B = B^{\frac{1}{2}}\, B^{\frac{1}{2}}$，逆算子 $B^{-\frac{1}{2}} := \left(B^{\frac{1}{2}}\right)^{-1}$。对应的范数

$$\begin{aligned}
&\left\|B^{\frac{1}{2}}\nu\right\|_X^2 = \langle B\nu,\nu\rangle_X \leq \left\|B\nu\right\|_X \left\|\nu\right\|_X \leq C_2^A \left\|\nu\right\|_X^2,\\
\hookrightarrow &\left\|B^{\frac{1}{2}}\nu\right\|_X \leq \sqrt{C_2^A}\left\|\nu\right\|_X, \quad \forall\, \nu \in X.
\end{aligned}$$

因此，对于任一 $f \in X'$，我们有

$$\begin{aligned}
\|f\|_{X'} &= \sup_{0\neq\nu\in X} \frac{\left|\langle f,\nu\rangle\right|}{\left\|\nu\right\|_X}\\
&= \sup_{0\neq\nu\in X} \frac{\left|\langle Jf,\nu\rangle_X\right|}{\left\|\nu\right\|_X}\\
&= \sup_{0\neq\nu\in X} \frac{\left|\langle B^{-\frac{1}{2}}Jf, B^{\frac{1}{2}}\nu\rangle_X\right|}{\left\|\nu\right\|_X}\\
&\leq \sup_{0\neq\nu\in X} \frac{\left\|B^{-\frac{1}{2}}Jf\right\|_X \left\|B^{\frac{1}{2}}\nu\right\|_X}{\left\|\nu\right\|_X}\\
&\leq \sqrt{C_2^A}\left\|B^{-\frac{1}{2}}Jf\right\|_X,
\end{aligned}$$

进而

$$\left\|f\right\|_{X'}^2 \leq C_2^A \left\|B^{-\frac{1}{2}}Jf\right\|_X^2 = C_2^A\langle B^{-1}Jf, Jf\rangle_X = C_2^A\langle A^{-1}f, f\rangle,$$

其中我们使用到了如下关系

$$\begin{aligned}
\left\|B^{-\frac{1}{2}}Jf\right\|_X^2 &= \left\|\left(B^{\frac{1}{2}}\right)^{-1} Jf\right\|_X^2\\
&= \langle B^{-1}Jf, Jf\rangle_X\\
&= \langle A^{-1}f, f\rangle.
\end{aligned}$$

□

8.3.3　算子与稳定性条件

设 Π 是一个巴拿赫空间 (Banach space)，设 $B : X \mapsto \Pi'$ 是一个有界的线性算子，满足条件

$$\left\|B\nu\right\|_{\Pi'} \leq C_2^B \left\|\nu\right\|_X, \quad \forall \nu \in X. \tag{8.79}$$

算子 B 意味着如下双线性形式 $b(.,.) : X \times \Pi \mapsto \mathbb{R}$

$$b(\nu, q) := \langle B\nu, q\rangle, \quad (\nu, q) \in X \times \Pi.$$

B 的核或称零空间 (kernel, null space)定义为

$$\ker B := \{\nu \in X : B\nu = 0\}. \tag{8.80}$$

$\ker B$ 在希尔伯特空间 X 中的正交补 (orthogonal complement)为

$$(\ker B)^{\perp} := \{w \in X : \langle w, \nu\rangle_X = 0, \quad \forall\, \nu \in \ker B\} \subset X. \tag{8.81}$$

进而我们有 $\ker B$ 的正交空间

$$(\ker B)^0 := \{f \in X' : \langle f, \nu \rangle = 0, \quad \forall \nu \in \ker B\} \subset X'. \tag{8.82}$$

对于某一给定的 $g \in \Pi'$，我们想要求得以下算子方程的解 $u \in X$

$$Bu = g. \tag{8.83}$$

将 $B : X \mapsto \Pi'$ 的值域或称像 (range, image) ，定义为

$$\mathrm{Im}_X B := \{B\nu \in \Pi', \quad \forall \nu \in X\}.$$

则算子方程(8.83)要求是可解的 (solvability condition)，即要求 g 在 B 的值域中

$$g \in \mathrm{Im}_X B. \tag{8.84}$$

将 B 的伴随算子 (adjoint operator)定义为 $B' : X \mapsto \Pi'$，满足

$$\langle \nu, B'q \rangle := \langle B\nu, q \rangle, \quad \forall (\nu, q) \in X \times \Pi.$$

由 B 的性质 (8.80)，(8.81)，(8.82)可得伴随算子 B' 的性质

$$\ker B' := \{q \in \Pi : \langle B\nu, q \rangle = 0, \forall \nu \in X\}, \tag{8.85}$$

$$(\ker B')^{\perp} := \{p \in \Pi : \langle p, q \rangle_{\Pi} = 0, \forall q \in \ker B'\}, \tag{8.86}$$

$$(\ker B')^0 := \{g \in \Pi' : \langle g, q \rangle = 0, \quad \forall q \in \ker B'\}. \tag{8.87}$$

$\mathrm{Im}_X B$ 的性质，由闭值域定理 (closed range theorem)给出

Theorem 8.15 (闭值域定理)**.** *设 X 和 Π 是巴拿赫空间，有界线性算子 $B : X \mapsto \Pi'$。则以下属性等价*

- *$\mathrm{Im}_X B$ 是 Π' 中的闭集，*
- *$\mathrm{Im}_{\Pi} B'$ 是 X' 中的闭集，*
- *$\mathrm{Im}_X B = (\ker B')^0$，*
- *$\mathrm{Im}_{\pi} B' = (\ker B)^0$。*

证明. 略。可参考 (Muscat, 2014, Proposition 11.30)。 □

可求解性条件(8.84)⇒

$$\langle g, q \rangle = 0, \forall q \in \ker B' \subset \pi. \tag{8.88}$$

若可求解性条件(8.84)或(8.88)得到满足，则算子方程(8.83)存在至少一个解 $u \in X$。但解并不唯一：我们可以加入任一 $u_0 \in \ker B$，使得 $u + u_0$ 也是 $B(u + u_0) = g$ 的解。因此我们需要引入额外的假设条件 $u \in (\ker B)^{\perp}$，以确保解的唯一性。

Theorem 8.16 (算子方程解的唯一存在性)**.** *设希尔伯特空间 X 和 Π。有界的线性算子 $B : X \mapsto \Pi'$。进而我们有稳定性条件*

$$C_S \|\nu\|_X \leq \sup_{0 \neq q \in \Pi} \frac{\langle B\nu, q \rangle}{\|q\|_{\Pi}}, \quad \forall \nu \in (\ker B)^{\perp}, \tag{8.89}$$

那么对于一个给定的 $g \in \mathrm{Im}_X B$，算子方程 $Bu = g$ 存在一个唯一的解 $u \in (\ker B)^{\perp}$，满足

$$\|u\|_X \leq \frac{1}{C_S} \|g\|_{\Pi'}.$$

证明. 已知根据假设条件 $g \in \mathrm{Im}_X B$，算子方程 $Bu = g$ 存在唯一的一个解 $u \in (\ker B)^{\perp}$，满足

$$\langle Bu, q\rangle = \langle g, q\rangle, \quad \forall\, q \in \Pi.$$

现在设存在第二个解 $\bar{u} \in (\ker B)^{\perp}$，满足

$$\langle B\bar{u}, q\rangle = \langle g, q\rangle, \quad \forall\, q \in \Pi,$$

则我们有

$$\langle B(u-\bar{u}), q\rangle = 0, \quad \forall q \in \Pi.$$

显然 $u - \bar{u} \in (\ker B)^{\perp}$ 也满足稳定性条件(8.89)

$$\begin{aligned} 0 \le C_S \left\|u-\bar{u}\right\|_X &\le \sup_{0\neq q\in\Pi} \frac{\langle B\left(u-\bar{u}\right), q\rangle}{\left\|q\right\|_{\Pi}}, \quad \forall\, \nu \in (\ker B)^{\perp} = 0, \\ \hookrightarrow u &= \bar{u}. \end{aligned}$$

把唯一解 u 代回(8.89)我们有

$$\begin{aligned} C_S \left\|u\right\|_X &\le \sup_{0\neq q\in\Pi} \frac{\langle Bu, q\rangle}{\left\|q\right\|_{\Pi}} \\ &= \sup_{0\neq q\in\Pi} \frac{\langle g, q\rangle}{\left\|q\right\|_{\Pi}} \\ &\le \left\|g\right\|_{\Pi'}. \end{aligned}$$

□

8.3.4　含有限制条件的算子方程

经验研究中我们常常需要求得带有约束条件 $Bu = g$ 的算子方程 $Au = f$ 的解。常见的求解思路分为四步。

1. 关于限定条件 $Bu = g$，假定可求解条件(8.84)成立

$$g \in \mathrm{Im}_X B := \{B\nu \in \Pi', \quad \forall\, \nu \in X\}.$$

对于给定的 $g \in \Pi'$，定义流形 V_g

$$V_g := \{\nu \in X : B\nu = g\}.$$

此外我们定义零空间 V_0

$$V_0 = \ker B := \{\nu \in X : B\nu = 0\}.$$

2. 关于算子方程 $Au = f$，同样假定可求解条件(8.84)成立

$$f \in \mathrm{Im}_{V_g} A := \{A\nu \in X', \quad \forall\, \nu \in V_g\}.$$

对于给定的 $f \in X'$，构建变分问题

$$\langle Au, \nu\rangle = \langle f, \nu\rangle, \quad \forall\, \nu \in V_0, \tag{8.90}$$

求解该问题，得到解 $u \in V_g$。

3. 所求得解 $u \in V_g$ 的唯一性，见 Theorem 8.17。

4. 对于唯一存在解 $u \in V_g$，可以将 u 的范数，和给定的 $f \in X', g \in \Pi'$ 的范数联系起来。

   假定对于给定的 $g \in \operatorname{Im}_X B$，$\exists\, u_g \in V_g$，我们有

$$\|u_g\|_X \leq C_B \|g\|_{\Pi'}, \quad C_B > 0 \in \mathbb{R}. \tag{8.91}$$

   ，则范数之间的关联见 Corollary 8.4。

Theorem 8.17 (带约束算子方程解的唯一存在性)**.** *设一个有界线性 V_0-椭圆算子 $A : X \mapsto X'$*

$$\langle A\nu, \nu\rangle \geq C_1^A \left\|\nu\right\|_X^2, \quad \forall /, \nu \in V_0 \coloneqq \ker B,$$

其中 $B : X \mapsto \Pi'$。并且有给定的 f, g

$$\begin{aligned} f &\in \operatorname{Im}_{V_g} A \coloneqq \{A\nu \in X', \quad \forall\, \nu \in V_g\}, \\ g &\in \operatorname{Im}_X B \coloneqq \{B\nu \in \Pi', \quad \forall\, \nu \in X\}. \end{aligned} \tag{8.92}$$

那么作为带有约束条件 $Bu\ g$ 的算子方程 $Au = f$，有且只有一个解 $u \in X$。

证明. 由已知条件 $g \in \operatorname{Im}_X B$ 可得，约束条件存在至少一个解 $u_g \in X$ 满足 $Bu_g = g$。除此而外，我们还需要求得 $u_0 \coloneqq u - u_g \in V_0$，作为以下算子方程的解

$$Au_0 = f - Au_g,$$

算子方程等价于以下变分问题

$$\langle Au_0, \nu\rangle = \langle f - Au_g, \nu\rangle, \quad \forall \nu \in V_0.$$

[存在性] 由已知条件 $f \in \operatorname{Im}_{V_g} A$ 可得，$f - Au_g \in \operatorname{Im}_{V_0} A$，那么方程 $Au_0 = f - Au_g$ 至少存在一个解 $u_0 \in V_0$。

[唯一性] 现在来证明 $u_0 \in V_0$ 是唯一的。设 $\bar{u}_0 \in V_0$ 是算子方程的另一个解，满足 $A\bar{u}_0 = f - Au_g$。由已知条件 A 的 V_0-椭圆特性可得

$$\begin{aligned} &0 \leq C_1^A \left\|u_0 - \bar{u}_0\right\|_X^2 \leq \langle A\left(u_0 - \bar{u}_0\right), u_0 - \bar{u}_0\rangle = \langle Au_0 - A\bar{u}_0, u_0 - \bar{u}_0\rangle = 0, \\ &\hookrightarrow u_0 \in X = \bar{u}_0 \in X. \end{aligned}$$

$u_g \in V_g$ 可能不是唯一的解，但 $u \in X = u_0 + u_g$ 却是变分问题的唯一最终解，并不受 (可能是多重的) $u_g \in V_g$ 的影响。这是由于，对于满足约束条件方程 $B\hat{u}_g = g$ 的某一个解 $\hat{u}_g \in X$ 而言，这意味着存在唯一一个 $\hat{u}_0 \in V_0$，构成算子方程 $A(\hat{u}_0 + \hat{u}_g) = f$ 的解，进而

$$\begin{aligned} &B(u_g - \hat{u}_g) = Bu_g - B\hat{u}_g = g - g = 0 \in \Pi', \\ &\hookrightarrow u_g - \hat{u}_g \in \ker B = V_0. \end{aligned}$$

由于

$$\begin{aligned} A(u_0 + u_g) &= f, \\ A(\hat{u}_0 + \hat{u}_g) &= f, \end{aligned}$$

我们因而有

$$A(u_0 + u_g - \hat{u}_0 - \hat{u}_g) = 0.$$

显然，$u_0 - \hat{u}_0 + (u_g - \hat{u_g}) \in V_0$，由 A 的 V_0-椭圆特性我们有

$$u_0 - \hat{u}_0 + (u_g - \hat{u}_g) = 0,$$

因此我们可得解的唯一性

$$u = u_0 + u_g = \hat{u}_0 + \hat{u}_g.$$

□

Corollary 8.4. *已知 $u \in V_g$ 是有限制条件的算子方程唯一解 (Theorem 8.17)，并且满足假定(8.91)。那么 $\|u\|_X$ 和给定的 $f \in X', g \in \Pi'$ 的范数之间关系为*

$$\|u\|_X \le \frac{1}{C_1^A}\|f\|_{X'} + \left(1 + \frac{C_2^A}{C_1^A}\right) C_B \|g\|_{\Pi'}.$$

证明. 由 Theorem 8.17得，算子方程 $Au = f$ 的解表现为 $u = u_g + u_0$ 的形式，其中 $u_0 \in V_0$ 是以下变分问题的唯一解

$$\langle Au_0, \nu\rangle = \langle f - Au_g, \nu\rangle, \quad \forall \nu \in V_0.$$

由算子 A 的 V_0-椭圆特性可得

$$\begin{aligned} C_1^A \|u_0\|_X^2 &\le \langle Au_0, u_0\rangle = \langle f - Au_0, u_0\rangle \le \|f - Au_g\|_{X'} \|u_0\|_X, \\ \hookrightarrow \|u_0\|_X &\le \frac{1}{C_1^A}\left[\|f\|_{X'} + C_2^A \|u_g\|_X\right], \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \|u\|_X &= \|u_0 + u_g\|_X \\ &\le \|u_0\|_X + \|u_g\|_X \\ &\le \frac{1}{C_1^A}\|f\|_{X'} + \left(1 + \frac{C_2^A}{C_1^A}\right) \|u_g\|_X \\ &\le \frac{1}{C_1^A}\|f\|_{X'} + \left(1 + \frac{C_2^A}{C_1^A}\right) C_B \|g\|_{\Pi'}. \end{aligned}$$

□

8.3.5　混合算子方程 (鞍点变分问题)

第8.3.4节讨论了如何构建带有限制条件的算子方程来求解变分问题。除此而外的另一种方法是引入拉格朗日乘子 $p \in \Pi$，构建扩展变分问题，对于 $\forall (\nu, q) \in X \times \Pi$，求解 $(u, p) \in X$，使其满足

$$\langle Au, \nu\rangle + \langle B\nu, p\rangle = \langle f, \nu\rangle, \tag{8.93a}$$

$$\langle Bu, q\rangle = \langle g, q\rangle, \tag{8.93b}$$

其中 $u \in V_g$ 是 $Au = f$ 的解。

与上节相同，(8.93b)同样可以用于描述限制条件 $Bu = g$。但不同的是(8.93a)可理解为另一个变分法问题：将 $\nu \in V_0$ 作为检验方程，求解 $u_0 \in V_0$。显然，这种混合算子求解变分问题的研究思路，可行前提之一是确保存在拉格朗日乘子 $p \in \Pi$，使得 $\forall \nu \in X$ 都满足等式(8.93a)。具体来说就是，对于一组 $(\nu, q) \in X \times \Pi$，定义一个拉格朗日泛函

$$\mathcal{L}(\nu, q) := \frac{1}{2}\langle A\nu, \nu\rangle - \langle f, \nu\rangle + \langle B\nu, q\rangle - \langle g, q\rangle,$$

应当使得 Theorem (8.18)成立。

Theorem 8.18. 设 $A: X \mapsto X'$ 为一个有界线性算子，$\forall \nu \in X$ 都具有自伴随 $\langle A\nu, nu\rangle = \langle \nu, A\nu\rangle$、正半定 $\langle A\nu, \nu\rangle \geq 0$ 的特性。设另一个有界线性算子 $B: X \mapsto \Pi'$。

当且仅当

$$\mathcal{L}(u,q) \leq \mathcal{L}(u,p) \leq \mathcal{L}(\nu,p) \quad \forall (\nu,q) \in X \times \Pi \tag{8.94}$$

时，(u,p) 成为变分问题(8.93a)-(8.93b)的一个解。

证明. 假设有一组解 $(u,p) \in X \times \Pi$。

1. 证(8.94)的后半部分。

$$\begin{aligned}
&\mathcal{L}(\nu,p) - \mathcal{L}(u,p) \\
&= \frac{1}{2}\langle A\nu,\nu\rangle - \langle f,\nu\rangle + \langle B\nu,p\rangle - \langle g,p\rangle - \frac{1}{2}\langle Au,u\rangle + \langle f,u\rangle - \langle Bu,p\rangle + \langle g,p\rangle \\
&= \frac{1}{2}\underbrace{\langle A(u-\nu),(u-\nu)\rangle}_{\text{正半定},\geq 0} + \underbrace{\langle Au,(u-\nu)\rangle + \langle B(u-\nu),p\rangle - \langle f,u-\nu\rangle}_{(8.93a),=0} \\
&\geq 0,
\end{aligned}$$

$$\therefore \mathcal{L}(u,p) < \mathcal{L}(\nu,p), \quad \forall \nu \in X.$$

2. 证(8.94)的前半部分。

$$\begin{aligned}
&\mathcal{L}(u,p) - \mathcal{L}(u,q) \\
&= \frac{1}{2}\langle Au,u\rangle - \langle f,u\rangle + \langle Bu,p\rangle - \langle g,p\rangle - \frac{1}{2}\langle Au,u\rangle + \langle f,u\rangle - \langle Bu,q\rangle + \langle g,p\rangle \\
&= \underbrace{\langle Bu,p-q\rangle - \langle g,p-q\rangle}_{(8.93b),=0} \\
&= 0,
\end{aligned}$$

$$\therefore \mathcal{L}(u,q) \leq \mathcal{L}(\nu,p), \quad \forall q \in \Pi.$$

3. 假设已知某个 $p \in \Pi$ 是方程的解。构建如下最小化问题

$$\mathcal{L}(u,p) \leq \mathcal{L}(\nu,p), \quad \forall \nu \in X, \tag{8.95}$$

求解 $u \in X$ 满足式(8.93a)。

设某个 $u \in X$ 是最小化问题的解，我们有 $\forall w \in X$ 满足以下两式

$$\frac{d}{dt}\mathcal{L}(u+tw,p)\,|_{t=0} = 0. \tag{8.96}$$

以及

$$\begin{aligned}
\mathcal{L}(u+tw,p) = &\underbrace{\frac{1}{2}\langle Au,u\rangle - \langle f,u\rangle + \langle Bu,p\rangle - \langle g,p\rangle}_{=\mathcal{L}(u,p)} \\
&+ \frac{1}{2}t^2\langle Aw,w\rangle + t\left[\langle Au,w\rangle + \langle Bw,p\rangle - \langle f,w\rangle\right],
\end{aligned} \tag{8.97}$$

进而(8.96)可得，(8.97)$\Rightarrow$

$$\langle Au,w\rangle + \langle Bw,p\rangle - \langle f,w\rangle = 0, \forall w \in X,$$

满足式(8.93a)。

4. 对于任一 $q \in \Pi$，证明(8.93b)。

   (a) 定义 $\tilde{q} := p + q$，进而

$$
\begin{aligned}
0 &\le \mathcal{L}(u,p) - \mathcal{L}(u,p+g) \\
&= \frac{1}{2}\langle Au,u\rangle - \langle f,u\rangle + \langle Bu,p\rangle - \langle g,p\rangle \\
&\quad - \frac{1}{2}\langle Au,u\rangle + \langle f,u\rangle - \langle Bu,p+q\rangle + \langle g,p+q\rangle \\
&= -\langle Bu,q\rangle + \langle g,q\rangle.
\end{aligned}
$$

   (b) 定义 $\tilde{q} := p - q$，进而

$$
0 \le \mathcal{L}(u,p) - \mathcal{L}(u,p-g) = \langle Bu,q\rangle - \langle g,q\rangle.
$$

   (c)

$$
\therefore \langle Bu,q\rangle = \langle g,q\rangle, \quad \forall\, q \in \Pi. \tag{8.98}
$$

□

由此可见，扩展变分问题(8.93a)-(8.93b)中，组合 $(u,p) \in X \times \Pi$ 是一个拉格朗日泛函 $\mathcal{L}(.,.)$ 的鞍点。从这意义上说，扩展变分问题也常称为鞍点变分问题。下面来探讨 (u,p) 解的唯一性。

Theorem 8.19 (混合算子方程 (鞍点变分问题) 的解)**.** *假设巴拿赫空间 X,Π，有界算子 $A : X \mapsto X'$, $B : X \mapsto \Pi'$。设 A 满足 V_0-椭圆特性*

$$
\langle A\nu,\nu\rangle \ge C_1^A \left\|\nu\right\|_X^2, \quad \forall \nu \in V_0 = \ker B,
$$

设稳定性条件

$$
C_s \left\|q\right\|_\Pi \le \sup_{0\ne\nu\in X} \frac{\langle B\nu,q\rangle}{\left\|\nu\right\|_X}, \quad \forall q \in \Pi. \tag{8.99}
$$

那么对于 $g \in \mathrm{Im}_X B, f \in \mathrm{Im}_{V_g} A$，扩展变分问题(8.93a)-(8.93b)都存在唯一的解 $(u,p) \in X \times \Pi$，满足如下关系

$$
\left\|u\right\|_X \le \frac{1}{C_1^A}\left\|f\right\|_{X'} + \left(1 + \frac{C_2^A}{C_1^A}\right) C_B \left\|g\right\|_{\Pi'}, \tag{8.100}
$$

$$
\left\|p\right\|_\Pi \le \frac{1}{C_S}\left(1 + \frac{C_2^A}{C_1^A}\right)\left\{\left\|f\right\|_{X'} + C_B C_2^A \left\|g\right\|_{\Pi'}\right\}. \tag{8.101}
$$

证明.　1. 证明对于变分问题

$$
\begin{aligned}
\langle Au,\nu\rangle &= \langle f,\nu\rangle, \quad \forall\, \nu \in V_0, \\
\langle Bu,q\rangle &= \langle g,q\rangle, \quad \forall q \in Pi
\end{aligned}
$$

   存在唯一的解 $u \in X$。

   唯一解的存在性由 Theorem 8.17证得。唯一解的范数不等式由 Corollary 8.4给出，对应式(8.100)。

2. 证明对于变分问题

$$
\langle B\nu,p\rangle = \langle f - Au,\nu\rangle, \quad \forall\, \nu \in X
$$

   存在唯一的解 $p \in \Pi$。

(a) 解的存在性。我们有 $f - Au \in (\ker B)^0$，进而根据闭值域定理 (Theorem 8.15) 我们有 $f - Au \in \mathrm{Im}_\Pi(B')$，进而变分问题的解是 $p \in \Pi$。

(b) 解的唯一性。假定变分问题有两个解 $p, \hat{p} \in \Pi$，满足

$$\begin{aligned}\langle B\nu, p\rangle &= \langle f - Au, \nu\rangle, \quad \forall \nu \in X,\\ \langle B\nu, \hat{p}\rangle &= \langle f - Au, \nu\rangle, \quad \forall \nu \in X.\end{aligned}$$

两式相减 $\Rightarrow$

$$\langle B\nu, p - \hat{p}\rangle = 0, \quad \forall \nu \in X.$$

代入稳定性条件(8.99)有

$$\begin{aligned}&0 \le C_s \left\|p - \hat{p}\right\|_\Pi \le \sup_{0\neq\nu\in X} \frac{\langle B\nu, p - \hat{p}\rangle}{\left\|\nu\right\|_X} \le 0, \quad \forall q \in \Pi,\\ &\hookrightarrow p = \hat{p} \in \Pi.\end{aligned}$$

(c) 计算唯一解 $p \in \Pi$ 的范数。再次代入稳定性条件(8.99)有

$$C_s \left\|p\right\|_\Pi \le \sup_{0\neq\nu\in X} \frac{\langle B\nu, p\rangle}{\left\|\nu\right\|_X} = \sup_{0\neq\nu\in X} \frac{\langle f - Au, \nu\rangle}{\left\|\nu\right\|_X} \le \left\|f\right\|_{X'} + C_2^A \left\|u\right\|_X,$$

代入(8.100)替换 $\left\|u\right\|_X$，我们得(8.100)。

□

设 $A : X \mapsto X'$ 是个 V_0-椭圆算子时，Theorem 8.19成立。此外，若假设 A 是个 X-椭圆算子，即

$$\langle A\nu, \nu\rangle \ge C_1^A \left\|\nu\right\|_X^2, \quad \forall \nu \in X, \tag{8.102}$$

Theorem 8.19依然成立。

Theorem 8.19探讨了求扩展变分问题的解 $(u, p) \in X \times \Pi$。事实上求解过程可以进一步简化，u 是一个关于 p 的方程。已知

$$\begin{aligned}&\langle Au, \nu\rangle + \langle B\nu, p\rangle = \langle f, \nu\rangle,\\ &\hookrightarrow \langle B\nu, p\rangle = \langle \nu, Bp\rangle = \langle B'p, \nu\rangle,\\ &\hookrightarrow \langle Au, \nu\rangle + \langle B'p, \nu\rangle = \langle f, \nu\rangle,\\ &\hookrightarrow Au + B'p = f,\\ &\hookrightarrow u = A^{-1}\left(f - B'p\right)\\ &\hookrightarrow \langle Bu, q\rangle = \langle g, q\rangle,\\ &\hookrightarrow \langle BA^{-1}\left(f - B'p\right), q\rangle = \langle g, q\rangle,\\ &\hookrightarrow \langle BA^{-1}f - g, q\rangle = \langle BA^{-1}B'p, q\rangle,\end{aligned}$$

即对于任一 $p \in \Pi$，都存在唯一的一个解 $u = A^{-1}\left(f - B'p\right) \in X$。这样一来，原本寻找 $(u, p) \in X \times \Pi$ 的变分问题，就变成了一个新的 (椭圆) 变分问题：寻找解 $p \in \Pi$ 使满足

$$\langle BA^{-1}f - g, q\rangle = \langle BA^{-1}B'p, q\rangle. \tag{8.103}$$

为了探讨 (椭圆) 变分问题(8.103)的解 u 的唯一性，首先我们检验它是否符合拉克斯—密格拉蒙定理 (Theorem 8.14) 的前提假；如果是，则应用该定理。检验过程见 Lemma 8.13.

Lemma 8.13. 设 *Theorem 8.19*的假设条件均得到满足。

那么算子 $S := BA^{-1}B'$ 有界，并且由稳定条件(8.99)可得 S 是 Π-椭圆的，满足

$$\langle Sq, q\rangle \geq C_1^S \|q\|_\Pi^2, \quad \forall q \in \pi. \tag{8.104}$$

证明. 已知即对于任一 $p \in \Pi$，对于如下变分问题

$$\langle Au, \nu\rangle = \langle B\nu, q\rangle \quad \forall \nu \in X,$$

都存在唯一的一个解 $u = A^{-1}(f - B'p) \in X$.

1. S 的有界性。已知 $A : X \mapsto X'$ 的 X-椭圆特性(8.102)，那么根据 Theorem 8.19可证得存在唯一解 $u \in X$ 满足

$$\begin{aligned} \|u\|_X &= \|A^{-1}B'q\|_X \\ &\leq \frac{1}{C_1^A}\|B'q\|_{X'} \\ &\leq \frac{C_2^B}{C_1^A}\|q\|_\Pi, \quad \forall q \in \Pi. \end{aligned} \tag{8.105}$$

   由(8.105)可得

$$\begin{aligned} \|Sq\|_{\Pi'} &= \|BA^{-1}B'q\|_{\Pi'} \\ &= \|Bu\|_{\Pi'} \\ &\leq C_2^B\|u\|_X \\ &\leq \frac{(C_2^B)^2}{C_1^A}\|q\|_\Pi, \quad \forall q \in \Pi, \end{aligned}$$

   即 $S : \Pi \mapsto \Pi'$ 有界。

2. S 的椭圆性。

$$\begin{aligned} \langle Sq, q\rangle &= \langle BA^{-1}B'q, q\rangle \\ &= \langle Bu, q\rangle \\ &= \langle Au, u\rangle \\ &\geq C_1^A\|u\|_X^2, \end{aligned} \tag{8.106}$$

   根据稳定条件(8.99)，

$$\begin{aligned} C_S\|q\|_\Pi &\leq \sup_{0\neq\nu\in X}\frac{\langle B\nu, q\rangle}{\|\nu\|_X} \\ &= \sup_{0\neq\nu\in X}\frac{\langle Au, \nu\rangle}{\|\nu\|_X} \\ &\leq C_2^A\|u\|_X, \end{aligned}$$

   代回(8.106)最后得

$$\langle Sq, q\rangle \geq C_1^A\|u\|_X^2 \geq \frac{(C_2^B)^2}{C_1^A}\|q\|_\Pi^2,$$

   由此证得(8.105)，$C_1^S := \frac{(C_2^B)^2}{C_1^A}$。

□

进而，我们可以使用拉克斯—密格拉蒙定理 (Theorem 8.14) 求得椭圆问题(8.103)的唯一解 $p \in \Pi$。

在此基础上，回到混合算子方程 (鞍点变分问题) (8.93a) -(8.93b)上来：

Theorem 8.20 (混合算子方程 (鞍点变分问题) 的解 (续))**.** *设巴拿赫空间 X, Π。有界算子 $A: X \mapsto X', B: X \mapsto \Pi'$。假设 A 是 X-椭圆的，满足稳定条件(8.99)。对于给定的 $f \in X', g \in \Pi'$，混合算子方程 (鞍点变分问题) (8.93a) -(8.93b)存在唯一的解 $(u, p) \in X \times \Pi$，满足*

$$\|p\|_{\Pi} \leq \frac{1}{C_1^S}\|BA^{-1}f - g\|_{\Pi'} \leq \frac{1}{C_1^S}\left[\frac{C_2^B}{C_1^A}\|f\|_{X'} + \|g\|_{\Pi'}\right], \tag{8.107}$$

$$\|u\|_X \leq \frac{1}{C_1^A}\left[1 + \frac{\left(C_2^B\right)^2}{C_1^A C_1^S}\right]\|f\|_{X'} + \frac{C_2^B}{C_1^A C_1^S}\|g\|_{\Pi'}. \tag{8.108}$$

证明. 1. 根据 Theorem 8.19可证得(8.107)。

2. 根据拉克斯—密格拉蒙定理 (Theorem 8.14)，可以证明混合算子方程 (鞍点变分问题)(8.93a) -(8.93b)存在唯一解，满足

$$\langle Au, \nu\rangle = \langle f - B'p, \nu\rangle, \quad \forall \nu \in X.$$

由算子 A 的 X-椭圆特性可得

$$C_1^A\|u\|_X^2 \leq \langle Au, u\rangle = \langle f - B'p, u\rangle \leq \|f - B'p\|_{X'}\|u\|_X,$$

由此我们有

$$\begin{aligned}\|u\|_X &\leq \frac{1}{C_1^A}\|f - B'p\|_{X'} \\ &\leq \frac{1}{C_1^A}\|f\|_{X'} + \frac{C_2^B}{C_1^A}\|p\|_{\Pi}.\end{aligned}$$

将(8.107)代入上式，可得(8.108)。

□

8.3.6 强制算子方程

前面介绍有界线性算子 $A: X \mapsto X'$，假设它具有 X-椭圆的特性，如(8.76)。这一假设过强。在多数情况下，我们用强制算子 (coercive operator) 予以替代。

Definition 8.42 (强制算子)**.** *如果存在一个紧凑算子 $C: X \mapsto X'$，使得和 $A: X \mapsto X'$ 一道满足 Gårding 不等式 (Gårding inequality)*

$$\langle (A + C)\nu, \nu\rangle \geq C_1^A\|\nu\|_X^2, \quad \forall \nu \in X. \tag{8.109}$$

那么我们称 A 是一个强制算子 (coercive operator)。

Gårding 不等式的详细介绍，可参考 (Jovanović and Süli, 2014, Theorem 2.4)。证明可见如 (Agranovich, 2015, Theorem 8.1.1)，(Renardy and Rogers, 2004, Therorem 9.17)。

紧凑算子的定义。对于 $C: X \mapsto Y$，若 X 中单位球体 (unit sphere) 的像 (image) 在 Y 中相对紧凑，则我们称 C 为紧凑算子 (compact operator)。一个值得关注的特性是，紧凑算子和有界线性算子的乘也是紧凑的。

根据 Riesz-Schauder 定理 (Riesz-Schauder theorem, (Yosida, 1978, Sec. X.5), (Muscat, 2014, Theorem 14.18))，我们有弗雷德霍姆二择一定理 (Fredholm alternative theorem)如下。

Theorem 8.21 (弗雷德霍姆二择一定理)**.** *设 $K : X \mapsto X$ 是一个紧凑算子。那么以下两种情况之一会出现：*

- *齐次方程 (homogeneous equation)*
$$(I - K)\, u = 0$$
有一个不平凡解 (nontrivial solution) $u \in X$，或

- *非其次方程*
$$(I - K)\, u = g$$
对于每一个给定的 $g \in X$，都有唯一对应的解 $u \in X$，满足关系
$$\|u\|_X \le c\|g\|_X.$$

证明. 略。可见 (Agranovich, 2015, Sec. 18.1)。 □

基于弗雷德霍姆二择一定理 (Theorem 8.21)，我们可以探讨当有界线性算子 A 是强制的时，算子方程 $Au = f$ 的解。

Theorem 8.22 (强制算子方程的解)**.** *假设一个有界线性算子 $A : X \mapsto X'$，具有强制性 (Definition 8.42)、内射性 (injective, 即 $Au = 0 \Rightarrow u = 0$)。那么强制算子方程 $Au = f$ 存在唯一的解 $u \in X$，并且满足条件*
$$\|u\|_X \le c\,\|f\|_{X'}.$$

证明. 定义一个线性算子 $D := A + C : X \mapsto X'$，由(8.109)可得，线性有界算子 D 也是 X-椭圆的。

通过拉克斯—密格拉蒙定理 (Theorem 8.14拉克斯—密格拉蒙定理) 可得，逆算子 $D^{-1} : X' \mapsto X$。这样，我们可以将原强制算子方程 $Au = f$ 转换为新的强制算子方程

$$\begin{aligned} &Bu = D^{-1}Au = D^{-1}f, \text{其中线性有界算子} \\ &B := D^{-1}A = D^{-1}(D - C) = I - D^{-1}C : X \mapsto X. \end{aligned} \tag{8.110}$$

由假设条件 $C : X \mapsto X'$ 是紧凑算子，和 $D : X \mapsto X'$、进而 $D^{-1} : X' \mapsto X$ 是线性有界算子，可得 $D^{-1}C : X \mapsto X$ 是紧凑算子。进而，可以根据弗雷德霍姆二择一定理 (Theorem 8.21) 证得强制算子方程(8.110)解的唯一存在性。

由假设条件 $A : X \mapsto X'$ 的内射性，可得齐次方程 $D^{-1}Au = 0$ 的所有解 $u \in X$ 都是平凡解 (trivial solution)。因此，非齐次方程 $Bu = D^{-1}f$ 存在唯一解 $u \in X$，满足

$$\|u\|_X \le c\|D^{-1}f\|_X \le \tilde{c}\,\|f\|_{X'}.$$

□

8.4　变分法求解边界值问题

在介绍了常见的变分方法 (第8.3节) 之后，本届关注如何利用变分法分析求解二阶椭圆边界值问题，尤其是位势方程 (potential equations，第8.1节) 的边界值。应用变分法构建的边界值问题的若方程，是有限元分析的重要基础之一。

8.4.1 位势方程基本介绍

如第8.1节所讨论，先介绍三个算子，分别为实值标量的偏微分算子(8.1)

$$(L\,u)\,(x) := -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_j}\left[a_{ji}(x)\frac{\partial}{\partial x_i}u(x)\right], \quad x \in \Omega \in \mathbb{R}^d, \tag{8.111}$$

内界迹算子(8.4)

$$\gamma_0^{int}u(x) := \lim_{\Omega \ni \tilde{x} \mapsto x \in \Gamma} u\left(\tilde{x}\right), \quad x \in \Gamma = \partial\Omega,$$

与之相对应的内部共形导数(8.7)

$$\gamma_1^{int}u(x) := \lim_{\Omega \ni \tilde{x} \mapsto x \in \Gamma}\left[\sum_{i,j=1}^{d} n_j(x)a_{ji}\left(\tilde{x}\right)\frac{\partial}{\partial \tilde{x}_i}u\left(\tilde{x}\right)\right], \quad x \in \Gamma = \partial\Omega. \tag{8.112}$$

当 $u, \nu \in H^1(\Omega)$ 时，$Lu \in \widetilde{H}^{-1}(\Omega)$，格林第一恒等式 (8.6)化简为

$$\begin{aligned} a\left(u,\nu\right) &= \int_\Omega \left(L\,u\right)(x)\,\nu(x)\,dx + \int_\Gamma \left[\gamma_1^{int}u(x)\right]\left[\gamma_0^{int}\nu(x)\right]\,ds_x \\ &= \langle Lu, \nu\rangle_\Omega + \langle \gamma_1^{\mathrm{int}}u, \gamma_0^{\mathrm{int}}\nu\rangle_\Gamma, \end{aligned} \tag{8.113}$$

其中对称双线性泛函 $a(.,.)$ 的定义

$$a\left(u,\nu\right) := \langle u, \nu\rangle = \sum_{i,j=1}^{d}\int_\Omega a_{ji}(x)\frac{\partial}{\partial x_i}u(x)\,\frac{\partial}{\partial x_j}\nu(x)\,dx \tag{8.114}$$

Lemma 8.14. 假设 $a_{ij} \in L^\infty(\Omega)$, $i, j = 1, \ldots, d$ 满足

$$\|a\|_{L^\infty(\Omega)} := \max_{i,j=1,\ldots,d}\sup_{x\in\Omega}\left|a_{ij}(x)\right|. \tag{8.115}$$

那么我们可得线性形泛函 $a(.,.) : H^1(\Omega) \times H^1(\Omega) \mapsto \mathbb{R}$ 有界，并且满足

$$\begin{aligned} &\left|a(u,\nu)\right| \le c_2^A\left|u\right|_{H^1(\Omega)}\left|\nu\right|_{H^1(\Omega)}, \quad \forall u, \nu \in H^1(\Omega), \\ &c_2^A := d\,\|a\|_{L^\infty(\Omega)}. \end{aligned} \tag{8.116}$$

证明. 由(8.114)，(8.115)可得

$$\begin{aligned} \left|a\left(u,\nu\right)\right| &= \Big|\sum_{i,j=1}^{d}\int_\Omega a_{ji}(x)\frac{\partial}{\partial x_i}u(x)\,\frac{\partial}{\partial x_j}\nu(x)\,dx\Big| \\ &\le \|a\|_{L^\infty(\Omega)}\int_\Omega\left\{\sum_{i=1}^{d}\left|\frac{\partial}{\partial x_i}u(x)\right|\sum_{j=1}^{d}\left|\frac{\partial}{\partial x_j}\nu(x)\right|\right\}dx, \end{aligned}$$

连续两次使用柯西——施瓦茨不等式 (Definition 8.21)，上式变为

$$\begin{aligned} \left|a\left(u,\nu\right)\right| &\le \|a\|_{L^\infty(\Omega)}\left\{\int_\Omega\left[\sum_{i=1}^{d}\left|\frac{\partial}{\partial x_i}u(x)\right|\right]^2 dx\right\}^{\frac{1}{2}}\left\{\int_\Omega\left[\sum_{j=1}^{d}\left|\frac{\partial}{\partial x_j}\nu(x)\right|\right]^2 dx\right\}^{\frac{1}{2}} \\ &\le \|a\|_{L^\infty(\Omega)}\left\{\int_\Omega d\sum_{i=1}^{d}\left|\frac{\partial}{\partial x_i}u(x)\right|^2 dx\right\}^{\frac{1}{2}}\left\{\int_\Omega d\sum_{j=1}^{d}\left|\frac{\partial}{\partial x_j}\nu(x)\right|^2 dx\right\}^{\frac{1}{2}} \\ &= \underbrace{d\,\|a\|_{L^\infty(\Omega)}}_{=:c_2^A}\|\nabla u\|_{L^2(\Omega)}\|\nabla\nu\|_{L^2(\Omega)}. \end{aligned}$$

□

可见，由 Lemma 8.14可得，(8.116)又进一步表示为

$$\left|a\left(u,\nu\right)\right| \leq c_2^A \left\|u\right\|_{H^1(\Omega)} \left\|\nu\right\|_{H^1(\Omega)}, \quad \forall u,\nu \in H^1(\Omega). \tag{8.117}$$

Lemma 8.15 (双线性算子的半椭圆特性)**.** 设 L 是一个如(8.111)所定义的一致椭圆偏微分算子。用双线性形式 $(a.,.)$ (8.114)表示，我们有

$$a(\nu,\nu) \geq \lambda_0 \left|\nu\right|^2_{H^1(\Omega)}, \quad \forall \nu \in H^1(\Omega), \tag{8.118}$$

其中常数 $\lambda_0 > 0$，见椭圆算子(8.2)。

证明. 设一个 $w_i(x)$

$$w_i(x) := \frac{\partial}{\partial x_i}\nu(x), \quad i = 1,\ldots,d,$$

$$\begin{aligned}
\hookrightarrow a(\nu,\nu) &= \sum_{i,j=1}^{d} \int_\Omega a_{ji}(x) \frac{\partial}{\partial x_i}\nu(x) \frac{\partial}{\partial x_i}\nu(x)\, dx \\
&= \int_\Omega \left[A\underline{w}(x), \underline{w}(x)\right]\, dx \\
&\geq \lambda_0 \int_\Omega \left[\underline{w}(x), \underline{w}(x)\right]\, dx \\
&= \lambda_0 \left\|\nabla\nu\right\|^2_{L^2(\Omega)} = \lambda_0 \left\|\nu\right\|^2_{W^{1,2}(\Omega)} = \lambda_0 \left\|\nu\right\|^2_{H^1(\Omega)}.
\end{aligned}$$

□

8.4.2　狄利克雷边界值问题 1

将狄利克雷边界值问题改写为变分问题

回顾一下第8.1节的狄利克雷边界值问题 (8.11)-(8.12a)：基于给定的 f 和 g，求解

$$\begin{aligned}
(Lu)(x) &= f(x), \quad x \in \Omega, \\
\gamma_0^{\text{int}} u(x) &= g(x), \quad x \in \Gamma,
\end{aligned} \tag{8.119}$$

对应弱形式下的流形空间

$$\begin{aligned}
V_g &:= \left\{\nu \in H^1(\Omega) : \gamma_0^{\text{int}}\nu(x) = g(x), \quad x \in \Gamma\right\}, \\
V_0 &= H_0^1(\Omega).
\end{aligned} \tag{8.120}$$

根据第8.3介绍的知识，我们可以将狄利克雷边界值问题(8.119)-(8.120)改写为变分问题：

$$a(u,\nu) = \langle f,\nu\rangle_\Omega, \quad \forall\, \nu \in V_0, \tag{8.121}$$

其中双线性泛函 $a(u,\nu)$ 由格林第一恒等式(8.113)定义。研究目标是寻找变分问题的解 $u \in V_g$。由此狄利克雷边界问题表现为流形 V_g 中的一个附属条件，因此又称为基本边界条件 (essential boundary conditions)。

解的存在性与唯一性

由上所述可以看出，狄利克雷边界值问题对应的变分问题，可以归入带限制条件的算子方程问题类型，可用相应的变分方法求解 (第(8.3.4)节)。借助 Theorem 8.17和 Corollary 8.4，我们可以证明 $u \in V_g$ 存在且唯一。

Theorem 8.23 (狄利克雷边界值问题的弱形式解)**.** 给定 $f \in H^{-1}(\Omega), g \in H^{\frac{1}{2}}(\Gamma)$，变分问题(8.119)存在唯一的解 $u \in H^1(\Omega)$，满足

$$\|u\|_{H^1(\Omega)} \leq \frac{1}{c_1^A}\|f\|_{H^{-1}(\Omega)} + \left(1+\frac{c_2^A}{c_1^A}\right) c_{IT}\|g\|_{H^{\frac{1}{2}}(\Gamma)}. \tag{8.122}$$

证明. 与 Theorem 8.17类似，假设解 u 由两部分相加而得，$u := u_0 + u_g$，其中 u_g 是 $\gamma_0^{\text{int}} u_g = g$ 的解，u_0 是 $\gamma_1^{\text{int}}(u_0+u_g) = f$ 的解。显然，u 的唯一存在，可由 u_g 和 u_0 的唯一存在所分别证得。

1. 证明 u_g 的唯一存在性。根据给定条件 $g \in H^{\frac{1}{2}}(\Gamma)$，应用逆迹定理 Theorem 8.11，可得存在唯一一个有界延拓 $u_g \in H^{\Omega}$，满足

$$\begin{aligned}
&\gamma_0^{\text{int}} u_g = g, \\
&\|u_g\|_{H^1(\Omega)} \leq c_{\text{IT}}\|g\|_{H^{\frac{1}{2}}(\Gamma)}.
\end{aligned}$$

2. 在求得唯一解 u_g 的基础上，继续求唯一解 u_0。变分问题(8.121)变为，寻找解 $u_0 \in V_0 = H_0^1(\Omega)$，使满足新的变分问题

$$a(u_0, \nu) = \langle f, \nu\rangle_\Omega - a(u_g, \nu), \quad \forall \nu \in V_0. \tag{8.123}$$

   (a) 由范数等价定理 (Theorem 8.3-8.4, (8.38)) 得，$H^1(\Omega)$ 中的等价范为

$$\|u_0\|_{W^{1,2}(\Omega),\Gamma} = \left\{\left[\int_\Gamma \gamma_0^{\text{int}} u_0(x)\, ds_x\right]^2 + \|\nabla u_0\|_{L^2(\Omega)}^2\right\}^{\frac{1}{2}},$$

   即 $a(.,.)$ 有界。

   (b) 由 Lemma 8.15 (8.118)可得

$$\begin{aligned}
a(u_0, u_0) &\geq \lambda_0 |u_0\|_{H^1(\Omega)} \\
&= \lambda_0 \|\nabla u_0\|_{W^{1,2}(\Omega),\Gamma}^2 \\
&\geq c_1^A \|u_0\|_{H^1(\Omega)}^2,
\end{aligned} \tag{8.124}$$

   即 $a(.,.)$ 是个 V_0-椭圆算子。

   (c) $a(.,.)$ 有界且 V_0-椭圆，满足拉克斯一密格拉蒙定理 (Theorem 8.14) 所需的前提条件，根据该定理，变分问题(8.123)有唯一解 $u_0 \in V_0$，满足

$$\begin{aligned}
c_1^A \|u_0\|_{H^1(\Omega)}^2 &\leq a(u_0, u_0) \\
&= \langle f, u_0\rangle_\Omega - a(u_g, u_0),
\end{aligned}$$

   代入 Lemma 8.10，上式进一步变为

$$\begin{aligned}
c_1^A \|u_0\|_{H^1(\Omega)}^2 &\leq \left(\|f\|_{H^{-1}(\Omega)} + c_2^A \|u_g\|_{H^1(\Omega)}\right) \|u_0\|_{H^1(\Omega)}, \\
\hookrightarrow \|u_0\|_{H^1(\Omega)} &\leq \frac{1}{c_1^A}\|f\|_{H^{-1}(\Omega)} + \frac{c_2^A}{c_1^A}\|u_g\|_{H^1(\Omega)}.
\end{aligned}$$

3. 在此基础上我们有

$$\begin{aligned}
\|u_0\|_{H^1(\Omega)} + \|u_g\|_{H^1(\Omega)} &\leq \frac{1}{c_1^A}\|f\|_{H^{-1}(\Omega)} + \left(1+\frac{c_2^A}{c_1^A}\right)\|u_g\|_{H^1(\Omega)}, \\
\hookrightarrow \|u\|_{H^1(\Omega)} &\leq \frac{1}{c_1^A}\|f\|_{H^{-1}(\Omega)} + \left(1+\frac{c_2^A}{c_1^A}\right) c_{\text{IT}}\|g\|_{H^{\frac{1}{2}}(\Gamma)}.
\end{aligned} \tag{8.125}$$

□

共形导数

变分问题(8.121)的唯一解 $u \in V_g$，常常又称为狄利克雷边界值问题(8.119)的弱形式解。在此基础上，对于 $f \in \widetilde{H}^{-1}(\Omega)$，可通过构建变分问题进一步求解共形导数 $\gamma_1^{\text{int}} u \in H^{-\frac{1}{2}(\Gamma)}$

$$\langle \gamma_1^{\text{int}} u, z\rangle_\Gamma = a(u, \varepsilon_z) - \langle f, \varepsilon_z\rangle_\Omega, \quad \forall z \in H^{\frac{1}{2}}(\Gamma), \tag{8.126}$$

其中 ε_z 是逆迹定理 Theorem 8.11所定义的有界延拓算子。(8.126)的唯一可解条件由 Theorem 8.16给出，由此我们可以假定如下稳定性条件

$$\|w\|_{H^{-\frac{1}{2}}(\Gamma)} = \sup_{0\neq z\in H^{\frac{1}{2}}(\Gamma)} \frac{\langle w, z\rangle_\Gamma}{\|z\|_{H^{\frac{1}{2}}(\Gamma)}}, \quad \forall w \in H^{-\frac{1}{2}(\Gamma)}. \tag{8.127}$$

在此基础上我们有共形导数的解

Lemma 8.16 (共形导数的解)**.** *设给定 $f \in \widetilde{H}^{-1}(\Omega), g \in H^{\frac{1}{2}}(\Gamma)$，我们有 $u \in H^1(\Gamma)$ 是狄利克雷边界值问题(8.121)的唯一解。*

那么相应的共形导数 $\gamma_1^{int} u \in H^{-\frac{1}{2}}(\Gamma)$ 满足

$$\|\gamma_1^{int} u\|_{H^{-\frac{1}{2}}(\Gamma)} \le c_{IT} \left\{ \|f\|_{\widetilde{H}^{-1}(\Omega)} + c_2^A |u|_{H^1(\Omega)} \right\}. \tag{8.128}$$

证明. 稳定性条件(8.127)⇒

$$\|\gamma_1^{\text{int}} u\|_{H^{-\frac{1}{2}}(\Gamma)} = \sup_{0\neq z\in H^{\frac{1}{2}}(\Gamma)} \frac{\langle \gamma_1^{\text{int}} u, z\rangle_\Gamma}{\|z\|_{H^{\frac{1}{2}}(\Gamma)}},$$

变分问题(8.126)、Lemma (8.10)⇒

$$\begin{aligned}\|\gamma_1^{\text{int}} u\|_{H^{-\frac{1}{2}}(\Gamma)} &= \sup_{0\neq z\in H^{\frac{1}{2}}(\Gamma)} \frac{\left|a(u, \varepsilon_z) - \langle f, \varepsilon_z\rangle_\Omega\right|_\Gamma}{\|z\|_{H^{\frac{1}{2}}(\Gamma)}}, \\ &\le \left\{ c_2^A |u|_{H^1(\Omega)} + \|f\|_{\widetilde{H}^{-1}(\Omega)} \right\} \underbrace{\sup_{0\neq z\in H^{\frac{1}{2}}(\Gamma)} \frac{\|\varepsilon_z\|_{H^1(\Omega)}}{\|z\|_{H^{\frac{1}{2}}(\Gamma)}}}_{=:\mathcal{A}}\end{aligned}$$

由逆迹定理 Theorem 8.11可得

$$\mathcal{A} \le c_{\text{IT}},$$

因此

$$\|\gamma_1^{\text{int}} u\|_{H^{-\frac{1}{2}}(\Gamma)} \le c_{\text{IT}} \left\{ \|f\|_{\widetilde{H}^{-1}(\Omega)} + c_2^A |u|_{H^1(\Omega)} \right\}.$$

□

带有齐次偏微分方程的狄利克雷边界值问题弱解

来考虑一类特殊的狄利克雷边界值问题，即含有齐次偏微分方程 $f \equiv 0$ 的情况，其解 u 对于边界积分算子的分析具有重要意义。

Corollary 8.5 (带有齐次偏微分方程的狄利克雷边界值问题弱解)**.** *设 $u \in H^1(\Omega)$ 是以下带有齐次偏微分方程的狄利克雷边界值问题的弱形式解*

$$\begin{aligned}(Lu)(x) &= 0, \quad x \in \Omega, \\ \gamma_0^{int} u(x) &= g(x), \quad x \in \Gamma,\end{aligned}$$

其中 L 是一个一致椭圆的二阶偏微分算子。

那么我们有

$$a(u, u) \ge c\|\gamma_1^{int} u\|^2_{H^{-\frac{1}{2}}(\Gamma)}. \tag{8.129}$$

证明. 1. 当 $f \equiv 0$ 时，共形导数算子和逆迹的关系，由 Lemma 8.16的 (8.128)变为

$$\left\|\gamma_1^{\mathrm{int}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)}^2 \leq \left[c_{\mathrm{IT}}\, c_2^A\right]^2 \left|u\right|_{H^1(\Omega)}^2.$$

2. 由双线性算子的半椭圆特性 Lemma 8.15 (8.118)可得

$$\lambda_0 \left|u\right|_{H^1(\Omega)}^2 \leq a(u,u).$$

□

狄利克雷边界值问题的强解简述

当 Ω 是利普希茨域时，基于给定的 f 和 g，我们可以引入更严格的假设条件，构建强形式的正则解如 u、$\gamma_1^{\mathrm{int}} u$ 等。

Theorem 8.24. 设 $\Omega \in \mathbb{R}^d$ 是个有界的利普希茨域，边界为 $\Gamma = \partial\Omega$。设 $u \in H^1(\Omega)$ 是狄利克雷边界值问题的弱形式解

$$\begin{aligned}(Lu)(x) &= f(x), \quad x \in \Omega,\\ \gamma_0^{int} u(x) &= g(x), \quad x \in \Gamma.\end{aligned}$$

如果给定的 f 和 g 满足 $f \in L^2(\Omega), g \in H^1(\Gamma)$，那么我们有 $u \in H^{\frac{3}{2}}(\Omega), \gamma_1^{int} u \in L^2(\Gamma)$，并且

$$\begin{aligned}\left\|u\right\|_{H^{\frac{3}{2}}(\Omega)} &\leq c_1 \left\{\left\|f\right\|_{L^2(\Omega)} + H^1(\Gamma)\right\},\\ \left\|\gamma_1^{int} u\right\|_{L^2(\Gamma)} &\leq c_2 \left\{\left\|f\right\|_{L^2(\Omega)} + H^1(\Gamma)\right\}.\end{aligned}$$

我们甚至可以使假定条件更加严格，如 $\Gamma = \partial\Omega$ 是平滑或分段平滑的边界，Ω 是凹的，$f \in L^2(\Omega)$ 等。若 $g = \gamma_0^{\mathrm{int}} u$ 是方程解 $u_g \in H^2(\Omega)$ 的迹，则我们有 $u \in H^2(\Omega)$。更多强形式解的讨论，可见Demkowicz (2006); Demkowicz et al. (2007)。

8.4.3 狄利克雷边界值问题 2

如第8.3.5节所述，狄利克雷边界值问题(8.119)也可以改写为鞍点变分问题，即混合算子方程，共形导数对应拉格朗日乘子 (Babuška, 1973; Bramble, 1981)。

从格林第一恒等式(8.113)入手，设拉格朗日乘子

$$\lambda := \gamma_1^{\mathrm{int}} u \in H^{-\frac{1}{2}}(\Gamma),$$

进而鞍点变分问题表示为，寻找解 $(u,\lambda) \in H^1(\Omega) \times H^{-\frac{1}{2}}(\Gamma)$，使得满足

$$\begin{aligned}a(u,\nu) - b(\nu,\lambda) &= \langle f, \nu\rangle_\Omega \quad \forall \nu \in H^1(\Omega),\\ b(u,\mu) &= \langle g, \mu\rangle_\Gamma \quad \forall \mu \in H^{-\frac{1}{2}}(\Gamma),\end{aligned} \tag{8.130}$$

其中定义了一个新的双线性泛函算子

$$b(\nu,\mu) := \langle \gamma_0^{\mathrm{int}} \nu, \mu\rangle_\Gamma, \quad (\nu,\mu) \in H^1(\Gamma) \times H^{-\frac{1}{2}}(\Gamma).$$

解的唯一存在性

鞍点变分形式的狄利克雷边界值问题(8.130)，解的存在性和唯一性，可由 Theorem 8.19证得。使用该定理之前，需要确保两个前提条件得到满足。一是双线性泛函 $a(.,.)$ 的椭圆特性，二是解的稳定性条件。

1. 椭圆性。类似于式(8.124)，由 Lemma 8.15 (8.118)可得

$$a(.,.) \geq c_1^A \left\| \cdot \right\|_{H^1(\Omega)}^2,$$

此外由于

$$\ker B := \left\{ \nu \in H^1(\Omega) : \langle \gamma_0^{\text{int}} \nu, \mu \rangle_\Gamma = 0, \quad \forall \mu \in H^{-\frac{1}{2}}(\Gamma) \right\} = H_0^1(\Omega),$$

我们因此有，$a(.,.)$ 是一个 $\ker B$-椭圆 (或 H_0^1-椭圆) 的双线性形。(通常来说，我们需要求得一个扩展双线性形 $\tilde{a}(.,.)$，使得满足 $H^1(\Omega)$-椭圆性质，相关讨论见第8.4.3节。)

2. 解的稳定性条件可以表示为

$$c_S \left\| \mu \right\|_{H^{-\frac{1}{2}}(\Gamma)} \leq \sup_{0 \neq \nu \in H^1(\Omega)} \frac{\langle \gamma_0^{\text{int}} \nu, \mu \rangle_\Gamma}{\left\| \nu \right\|_{H^1(\Omega)}}, \quad \forall \mu \in H^{-\frac{1}{2}}(\Gamma), \tag{8.131}$$

其证明见可见 Lemma 8.17。

3. 应用定理 Theorem 8.19，求得鞍点变分问题的唯一解。

Lemma 8.17 (鞍点变分形式狄利克雷边界值问题解的稳定条件)**.** 稳定条件(8.131)成立。

证明. 已知给定的任一 $\mu \in H^{-\frac{1}{2}}(\Gamma)$。由里兹表现定理 (Theorem 8.13) 可得，存在唯一的一个 $u_\mu \in H^{\frac{1}{2}}(\Gamma)$，满足

$$\begin{aligned} \langle u_\mu, \nu \rangle_{H^{\frac{1}{2}}(\Gamma)} &= \langle \mu, \nu \rangle_\Gamma \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma), \\ \left\| u_\mu \right\|_{H^{\frac{1}{2}}(\Gamma)} &= \left\| \mu \right\|_{H^{\frac{1}{2}}(\Gamma)}. \end{aligned}$$

由逆迹定理 Theorem 8.11可得，存在一个延拓算子 $\varepsilon u_\mu \in H^1(\Omega)$，满足

$$\left\| \varepsilon u_\mu \right\|_{H^1(\Omega)} \leq c_{\text{IT}} \left\| u_\mu \right\|_{H^{\frac{1}{2}}(\Gamma)}.$$

那么，对于 $\nu = \varepsilon u_\mu \in H^1(\Omega)$ 我们有

$$\begin{aligned} \frac{\langle \nu, \mu \rangle_\Gamma}{\left\| \nu \right\|_{H^1(\Omega)}} &= \frac{\langle u_\mu, \mu \rangle_\Gamma}{\left\| \varepsilon u_\mu \right\|_{H^1(\Omega)}} = \frac{\langle u_\mu, u_\mu \rangle_{H^{\frac{1}{2}}(\Gamma)}}{\left\| \varepsilon u_\mu \right\|_{H^1(\Omega)}} \\ &\geq \frac{1}{c_{\text{IT}}} \left\| u_\mu \right\|_{H^{\frac{1}{2}}(\Gamma)} = \frac{1}{c_{\text{IT}}} \left\| u \right\|_{H^{-\frac{1}{2}}(\Gamma)} \end{aligned}$$

∴ 稳定性条件(8.131)成立。 □

调整鞍点变分问题

需要注意的是，在鞍点变分问题(8.119)中的双线性形式算子 $a(.,.)$ 是 $H_0^1(\Omega)$-椭圆的。我们常常需要将它扩展为一个 $H^1(\Omega)$-椭圆的算子 $\tilde{a}(.,.)$，对应新的调整鞍点变分问题。调整思路如下：

已知，用拉格朗日乘子 $\lambda := \gamma_1^{\text{int}} u \in H^{-\frac{1}{2}}(\Gamma)$ 来描述问题解 u 的共形导数，那么利用格林第二恒等式(8.8)，可得狄利克雷边界值问题的正交条件(8.15)

$$\begin{aligned} &\int_\Omega f(x)\,dx + \int_\Gamma \lambda(x) ds_x = 0, \\ \hookrightarrow &\int_\Gamma \lambda(x) ds_x \int_\Gamma \mu(x)\,ds_x = -\int_\Omega f(x)\,dx \int_\Gamma \mu(x)\,ds_x, \quad \forall \mu \in H^{-\frac{1}{2}}(\Gamma). \end{aligned} \tag{8.132}$$

另一方面，根据狄利克雷边界条件有

$$\begin{aligned} &\gamma_0^{\text{int}} u = g, \\ \hookrightarrow &\int_\Gamma \gamma_0^{\text{int}} u(x)\,ds_x = \int_\Gamma g(x)\,ds_x, \\ \hookrightarrow &\int_\Gamma \gamma_0^{\text{int}} u(x)\,ds_x \int_\Gamma \gamma_0^{\text{int}} \nu(x)\,ds_x = \int_\Gamma g(x)\,ds_x \int_\Gamma \gamma_0^{\text{int}} \nu(x)\,ds_x, \quad \forall \nu \in H^1(\Omega). \end{aligned} \tag{8.133}$$

将(8.132)、(8.133)代入(8.130)，得到调整鞍点变分问题：寻找解 $(u,\lambda) \in H^1(\Omega) \times H^{-\frac{1}{2}}(\Gamma)$，使得 $\forall (\nu,\mu) \in H^1(\Omega) \times H^{-\frac{1}{2}}(\Gamma)$ 均满足

$$\begin{aligned} &\underbrace{\int_\Gamma \gamma_0^{\text{int}} u(x)\,ds_x \int_\Gamma \gamma_0^{\text{int}} \nu(x)\,ds_x + a(u,\nu)}_{=:\tilde{a}(u,\nu)} - b(\nu,\lambda) = \langle f,\nu\rangle_\Omega + \int_\Gamma g(x)\,ds_x \int_\Gamma \gamma_0^{\text{int}} \nu(x)\,ds_x, \\ &b(u,\mu) + \int_\Gamma \lambda(x) ds_x \int_\Gamma \mu(x)\,ds_x = \langle g,\mu\rangle_\Gamma - \int_\Omega f(x)\,dx \int_\Gamma \mu(x)\,ds_x. \end{aligned} \tag{8.134}$$

下面的问题就是，调整鞍点变分问题(8.134)的解是否存在，是否唯一，以及是否与原鞍点变分问题(8.130)的解一致。换句话说，两个鞍点变分问题是否等价。

Theorem 8.25 (变分问题等价)**.** 调整鞍点变分问题(8.134)有唯一的解 $(u,\lambda) \in H^1(\Omega) \times H^{-\frac{1}{2}}(\Gamma)$，并且与鞍点变分问题(8.130)的解一致。即，两个变分问题等价。

证明. 1. 证明扩展双线性形 $\tilde{a}(u,\nu)$ 有界。

$a(u,\nu)$ 有界 $\Rightarrow$

$$\tilde{a}(u,\nu) := \int_\gamma \gamma_0^{\text{int}} u(x) ds_x \int_\gamma \gamma_0^{\text{int}} \nu(x) ds_x + a(u,\nu), \quad \forall u,\nu \in H^1(\Omega)$$

有界。

2. 由 $a(.,.)$ 的半椭圆性质 (Lemma 8.15) 和(8.38)有

$$\begin{aligned} \tilde{a}(\nu,\nu) &= \left[\int_\Gamma \gamma_0^{\text{int}} \nu(x)\,ds_x\right]^2 + a(\nu,\nu) \\ &\geq \min\{1,\lambda_0\} \left\|\nu\right\|^2_{W^{1,2}(\Omega),\Gamma} \\ &\geq c_1^{\tilde{A}} \left\|\nu\right\|^2_{H^1(\Omega)}, \quad \forall \nu \in H^1(\Omega), \end{aligned}$$

由此可得 $\tilde{a}(u,\nu)$ 是 $H^1(\Omega)$-椭圆。

3. 前提条件得到满足，可通过 Theorem 8.19、Theorem 8.20证得，调整鞍点变分问题(8.134)有唯一解。

4. 对于 $(\nu,\mu) \equiv (1,1)$ 的特殊情况，(8.134)变为

$$\begin{aligned} &|\Gamma| \int_\Gamma \gamma_0^{\text{int}} u(x)\,ds_x - \int_\Gamma \lambda(x)\,ds_x = \int_\Omega f(x)\,dx + |\Gamma| \int_\Gamma g(x)\,ds_x, \\ &\int_\Gamma \gamma_0^{\text{int}} u(x)\,ds_x + |\Gamma| \int_\Gamma \lambda(x)\,ds_x = \int_\Gamma g(x)\,ds_x, -|\Gamma| \int_\Omega f(x)\,dx. \end{aligned}$$

进而

$$\left(1+|\Gamma|^2\right)\int_\Gamma \gamma_0^{\text{int}}u(x)ds_x=\left(1+|\Gamma|^2\right)\int_\Gamma g(x)\,ds_x$$

即(8.133)。从而有

$$|\Gamma|\int_\Gamma \lambda(x)\,ds_x=-|\Gamma|\int_\Gamma f(x)\,ds_x,$$

即(8.132)。

因此可见，(u,λ) 也是鞍点变分问题(8.130)的解的解；两个问题等价。

□

8.4.4　诺依曼边界值问题

将诺依曼边界值问题改写为变分问题

回顾一下第8.1节的诺依曼边界值问题(8.11)-(8.12b)：基于给定的 f 和 g，求解

$$\begin{aligned}(Lu)(x)&=f(x),\quad x\in\Omega,\\ \gamma_1^{\text{int}}u(x)&=g(x),\quad x\in\Gamma.\end{aligned}\tag{8.135}$$

假定 f 和 g 满足可求解性条件 (8.17)

$$\int_\Omega f(x)\,dx+\int_\Gamma g(x)\,ds_x=0.\tag{8.136}$$

基于前文的分析可见，诺依曼边界值问题(8.135)的解 $u\in H^1(\Omega)$ 将与某一个常数有关。为了将此常数予以确定，可以在 $H^1(\Omega)$ 中定义一个测试空间 $H_*^1(\Omega)$，测试方程 $\nu(x)$ 用于规模调节。

$$H_*^1(\Omega):=\left\{\nu\in H^1(\Omega):\int_\Omega \nu(x)\,dx=0\right\}.\tag{8.137}$$

从而构建变分问题，求解 $u\in H_*^1(\Omega)$ 使得满足

$$a(u,\nu)=\langle f,\nu\rangle_\Omega+\langle g,\gamma_0^{\text{int}}\nu\rangle_\Gamma,\quad \forall\,\nu\in H_*^1(\Omega).\tag{8.138}$$

解的存在性与唯一性

Theorem 8.26 (诺依曼边界值问题的变分法求解)**.** *给定 $f\in\tilde{H}^{-1}(\Omega),g\in H^{-\frac{1}{2}}(\Gamma)$，满足可求解性条件(8.136)。*

则诺依曼边界值的变分问题(8.138)存在唯一的解 $u\in H_*^1(\Omega)$，满足

$$\|u\|_{H^1(\Omega)}\le\frac{1}{\tilde{c}_1^A}\left\{\|f\|_{\tilde{H}^{-1}(\Omega)}+c_T\|g\|_{H^{-\frac{1}{2}}(\Gamma)}.\right\}$$

证明.　1. 证 $a(.,.)$ 有界且椭圆。

对于 $\nu\in H_*^1(\Omega)$，它在 $H^1(\Omega)$ 中的等价范，可由(8.38)求得

$$\|\nu\|_{W^{1,2}(\Omega),\Omega}=\left\{\left[\int_\Omega \nu(x)\,dx\right]^2+\|\nabla\nu\|_{L^2(\Omega)}^2\right\}^{\frac{1}{2}},$$

进而利用 Lemma 8.15可得，

$$a(\nu,\nu)\ge\lambda_0\|\nabla\nu\|_{W^{1,2}(\Omega),\Omega}^2\ge\tilde{c}_1^A\|\nu\|_{H^1(\Omega)}^2,\quad\forall\nu\in H_*^1(\Omega),\tag{8.139}$$

因此可得双线性形式 $a(.,.)$ 的 H_*^1-椭圆特性。

2. 根据拉克斯—密格拉蒙定理 (Theorem 8.14)，证明变分问题(8.138)存在唯一解。

3. 将求得的变分问题唯一解 $u \in H^1_*(\Omega)$ 代回椭圆条件 (8.139)

$$\begin{aligned}\widetilde{c}_1^A \left\|u\right\|^2_{H^1(\Omega)} &\leq a(u,u) = \langle f, u\rangle_\Omega + \langle g, \gamma_0^{\text{int}} u\rangle_\Gamma \\ &\leq \left\|f\right\|_{\widetilde{H}^{-1}(\Omega)} \left\|u\right\|_{H^1(\Omega)} + \left\|g\right\|_{H^{-\frac{1}{2}}(\Gamma)} \left\|\gamma_0^{\text{int}} u\right\|_{H^{\frac{1}{2}}(\Gamma)}.\end{aligned}$$

由迹定理 Theorem 8.10得，上式变为

$$\widetilde{c}_1^A \left\|u\right\|_{H^1(\Omega)} \leq \left\|f\right\|_{\widetilde{H}^{-1}(\Omega)} + \left\|g\right\|_{H^{-\frac{1}{2}}(\Gamma)} c_T \left\|u\right\|_{H^1(\Gamma)}.$$

□

鞍点变分问题

类似地，我们也可以构建一个与(8.138)等价的鞍点变分问题。此时，用于规模调节的测试空间 $H^1_*(\Omega)$ 以副条件 (side condition) 的情况出现。使用一个拉格朗日乘子，寻找解 $(u,\lambda) \in H^1(\Omega) \times \mathbb{R}$，满足

$$\begin{aligned}&a(u,\nu) + \lambda \int_\Omega \nu(x)\,dx = \langle f, \nu\rangle_\Omega + \langle g, \gamma_0^{\text{int}} \nu\rangle_\Gamma, \\ &\int_\Omega u(x)\,dx = 0, \quad \forall\, \nu \in H^1(\Omega).\end{aligned} \tag{8.140}$$

Theorem 8.27 (诺依曼边界值问题的鞍点变分法求解)**.** 诺依曼边界值的鞍点变分问题(8.140)有唯一解 $(u,\lambda) \times H^1\Omega \times \mathbb{R}$。

证明. 1. 双线性形式 $b(.,.)$ 有界

$$b(\nu,\mu) := \mu \int_\Omega \nu(x)\,dx, \quad \forall\, \nu \in H^1(\Omega), \mu \in \mathbb{R},$$

并且有 $\ker B = H^1_*(\Omega)$。

2. 进而由椭圆性(8.139)得，双线性形 $a(.,.)$ 是 $\ker B$-椭圆的。

3. 证明满足稳定性条件

$$c_S \left|\mu\right| \leq \sup_{0\neq\nu\in H^1(\Omega)} \frac{b(\nu,\mu)}{\left\|\nu\right\|_{H^1(\Omega)}}, \quad \forall\, \mu \in \mathbb{R}. \tag{8.141}$$

对于任一给定的 $\mu \in \mathbb{R}$，定义 $\nu^* := \nu \in H^1(\Omega)$，可以证得(8.141)，其中 $c_S = |\Omega|^{-\frac{1}{2}}$。

4. 根据定理 Theorem 8.19可求得，诺依曼边界值的鞍点变分问题存在唯一解 $(u,\lambda) \in H^1(\Omega)\times\mathbb{R}$。

5. 对于测试方程 $\nu \equiv 1$ 的特殊情况，根据诺依曼边界值问题的可求解条件(8.136)我们有拉格朗日乘子的值

$$\lambda = 0.$$

□

调整鞍点变分问题

如前所述，鞍点变分问题(8.140)也可以改写为一个新的鞍点变分问题：基于给定的任一 $f \in \widetilde{H}^{-1}(\Omega), g \in H^{-\frac{1}{2}}(\Gamma)$，寻找 $(u, \lambda) \in H^1(\Omega) \times \mathbb{R}$，使满足

$$\begin{aligned} a(u,\nu) + \lambda \int_\Omega \nu(x)\,dx &= \langle f, \nu\rangle_\Omega + \langle g, \gamma_0^{\text{int}}\nu\rangle_\Gamma, \\ \int_\Gamma u(x)\,dx - \lambda &= 0, \quad \forall \nu \in H^1(\Omega). \end{aligned} \tag{8.142}$$

利用第二行等式求得拉格朗日乘子 $\lambda \in \mathbb{R}$ 代入第一行，我们得到一个新的调整变分问题，寻找 $u \in H^1(\Omega)$ 使满足

$$a(u,\nu) + \int_\Omega u(x)dx \int_\Omega \nu(x)dx = \langle f, \nu\rangle_\Omega + \langle g, \gamma_0^{\text{int}}\nu\rangle_\Gamma, \quad \forall \nu \in H^1(\Omega). \tag{8.143}$$

Theorem 8.28 (诺依曼边界值的调整变分问题解)**.** 基于给定的任一 $f \in \widetilde{H}^{-1}(\Omega), g \in H^{-\frac{1}{2}}(\Gamma)$，调整变分问题(8.143)有唯一的解 $u \in H^1(\Omega)$。

如果给定的 $f \in \widetilde{H}^{-1}(\Omega), g \in H^{-\frac{1}{2}}(\Gamma)$ 满足可求解条件(8.136)，那么调整变分问题(8.143)的解 $u \in H^1_*(\Omega)$；换句话说，调整变分问题(8.143)和变分问题(8.140)等价。

证明.　1. 由(8.143)可见，调整双线性形 $\widetilde{a}(.,.)$ 写为

$$\widetilde{a}(u,\nu) := a(u,\nu) + \int_\Omega u(x)\,dx \int_\Omega \nu(x)\,dx.$$

由 $a(.,.)$ 的半椭圆属性 Lemma 8.15我们有

$$\begin{aligned} \widetilde{a}(\nu,\nu) &\geq \lambda_0 \left\|\nabla\nu\right\|^2_{L^2(\Omega)} + \left[\int_\Omega \nu(x)dx\right]^2 \\ &\geq \min\{\lambda_0, 1\} \left\|\nu\right\|^2_{W^{1,2}(\Omega),\Omega} \\ &\geq \hat{c}_1^A \left\|\nu\right\|^2_{W^{1,2}(\Omega),\Omega}, \quad \forall \nu \in H^1(\Omega), \end{aligned} \tag{8.144}$$

可得 $\widetilde{a}(.,.)$ 是 $H^1(\Omega)$-椭圆且有界的。

2. 满足前提条件后，可由拉克斯—密格拉蒙定理 (Theorem 8.14) 证得，调整变分问题(8.143)有唯一解 $u \in H^1(\Omega)$。

3. 设测试方程 $\nu(x) \equiv 1$。调整变分问题(8.143)变为

$$\begin{aligned} \left|\Omega\right| \int_\Omega u(x)\,dx &= \langle f, 1\rangle_\Omega + \langle g, 1\rangle_\Gamma, \\ &= \int_\Omega f(x)\,dx + \int_\Gamma g(x)\,ds_x = 0, \end{aligned}$$

其中最后一个等式用到可求解性条件(8.136)。由此我们有 $u \in H^1_*(\Omega)$，也是变分问题(8.140)的解；换句话说，两个问题等价。

□

诺依曼边界值问题的通解

利用变分法求解诺依曼边界值问题(8.135)，所得到的 $u \in H^1_*(\Omega)$ 是弱形式解。更一般意义上的通解 $\widetilde{u} \in H^1(\Omega)$ 可写为

$$\widetilde{u} := u + \alpha,$$

其中 $\alpha \in \mathbb{R}$ 是任意常数。

8.4.5 混合边界值问题

回顾一下第8.1节的混合边界值问题(8.11)，(8.12a)-(8.12b)

$$\begin{aligned}(Lu)(x) &= f(x), \quad x \in \Omega,\\ \gamma_0^{\text{int}} u(x) &= g_D(x), \quad x \in \Gamma_D,\\ \gamma_1^{\text{int}} u(x) &= g_N(x), \quad x \in \Gamma_N,\end{aligned} \tag{8.145}$$

假定 $\Gamma = \overline{\Gamma}_D \cup \overline{\Gamma}_N$ 。对应地，可由格林第一恒等式(8.6)建立变分问题：求解 $u \in H^1(\Omega), \gamma_0^{\text{int}} u(x) = g_D(x), x \in \Gamma_D$，使满足

$$a(u,\nu) = \langle f,\nu\rangle_\Omega + \langle g_N, \gamma_0^{\text{int}}\nu\rangle_{\Gamma_N}, \quad \forall \nu \in H_0^1(\Omega,\Gamma_D), \tag{8.146}$$

其中

$$H_0^1(\Omega,\Gamma_D) := \left\{\nu \in H^1(\Omega) : \gamma_0^{\text{int}}\nu(x) = 0, \quad x \in \Gamma_D\right\}.$$

混合边界值变分问题(8.146)的唯一解，可由以下定理证明

Theorem 8.29 (混合边界值变分问题的唯一解)**.** *给定 $f \in \widetilde{H}^{-1}(\Omega), g_D \in H^{\frac{1}{2}}(\Gamma_D), g_N \in H^{-\frac{1}{2}}(\Gamma_N)$，那么混合边界值变分问题(8.146)有唯一解 $u \in H^1(\Omega)$，满足*

$$\|u\|_{H^1(\Omega)} \le c\left[\|f\|_{\widetilde{H}^{-1}(\Omega)} + \|g_D\|_{H^{\frac{1}{2}}(\Gamma_D)} + \|g_N\|_{H^{-\frac{1}{2}}(\Gamma_N)}\right]. \tag{8.147}$$

证明. 1. 定义两个延拓算子。

(a) 给定 $g_D \in H^{\frac{1}{2}}(\Gamma_D)$，定义一个有界的延拓 $\widetilde{g}_D \in H^{\frac{1}{2}}(\Gamma)$，满足

$$\|\widetilde{g}_D\|_{H^{\frac{1}{2}}(\Gamma)} \le c\,\|g_D\|_{H^{\frac{1}{2}}(\Gamma_D)}.$$

(b) 定义第二个延拓 $u_{\widetilde{g}_D} \in H^{\frac{1}{2}}(\Omega)$，使得 $\gamma_0^{\text{int}} u_{\widetilde{g}_D} = \widetilde{g}_D$，并利用逆迹定理 Theorem 8.11得

$$\|u_{\widetilde{g}_D}\|_{H^1(\Omega)} \le c_{\text{IT}}\,\|\widetilde{g}_D\|_{H^{\frac{1}{2}}(\Gamma)}.$$

2. 基于(8.146)，构建新的变分问题。

$$a(u_0,\nu) = \langle f,\nu\rangle_\Omega + \langle g_N, \gamma_0^{\text{int}}\nu\rangle_{\Gamma_N} - a(u_{\widetilde{g}_D},\nu), \quad \forall \nu \in H_0^1(\Omega,\Gamma_D),$$

研究目标是求得唯一解 $u_0 \in H_0^1(\Omega,\Gamma_D)$。

根据(8.38)可定义 $H^1(\Omega)$ 中的等价范

$$\|\nu\|_{W^{1,2}(\Omega),\Gamma_D} := \left\{\left[\int_{\Gamma_D} \gamma_0^{\text{int}}\nu(x)\,ds_x\right]^2 + \|\nabla\nu\|_{L^2(\Omega)}^2\right\}^{\frac{1}{2}}.$$

由双线性算子 $a(.,.)$ 的半椭圆特性 (Lemma 8.15) 可得

$$\begin{aligned}a(\nu,\nu) &\ge \lambda_0\,\|\nabla\nu\|_{L^2(\Omega)}^2\\ &= \lambda_0\,|\nu|_{W^{1,2}(\Omega),\Gamma_D}^2\\ &\ge c_1^A\,\|\nu\|_{W^{1,2}(\Omega)}^2.\end{aligned}$$

有界，椭圆，假设条件满足。由拉克斯—密格拉蒙定理 (Theorem 8.14) 可得，变分问题(8.146)有唯一解 $u_0 \in H_0^1(\Omega,\Gamma_D)$。

3. 将求得的唯一解 u_0 分别代回半椭圆 (Lemma 8.15) 和变分问题，可得

$$\begin{aligned} c_1^A \left\|u_0\right\|_{H^1(\Omega)}^2 &\leq a(u_0, u_0) \\ &= \langle f, u_0\rangle_\Omega + \langle g_N, \gamma_0^{\text{int}} u_0\rangle_{\Gamma_N} - a(u_{\widetilde{g}_D}, u_0) \\ &= \left[\left\|f\right\|_{\widetilde{H}^{-1}(\Omega)} + c_2^A \left\|u_{\widetilde{g}_D}\right\|_{H^1(\Omega)}\right] \left\|u_0\right\|_{H^1(\Omega)} + \left\|g_N\right\|_{H^{-\frac{1}{2}}(\Gamma_N)} \left\|\gamma_0^{\text{int}} u_0\right\|_{\widetilde{H}^{\frac{1}{2}}(\Gamma_N)}, \end{aligned}$$

证得(8.147)。

□

8.4.6　罗宾边界值问题

回顾一下第8.1节的罗宾边界值问题：基于给定的 f 和 g，求解

$$\begin{aligned} &(Lu)(x) = f(x), \quad x \in \Omega, \\ &\gamma_1^{\text{int}} u(x) + \kappa(x)\gamma_0^{\text{int}} u(x) = g(x), \quad x \in \Gamma. \end{aligned} \tag{8.148}$$

由格林第一恒等式(8.6)建立变分问题：求解 $u \in H^1(\Omega)$，使满足

$$a(u, \nu) + \int_\Gamma \kappa(x)\gamma_0^{\text{int}} u(x)\gamma_0^{\text{int}} \nu(x)\, ds_x = \langle f, \nu\rangle_\Omega + \langle g, \gamma_0^{\text{int}} \nu\rangle_\Gamma, \quad \forall \nu \in H^1(\Omega). \tag{8.149}$$

Theorem 8.30 (罗宾边界值变分问题的唯一解)**.** *给定 $f \in \widetilde{H}^{-1}(\Omega), g \in H^{-\frac{1}{2}}(\Gamma)$，假定 $\kappa(x) \geq \kappa_0 > 0, \forall x \in \Gamma$。那么罗宾边界值变分问题(8.149)存在唯一解 $u \in H^1(\Omega)$，满足*

$$\left\|u\right\|_{H^1(\Omega)} \leq c \left[\left\|f\right\|_{\widetilde{H}^{-1}(\Omega)} + \left\|g\right\|_{H^{-\frac{1}{2}}(\Gamma)}\right]. \tag{8.150}$$

证明. 根据(8.38)，定义 $H^1(\Omega)$ 中的等价范

$$\left\|\nu\right\|_{W^{1,2}(\Omega),\Gamma} := \left\{\left\|\gamma_0^{\text{int}} \nu\right\|_{L^2(\Gamma)}^2 + \left\|\nabla \nu\right\|_{L^2(\Omega)}^2\right\}^{\frac{1}{2}}$$

结合 $\kappa(x) \geq \kappa_0 > 0, \forall x \in \Gamma$，以及半椭圆特性 Lemma 8.15，有

$$\begin{aligned} a(\nu, \nu) + \int_\Gamma \kappa(x) \left[\gamma_0^{\text{int}} \nu(x)\right]^2 ds_x &\geq \lambda_0 \left\|\nabla \nu\right\|_{L^2(\Omega)}^2 + \kappa_0 \left\|\gamma_0^{\text{int}} \nu\right\|_{L^2(\Gamma)}^2 \\ &\geq \min\{\lambda_0, \kappa_0\} \left\|\nu\right\|_{H^1(\Omega),\Gamma}^2 \\ &\geq c_1^A \left\|\nu\right\|_{H^1(\Omega)}^2. \end{aligned}$$

进而，根据拉克斯一密格拉蒙定理 (Theorem 8.14)，可得罗宾边界值变分问题(8.149)的唯一解。由此可得(8.150)。 □

8.5　拉普拉斯算子的基本解

8.5.1　基本解

回顾一下偏微分方程(8.11)

$$(Lu)(x) = f(x), \quad x \in \Omega \subset \mathbb{R}^d,$$

其中 L 是椭圆线性二阶偏微分算子(8.1)

$$(L\,u)(x) = -\sum_{i,j=1}^d \frac{\partial}{\partial x_j}\left[a_{ji}(x)\frac{\partial}{\partial x_i} u(x)\right].$$

对应的内部共形导数(8.7)

$$\gamma_1^{int}u(x)=\sum_{i,j=1}^{d}n_j(x)a_{ji}(x)\frac{\partial}{\partial x_i}u(x),\quad x\in\Gamma.$$

由格林第二恒等式(8.8)可得，$\forall\, y\in\Omega\in\mathbb{R}^d$ 测试方程 $\nu(y)$ 和对应 y 的偏微分方程解 $u(y)$ 满足

$$\begin{aligned}&a(u,\nu)=a(\nu,u)\Rightarrow\\&\int_\Omega(L\nu)(y)u(y)\,dy=\int_\Gamma\gamma_1^{\text{int}}u(y)\gamma_0^{\text{int}}\nu(y)\,ds_y-\int_\Gamma\gamma_1^{\text{int}}\nu(y)\gamma_0^{\text{iunt}}u(y)\,ds_y+\int_\Gamma f(y)\nu(y)\,dy.\end{aligned}$$

将测试方程定义为 $\nu(y):=U^*(x,y)$，$u(x)$ 和 $u(y)$ 的关系可以表示如下

$$\int_\Omega\left(L_yU^*\right)(x,y)u(y)\,dy=u(x),\quad\forall\, x\in\Omega,$$

那么偏微分方程(8.11)的解 $u(x),x\in\Omega$，根据里兹表现定理8.13，可以改写为如下表现方程

$$u(x)=\int_\Omega U^*(x,y)f(y)\,dy+\int_\Gamma U^*(x,y)\gamma_1^{\text{int}}u(y)\,ds_y-\int_\Gamma U^*(x,y)\gamma_0^{\text{int}}u(y)\,ds_y.\tag{8.151}$$

由此可见，为了基于表现方程(8.151)求得任何偏微分形式方程(8.11)的解 $u(x)$，我们需要以下两方面的信息

- $x\in\Gamma$ 对应的柯西数列 $\left[\gamma_0^{\text{int}}u(x),\gamma_1^{\text{int}}u(x)\right]$，
- 连接 $u(x)$ 和 $u(y)$ 线性二次偏微分算子 $\left(L_yU^*\right)(x,y)$，对应(8.151)。

对于前者，关键在于构建合适的边界积分式 (boundary integral equations)以生成完整柯西数列 (complete cauchy data)，我们将在第8.6节讨论。

对于后者，从分布意义上来看，偏微分方程(8.11)的解 $u(x)$ 可以理解为

$$u(x)=\int_\Gamma\delta_0(y-x)u(y)\,dy,\quad x\in\Omega,$$

那么定义

$$\left(L_yU^*\right)(x,y)=\delta_0(y-x),\quad x,y\in\mathbb{R}^d,\tag{8.152}$$

我们将(8.152)中的解 $U^*(x,y)$ 称为基本解 (fundamental solution)。问题的关键就在于，利用(8.152)求得唯一的基本解 $U^*(x,y)$，进而带回到表现方程(8.151)中。

关于不同形式的偏微分算子 (尤其是存在分段常系数的偏微分算子) 是否存在基本解的问题，相关证明可见如Hörmander (1983a,b, 1994a,b)，我们不做过多讨论。对经济学应用研究来说，我们更多关注当偏微分算子是拉普拉斯算子 (Laplace operator)时对应的基本解。

8.5.2 拉普拉斯算子

拉普拉斯算子

$$(Lu)(x):=-\Delta u(x),\quad x\in\mathbb{R}^d,d=2,3.$$

对应的基本解 $U^*(x,y)$ 就是下述偏微分方程的分布解

$$-\Delta_yU^*(x,y)=\delta_0(y-x),\quad x,y\in\mathbb{R}^d.$$

由于拉普拉斯算子与旋转 (rotation) 和平移 (translation) 无关，可以定义 $z := y-x, U^*(x,y) = \nu(z)$。则 (8.152)变为

$$-\Delta\nu(z) = \delta_0(z), \quad z \in \mathbb{R}^d, \tag{8.153}$$

我们的任务是根据(8.153)求得基本解。

由 Definition 8.32的傅里叶变换(8.50)和(8.52b)我们有

$$\begin{aligned} \left|\xi\right|^2 \widehat{\nu}(\xi) &= (2\pi)^{-\frac{d}{2}}, \\ \hookrightarrow \widehat{\nu}(\xi) &= (2\pi)^{-\frac{d}{2}} \left|\xi\right|^{-\frac{1}{2}} \in \mathcal{S}'(\mathbb{R}^d), \end{aligned}$$

其中 $\mathcal{S}'(\mathbb{R}^d)$ 表示缓增分布空间(Definition 8.31)。

可见，一个缓增空间中的 $\nu(\xi) \in \mathcal{S}'(\mathbb{R}^d)$，和它所对应的傅里叶变换 $\widehat{\nu}(\xi)$ 的关系为

$$\langle \widehat{\nu}, \varphi \rangle_{L^2(\mathbb{R}^d)} = \langle \nu, \widehat{\varphi} \rangle_{L^2(\mathbb{R}^d)}, \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^d).$$

已知

$$\varphi(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left[i\langle z, \xi\rangle\right] \widehat{\varphi}(z)\, dz,$$

那么

$$\langle \widehat{\nu}, \varphi \rangle_{L^2(\mathbb{R}^d)} = (2\pi)^{-d} \int_{\mathbb{R}^d} \left|\xi\right|^{-\frac{1}{2}} \int_{\mathbb{R}^d} \exp\left[i\langle z, \xi\rangle\right] \widehat{\varphi}(z)\, dz d\xi.$$

上式的问题在于，积分 $\int_{\mathbb{R}^d} \left|\xi\right|^{-\frac{1}{2}} d\xi$ 不存在，因此无法将两个求积操作合并。一个解决方案是，利用

$$\Delta_z \exp\left(i\langle z, \xi\rangle\right) = -|\xi|^2 \exp\left(i\langle z, \xi\rangle\right),$$

首先根据 $\left|\xi\right|$ 的值作分步骤积分，进而调整两个积分操作的顺序，随后重复分步骤积分操作

$$\begin{aligned} &\langle \nu, \widehat{\varphi} \rangle_{L^2(\mathbb{R}^d)} \\ &= \langle \widehat{\nu}, \varphi \rangle_{L^2(\mathbb{R}^d)} = (2\pi)^{-d} \int_{\mathbb{R}^d} \left|\xi\right|^{-\frac{1}{2}} \int_{\mathbb{R}^d} \exp\left[i\langle z, \xi\rangle\right] \widehat{\varphi}(z)\, dz d\xi \\ &= (2\pi)^{-d} \int_{|\xi|\le 1} |\xi|^{-\frac{1}{2}} \int_{\mathbb{R}^d} \exp\left[i\langle z, \xi\rangle\right] \widehat{\varphi}(z) dz d\xi + (2\pi)^{-d} \int_{|\xi|>1} |\xi|^{-\frac{1}{2}} \int_{\mathbb{R}^d} \left[-\Delta_z \frac{\exp\left(i\langle z, \xi\rangle\right)}{|\xi|^2}\right] \widehat{\varphi}(z) dz d\xi \\ &= (2\pi)^{-d} \int_{|\xi|\le 1} |\xi|^{-\frac{1}{2}} \int_{\mathbb{R}^d} \exp\left[i\langle z, \xi\rangle\right] \widehat{\varphi}(z) dz d\xi + (2\pi)^{-d} \int_{|\xi|>1} |\xi|^{-\frac{1}{4}} \int_{\mathbb{R}^d} \exp\left(i\langle z, \xi\rangle\right) \left[-\Delta_z \widehat{\varphi}(z)\right] dz d\xi \\ &= (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{\varphi}(z) \int_{|\xi|\le 1} \frac{\exp\left[i\langle z, \xi\rangle\right]}{|\xi|^2} d\xi dz + (2\pi)^{-d} \int_{\mathbb{R}^d} \left[-\Delta_z \widehat{\varphi}(z)\right] \int_{|\xi|>1} \frac{\exp\left(i\langle z, \xi\rangle\right)}{|\xi|^4} d\xi dz \\ &= \int_{\mathbb{R}^d} \widehat{\varphi}(z) (2\pi)^{-d} \left[\int_{|\xi|\le 1} \frac{\exp\left[i\langle z, \xi\rangle\right]}{|\xi|^2} d\xi - \Delta_z \int_{|\xi|>1} \frac{\exp\left[i\langle z, \xi\rangle\right]}{|\xi|^4} d\xi\right] dz. \end{aligned} \tag{8.154}$$

8.5.3　三维空间的基本解

对于三维空间 $d = 3$ 的情况，我们可以建立三维坐标系

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} r\,\cos\varphi\,\sin\theta \\ r\,\sin\varphi\,\sin\theta \\ r\,\cos\theta \end{pmatrix}, \quad r \in (0,\infty), \varphi \in (0, 2\pi), \theta \in (0, \pi).$$

将傅里叶变换的旋转对称 (Lemma 8.2) 代回(8.154)有

$$\nu(z)=\nu(|z|)=(2\pi)^{-3}\left[\underbrace{\int_{|\xi|\le 1}\frac{\exp\left(i\langle z,\xi\rangle\right)}{|\xi|^2}d\xi}_{=:\mathcal{A}}-\underbrace{\Delta_z\int_{|\xi|>1}\frac{\exp\left(i\langle z,\xi\rangle\right)}{|\xi|^4}d\xi}_{=:\mathcal{B}}\right],$$

$$\begin{aligned}\mathcal{A}&=\int_0^{2\pi}\int_0^{\pi}\int_0^1\exp\left[i\,|z|\,r\cos\theta\right]\sin\theta\,dr\,d\theta\,d\varphi,\\ \mathcal{B}&=\Delta_z\int_0^{2\pi}\int_0^{\pi}\int_1^{\infty}\frac{\exp\left[i\,|z|\,r\cos\theta\right]}{r^2}\sin\theta\,dr\,d\theta\,d\varphi,\end{aligned}$$

$$\hookrightarrow\nu(z)=\nu(|z|)=(2\pi)^{-2}\left[\int_0^{\pi}\int_0^1\exp\left[i|z|r\cos\theta\right]\sin\theta drd\theta-\Delta_z\int_0^{\pi}\int_1^{\infty}\frac{\exp\left[i|z|r\cos\theta\right]}{r^2}\sin\theta drd\theta\right].$$

设 $\iota:=\cos\theta\in(-1,1)$，那么

$$\begin{aligned}&\int_0^{\pi}\exp\left[i|z|r\cos\theta\right]\sin\theta d\theta\\ &=\int_{-1}^{1}\exp\left[i|z|r\iota\right]d\iota\\ &=\frac{\exp\left[i|z|r\right]-\exp\left[-i|z|r\right]}{i|z|r}\\ &=\frac{2}{|z|r}\sin|z|r\end{aligned}$$

$$\hookrightarrow\nu(z)=(2\pi)^{-2}\left[\underbrace{\int_0^1\frac{\sin|z|r}{|z|r}dr}_{=:\mathcal{C}}-\underbrace{\Delta_z\int_1^{\infty}\frac{\sin|z|r}{|z|r^3}dr}_{=:\mathcal{D}}\right]$$

先来看 $\mathcal{C}$。设 $\zeta:=|z|r$ 我们有

$$\begin{aligned}\mathcal{C}&=\int_0^1\frac{\sin|z|r}{|z|r}dr\\ &=|z|^{-1}\int_0^{|z|}\frac{\sin\zeta}{\zeta}d\zeta\\ &=\frac{\mathrm{Si}(|z|)}{|z|},\end{aligned}$$

其中 Si 表示三角积分 (trigonometric integral)中正弦积分 (sine integral)的一种，定义为

$$\mathrm{Si}(x):=\int_0^x\frac{\sin t}{t}dt,$$

以及

$$\mathrm{sinc}(x):=\begin{cases}\frac{\sin x}{x}\\ \frac{\sin\pi x}{\pi x}\end{cases}$$

分别称非标准化 sinc 方程 (unnormalized sinc finction)和标准化 sinc 方程 (normalized sinc finction)。

再来看 $\mathcal{D}$。由于

$$\begin{aligned}\int \frac{\sin ax}{x^3}\,dx &= -\frac{1}{2}\frac{\sin ax}{x^2} + \frac{a}{2}\int \frac{\cos ax}{x^2}\,dx \\ &= -\frac{1}{2}\frac{\sin ax}{x^2} - \frac{a}{2}\frac{\cos ax}{x} - \frac{a^2}{2}\int \frac{\sin ax}{x}\,dx,\end{aligned}$$

$$\begin{aligned}\hookrightarrow \mathcal{D} &= \int_1^\infty \frac{\sin|z|r}{|z|r^3}\,dr \\ &= \left[-\frac{1}{2}\frac{\sin|z|r}{|z|r^2} - \frac{1}{2}\frac{\cos|z|r}{r}\right]_1^\infty - \frac{|z|}{2}\int_1^\infty \frac{\sin|z|r}{r}\,dr \\ &= \frac{1}{2}\frac{\sin|z|}{|z|} + \frac{1}{2}\cos|z| - \frac{|z|}{2}\left[\frac{\pi}{2} - \mathrm{Si}(|z|)\right].\end{aligned}$$

$$\begin{aligned}\hookrightarrow \nu(z) &= (2\pi)^{-2}\left[\mathcal{C} - \Delta_z\mathcal{D}\right] \\ &= (2\pi)^{-2}\left\{\frac{\mathrm{Si}(|z|)}{|z|} - \Delta_z\left[\frac{1}{2}\frac{\sin|z|}{|z|} + \frac{1}{2}\cos|z| - \frac{\pi}{4}|z| + \frac{|z|}{2}\mathrm{Si}(|z|)\right]\right\} \\ &= \frac{1}{8\pi}\Delta_z|z| + \frac{1}{2\pi^2}\underbrace{\left\{\frac{\mathrm{Si}(|z|)}{|z|} - \Delta_z\left[\frac{1}{2}\frac{\sin|z|}{|z|} + \frac{1}{2}\cos|z| + \frac{1}{2}|z|\,\mathrm{Si}(|z|)\right]\right\}}_{=0} \\ &= \frac{1}{4\pi}\frac{1}{|z|}.\end{aligned}$$

这样我们有三维空间中拉普拉斯算子的基本解为

$$U^*(x,y) = \nu(z) = \frac{1}{4\pi}\frac{1}{|x-y|}, \quad x,y \in \mathbb{R}^3. \tag{8.155}$$

8.5.4　二维空间的基本解

对于二维空间 $d = 2$ 的情况，首先要对基本解的逆傅里叶变换作一定的正则化处理 (Vladimirov, 1971)。定义一个缓增空间中的分布 $\mathcal{P}\frac{1}{|x|^2} \in \mathcal{S}'(\mathbb{R}^2)$，满足

$$\langle \mathcal{P}\frac{1}{|x|^2}, \varphi(x)\rangle_{L^2(\mathbb{R}^2)} = \int_{x\in\mathbb{R}^2:|x|\le 1} \frac{\varphi(x)-\varphi(0)}{|x|^2}\,dx + \int_{x\in\mathbb{R}^2:|x|\ge 1} \frac{\varphi(x)}{|x|^2}\,dx, \quad \forall\,\varphi \in \mathcal{S}(\mathbb{R}^2).$$

那么

$$\begin{aligned}2\pi\langle \nu, \widehat{\varphi}\rangle_{L^2(\mathbb{R}^2)} &= \langle \mathcal{P}\frac{1}{|\xi|^2}, \varphi\rangle_{L^2(\mathbb{R}^2)} \\ &= \int_{\xi\in\mathbb{R}^2:|\xi|\le 1} \frac{\varphi(\xi)-\varphi(0)}{|\xi|^2}\,d\xi + \int_{\xi\in\mathbb{R}^2:|\xi|\ge 1} \frac{\varphi(\xi)}{|\xi|^2}\,d\xi, \quad \forall\,\varphi \in \mathcal{S}(\mathbb{R}^2),\end{aligned}$$

其中

$$\begin{aligned}\varphi(\xi) &= \frac{1}{2\pi}\int_{\mathbb{R}^2} \exp\left(i\langle z,\xi\rangle\right)\widehat{\varphi}(z)\,dz, \\ \varphi(0) &= \frac{1}{2\pi}\int_{\mathbb{R}^2} \widehat{\varphi}(z)\,dz,\end{aligned}$$

$$\begin{aligned}&\hookrightarrow (2\pi)^2\,\langle \nu, \widehat{\varphi}\rangle_{L^2(\mathbb{R}^2)} \\ &= \int_{\xi\in\mathbb{R}^2:|\xi|\le 1} \frac{1}{|\xi|^2}\int_{\mathbb{R}^2}\left[\exp\left(i\langle z,\xi\rangle\right) - 1\right]\widehat{\varphi}(z)dzd\xi + \int_{\xi\in\mathbb{R}^2:|\xi|>1} \frac{1}{|\xi|^2}\int_{\mathbb{R}^2} \exp\left(i\langle z,\xi\rangle\right)\widehat{\varphi}(z)dzd\xi.\end{aligned}$$

同样地，我们无法调整两个积分的先后顺序。采用类似于(8.154)的思路，根据 $|\xi|$ 作分步骤积分，进而调整两个积分操作的顺序，然后重复分步骤积分运算，得

$$(2\pi)^2 \langle \nu, \widehat{\varphi} \rangle_{L^2(\mathbb{R}^2)} = \int_{\mathbb{R}^2} \widehat{\varphi}(z) \left[\int_{\xi\in\mathbb{R}^2:|\xi|\le 1} \frac{\exp\left[i\langle z,\xi\rangle\right]-1}{|\xi|^2} d\xi + \int_{\xi\in\mathbb{R}^2:|\xi|>1} \frac{\exp\left[i\langle z,\xi\rangle\right]}{|\xi|^2} d\xi \right] dz.$$

根据傅里叶变换的旋转对称 (Lemma 8.2)，由上式可得

$$\nu(z) = \nu(|z|) = (2\pi)^{-2} \left[\int_{\xi\in\mathbb{R}^2:|\xi|\le 1} \frac{\exp\left[i\langle z,\xi\rangle\right]-1}{|\xi|^2} d\xi + \int_{\xi\in\mathbb{R}^2:|\xi|>1} \frac{\exp\left[i\langle z,\xi\rangle\right]}{|\xi|^2} d\xi \right].$$

建立二维坐标系

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}, \quad r\in(0,\infty), \varphi = (0,2\pi),$$

上式变为

$$\begin{aligned}
\nu(z) = \nu(|z|) &= (2\pi)^{-2} \left\{ \int_0^1 \int_0^{2\pi} \frac{1}{r} \left[\exp\left(ir|z|\cos\varphi\right)-1\right] d\varphi dr + \int_1^\infty \int_0^{2\pi} \frac{1}{r} \exp\left(ir|z|\cos\varphi\right) d\varphi dr \right\} \\
&= (2\pi)^{-1} \left\{ \int_0^1 \frac{1}{r} \left[J_0(r|z|)-1\right] dr + \int_1^\infty \frac{1}{r} \left[J_0(r|z|)-1\right] dr \right\},
\end{aligned}$$

其中我们用到了一阶贝塞尔方程 (Bessel function, Definition 8.34, 或参考 (Gradshteyn and Ryzhik, 2014, Sec. 8.411))

$$J_0(s) := (2\pi)^{-1} \int_0^{2\pi} \exp\left(is\cos\varphi\right) d\varphi.$$

定义 $r := \frac{s}{\varrho}$，上式进一步变为

$$\begin{aligned}
\nu(z) &= \frac{1}{2\pi} \int_0^\varrho \frac{J_0(s)-1}{s} ds + \frac{1}{2\pi} \int_\varrho^\infty \frac{J_0(s)}{s} ds \\
&= \frac{1}{2\pi} \int_0^1 \frac{J_0(s)-1}{s} ds + \frac{1}{2\pi} \int_1^\infty \frac{J_0(s)}{s} ds + \frac{1}{2\pi} \int_\varrho^1 \frac{1}{s} ds \\
&= -\frac{1}{2\pi} \log|z| - \frac{c_0}{2\pi},
\end{aligned}$$

其中 c_0 是个常数

$$c_0 := \int_0^1 \frac{1-J_0(s)}{s} ds - \int_1^\infty \frac{J_0(s)}{s} ds.$$

由于任何常数都满足齐次拉普拉斯方程的条件，在计算二维空间拉普拉斯方程的基本解时，我们可以忽略常数项。因此基本解变为

$$U^*(x,y) = \nu(z) = -\frac{1}{2\pi} \log|x-y|, \quad x,y\in\mathbb{R}^2. \tag{8.156}$$

8.5.5 基本解总结

小结：拉普拉斯算子在二维、三维空间中的基本解(8.155)，(8.156)为

$$U^*(x,y) = \begin{cases} -\frac{1}{2\pi}\log|x-y| & d=2, \\ \frac{1}{4\pi}\frac{1}{|x-y|} & d=3. \end{cases} \tag{8.157}$$

对应地，对于给定的 $x\in\Omega$，下述形式偏微分方程

$$-\Delta u(x) = f(x), \quad x\in\Omega\subset\mathbb{R}^d$$

的解 $u(x)$，都以表现方程的形式出现

$$u(x) = \int_\Omega U^*(x,y) f(y)\, dy + \int_\Gamma U^*(x,y) \frac{\partial}{\partial n_y} u(y)\, ds_y - \int_\Gamma \frac{\partial}{\partial n_y} U^*(x,y) u(y)\, ds_y. \tag{8.158}$$

8.6　边界积分算子

回顾一下第8.5节提出的问题，我们的目标是求解 $d=2,3$ 下的泊松方程 (Poisson equation)

$$-\Delta u(x)=f(x),\quad x\in\Omega\subset\mathbb{R}^d,$$

$f(x)=0$ 时的齐次形式泊松方程又称拉普拉斯方程 (Laplace function)。

泊松方程的解满足表现形式 (representation formula) (8.151)

$$u(x)=\int_\Omega U^*(x,y)f(y)\,dy+\int_\Gamma U^*(x,y)\gamma_1^{\mathrm{int}}u(y)\,ds_y-\int_\Gamma U^*(x,y)\gamma_0^{\mathrm{int}}u(y)\,ds_y,\tag{8.159}$$

为了求解(8.159)，首先要求得拉普拉斯算子的基本解 $U^*(x,y)$，见(8.157)给出

$$U^*(x,y)=\begin{cases}-\frac{1}{2\pi}\log|x-y| & d=2,\\ \frac{1}{4\pi}\frac{1}{|x-y|} & d=3.\end{cases}$$

在此基础上，需要构建适宜的边界积分方程，以 $x\in\Gamma$ 生成完整的柯西数 $\left[\gamma_0^{\mathrm{int}}u(x),\gamma_1^{\mathrm{int}}u(x)\right]$。而这就需要我们探讨表面积位势 (surface potential) 和体积位势 (volume potental) 的映射特征。

8.6.1　牛顿位势

泊松方程表现式(8.159)中，对于某一给定的方程 $f(y),y\in\Omega$，我们可以定义 $f(y)$ 的体积位势，或称牛顿位势算子 (Newton potential)为 $\widetilde{N}_0f$，满足

$$\left(\widetilde{N}_0f\right)(x):=\int_\Omega U^*(x,y)f(y)\,dy,\quad x\in\mathbb{R}^d.\tag{8.160}$$

由内积形式可得

$$\langle\widetilde{N}_0\varphi,\psi\rangle_\Omega=\int_\Omega\psi(x)\int_\Omega U^*(x,y)\varphi(y)dydx=\langle\varphi,\widetilde{N}_0\psi\rangle_\Omega,\quad \varphi,\psi\in\mathcal{S}(\mathbb{R}^d),$$

可见 $\widetilde{N}_0\varphi\in\mathcal{S}(\mathbb{R}^d)$。

进而，牛顿位势算子 $\widetilde{N}_0:\mathcal{S}'(\mathbb{R}^d)\mapsto\mathcal{S}'(\mathbb{R}^d)$ 可以定义为

$$\langle\widetilde{N}_0f,\psi\rangle_\Omega:=\langle f,\widetilde{N}_0\psi\rangle_\Omega,\quad\forall\,\psi\in\mathcal{S}(\mathbb{R}^d).$$

Theorem 8.31 (牛顿位势算子的映射)**.** *牛顿位势算子 $\widetilde{N}_0:\widetilde{H}^{-1}(\Omega)\mapsto H^1(\Omega)$ 定义了一个连续映射*

$$\left\|\widetilde{N}_0f\right\|_{H^1(\Omega)}\le c\left\|f\right\|_{\widetilde{H}^{-1}(\Omega)}.\tag{8.161}$$

证明. 对于 $\varphi\in C_0^\infty(\Omega)$ 我们可得

$$\left\|\varphi\right\|^2_{H^{-1}(\mathbb{R}^d)}=\int_{\mathbb{R}^d}\frac{|\widehat{\varphi}(\xi)|^2}{1+|\xi|^2}d\xi,$$

其中 $\widehat{\varphi}(\xi)$ 表示傅里叶变换

$$\widehat{\varphi}(\xi)=(2\pi)^{-\frac{d}{2}}\int_{\mathbb{R}^d}\exp\left[-i\langle x,\xi\rangle\right]\varphi(x)\,dx.$$

由 $\operatorname{supp}\varphi\subset\Omega$ 可得 (支撑集 supp 的定义，见242页脚注)

$$\begin{aligned}\left\|\varphi\right\|_{H^{-1}(\mathbb{R}^d)}&=\sup_{0\neq\nu\in H^1(\mathbb{R}^d)}\frac{\langle\varphi,\nu\rangle_{L^2(\mathbb{R}^d)}}{\left\|\nu\right\|_{H^1(\mathbb{R}^d)}}\\&\le\sup_{0\neq\nu\in H^1(\Omega)}\frac{\langle\varphi,\nu\rangle_{L^2(\Omega)}}{\left\|\nu\right\|_{H^1(\Omega)}}\\&=\left\|\varphi\right\|_{\widetilde{H}^{-1}(\Omega)}.\end{aligned}\tag{8.162}$$

由定义(8.160)，我们定义一个 $u(x)$

$$u(x) \coloneqq \left(\widetilde{N}_0\varphi\right)(x) = \int_\Omega U^*(x,y)\varphi(y)\,dy, \quad x \in \mathbb{R}^d. \tag{8.163}$$

设 $\Omega \subset B_R(0)$，以及一个有紧支撑的非负单调递增 cutoff 方程 $\mu \in C_0^\infty([0,\infty))$，满足 $\mu(r) = 1, r \in [0,2R]$。进而定义

$$u_\mu(x) \coloneqq \int_\Omega \mu\left(|x-y|\right) U^*(x,y)\varphi(y)\,dy, \quad x \in \mathbb{R}^d. \tag{8.164}$$

$$\begin{aligned}
&x, y \in \Omega,\\
\hookrightarrow &|x-y| \geq 0,\\
\hookrightarrow &\mu(|x-y|) = 1,\\
\hookrightarrow &u_\mu(x) = u(x), \quad x \in \Omega,\\
\hookrightarrow &\|u\|_{H^1(\Omega)} = \|u_\mu\|_{H^1(\Omega)} \leq \|u_\mu\|_{H^1(\mathbb{R}^d)}
\end{aligned}$$

以及

$$\|u_\mu\|^2_{H^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \left(1+|\xi|^2\right) |\widehat{u}_\mu(\xi)|^2\,d\xi. \tag{8.165}$$

现在来计算 $u_\mu(x)$ 的傅里叶变换

$$\begin{aligned}
\widehat{u}_\mu(x) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left[-i\langle x,\xi\rangle\right] u_\mu(x)\,dx\\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left[-i\langle x,\xi\rangle\right] \int_{\mathbb{R}^d} \mu(|x-y|)U^*(x,y)\varphi(y)dydx\\
&= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \exp\left[-i\langle z+y,\xi\rangle\right] \mu(|z|)U^*(z+y,y)\varphi(y)dydz\\
&= \underbrace{(2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left[-i\langle y,\xi\rangle\right] \varphi(y)dy}_{=\widehat{\varphi}(\xi)} \underbrace{\int_{\mathbb{R}^d} \exp\left[-i\langle z,\xi\rangle\right] \mu(|z|)U^*(z,0)dz}_{=:I(|\xi|)}.
\end{aligned} \tag{8.166}$$

求解(8.166)需要进一步求 $I(|\xi|)$ 的值。已知 $u(|z|)$ 和 $U^*(z,0)$ 都是只与 $|z|$ 有关的方程，我们可以利用傅里叶变换的旋转对称 Lemma 8.2，在三维坐标$\xi = (0,0,|\xi|)^\top$ 中测算 $I(|\xi|)$。以 $d=3$ 为例，建立坐标系

$$z = \begin{pmatrix} z_1\\ z_2\\ z_3 \end{pmatrix} = \begin{pmatrix} r\cos\phi\sin\theta\\ r\sin\phi\sin\theta\\ r\cos\theta, \end{pmatrix} \quad r \in [0,\infty), \phi \in [0,2\pi), \theta \in [0,\pi).$$

将(8.156)代入 $I(|\xi|)$

$$\begin{aligned}
I(|\xi|) &= \frac{1}{4\pi}\int_{\mathbb{R}^d} \exp\left[-i\langle z,\xi\rangle\right] \frac{\mu(|z|)}{|z|}dz\\
&= \frac{1}{4\pi}\int_0^\infty \int_0^{2\pi} \int_0^\pi \exp\left[-i|\xi|r\cos\theta\right] \frac{\mu(r)}{r} r^2 \sin\theta d\theta d\phi dr\\
&= \frac{1}{2}\int_0^\infty r\mu(r)\,dr \underbrace{\int_0^\pi \exp\left[-ir|\xi|\cos\theta\right]\sin\theta\,d\theta}\,dr.
\end{aligned}$$

定义 $\iota := \cos\theta$，我们有

$$\begin{aligned}\int_0^\pi \exp\left[-ir|\xi|\cos\theta\right]\sin\theta\,d\theta &= \int_{-1}^1 \exp\left[-ir|\xi|\iota\right]d\iota \\ &= \left[-\frac{1}{ir|\xi|}\exp\left[-ir|\xi|\iota\right]\right]_{-1}^1 = \frac{2\sin r|\xi|}{r|\xi|}.\end{aligned}$$

由此可得

$$I(|\xi|) = \frac{1}{|\xi|^2}\int_0^\infty \mu(r)\sin r|\xi|\,dr.$$

下面根据 $|\xi|$ 的值，作分部求积。

1. 来看 $|\xi| > 1$ 的情况。定义 $s := r|\xi|$ 可得

$$I(|\xi|) = \frac{1}{|\xi|^2}\int_0^\infty \mu\left(\frac{s}{|\xi|}\right)\sin s\,ds.$$

根据定义可知 $0 \le \mu(r) \le 1$ 且有紧支撑，则

$$I(|\xi|) \le c_1(R)\frac{1}{|\xi|^2}, \quad |\xi| \ge 1. \tag{8.167}$$

此外考虑到

$$\left(1+|\xi|^2\right)^2 \le 4|\xi|^4, \quad |\xi| \ge 1,$$

结合(8.165)，(8.166)，(8.167)可得

$$\begin{aligned}\left\|u_\mu\right\|^2_{H^1(\mathbb{R}^d),|\xi|>1} &= \int_{|\xi|>1}\left(1+|\xi|^2\right)\,|\widehat{u}_\mu(\xi)|^2\,d\xi \\ &= \int_{|\xi|>1}\left(1+|\xi|^2\right)\,\left|\widehat{\varphi}(\xi)I(|\xi|)\right|^2 d\xi \\ &\le \left[c_1(R)\right]^2\int_{|\xi|>1}\frac{1+|\xi|^2}{|\xi|^4}\left|\widehat{\varphi}(\xi)\right|^2 d\xi \\ &\le 4\left[c_1(R)\right]^2\int_{|\xi|>1}\frac{1}{1+|\xi|^2}\left|\widehat{\varphi}(\xi)\right|^2 d\xi.\end{aligned} \tag{8.168}$$

2. 来看 $|\xi| \le 1$ 的情况。

$$\begin{aligned}I(|\xi|) &= \int_0^\infty \mu(r)\frac{\sin r|\xi|}{|\xi|}\,dr \\ &\le c_2(R), \quad |\xi| \le 1.\end{aligned} \tag{8.169}$$

结合(8.165)，(8.166)，(8.169)可得

$$\begin{aligned}\left\|u_\mu\right\|^2_{H^1(\mathbb{R}^d),|\xi|\le 1} &= \int_{|\xi|\le 1}\left(1+|\xi|^2\right)\,|\widehat{u}_\mu(\xi)|^2\,d\xi \\ &= \int_{|\xi|\le 1}\left(1+|\xi|^2\right)\,\left|\widehat{\varphi}(\xi)I(|\xi|)\right|^2 d\xi \\ &\le 2\left[c_2(R)\right]^2\int_{|\xi|\le 1}\left|\widehat{\varphi}(\xi)\right|^2 d\xi \\ &\le 4\left[c_2(R)\right]^2\int_{|\xi|\le 1}\frac{1}{1+|\xi|^2}\left|\widehat{\varphi}(\xi)\right|^2 d\xi.\end{aligned} \tag{8.170}$$

将(8.168)，(8.170)汇总，代回(8.165)，(8.166)可得

$$\begin{aligned}\|u_\mu\|^2_{H^1(\mathbb{R}^d)} &= \int_{\xi\in\mathbb{R}^d}(1+|\xi|^2)\left|\widehat{u}_\mu(\xi)\right|^2 d\xi \\ &\le c\int_{\xi\in\mathbb{R}^d}\frac{1}{1+|\xi|^2}\left|\widehat{\varphi}(\xi)\right|^2 d\xi \\ &= c\|\varphi\|^2_{H^{-1}(\mathbb{R}^d)}.\end{aligned} \tag{8.171}$$

将(8.164)，(8.162)代入上式，有

$$\begin{aligned}\|\widetilde{N}_0\varphi\|^2_{H^1(\mathbb{R}^d)} = \|u_\mu\|^2_{H^1(\mathbb{R}^d)} &\le c\|\varphi\|^2_{H^{-1}(\mathbb{R}^d)} \\ &\le c\|\varphi\|^2_{\widetilde{H}^{-1}(\Omega)}.\end{aligned} \tag{8.172}$$

因此

$$\frac{|\langle\widetilde{N}_0 f,\varphi\rangle_\Omega|}{\|\varphi\|_{\widetilde{H}^{-1}(\Omega)}} = \frac{|\langle f,\widetilde{N}_0\varphi\rangle_\Omega|}{\|\varphi\|_{\widetilde{H}^{-1}(\Omega)}} \le \frac{\|f\|_{\widetilde{H}^{-1}(\Omega)}\|\widetilde{N}_0\varphi\|_{H^1(\Omega)}}{\|\varphi\|_{\widetilde{H}^{-1}(\Omega)}} \le c\,\|f\|_{\widetilde{H}^{-1}(\Omega)}, \quad \forall\varphi\in C_0^\infty(\Omega).$$

对 $\|\cdot\|_{\widetilde{H}^{-1}(\Omega)}$ 取闭包，由对偶配对可得(8.161)。 □

Theorem 8.32 ($\mathbb{R}^d$ 中的牛顿位势通解)**.** 牛顿位势 $\widetilde{N}_0\widetilde{f}$ 是下述偏微分方程的通解

$$-\Delta_x\left(\widetilde{N}_0\widetilde{f}\right) = \widetilde{f}(x) = \begin{cases} f(x) & x\in\Omega, \\ 0 & x\in\mathbb{R}^d\backslash\overline{\Omega}. \end{cases} \tag{8.173}$$

证明. 对于 $\varphi\in C_0^\infty(\mathbb{R}^d)$，依次作分段积分，改变积分符顺序，使用基本解的对称特性，可得

$$\begin{aligned}\int_{\mathbb{R}^d}\left[-\Delta_x\left(\widetilde{N}_0\widetilde{f}\right)(x)\right]\varphi(x)\,dx &= \int_{\mathbb{R}^d}\underbrace{\left(\widetilde{N}_0\widetilde{f}\right)(x)}_{(8.160)}\left[-\Delta_x\varphi(x)\right]dx \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}U^*(x,y)\widetilde{f}(y)\,dy\left[-\Delta_x\varphi(x)\right]dx \\ &= \int_{\mathbb{R}^d}\widetilde{f}(y)\int_{\mathbb{R}^d}U^*(x,y)\left[-\Delta_x\varphi(x)\right]dxdy \\ &= \int_{\mathbb{R}^d}\widetilde{f}(y)\int_{\mathbb{R}^d}\underbrace{\left[-\Delta_x U^*(x,y)\right]}_{=:\delta_0(x-y)}\varphi(x)\,dxdy \\ &= \int_{\mathbb{R}^d}\widetilde{f}(y)\underbrace{\int_{\mathbb{R}^d}\delta_0(x-y)\varphi(x)\,dx}\,dy \\ &= \int_{\mathbb{R}^d}\widetilde{f}(y)\varphi(y)\,dy.\end{aligned}$$

对 $\|\cdot\|_{H^1(\mathbb{R}^d)}$ 取 $C_0^\infty(\mathbb{R}^d)$ 的闭包，再使用配偶配对，可得(8.173)成立。 □

Corollary 8.6 (Ω 中的牛顿位势通解)**.** 将 *Theorem 8.32*中的牛顿位势通解进一步限定到有界域 $\Omega\subset\mathbb{R}^d$ 的情况。

此时位势方程 $\widetilde{N}_0 f$ 是以下偏微分方程的通解

$$-\Delta_x\widetilde{N}_0 f(x) = f, \quad x\in\Omega.$$

内界迹算子 $\gamma_0^{int}(\widetilde{N}_0 f)$ 可以表示为

$$\gamma_0^{int}(\widetilde{N}_0 f)(x) = \lim_{\Omega \ni \widetilde{x} \mapsto x \in \Gamma} \left(\widetilde{N}_0 f\right)(\widetilde{x}). \tag{8.174}$$

内界迹算子 $\gamma_0^{int}(\widetilde{N}_0 f)$ 定义了一个有界线性算子 $N_0 f$

$$N_0 = \gamma_0^{int}\widetilde{N}_0 : \widetilde{H}^{-1}(\Omega) \mapsto H^{\frac{1}{2}}(\Gamma), \tag{8.175}$$

$N_0 f$ 的范数满足

$$\|N_0 f\|_{H^{\frac{1}{2}}(\Gamma)} \le c_2^N \|f\|_{\widetilde{H}^{-1}(\Omega)} \quad \forall f \in \widetilde{H}^{-1}(\Omega). \tag{8.176}$$

Lemma 8.18 (牛顿位势算子方程的表现式)**.** 设给定 $f \in L^\infty$。那么牛顿位势表示如下等式，又称弱奇异表面积分 *(weakly singular surface integral)*

$$(N_0 f)(x) = \gamma_0^{int}\left(\widetilde{N}_0 f\right)(x) = \int_\Omega U^*(x,y) f(y)\, dy, \quad x \in \Gamma.$$

证明. 关于奇异积分、弱奇异积分，可参考Neri (1971); Vainikko (1993)。关于表面积分，可参考(Callahan, 2010, Ch.10)。

对于给定的 $\widetilde{x} \in \Omega, x \in \Gamma$，设 $\exists \varepsilon > 0$，满足 $|\widetilde{x} - x| < \varepsilon$。则

$$\begin{aligned}
&\left|\int_\Omega U^*(\widetilde{x},y) f(y)\, dy - \int_{y\in\Omega:|y-x|>\varepsilon} U^*(x,y) f(y)\, dy\right| \\
&\le \underbrace{\left|\int_{y\in\Omega:|y-x|>\varepsilon} \left[U^*(\widetilde{x},y) - U^*(x,y)\right] f(y) dy\right|}_{=:\mathcal{A}} + \underbrace{\left|\int_{y\in\Omega:|y-x|\le\varepsilon} U^*(\widetilde{x},y) f(y) dy\right|}_{=:\mathcal{B}},
\end{aligned} \tag{8.177}$$

其中

1. $\mathcal{A} \Rightarrow$

$$\lim_{\Omega \ni \widetilde{x} \mapsto x \in \Gamma} \left|\int_{y\in\Omega:|y-x|>\varepsilon} \left[U^*(\widetilde{x},y) - U^*(x,y)\right] f(y) dy\right| = 0.$$

2. $\mathcal{B} \Rightarrow$

$$\begin{aligned}
\left|\int_{y\in\Omega:|y-x|\le\varepsilon} U^*(\widetilde{x},y) f(y) dy\right| &\le \|f\|_{L^\infty(\Omega\cap B_\varepsilon(x))} \int_{\Omega\cap B_\varepsilon(x)} |U^*(\widetilde{x},y)|\, dy \\
&\le \|f\|_{L^\infty(\Omega)} \underbrace{\int_{B_\varepsilon(x)} |U^*(\widetilde{x},y)|\, dy}_{=:\mathcal{C}},
\end{aligned}$$

3. 分 $d = 2, 3$ 两种情况来讨论 $\mathcal{C}$ 的值

   (a) $d = 2 \Rightarrow$

$$\begin{aligned}
\mathcal{C} := \int_{B_\varepsilon(x)} |U^*(\widetilde{x},y)|\, dy &= \frac{1}{2\pi} \int_{|y-\widetilde{x}|<2\varepsilon} |\log|y-\widetilde{x}||\, dy \\
&= \frac{1}{2} \int_0^{2\pi} \int_0^{2\varepsilon} |\log r|\, r dr d\varphi \\
&= \varepsilon^2 \left[1 - 2\log(2\varepsilon)\right].
\end{aligned} \tag{8.178}$$

(b) $d=3 \Rightarrow$

$$\begin{aligned}\mathcal{C} := \int_{B_\varepsilon(x)} |U^*(\widetilde{x},y)|\, dy &= \frac{1}{4\pi}\int_{|y-\widetilde{x}|<2\varepsilon} \frac{1}{|y-\widetilde{x}|} dy \\ &= \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\varepsilon} \frac{1}{r} r^2 \sin\psi dr d\psi d\varphi \\ &= 2\varepsilon^2.\end{aligned} \tag{8.179}$$

取极限 $\widetilde{x} \to x, \varepsilon \to 0$，可证得。 □

Lemma 8.19. 算子

$$N_1 := \gamma_1^{int}\widetilde{N}_0 : \widetilde{H}^{-1}(\Omega) \mapsto H^{-\frac{1}{2}}(\Gamma)$$

是有界算子，即是说，满足不等式条件

$$\|N_1 f\|_{H^{-\frac{1}{2}}(\Gamma)} = \left\|\gamma_1^{int}\widetilde{N}_0 f\right\|_{H^{-\frac{1}{2}}(\Gamma)} \le c\, \|f\|_{\widetilde{H}^{-1}(\Omega)}.$$

证明. 偏微分方程

$$-\Delta u(x) = f(x), \quad x \in \Omega$$

的通解是 $u = \widetilde{N}_0 f \in H^1(\Omega)$。

对于任意给定的 $w \in H^{\frac{1}{2}}(\Gamma)$，构建一个有界延拓 $\varepsilon w \in H^1(\Omega)$，满足

$$\|\varepsilon w\|_{H^1(\Omega)} \le c_{\mathrm{IT}} \|w\|_{H^{\frac{1}{2}}(\Gamma)}.$$

□

由格林第一恒等式(8.6)可得

$$\begin{aligned}\langle \gamma_1^{\mathrm{int}} u, w\rangle_\Gamma &= \langle u, \varepsilon_w w\rangle \\ &= \int_\Omega \nabla u(x) \nabla \varepsilon_w(x)\, dx - \langle f, \varepsilon_w\rangle_\Omega.\end{aligned}$$

代入牛顿位势算子的映射(8.161)(Theorem 8.31) 中，我们有

$$\begin{aligned}\left|\langle \gamma_1^{\mathrm{int}} u, w\rangle_\Gamma\right| &\le \left\{\|u\|_{H^1(\Omega)} + \|f\|_{\widetilde{H}^{-1}(\Omega)}\right\} \|\varepsilon_w\|_{H^1(\Omega)} \\ &\le (c+1) c_{\mathrm{IT}} \|f\|_{\widetilde{H}^{-1}(\Omega)} \|w\|_{H^{-\frac{1}{2}}(\Gamma)}.\end{aligned}$$

8.6.2 单层位势

设 $w \in H^{-\frac{1}{2}}(\Gamma)$ 为一个给定的密度方程。考虑如下单层位势算子 (single layer potential)

$$u(x) := \left(\widetilde{V} w\right)(x) := \int_\Gamma U^*(x,y) w(y)\, ds_y, \quad x \in \Omega \cup \Omega^c, \forall\, y \in \Gamma. \tag{8.180}$$

则基于给定的密度 $w \in H^{-\frac{1}{2}}(\Gamma)$，单层位势算子 $\left(\widetilde{V} w\right)(x) \in H^1(\Omega)$ 就是齐次偏微分方程的解，如下

Lemma 8.20. (8.180)定义的单层位势算子 $u(x) = \left(\widetilde{V} w\right)(x)$，构成如下齐次偏微分方程的解

$$-\Delta u(x) = 0, \quad x \in \Omega \cup \Omega^c.$$

对于给定的密度方程 $w \in H^{-\frac{1}{2}}(\Gamma)$，我们有对应的 $u \in H^1(\Omega)$，满足

$$\|u\|_{H^1(\Omega)} = \left\|\widetilde{V} w\right\|_{H^1(\Omega)} \le c\, \|w\|_{H^{-\frac{1}{2}}(\Gamma)}. \tag{8.181}$$

证明. 对于 $x \in \Omega \cup \Omega^c$ 和 $y \in \Gamma$，可得齐次偏微分方程的基本解 $U^*(x,y) \in C^\infty$。因此，我们可以调整差分和积分的顺序

$$
\begin{aligned}
-\Delta_x u(x) &= -\Delta_x \int_\Omega U^*(x,y) f(y)\, dy \\
&= \int_\Omega \left[-\Delta_x U^*(x,y)\right] f(y)\, dy = 0.
\end{aligned}
$$

在此基础上，对于 $\varphi \in C^\infty(\Omega)$，可得

$$
\begin{aligned}
\int_\Omega u(x)\varphi(x)\, dx &= \int_\Omega \underbrace{\int_\Gamma U^*(x,y) w(y) ds_y}_{:=u(x)} \varphi(x) dx \\
&= \int_\Gamma w(y) \underbrace{\int_\Omega U^*(x,y)\varphi(x)\, dx}_{=:(N_0\varphi)(y)}\, ds_y, \\
&= \int_\Gamma w(y) \left(N_0\varphi\right)(y)\, ds_y
\end{aligned}
$$

其中定义 $\left(N_0\varphi\right)(y) ds_y$ 为

$$
\left(N_0\varphi\right)(y) ds_y := \gamma_0^{\text{int}} \int_\Omega U^*(x,y)\varphi(x)\, dx, \quad x \in \Gamma.
$$

代入牛顿位势在 Ω 中的通解(8.176)(Corollary 8.6)，上式进一步变为

$$
\begin{aligned}
\int_\Omega u(x)\varphi(x)\, dx &= \int_\Gamma w(y) \left(N_0\varphi\right)(y)\, ds_y \\
&\le \|w\|_{H^{-\frac{1}{2}}(\Gamma)} \|N_0\varphi\|_{H^{\frac{1}{2}}(\Gamma)} \\
&\le c_2^N \|w\|_{H^{-\frac{1}{2}}(\Gamma)} \|\varphi\|_{\widetilde{H}^{-1}(\Omega)}
\end{aligned}
$$

对 $\|\cdot\|_{\widetilde{H}^{-1}(\Omega)}$ 取 $c^\infty(\Omega)$ 的闭包，再使用配偶配对，可证(8.181)成立。 □

单层位势算子(8.180)定义了一个有界的线性映射 $\widetilde{V} : H^{-\frac{1}{2}}(\Gamma) \mapsto H^1(\Omega)$。因此，内界迹算子 $\widetilde{V}w \in H^{-1}(\Omega)$ 定义良好。对应地，我们也可以写出线性算子 $V = \gamma_0^{\text{int}}\widetilde{V} : H^{-\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$，其中 V 满足

$$
\|Vw\|_{H^{\frac{1}{2}}(\Gamma)} \le c_2^V \|w\|_{W^{-\frac{1}{2}}(\Gamma)}, \quad \forall\, w \in H^{-\frac{1}{2}(\Gamma)}. \tag{8.182}
$$

Lemma 8.21 (单层位势算子方程的表现式)**.** *设给定 $w \in L^\infty(\Gamma)$，则单层位势算子方程的表现式 (representation formula) 是一个弱奇异表面积分*

$$
\left(Vw\right)(x) = \gamma_0^{int}\left(\widetilde{V}w\right)(x) = \int_\Gamma U^*(x,y) w(y)\, ds_y, \quad x \in \Gamma. \tag{8.183}
$$

证明. 对于给定的 $\widetilde{x} \in \Omega, x \in \Gamma$，假定 $\exists \varepsilon > 0$ 满足 $|x - \widetilde{x}| < \varepsilon$，那么

$$
\begin{aligned}
&\left| \int_\Gamma U^*(\widetilde{x},y) w(y)\, ds_y - \int_{y\in\Gamma:|y-x|>\varepsilon} U^*(x,y) w(y)\, ds_y \right| \\
&\le \underbrace{\left| \int_{y\in\Gamma:|y-x|>\varepsilon} \left[U^*(\widetilde{x},y) - U^*(x,y)\right] w(y) ds_y \right|}_{=:\mathcal{A}} + \underbrace{\left| \int_{y\in\Gamma:|y-x|\le\varepsilon} U^*(\widetilde{x},y) w(y) ds_y \right|}_{=:\mathcal{B}},
\end{aligned} \tag{8.184}
$$

其中

1. $\mathcal{A} \Rightarrow$

$$\lim_{\Omega \ni \widetilde{x} \mapsto x \in \Gamma} \left| \int_{y \in \Gamma : |y-x| > \varepsilon} \left[U^*(\widetilde{x}, y) - U^*(x, y) \right] w(y) ds_y \right| = 0.$$

2. $\mathcal{B} \Rightarrow$

$$\begin{aligned} \left| \int_{y \in \Gamma : |y-x| \le \varepsilon} U^*(\widetilde{x}, y) w(y) ds_y \right| &\le \|w\|_{L^\infty(\Gamma \cap B_\varepsilon(x))} \int_{\Gamma \cap B_\varepsilon(x)} |U^*(\widetilde{x}, y)| \, ds_y \\ &\le \|w\|_{L^\infty(\Gamma)} \underbrace{\int_{B_\varepsilon(x)} |U^*(\widetilde{x}, y)| \, ds_y}_{=:\mathcal{C}}, \end{aligned}$$

3. 类似地，分 $d = 2, 3$ 两种情况来讨论 $\mathcal{C}$ 的值

   (a) $d = 2 \Rightarrow$(8.178)

$$\begin{aligned} \mathcal{C} := \int_{B_\varepsilon(x)} |U^*(\widetilde{x}, y)| \, dy &= \frac{1}{2\pi} \int_{|y-\widetilde{x}| < 2\varepsilon} |\log |y - \widetilde{x}|| \, dy \\ &= \frac{1}{2} \int_0^{2\pi} \int_0^{2\varepsilon} |\log r| \, r dr d\varphi \\ &= \varepsilon^2 \left[1 - 2 \log (2\varepsilon) \right]. \end{aligned}$$

   (b) $d = 3 \Rightarrow$ (8.179)

$$\begin{aligned} \mathcal{C} := \int_{B_\varepsilon(x)} |U^*(\widetilde{x}, y)| \, dy &= \frac{1}{4\pi} \int_{|y-\widetilde{x}| < 2\varepsilon} \frac{1}{|y - \widetilde{x}|} dy \\ &= \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} \int_0^{2\varepsilon} \frac{1}{r} r^2 \sin \psi dr d\psi d\varphi \\ &= 2\varepsilon^2. \end{aligned}$$

取极限 $\widetilde{x} \to x, \varepsilon \to 0$，可证得。 □

类似地，我们也可以得到单层位势算子的外界迹 (exterior trace)

$$(Vw)(x) = \gamma_0^{\text{ext}} \left(\widetilde{V} w \right)(x) := \lim_{\Omega^c \ni \widetilde{x} \mapsto x \in \Gamma} \left(\widetilde{V} w \right)(\widetilde{x}), \quad x \in \Gamma. \tag{8.185}$$

单层位势的跃动关系 (jump relation) 因此可以表示为

$$\left[\gamma_0 \widetilde{V} w \right] := \gamma_0^{\text{ext}} \left(\widetilde{V} w \right)(x) - \gamma_0^{\text{int}} \left(\widetilde{V} w \right)(x) = 0, \quad x \in \Gamma. \tag{8.186}$$

8.6.3 伴随双层位势

由第8.6.2可知，基于给定密度方程 $w \in H^{-\frac{1}{2}}(\Gamma)$，我们可以定义一个位势算子 $\left(\widetilde{V} w \right) \in H^1(\Omega)$，作为齐次偏微分方程的解，详见 Lemma 8.20。

在此基础上，可以进一步将对应的内界共形导数定义成以下有界线性算子 $\gamma_1^{\text{int}} \widetilde{V} : H^{-\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma)$，满足

$$\left\| \gamma_1^{\text{int}} \widetilde{V} w \right\|_{H^{-\frac{1}{2}}(\Omega)} \le c \, \|w\|_{H^{-\frac{1}{2}}(\Gamma)}, \quad \forall w \in H^{-\frac{1}{2}}(\Gamma).$$

于是有伴随双层位势算子 (adjoint double layer potential)如下

Lemma 8.22 (伴随双层位势算子). 给定密度方程 $w \in H^{-\frac{1}{2}}(\Gamma)$，对应内部共形导数 $\gamma_1^{int}\left(\widetilde{V}w\right)(x) \in H^{-\frac{1}{2}}(\Gamma)$ 的表现式为

$$\gamma_1^{int}\left(\widetilde{V}w\right)(x) = \sigma(x)w(x) + (K'w)(x), \quad x \in \gamma,$$

采用内积形式可表现为

$$\langle \gamma_1^{int}\left(\widetilde{V}w\right)(x), \nu\rangle_\Gamma = \langle \sigma w + K'w, \nu\rangle_\Gamma, \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma),$$

其中与伴随双层位势相关的算子 $(K'w)(x)$ 和 $\sigma(x)$ 分别定义为

$$(K'w)(x) := \lim_{\varepsilon\to 0}\int_{y\in\Gamma:|y-x|\geq\varepsilon} \gamma_{1,x}^{int}U^*(x,y)w(y)\,ds_y, \tag{8.187}$$

$$\sigma(x) := \lim_{\varepsilon\to 0}\frac{1}{2(d-1)\pi}\frac{1}{\varepsilon^{d-1}}\int_{y\in\Omega:|y-x|=\varepsilon} ds_y, \quad x \in \Gamma. \tag{8.188}$$

证明. 已知给定方程 $w \in H^{-\frac{1}{2}}(\Gamma)$，齐次偏微分方程的解，$u(x) = \left(\widetilde{V}w\right) \in H^1(\Omega)$，见 Lemma 8.20。

进而由格林第一恒等式(8.6)可得，对某一 $\varphi \in C^\infty(\Omega)$ 我们有

$$\begin{aligned}\int_\Gamma u(x)\gamma_0^{\text{int}}\varphi(x)\,ds_x &= \int_\Omega \nabla_x u(x)\nabla_x\varphi(x)\,dx \\ &= \int_\Omega \nabla_x \underbrace{\int_\Gamma U^*(x,y)w(y)\,ds_y}_{=:u(x)} \nabla\varphi(x)\,dx\end{aligned}$$

由弱奇异表面积分 Lemma 8.21可见

$$u(x) = (Vw) = \gamma_0^{\text{int}}\left(\widetilde{V}w\right) = \int_\Gamma U^*(x,y)w(y)\,ds_y,$$

上式因此可调整为

$$\begin{aligned}\int_\Gamma \gamma_1^{\text{int}}u(x)\gamma_0^{\text{int}}\varphi(x)\,ds_x &= \int_\Omega \nabla_x\left[\lim_{\varepsilon\to 0}\int_{y\in\Gamma:|y-x|\geq\varepsilon} U^*(x,y)w(y)\,ds_y\right]\nabla_x\varphi(x)\,dx \\ &= \int_\Gamma w(y)\lim_{\varepsilon\to 0}\underbrace{\int_{x\in\Gamma:|x-y|\geq\varepsilon}\nabla_x U^*(x,y)\nabla_x\varphi(x)\,dx}_{=:\mathcal{A}}\,ds_y,\end{aligned}$$

对 $\mathcal{A}$ 再一次使用格林第一恒等式(8.6)

$$\begin{aligned}\mathcal{A} &:= \int_{x\in\Gamma:|x-y|\geq\varepsilon}\nabla_x U^*(x,y)\nabla_x\varphi(x)\,dx \\ &= \underbrace{\int_{x\in\Omega:|x-y|=\varepsilon}\gamma_{1,x}^{\text{int}}U^*(x,y)\varphi(x)\,ds_x}_{=:\mathcal{B}} + \underbrace{\int_{x\in\Gamma:|x-y|\geq\varepsilon}\gamma_{1,x}^{\text{int}}U^*(x,y)\gamma_0^{\text{int}}\varphi(x)\,ds_x}_{=:\mathcal{C}},\end{aligned}$$

其中

1. $\mathcal{C}$ 对应伴随双层位势算子 K'，如式(8.187)。
2. $\mathcal{B}$ 计算如下

$$\begin{aligned}\mathcal{B} &:= \int_{x\in\Omega:|x-y|=\varepsilon}\gamma_{1,x}^{\text{int}}U^*(x,y)\varphi(x)\,ds_x \\ &= \underbrace{\int_{x\in\Omega:|x-y|=\varepsilon}\gamma_{1,x}^{\text{int}}U^*(x,y)\left[\varphi(x)-\varphi(y)\right]ds_x}_{=:\mathcal{B}_1} + \varphi(y)\underbrace{\int_{x\in\Omega:|x-y|=\varepsilon}\gamma_{1,x}^{\text{int}}U^*(x,y)\,ds_x}_{=:\mathcal{B}_2},\end{aligned}$$

并且

$$\|\mathcal{B}\| \leq \max_{x\in\Omega:|x-y|=\varepsilon} |\varphi(x)-\varphi(y)| \underbrace{\int_{x\in\Omega:|x-y|=\varepsilon} \left|\gamma_{1,x}^{\text{int}} U^*(x,y)\right| ds_x}_{=:\mathcal{D}}$$

(a) 对于 $\mathcal{B}_1$，取 $\varepsilon \to 0$ 时的极值

$$\lim_{\varepsilon\to 0} |\mathcal{B}_1| = \lim_{\varepsilon\to 0} \left| \int_{x\in\Omega:|x-y|=\varepsilon} \gamma_{1,x}^{\text{int}} U^*(x,y) \left[\varphi(x)-\varphi(y)\right] ds_x \right| = 0,$$

(b) 对于 $\mathcal{B}_2$，设

$$n_x := \frac{y-x}{|y-x|}, x\in\Omega, \quad \varepsilon := |y-x|,$$

$$\begin{aligned}
\hookrightarrow \mathcal{B}_2 &= \int_{x\in\Omega:|x-y|=\varepsilon} \gamma_{1,x}^{\text{int}} U^*(x,y)\, ds_x \\
&= -\frac{1}{2(d-1)\pi} \int_{x\in\Omega:|x-y|=\varepsilon} \frac{(nx, x-y)}{|x-y|^d} ds_x \\
&= \frac{1}{2(d-1)\pi} \int_{x\in\Omega:|x-y|=\varepsilon} \frac{1}{|x-y|^{d-1}} ds_x \\
&= \frac{1}{2(d-1)\pi} \frac{1}{\varepsilon^{d-1}} \int_{x\in\Omega:|x-y|=\varepsilon} ds_x
\end{aligned}$$

(c) 分 $d=2,3$ 两种情况来讨论 $\mathcal{D}$ 的值

i. $d=2 \Rightarrow$

$$\begin{aligned}
\mathcal{D} &\leq \int_{x\in\mathcal{R}^2:|x-y|=\varepsilon} \left|\gamma_{1,x}^{\text{int}} U^*(x,y)\right| ds_x \\
&= \frac{1}{2\pi} \int_{x\in\mathcal{R}^2:|x-y|=\varepsilon} \frac{1}{|x-y|} ds_x = 1,
\end{aligned}$$

ii. $d=3 \Rightarrow$

$$\begin{aligned}
\mathcal{D} &\leq \int_{x\in\mathcal{R}^3:|x-y|=\varepsilon} \left|\gamma_{1,x}^{\text{int}} U^*(x,y)\right| ds_x \\
&= \frac{1}{4\pi} \int_{x\in\mathcal{R}^3:|x-y|=\varepsilon} \frac{1}{|x-y|^2} ds_x = 1.
\end{aligned}$$

3. 由此我们可得

$$\begin{aligned}
&\int_\Gamma \gamma_1^{\text{int}} u(x) \gamma_0^{\text{int}} \varphi(x)\, ds_x \\
&= \int_\Gamma \left\{ \lim_{\varepsilon\to 0} \int_{x\in\Gamma:|x-y|\geq\varepsilon} \gamma_{1,x}^{\text{int}} U^*(x,y) \gamma_0^{\text{int}} \varphi(x) ds_x + \gamma_0^{\text{int}} \varphi(y) \sigma(y) \right\} ds_y \\
&= \left[\int_\Gamma \gamma_0^{\text{int}} \varphi(x) \lim_{\varepsilon\to 0} \int_{x\in\Gamma:|x-y|\geq\varepsilon} \gamma_{1,x}^{\text{int}} U^*(x,y) w(y) ds_y ds_x \right] + \left[\int_\Gamma w(y)\sigma(y)\varphi(y) ds_y \right] \\
&= \int_\gamma \gamma_0^{\text{int}} \varphi(x) \left[\sigma(x) w(x) + (K'w)(x)\right] ds_x.
\end{aligned}$$

□

设 $x\in\gamma=\partial\Omega$ 是可积区间 (或至少在 Γ 的某一段邻域中可积)，此时定义式(8.188)由

$$\sigma(x) = \frac{1}{2}, \quad \text{对于 (几乎所有)} x\in\Gamma.$$

伴随着单层位势出现的边界积分算子 K' 是伴随双层位势 (adjoint double layer potential)。K' 线性且有界，即

$$\|K'w\|_{H^{-\frac{1}{2}}(\Gamma)} \le c_2^{K'} \|w\|_{H^{-\frac{1}{2}}(\Gamma)}, \quad w \in H^{-\frac{1}{2}(\Gamma)}. \tag{8.189}$$

由 Lemma 8.22进一步可知, 单层位势算子 $\widetilde{V}$ 在 $H^{-\frac{1}{2}}(\Gamma)$ 中的外部共形导数 (exterior conormal derivative)

$$\left(\gamma_1 \widetilde{V} w\right)(x) = [\sigma(x) - 1]\, w(x) + (K'w)(x), \quad x \in \Gamma.$$

Lemma 8.23 (单层位势外部共形导数的跃动关系)**.** *在 $H^{-\frac{1}{2}}(\Gamma)$ 中，单层位势 $\widetilde{V}$ 的外部共法导数满足跃动关系 (jump relation)*

$$\gamma_1\left(\widetilde{V} w\right) := \gamma_1^{ext}(Vw) - \gamma_1^{int}\left(\widetilde{V} w\right) = -w(x), \quad x \in \Gamma. \tag{8.190}$$

证明. 对于 $u = \widetilde{V} w, \varphi \in C_0^\infty(\mathbb{R}^d)$,

1. 一方面我们有

$$\begin{aligned}
\int_{\mathbb{R}^d} [-\Delta u(x)]\, \varphi(x)\, dx &= \int_{\mathbb{R}^d} -\Delta_x \int_\Gamma U^*(x,y) w(y) ds_y \varphi(x) dx \\
&= \int_\Gamma w(y) \int_{\mathbb{R}^d} -\Delta_x U^*(x,y) \varphi(x) dx ds_y \\
&= \int_\Gamma w(y) \int_{\mathbb{R}^d} \delta_0(x-y) \varphi(x) dx ds_y \\
&= \int_\Gamma w(y) \gamma_0^{\text{int}} \varphi(y) ds_y.
\end{aligned}$$

2. 另一方面我们有

$$\begin{aligned}
\int_{\mathbb{R}^d} [-\Delta u(x)]\, \varphi(x)\, dx &= a_{\mathbb{R}^d}(u,\varphi) = a_\Omega(u,\varphi) + a_{\Omega^c}(u,\varphi) \\
&= \int_\Gamma \gamma_1^{\text{int}} u(x) \gamma_0^{\text{int}} \varphi(x) ds_x + \int_\Gamma \gamma_1^{\text{ext}} u(x) \gamma_0^{\text{ext}} \varphi(x) ds_x.
\end{aligned}$$

3. 结合上两式可得

$$\int_\Gamma w(x)\varphi(x) ds_x = \int_\Gamma \gamma_1^{\text{int}} u(x) \gamma_0^{\text{int}} \varphi(x) ds_x - \int_\Gamma \gamma_1^{\text{ext}} u(x) \gamma_0^{\text{int}} \varphi(x)\, ds_x, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^d).$$

对 $\|\cdot\|_{H^{\frac{1}{2}}(\Gamma)}$ 取 $C_0^\infty(\mathbb{R}^d)$ 的闭包，再使用配偶配对，可证(8.190)成立。

□

8.6.4　双层位势

对于给定的密度方程 $\nu(x) \in H^{\frac{1}{2}}(\Gamma)$，我们将双层位势 (double layer potential)定义如下

$$u(x) = (W\nu)(x) := \int_\Gamma \left[\gamma_{1,y}^{\text{int}} U^*(x,y)\right] \nu(y)\, dy, \quad x \in \Omega \cup \Omega^c. \tag{8.191}$$

Lemma 8.24. *$u(x) = (W\nu)(x), x \in \Omega \cup \Omega^c$ 构成齐次偏微分方程*

$$-\Delta_x u(x) = 0, \quad x \in \Omega \cup \Omega^c$$

的解。若 $\nu \in H^{\frac{1}{2}}(\Gamma)$，则我们有 $u \in H^1(\Omega)$，满足关系

$$\|u\|_{H^1(\Omega)} = \|W\nu\|_{H^1(\Omega)} \le c \|\nu\|_{H^{\frac{1}{2}}(\Gamma)}.$$

证明. 由 $x \in \Omega \cup \Omega^c, y \in \Gamma$ 可得 $x \neq y$，进而可以通过改动差分、积分的顺序，调整泊松方程(8.159)中的基本解 $U(x,y)$：对于 $\varphi \in C^\infty(\Omega)$ 我们有

$$\begin{aligned}\langle W\nu, \varphi\rangle_\Omega &= \int_\Omega \int_\Gamma \left[\gamma_{1,y}^{\text{int}} U^*(x,y)\right] \nu(y) ds_y \varphi(x) dx \\ &= \int_\Gamma \nu(y) \gamma_{1,y}^{\text{int}} \int_\Gamma U^*(x,y)\varphi(x) dx ds_y \\ &= \int_\Gamma \nu(y) \gamma_{1,y}^{\text{int}} \left(\widetilde{N}_0, \varphi\right)(y) ds_y \\ &= \langle \nu, \gamma_1^{\text{int}} \left(\widetilde{N}_0, \varphi\right)\rangle_\Gamma.\end{aligned}$$

对于 $f \in \widetilde{H}^{-1}(\Omega)$ 而言，上式意味着

$$\langle W\nu, f\rangle_\Omega = \langle \nu, \gamma_1^{\text{int}} \widetilde{N}_0 f\rangle_\Gamma.$$

根据 Corollary 8.6可知，$\widetilde{N}_0 f \in H^1(\Omega)$ 是以下非齐次偏微分方程的解

$$-\Delta_x \left(\widetilde{N}_0 f\right)(x) = f(x), \quad x \in \Omega.$$

由 Lemma 8.19可得

$$\gamma_1^{\text{int}} \widetilde{N}_0 f \in H^{-\frac{1}{2}}(\Gamma),$$

进而

$$\begin{aligned}\|W\nu\|_{H^1(\Omega)} &= \sup_{0 \neq f \in \widetilde{H}^{-1}(\Omega)} \frac{\langle W\nu, f\rangle_\Omega}{\|f\|_{\widetilde{H}^{-1}(\Omega)}} \\ &= \sup_{0 \neq f \in \widetilde{H}^{-1}(\Omega)} \frac{\langle \nu, \gamma_1^{\text{int}} \widetilde{N}_0 f\rangle_\Gamma}{\|f\|_{\widetilde{H}^{-1}(\Omega)}} \\ &\leq c \|\nu\|_{H^{\frac{1}{2}}(\Gamma)}.\end{aligned}$$

□

由此可见，双层位势(8.191)定义了一个有界的线性算子 $W : H^{\frac{1}{2}}(\Gamma) \mapsto H^1(\Omega)$。

类似地，将内界迹算子 $\gamma_0\text{int}$ 作用于双层位势算子 $u = W\nu \in H^1(\Omega), \nu \in H^{\frac{1}{2}}(\Gamma)$，我们得到一个有界线性算子

$$\gamma_0^{\text{int}} W : H^{\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma),$$

满足

$$\left\|\gamma_0^{\text{int}} W\nu\right\|_{H^{\frac{1}{2}}(\Gamma)} \leq c \|\nu\|_{H^{\frac{1}{2}}(\Gamma)}, \quad \nu \in H^{\frac{1}{2}}(\Gamma).$$

Lemma 8.25 (双层位势的跃动关系)**.** 对于 $\nu \in H^{\frac{1}{2}}(\Gamma)$，可以将 $\gamma_0^{int}(W\nu)(x)$ 的表现式写为

$$\gamma_0^{int}(W\nu)(x) = [-1 + \sigma(x)]\nu(x) + (K\nu)(x), \quad x \in \Gamma, \tag{8.192}$$

其中 $\sigma(x), (K\nu)(x)$ 分别由伴随双层位势算子定义(8.188)(8.187)而来

$$\sigma(x) := \lim_{\varepsilon \to 0} \frac{1}{2(d-1)\pi} \frac{1}{\varepsilon^{d-1}} \int_{y \in \Omega : |y-x| = \varepsilon} ds_y, \tag{8.193}$$

$$(K\nu)(x) := \lim_{\varepsilon \to 0} \int_{y \in \Omega : |y-x| \geq \varepsilon} \left[\gamma_{1,y}^{int} U^*(x,y)\right] \nu(y) ds_y, \quad x \in \Gamma. \tag{8.194}$$

证明. 设任一给定的 $\varepsilon > 0$，现在要证明算子满足关系

$$(K_\varepsilon \nu) = \int_{y\in\Gamma:|y-x|\geq\varepsilon} \left[\gamma_{1,y}^{\text{int}} U^*(x,y)\right] \nu(y) ds_y.$$

来考虑极限值 $\Omega \ni \widetilde{x} \to x \in \Gamma$ 的情况。此时需假设 $|\widetilde{x} - x| < \varepsilon$，根据

$$\begin{aligned}
&(W\nu)(\widetilde{x}) - (K_\varepsilon \nu)(x) \\
&= \int_{y\in\Gamma:|y-x|\geq\varepsilon} \left[\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y) - \gamma_{1,y}^{\text{int}} U^*(x,y)\right] \nu(y) ds_y + \int_{y\in\Gamma:|y-x|<\varepsilon} \left[\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y)\right] \nu(y) ds_y \\
&= \underbrace{\int_{y\in\Gamma:|y-x|\geq\varepsilon} \left[\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y) - \gamma_{1,y}^{\text{int}} U^*(x,y)\right] \nu(y) ds_y}_{=:\mathcal{A}} \\
&\quad + \underbrace{\int_{y\in\Gamma:|y-x|<\varepsilon} \left[\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y)\right] [\nu(y) - \nu(x)] ds_y}_{=:\mathcal{B}} \\
&\quad + \nu(x) \underbrace{\int_{y\in\Gamma:|y-x|<\varepsilon} \gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y) ds_y}_{=:\mathcal{C}}.
\end{aligned}$$

1. $\mathcal{A} \Rightarrow$

$$\lim_{\Omega\ni\widetilde{x}\to x\in\Gamma} \mathcal{A} = \lim_{\Omega\ni\widetilde{x}\to x\in\Gamma} \int_{y\in\Gamma:|y-x|\geq\varepsilon} \left[\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y) - \gamma_{1,y}^{\text{int}} U^*(x,y)\right] \nu(y) ds_y = 0, \quad \forall \varepsilon > 0.$$

2. $\mathcal{B} \Rightarrow$

$$\begin{aligned}
|\mathcal{B}| &= \left| \int_{y\in\Gamma:|y-x|<\varepsilon} \left[\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y)\right] [\nu(y) - \nu(x)] ds_y \right| \\
&\leq \sup_{y\in\Gamma:|y-x|<\varepsilon} |\nu(x) - \nu(y)| \underbrace{\int_{y\in\Gamma:|y-x|<\varepsilon} \left|\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y)\right| ds_y}_{=:\mathcal{B}_1},
\end{aligned}$$

由于

$$\int_\Gamma \left|\gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y)\right| ds_y \leq M, \quad \widetilde{x} \in \Omega,$$

因此随着 $\varepsilon \to 0$，$\mathcal{B}_1 \to 0$，$\mathcal{B} \to 0$。

3. 来考虑

$$B_\varepsilon = \{y \in \Omega : |y - x| < \varepsilon\},$$

那么我们有

$$\mathcal{C} = \underbrace{\int_{\partial B_\varepsilon(x)} \gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y) ds_y}_{=:\mathcal{C}_1} - \underbrace{\int_{y\in\Omega:|y-x|=\varepsilon} \gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y) ds_y}_{=:\mathcal{C}_2}.$$

(a) 利用 $u(x) = 1$ 时的泊松方程表现式(8.159)，以及 $\widetilde{x} \in B_\varepsilon(x)$，我们有[15]

$$\mathcal{C}_1 = \int_{\partial B_\varepsilon(x)} \gamma_{1,y}^{\text{int}} U^*(\widetilde{x},y) ds_y = -1.$$

[15] cf. 第8.6.5节。

(b) 设 $\underline{n}(y)=\frac{y-x}{\varepsilon}$,

$$\begin{aligned}\lim_{\varepsilon\to 0}\lim_{\Omega\ni\widetilde{x}\to x\in\Gamma}\mathcal{C}_2&=\lim_{\varepsilon\to 0}\int_{y\in\Omega:|y-x|=\varepsilon}\gamma_{1,y}^{\mathrm{int}}U^*(\widetilde{x},y)ds_y\\&=-\lim_{\varepsilon\to 0}\frac{1}{2(d-1)\pi}\int_{y\in\Omega:|y-x|=\varepsilon}\frac{(n_y,y-x)}{|x-y|^d}ds_y\\&=-\lim_{\varepsilon\to 0}\frac{1}{2(d-1)\pi}\frac{1}{\varepsilon^{d-1}}\int_{y\in\Omega:|y-x|=\varepsilon}ds_y\\&=-\sigma(x).\end{aligned}$$

由此可证。 □

类似地，我们有外界迹

$$\gamma_0^{\mathrm{ext}}\left(W\nu\right)(x)=\sigma(x)\nu(x)+\left(K\nu\right)(x),\quad x\in\Gamma. \tag{8.195}$$

结合(8.195)和(8.192)，可得双层位势的跃动关系

$$[\gamma_0 W\nu]\coloneqq\gamma_0^{\mathrm{ext}}\left(W\nu\right)(x)-\gamma_0^{\mathrm{int}}\left(W\nu\right)(x)=\nu(x),\quad x\in\Gamma. \tag{8.196}$$

Lemma 8.26. 双层位势 $u(x)=\left(W\nu\right)(x)$ 的共形导数 $\gamma_1 W\nu$，其跃动关系可以表示为

$$[\gamma_1 W\nu]\coloneqq\gamma_1^{ext}\left(W\nu\right)(x)-\gamma_1^{int}\left(W\nu\right)(x)=0,\quad x\in\Gamma. \tag{8.197}$$

证明. 1. 一方面，对于双层位势 $u(x)=\left(W\nu\right)(x), x\in\mathbb{R}^n,\varphi\in C_0^\infty(\mathbb{R}^d)$ 我们有

$$\begin{aligned}\int_{\mathbb{R}^d}\left[-\Delta u(x)\right]\varphi(x)\,dx&=\int_{\mathbb{R}^d}-\Delta_x\int_\Gamma U^*(x,y)w(y)ds_y\varphi(x)dx\\&=\int_\Gamma w(y)\gamma_{1,y}^{\mathrm{int}}\int_{\mathbb{R}^d}-\Delta_x U^*(x,y)\varphi(x)dxds_y\\&=\int_\Gamma w(y)\gamma_{1,y}^{\mathrm{int}}\int_{\mathbb{R}^d}\delta_0(x-y)\varphi(x)dxds_y=0.\end{aligned}$$

2. 另一方面，根据格林第一恒等式(8.6)我们有

$$\begin{aligned}a_{\mathbb{R}^d}(u,\varphi)&=a_\Omega(u,\varphi)+a_{\Omega^c}(u,\varphi)\\&=\int_\Gamma\gamma_1^{\mathrm{int}}u(x)\gamma_0^{\mathrm{int}}\varphi(x)ds_x-\int_\Gamma\gamma_1^{\mathrm{ext}}u(x)\gamma_0^{\mathrm{ext}}\varphi(x)ds_x.\end{aligned}$$

3. 两式联立 $\Rightarrow$

$$\int_\Gamma\gamma_1^{\mathrm{int}}u(x)\gamma_0^{\mathrm{int}}\varphi(x)ds_x=\int_\Gamma\gamma_1^{\mathrm{ext}}u(x)\gamma_0^{\mathrm{ext}}\varphi(x)ds_x.$$

对 $\|\cdot\|_{H^1(\mathbb{R}^d)}$ 取 $C_0^\infty(\mathbb{R}^d)$ 的闭包，我们可得 $\gamma_0^{\mathrm{int}}\varphi=\gamma_0^{\mathrm{ext}}\varphi$，进而证得

$$\gamma_1^{\mathrm{ext}}u-\gamma_1^{\mathrm{int}}u=0.$$

□

8.6.5 超奇异边界积分算子

由第8.6.4节关于双层位势的介绍可见，给定密度方程 $\nu \in H^{\frac{1}{2}}(\Gamma)$，它的双层位势 $(W\nu)$ 的余法向导数 (conormal derivative)$\gamma_1^{\text{int}}(W\nu)(x)$ 定义了一个有界算子

$$\gamma_1^{\text{int}} W : H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma),$$

可表示为

$$(D\nu)(x) := -\gamma_1^{\text{int}}(W\nu)(x) = -\lim_{\Omega \ni \widetilde{x} \to x \in \Gamma} n_x \nabla_{\widetilde{x}} (W\nu)(\widetilde{x}), \quad x \in \Gamma, \tag{8.198}$$

满足有界性

$$\|D\nu\|_{H^{-\frac{1}{2}}(\Gamma)} \le c_2^D \ \|\nu\|_{H^{\frac{1}{2}}(\Gamma)}, \quad \nu \in H^{\frac{1}{2}}(\Gamma). \tag{8.199}$$

具体说来，由(8.198)可见

1. 对于 $d = 2$ 的二维系统，双层位势表示为

$$(W\nu)(\widetilde{x}) = \frac{1}{2\pi} \lim_{\varepsilon \to 0} \int_{y \in \Gamma : |y-x| \ge \varepsilon} \frac{(\widetilde{x} - y, n_y)}{|\widetilde{x} - y|^2} \nu(y) ds_y, \quad \widetilde{x} \in \Omega. \tag{8.200}$$

对于任一给定的 $\varepsilon > 0$，我们取极限值 $\Omega \ni \widetilde{x} \to x \in \Gamma$，将(8.200)对 $\widetilde{x}$ 求偏导，代入(8.198)得

$$(D_\varepsilon \nu)(x) = \frac{1}{2\pi} \lim_{\varepsilon \to 0} \int_{y \in \Gamma : |y-x| \ge \varepsilon} \left[-\frac{(n_x, n_y)}{|x-y|^2} + 2\frac{(x-y, n_x)(x-y, n_y)}{|x-y|^4} \right] \nu(y)\, ds_y. \tag{8.201}$$

2. 类似地，对于 $d = 3$ 的三维系统，可得

$$(D_\varepsilon \nu)(x) = \frac{1}{4\pi} \int_{y \in \Gamma : |y-x| \ge \varepsilon} \left[-\frac{(n_x, n_y)}{|x-y|^3} + 3\frac{(y-x, n_x)(y-x, n_y)}{|x-y|^5} \right] \nu(y)\, ds_y. \tag{8.202}$$

然而，无论是二维系统(8.201)还是三维系统(8.202)的情况，极限 $\varepsilon \to 0, x \in \Gamma$ 下的积分都不存在柯西主值 (Cauchy principal value)。我们将此类算子 D 称为超奇异边界积分算子 (hypersingular boundary integral operator)，作为柯西积分 (Cauchy integral) 的一个扩展[16]。

在不存在柯西主值的情况下，为了求得 D 的线性表现式，往往需引入一些正则设定，如将 $u_0(x) \equiv 1$ 引入泊松方程表现式(8.159)中可得 (更详细介绍见305页第8.6.8节。)

$$1 = -\int_\Gamma \gamma_{1,y}^{\text{int}} U^*(\widetilde{x}, y) ds_y, \quad \widetilde{x} \in \Omega,$$

$$\hookrightarrow \nabla_x (Wu_0)(\widetilde{x}) = \underline{0}, \quad \widetilde{x} \in \Omega,$$

$$\hookrightarrow (Du_0)(x) = 0, \quad x \in \Gamma. \tag{8.203}$$

进而可以将(8.198)写为

$$\begin{aligned}(D\nu)(x) &= -\lim_{\Omega \ni \widetilde{x} \to x \in \Gamma} n_x \nabla_{\widetilde{x}} (W\nu)(\widetilde{x}) \\ &= -\lim_{\Omega \ni \widetilde{x} \to x \in \Gamma} n_x \nabla_{\widetilde{x}} \int_\Gamma \gamma_{1,y}^{\text{int}} U^*(\widetilde{x}, y) [\nu(y) - \nu(x)]\, ds_y, \quad x \in \Gamma.\end{aligned}$$

根据上式，如果密度方程是连续的，我们就可以将超奇异边界积分算子 D 的表现式写为

$$(D\nu)(x) = -\int_\Gamma \gamma_{1,x}^{\text{int}} \gamma_{1,y}^{\text{int}} U^*(x, y) [\nu(y) - \nu(x)]\, ds_y, \quad x \in \Gamma. \tag{8.204}$$

下面分二维、三维空间两个情况继续展开分析。

[16]关于超奇异边界积分方程 (hypersingular boundary integral equation)，一个简要介绍可见 (Mason and Handscomb, 2003, Ch.9)，更多介绍可见Lifanov et al. (2003); Ang (2013)，数值近似算法的探讨可见Guiggiani et al. (1992)。

二维空间

以 $d=2$ 为例。假定 $\Gamma=\partial\Omega$ 是分段平滑的，满足

$$\Gamma=\cup_{k=1}^{p}\Gamma_k,$$

每一分段 Γ_k 都可由局部参数化表示

$$\Gamma_k:=y=y(t)=\begin{pmatrix}y_1(t)\\y_2(t)\end{pmatrix},\quad t\in(t_k,t_{k+1}),\tag{8.205}$$

其中 $y_i(t),i=1,2$ 为连续可导方程。

结合(8.204)、(8.205)我们有

$$\begin{aligned}ds_y&=\sqrt{[y_1'(t)]^2+[y_2'(t)]^2}\,dt\\n(y)&=\frac{1}{\sqrt{[y_1'(t)]^2+[y_2'(t)]^2}}\begin{pmatrix}y_2'(t)\\-y_1'(t)\end{pmatrix},\quad y\in\Gamma_k,\end{aligned}$$

其中 $n(y)$ 表示外部法向量 (exterior normal vector)。

二维系统中对于 $x\in\mathbb{R}^2$ 的情况，某一标量方程 $\widetilde{\nu}(x)$ 的旋转 (rotation, 或称 curl)可以定义为

$$\underline{\mathrm{curl}}\,\widetilde{\nu}(x):=\begin{pmatrix}\frac{\partial}{\partial x_2}\widetilde{\nu}(x)\\-\frac{\partial}{\partial x_1}\widetilde{\nu}(x)\end{pmatrix},$$

现在来看 $\widetilde{\nu}(x)$ 的计算：给定某一密度方程 $\nu(x),x\in\Gamma_k$，则我们可以考虑一个在 Γ_k 的适宜邻域上的延拓 $\widetilde{\nu}$

$$\widetilde{\nu}(\widetilde{x})=\nu(x),\quad \widetilde{x}=x+(\widetilde{x}-x,\underline{n}(x))\,\underline{n}(x).\tag{8.206}$$

也即，对于 $x\in\Gamma_k$，我们可以定义 $\nu(x)$ 的旋转 $\mathrm{curl}_{\Gamma_k}\nu(x)$ 为

$$\begin{aligned}\mathrm{curl}_{\Gamma_k}\nu(x)&:=\underline{n}(x)\,\underline{\mathrm{curl}}\widetilde{\nu}(x)\\&=n_1(x)\frac{\partial}{\partial x_2}\widetilde{\nu}(x)-n_2(x)\frac{\partial}{\partial x_1}\widetilde{\nu}(x),\quad x\in\Gamma_k.\end{aligned}$$

根据上式我们有

$$\begin{aligned}\int_{\Gamma_k}\mathrm{curl}_{\Gamma_k}\nu(y)ds_y&=\int_{\Gamma_k}\left[n_1(y)\frac{\partial}{\partial y_2}\widetilde{\nu}(y)-n_2(y)\frac{\partial}{\partial y_1}\widetilde{\nu}(y)\right]ds_y\\&=\int_{t_k}^{t_{k+1}}\left[y_2'(t)\frac{\partial}{\partial y_2}\nu(y(t))+y_1'(t)\frac{\partial}{\partial y_1}\nu(y(t))\right]dt\\&=\int_{t_k}^{t_{k+1}}\frac{d}{dt}\nu(y(t))dt,\end{aligned}$$

通过该式可见，$\mathrm{curl}_{\Gamma_k}\nu$ 与预先选定的延拓 $\widetilde{\nu}$ 无关。

Lemma 8.27. 设一个开放的有界区间 Γ_k，Γ_k 由局部参数式(8.205)所定义。对应的连续可导方程为 $y_i(t),i=1,2$。

如果 ν 和 w 连续可导，则我们有分部积分式

$$\int_{\Gamma_k}\nu(y)\,\mathrm{curl}_{\Gamma_k}w(y)\,ds_y=-\int_{\Gamma_k}\mathrm{curl}_{\Gamma_k}\nu(y)w(y)ds_y+\nu(y(t))w(y(t))\Big|_{t_k}^{t_{k+1}}.$$

证明.

$$\int_{\Gamma_k} \operatorname{curl}_{\Gamma_k} [\nu(y)w(y)]\, ds_y = \int_{t_k}^{t_{k+1}} \frac{d}{dt} [\nu(y(t))w(y(t))]\, dt$$
$$= [\nu(y(t))w(y(t))]_{t=t_k}^{t=t_{k+1}}.$$

□

若 ν 是沿着闭曲线 (closed curve) Γ 而定义的[17]，则我们可将 ν 的旋转 $\operatorname{curl}_\Gamma \nu(x)$ 定义为

$$\operatorname{curl}_\Gamma \nu(x) := \operatorname{curl}_{\Gamma_k} \nu(x), \quad x \in \Gamma_k, k = 1, \ldots, p,$$

那么根据 Lemma 8.27，我们有

Corollary 8.7. 设 Γ 是一个分段平滑的闭曲线，如果方程 ν 和 w 均分段连续可微，则我们有

$$\int_\Gamma \nu(y) \operatorname{curl}_\Gamma w(y) ds_y = -\int_\Gamma \operatorname{curl}_\Gamma \nu(y) w(y) ds_y + \sum_{k=1}^{p} \nu(y(t))w(y(t))\Big|_{t_k}^{t_{k+1}}, \tag{8.207}$$

进一步，如果 ν 和 w 全局连续，则

$$\int_\Gamma \nu(y) \operatorname{curl}_\Gamma w(y) ds_y = -\int_\Gamma \operatorname{curl} \nu(y) w(y) ds_y. \tag{8.208}$$

证明. 略。 □

利用分部积分法，我们可以将超奇异边界积分算子 D 的双线性形式，改写为单层位势算子 V 的双线性形式。在二维系统中，这种改写可以表示如下。

Theorem 8.33. 设 Γ 是一个分段平滑的封闭曲线。$u, \nu \in \Gamma$ 是全局连续方程。进一步，设 u, ν 在 Γ 的某一段 Γ_k 上连续可微。那么我们有

$$\langle Du, \nu \rangle_\Gamma = -\frac{1}{2\pi} \int_\Gamma \operatorname{curl}_\Gamma \nu(x) \int_\Gamma \log|x-y| \operatorname{curl}_\Gamma u(y) ds_y ds_x. \tag{8.209}$$

证明. 根据定义式(8.198)，超奇异边界积分算子 D 可以定义为双层位势算子 W 的法导数 $\gamma_1^{\text{int}} W$ 的负数。对于 $\widetilde{x} \in \Omega$，定义

$$w(\widetilde{x}) := (Wu)(\widetilde{x}) = -\frac{1}{2\pi} \int_\Gamma u(y) \frac{\partial}{\partial n_y} \log|\widetilde{x} - y|\, ds_y, \quad \widetilde{x} \in \Omega.$$

一方面 $\widetilde{x} \in \Omega, y \in \Gamma \Rightarrow \widetilde{x} \neq y$，另一方面

$$\frac{\partial}{\partial y_i} \log|\widetilde{x} - y| = \frac{y_i - \widetilde{x}_i}{|\widetilde{x} - y|^2} = -\frac{\widetilde{x}_i - y_i}{|\widetilde{x} - y|^2} = -\frac{\partial}{\partial \widetilde{x}_i} \log|\widetilde{x} - y|,$$

结合这两方面可得

$$\frac{\partial}{\partial \widetilde{x}_i} \left(\frac{\partial}{\partial n_y} \log|\widetilde{x} - y| \right) = -\underline{n}(y) \nabla_y \left(\frac{\partial}{\partial y_i} \log|\widetilde{x} - y| \right).$$

由于 $\Delta_y \log|\widetilde{x} - y| = 0$, 对于$y \neq \widetilde{x}$，则我们有以下两个等式

$$\begin{aligned}
\operatorname{curl}_{\Gamma,y} \left(\frac{\partial}{\partial y_1} \log|\widetilde{x} - y| \right) &= n_1(y) \frac{\partial}{\partial y_2} \frac{\partial}{\partial y_1} \log|\widetilde{x} - y| - n_2(y) \frac{\partial}{\partial y_1} \frac{\partial}{\partial y_1} \log|\widetilde{x} - y| \\
&= n_1(y) \frac{\partial}{\partial y_2} \frac{\partial}{\partial y_1} \log|\widetilde{x} - y| + n_2(y) \frac{\partial}{\partial y_2} \frac{\partial}{\partial y_2} \log|\widetilde{x} - y| \\
&= \underline{n}(y) \nabla y \left(\frac{\partial}{\partial y_2} \log|\widetilde{x} - y| \right),
\end{aligned}$$

[17]闭曲线是指这样的一种曲线，其起始点同时也是其终结点。

$$\operatorname{curl}_{\Gamma,y}\left(\frac{\partial}{\partial y_2}\log|\widetilde{x}-y|\right)=\underline{n}(y)\nabla y\left(\frac{\partial}{\partial y_1}\log|\widetilde{x}-y|\right).$$

因此，可以将全局连续方程 u 的双层位势 $w=(Wu)$ 的偏导数写为如下两个分部积分形式

$$\begin{aligned}\frac{\partial}{\partial\widetilde{x}_1}w(\widetilde{x})&=-\frac{1}{2\pi}\int_\Gamma u(y)\frac{\partial}{\partial\widetilde{x}}\frac{\partial}{\partial n_y}\log|\widetilde{x}-y|\ ds_y\\&=\frac{1}{2\pi}\int_\Gamma u(y)n(y)\nabla_y\left(\frac{\partial}{\partial y_1}\log|\widetilde{x}-y|\right)\ ds_y\\&=-\frac{1}{2\pi}\int_\Gamma u(y)\operatorname{curl}_{\Gamma,y}\left(\frac{\partial}{\partial y_2}\log|\widetilde{x}-y|\right)\ ds_y\\&=\frac{1}{2\pi}\int_\Gamma \operatorname{curl}_\Gamma u(y)\frac{\partial}{\partial y_2}\log|\widetilde{x}-y|\ ds_y,\end{aligned}$$

$$\frac{\partial}{\partial\widetilde{x}_2}w(\widetilde{x})=-\frac{1}{2\pi}\int_\Gamma \operatorname{curl}_\Gamma u(y)\frac{\partial}{\partial y_1}\log|\widetilde{x}-y|\ ds_y.$$

根据定义，双层位势 $w=Wu$ 的法导数

$$\begin{aligned}&\underline{n}(x)\nabla_{\widetilde{x}}w(\widetilde{x})\\&=\frac{1}{2\pi}\int_\Gamma \operatorname{curl}_\Gamma u(y)\left[n_1(x)\frac{\partial}{\partial y_2}\log|\widetilde{x}-y|-n_2(x)\frac{\partial}{\partial y_1}\log|\widetilde{x}-y|\right]\ ds_y\\&=\frac{1}{2\pi}\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\operatorname{curl}_\Gamma u(y)\left[n_1(x)\frac{\partial}{\partial y_2}\log|\widetilde{x}-y|-n_2(x)\frac{\partial}{\partial y_1}\log|\widetilde{x}-y|\right]\ ds_y\end{aligned}$$

取极限 $\Omega\ni\widetilde{x}\to x\in\Gamma$，上式变为

$$\begin{aligned}&\frac{\partial}{\partial n_x}w(x)\\&=\frac{1}{2\pi}\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\operatorname{curl}_\Gamma u(y)\left[n_1(x)\frac{\partial}{\partial y_2}\log|x-y|-n_2(x)\frac{\partial}{\partial y_1}\log|x-y|\right]\ ds_y\\&=-\frac{1}{2\pi}\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\operatorname{curl}_\Gamma u(y)\left[n_1(x)\frac{\partial}{\partial x_2}\log|x-y|-n_2(x)\frac{\partial}{\partial x_1}\log|x-y|\right]\ ds_y\\&=-\frac{1}{2\pi}\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\operatorname{curl}_\Gamma u(y)\operatorname{curl}_{\Gamma,x}\log|x-y|\ ds_y.\end{aligned}$$

进而

$$\begin{aligned}&\int_\Gamma \nu(x)\frac{\partial}{\partial n_x}w(x)ds_x\\&=-\frac{1}{2\pi}\int_{x\in\Gamma}\nu(x)\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\operatorname{curl}_\Gamma u(y)\operatorname{curl}_{\Gamma,x}\log|x-y|\ ds_yds_x\\&=-\frac{1}{2\pi}\int_{y\in\Gamma}\operatorname{curl}_\Gamma u(y)\lim_{\varepsilon\to0}\int_{x\in\Gamma:|x-y|\geq\varepsilon}\nu(x)\operatorname{curl}_{\Gamma,x}\log|x-y|\ ds_xds_y.\end{aligned}$$

由此可证。 □

三维空间

超奇异积分算子 D 的分部积分表现式也可以用于三维系统 $d=3$ 的分析中 (Dautray and Lions, 1990)。

设一个分段平滑表面 $\Gamma = \cup_{k=1}^{p}$，其中每一个分段 Γ_k 都可以用参数表示如下

$$y \in \Gamma_k : y(s,t) = \begin{pmatrix} y_1(s,t) \\ y_2(s,t) \\ y_3(s,t) \end{pmatrix}, \quad (s,t) \in \tau,$$

τ 指某个参考元 (reference element)。

$\mathbb{R}^3$ 中某个向量方程 $\underline{\nu}$ 的旋转可以定义为

$$\underline{\mathrm{curl}}\underline{\nu}(x) := \nabla \times \underline{\nu}(x), \quad x \in \mathbb{R}^3.$$

如果 u 是个 Γ_k 上的标量方程，其表面旋转 (surface curl) 可定义为

$$\underline{\mathrm{curl}}_{\Gamma_k} u(x) := \underline{n}(x) \times \nabla \widetilde{u}(x), \quad x \in \Gamma_k,$$

其中 $\widetilde{u} \in \Gamma_k$ 是一个给定的 $u \in \Gamma_k$ 在 Γ_k 的三维邻域上的适宜延拓。

此外我们有

$$\mathrm{curl}_{\Gamma_k} \underline{\nu}(x) := \underline{n}(x) \times \underline{\mathrm{curl}}\, \widetilde{\underline{\nu}}(x), \quad x \in \Gamma_k.$$

Lemma 8.28. 设 Γ 为一个 $\mathbb{R}^3$ 中的分段平滑封闭利普希茨表面。假定每个表面分段 Γ_k 都是平滑的，对应分段平滑边界曲线 $\partial\Gamma_k$。设 u 和 ν 均在 Γ 中全局连续，并且在每个分段 Γ_k 中都局部有节且平滑。那么由分部积分可得

$$\int_\Gamma \underline{\mathrm{curl}}_\Gamma u(x) \underline{\nu}(x)\, ds_x = -\int_\Gamma u(x)\, \mathrm{curl}_\Gamma\, \underline{\nu}(x)\, ds_x. \tag{8.210}$$

证明. 根据乘积法则可得

$$\nabla \times [\widetilde{u}(x)\underline{\nu}(x)] \equiv \nabla \widetilde{u}(x) \times \underline{\nu}(x) + \widetilde{u}(x)\, [\nabla \times \underline{\nu}(x)],$$

LHS (8.210)⇒

$$\begin{aligned}
\int_{\Gamma_k} \underline{\mathrm{curl}}_{\Gamma_k} u(x) \underline{\nu}(x)\, ds_x &= \int_{\Gamma_k} [\underline{n}(x) \times \nabla \widetilde{u}(x)]\, \nu(x)\, ds_x \\
&= \int_{\Gamma_k} [\nabla \widetilde{u}(x) \times \underline{\nu}(x)]\, n(x)\, ds_x \\
&= \int_{\Gamma_k} \{[\nabla \times (\widetilde{u}(x) \times \underline{\nu}(x))] - [\widetilde{u}(x)\, (\nabla \times \underline{\nu}(x))]\}\, n(x)\, ds_x \\
&= \underbrace{\int_{\partial\Gamma_k} u(x) \underline{\nu}(x) \underline{t}(x)\, d\sigma}_{=:0} - \int_{\Gamma_k} u(x)\, \mathrm{curl}_{\Gamma_k}\, \underline{\nu}(x)\, ds_x \\
&= -\int_{\Gamma_k} u(x)\, \mathrm{curl}_{\Gamma_k}\, \underline{\nu}(x)\, ds_x
\end{aligned}$$

□

Theorem 8.34 (超奇异积分算子的双线性表现式)**.** 设一个分段平滑的封闭表面 Γ。在 Γ 中定义全局连续方程 u 和 ν，且在 Γ_k 中可微。那么超奇异积分算子 D 的双线性表现式为

$$\langle Du, \nu \rangle_\Gamma = \frac{1}{4\pi} \int_\Gamma \int_\Gamma \frac{\underline{\mathrm{curl}}_\Gamma u(y) \underline{\mathrm{curl}}_\Gamma \nu(x)}{|x-y|}\, ds_x ds_y. \tag{8.211}$$

证明. 与二维系统的证明过程 (Theorem 8.33) 类似。根据定义式(8.198)，超奇异边界积分算子 D 可以定义为双层位势算子 W 的法导数 $\gamma_1^{\text{int}}W$ 的负数。对于 $\widetilde{x}\in\Omega$，定义

$$w(\widetilde{x}) := -\frac{1}{4\pi}\int_\Gamma u(y)\underbrace{\frac{\partial}{\partial n_y}\frac{1}{|\widetilde{x}-y|}}_{\text{核方程}}\,ds_y.$$

1. 求核方程的偏导数。根据

$$\frac{\partial}{\partial y_i}\frac{1}{|\widetilde{x}-y|}=\frac{\widetilde{x}-y_i}{|\widetilde{x}-y|^3}=-\frac{y_i-\widetilde{x}}{|\widetilde{x}-y|^3}=-\frac{\partial}{\partial x_i}\frac{1}{|\widetilde{x}-y|},$$

则

$$\frac{\partial}{\partial \widetilde{x}_i}\left(\frac{\partial}{\partial y_i}\frac{1}{|\widetilde{x}-y|}\right)=-\underline{n}(y)\nabla_y\left(\frac{\partial}{\partial y_i}\frac{1}{|\widetilde{x}-y|}\right).$$

2. 设 $\underline{e}_i$ 为 $\mathbb{R}^3$ 中的第 i-th 个单位向量 (unit vector)。由于 $\widetilde{x}\neq y$ 可得向量乘

$$\begin{aligned}\underline{\text{curl}}_y\left(\underline{e}_i\times\nabla_y\frac{1}{|\widetilde{x}-y|}\right)&=\nabla_y\times\left(\underline{e}_i\times\nabla_y\frac{1}{|\widetilde{x}-y|}\right)\\&=\left(\nabla_y\cdot\nabla_y\frac{1}{|\widetilde{x}-y|}\right)\underline{e}_i-\left(\nabla_y\cdot\underline{e}_i\right)\nabla_y\frac{1}{|\widetilde{x}-y|}\\&=\underbrace{\Delta_y\frac{1}{|\widetilde{x}-y|}\underline{e}_i}_{=0}-\frac{\partial}{\partial y_i}\left(\nabla_y\frac{1}{|\widetilde{x}-y|}\right)\\&=-\frac{\partial}{\partial y_i}\left(\nabla_y\frac{1}{|\widetilde{x}-y|}\right).\end{aligned}$$

3. 对双层位势 $w(\widetilde{x})$ 作关于 $\widetilde{x}_i$ 的偏微分，调整微分、积分顺序得

$$\begin{aligned}\frac{\partial}{\partial\widetilde{x}_i}w(\widetilde{x})&=-\frac{1}{4\pi}\int_\Gamma u(y)\frac{\partial}{\partial\widetilde{x}_i}\left(\frac{\partial}{\partial n_y}\frac{1}{|\widetilde{x}-y|}\right)ds_y\\&=\frac{1}{4\pi}\int_\Gamma u(y)n_y\cdot\nabla_y\left(\frac{\partial}{\partial y_i}\frac{1}{|\widetilde{x}-y|}\right)ds_y\\&=-\frac{1}{4\pi}\int_\Gamma u(y)n(y)\cdot\underline{\text{curl}}_y\left(\underline{e}_i\times\nabla_y\frac{1}{|\widetilde{x}-y|}\right)ds_y\\&=-\frac{1}{4\pi}\int_\Gamma u(y)\cdot\text{curl}_{\Gamma,y}\left(\underline{e}_i\times\nabla_y\frac{1}{|\widetilde{x}-y|}\right)ds_y.\end{aligned}$$

根据 Lemma 8.28，上式进一步改写为

$$\begin{aligned}\frac{\partial}{\partial\widetilde{x}_i}w(\widetilde{x})&=\frac{1}{4\pi}\int_\Gamma\underline{\text{curl}}_{\Gamma,y}u(y)\cdot\left(\underline{e}_i\times\nabla_x\frac{1}{|\widetilde{x}-y|}\right)ds_y\\&=-\frac{1}{4\pi}\int_\Gamma\underline{e}_i\left(\underline{\text{curl}}_{\Gamma,y}u(y)\times\nabla_y\frac{1}{|\widetilde{x}-y|}\right)ds_y.\end{aligned}$$

由此可得双层位势的梯度 (gradient)

$$\begin{aligned}\nabla_{\widetilde{x}}w(\widetilde{x})&=-\frac{1}{4\pi}\int_\Gamma\left(\underline{\text{curl}}_{\Gamma,y}u(y)\times\nabla_y\frac{1}{|\widetilde{x}-y|}\right)ds_y\\&=\frac{1}{4\pi}\int_\Gamma\left(\underline{\text{curl}}_{\Gamma,y}u(y)\times\nabla_{\widetilde{x}}\frac{1}{|\widetilde{x}-y|}\right)ds_y.\end{aligned}$$

两侧同时乘以法向量 (normal vector) $\underline{n}(x)$ 得

$$\begin{aligned}\underline{n}(x)\cdot\nabla_{\widetilde{x}}w(\widetilde{x}) &= \frac{1}{4\pi}\int_\Gamma\left(\underline{\mathrm{curl}}_{\Gamma,y}u(y)\times\nabla_{\widetilde{x}}\frac{1}{|\widetilde{x}-y|}\right)\cdot\underline{n}(x)\,ds_y\\ &= -\frac{1}{4\pi}\int_\Gamma\underline{\mathrm{curl}}_{\Gamma,y}u(y)\cdot\left(\underline{n}(x)\times\nabla_{\widetilde{x}}\frac{1}{|\widetilde{x}-y|}\right)ds_y\\ &= -\frac{1}{4\pi}\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\underline{\mathrm{curl}}_{\Gamma,y}u(y)\cdot\left(\underline{n}(x)\times\nabla_{\widetilde{x}}\frac{1}{|\widetilde{x}-y|}\right)ds_y.\end{aligned}$$

取极限 $\Omega\ni\widetilde{x}\to x\in\Gamma$，上式变为

$$\begin{aligned}(Du) &= -\frac{1}{4\pi}\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\underline{\mathrm{curl}}_{\Gamma,y}u(y)\cdot\left(\underline{n}(x)\times\nabla_x\frac{1}{|x-y|}\right)ds_y\\ &= -\frac{1}{4\pi}\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\underline{\mathrm{curl}}_{\Gamma,y}u(y)\cdot\underline{\mathrm{curl}}_{\Gamma,x}\frac{1}{|x-y|}\,ds_y\end{aligned}$$

4. 进而，双线性形式

$$\begin{aligned}\langle Du,\nu\rangle_\Gamma &= -\frac{1}{4\pi}\int_\Gamma\nu(x)\lim_{\varepsilon\to0}\int_{y\in\Gamma:|y-x|\geq\varepsilon}\underline{\mathrm{curl}}_{\Gamma,y}u(y)\cdot\underline{\mathrm{curl}}_{\Gamma,x}\frac{1}{|x-y|}\,ds_yds_x\\ &= -\frac{1}{4\pi}\int_\Gamma\lim_{\varepsilon\to0}\int_{x\in\Gamma:|x-y|\geq\varepsilon}\left(\nu(x)\underline{\mathrm{curl}}_{\Gamma,y}u(y)\right)\underline{\mathrm{curl}}_{\Gamma,x}\frac{1}{|x-y|}\,ds_xds_y\\ &= \frac{1}{4\pi}\int_\Gamma\lim_{\varepsilon\to0}\int_{x\in\Gamma:|x-y|\geq\varepsilon}\underbrace{\mathrm{curl}_{\Gamma,x}\left(\nu(x)\underline{\mathrm{curl}}_{\Gamma,y}u(y)\right)}\frac{1}{|x-y|}\,ds_xds_y,\end{aligned}$$

其中

$$\begin{aligned}\mathrm{curl}_{\Gamma,x}\left(\nu(x)\underline{\mathrm{curl}}_{\Gamma,y}u(y)\right) &= \underline{n}(x)\cdot\left[\nabla_x\times\left(\nu(x)\underline{\mathrm{curl}}_{\Gamma,y}u(y)\right)\right]\\ &= \underline{n}(x)\cdot\left[\nabla_x\nu(x)\times\underline{\mathrm{curl}}_{\Gamma,y}u(y)\right]\\ &= \left[n(x)\times\nabla_x\nu(x)\right]\cdot\underline{\mathrm{curl}}_{\Gamma,y}u(y)\\ &= \underline{\mathrm{curl}}_{\Gamma,x}\nu(x)\underline{\mathrm{curl}}_{\Gamma,y}u(y).\end{aligned}$$

因此

$$\langle Du,\nu\rangle_\Gamma = \frac{1}{4\pi}\int_\Gamma\lim_{\varepsilon\to0}\int_{x\in\Gamma:|x-y|\geq\varepsilon}\underline{\mathrm{curl}}_{\Gamma,x}\nu(x)\underline{\mathrm{curl}}_{\Gamma,y}\nu(y)\frac{1}{|x-y|}\,ds_xds_y$$

对上式取极限 $\varepsilon\to0$，证毕。

□

8.6.6 边界积分算子之间的关系

回顾一下泊松方程的表现式(8.159)

$$u(\widetilde{x}) = \int_\Omega U^*(\widetilde{x},y)f(y)\,dy + \int_\Gamma U^*(\widetilde{x},y)\gamma_1^{\mathrm{int}}u(y)\,ds_y - \int_\Gamma U^*(\widetilde{x},y)\gamma_0^{\mathrm{int}}u(y)\,ds_y,\quad \widetilde{x}\in\Omega.$$

取极限 $\Omega\ni\widetilde{x}\to x\in\Gamma$，由上式可以得出一系列边界和体积位势的性质。其中需要关注的是两个边界积分方程。第一个是将内界迹算子 γ_0^{int} 应用到解 $u(x),x\in\Gamma$ 中

$$\gamma_0^{\mathrm{int}}u(x) = \left(V\gamma_1^{\mathrm{int}}u\right)(x) + (1-\sigma(x))\,\gamma_0^{\mathrm{int}}u(x) - \left(K\gamma_0^{\mathrm{int}}u\right)(x) + N_0f(x). \tag{8.212}$$

第二个是将共法导数算子 γ_1^{int} 应用到解 $u(x), x \in \Gamma$ 中

$$\gamma_1^{\text{int}} u(x) = \sigma(x)\gamma_1^{\text{int}} u(x) + \left(K'\gamma_1^{\text{int}} u\right)(x) + \left(D\gamma_0^{\text{int}} u\right)(x) + N_1 f(x). \tag{8.213}$$

联立(8.212)和(8.213)，构成一个边界积分方程系统

$$\begin{pmatrix} \gamma_0^{\text{int}} u(x) \\ \gamma_1^{\text{int}} u(x) \end{pmatrix} = \underbrace{\begin{pmatrix} (1-\sigma) & V \\ D & \sigma I + K' \end{pmatrix}}_{=:\mathcal{C}} \begin{pmatrix} \gamma_0^{\text{int}} u(x) \\ \gamma_1^{\text{int}} u(x) \end{pmatrix} + \begin{pmatrix} N_0 f(x) \\ N_1 f(x) \end{pmatrix}. \tag{8.214}$$

我们将系数矩阵定义为卡尔德隆投影 (Calderón projection)

$$\mathcal{C} := \begin{pmatrix} (1-\sigma) & V \\ D & \sigma I + K' \end{pmatrix}. \tag{8.215}$$

Lemma 8.29 (卡尔德隆投影)**.** *式(8.215)算子 $\mathcal{C}$ 是一个投影，满足 $\mathcal{C} = \mathcal{C}^2$。*

证明. 设任一给定的 $(\psi, \varphi) \in H^{-\frac{1}{2}}(\Gamma) \times H^{\frac{1}{2}}(\Gamma)$。那么齐次偏微分方程的解 $u(\widetilde{x}), \widetilde{x} \in \Omega$ 可以表示为

$$u(\widetilde{x}) := (V\psi)(\widetilde{x}) - (W\varphi)(\widetilde{x}), \quad \widetilde{x} \in \Omega.$$

由 u 的内界迹和共法导数，对应 $x \in \Gamma$，我们有边界位势的性质

$$\begin{aligned} \gamma_0^{\text{int}} &= (V\psi)(x) + [1-\sigma(x)]\,\varphi(x) - (K\varphi)(x), \\ \gamma_1^{\text{int}} &= \sigma\psi(x) + (K'\psi)(x)\,(D\varphi)(x). \end{aligned}$$

不难看出，u 是齐次偏微分方程的解。通过 $\left[\gamma_0^{\text{int}}, \gamma_1^{\text{int}}\right], x \in \Gamma$，可以生成一组对应的柯西数 (Cauchy data)。这些柯西数构成边界积分方程系统(8.212)-(8.213)的解，即

$$\begin{aligned} \left(V\gamma_1^{\text{int}} u\right)(x) &= (\sigma I + K)\,\gamma_0^{\text{int}} u(x), \\ \left(D\gamma_0^{\text{int}} u\right)(x) &= ((1-\sigma)\,I - K')\,\gamma_1^{\text{int}} u(x). \end{aligned} \tag{8.216}$$

上式等价于

$$\begin{pmatrix} \gamma_0^{\text{int}} u(x) \\ \gamma_1^{\text{int}} u(x) \end{pmatrix} = \begin{pmatrix} (1-\sigma) & V \\ D & \sigma I + K' \end{pmatrix} \begin{pmatrix} \gamma_0^{\text{int}} u(x) \\ \gamma_1^{\text{int}} u(x) \end{pmatrix}. \tag{8.217}$$

联系

$$\begin{pmatrix} \gamma_0^{\text{int}} u(x) \\ \gamma_1^{\text{int}} u(x) \end{pmatrix} = \begin{pmatrix} (1-\sigma) & V \\ D & \sigma I + K' \end{pmatrix} \begin{pmatrix} \varphi(x) \\ \psi(x) \end{pmatrix}$$

可见，$\mathcal{C}$ 是一个投影。根据定义，投影矩阵，又称幂等矩阵 (idempotent matrix)，满足 $\mathcal{C} = \mathcal{C}^2$ (Seber and Lee, 2003, p.464)。 □

在得到卡尔德隆投影之后，我们有以下有界积分算子之间的关系。

Corollary 8.8 (有界积分算子之间的关系)**.** *所有有界积分算子均满足以下关系*

$$VD = (\sigma I + K)\,[(1-\sigma)\,I - K], \tag{8.218}$$

$$DV = (\sigma I + K')\,[(1-\sigma)\,I - K'], \tag{8.219}$$

$$VK' = KV, \tag{8.220}$$

$$K'D = DK. \tag{8.221}$$

由(8.220)可见，双层位势算子 K 是对称的——并不是自伴随的对称，而是随着单层位势算子 V 的对称。

ction 6.1.1

此外，边界积分方程系统中，假定单层位势 V 是可逆的 (更多讨论见)，则我们也可得到牛顿位势 $N_1 f$ 的表现式如下

Lemma 8.30 (牛顿位势算子与其他算子的关系)**.** 牛顿位势算子 $(N_1 f)(x), x \in \Gamma$ 的表现式为

$$(N_1 f)(x) = ((\sigma - 1)I + K')\, V^{-1}\, (N_0 f)(x) \tag{8.222}$$

证明. 由边界积分方程系统(8.214)第一行可知，假设 V 是可逆的

$$\gamma_1^{\text{int}} u(x) = V^{-1}\,(\sigma I + K')\,\gamma_0^{\text{int}} u(x) - V^{-1}\,(N_0 f)(x), \quad x \in \Gamma.$$

代入第二行，可得

$$\begin{aligned}
\gamma_1^{\text{int}} u(x) &= \left(D\gamma_0^{\text{int}} u\right)(x) + (\sigma I + K')\,\gamma_1^{\text{int}} u(x) + (N_1 f)(x) \\
&= \left(D\gamma_0^{\text{int}} u\right)(x) + (\sigma I + K')\left[V^{-1}\,(\sigma I + K)\,\gamma_0^{\text{int}} u(x) - V^{-1}\,(N_0 f)(x)\right] + (N_1 f)(x) \\
&= \left[D + (\sigma I + K')\,V^{-1}\,(\sigma I + K)\right]\gamma_0^{\text{int}}(x) - (\sigma I + K')\,V^{-1}\,(N_0 f)(x) + (N_1 f)(x), \quad x \in \Gamma.
\end{aligned}$$

整理可得

$$-V^{-1}\,(N_0 f)(x) = -\,(\sigma I + K')\,V^{-1}\,(N_0 f)(x) + (N_1 f)(x),$$

证毕。 □

8.6.7　单层位势算子的椭圆性和可逆性

单层位势算子 $V : H^{-\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$ 的可逆性可由拉克斯—密格拉蒙定理 Theorem 8.14证得。在此基础上，进一步讨论 V 的 $H^{\frac{1}{2}}(\Gamma)$-椭圆特性。

回顾，已知方程 $u(x) = \left(\widetilde{V} w\right)(x), x \in \Gamma$ 是以下内界狄利克雷边界值问题的解

$$\begin{aligned}
-\Delta u(x) &= 0, \quad x \in \Omega, \\
u(x) &= \gamma_0^{\text{int}}\left(\widetilde{V} w\right)(x) = (Vw)(x), \quad x \in \Gamma.
\end{aligned}$$

假定 $w \in H^{-\frac{1}{2}}(\Gamma)$，则 $u = \widetilde{V} w \in H^1(\Omega)$。通过选取方程 $\nu \in H^1(\Omega)$，由(8.6)可得内界狄利克雷边界值问题的格林第一恒等式

$$a_\Omega(u, \nu) := \int_\Omega \nabla u(x) \nabla \nu(x)\, dx = \langle \gamma_1^{\text{int}} u, \gamma_0^{\text{int}} \nu \rangle_\Gamma. \tag{8.223}$$

代入(8.129) (Corollary 8.5)，上式改写为不等式形式

$$a_\Omega(u, u) \geq c_1^{\text{int}} \left\| \gamma_1^{\text{int}} u \right\|_{H^{-\frac{1}{2}}(\Gamma)}. \tag{8.224}$$

随后我们需要计算外界共法导数 $\gamma_1^{\text{int}} u \in H^{-\frac{1}{2}}(\Gamma)$，这可以通过分析 $|x| \to \infty$ 时单层位势 $\left(\widetilde{V} w\right)(x)$ 的远端特征而得。为此，我们需要引入子空间 $H_*^{-\frac{1}{2}}(\Gamma)$，子空间中的方程正交于常数

$$H_*^{-\frac{1}{2}}(\Gamma) := \left\{ w \in H^{-\frac{1}{2}}(\Gamma) : \langle w, 1 \rangle_\Gamma = 0 \right\}. \tag{8.225}$$

Lemma 8.31. 假定

$$|x - y_0| > \max\{1, 2\,\text{diam}(\Omega)\}, \quad y_0 \in \Omega, x \in \mathbb{R}^d.$$

设

$$w \in \begin{cases} H^{-\frac{1}{2}}(\Gamma) & d=3, \\ H_*^{-\frac{1}{2}}(\Gamma) & d=2. \end{cases}$$

那么 $u=\widetilde{V}w$ 的边界可以表示为

$$|u(x)| = \left|\left(\widetilde{V}w\right)(x)\right| \le c_1(w)\,\frac{1}{|x-y_0|}, \tag{8.226}$$

$$|\nabla u(x)| = \left|\nabla\left(\widetilde{V}w\right)(x)\right| \le c_2(w)\,\frac{1}{|x-y_0|^2}. \tag{8.227}$$

证明. 由三角不等式我们有

$$\begin{aligned} |x-y_0| &\le |x-y| + |y-y_0| \\ &\le |x-y| + \operatorname{diam}(\Omega) \\ &\le |x-y| + \frac{1}{2}|x-y_0|. \end{aligned}$$

$$\hookrightarrow |x-y| \ge \frac{1}{2}|x-y_0|.$$

下面分 $d=3,2$ 分别证明。

1. 三维空间 $d=3 \Rightarrow$

$$\begin{aligned} \|U^*(x,\cdot)\|_{H^1(\Omega)}^2 &= \frac{1}{16\pi^2}\int_\Omega \frac{1}{|x-y|^2}\,d_y + \frac{1}{16\pi^2}\int_\Omega \frac{1}{|x-y|^4}\,d_y \\ &\le \frac{1}{4\pi^2}\int_\Omega \frac{1}{|x-y_0|^2}\,d_y + \frac{1}{\pi^2}\int_\Omega \frac{1}{|x-y_0|^4}\,d_y \\ &\le \frac{5}{4}\frac{|\Omega|}{\pi^2}\frac{1}{|x-y_0|^2}\,d_y, \end{aligned}$$

证得(8.226)。

$$\begin{aligned} |u(x)| &= \left|\langle \gamma_0^{\mathrm{int}} U^*(x,\cdot), w\rangle_\Gamma\right| \\ &\le \left\|\gamma_0^{\mathrm{int}} U^*(x,\cdot)\right\|_{H^{\frac{1}{2}}(\Gamma)} \|w\|_{H^{-\frac{1}{2}}(\Gamma)} \\ &\le c_{\mathrm{T}} \left\|\gamma_0^{\mathrm{int}} U^*(x,\cdot)\right\|_{H^1(\Omega)} \|w\|_{H^{-\frac{1}{2}}(\Gamma)}. \end{aligned}$$

$$\hookrightarrow \frac{\partial}{\partial x_i}u(x) = \frac{1}{4\pi}\int_\Gamma \frac{y_i - x_i}{|x-y|^3} w(y)\,ds_y,$$

证得(8.227)。

2. 二维空间 $d=2$ 中，沿着某个适宜的定值 $\bar{y}\in\Omega$ 作泰勒展开

$$\log|y-x| \approx \log|y_0-x| + \frac{(y-y_0, \bar{y}-x)}{|\bar{y}-x|^2}.$$

进而根据 $w \in H_*^{-\frac{1}{2}}(\Gamma)$ 可得

$$u(x) = -\frac{1}{2\pi}\int_\Gamma \frac{(y-y_0, \bar{y}-x)}{|\bar{y}-x|^2} w(y)\,ds_y,$$

证得(8.226)。在此基础上证得(8.227)。

□

对于 $y_0 \in \Omega, R \geq 2\,\mathrm{diam}(\Omega)$，定义球体

$$B_R(y_0) := \left\{x \in \mathbb{R}^d : |x - y_0| < R\right\}.$$

那么此时 $u(x) = (Vw)(x), x \in \Omega^c$ 就成为以下狄利克雷边界值问题的唯一解

$$-\Delta u(x) = 0, \quad x \in B_R(y_0)\backslash\overline{\Omega}, \tag{8.228}$$

$$u(x) = \gamma_0^{\mathrm{int}}\left(\widetilde{V}w\right)(x) = (Vw)(x), \quad x \in \Gamma, \tag{8.229}$$

$$u(x) = \left(\widetilde{V}w\right)(x), \quad x \in \partial B_R(y_0). \tag{8.230}$$

其中，对于有界域 $B_R(y_0)\backslash\overline{\Omega}$ (8.228) 中的情况，利用格林第一恒等式(8.6)可得

$$a_{B_R(y_0)\backslash\overline{\Omega}}(u,\nu) = \underbrace{-\langle\gamma_1^{\mathrm{ext}}u, \gamma_0^{\mathrm{ext}}\nu\rangle_\Gamma}_{=:\mathcal{A}} + \underbrace{\langle\gamma_1^{\mathrm{int}}u, \gamma_0^{\mathrm{int}}\nu\rangle_{\partial B_R(y_0)}}_{=:\mathcal{B}}. \tag{8.231}$$

1. 先来看 $\mathcal{B}$。设 $u = \nu$。由 Lemma 8.31可得

$$\begin{aligned}\left|\langle\gamma_1^{\mathrm{int}}u, \gamma_0^{\mathrm{int}}u\rangle_{\partial B_R(y_0)}\right| &\leq c_1(w)c_2(w)\int_{|x-y_0|=R}\frac{1}{|x-y_0|^3}\,ds_x \\ &\leq c\,R^{d-4}.\end{aligned}$$

2. 再来看极限 $R \to \infty$ 的情况。在外界域 (exterior domain) $\partial B_R(y_0)$ (8.230)中，根据格林第一恒等式(8.6)可得

$$a_{\Omega^c(u,u)} := \int_{\Omega^c}\nabla u(x)\nabla u(x)\,dx = -\langle\gamma_1^{\mathrm{ext}}u, \gamma_0^{\mathrm{ext}}u\rangle_\Gamma. \tag{8.232}$$

在二维空间中，假设条件 $w \in H_*^{-\frac{1}{2}}(\Gamma)$ 即可保证上述结论成立。以类似于 Corollary 8.5，(8.129)的方法，对于(8.232)的外界狄利克雷边界值问题，满足不等式

$$a_{\Omega^c(u,u)} \geq c_1^{\mathrm{ext}}\,\left\|\gamma_1^{\mathrm{ext}}u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)}. \tag{8.233}$$

Theorem 8.35 (单层位势算子的椭圆性 (三维空间))**.** 设

$$w \in \begin{cases} H^{-\frac{1}{2}}(\Gamma) & d = 3, \\ H_*^{-\frac{1}{2}}(\Gamma) & d = 2. \end{cases}$$

那么有以下不等式

$$\langle Vw, w\rangle_\Gamma \geq c_1^V\,\|w\|^2_{H^{-\frac{1}{2}}(\Gamma)}, \tag{8.234}$$

其中常数 $c_1^V \geq 0$。

证明. 根据 $u = \widetilde{V}w$，由内界、外界狄利克雷边界值问题的格林第一恒等式(8.223)、(8.232)可知

$$\begin{aligned}a_\Omega(u,u) &= \langle\gamma_1^{\mathrm{int}}u, \gamma_0^{\mathrm{int}}u\rangle_\Gamma, \\ a_{\Omega^c}(u,u) &= -\langle\gamma_1^{\mathrm{ext}}u, \gamma_0^{\mathrm{ext}}u\rangle_\Gamma.\end{aligned}$$

两式相加

$$\underbrace{a_\Omega(u,u) + a_{\Omega^c}(u,u)}_{=:\mathcal{C}} = \langle\underbrace{\left[\gamma_1^{\mathrm{int}}u - \gamma_1^{\mathrm{ext}}u\right]}_{=:\mathcal{A}}, \underbrace{\left[\gamma_0^{\mathrm{int}}u - \gamma_0^{\mathrm{ext}}u\right]}_{=:\mathcal{B}}\rangle_\Gamma.$$

1. 由单层位势算子的跃动关系(8.186)可得

$$\mathcal{B} := \gamma_0^{\text{int}} u - \gamma_0^{\text{ext}} u := \gamma_0 u,$$

2. 由单层位势共法导数的跃动关系(8.190)可得

$$\mathcal{A} := \gamma_1^{\text{int}} u - \gamma_1^{\text{ext}} u := w,$$

3. 将(8.224)，(8.233)代入 $\mathcal{C}$ 得

$$\begin{aligned}\mathcal{C} &:= a_\Omega(u,u) + a_{\Omega^c}(u,u)\\ &\geq c_1^{\text{int}} \left\|\gamma_1^{\text{int}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)} + c_1^{\text{ext}} \left\|\gamma_1^{\text{ext}} u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)}\\ &\geq \min\left\{c_1^{\text{int}}, c_2^{\text{ext}}\right\}\left[\left\|\gamma_1^{\text{int}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)} + \left\|\gamma_1^{\text{ext}} u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)}\right]\end{aligned}$$

$$\begin{aligned}\langle \mathcal{A}, \mathcal{B}\rangle_\Gamma &= \langle w, \gamma_0 u\rangle_\Gamma = \langle Vw, w\rangle_\Gamma\\ &= \mathcal{C} \geq \min\left\{c_1^{\text{int}}, c_2^{\text{ext}}\right\}\left[\left\|\gamma_1^{\text{int}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)} + \left\|\gamma_1^{\text{ext}} u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)}\right].\end{aligned}$$

此外根据(8.23)有

$$\begin{aligned}\|w\|^2_{H^{-\frac{1}{2}}(\Gamma)} &= \left\|\gamma_1^{\text{int}} u - \gamma_1^{\text{ext}} u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)}\\ &\leq \left[\left\|\gamma_1^{\text{int}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)} + \left\|\gamma_1^{\text{ext}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)}\right]^2\\ &\leq 2\left[\left\|\gamma_1^{\text{int}} u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)} + \left\|\gamma_1^{\text{ext}} u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)}\right],\end{aligned}$$

$$\begin{aligned}\langle Vw, w\rangle_\Gamma &\geq \min\left\{c_1^{\text{int}}, c_2^{\text{ext}}\right\}\left[\left\|\gamma_1^{\text{int}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)} + \left\|\gamma_1^{\text{ext}} u\right\|^2_{H^{-\frac{1}{2}}(\Gamma)}\right]\\ &\geq \underbrace{\min\left\{c_1^{\text{int}}, c_2^{\text{ext}}\right\}}_{=:c_1^V \geq 0} \|w\|^2_{H^{-\frac{1}{2}}(\Gamma)}.\end{aligned}$$

□

对于二维系统 $d=2$ 的情况，我们也可从 Theorem 8.35推得，单层位势 V 的 $H_*^{-\frac{1}{2}}(\Gamma)$-椭圆特性。具体来说，有两种证明思路。

鞍点变分问题求证椭圆性

第一种证明思路是，汇总 $d=2,3$ 的情况，构建鞍点问题

$$\langle Vt, \tau\rangle_\Gamma - \lambda\langle 1, \tau\rangle_\Gamma = 0, \quad \forall \tau \in H^{-\frac{1}{2}}(\Gamma), \tag{8.235}$$

$$\langle t, 1\rangle_\Gamma = 1, \tag{8.236}$$

寻找解 $(t,\lambda) \in H^{-\frac{1}{2}}(\Gamma) \times \mathbb{R}$。

将 t 定义如下

$$t := \widetilde{t} + \frac{1}{|\Gamma|}, \quad \tau \in H_*^{-\frac{1}{2}}(\Gamma),$$

不难看出，如果对于某一 $\widetilde{t} \in H_*^{-\frac{1}{2}}(\Gamma)$，(8.235)的解 t 同时也满足关系式(8.236)。那么求解上述按点问题的关键就变为，寻找合适的 $\widetilde{t}$，使满足如下关系

$$\langle V\widetilde{t}, \tau\rangle_\Gamma = -\frac{1}{|\Gamma|}\langle V1, \tau\rangle_\Gamma, \quad \forall \tau \in H_*^{-\frac{1}{2}}(\Gamma).$$

多自然密度的
绍见后

由 Theorem 8.35可见，上式中单层位势是 $H_*^{-\frac{1}{2}}(\Gamma)$-椭圆的。因此该变分问题可解，解唯一。设这个唯一解为 w_{eq}，表示自然密度方程 (natural density function)。

$$w_{\mathrm{eq}} := \widetilde{t} + \frac{1}{|\Gamma|}, \tag{8.237}$$

设 $\tau = w_{\mathrm{eq}}$，则拉格朗日乘子 λ 满足

$$\lambda = \langle V w_{\mathrm{eq}}, w_{\mathrm{eq}} \rangle_\Gamma.$$

在三维空间 $d = 3$ 中，由 Theorem 8.35可得拉，拉格朗日乘子严格为正 $\lambda > 0$，定义为 Γ 的容度 (capacity)，用 cap_Γ 表示。

在二维空间 $d = 2$ 中，cap_Γ 可定义为指数形式

$$\mathrm{cap}_\Gamma := \exp(-2\pi\lambda),$$

基于参数的基本解为

$$U^*(x, y) := \frac{1}{2\pi} \log r - \frac{1}{2\pi} \log |x - y|, \quad r \in \mathbb{R}^+,$$

对应的边界积分算子

$$(V_r w)(x) := \int_\Gamma U_r^*(x, y) w(y)\, ds_y, \quad x \in \Gamma.$$

由此可得

$$(V_r w)(x) = \frac{1}{2\pi} \log r + \lambda = \frac{1}{2\pi} \log \frac{r}{\mathrm{cap}_\Gamma},$$

尤其是当 $r = 1$ 时，有

$$\lambda := \frac{1}{2} \log \frac{1}{\mathrm{cap}_\Gamma}. \tag{8.238}$$

由(8.238)可见，在二维系统 $d = 2$ 中，$r = 1$ 的前提下，只有当指数容度 $\mathrm{cap}_\Gamma < 1$ 时，才能确保 $\lambda > 0$，进而使得解存在。Hsiao and Wendland (1977) 论证了，使满足 $\mathrm{cap}_\Gamma < 1$ 的一个充分条件是假定 $\mathrm{diam}(\Omega) < 1$——有时这便需要我们对给定的域 $\Omega \in \mathbb{R}^2$ 作规模调节，使得直径小于 1 的假定得到满足 (见304页第8.6.7节)。

Theorem 8.36 (单层位势算子的椭圆性 (二维空间))**.** *二维空间 $d = 2$ 中，设 $\mathrm{diam}(\Omega) < 1$，从而 $\lambda > 0$。单层位势 V 是 $H^{-\frac{1}{2}}(\Gamma)$-椭圆的，即*

$$\langle V w, w \rangle_\Gamma \geq \widetilde{c}_1^V \, \|w\|^2_{H^{-\frac{1}{2}}(\Gamma)}, \quad \forall w \in H^{-\frac{1}{2}}(\Gamma). \tag{8.239}$$

证明. 1. 看 RHS。对于任一给定 $w \in H^{-\frac{1}{2}}(\Gamma)$，都可以对 w 作唯一的分解

$$w = \widetilde{w} + \alpha w_{\mathrm{eq}}, \quad \widetilde{w} \in H_*^{-\frac{1}{2}}(\Gamma), \alpha = \langle w, 1 \rangle_\Gamma.$$

$$\begin{aligned}
\hookrightarrow \|w\|^2_{H^{-\frac{1}{2}}(\Gamma)} &= \|\widetilde{w} + \alpha w_{\mathrm{eq}}\|^2_{H^{-\frac{1}{2}}(\Gamma)} \\
&\leq \left[\|\widetilde{w}\|_{H^{-\frac{1}{2}}(\Gamma)} + \alpha \|w_{\mathrm{eq}}\|_{H^{-\frac{1}{2}}(\Gamma)} \right]^2 \\
&\leq 2 \left[\|\widetilde{w}\|^2_{H^{-\frac{1}{2}}(\Gamma)} + \alpha^2 \|w_{\mathrm{eq}}\|^2_{H^{-\frac{1}{2}}(\Gamma)} \right] \\
&\leq 2 \max \left\{ 1, \|w_{\mathrm{eq}}\|^2_{H^{-\frac{1}{2}}(\Gamma)} \right\} \left[\|\widetilde{w}\|^2_{H^{-\frac{1}{2}}(\Gamma)} + \alpha^2 \right].
\end{aligned}$$

2. 看 LHS。

$$\begin{aligned}
\langle V w, w \rangle_\Gamma &= \langle V (\widetilde{w} + \alpha w_{\mathrm{eq}}), \widetilde{w} + \alpha w_{\mathrm{eq}} \rangle_\Gamma \\
&= \underbrace{\langle V \widetilde{w}, \widetilde{w} \rangle}_{=:\mathcal{A}} + 2\alpha \underbrace{\langle \widetilde{V} w_{\mathrm{eq}}, \widetilde{w} \rangle_\Gamma}_{=:\mathcal{B}} + \alpha^2 \underbrace{\langle V w_{\mathrm{eq}}, w_{\mathrm{eq}} \rangle_\Gamma}_{=:\mathcal{C}},
\end{aligned}$$

其中

(a) 根据 Theorem 8.35

$$\mathcal{A} := \langle \widetilde{V} w_{\mathrm{eq}}, \widetilde{w} \rangle_\Gamma \geq c_1^{\mathrm{V}} \ \|\widetilde{w}\|^2_{H^{-\frac{1}{2}}(\Gamma)},$$

(b)

$$\mathcal{B} := \langle \widetilde{V} w_{\mathrm{eq}}, \widetilde{w} \rangle_\Gamma = 0,$$

(c) 根据定义有

$$\mathcal{C} := \langle V w_{\mathrm{eq}}, w_{\mathrm{eq}} \rangle_\Gamma = \lambda,$$

因此上式变为

$$\begin{aligned}\langle V w, w \rangle_\Gamma &= \mathcal{A} + 2\alpha\mathcal{B} + \alpha^2\mathcal{C} \\ &\geq c_1^{\mathrm{V}} \ \|\widetilde{w}\|^2_{H^{-\frac{1}{2}}(\Gamma)} + \alpha^2\lambda \\ &\geq \min\{c_1^{\mathrm{V}}, \lambda\} \left[\|\widetilde{w}\|^2_{H^{-\frac{1}{2}}(\Gamma)} + \alpha^2\right]\end{aligned}$$

3. 联立 LHS 和 RHS，可证。

□

第二种思路是从从自然密度方程 $w_{\mathrm{eq}} \in H^{-\frac{1}{2}}(\Gamma)$ 入手。已知 w_{eq} 是以下带有约束条件的算子方程的解

$$\begin{aligned}(V w_{\mathrm{eq}})(x) &= \lambda, \quad x \in \Gamma, \\ \langle w_{\mathrm{eq}}, 1 \rangle_\Gamma &= 1.\end{aligned}$$

我们首先对 w_{eq} 作规模调节：

$$w_{\mathrm{eq}} := \lambda\, \widetilde{w}_{\mathrm{eq}},$$

新的算子方程变为

$$\begin{aligned}(V \widetilde{w}_{\mathrm{eq}})(x) &= 1, \quad x \in \Gamma, \\ \langle \widetilde{w}_{\mathrm{eq}}, 1 \rangle_\Gamma &= \frac{1}{\lambda}.\end{aligned} \tag{8.240}$$

求解思路略。大致说来，第一步求解边界积分方程，对应(8.240)第一行，以求得作为自然密度的解 w_{eq}。第二步，将规模调整得到的 $\widetilde{w}_{\mathrm{eq}}$ 引入第二行，计算容度。

单层位势算子的可逆性

已知单层位势边界积分算子 $V : H^{-\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$ 第一有界(8.182)，第二 $H^{-\frac{1}{2}}(\Gamma)$-椭圆 ($d = 2$ 的证明见 Theorem 8.36，$d = 3$ 的证明见 Theorem 8.35)。那么可应用拉克斯一密格拉蒙定理 (Theorem 8.14) 证得单层位势 V 可逆，即 $V^{-1} : H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma)$ 有界，并且满足

$$\left\|V^{-1}\nu\right\|_{H^{-\frac{1}{2}}(\Gamma)} \leq \frac{1}{c_1^{\mathrm{V}}} \ \|\nu\|_{H^{\frac{1}{2}}(\Gamma)}, \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma).$$

在此基础上，双线性内积形式的单层位势 V

$$\begin{aligned}\langle V w, w_{\mathrm{eq}} \rangle_\Gamma &= \langle w, V w_{\mathrm{eq}} \rangle_\Gamma \\ &= \lambda \langle w, 1 \rangle_\Gamma = 0, \quad \forall w \in H_*^{\frac{1}{2}}(\Gamma).\end{aligned}$$

$$\begin{aligned}&\therefore V w \in H_*^{\frac{1}{2}}(\Gamma), \\ &H_*^{\frac{1}{2}}(\Gamma) := \left\{\nu \in H^{\frac{1}{2}}(\Gamma) : \langle \nu, w_{\mathrm{eq}} \rangle_\Gamma = 0\right\}.\end{aligned}$$

因此可见，单层位势 $V : H_*^{-\frac{1}{2}}(\Gamma) \mapsto H_*^{\frac{1}{2}}(\Gamma)$ 是一个同构 (isomorphism)[18]。

[18]同构的概念，可参考 (Muscat, 2014, pp.34, Definition 3.11)。

8.6.8 超奇异边界积分算子的椭圆性

第8.6.5节介绍了超奇异边界积分算子 D。由(8.203)可得，若引入特征解 (eigensolution) $u_0 \equiv 1, x \in \Gamma$，那么 $(Du_0)(x) = 0$。可见我们无法保证超奇异边界积分算子 D 在 $H^{\frac{1}{2}}(\Gamma)$ 中全部都是椭圆的，而只是半椭圆 (semi-ellipticity)的。在本节中，首先我们介绍半椭圆性。随后介绍替代方案：在 $H^{\frac{1}{2}}(\Gamma)$ 中构建满足椭圆特征的子空间，两种常见的子空间如下文所示。

超奇异边界积分算子的半椭圆性

Theorem 8.37 (超奇异边界积分算子的半椭圆性)**.** *超奇异边界积分算子 D 是半椭圆的，即*

$$\langle D\nu, \nu\rangle_\Gamma \geq c_1^D \left\|\nu\right\|^2_{H^{\frac{1}{2}}(\Gamma)}, \quad \forall \nu \in H_*^{\frac{1}{2}}(\Gamma).$$

证明. 1. 对于 $\nu \in H_*^{\frac{1}{2}}(\Gamma)$，考虑齐次偏微分方程的解 $u(x)$ 可以表示为如下双层位势

$$u(x) := -(Wu)(x), \quad x \in \Omega \cup \Omega c.$$

对于 $x \in \Gamma$，$u(x)$ 的迹算子和共法导数算子以伴随双层位势的形式表示如下

$$\begin{cases} \gamma_0^{\text{int}} u(x) = (1-\sigma(x))\,\nu(x) - (K\nu)(x), \\ \gamma_1^{\text{int}} u(x) = (D\nu)(x), \quad x \in \Gamma, \end{cases}$$

$$\begin{cases} \gamma_0^{\text{ext}} u(x) = -\sigma(x)\nu(x) - (K\nu)(x), \\ \gamma_1^{\text{ext}} u(x) = (D\nu)(x), \quad x \in \Gamma. \end{cases}$$

2. 由此可见，方程 $u = -Wu$ 构成内界狄利克雷边界值问题的唯一解

$$\begin{cases} -\Delta u(x) = 0 & x \in \Gamma, \\ \gamma_0^{\text{int}} u(x) = (1-\sigma(x))\,\nu(x) - (K\nu)(x) & x \in \Gamma. \end{cases}$$

由格林第一恒等式(8.6)可得，解 $u(x)$ 满足如下关系

$$\int_\Gamma \nabla u(x) \nabla w(x)\, dx = \langle \gamma_1^{\text{int}} u, \gamma_0^{\text{int}} w\rangle_\Gamma, \quad \forall w \in H^1(\Omega).$$

3. 设一个包围 Ω 的球体 $B_R(y_0)$，以 $y_0 \in \Omega$ 为圆心，以 $R > 2\,\text{diam}(\Omega)$ 为半径。即 $\Omega \subset B_R(y_0)$。那么 $u = -Wu$ 同时也是以下狄利克雷边界值问题的唯一解

$$\begin{cases} -\Delta u(x) = 0 & x \in B_R(y_0)\backslash\overline{\Omega}, \\ \gamma_0^{\text{ext}} u(x) = -\sigma(x)\nu(x) - (K\nu)(x) & x \in \Gamma = \partial\Omega, \\ \gamma_0 u(x) = \gamma_0^{\text{int}} u(x) + \gamma_0^{\text{ext}} u(x) = -(W\nu)(x) & x \in \partial B_r(y_0). \end{cases}$$

根据格林第一恒等式可得(8.6)可得，解 $u(x)$ 表示为如下双线性形式

$$\int_{B_R(y_0)\backslash\overline{\Omega}} \nabla u(x) \nabla w(x)\, dx = -\langle \gamma_1^{\text{ext}} u, \gamma_0^{\text{ext}} w\rangle_\Gamma + \langle \gamma_1 u, \gamma_0 w\rangle_{\partial B_r(y_0)}, \quad \forall w \in H^1(B_R(y_0)\backslash\overline{\Omega}).$$

(a) 对于 $x \in \Gamma$，根据定义有

$$u(x) = \frac{1}{2(d-1)\pi} \int_\Gamma \frac{(y-x, n_y)}{|x-y|^d} \nu(y)\, ds_y.$$

(b) 对于 $x \in \partial B_R(y_0)$，有

$$|u(x)| \le c_1(\nu)\, R^{1-d},$$
$$|\nabla u(x)| \le c_2(\nu) R^{-d}.$$

那么选取 $w = u = -W\nu$，取极限 $\lim R \to \infty$，可得外界域中的格林第一恒等式

$$\int_{\Omega^c} |\nabla u(x)|^2 \, dx = -\langle \gamma_1^{\text{ext}} u, \gamma_0^{\text{ext}} u \rangle_\Gamma.$$

4. 将内、外界域的格林第一恒等式相加求和，并考虑不同边界积分算子的跃动关系，我们有超奇异边界积分算子 D 的双线性形式

$$\begin{aligned}
\underbrace{\langle D\nu, \nu \rangle_\Gamma}_{=:\mathcal{C}} &= \langle \gamma_1^{\text{int}} u, \left[\gamma_0^{\text{int}} - \gamma_0^{\text{ext}} u\right] \rangle_\Gamma \\
&= \langle \gamma_1^{\text{int}} u, \gamma_0^{\text{int}} u \rangle_\Gamma - \langle \gamma_1^{\text{int}} u, \gamma_0^{\text{ext}} u \rangle_\Gamma \\
&= \int_\Omega |\nabla u(x)|^2 \, dx + \int_{\Omega^c} |\nabla u(x)|^2 \, dx \\
&= \underbrace{|u|^2_{H^1(\Omega)}}_{=:\mathcal{B}} + \underbrace{|u|^2_{H^1(\Omega^c)}}_{=:\mathcal{A}}.
\end{aligned}$$

(a) 来看 $\mathcal{A}$。外界域 Ω^c 中，双层位势 $u(x) = -(Wu)(x)$，取 $\lim |x| \to \infty$，有

$$c_1 \, \|u\|^2_{H^1(\Omega^c)} \le |u|^2_{H^1(\Omega^c)}.$$

(b) 来看 $\mathcal{B}$。

i. 由以下三点

- 对于 $\nu \in H^{\frac{1}{2}}_*(\Gamma)$，有自然密度 $w_{\text{eq}} \in H^{-\frac{1}{2}}(\Gamma)$,
- $V w_{\text{eq}} = 1$,
- 有界积分算子呈对称关系(8.220) 成立

可得

$$\begin{aligned}
\langle \gamma_0^{\text{int}} u, w_{\text{eq}} \rangle_\Gamma &= \langle \left(\frac{1}{2} I - K\right) \nu, w_{\text{eq}} \rangle_\Gamma \\
&= \langle \nu, w_{\text{eq}} \rangle_\Gamma - \langle \left(\frac{1}{2} I + K\right) \nu, w_{\text{eq}} \rangle_\Gamma \\
&= -\langle \left(\frac{1}{2} I + K\right) \nu, V^{-1} \rangle_\Gamma \\
&= -\langle V^{-1} \left(\frac{1}{2} I + K\right) \nu, 1 \rangle_\Gamma \\
&= -\langle \left(\frac{1}{2} I + K\right) V^{-1} \nu, 1 \rangle_\Gamma \\
&= -\langle V^{-1} \nu, \left(\frac{1}{2} I + K\right) \rangle_\Gamma \\
&= 0.
\end{aligned}$$

ii. 进而，根据索伯列夫空间中的范数等价定理(8.37)(Theorem 8.4) 可得 $H^1(\Omega)$ 空间中的一个等价范

$$\|u\|_{H^1_*(\Omega)} := \left\{ \left[\langle \gamma_0^{\text{int}} u, w_{\text{eq}} \rangle_\Gamma\right]^2 + \|\nabla u\|^2_{L^2(\Omega)} \right\}^{\frac{1}{2}}.$$

(c) 来看 $\mathcal{C}$。由迹定理和双层位势的跃动关系可得

$$\begin{aligned}\langle D\nu,\nu\rangle_\Gamma &\geq c\left\{\|u\|^2_{H^1(\Omega)}+\|u\|^2_{H^1(\Omega^c)}\right\}\\ &\geq \widetilde{c}\left\{\left\|\gamma_0^{\mathrm{int}}u\right\|^2_{H^{\frac12}(\Gamma)}+\left\|\gamma_0^{\mathrm{ext}}u\right\|^2_{H^{\frac12}(\Gamma)}\right\}\\ &\geq \frac12\widetilde{c}\left\|\gamma_0^{\mathrm{int}}u-\gamma_0^{\mathrm{ext}}u\right\|^2_{H^{\frac12}(\Gamma)}\\ &= c_1^{\mathrm{D}}\ \|\nu\|^2_{H^{\frac12}(\Gamma)},\quad \forall\,\nu\in H_*^{\frac12}(\Gamma).\end{aligned}$$

由此可证得超奇异边界积分算子 D 的 $H_*^{\frac12(\Gamma)}$-椭圆性。 □

椭圆性所处子空间的替代方案

Theorem 8.37对超奇异边界积分算子 D 的 $H_*^{\frac12(\Gamma)}$-椭圆性的证明，涉及到对解方程所处空间的限定：需要将其限定在一个更适宜的子空间之内——即正交于常数的子空间。显然，随着算子 D 双线性内积类型的不同，对应正交子空间的种类也有所不同。具体说来，根据索伯列夫空间中的范数等价定理(8.37)(Theorem 8.4)，可定义一个 $H^{\frac12}(\Gamma)$ 中的等价范

$$\|\nu\|_{H_*^{\frac12}(\Gamma)}:=\left\{[\langle\nu,w_{\mathrm{eq}}\rangle_\Gamma]^2+|\nu|^2_{H^{\frac12}(\Gamma)}\right\}^{\frac12},$$

其中 $w_{\mathrm{eq}}\in H^{-\frac12}(\Gamma)$ 由自然密度的定义式 (8.240)给出。那么

Corollary 8.9 (超奇异边界积分算子的半椭圆性)**.** *超奇异积分算子 D 是 $H^{\frac12}(\Gamma)$-半椭圆 (semi-elliptic)的，即*

$$\langle D\nu,\nu\rangle_\Gamma\geq \overline{c}_1^D\ |\nu|^2_{H^{\frac12}(\Gamma)},\quad \forall\nu\in H^{\frac12}(\Gamma). \tag{8.241}$$

Corollary 8.9中，对子空间 $H_*^{\frac12}(\Gamma)$ 的定义，涉及到边界积分方程(8.240)的唯一解，即自然密度 $w_{\mathrm{eq}}\in H^{-\frac12}(\Gamma)$。然而在实际计算过程中，该定义往往较难直接处理。有鉴于此，需要另外寻求替代方案。

常见的替代方案之一是定义另一个子空间 $H_{**}^{\frac12}(\Gamma)$，从而简化内积结构

$$H_{**}^{\frac12}(\Gamma):=\left\{\nu\in H^{\frac12}(\Gamma):\langle\nu,1\rangle_\Gamma=0\right\}.$$

则由(8.241)得，对于 $\nu\in H_{**}^{\frac12}(\Gamma)$，内积可以简化为

$$\begin{aligned}\langle D\nu,\nu\rangle_\Gamma &\geq \overline{c}_1^{\mathrm{D}}\ |\nu|^2_{H^{\frac12}(\Gamma)}\\ &= \overline{c}_1^{\mathrm{D}}\left\{|\nu|^2_{H^{\frac12}(\Gamma)}+[\langle\nu,1\rangle_\Gamma]^2\right\}\\ &\geq \widetilde{c}_1^{\mathrm{D}}\ \|\nu\|^2_{H^{\frac12}(\Gamma)}.\end{aligned} \tag{8.242}$$

若是采用子空间 $H_{**}^{\frac12}(\Gamma)$，我们也可以证明算子 D 的 $H_{**}^{\frac12}(\Gamma)$-椭圆性。(证明略)。

另一个替代方案是，考虑一个子空间 $\Gamma_0\subset\Gamma$。对于某一给定的 $\nu\in\widetilde{H}^{\frac12}(\Gamma_0)$，定义其延拓 $\widetilde{\nu}\in H^{\frac12}(\Gamma)$ 如下

$$\widetilde{\nu}(x)=\begin{cases}\nu(x) & x\in\Gamma_0,\\ 0 & \text{否则}.\end{cases}$$

根据范数等价定理，$H^{\frac12}(\Gamma)$ 上的等价范可定义为

$$\|w\|_{H^{\frac12}(\Gamma),\Gamma_0}:=\left\{\|w\|^2_{L^2(\Gamma\setminus\Gamma_0)}+|w|^2_{H^{\frac12}(\Gamma)}\right\}^{\frac12}.$$

那么，对于 $\nu \in \widetilde{H}^{\frac{1}{2}}(\Gamma_0)$，内积简化为

$$\begin{aligned}\langle D\nu, \nu\rangle_{\Gamma_0} &= \langle D\widetilde{\nu}, \widetilde{\nu}\rangle_{\Gamma} \\ &\geq \overline{c}_1^{\mathrm{D}} \left[\|\widetilde{\nu}\|^2_{L^2(\Gamma\setminus\Gamma_0)} + |\widetilde{\nu}|^2_{H^{\frac{1}{2}}(\Gamma)}\right] \\ &= \overline{c}_1^{\mathrm{D}} \ \|\widetilde{\nu}\|^2_{H^{\frac{1}{2}}(\Gamma),\Gamma_0} \\ &\geq \widehat{c}_1^{\mathrm{D}} \ \|\widetilde{\nu}\|^2_{H^{\frac{1}{2}}(\Gamma)} \\ &= \widehat{c}_1^{\mathrm{D}} \ \|\nu\|^2_{\widetilde{H}^{\frac{1}{2}}(\Gamma_0)},\end{aligned} \tag{8.243}$$

即算子 D 是 $\widetilde{H}^{\frac{1}{2}}(\Gamma_0)$-椭圆的。

8.6.9 Steklov-Poincaré 算子

在求解边界值问题的过程中，柯西数 $\gamma_0^{\mathrm{int}}u$，$\gamma_1^{\mathrm{int}}u$ 之间的关系变得非常重要。以齐次偏微分边界方程系统为例，即将 $f \equiv 0$ 引入(8.214)

$$\begin{pmatrix}\gamma_0^{\mathrm{int}}u(x) \\ \gamma_1^{\mathrm{int}}u(x)\end{pmatrix} = \underbrace{\begin{pmatrix}(1-\sigma)\,I - K & V \\ D & \sigma I + K'\end{pmatrix}}_{=:\mathcal{C}} \begin{pmatrix}\gamma_0^{\mathrm{int}}u(x) \\ \gamma_1^{\mathrm{int}}u(x)\end{pmatrix}. \tag{8.244}$$

单层位势算子 V 可逆。由(8.244)第一行可得

$$\begin{aligned}\gamma_0^{\mathrm{int}}u &= ((1-\sigma)\,I - K)\,\gamma_0^{\mathrm{int}}u + V\gamma_1^{\mathrm{int}}u, \\ \hookrightarrow \gamma_1^{\mathrm{int}}u &= \underbrace{V^{-1}\,(\sigma I + K)}_{=:S}\,\gamma_0^{\mathrm{int}}u.\end{aligned} \tag{8.245}$$

这构成一个由狄利克雷到诺依曼的映射 (Dirichlet to Neumann map)(Behrndt and ter Elst, 2015)。由此我们可以定义 Steklov-Poincaré 算子如下

$$S := V^{-1}\,(\sigma I + K) : H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma). \tag{8.246}$$

(8.245)代回第二行

$$\begin{aligned}\gamma_1^{\mathrm{int}}u(x) &= \left(D\gamma_0^{\mathrm{int}}u\right)(x) + (\sigma I + K')\,\gamma_1^{\mathrm{int}}u(x) \\ &= \underbrace{\left[D + (\sigma I + K')\,V^{-1}\,(\sigma I + K)\right]}_{=:S}\,\gamma_0^{\mathrm{int}}u(x), \quad x \in \Gamma,\end{aligned} \tag{8.247}$$

可见 Steklov-Poincaré 算子也可表示为如下对称形式

$$S := D + (\sigma I + K')\,V^{-1}\,(\sigma I + K) : H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma). \tag{8.248}$$

结合(8.245)和 (8.247)，狄利克雷到诺依曼的映射表达式表示为

$$\gamma_1^{\mathrm{int}}u(x) = \left(S\gamma_0^{\mathrm{int}}u\right)(x), \quad x \in \Gamma, \tag{8.249}$$

即狄利克雷到诺依曼的映射将某一给定的狄利克雷数 $\gamma_0^{\mathrm{int}}u \in H^{\frac{1}{2}}(\Gamma)$ 映射至对应的诺依曼数 $\gamma_1^{\mathrm{int}}u \in H^{-\frac{1}{2}}(\Gamma)$，其中和谐方程 (harmonic function)$u \in H^1(\Omega)$ 满足 $Lu \equiv 0$。

由逆单层位势算子 V^{-1} 的 $H^{\frac{1}{2}}(\Omega)$-椭圆属性可得

$$\begin{aligned}\langle S\nu, \nu\rangle_{\Gamma} &= \langle D\nu, \nu\rangle_{\Gamma} + \langle V^{-1}\,(\sigma I + K)\,\nu, (\sigma I + K)\,\nu\rangle_{\Gamma} \\ &\geq \langle D\nu, \nu\rangle_{\Gamma}, \quad \forall\,\nu \in H^{\frac{1}{2}}(\Gamma),\end{aligned} \tag{8.250}$$

可见 Steklov-Poincaré 算子继承了超奇异边界积分算子的椭圆特性，随着不同的子空间，可表示如下

$$\langle S\nu,\nu\rangle_\Gamma \geq c_1^D \left\|\nu\right\|^2_{H^{\frac{1}{2}}(\Gamma)}, \quad \forall\, \nu \in H_*^{\frac{1}{2}}(\Gamma). \tag{8.251}$$

$$\langle S\nu,\nu\rangle_\Gamma \geq \widetilde{c}_1^D \left\|\nu\right\|_{H^{\frac{1}{2}}(\Gamma)}, \quad \forall\, \nu \in H_{**}^{\frac{1}{2}}(\Gamma). \tag{8.252}$$

$$\langle S\nu,\nu\rangle_{\Gamma_0} \geq \hat{c}_1 \left\|\nu\right\|_{\widetilde{H}^{\frac{1}{2}}(\Gamma_0)}, \quad \forall\, \nu \in \widetilde{H}^{\frac{1}{2}}(\Gamma_0). \tag{8.253}$$

8.6.10　双层位势的收缩属性

在证得单层位势 V(以及相伴随的超奇异积分算子 D) 的椭圆性后，我们科进一步证明对应的双层位势 $\sigma I + K : H^{\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$ 也是椭圆的 (Steinbach and Wendland, 2001)。

由于算子 $V : H^{-\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$ 有界且 $H^{-\frac{1}{2}}(\Gamma)$ 椭圆，我们可以定义一个 $H^{\frac{1}{2}}(\Gamma)$ 中的等价范

$$\left\|u\right\|_{V^{-1}} := \sqrt{\langle V^{-1}u,u\rangle_\Gamma}, \quad \forall\, u \in H^{\frac{1}{2}}(\Gamma). \tag{8.254}$$

Theorem 8.38 (双层位势的收缩属性)**.** *对于 $u \in H_*^{\frac{1}{2}}(\Gamma)$ 我们有*

$$(1 - c_K)\left\|u\right\|_{V^{-1}} \leq \left\|(\sigma I + K)\, u\right\|_{V^{-1}} \leq c_K \left\|u\right\|_{V^{-1}}, \tag{8.255}$$

其中常数 c_K 满足

$$c_K = \frac{1}{2} + \sqrt{\frac{1}{4} - c_1^V\, c_1^D} < 1, \tag{8.256}$$

c_1^V, c_1^D 分别是单层位势 V 和超奇异边界积分算子 D 的椭圆常数。

证明. 由 Steklov-Poincaré 算子的对称表现式 (8.248)可得

$$\begin{aligned}\left\|(\sigma I + K)\, u\right\|^2_{V^{-1}} &= \langle V^{-1}(\sigma I + K), (\sigma I + K)\rangle_\Gamma \\ &= \underbrace{\langle Su,u\rangle_\Gamma}_{=:\mathcal{A}} - \underbrace{\langle Du,u\rangle_\Gamma}_{=:\mathcal{B}},\end{aligned}$$

1. $\mathcal{A}$。设算子 $A := JV^{-1} : H^{\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$ 自伴随且 $H^{\frac{1}{2}}(\Gamma)$-椭圆。考虑 $A = A^{\frac{1}{2}}A^{\frac{1}{2}}$，以及 S 的表现式(8.246)得

$$\begin{aligned}\mathcal{A} := \langle Su,u\rangle_\Gamma &= \langle V^{-1}(\sigma I + K)\, u, u\rangle_\Gamma \\ &= \langle JV^{-1}(\sigma I + K)\, u, u\rangle_\Gamma \\ &= \langle A^{\frac{1}{2}}(\sigma I + K)\, u, A^{\frac{1}{2}}u\rangle_\Gamma \\ &\leq \left\|A^{\frac{1}{2}}(\sigma I + K)\, u\right\|_{H^{\frac{1}{2}}(\Gamma)} \underbrace{\left\|A^{\frac{1}{2}}u\right\|_{H^{\frac{1}{2}}(\Gamma)}}_{=:\mathcal{A}_1},\end{aligned}$$

由于

$$\begin{aligned}\left\|A^{\frac{1}{2}}u\right\|_{H^{\frac{1}{2}}(\Gamma)} &= \langle A^{\frac{1}{2}}\nu, A^{\frac{1}{2}}\nu\rangle_{H^{\frac{1}{2}}(\Gamma)} \\ &= \langle JV^{-1}\nu,\nu\rangle_{H^{\frac{1}{2}}(\Gamma)} \\ &= \langle V^{-1}\nu,\nu\rangle_\Gamma \\ &= \left\|\nu\right\|_{V^{-1}},\end{aligned}$$

$$\hookrightarrow \mathcal{A}_1 = \|u\|_{V^{-1}}, \hookrightarrow \mathcal{A} = \langle Su, u\rangle_\Gamma \le \|(\sigma I + K)\, u\|_{V^{-1}} \, \|u\|_{V^{-1}}.$$

2. $\mathcal{B}$。已知对于 $u \in H_*^{\frac{1}{2}}(\Gamma)$，算子 D 椭圆。则逆单层位势 V^{-1} 的映射特征为

$$\begin{aligned}\mathcal{B} := \langle Du, u\rangle_\Gamma &\ge c_1^{\mathrm{D}} \, \|u\|^2_{H^{\frac{1}{2}}(\Gamma)} \\ &\ge c_1^{\mathrm{D}} c_1^{\mathrm{V}} \langle V^{-1}u, u\rangle_\Gamma \\ &= c_1^{\mathrm{D}} c_1^{\mathrm{V}} \, \|u\|_{V^{-1}}.\end{aligned}$$

3. $\therefore$

$$\begin{aligned}\|(\sigma I + K)\, u\|^2_{V^{-1}} &= \mathcal{A} - \mathcal{B} \\ &\le \underbrace{\|(\sigma I + K)\, u\|_{V^{-1}}}_{=:\mathcal{C}} \underbrace{\|u\|_{V^{-1}}}_{=:\mathcal{D}} - c_1^{\mathrm{D}} c_1^{\mathrm{V}} \, \|u\|_{V^{-1}}.\end{aligned}$$

定义

$$\begin{cases}\mathcal{C} := \|(\sigma I + K)\, u\|_{V^{-1}} & \mathcal{C} > 0, \\ \mathcal{D} := \|u\|_{V^{-1}} & \mathcal{D} > 0,\end{cases}$$

上式变为

$$\left(\frac{\mathcal{C}}{\mathcal{D}}\right)^2 - \left(\frac{\mathcal{C}}{\mathcal{D}}\right) + c_1^{\mathrm{D}} c_1^{\mathrm{V}} \le 0,$$

求解得

$$\frac{1}{2} - \sqrt{\frac{1}{4} - c_1^{\mathrm{D}} c_1^{\mathrm{V}}} \le \frac{\mathcal{C}}{\mathcal{D}} \le \frac{1}{2} + \sqrt{\frac{1}{4} - c_1^{\mathrm{D}} c_1^{\mathrm{V}}}$$

证毕。 □

双层位势 $\sigma I + K$ 的收缩特性 (contraction property)，尤其是 $H_*^{\frac{1}{2}}(\Gamma)$ 中的上界，可以延拓至 $H^{\frac{1}{2}}(\Gamma)$ 空间中，并且仍然成立。

Corollary 8.10 (双层位势收缩比率)**.** *对于 $u \in H^{\frac{1}{2}}(\Gamma)$，双层位势 $\sigma I + K$ 满足如下不等式*

$$\|(\sigma I + K)\|_{V^{-1}} \le c_K \, \|u\|_{V^{-1}}, \tag{8.257}$$

其中正常数 $0 \le c_K < 1$ 表示收缩比率 (contraction rate)，其值由(8.256)给出。

证明. 对于任一给定 $u \in H^{\frac{1}{2}}(\Gamma)$，我们有

$$u = \widetilde{u} + \frac{\langle u, w_{\mathrm{eq}}\rangle_\Gamma}{\langle 1, w_{\mathrm{eq}}\rangle_\Gamma} u_0, \quad \widetilde{u} \in H_*^{\frac{1}{2}}(\Gamma), u_0 \equiv 1.$$

1. 已知 $(\sigma I + K)\, u_0 = 0$，代入 Theorem 8.38我们有

$$\|(\sigma I + K)\, u\|_{V^{-1}} = \|(\sigma I + K)\, \widetilde{u}\|_{V^{-1}} \le c_K \, \|\widetilde{u}\|_{V^{-1}}.$$

2. 此外我们有

$$\|u\|^2_{V^{-1}} = \|\widetilde{u}\|^2_{V^{-1}} + \frac{[\langle u, w_{\mathrm{eq}}\rangle_\Gamma]^2}{\langle 1, w_{\mathrm{eq}}\rangle_\Gamma} \ge \|\widetilde{u}\|_{V^{-1}}.$$

□

Corollary 8.11 (位移双层位势的收缩比率). 对于 $\nu \in H_*^{\frac{1}{2}}(\Gamma)$，以下不等式成立

$$(1-c_K)\,\|\nu\|_{V^{-1}} \le \|[(1-\sigma)-K]\,\nu\|_{V^{-1}} \le c_K\,\|\nu\|_{V^{-1}}\,, \tag{8.258}$$

其中收缩比值 c_K 由(8.256)给出。

证明. 1. 由三角不等式可得

$$\begin{aligned}\|\nu\|_{V^{-1}} &= \|[(1-\sigma)\,I-K]\,\nu+(\sigma I+K)\,\nu\|_{V^{-1}}\\ &\le \|[(1-\sigma)\,I-K]\,\nu\|_{V^{-1}}+\|(\sigma I+K)\,\nu\|_{V^{-1}}\\ &\le \|[(1-\sigma)\,I-K]\,\nu\|_{V^{-1}}+c_K\,\|\nu\|_{V^{-1}}\,,\end{aligned}$$

$$\hookrightarrow (1-c_K)\,\|\nu\|_{V^{-1}} \le \|[(1-\sigma)\,I-K]\,\nu\|_{V^{-1}}\,.$$

2. 同时使用 Steklov-Poincaré 算子的两种表现式 (8.246)，(8.248)可得

$$\begin{aligned}\|[(1-\sigma)\,I-K]\,\nu\|_{V^{-1}}^2 &= \|[I-(\sigma I+K)]\,\nu\|_{V^{-1}}^2\\ &= \|\nu\|_{V^{-1}}^2+\|\sigma I+K\|_{V^{-1}}^2-2\langle V^{-1}\,(\sigma I+K)\,\nu,\nu\rangle_\Gamma\\ &= \|\nu\|_{V^{-1}}^2+\|\sigma I+K\|_{V^{-1}}^2-2\langle S\nu,\nu\rangle_\Gamma\\ &= \|\nu\|_{V^{-1}}^2-\|\sigma I+K\|_{V^{-1}}^2-2\langle D\nu,\nu\rangle_\Gamma\\ &\le \left[1-(1-c_K)^2-2c_1^{\mathrm{D}}c_1^{\mathrm{V}}\right]\,\|\nu\|_{V^{-1}}^2\\ &= c_K^2\,\|\nu\|_{V^{-1}}^2\,.\end{aligned}$$

$$\hookrightarrow \|[(1-\sigma)-K]\,\nu\|_{V^{-1}} \le c_K\,\|\nu\|_{V^{-1}}\,.$$

□

现在来看子空间 $H_*^{\frac{1}{2}}(\Gamma)$ 的情况，定义如 (8.225)。根据单层位势的性质 $V: H_*^{-\frac{1}{2}}(\Gamma)\mapsto H_*^{\frac{1}{2}}(\Gamma)$，我们可以将双层位势 $\sigma I+K: H^{\frac{1}{2}}(\Gamma)\mapsto H^{\frac{1}{2}}(\Gamma)$ 的测度 Theorem 8.38, (8.255)，转变为伴随双层位势 $\sigma I+K': H^{-\frac{1}{2}}(\Gamma)\mapsto H^{-\frac{1}{2}}(\Gamma)$ 的测度。

Corollary 8.12 (伴随双层位势的收缩属性). 对于 $w\in H_*^{-\frac{1}{2}}(\Gamma)$ 的伴随双层位势，满足

$$(1-c_K)\,\|w\|_{V^{-1}} \le \|(\sigma I+K')\,w\|_V \le c_K\,\|w\|_V\,. \tag{8.259}$$

证明. 1. 已知对于 $w\in H_*^{\frac{1}{2}}(\Gamma)$，存在唯一确定的 $\nu\in H_*^{\frac{1}{2}}(\Gamma)$，满足 $\nu=Vw$ 或是 $\nu=V^{-1}\nu$。由有界积分算子的对称性(8.220)可得

$$\begin{aligned}\|(\sigma I+K')\,w\|_V^2 &= \langle V\,(\sigma I+K')\,V^{-1}\nu,(\sigma I+K')\,V^{-1}\nu\rangle_\Gamma\\ &= \langle V^{-1}\,(\sigma I+K)\,\nu,(\sigma I+K)\,\nu\rangle_\Gamma\\ &= \|(\sigma I+K)\,\nu\|_{V^{-1}}\,.\end{aligned}$$

2.

$$\|w\|_V^2 = \langle Vw,w\rangle_\Gamma = \langle V^{-1}\nu,\nu\rangle_\Gamma = \|\nu\|_{V^{-1}}^2\,.$$

3. 代入 Corollary 8.11，可证。

□

与在双层位势 $(\sigma I + K)$ 延拓收缩属性 (Corollary 8.10) 相类似，我们也可以将伴随双层位势 $(\sigma I + K')$ 的收缩属性延拓到 $H^{-\frac{1}{2}}(\Gamma)$ 中。

Corollary 8.13 (伴随双层位势收缩比率)**.** 对于 $w \in H^{-\frac{1}{2}}(\Gamma)$，伴随双层位势 $(\sigma I + K')$ 满足如下不等式

$$\|(\sigma I + K')\, w\|_V \le c_K \, \|w\|_V \,. \tag{8.260}$$

证明. 略。 □

同样地

Corollary 8.14 (位移伴随双层位势收缩比率)**.** 对于 $w \in H_*^{-\frac{1}{2}}(\Gamma)$，位移伴随双层位势 $(1-\sigma)\, I - K'$ 满足如下不等式

$$(1 - c_K)\, \|w\|_V \le \|[(1-\delta)\, I - K']\, w\|_V \le c_K \, \|w\|_V \,, \tag{8.261}$$

收缩比率 c_K 的值由(8.256)给出。

证明. 略。 □

8.6.11 边界积分算子的映射特性

至今我们分析的一系列边界积分算子的特性，都是从牛顿位势的映射 $\widetilde{N}_0 : \widetilde{H}^{-1}(\Omega) \mapsto H^1(\Omega)$，以及逆定理和对偶配对而展开的。对于利普希茨域，我们可以得到更为通用的结论。

Theorem 8.39 (牛顿位势在利普希茨域中的映射特性)**.** 对于 $s \in [-2, 0]$，牛顿位势 $\widetilde{N}_0 : \widetilde{H}^s(\Omega) \mapsto H^{s+1}(\Omega)$ 是一个连续映射，即

$$\left\|\widetilde{N}_0 f\right\|_{H^{s+2}(\Omega)} \le c \, \|f\|_{\widetilde{H}^s(\Omega)} \,, \quad \forall\, f \in \widetilde{H}^s(\Omega).$$

证明. 1. 设 $s \in [-1, 0]$。牛顿位势 $f \in H^s(\Omega)$ 的定义如(8.160)。$\widetilde{f} \in H^s(\mathbb{R}^d)$ 是 f 的延拓，我们有

$$\begin{aligned}
\left\|\widetilde{f}\right\|_{H^s(\mathbb{R}^d)} &= \sup_{0 \ne \nu \in H^{-s}(\mathbb{R}^d)} \frac{\langle \widetilde{f}, \nu \rangle_{\mathbb{R}^d}}{\|\nu\|_{H^{-s}(\mathbb{R}^d)}} \\
&\le \sup_{0 \ne \nu \in H^{-s}(\Omega)} \frac{\langle f, \nu \rangle_{\Omega}}{\|\nu\|_{H^{-s}(\Omega)}} \\
&= \|f\|_{H^s(\Omega)} \,,
\end{aligned}$$

由牛顿位势算子的映射定理8.31得

$$\left\|\widetilde{N}_0 f\right\|_{H^{s+2}(\Omega)} \le c \, \|f\|_{\widetilde{H}^s(\Omega)} \,, \quad \forall\, f \in \widetilde{H}^s(\Omega).$$

2. 设 $s \in [-2, -1)$。由于牛顿位势 $\widetilde{N}_0$ 是自伴随算子，根据配偶配对我们有

$$\begin{aligned}
\left\|\widetilde{N}_0 f\right\|_{H^{s+2}(\Omega)} &= \sup_{0 \ne g \in \widetilde{H}^{-2-s}(\Omega)} \frac{\langle \widetilde{N}_0 f, g \rangle_{\Omega}}{\|g\|_{\widetilde{H}^{-2-s}(\Omega)}} \\
&= \sup_{0 \ne g \in \widetilde{H}^{-2-s}(\Omega)} \frac{\langle f, \widetilde{N}_0 g \rangle_{\Gamma}}{\|g\|_{\widetilde{H}^{-2-s}(\Omega)}} \\
&\le \|f\|_{\widetilde{H}^s(\Omega)} \sup_{0 \ne g \in \widetilde{H}^{-2-s}(\Omega)} \frac{\left\|\widetilde{N}_0 g\right\|_{H^{-s}(\Omega)}}{\|g\|_{\widetilde{H}^{-2-s}(\Omega)}} \\
&\le c \, \|f\|_{\widetilde{H}^s(\Omega)}
\end{aligned}$$

□

利用牛顿位势在利普希茨域中的映射特性 (定理8.39)，我们可以得到单层位势 $\widetilde{V}$(8.180)的映射特性。进而，利用 $\widetilde{V}$ 的迹，可得边界积分算子 $V := \gamma_0^{\text{int}}\widetilde{V}$ 的映射特性。

Theorem 8.40 (单层位势算子的映射特性)**.** *对于 $|s| < \frac{1}{2}$，单层位势 $V : H^{-\frac{1}{2}+s} \mapsto H^{\frac{1}{2}+s}$ 有界，即*

$$\|Vw\|_{H^{\frac{1}{2}+s}(\Gamma)} \le c \|w\|_{H^{-\frac{1}{2}+s}(\Gamma)}, \quad \forall w \in H^{-\frac{1}{2}+s}(\Gamma).$$

证明. 1. 对于 $\varphi \in C^\infty(\Omega)$，有

$$\begin{aligned}
\langle \widetilde{V}w, \varphi\rangle_\Omega &= \int_\Omega \varphi(x) \int_\Gamma U^*(x,y)w(y)\, ds_y dx \\
&= \int_\Gamma w(y) \int_\Gamma U^*(x,y)\varphi(x)\, dx ds_y \\
&= \langle w, \gamma_0^{\text{int}} \widetilde{N}_0 \varphi\rangle_\Gamma \\
&\le \|w\|_{H^{-\frac{1}{2}+s}(\Gamma)} \left\|\gamma_0^{\text{int}} \widetilde{N}_0 \varphi\right\|_{H^{\frac{1}{2}-s}(\Gamma)}.
\end{aligned}$$

根据迹定理8.10可得

$$\begin{aligned}
\langle \widetilde{V}w, \varphi\rangle_\Omega &\le \|w\|_{H^{-\frac{1}{2}+s}(\Gamma)} \left\|\gamma_0^{\text{int}} \widetilde{N}_0 \varphi\right\|_{H^{\frac{1}{2}-s}(\Gamma)} \\
&\le c_T \|w\|_{H^{-\frac{1}{2}+s}(\Gamma)} \left\|\widetilde{N}_0 \varphi\right\|_{H^{1-s}(\Gamma)}.
\end{aligned}$$

2. 利用 Theorem 8.39，可得

$$\begin{aligned}
\langle \widetilde{V}w, \varphi\rangle_\Omega &\le c_T \|w\|_{H^{-\frac{1}{2}+s}(\Gamma)} \left\|\widetilde{N}_0 \varphi\right\|_{H^{1-s}(\Gamma)} \\
&\le c \|w\|_{H^{-\frac{1}{2}+s}(\Gamma)} \|\varphi\|_{\widetilde{H}^{-1-s}(\Omega)}, \forall \varphi \in C^\infty(\Omega).
\end{aligned}$$

3. 将 w 视为密度方程，根据定义可得 $\widetilde{V}w \in H^{1+s}(\Omega)$。对 $\widetilde{V}w$ 取迹，在满足 $\frac{1}{2}+s>0$ 的情况下，我们有 $Vw : \gamma_0^{\text{int}}\widetilde{V}w \in H^{\frac{1}{2}}(\Gamma)$。

□

对于利普希茨域 Ω 的情况，可由 Theorem 8.40求得全部边界积分算子的映射特性。

Theorem 8.41 (边界积分算子的映射特性)**.** *设利普希茨域 Ω，对应边界 $\Gamma := \partial\Omega$。则对于 $S \in \left[-\frac{1}{2}, \frac{1}{2}\right]$，以下边界积分算子都有界，且满足映射关系 (Costabel, 1988)*

$$\begin{aligned}
V &: H^{-\frac{1}{2}+s}(\Gamma) \mapsto H^{\frac{1}{2}+s}(\Gamma), \\
K &: H^{\frac{1}{2}+s}(\Gamma) \mapsto H^{\frac{1}{2}+s}(\Gamma), \\
K' &: H^{-\frac{1}{2}+s}(\Gamma) \mapsto H^{-\frac{1}{2}+s}(\Gamma), \\
D &: H^{\frac{1}{2}+s}(\Gamma) \mapsto H^{-\frac{1}{2}+s}(\Gamma).
\end{aligned}$$

证明. 1. $|s| < \frac{1}{2}$ 时单层位势 V 的映射特性可由 Theorem 8.40证得。$|s| = \frac{1}{2}$ 时 V 的映射特性，见Verchota (1984); McLean (2000)。

2. 在此基础上，K, K', D 的映射特性可由共法导数算子推得。已知单层位势 $u(x) = \widetilde{V}w(x), x \in \Omega$ 构成齐次偏微分方程的解，对应狄利克雷数 $\gamma_0^{\text{int}}u(x) = (Vw)(x), x \in \Gamma$。

   已知单层位势 $V : L^2(\Gamma) \mapsto L^1(\Gamma)$ 是连续的，根据 Theorem 8.24我们有

$$\left\|\gamma_1^{\text{int}} u\right\|_{L^2(\Gamma)} \le c \|Vw\|_{H^1(\Gamma)} \le \widetilde{c} \|w\|_{L^2(\Gamma)},$$

进而可证得伴随双层位势算子 $\gamma_1^{\text{int}}\widetilde{V}=\sigma I+K':L^2(\Gamma)\mapsto L^2(\Gamma)$ 的连续性。另一方面，由对偶配对可得

$$\left\|\gamma_1^{\text{int}}u\right\|_{L^2(\Gamma)}=\sup_{0\neq\varphi\in H^1(\Gamma)}\frac{\langle\gamma_1^{\text{int}},\varphi\rangle_\Gamma}{\|\varphi\|_{H^1(\Gamma)}}.$$

对于任一 $\varphi\in H^1(\Gamma)$，狄利克雷边界值问题

$$\begin{cases}L\nu(x)=0 & x\in\Gamma,\\ \gamma_0^{\text{int}}\nu(x)=\varphi(x) & x\in\Gamma\end{cases}$$

有唯一解 $\nu\in H^{\frac{3}{2}}(\Omega)$。在此基础上，由 Theorem 8.24可得

$$\left\|\gamma_1^{\text{int}}\nu\right\|_{L^2(\Gamma)}\leq c\left\|\gamma_0^{\text{int}}\nu\right\|_{H^1(\Gamma)}\leq c\left\|\varphi\right\|_{H^1(\Gamma)}.$$

既然 $u=\widetilde{V}w$ 和 ν 都是齐次偏微分方程的解，连续两次使用格林第一恒等式(8.6)有

$$\begin{aligned}&\langle\gamma_1^{\text{int}}u,\varphi\rangle_\Gamma=a\left(u,\nu\right)\\&=a\left(\nu,u\right)=\langle\gamma_1^{\text{int}}\nu,\gamma_0^{\text{int}}u\rangle_\Gamma\\&\leq\left\|\gamma_1^{\text{int}}\nu\right\|_{L^2(\Gamma)}\left\|\gamma_0^{\text{int}}u\right\|_{L^2(\Gamma)}\\&\leq c\left\|\varphi\right\|_{H^1(\Gamma)}\left\|Vw\right\|_{L^2(\Gamma)}.\end{aligned}$$

可见，由单层位势 $V:H^{-1}(\Gamma)\mapsto L^2(\Gamma)$ 的连续性，我们有

$$\left\|\gamma_1^{\text{int}}u\right\|_{H^{-1}(\Gamma)}\leq c\left\|Vw\right\|_{L^2(\Gamma)}\leq\widetilde{c}\|w\|_{H^{-1}(\Gamma)},$$

进而可得共法导数算子 $\gamma_1^{\text{int}}\widetilde{V}:\sigma I+K':H^{-1}(\Gamma)\mapsto H^{-1}(\Gamma)$ 的连续性。在此基础上，若满足 $s\leq\frac{1}{2}$，那么利用插值定理可得 $K':H^{-\frac{1}{2}+s}(\Gamma)\mapsto H^{-\frac{1}{2}+s}(\Gamma)$。

3. 由于

$$\begin{aligned}\|K\nu\|_{H^{\frac{1}{2}+s}(\Gamma)}&=\sup_{0\neq w\in H^{-\frac{1}{2}-s}(\Gamma)}\frac{\langle K\nu,w\rangle_\Gamma}{\|w\|_{H^{-\frac{1}{2}-s}(\Gamma)}}\\&=\sup_{0\neq w\in H^{-\frac{1}{2}-s}(\Gamma)}\frac{\langle\nu,K'w\rangle_\Gamma}{\|w\|_{H^{-\frac{1}{2}-s}(\Gamma)}}\\&=\|\nu\|_{H^{\frac{1}{2}+s}(\Gamma)}\sup_{0\neq w\in H^{-\frac{1}{2}-s}(\Gamma)}\frac{\|K'w\|_{H^{-\frac{1}{2}-s}(\Gamma)}}{\|w\|_{H^{-\frac{1}{2}-s}(\Gamma)}}\\&\leq c\left\|\nu\right\|_{H^{\frac{1}{2}+s}(\Gamma)},\end{aligned}$$

进而可得对于 $|s|\leq\frac{1}{2}$，有 $K:H^{\frac{1}{2}+s}(\Gamma)\mapsto H^{\frac{1}{2}+s}(\Gamma)$。

4. 由双层位势的定义 $u(x)=\left(Wu\right)(x)$ 可得，利用 Theorem 8.24，有

$$\begin{aligned}\|D\nu\|_{L^2(\Gamma)}&=\left\|\gamma_1^{\text{int}}u\right\|_{L^2(\Gamma)}\\&\leq c\left\|\gamma_0^{\text{int}}u\right\|_{H^1(\Gamma)}\\&=c\left\|\left[\left(\sigma-1\right)I+K\right]\nu\right\|_{H^1(\Gamma)},\end{aligned}$$

进而可得 $D:H^1(\Gamma)\mapsto L^2(\Gamma)$。进一步使用配偶配对以及插值法，得 $D:H^{\frac{1}{2}}(\Gamma)\mapsto H^{-\frac{1}{2}}(\Gamma)$。

□

如果有界域 $\Omega \subset \mathbb{R}^d$ 的边界 $\gamma = \partial\Omega$ 是分段平滑的，那么即使 $|s|$ 延展到 $\frac{1}{2}$ 之外，Theorem 8.41也可能成立。举例来说，若 $\Omega \subset \mathbb{R}^2$ 是个有界的多边形，有 J 个角，内夹角 α_j，可定义

$$\sigma_0 := \min_{j=1,\dots,J} \left\{ \min \left[\frac{\pi}{\alpha_j}, \frac{\pi}{2\pi - \alpha_j} \right] \right\},$$

此时 Theorem 8.41对于所有 $|s| < \sigma_0$ 均成立 (Costabel and Stephan, 1985)；如果边界 Γ 是 C^∞ 的，则 Theorem 8.41对于所有 $s \in \mathbb{R}$ 都成立。

8.7　边界积分方程

在对一系列边界积分算子有了初步了解后 (第8.1) 节，本届来分析边界积分方程 (BIEs, boundary integral equations)。具体来说，考虑由标量齐次偏微分方程表示的边界值问题

$$(Lu)(x) = 0, \quad x \in \Gamma, \tag{8.262}$$

其中 L 是一个椭圆、自伴随的二阶偏微分算子。Ω 是一个有界的单连通域 (simply connected domain)，其利普希茨边界定义为 $\Gamma := \partial\Omega$。[19]

对于某个 $\widetilde{x} \in \Omega$，齐次偏微分方程(8.262)的解 $u(\widetilde{x})$ 可由表现式(8.151)给出

$$u(\widetilde{x}) = \int_\Gamma U^*(\widetilde{x}, y)\gamma_1^{\text{int}} u(y)\, ds_y - \int_\Gamma \gamma_{1,y}^{\text{int}} U^*(\widetilde{x}, y)\gamma_0^{\text{int}} u(y) ds_y. \tag{8.263}$$

可见为了通过(8.263)求解(8.262)，一方面需要求得基本解 $U^*(\widetilde{x}, y)$，另一方面我们需要完整柯西数 $\left[\gamma_0^{\text{int}} u(\widetilde{x}), \gamma_1^{\text{int}} u(\widetilde{x})\right]$。前者见第8.5节。就后者而言，计算完整柯西数所需的全部信息，可通过构建适宜的边界积分方程求得。以下述边界积分方程系统为例，

$$\begin{pmatrix} \gamma_0^{\text{int}} u \\ \gamma_1^{\text{int}} u \end{pmatrix} = \begin{pmatrix} (1-\sigma) I - K & V \\ D & \sigma I + K' \end{pmatrix} \begin{pmatrix} \gamma_0^{\text{int}} u \\ \gamma_1^{\text{int}} u \end{pmatrix}. \tag{8.264}$$

上式中，如果所有边界积分算子的密度方程都恰好是柯西数 $\left[\gamma_0^{\text{int}} u(x), \gamma_1^{\text{int}} u(x)\right], x \in \Gamma$，则我们称这种求解思路为直接法。

与之相对应，另一种求解思路为间接法，是指用适宜的位势算子求解边界值问题。例如，对于给定的 $\widetilde{x} \in \Omega$，齐次偏微分方程(8.262)的解 $u(\widetilde{x})$，可以由单层位势求解而得

$$u(\widetilde{x}) = \int_\Gamma U^*(\widetilde{x}, y) w(y)\, ds_y, \quad \widetilde{x} \in \Omega, \tag{8.265}$$

或由双层位势求解而得

$$u(\widetilde{x}) = -\int_\Gamma \gamma_{1,y}^{\text{int}} U^*(\widetilde{x}, y)\nu(y)\, ds_y, \quad \widetilde{x} \in \Omega, \tag{8.266}$$

需要指出的是，间接法中的密度方程 $w(y)$ 和 $\nu(y)$ 往往并无明确的经济学含义。

接下来我们将分别介绍几种不同的边界积分方程，以及如何用这些边界积分方程生成对应的柯西数，来描述带有不同边界条件的边界值问题。

8.7.1　狄利克雷边界值问题

考虑如下狄利克雷边界值问题

$$\begin{aligned} (Lu)(x) &= 0, \quad x \in \Omega, \\ \gamma_0^{\text{int}} u(x) &= g(x), \quad x \in \Gamma. \end{aligned} \tag{8.267}$$

[19]在这里我们以齐次泊松方程，即拉普拉斯方程为例来说明。求解非齐次偏微分方程，可引入牛顿位势，将非齐次偏微分方程中的全部牛顿位势化简为表面积位势，如见Jung and Steinbach (2002); Of et al. (2010)。

直接法求解狄利克雷边界值问题

使用直接法求解(8.267)，$u(x)$ 的表现式由(8.263)改写为

$$u(\widetilde{x}) = \int_\Gamma U^*(\widetilde{x}, y)\gamma_1^{\text{int}} u(y)\, ds_y - \int_\Gamma \gamma_{1,y}^{\text{int}} U^*(\widetilde{x}, y) g(y) ds_y, \quad \widetilde{x} \in \Omega. \tag{8.268}$$

求解(8.268)需要首先计算未知的诺依曼数 $\gamma_1^{\text{int}} u \in H^{-\frac{1}{2}}(\Gamma)$。由边界积分方程系统 (卡尔德隆系统)(8.264)的第一行我们有

$$\left(V\gamma_1^{\text{int}} u\right)(x) = \sigma(x) g(x) + (Kg)(x), \quad x \in \Gamma, \tag{8.269}$$

我们称之为第一类弗雷德霍姆边界积分方程 (the first kind of Fredholm boundary integral function)(Atkinson, 1996, 1997)。

(8.269)的解 $u(x)$，其唯一性可由以下方式证得

1. 单层位势算子 $V : H^{-\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$ 有界(8.182)，
2. V 是 $H^{-\frac{1}{2}}(\Gamma)$-椭圆的。$d = 3$ 的证明见 Theorem 8.35；$d = 2$ 的证明见 Theorem 8.36。
3. 已知 V 有界且椭圆，由拉克斯一密格拉蒙定理 Theorem 8.14可得 $u(x)$ 的唯一可解性。

唯一可解的 $\gamma_1^{\text{int}} u(x) \in H^{-\frac{1}{2}}(\Gamma)$ 满足

$$\begin{aligned}\left\|\gamma_1^{\text{int}} u\right\|_{H^{-\frac{1}{2}}(\Gamma)} &\le \frac{1}{c_1^V} \left\|(\sigma I + K) g\right\|_{H^{\frac{1}{2}}(\Gamma)} \\ &\le \frac{c_2^W}{c_1^V} \|g\|_{H^{\frac{1}{2}}(\Gamma)}.\end{aligned}$$

由于边界积分方程(8.269)存在于 $H^{\frac{1}{2}}(\Gamma)$ 空间中，则我们有

$$\begin{aligned}0 &= \left\|V\gamma_1^{\text{int}} u - (\sigma I + K) g\right\|_{H^{\frac{1}{2}}(\Gamma)} \\ &= \sup_{0 \ne \tau \in H^{-\frac{1}{2}}(\Gamma)} \frac{\langle V\gamma_1^{\text{int}} u - (\sigma I + K) g, \tau\rangle_\Gamma}{\|\tau\|_{H^{-\frac{1}{2}}(\Gamma)}},\end{aligned}$$

由此可见，除了直接求解(8.269)之外，我们也可以考虑如下等价的变分问题：寻找解 $\gamma_1^{\text{int}} u \in H^{-\frac{1}{2}}(\Gamma)$，使其满足

$$\begin{aligned}\langle V\gamma_1^{\text{int}} u, \tau\rangle_\Gamma &= \langle (\sigma I + K) g, \tau\rangle_\Gamma \\ &= \langle \left(\frac{1}{2} I + K\right) g, \tau\rangle_\Gamma, \quad \forall \tau \in H^{-\frac{1}{2}}(\Gamma),\end{aligned} \tag{8.270}$$

其中第二行等式是由于，根据(8.188)，对于几乎所有 $x \in \Gamma$ 都有 $\sigma(x) = \frac{1}{2}$。

除了根据卡尔德隆系统(8.264)的第一行，构建(8.269)来求解狄利克雷边界值之外，另一种方法是根据第二行计算未知的诺依曼数 $\gamma_1^{\text{int}} u \in H^{-\frac{1}{2}}(\Gamma)$，即

$$\left[(1 - \sigma) I - K'\right] \gamma_1^{\text{int}} u(x) = (Dg)(x), \quad x \in \Gamma, \tag{8.271}$$

我们称之为第二类弗雷德霍姆边界积分方程 (the second kind of Fredholm boundary integral function)(Atkinson, 1996, 1997)。方程的解 $\gamma_1^{\text{int}} u(x)$ 由诺依曼级数 (Neumann series)给出

$$\gamma_1^{\text{int}} u(x) = \sum_{\ell=0}^{\infty} (\sigma I + K')^\ell (Dg)(x), \quad x \in \Gamma. \tag{8.272}$$

已知伴随双层位势 $(\sigma I + K')$ 的收缩特性可由单层位势 V 中等价范的形式 $\|\cdot\|_V$ 来表示，见 Corollary 8.13式(8.260)，则诺依曼级数(8.272)在 $H^{-\frac{1}{2}}(\Gamma)$ 中收敛。

间接法求解狄利克雷边界值问题

利用间接法，通过单层位势 V(8.265)求解狄利克雷边界值问题 (齐次偏微分方程)，前提是需要计算未知的密度方程 $w \in H^{-\frac{1}{2}}(\Gamma)$，其关键是求解如下边界积分方程

$$(Vw)(x) = g(x), \quad x \in \Gamma. \tag{8.273}$$

将间接法之一(8.273)，与直接法之一 (8.269)相比，只是在 RHS 的定义上有所不同。换句话说，(8.273)解的唯一存在性，可由(8.269)解的唯一存在性而证得。

除了利用单层位势 V，也可以利用双层位势 $(\sigma I + K)$(8.266)。根据双层位势的跃动关系 Lemma 8.25式(8.192)，可建立如下边界积分方程

$$[1 - \sigma(x)]\,\nu(x) - (K\nu)(x) = g(x), \quad x \in \Gamma, \tag{8.274}$$

其密度方程 $\nu(x) \in H^{\frac{1}{2}}(\Gamma)$ 由诺依曼级数 (Neumann series)算得

$$\nu(x) = \sum_{\ell=0}^{\infty} (\sigma I + K')^{\ell}\, g(x), \quad \forall\, x \in \Gamma. \tag{8.275}$$

类似地，已知双层位势 $(\sigma I + K)$ 的收缩属性可由逆单层位势 V^{-1} 中等价范的形式 $\|\cdot\|_{V^{-1}}$ 来表示，见 Corollary 8.10式(8.257)，则诺依曼级数(8.275)在 $H^{\frac{1}{2}}(\Gamma)$ 中收敛，进而

$$\begin{aligned} 0 &= \|(1-\sigma)\,\nu - K\nu - g\|_{H^{\frac{1}{2}}(\Gamma)} \\ &= \sup_{0 \neq \tau \in H^{-\frac{1}{2}}(\Gamma)} \frac{\langle \frac{1}{2}\nu - K\nu - g, \tau\rangle_{\Gamma}}{\|\tau\|_{H^{-\frac{1}{2}}(\Gamma)}}, \end{aligned}$$

其中根据(8.188)，对于几乎所有 $x \in \Gamma$ 都有 $\sigma(x) = \frac{1}{2}$。

因此除了直接求解边界积分方程(8.273)，也可以考虑如下等价变分问题：寻找解 $\nu \in H^{\frac{1}{2}}(\Gamma)$，使满足

$$\langle \left(\frac{1}{2}I - K\right)\nu, \tau\rangle_{\Gamma} = \langle g, \tau\rangle_{\Gamma}, \quad \forall\, \tau \in H^{-\frac{1}{2}}(\Gamma). \tag{8.276}$$

Lemma 8.32 (变分问题的稳定条件)**.** *变分问题(8.276)有稳定条件*

$$c_S\, \|\nu\|_{H^{\frac{1}{2}}(\Gamma)} \leq \sup_{0 \neq \tau \in H^{-\frac{1}{2}}(\Gamma)} \frac{\langle \left(\frac{1}{2}I - K\right)\nu, \tau\rangle_{\Gamma}}{\|\tau\|_{H^{-\frac{1}{2}}(\Gamma)}}, \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma), \tag{8.277}$$

其中 c_S 是个正常数。

证明. 设任一 $\nu \in H^{\frac{1}{2}}(\Gamma)$，定义 $\tau_\nu := V^{-1}\nu \in H^{-\frac{1}{2}}(\Gamma)$，则有

$$\|\tau_\nu\|_{H^{-\frac{1}{2}}(\Gamma)} = \left\|V^{-1}\nu\right\|_{H^{-\frac{1}{2}}(\Gamma)} \leq \frac{1}{c_1^V}\, \|\nu\|_{H^{\frac{1}{2}}(\Gamma)}.$$

由双层位势的收缩比率 Corollary 8.10式(8.257)，以及单层位势 V 的映射特性 Theorem 8.41，

我们有

$$
\begin{aligned}
\langle\left(\frac{1}{2}I-K\right)\nu,\tau_\nu\rangle_\Gamma &= \langle\left(\frac{1}{2}I-K\right)\nu,V^{-1}\nu\rangle_\Gamma \\
&= \langle V^{-1}\nu,\nu\rangle_\Gamma - \langle V^{-1}\left(\frac{1}{2}I+K\right)\nu,\nu\rangle_\Gamma \\
&\geq \|\nu\|_{V^{-1}}^2 - \left\|\left(\frac{1}{2}I+K\right)\nu\right\|_{V^{-1}}\|\nu\|_{V^{-1}} \\
&\geq (1-c_K)\,\|\nu\|_{V^{-1}}^2 \\
&\geq (1-c_K)\,\langle V^{-1}\nu,\nu\rangle_\Gamma \\
&\geq (1-c_K)\,\frac{1}{c_2^V}\,\|\nu\|_{H^{\frac{1}{2}}(\Gamma)}^2 \\
&\geq (1-c_K)\,\frac{c_1^V}{c_2^V}\,\|\nu\|_{H^{\frac{1}{2}}(\Gamma)}\,\|\tau_\nu\|_{H^{-\frac{1}{2}}(\Gamma)},
\end{aligned}
$$

证毕。 □

在此基础上，变分问题(8.276)的唯一可解性，可由 Theorem 8.16证得。

Remark 8.1. *对于狄利克雷边界值问题(8.267)，我们介绍了 4 种求解方法，对应 4 个边界积分方程。这四种解都是唯一的。随着研究对象的不同和所采用的离散化方法 (第7.D节) 的不同，这些方法各有利弊。在随后的分析中，我们以(8.270)方法为主，作进一步分析。*

8.7.2 诺依曼边界值问题

考虑如下诺依曼边界值问题

$$
\begin{aligned}
Lu(x) &= 0, \quad x\in\Gamma, \\
\gamma_1^{\text{int}}u(x) &= g(x), \quad x\in\Gamma.
\end{aligned}
\tag{8.278}
$$

在满足假定可解性条件(8.17)

$$
\int_\Gamma g(x)ds_x = 0. \tag{8.279}
$$

的情况下，诺依曼边界值问题 (8.278)的解可由表现式(8.263)改写为

$$
u(\widetilde{x}) = \int_\Gamma U^*(\widetilde{x},y)g(y)\,ds_y - \int_\Gamma \gamma_{1,y}^{\text{int}}U^*(\widetilde{x},y)\,\gamma_0^{\text{int}}u(y)\,ds_y, \quad \widetilde{x}\in\Omega. \tag{8.280}
$$

为了求解(8.280)，我们需要计算诺依曼数 $\gamma_0^{\text{int}}u\in H^{\frac{1}{2}}(\Gamma)$。由卡尔德隆系统(8.264)的第二行可得第一类弗雷德霍姆边界积分方程

$$
\left(D\gamma_0^{\text{int}}u\right)(x) = (1-\sigma(x))\,g(x) - (K'g)(x), \quad x\in\Gamma. \tag{8.281}
$$

先来证明(8.281)存在唯一解 $\gamma_0^{\text{int}}u\in H_*^{\frac{1}{2}}(\Gamma)$。

1. 证明可解。根据(8.203)，$u_0\equiv 1$ 是超奇异边界积分算子 D 的特征解，满足 $(Du_0)(x)=0$。因此可得 $\ker D = span\{u_0\}$。根据闭值域定理 Theorem 8.15，我们需要假设如下可解条件

$$
(1-\sigma)\,g - K'g \in \operatorname{Im}(D) = (\ker D)^0,
$$

其中 $(\ker D)^0$ 是 v 的正交空间，定义见(8.82)。由正交属性可得

$$
\begin{aligned}
\langle(1-\sigma)\,g-K'g,u_0\rangle_\Gamma &= \langle g,1\rangle_\Gamma - \langle(\sigma I+K')\,g,u_0\rangle_\Gamma \\
&= \langle g,1\rangle_\Gamma - \langle g,(\sigma I+K)\,u_0\rangle_\Gamma \\
&= 0,
\end{aligned}
\tag{8.282}
$$

可证的边界积分方程(8.281)可解。

2. 超奇异积分算子 $D: H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma)$ 有界(8.199)，且 $H_*^{\frac{1}{2}}(\Gamma)$-椭圆 (Theorem 8.14)。

3. 由拉克斯一密格拉蒙定理 Theorem 8.14可得，边界积分方程(8.281)存在唯一解 $\gamma_0^{\text{int}} u \in H_*^{\frac{1}{2}}(\Gamma)$。

在此基础上，求解诺依曼边界值问题(8.278)，便等价于构建一个带有约束条件的变分问题：寻找解 $\gamma_0^{\text{int}} u \in H_*^{\frac{1}{2}}(\Gamma)$，使满足

$$\langle D\gamma_0^{\text{int}} u, \nu\rangle_\Gamma = \langle \left(\frac{1}{2}I - K'\right) g, \nu\rangle_\Gamma, \quad \forall \nu \in H_*^{\frac{1}{2}}(\Gamma), \tag{8.283}$$

带约束条件的变分问题又等价于以下鞍点变分问题：寻找解 $\left(\gamma_0^{\text{int}} u, \lambda\right) \in H^{\frac{1}{2}}(\Gamma) \times \mathbb{R}$，使满足

$$\begin{aligned} \langle D\gamma_0^{\text{int}} u, \nu\rangle_\Gamma + \lambda\langle \nu, w_{\text{eq}}\rangle_\Gamma &= \langle \left(\frac{1}{2}I - K'\right) g, \nu\rangle_\Gamma, \\ \langle \gamma_0^{\text{int}} u, w_{\text{eq}}\rangle_\Gamma &= 0, \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma). \end{aligned} \tag{8.284}$$

在(8.284)的第一行中，设 $\nu = u_0 \in H^{\frac{1}{2}}(\Gamma)$ 作为检测方程，我们有 $Du_0 = 0$。进而，由正交关系(8.282)可得

$$0 = \lambda \langle 1, w_{\text{eq}}\rangle_\Gamma = \lambda \langle 1, V^{-1} 1\rangle_\Gamma, \tag{8.285}$$

由于逆单层位势 V^{-1} 是椭圆的，根据上式我们有 $\lambda = 0$。进而鞍点变分问题(8.284)变为调整鞍点变分问题：寻找解 $\left(\gamma_0^{\text{int}} u, \lambda\right) \in H^{\frac{1}{2}}(\Gamma) \times \mathbb{R}$，使满足

$$\begin{aligned} \langle D\gamma_0^{\text{int}} u, \nu\rangle_\Gamma + \lambda\langle \nu, w_{\text{eq}}\rangle_\Gamma &= \langle \left(\frac{1}{2}I - K'\right) g, \nu\rangle_\Gamma, \\ \langle \gamma_0^{\text{int}} u, w_{\text{eq}}\rangle_\Gamma - \frac{\lambda}{\alpha} &= 0, \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma), \end{aligned} \tag{8.286}$$

其中 $\alpha \in \mathbb{R}_+$ 是某个待选参数。将第二行代回第一行以消除拉格朗日乘数 λ，从而得到最终的调整变分问题：寻找解 $\gamma_0^{\text{int}} u \in H^{\frac{1}{2}}(\Gamma)$，使满足

$$\langle D\gamma_0^{\text{int}} u, \nu\rangle_\Gamma + \underbrace{\alpha\langle \gamma_0^{\text{int}} u, w_{\text{eq}}\rangle_\Gamma}_{=:\lambda} \langle \nu, w_{\text{eq}}\rangle_\Gamma = \langle \left(\frac{1}{2}I - K'\right) g, \nu\rangle_\Gamma, \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma). \tag{8.287}$$

来看调整鞍点变分问题(8.287)。定义一个调整超奇异边界积分算子 $\widetilde{D}: H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma)$

$$\langle \widetilde{D} w, \nu\rangle_\Gamma := \langle Dw, \nu\rangle_\Gamma + \alpha \langle w, w_{\text{eq}}\rangle_\Gamma \langle \nu, w_{\text{eq}}\rangle_\Gamma,$$

由于

$$\begin{aligned} \langle \widetilde{D}\nu, \nu\rangle_\Gamma &= \langle D\nu, \nu\rangle_\Gamma + \alpha \left[\langle \nu, w_{\text{eq}}\rangle_\Gamma\right]^2 \\ &\geq \bar{c}_1^D \ |\nu|^2_{H^{\frac{1}{2}}(\Gamma)} + \alpha \left[\langle \nu, w_{\text{eq}}\rangle_\Gamma\right]^2 \\ &\geq \min\left\{\bar{c}_1^D, \alpha\right\} \left\{|\nu|^2_{H^{\frac{1}{2}}(\Gamma)} + \left[\langle \nu, w_{\text{eq}}\rangle_\Gamma\right]^2\right\} \\ &= \min\left\{\bar{c}_1^D, \alpha\right\} \ \|\nu\|_{H_*^{\frac{1}{2}}(\Gamma)} \\ &\geq \hat{c}_1^D \ \|\nu\|_{H_*^{\frac{1}{2}}(\Gamma)}, \quad \forall \nu \in H^{\frac{1}{2}}(\Gamma), \end{aligned}$$

我们得 $\widetilde{D}$ 有界且 $H^{\frac{1}{2}}(\Gamma)$-椭圆 $\forall w, \nu \in H^{\frac{1}{2}}(\Gamma)$。这意味着，在消除拉格朗日乘数 λ，从调整鞍点变分问题(8.286) 变为(8.287)的过程中，任何给定的 RHS，在左侧都有唯一一个 α 值与之相对应。换句话说，对于任何的诺依曼数 $\gamma_1^{\text{int}} u = g \in H^{-\frac{1}{2}}(\Gamma)$，(8.287)都存在唯一解 $\gamma_0^{\text{int}} u \in H^{\frac{1}{2}}(\Gamma)$。

现在假定有一组给定的诺依曼数 $\gamma_1^{\text{int}}u = g \in H^{-\frac{1}{2}}(\Gamma)$，满足可解性条件(8.279)。那么，通过引入检测方程 $\nu = u_0 \equiv 1$，调整鞍点变分问题(8.287)可以变为

$$\alpha\langle\gamma_0^{\text{int}}u, w_{\text{eq}}\rangle_\Gamma \langle 1, w_{\text{eq}}\rangle_\Gamma = 0,$$
$$\langle 1, w_{\text{eq}}\rangle_\Gamma = \langle 1, V^{-1}1\rangle_\Gamma > 0,$$

因此我们有 $\gamma_0^{\text{int}}u \in H_*^{\frac{1}{2}}(\Gamma)$。可见调整变分问题(8.287)与原变分问题(8.284)等价。

根据 Corollary 8.9，超奇异边界积分算子是 $H_{**}^{\frac{1}{2}(\Gamma)}$-椭圆的，则(8.281)也存在唯一的解 $\gamma_0^{\text{int}} \in H_{**}^{\frac{1}{2}(\Gamma)}$。这让我们可以用与上文相近的思路，构建一个子空间中的调整变分问题：寻找解 $\gamma_0^{\text{int}} \in H^{\frac{1}{2}(\Gamma)}$，使满足

$$\langle D\gamma_0^{\text{int}}u, \nu\rangle_\Gamma + \overline{\alpha}\langle\gamma_0^{\text{int}}u, 1\rangle_\Gamma \langle \nu, 1\rangle_\Gamma = \langle\left(\frac{1}{2}I - K'\right)g, \nu\rangle_\Gamma, \quad \forall\, \nu \in H^{\frac{1}{2}}(\Gamma), \tag{8.288}$$

同样地，$\overline{\alpha}$ 是个待选参数。在此基础上，如果假定可解性条件(8.279)得到满足，那么变分问题(8.288)的解 $\gamma_0^{\text{int}} \in H_{**}^{\frac{1}{2}}(\Gamma)$。

定义一个调整超奇异积分算子 $\hat{D} : H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma)$ 如下

$$\langle\hat{D}w, \nu\rangle_\Gamma := \langle Dw, \nu\rangle_\Gamma + \overline{\alpha}\,\langle w, 1\rangle_\Gamma \langle \nu, 1\rangle_\Gamma, \quad \forall\, w, \nu \in H^{\frac{1}{2}}(\Gamma), \tag{8.289}$$

可见 $\hat{D}$ 也是有界且 $H^{\frac{1}{2}}(\Gamma)$-椭圆的算子。

如果我们使用间接双层位势(8.266)来计算未知的密度方程 $\nu \in H_*^{\frac{1}{2}}(\Gamma)$，则我们可以得到超奇异边界积分算子形式的边界积分方程

$$(D\nu)(x) = g(x), \quad x \in \gamma, \tag{8.290}$$

进而我们可以采取与分析 BIE(8.281)相近似的方法，来分析 BIE(8.290)。

如果我们使用直接法的表现式(8.265)来计算 $\nu \in H_*^{\frac{1}{2}}(\Gamma)$，那么代入卡尔德隆系统(8.264)第一行可得边界积分方程

$$(\sigma I + K)\,\gamma_0^{\text{int}}u(x) = \left(V\gamma_1^{\text{int}}u\right)(x) = (Vg)(x), \quad x \in \Gamma, \tag{8.291}$$

其中待求解为狄利克雷数 $\gamma_0^{\text{int}}u(x), x \in \Gamma$。(8.291)的解，可由诺依曼级数求得

$$\gamma_0^{\text{int}}u(x) = \sum_{\ell=0}^{\infty}\left[(1-\sigma)\,I - K\right]^\ell (Vg)(x), \quad \forall x \in \Gamma. \tag{8.292}$$

诺依曼级数(8.292)的收敛性证明：已知位移双层位势 $[(1-\sigma)\,I - K]$ 的收缩特性可由逆单层位势 V^{-1} 中的等价范形式 $\|\cdot\|_{V^{-1}}$ 来表示，见 Corollary 8.10式(8.257)。则可以在 $H^{\frac{1}{2}}(\Gamma)$ 中构建基于(8.291)的变分问题：寻找解 $\gamma_0^{\text{int}}u \in H_*^{\frac{1}{2}}(\Gamma)$，使满足

$$\langle(\sigma I + K)\,\gamma_0^{\text{int}}u, \nu\rangle_{V^{-1}} = \langle Vg, \nu\rangle_{V^{-1}}, \quad \forall\, \nu \in H_*^{\frac{1}{2}}(\Gamma), \tag{8.293}$$

其中已知单层位势 $V : H^{-\frac{1}{2}}(\Gamma) \mapsto H^{\frac{1}{2}}(\Gamma)$ 有界且 $H^{-\frac{1}{2}}(\Gamma)$-椭圆，$\langle .,. \rangle_{V^{-1}}$ 定义为 $H^{\frac{1}{2}}(\Gamma)$ 中的内积形式

$$\langle w, \nu\rangle_{V^{-1}} := \langle V^{-1}w, \nu\rangle_\Gamma, \quad w, \nu \in H^{\frac{1}{2}}(\Gamma),$$

那么变分问题(8.293)可以调整为如下变分问题：寻找 $\gamma_0^{\text{int}}u \in H_*^{\frac{1}{2}}(\Gamma)$，使满足

$$\begin{aligned}\langle S\gamma_0^{\text{int}}u, \nu\rangle_\Gamma &= \langle V^{-1}\,(\sigma I + K)\,\gamma_0^{\text{int}}u, \nu\rangle_\Gamma \\ &= \langle g, \nu\rangle_\Gamma, \quad \forall \nu \in H_*^{\frac{1}{2}}(\Gamma),\end{aligned} \tag{8.294}$$

其中 $S: H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma)$ 是 Steklov-Poincaré 算子。

由于 Steklov-Poincaré 算子 S 和超奇异边界积分算子 $D: H^{\frac{1}{2}}(\Gamma) \mapsto H^{-\frac{1}{2}}(\Gamma)$ 有同样的映射特性，见式(8.245)和 Theorem 8.41。那么对变分问题(8.294)的唯一可解性的证明，便类似于对变分问题(8.283)唯一可解性的证明过程。

将 Steklov-Poincaré 算子的对称型表现式(8.248)代入调整变分问题(8.294)得

$$\begin{aligned}\langle S\gamma_0^{\text{int}}u, \nu\rangle_\Gamma &= \langle \left[D + (\sigma I + K')\, V^{-1}\, (\sigma I + K)\right] \gamma_0^{\text{int}}u, \nu\rangle_\Gamma \\ &= \langle g, \nu\rangle_\Gamma, \quad \forall\, \nu \in H_*^{\frac{1}{2}}(\Gamma),\end{aligned} \tag{8.295}$$

不难看出，在根据单层位势(8.265)，使用间接法求解诺依曼边界值问题是，我们最终得到以下边界积分方程

$$(\sigma I + K')\, w(x) = g(x), \quad x \in \Gamma, \tag{8.296}$$

其中未知的密度方程 $w \in H^{-\frac{1}{2}}(\Gamma)$ 可由诺依曼级数求得

$$w(x) = \sum_{\ell=0}^{\infty} \left[(1-\sigma)\, I - K'\right]^{\ell} g(x), \quad x \in \Gamma, \tag{8.297}$$

在 $H^{-\frac{1}{2}}(\Gamma)$ 空间中，转移伴随双层位势 $[(1-\sigma)\, I - K']$ 的收缩特性，可由逆单层位势 V^{-1} 形式的等价范 $\|\cdot\|_{V^{-1}}$ Corollary (8.14)式(8.261)求得。由此可得诺依曼级数(8.297)的收敛特性。

Remark 8.2. *对于诺依曼边界值问题(8.278)，我们介绍了 4 种求解方法，对应 4 个边界积分方程。这四种解都是唯一的。随着研究对象的不同和所采用的离散化方法 (第 7.D节) 的不同，这些方法各有利弊。在随后的分析中，我们以(8.288)方法为主，作进一步分析。*

8.7.3　混合边界值问题

实际应用中的边界值问题，常常是两种边界条件，即狄利克雷边界条件和诺依曼边界条件的混合，

$$\begin{aligned}Lu(x) &= 0, \quad x \in \Omega, \\ \gamma_0^{\text{int}}u(x) &= g_D(x), \quad x \in \Gamma_D, \\ \gamma_1^{\text{int}}u(x) &= g_N(x), \quad x \in \Gamma_N.\end{aligned} \tag{8.298}$$

根据表现式(8.262)，混合边界值问题的解 $u(\widetilde{x}), \widetilde{x} \in \Omega$ 可以写成如下表现形式

$$\begin{aligned}u(\widetilde{x}) = &\int_{\Gamma_N} U^*(\widetilde{x}, y) g_N(y)\, ds_y + \int_{\Gamma_D} U^*(\widetilde{x}, y) \gamma_{1,y}^{\text{int}} u(y)\, ds_y \\ &- \int_{\Gamma_D} \gamma_{1,y}^{\text{int}} U^*(\widetilde{x}, y) g_D(y)\, ds_y - \int_{\Gamma_N} \gamma_{1,y}^{\text{int}} U^*(\widetilde{x}, y) \gamma_0^{\text{int}} u(y)\, ds_y,\end{aligned} \tag{8.299}$$

可见为了求解混合边界值问题(8.298)，需要计算的未知数有两个，分别是狄利克雷数 $\gamma_0^{\text{int}}u(x), x \in \Gamma_D$ 和诺依曼数 $\gamma_1^{\text{int}}u(x), x \in \Gamma_N$。如前文所介绍的，狄利克雷边界值问题和诺依曼边界值问题分别有几种求解方法可供选择。对应地，可将这些方法之间搭配组合来求混合边界值问题。举两个例子。

直接法

采取两种直接法的组合，构建如下对称公式：对于 $x \in \Gamma_D$，采用卡尔德隆系统(8.264)的第一行求解；对于 $x \in \Gamma_N$，采用第二行求解，即

$$\begin{aligned}
&\underbrace{\left(V\gamma_1^{\text{int}}u\right)(x)}_{=:(Vg_N)(x)} = (\sigma I + K)\underbrace{\gamma_0^{\text{int}}u(x)}_{=:g_D(x)}, \quad x \in \Gamma D, \\
&\underbrace{\left(D\gamma_0^{\text{int}}u\right)(x)}_{=:(Dg_D)(x)} = [(1-\sigma)I - K']\underbrace{\gamma_1^{\text{int}}u(x)}_{=:g_N(x)}, \quad x \in \Gamma_N.
\end{aligned} \tag{8.300}$$

对于已知的 $g_D \in H^{\frac{1}{2}}(\Gamma_D), g_N \in H^{-\frac{1}{2}}(\Gamma_N)$，分别定义二者在 Γ 中的适宜的延拓 $\widetilde{g}_D \in H^{\frac{1}{2}}(\Gamma_D), \widetilde{g}_N \in H^{-\frac{1}{2}}(\Gamma_N)$，满足

$$\begin{aligned}
\widetilde{g}_D(x) &= g_D(x), \quad x \in \Gamma_D, \\
\widetilde{g}_N(x) &= g_N(x), \quad x \in \Gamma_N.
\end{aligned}$$

设

$$\begin{aligned}
\widetilde{u} &:= \gamma_0^{\text{int}}u - \widetilde{g}_D \in \widetilde{H}^{\frac{1}{2}}(\Gamma_N), \\
\widetilde{t} &:= \gamma_1^{\text{int}}u - \widetilde{g}_N \in \widetilde{H}^{-\frac{1}{2}}(\Gamma_D),
\end{aligned}$$

则边界积分方程系统(8.300)变为

$$\begin{aligned}
\left(V\widetilde{t}\right)(x) - (K\widetilde{u})(x) &= (\sigma I + K)\widetilde{g}_D(x) - (V\widetilde{g}_N)(x), \quad x \in \Gamma_D, \\
(D\widetilde{u})(x) + \left(K'\widetilde{t}\right)(x) &= [(1-\sigma)I - K']\widetilde{g}_N(x) - (D\widetilde{g}_D)(x), \quad x \in \Gamma_N.
\end{aligned} \tag{8.301}$$

显然，新的边界积分方程系统(8.301)可以用一个变分问题来表示：寻找解 $\left(\widetilde{t}, \widetilde{u}\right) \in \widetilde{H}^{-\frac{1}{2}}(\Gamma_D) \times \widetilde{H}^{\frac{1}{2}}(\Gamma_N)$，使得 $\forall\,(\tau, \nu) \in \widetilde{H}^{-\frac{1}{2}}(\Gamma_D) \times \widetilde{H}^{\frac{1}{2}}(\Gamma_N)$，均满足

$$\begin{aligned}
&a(\widetilde{t}, \widetilde{u}; \tau, \nu) = F(\tau, \nu), \\
&a(\widetilde{t}, \widetilde{u}; \tau, \nu) := \langle V\widetilde{t}, \tau\rangle_{\Gamma_D} - \langle K\widetilde{u}, \tau\rangle_{\Gamma_D} + \langle K'\widetilde{t}, \nu\rangle_{\Gamma_N} + \langle D\widetilde{u}, \nu\rangle_{\Gamma_N}, \\
&F(\tau, \nu) = \left\langle \left(\frac{1}{2}I + K\right)\widetilde{g}_D - (V, \widetilde{g}_D), \tau \right\rangle_{\Gamma_D} + \left\langle \left(\frac{1}{2}I - K'\right)\widetilde{g}_N - (D\widetilde{g}_D), \nu \right\rangle_{\Gamma_N}.
\end{aligned} \tag{8.302}$$

Lemma 8.33. 式(8.302)中，对称边界积分算子的双线性形式 $a(.,.;.,.)$ 有界，并且 $\widetilde{H}^{-\frac{1}{2}(\Gamma_D)} \times \widetilde{H}^{\frac{1}{2}(\Gamma_N)}$-椭圆，即

$$\begin{aligned}
a(t, u; \tau, \nu) &\le c_2^A \left\|(t, u)\right\|_{\widetilde{H}^{-\frac{1}{2}(\Gamma_D)} \times \widetilde{H}^{\frac{1}{2}(\Gamma_N)}} \left\|(\tau, \nu)\right\|_{\widetilde{H}^{-\frac{1}{2}(\Gamma_D)} \times \widetilde{H}^{\frac{1}{2}(\Gamma_N)}} \\
&\forall\,(t, u), (\tau, \nu) \in \widetilde{H}^{-\frac{1}{2}(\Gamma_D)} \times \widetilde{H}^{\frac{1}{2}(\Gamma_N)},
\end{aligned} \tag{8.303}$$

并且

$$a(\tau, \nu; \tau, \nu) \ge \min\left\{c_1^V, \hat{c}_1^D\right\} \left\|(\tau, \nu)\right\|^2_{\widetilde{H}^{-\frac{1}{2}(\Gamma_D)} \times \widetilde{H}^{\frac{1}{2}(\Gamma_N)}}, \tag{8.304}$$

其中范数形式定义为

$$\left\|(\tau, \nu)\right\|^2_{\widetilde{H}_{-\frac{1}{2}(\Gamma_D)} \times \widetilde{H}_{\frac{1}{2}(\Gamma_N)}} := \left\|\tau\right\|^2_{\widetilde{H}^{-\frac{1}{2}}(\Gamma_D)} + \left\|\nu\right\|^2_{\widetilde{H}^{\frac{1}{2}}(\Gamma_N)}.$$

证明. 根据定义式(8.302)

$$\begin{aligned}
a(\tau, \nu; \tau, \nu) &= \langle V\tau, \tau\rangle_{\Gamma_D} - \langle K\nu, \tau\rangle_{\Gamma_D} + \langle K'\tau, \nu\rangle_{\Gamma_N} + \langle D\nu, \nu\rangle_{\Gamma_N} \\
&= \langle V\tau, \tau\rangle_{\Gamma_D} + \langle D\nu, \nu\rangle_{\Gamma_N} \\
&\ge \underbrace{c_1^V \left\|\tau\right\|^2_{\widetilde{H}^{-\frac{1}{2}}(\Gamma_D)}}_{=:\mathcal{A}} + \underbrace{\hat{c}_1^D \left\|\nu\right\|^2_{\widetilde{H}^{\frac{1}{2}}(\Gamma_N)}}_{=:\mathcal{B}}.
\end{aligned}$$

最后一行中，$\mathcal{A}$ 由单层位势算子 V 的椭圆性求得，见 Theorem 8.36-8.35；$\mathcal{B}$ 由超奇异边界积分算子 D 的椭圆性求得，见(8.243)。

由 V, D 的有界特性可得 $a(.,.;.,.)$ 有界。 □

Lemma 8.34. 变分问题(8.302)是唯一可解的。

证明. $\forall(\tau,\nu)\in\widetilde{H}^{-\frac{1}{2}}(\Gamma_D)\times\widetilde{H}^{\frac{1}{2}}(\Gamma_N)$，$F(\tau,\nu)=$ 都有界。并且根据 Lemma 8.303，$a(.,.;.,.)$ 有界且椭圆。那么可根据拉克斯一密格拉蒙定理 Theorem 8.14证得唯一可解性。 □

间接映射法

利用狄利克雷到诺依曼的映射(8.249)，则混合边界值问题(8.298)变为，寻找解 $\gamma_0^{\text{int}}\in H^{\frac{1}{2}}(\Gamma)$，使满足

$$\begin{aligned}\gamma_0^{\text{int}}u(x)&=g_D(x),\quad x\in\Gamma_D,\quad x\in\Gamma_D,\\ \gamma_1^{\text{int}}u(x)&=\left(S\gamma_0^{\text{int}}u\right)(x)=g_N(x),\quad x\in\Gamma_N.\end{aligned}$$

对于给定的狄利克雷数 $g_D\in H^{\frac{1}{2}}(\Gamma_D)$，假设存在任一延拓 $\widetilde{g}_D\in H^{\frac{1}{2}}(\Gamma)$，则可以表示为如下变分问题：寻找 $\widetilde{u}:=\gamma_0^{\text{int}}u-\widetilde{g}_D\in\widetilde{H}$，使满足

$$\langle S\widetilde{u},\nu\rangle_{\Gamma_N}=\langle g_N-S\widetilde{g}_D,\nu\rangle_{\Gamma_N},\quad\forall\nu\in\widetilde{H}^{\frac{1}{2}}(\Gamma_N).\tag{8.305}$$

类似地，由于 Steklov-Poincaré 算子 $S:H^{\frac{1}{2}}(\Gamma)\mapsto H^{-\frac{1}{2}}(\Gamma)$ 有界并且 $\widetilde{H}^{\frac{1}{2}}(\Gamma_N)$-椭圆(8.253)，可以利用拉克斯一密格拉蒙定理 Theorem 8.14证得 $\widetilde{u}$ 的唯一可解性。在计算得到狄利克雷数 $\gamma_0^{\text{int}}u(x)$ 的值后，我们可以通过求解狄利克雷边界值问题得到完整的诺依曼数 $\gamma_1^{\text{int}}u\in H^{-\frac{1}{2}}(\Gamma)$。

8.7.4　罗宾边界值问题

考虑如下罗宾边界值问题

$$\begin{aligned}(Lu)(x)&=0,\quad x\in\Omega,\\ \gamma_1^{\text{int}}u(x)+\kappa(x)\gamma_0^{\text{int}}u(x)&=g(x),\quad x\in\Gamma.\end{aligned}$$

利用狄利克雷到诺依曼的映射(8.249)，我们可以构建如下边界积分方程，用于求解未知的狄利克雷数 $\gamma_0^{\text{int}}u(x)\in H^{\frac{1}{2}}(\Gamma)$

$$\gamma_1^{\text{int}}u(x)=\left(S\gamma_0^{\text{int}}u\right)(x)=g(x)-\kappa(x)\gamma_0^{\text{int}}u(x),\quad x\in\Gamma.$$

因此可以构建如下变分问题：寻找解 $\gamma_0^{\text{int}}u\in H^{\frac{1}{2}}(\Gamma)$，使满足

$$\langle S\gamma_0^{\text{int}}u,\nu\rangle_\Gamma+\langle\kappa\gamma_0^{\text{int}}u,\nu\rangle_\Gamma=\langle g,\nu\rangle_\Gamma,\quad\forall\nu\in H^{\frac{1}{2}}(\Gamma).\tag{8.306}$$

已知 Steklov-Poincaré 算子 S 和超奇异边界积分算子 D 之间的关系满足不等式(8.250)

$$\langle S\nu,\nu\rangle_\Gamma\geq\langle D\nu,\nu\rangle_\Gamma,\quad\forall\nu\in H^{\frac{1}{2}}(\Gamma)$$

根据 D 的半椭圆，见 Corollary 8.9式(8.241)，我们有

$$\langle D\nu,\nu\rangle_\Gamma\geq\bar{c}_1^D\ |\nu|^2_{H^{\frac{1}{2}}(\Gamma)},\quad\forall\nu\geq H^{\frac{1}{2}}(\Gamma),$$

此外假定 $\kappa(x) \geq \kappa_0, x \in \Gamma$。那么定义一个双线性形式 $a(.,.)$，满足性质

$$\begin{aligned} a(.,.) &\coloneqq \langle S\nu, \nu\rangle_\Gamma + \langle \kappa\nu, \nu\rangle_\Gamma \\ &\geq \langle D\nu, \nu\rangle_\Gamma + \kappa_0 \left\|\nu\right\|^2_{H^{\frac{1}{2}}(\Gamma)} \\ &\geq \bar{c}_1^D \left|\nu\right|^2_{H^{\frac{1}{2}}(\Gamma)} + \kappa_0 \left\|\nu\right\|^2_{H^{\frac{1}{2}}(\Gamma)} \\ &\geq \min\left\{\bar{c}_1^D, \kappa_0\right\} \left\|\nu\right\|^2_{H^{\frac{1}{2}}(\Gamma)}, \end{aligned}$$

可见双线性形式 $a(.,.)$ 是 $H^{\frac{1}{2}}(\Gamma)$-椭圆的。

在此基础上，变分问题(8.306)的唯一可解性可由拉克斯—密格拉蒙定理 Theorem 8.14证得。

8.8 近似方法

本节主要介绍含有算子方程的变分问题 (第8.3) 的常见近似求解方法。这些近似法有所不同，大体上来说都是通过有限维的协调检测空间 (conforming trial spaces)，将原本非线性方程系统作近似线性化的。

8.8.1 伽辽金——布博诺夫法

设一个算子 $A : X \mapsto X'$ 且 X-椭圆，对于所有 $\nu \in X$ 满足

$$\begin{aligned} \langle A\nu, \nu\rangle &\geq c_1^A \left\|\nu\right\|^2_X, \\ \left\|A\nu\right\|_{X'} &\leq c_1^A \left\|\nu\right\|_X. \end{aligned}$$

假定对于某一给定的方程 $f \in X'$，我们想要求得变分问题(8.69)的解 $u \in X$

$$\langle Au, \nu\rangle = \langle f, \nu\rangle, \quad \forall \nu \in X. \tag{8.307}$$

由拉克斯—密格拉蒙定理 Theorem 8.14可知，(8.307)存在唯一的解 $u \in X$，满足

$$\left\|u\right\|_X \leq \frac{1}{c_1^A} \left\|f\right\|_{X'}.$$

考虑如下一个序列 (sequence) $\{X_M\}$, $M \in \mathbb{N}$，序列由一系列协调检测空间 (conforming trial spaces)构成，定义如下

$$X_M \coloneqq \operatorname{span}\{\varphi_k\}_{k=1}^M \subset X.$$

那么可以根据(8.307)构建伽辽金——布博诺夫变分问题 (Galerkin-Bubnov variational problem)，寻找近似解 $u_M \in X_M$，使满足

$$\langle Au_M, \nu_M\rangle = \langle f, \nu_M\rangle, \quad \forall u_M \in X_M, \tag{8.308}$$

其中近似解 $u_M \in X_M$ 定义为

$$u_M \coloneqq \sum_{k=1}^M u_k\varphi_k \in X_M. \tag{8.309}$$

为了证明伽辽金——布博诺夫变分法，是对原变分问题(8.307)的有效近似方法，需要分别证明以下三方面

- 近似解 $u_M \in X_M$ 是唯一可解的，
- 近似解 $u_M \in X_M$ 的稳定性，
- 近似解 $u_M \in X_M$ 的收敛性。

近似解的唯一可解性

由于 $X_M \subset X$，选取 $\nu = \nu_M \in X_M$ 代入原变分问题(8.307)中，进而减去(8.308)可得

$$\langle A(u - u_M), \nu_M \rangle = 0, \quad \forall \nu_M \in X_M, \tag{8.310}$$

这称为伽辽金正交 (Galerkin Orthogonality)。

将近似解 u_M 的定义式(8.309)代入伽辽金——布博诺夫变分式(8.308)得

$$\left\langle A \sum_{k=1}^{M} u_k \varphi_k, \nu_M \right\rangle = \langle f, \nu_M \rangle,$$

由算子 A 的线性定义，上式可以进一步调整为有限维度的变分问题

$$\sum_{k=1}^{M} u_k \langle A\varphi_k, \varphi_\ell \rangle = \langle f, \varphi_\ell \rangle, \quad \ell = 1, \ldots, M. \tag{8.311}$$

利用伽辽金——布博诺夫近似法，有限维变分问题(8.311)，进而原变分问题(8.307)，等价于一个线性方程系统：求解系数向量 $\underline{u} \in \mathbb{R}^M$，使满足

$$A_M \underline{u} = \underline{f}, \tag{8.312}$$

其中

- $A_M \in \mathbb{R}^{M \times M}$ 称为刚度矩阵 (stiffness matrix)，矩阵中的元素 $A_M[\ell, k]$ 定义为

$$A_M[\ell, k] := \langle A\varphi_k, \varphi_l \rangle, \quad k, \ell = 1, \ldots, M.$$

- 向量 $\underline{f}$ 由一系列方程 f_ℓ 构成，f_ℓ 定义为

$$f_\ell := \langle f, \varphi_\ell \rangle, \quad \ell = 1, \ldots, M.$$

已知对于某一向量 $\underline{\nu} \in \mathbb{R}^M$，可定义方程如下

$$\nu_M = \sum_{k=1}^{M} \nu_k \varphi_k \in X_M,$$

那么对于任意给定的 $\underline{u}, \underline{\nu} \in \mathbb{R}^M$，我们有

$$\begin{aligned}
\langle A_M \underline{u}, \underline{\nu} \rangle &= \sum_{k=1}^{M} \sum_{\ell=1}^{M} A_M[\ell, k] u_k \nu_\ell \\
&= \sum_{k=1}^{M} \sum_{\ell=1}^{M} \langle A\varphi_k, \varphi_\ell \rangle u_k \nu_\ell \\
&= \left\langle A \sum_{k=1}^{M} u_k \varphi k, \sum_{\ell=1}^{M} \nu_\ell \varphi_\ell \right\rangle \\
&= \langle A u_M, \nu_M \rangle.
\end{aligned}$$

由此可见，刚度矩阵 $A_M \in \mathbb{R}^{M \times M}$ 继承了线性椭圆算子 $A : X \mapsto X'$ 的所有特性，其中尤其重要的是：

- A 是自伴随的 $\to A_M$ 是对称的

- A 是 X-椭圆的 $\to$ A_M 是正定的

$$\begin{aligned}\langle A_M \underline{\nu}, \nu\rangle &= \langle A\nu_M, \nu_M\rangle \\ &\geq c_1^A \left\|\nu_M\right\|_X^2, \quad \forall \nu \in \mathbb{R}^M \leftrightarrow \nu_M \in X_M.\end{aligned}$$

由此我们有：A 的椭圆性 $\to$ 变分问题(8.307)的唯一可解性 $\to$ 伽辽金——布博诺夫变分问题(8.308)的唯一可解性 $\to$ 线性方程系统(8.312)的唯一可解性。

近似解的稳定性

近似解的稳定性可由齐亚引理 (Céa's lemma)予以证明。

Theorem 8.42 (齐亚引理)**.** *设 $A: X \mapsto X'$ 是一个有界且 X-椭圆的算子。对于伽辽金——布博诺夫变分问题(8.308)的唯一近似解 $u_M \in X_M$ 而言，满足稳定性条件*

$$\left\|u_M\right\|_X \leq \frac{1}{c_1^A}\left\|f\right\|_{X'}, \tag{8.313}$$

并且误差测度项为

$$\left\|u-u_M\right\|_X \leq \frac{c_2^A}{c_1^A} \inf_{\nu_M \in X_M}\left\|u-\nu_M\right\|_X. \tag{8.314}$$

证明. 近似解 u_M 的唯一可解性由第325页第8.8.1节给出。对于唯一的近似解 $u_M \in X_M$，由算子 A 的 X-椭圆性有

$$c_1^A \left\|u_M\right\|_X^2 \leq \langle Au_M, u_M\rangle \langle f, u_M\rangle \leq \left\|f\right\|_{X'} \left\|u_M\right\|_X,$$

进而得到稳态条件(8.313)。

由线性算子 A 有界且 X-椭圆，以及伽辽金正交(8.310)可得，对于任一 $\nu_M \in X_M$ 都有

$$\begin{aligned}c_1^A \left\|u-u_M\right\|^2 &\leq \langle A(u-u_M), u-u_M\rangle \\ &= \langle A(u-u_M), u-\nu_M\rangle + \langle A(u-u_M), \nu_M-u_M\rangle \\ &= \langle A(u-u_M), u-\nu_M\rangle \\ &\leq c_2^A \left\|u-u_M\right\|_X \left\|u-\nu_M\right\|_X,\end{aligned}$$

进而得到误差测度项(8.314)。 □

近似解的收敛性

近似解的收敛性是指，随着 $M \to \infty$，近似解 $u_M \to u$，这与检测空间 X_M 的近似属性有关

$$\lim_{M\to\infty} \inf_{\nu_M \in X_M}\left\|\nu-\nu_M\right\|_X = 0, \quad \forall \nu \in X_M. \tag{8.315}$$

在构建协调检测空间的序列 $\{X_M\}_{M\in\mathbb{N}} \subset X$ 的过程中，需要确保每个检测空间 X_M 均满足近似属性(8.315)：其方法之一是利用局域多项式作基方程构建序列，第九、十分别讨论有限元、有界元情况下的基方程。此外，随着近似解的类型不同，我们还要考虑其他一些近似属性，见下。

reference

8.8.2 线性形近似

在构建经济学模型时，变分问题(8.307)右侧的 $f \in X'$ 有时表示为以下线性形式 $f = Bg$，其中 $g \in Y$ 为预先给定的方程，有界线性算子 $B : Y \mapsto X'$ 满足

$$\|Bg\|_{X'} \le c_2^B \|g\|_Y, \quad \forall g \in Y.$$

对应的，变分问题调整为：寻找解 $u \in X$，使满足

$$\langle Au, \nu \rangle = \langle Bg, \nu \rangle, \quad \forall \nu \in X. \tag{8.316}$$

利用伽辽金——布博诺夫法(8.308)构建近似变分问题：寻找唯一的近似解 $u_M \in X_M$，使满足

$$\langle Au_M, \nu_M \rangle = \langle Bg, \nu_M \rangle, \quad \forall \nu_M \in X_M. \tag{8.317}$$

随后的任务是生成类似(8.312)的线性方程系统。通过测度算子 $B : Y \mapsto X'$ 或伴随算子 $B' : X \mapsto Y'$ 来计算

$$f_\ell = \langle Bg, \varphi_\ell \rangle = \langle g, B'\varphi_\ell \rangle, \quad \ell = 1, \ldots, M,$$

用 g_N 对上式中的 g 作近似替代

$$g_N = \sum_{i=1}^{N} g_i \psi_i \in Y_N = \operatorname{span}\{\psi_i\}_{i=1}^N \subset Y, \quad i = 1, \ldots, N.$$

这样我们有了一个新的扰动变分问题 (perturbed variational problem)：寻找近似解 $\tilde{u}_M \in X_M$，使满足

$$\langle A\tilde{u}_M, \nu_M \rangle = \langle Bg_N, \nu_M \rangle, \quad \forall \nu_M \in X_M. \tag{8.318}$$

该问题等价于如下线性方程系统

$$A_M \underline{\tilde{u}}_M = B_N \underline{g}, \tag{8.319}$$

其中

- 矩阵 A_M, B_N 中的元素分别为

$$A_M[\ell, k] = \langle A\varphi_k, \varphi_\ell \rangle, \quad k, \ell = 1, \ldots, M,$$
$$B_N[\ell, i] = \langle B\psi_i, \varphi_\ell \rangle, \quad i = 1, \ldots, N.$$

- 向量 $\underline{g}$ 中包含 g_N，是对 g 的近似。

借助线性方程系统(8.319)，我们可以计算矩阵 B_N，并且 B_N 不受给定的近似方程 g_N 的影响。

根据有界线性算子 A 的 X-椭圆性，可得 A_M 是正定矩阵，进而(8.319)唯一可解，进而扰动变分问题(8.318)唯一可解。

唯一近似解 $u_M \in X_M$ 对唯一解 $u \in X$ 的近似精度 (误差)，主要受到(8.319)中 g_N 对 g 的近似精度 (误差) 的影响，误差测度见以下斯特朗第一引理 (Strang's First Lemma)。

Theorem 8.43 (斯特朗第一引理)**.** *设一个线性算子 $A : X \mapsto X'$ 有界且 X-椭圆。设 $u \in X$ 是连续变分问题(8.316)的唯一解，那么对于扰动变分问题(8.318)的唯一近似解 $\tilde{u}_M \in X_M$，近似的误差测度为*

$$\|u - \tilde{u}_M\|_X \le \frac{1}{c_1^A} \left\{ c_2^A \inf_{\nu_M \in X_M} \|u - \nu_M\|_X + c_2^B \|g - g_N\|_X \right\}. \tag{8.320}$$

证明. 将两个变分问题(8.317)，(8.318)相减得

$$\langle A\left(u_M-\tilde{u}_M\right),\nu_M\rangle=\langle B\left(g-g_N\right),\nu_M\rangle,\quad \forall\,\nu_M\in X_M.$$

现在定义检测方程为 $\nu_M:=u_M-\tilde{u}_M\in X_M$，再考虑到 A 和 B 的有界性，以及 A 的 X-椭圆性，上式变为

$$\begin{aligned}c_1^A\,\|u_M-\tilde{u}_M\|_X^2&\le\langle A\left(u_M-\tilde{u}_M\right),\left(u_M-\tilde{u}_M\right)\rangle\\&=\langle B\left(g-g_N\right),\left(u_M-\tilde{u}_M\right)\rangle\\&\le\|B\left(g-g_N\right)\|_{X'}\,\|u_M-\tilde{u}_M\|_X\\&\le c_2^B\,\|g-g_N\|_Y\,\|u_M-\tilde{u}_M\|_X,\end{aligned}$$

$$\hookrightarrow\|u_M-\tilde{u}_M\|_X\le\frac{c_2^B}{c_1^A}\,\|g-g_N\|_Y,$$

进一步应用三角不等式，结合齐亚引理 Theorem 8.42式(8.314)有

$$\begin{aligned}\|u-\tilde{u}_M\|_X&\le\|u-u_M\|_X+\|u_M-\tilde{u}_M\|_X\\&\le\frac{c_2^A}{c_1^A}\inf_{\nu_M\in X_M}\|u-\nu_M\|_X+\frac{c_2^B}{c_1^A}\,\|g-g_N\|_Y,\end{aligned}$$

证毕。 □

8.8.3 算子的近似

除了对变分问题(8.307)右侧的近似之外，在进行求积分的数值近似计算时，有时我们也要考虑对给定算子作近似。举例来说，将伽辽金变分问题(8.308)转化为如下扰动变分问题：寻求解 $\tilde{u}_M\in X_M$，使满足

$$\langle\widetilde{A}\tilde{u}_M,\nu_M\rangle=\langle f,\nu_M\rangle,\quad\forall\,\nu_M\in X_M,\tag{8.321}$$

其中近似算子 $\widetilde{A}:X\mapsto X'$ 是个有界的线性算子，满足

$$\left\|\widetilde{A}\nu\right\|_{X'}\le\tilde{c}_2^A\,\|\nu\|_X,\quad\forall\,\nu\in X.\tag{8.322}$$

扰动变分问题(8.321)与伽辽金变分问题(8.308)相减，得伽辽金正交条件

$$\langle Au_M-\widetilde{A}\tilde{u}_M,\nu_M\rangle=0,\quad\forall\,\nu_M\in X_M.\tag{8.323}$$

为了确保变分问题(8.321)唯一可解，我们假定近似算子 $\widetilde{A}$ 的稳定性。在此基础上，唯一近似解 $\tilde{u}_M\in X_M$ 的误差测度项见斯特朗第二引理 (Strang's Second Lemma)。。

Theorem 8.44 (斯特朗第二引理)**.** 假设近似算子 $\widetilde{A}:X\mapsto X'$ 线性，有界，X-椭圆，即

$$\langle\widetilde{A}\nu_M,\nu_M\rangle\ge\tilde{c}_1^A\,\|\nu_M\|_X^2.\tag{8.324}$$

那么扰动变分问题(8.321)存在唯一解 $\tilde{u}_M\in X_M$，满足误差测度项

$$\|u_M-\tilde{u}_M\|_X\le\left[1+\frac{1}{\tilde{c}_1^A}\left(c_2^A+\tilde{c}_2^A\right)\right]\frac{c_2^A}{c_1^A}\inf_{\nu_M\in X_M}\|u-\nu_M\|_X+\frac{1}{\tilde{c}_1^A}\left\|\left(A-\widetilde{A}\right)u\right\|_{X'}.\tag{8.325}$$

证明. 第一步。由假设条件近似算子 $\widetilde{A}$ 的 X-椭圆性可得，对应的刚度矩阵 $\widetilde{A}_M$ 正定，进而扰动变分问题(8.321)存在唯一解 $\tilde{u}_M \in X_M$。

第二步。对于伽辽金变分问题(8.308)的唯一解 $u_M \in X_M$，由 $\widetilde{A}$ 的 X_M-椭圆性，以及正交条件(8.323)可得

$$
\begin{aligned}
\tilde{c}_1^A \left\| u_M - \tilde{u}_M \right\|_X^2 &\leq \left\langle \widetilde{A}\left(u_M - \tilde{u}_M\right), u_M - \tilde{u}_M \right\rangle \\
&= \left\langle \left(\widetilde{A} - A\right) u_M, u_M - \tilde{u}_M \right\rangle + \left\langle A u_M - \widetilde{A}_M, u_M - \tilde{u}_M \right\rangle \\
&= \left\langle \left(\widetilde{A} - A\right) u_M, u_M - \tilde{u}_M \right\rangle \\
&\leq \left\| \left(\widetilde{A} - A\right) u_M \right\|_{X'} \left\| u_M - \tilde{u}_M \right\|_X ,
\end{aligned}
$$

$$
\hookrightarrow \left\| u_M - \tilde{u}_M \right\|_X \leq \frac{1}{\tilde{c}_1^A} \left\| \left(\widetilde{A} - A\right) u_M \right\|_{X'} .
$$

由于两个算子 $A, \widetilde{A} : X \mapsto X'$ 均有界，

$$
\begin{aligned}
\left\| \left(A - \widetilde{A}\right) u_M \right\|_{X'} &\leq \left\| \left(A - \widetilde{A}\right) u \right\|_{X'} + \left\| \left(A - \widetilde{A}\right) \left(u - u_M\right) \right\|_{X'} \\
&\leq \left\| \left(A - \widetilde{A}\right) u \right\|_{X'} + \left[c_2^A + \tilde{c}_2^A\right] \left\| u - u_M \right\|_X .
\end{aligned}
$$

第三步。在此基础上

$$
\begin{aligned}
\left\| u_M - \tilde{u}_M \right\|_X &\leq \left\| u - u_M \right\|_X + \left\| u_M - \widetilde{u}_M \right\|_X \\
&\leq \left\| u - u_M \right\|_X + \frac{1}{\tilde{c}_1^A} \left\| \left(\widetilde{A} - A\right) u_M \right\|_{X'} \\
&\leq \left\| u - u_M \right\|_X + \frac{1}{\tilde{c}_1^A} \left\| \left(\widetilde{A} - A\right) u \right\|_{X'} + \frac{c_2^A + \tilde{c}_2^A}{\tilde{c}_1^A} \left\| u - u_M \right\|_X \\
&\leq \left[1 + \frac{1}{\tilde{c}_1^A}\left(c_2^A + \tilde{c}_2^A\right)\right] \frac{c_2^A}{c_1^A} \inf_{\nu_M \in X_M} \left\| u - \nu_M \right\|_X + \frac{1}{\tilde{c}_1^A} \left\| \left(A - \widetilde{A}\right) u \right\|_{X'} .
\end{aligned}
$$

□

8.8.4　伽辽金——佩特洛夫法

设一个有界线性算子 $B : X \mapsto \Pi'$ 满足 (连续) 稳定性条件

$$
c_S \left\| \nu \right\| \leq \sup_{0 \neq q \in \Pi} \frac{\langle B\nu, q \rangle}{\left\| q \right\|_\Pi}, \forall \nu \in (\ker B)^\top \subset X, \tag{8.326}
$$

那么根据定理8.16，对于某一给定的 $g \operatorname{Im}_X(B)$，算子方程 $Bu = g$ 有唯一解 $u \in (\ker B)^\top$，满足

$$
\langle Bu, q \rangle = \langle g, q \rangle_\Pi, \quad \forall q \in \Pi.
$$

对于 $M \in \mathbb{R}$，引入两个协调检测空间序列

$$
\begin{aligned}
X_M &= \operatorname{span}\left\{\varphi_k\right\}_{k=1}^M \subset (\ker B)^\top , \\
\Pi_M &= \operatorname{span}\left\{\psi_k\right\}_{k=1}^M \subset \Pi.
\end{aligned}
$$

则伽辽金——佩特洛夫变分问题 (Galerkin-Petrov varational problem)可以定义如下：寻找近似解 $u_M \in X_M$，u_M 的定义如(8.309)，使满足

$$
\langle Bu_M, q_M \rangle = \langle g, q_M \rangle, \quad \forall q_M \in \Pi_M. \tag{8.327}
$$

与伽辽金——布博诺夫法(8.307)相比，在伽辽金——佩特洛夫法(8.327)中有两组检测空间。由 $\Pi_M \subset \Pi$ 可得伽辽金正交条件

$$\langle B(u-u_M), q_M\rangle = 0, \quad \forall q_M \in \Pi_M. \tag{8.328}$$

那么伽辽金——佩特洛夫变分问题(8.327)等价于如下线性方程系统

$$B_M \underline{u}_M = g, \tag{8.329}$$

- 刚度矩阵 B_M 中的元素定义为

$$B_M[\ell, k] = \langle B\varphi_k, \psi_\ell\rangle, \quad k, \ell = 1, \ldots, M,$$

- 向量 $\underline{g}$ 由 g_ℓ 组成

$$g_\ell = \langle g, \psi_\ell\rangle, \quad \ell = 1, \ldots, M.$$

假定线性方程系统(8.329)满足离散的稳定条件（我们将在下文中证明这一假设，见 Lemma 8.35）

$$\tilde{c}_S \|\nu_M\|_X \le \sup_{0\neq q_M\in\Pi_M} \frac{\langle B\nu_M, q_M\rangle}{\|q_M\|_\Pi}, \quad \forall \nu_M \in X_M, \tag{8.330}$$

那么可得系统存在唯一解，见如下定理。

Theorem 8.45 (伽辽金——佩特洛夫变分问题的唯一近似解)**.** *设算子方程 $Bu = g$ 存在唯一解 $u \in (\ker B)^\top$，则伽辽金——佩特洛夫变分问题(8.327)有唯一近似解 $u_M \in X_M$。*

假定离散稳定条件(8.330)成立，则误差测度项表示为

$$\|u-u_M\|_X \le \left(1+\frac{c_2^B}{\tilde{c}_B}\right) \inf_{\nu_M\in X_M} \|u-\nu_M\|_X. \tag{8.331}$$

证明. 1. 对于任一给定 $\nu \in (\ker B)^\top$，以下变分问题都存在唯一确定的近似解 $\nu_M = P_M\nu \in X_M$，满足

$$\langle B\nu_M, q_M\rangle = \langle B\nu, q_M\rangle, \quad \forall q_M \in \Pi_M.$$

2. 由离散稳定性条件(8.330)可得，唯一近似解 $\nu_M \in X_M$ 满足

$$\begin{aligned}
\tilde{c}_S \|\nu_M\|_X &\le \sup_{0\neq q_M\in\Pi_M} \frac{\langle B\nu_M, q_M\rangle}{\|q_M\|_\Pi} \\
&= \sup_{0\neq q_M\in\Pi_M} \frac{\langle B\nu, q_M\rangle}{\|q_M\|_\Pi} \\
&\le c_2^B \|\nu\|_X,
\end{aligned}$$

由此可见，对于任一 $\nu \in (\ker B)^\top$，都有唯一的 $\nu_M = P_M\nu \in \Pi_M$ 与之相对应，满足

$$\|P_M\nu\|_X \le \frac{c_2^B}{\tilde{c}_2^B} \|\nu\|_X.$$

3. 一方面伽辽金——佩特洛夫变分问题(8.327)有唯一解 $u_M \in X_M$，据此可得 $u_M = p_M u$。另一方面，对于所有 ν_M 都有 $\nu_M \in P_M\nu_M$。因此，对于某一任意的 $\nu_M \in X_M$，都有

$$\begin{aligned}
\|u-u_M\|_X &= \|(u-\nu_M) + (\nu_M - u_M)\|_X \\
&= \|(u-\nu_M) - P_M(u-\nu_M)\|_X \\
&\le \|u-\nu_M\|_X + \|P_M(u-\nu_M)\|_X \\
&\le \left(1+\frac{c_2^B}{\tilde{c}_S}\right) \|u-\nu_M\|_X \\
&\le \left(1+\frac{c_2^B}{\tilde{c}_S}\right) \inf_{\nu_M\in X_M} \|u-\nu_M\|_X.
\end{aligned}$$

□

随着 $M \to \infty$，伽辽金——佩特洛夫法(8.327)近似解 $u_M \to u$，收敛属性的证明过程同伽辽金——布博诺夫法，见第326页第8.8.1节。

对伽辽金——佩特洛夫法(8.327)近似解 $u_M \in X_M$ 的离散稳定性条件(8.330)的证明，可见Fortin (1977)。

Lemma 8.35 (离散稳定性条件)**.** 设一个有界线性算子 $B : X \mapsto \Pi'$，满足连续稳定条件(8.326)。如果存在一个有界的投影算子 $R_M : \Pi \mapsto \Pi_M$，满足

$$\begin{aligned} \langle B\nu_M, q - R_M q\rangle = 0, \quad \forall\, \nu_M \in X_M, \\ \|R_M q\|_\Pi \le c_R \, \|q\|_\Pi, \quad \forall\, q \in \Pi, \end{aligned}$$

则离散稳定条件(8.330)成立，对应 $\tilde{c}_S = \frac{c_S}{c_R}$。

证明. 由连续稳定条件(8.326)可得，对于某一 $q_N \in \Pi_N \subset \Pi$ 有

$$\begin{aligned} c_S \, \|\nu_M\|_X &\le \sup_{0\neq q\in\Pi} \frac{\langle B\nu_M, q\rangle}{\|q\|_\Pi} \\ &= \sup_{0\neq q\in\Pi} \frac{\langle B\nu_M, R_M q\rangle}{\|q\|_\Pi} \\ &\le c_R \sup_{0\neq q\in\Pi} \frac{\langle B\nu_M, R_M q\rangle}{\|R_M q\|_\Pi} \\ &\le c_R \sup_{0\neq q_M\in\Pi_M} \frac{\langle B\nu_M, q_M\rangle}{\|q_M\|_\Pi}, \end{aligned}$$

$$\hookrightarrow \underbrace{\frac{c_R}{c_S}}_{=:\tilde{c}_S} \|\nu_M\|_X \le \frac{\langle B\nu_M, q_M\rangle}{\|q_M\|_\Pi}$$

□

8.8.5　混合边界值问题

混合算子方程的鞍点变分问题

现在来看关于混合算子方程的鞍点变分问题(8.93a)-(8.93b)：寻找解 $(u, p) \in X \times \Pi$，使满足

$$\begin{aligned} \langle Au, \nu\rangle + \langle B\nu, p\rangle &= \langle f, \nu\rangle, \\ \langle Bu, q\rangle &= \langle g, q\rangle, \quad \forall\, (\nu, q) \in X \times \Pi. \end{aligned} \tag{8.332}$$

假定两个有界线性算子 $A : X \mapsto X', B : X \mapsto \Pi'$。假定 A 是 X-椭圆的。假定连续稳定条件(8.326)成立，那么 Theorem 8.19的所有前提条件均得到满足，混合算子方程的鞍点变分问题(8.332)存在唯一的解 $(u, p) \in X \times \Pi$。

伽辽金变分问题

对于 $M, N \in \mathbb{N}$，定义两个协调检测空间

$$\begin{aligned} X_M &= \operatorname{span}\{\varphi_k\}_{k=1}^M \subset X, \\ \Pi_N &= \operatorname{span}\{\psi_i\}_{i=1}^N \subset \Pi. \end{aligned}$$

则鞍点变分问题(8.332)可以表示为如下伽辽金变分问题：寻求解 $(u_M, p_N) \in X_M \times \Pi_N$，使满足

$$\begin{aligned} \langle Au_M, \nu_M \rangle + \langle B\nu_M, p_N \rangle &= \langle f, \nu_M \rangle, \\ \langle Bu_M, q_N \rangle &= \langle g, q_N \rangle, \quad \forall\, (\nu_M, q_N) \in X_M \times \Pi_N. \end{aligned} \tag{8.333}$$

线性方程系统

定义矩阵 A_M, B_N，其中元素

$$\begin{aligned} A_M[\ell, k] &= \langle A\varphi_k, \varphi_\ell \rangle, \quad k, \ell = 1, \ldots, M, \\ B_N[j, k] &= \langle B\varphi_k, \psi_j \rangle, \quad j = 1, \ldots, N. \end{aligned}$$

定义向量 $\underline{f}, \underline{g}$，向量中的元素 f, g 分别为

$$\begin{aligned} f_\ell &= \langle f, \varphi_\ell \rangle, \\ g_j &= \langle g, \psi_j \rangle. \end{aligned}$$

则伽辽金变分问题(8.333)等价于下述线性方程系统

$$\begin{pmatrix} A_M & B_N^\top \\ B_N & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{g} \end{pmatrix}. \tag{8.334}$$

舒尔补系统

先来看线性方程系统(8.334)唯一可解性的证明。首先，唯一可解性的必要条件是 $M \geq N$。设系统矩阵 K 的维度 $\dim K = M + N$，根据

$$\operatorname{range} A_M \leq M, \quad \operatorname{range} B_N \leq \min\{M, N\}$$

可得

$$\operatorname{range} K \leq M + \min\{M, N\}.$$

若 $M < N$，会出现以下情况

$$\operatorname{range} K \leq 2M < M + N = \dim K,$$

换句话说，线性方程系统(8.334)不可解。为了使解存在，必要条件是要审慎定义检测空间 X_M 和 Π_N，使 $M \geq N$，换句话说，检测空间 X_M 需要大于另一个检测空间 Π_N。

在满足必要条件 $M \geq N$ 后，伽辽金变分问题(8.333)唯一可解性的证明，可由 Theorem 8.20证得。

如果 A_M 是可逆的，则可将线性方程系统(8.334)转换成舒尔补系统 (Schur complement system)，见第6.A.1节。

由于协调检测空间 $X_M \subset X$，由有界线性算子 A 的 X-椭圆性可得，刚度矩阵 A_M 是正定的，即

$$\langle A_M \underline{\nu}, \underline{\nu} \rangle = \langle A\nu_M, \nu_M \rangle \geq c_1^A \|\nu_M\|_X^2 > 0, \quad \forall\, \underline{0} \neq \underline{\nu} \in \mathbb{R}^M \leftrightarrow \nu_M \in X_M,$$

因此 A_M 可逆。

线性方程系统(8.334)转换成舒尔补系统 (Schur complement system) 如下

$$B_N A_M^{-1} B_N^\top \underline{p} = B_N A_M^{-1} \underline{f} - \underline{g}, \tag{8.335}$$

唯一可解性

关于舒尔补系统(8.335)唯一可解性的证明。

假定系统满足离散稳定条件

$$\tilde{c}_S \, \|q_N\|_\Pi \leq \sup_{0\neq\nu_M\in X_M} \frac{\langle B\nu_M, q_N\rangle}{\|\nu_M\|_X}, \quad \forall\, q_N \in \Pi_N, \tag{8.336}$$

上式又称 Babuška-Brezzi-Ladyshenskaya (BBL) 条件。通常来说我们无法直接从 Theorem 8.18式8.94直接推得离散稳定 (BBL) 条件(8.336)的成立，一个间接的证明方法见下文 Lemma 8.37节。

Lemma 8.36 (舒尔补系统的唯一可解性)**.** 设两个有界线性算子 $A : X \mapsto X', B : X \mapsto \Pi'$，且 A 是 X-椭圆的。对于协调检测空间 $X_M \subset X, \Pi_N \subset \Pi$，假定离散稳定条件(8.336)成立，那么在舒尔补系统(8.335)中，定义舒尔补矩阵

$$S_N =: B_N A_M^{-1} B_N^\top,$$

则 S_N 是正定矩阵，满足

$$\left(S_N \underline{q}, \underline{q}\right) \geq c_1^{S_N} \, \|q_N\|_\Pi^2, \quad \forall\, 0 \neq \underline{q} \in \Pi_N \leftrightarrow q_N \in \Pi_N. \tag{8.337}$$

证明. 对于任意给定的定值 $\underline{q} \in \mathbb{R}^N$，定义 $\underline{\bar{u}} := A_M^{-1} B_N^\top \underline{q}$，即对于方程 $q_N \in \pi_N, \bar{u}_M \in X_M$，我们有

$$\langle A\bar{u}_M, \nu_M\rangle = \langle B\nu_M, q_N\rangle, \quad \forall\, \nu_M \in X_M.$$

由有界线性算子 A 的 X-椭圆性，可得

$$\begin{aligned} c_1^A \, \|\bar{u}_M\|_X^2 &\leq \langle A\bar{u}_M, \bar{u}_M\rangle \\ &= \langle B\bar{u}_M, q_N\rangle \\ &= \left(B_N \underline{\bar{u}}, \underline{q}\right) \\ &= \left(\underbrace{B_N A_M^{-1} B_N^\top}_{=:S_N} \underline{q}, \underline{q}\right). \end{aligned}$$

另一方面，由离散稳定性条件(8.336)可得

$$\begin{aligned} c_S \, \|q_N\|_\Pi &\leq \sup_{\underline{0}\neq\underline{\nu}_M\in\mathbb{R}^M} \frac{\langle B\nu_M, q_N\rangle}{\|\nu_M\|_X} \\ &= \sup_{\underline{0}\neq\underline{\nu}_M\in\mathbb{R}^M} \frac{\langle A\bar{u}_M, \nu_M\rangle}{\|\nu_M\|_X} \\ &\leq c_2^A \, \|\bar{u}_M\|_X, \end{aligned}$$

$$\begin{aligned} \hookrightarrow \|q_N\|_\Pi^2 &\leq \left(\frac{c_2^A}{c_S}\right)^2 \|\bar{u}_M\|_X^2 \\ &\leq \left(\frac{1}{c_1^A}\right)\left(\frac{c_2^A}{c_S}\right)^2 \left(B_N A_M^{-1} B_N^\top \underline{q}, \underline{q}\right), \end{aligned}$$

$$\hookrightarrow \left(\underbrace{B_N A_M^{-1} B_N^\top}_{=:S_N} \underline{q}, \underline{q}\right) \underbrace{\geq c_1^A \left(\frac{c_S}{c_2^A}\right)^2}_{=:c_1^{S_N}} \|q_N\|_\Pi^2.$$

□

由舒尔补系统(8.335)唯一可解，可得线性方程系统(8.334)唯一可解。

稳定性测度

线性方程系统(8.334)唯一解的稳定性测度，见下定理。

Theorem 8.46 (唯一解的稳定性测度)**.** 设有界线性算子 $A: X \mapsto X', B: X \mapsto \Pi$，$A$ 是 X-椭圆的。对于协调检测空间 $X_M \subset X, \Pi_N \subset \Pi$，假设离散稳定条件(8.336)成立。则鞍点变分问题(8.333)存在唯一近似解 $(u_M, p_N) \in X_M \times \Pi_N$，对应的稳定性测度如下

$$\|p_N\|_\Pi \leq \frac{1}{c_1^{S_N}} \frac{c_2^B}{c_1^A} \|f\|_{X'} + \frac{1}{c_1^{S_N}} \|g\|_{\Pi'}, \tag{8.338}$$

$$\|u_M\|_X \leq \left(1 + \frac{c_2^B}{c_1^{S_N}} \frac{c_2^B}{c_1^A}\right) \|f\|_{X'} + \frac{1}{c_1^{S_N}} \frac{c_2^B}{c_1^A} \|g\|_{\Pi'}. \tag{8.339}$$

证明. 设线性方程系统(8.334)和伽辽金变分问题(8.333)的唯一 (近似) 解分别为

$$(\underline{u}, \underline{p}) \in \mathbb{R}^M \times \mathbb{R}^N \leftrightarrow (u_M, p_N) \in X_M \times \Pi_N.$$

1. 由 Lemma 8.36可得

$$\begin{aligned} c_1^{S_N} \|p_N\|_\Pi^2 &\leq (S_N \underline{p}, \underline{p}) \\ &= (B_N A_M^{-1} B_N^\top \underline{p}, \underline{p}) \\ &= (B_N A_M^{-1} \underline{f} - \underline{g}, \underline{p}) \\ &= \langle B\bar{u}_M - g, p_N \rangle \\ &\leq \left[c_2^B \|\bar{u}_M\|_X + \|g\|_{\Pi'}\right] \|p_N\|_\Pi, \end{aligned}$$

$$\hookrightarrow \|p_N\|_\Pi \leq \frac{1}{c_1^{S_N}} \left[c_2^B \|\bar{u}_M\|_X + \|g\|_{\Pi'}\right].$$

在上式中，$\underline{\bar{u}} = A_M^{-1} \underline{f} \in \mathbb{R}^M \leftrightarrow \bar{u}_M \in X_M$ 是如下变分问题的唯一解

$$\langle A\bar{u}_M, \nu_M \rangle = \langle f, \nu_M \rangle, \quad \forall \nu_M \in X_M,$$

由 A 的 X-椭圆性有

$$\|\bar{u}_M\|_X \leq \frac{1}{c_1^A} \|f\|_{X'},$$

则我们有(8.338)成立。

2.

$$\begin{aligned} c_1^A \|u_M\|_X^2 &\leq \langle A u_M, u_M \rangle \\ &= \langle f, u_M \rangle - \langle B u_M, p_N \rangle \\ &\leq \left[\|f\|_{X'} + c_2^B \|p_N\|_\Pi\right] \|u_M\|_X, \end{aligned}$$

$$\hookrightarrow \|u_M\|_X \leq \frac{1}{c_1^A} \left[\|f\|_{X'} + c_2^B \|p_N\|_\Pi\right],$$

将(8.338)代入上式，我们有(8.339)成立。

□

误差测度项

由唯一近似解 $(u_M, p_N) \in X_M \times \Pi_N$ 的稳定性测度 (8.339)-(8.338)可进一步求得误差测度项，见如下定理。

Theorem 8.47 (误差测度项)**.** *设 Theorem 8.46的所有假设条件都成立。则鞍点变分问题(8.333)的唯一近似解 $(u_M, p_N) \in X_M \times \Pi_N$，对应的误差测度项为*

$$\|u - u_M\|_X + \|p - p_N\|_\Pi \le c \left\{ \inf_{\nu_M \in X_M} \|u - \nu_M\|_X + \inf_{q_N \in \Pi_N} \|p - q_N\|_\Pi \right\}. \tag{8.340}$$

证明. 利用协调测试空间 $X_M \subset X, \Pi_N \subset \Pi$，将伽辽金变分式(8.333)与连续鞍点变分式(8.332)相减，可得伽辽金正交

$$\begin{aligned} \langle A(u - u_M), \nu_M \rangle + \langle B\nu_M, p - p_N \rangle &= 0, \\ \langle B(u - u_M), q_N \rangle &= 0, \quad \forall (\nu_M, q_N) \in X_M \times \Pi_N. \end{aligned}$$

那么，对于任意给定的 $(\bar{u}_M, \bar{p}_N) \in X_M \times \Pi_N$，我们有

$$\begin{aligned} \langle A(\bar{u}_M - u_M), \nu_M \rangle + \langle B\nu_M, \bar{p}_N - p_N \rangle &= \langle A(\bar{u}_M - u), \nu_M \rangle + \langle B'(\bar{p}_N - p), \nu_M \rangle, \\ \langle B(\bar{u}_M - u_M), q_N \rangle &= \langle B(\bar{u}_M - u), q_N \rangle. \end{aligned}$$

同样地，由 Theorem (8.46)可得，存在唯一解 $(\bar{u}_M - u_M, \bar{p}_N - p_N) \in X_M \times \Pi_N$，以及对应的稳定性测度。那么对于任意 $(\bar{u}_M, \bar{p}_N) \in X_M \times \Pi_N$，我们有

$$\begin{aligned} \|\bar{p}_N - p_N\|_\Pi &\le c_1 \|A(\bar{u}_M - u_M) + B'(\bar{p}_N - p_N)\|_{X'} + c_2 \|B(\bar{u}_M - u)\|_{\Pi'}, \\ \|\bar{u}_M - u_M\|_X &\le c_3 \|A(\bar{u}_M - u_M) + B'(\bar{p}_N - p_N)\|_{X'} + c_4 \|B(\bar{u}_M - u)\|_{\Pi'}. \end{aligned}$$

根据有界线性算子 A, B, B' 的映射属性，利用三角不等式关系可得，对于任意 $(\bar{u}_M, \bar{p}_N) \in X_M \times \Pi_N$ 都有

$$\begin{aligned} \|\bar{p}_N - p_N\|_\Pi &\le \|p - \bar{p}_N\|_\Pi + \|\bar{p}_N - p_N\|_\Pi \\ &\le \left(1 + c_1 c_2^B\right) \|p - \bar{p}_N\|_\Pi + \left(c_1 c_2^A + c_2 c_2^B\right) \|u - \bar{u}_M\|_X. \end{aligned}$$

采用类似的方式，我们将也可求得 $\|u - u_M\|$ 的不等式关系。二者相加，证得(8.340)。 □

离散稳定条件

采用如(8.35)类似的思路，利用Fortin (1977)，可证得离散稳定条件(8.336)。

Lemma 8.37 (离散稳定条件的证明)**.** *设有界线性算子 $B : X \mapsto \Pi'$，满足连续稳定性条件 Theorem 8.18式8.94。*

如果存在一个有界的投影算子 $P_M : X \mapsto X_M$，满足

$$\begin{aligned} \langle B(\nu - P_M\nu), q_N \rangle &= 0, \quad \forall q_N \in \Pi_N, \\ \|P_M\nu\|_X &\le c_p \|\nu\|_X, \quad \forall \nu \in X, \end{aligned}$$

那么离散稳定条件(8.336)成立，对应 $\tilde{c}_S = \frac{c_S}{c_P}$。

8.8.6 强制算子

对于算子方程 $Au = f, u \in X, f \in X'$ 的近似解，如果线性有界算子 $A : X \mapsto X'$ 不满足 X-椭圆性，那么上述近似解法可能不适用。一个替代的弱假定是 A 的强制性 (coercive operator)，即存在一个紧凑算子 (compact operator)$C : X \mapsto X'$，满足 Gårding 不等式 Definition 8.42式(8.109)

$$\langle (A + C)\,\nu, \nu\rangle \geq C_1^A \left\|\nu\right\|_X^2, \quad \forall \nu \in X.$$

设一个由有限维检测空间组成的序列。对应地，强制算子方程可以改写为如下伽辽金变分问题：寻找解 $u_M \in X_M$，使满足

$$\langle Au_M, \nu_M\rangle = \langle f, \nu_M\rangle, \quad \forall\, \nu \in X_M. \tag{8.341}$$

注意：(8.341)与(8.307)的方程形式一致，但不同之处在于线性有界算子算子 A 的特性从椭圆性扩展到更一般的强制性上来，从而在计算数值解的稳定性和误差测度项时有所不同。

Theorem 8.48 (齐亚引理)**.** *设有界线性强制算子 $A : X \mapsto X_M$，假定满足离散稳定性条件。离散稳定性条件的证明见 Theorem 8.49。*

$$c_S \left\|w_M\right\|_X \leq \sup_{\nu_M \in X_M, \|\nu_M\|_X > 0} \frac{\langle Aw_M, \nu_M\rangle}{\left\|\nu_M\right\|_X}, \quad \forall\, w_M \in X_M, \tag{8.342}$$

那么伽辽金变分问题(8.341)有唯一解 $u_M \in X_M$，对应的稳定性测度和误差项测度分别如下

$$\|u_M\| \leq \frac{1}{c_S}\, \|f\|_{X'}, \tag{8.343}$$

$$\|u - u_M\|_X \leq \left(1 + \frac{c_2^A}{c_S}\right) \inf_{\nu_M \in X_M} \|u - u_M\|_X. \tag{8.344}$$

证明. 1. 考虑一个齐次线性方程系统如下

$$A_M\, \underline{\bar{w}} = 0,$$

对应的齐次变分问题为，寻找解 $\overline{w}_M \in X_M$，使满足

$$\langle A_M \overline{w}_M, \nu_M\rangle = 0, \quad \forall\, \nu_M \in X_M.$$

根据离散稳定性条件(8.342)可得

$$c_S \left\|\bar{w}_M\right\|_X \leq \sup_{\nu_M \in X_M, \|\nu_M\|_X > 0} \frac{\langle A\bar{w}_M, \nu_M\rangle}{\left\|w_M\right\|_X}, \quad \forall\, w_M \in X_M,$$

由此可见

$$\bar{w}_M = 0 \leftrightarrow \underline{\bar{w}}_M \in X_M,$$

因此我们有：线性方程系统 $A_M \underline{u} = \underline{f}$，以及伽辽金变分问题(8.341)都是唯一可解的。

2. 设这个唯一解为 $\underline{u} \in \mathbb{R}^M \leftrightarrow u_M \in X_M$。再一次根据根据离散稳定性条件(8.342)有

$$\begin{aligned} c_S \left\|u_M\right\|_X &\leq \sup_{\nu_M \in X_M, \|\nu_M\|_X > 0} \frac{\langle Au_M, \nu_M\rangle}{\left\|\nu_M\right\|_X} \\ &= \sup_{\nu_M \in X_M, \|\nu_M\|_X > 0} \frac{\langle f, \nu_M\rangle}{\left\|\nu_M\right\|_X}, \end{aligned}$$

证得稳定性测度(8.343)。

3. 对于任意 $w \in X$，定义如下伽辽金变分问题，使得存在唯一近似解 $w_M \in X_M$ 满足

$$\langle Aw_M, \nu_M \rangle = \langle Aw, \nu_M \rangle, \quad \forall \nu_M \in X_M.$$

根据上式，定义一个投影算子 $G_M : X \mapsto X_M$，使 $w_M = G_M w$，满足

$$\begin{aligned} \|G_M w\|_X = \|w_M\|_X &\leq \frac{1}{c_S} \|f\|_{X'} \\ &= \frac{1}{c_S} \|Aw\|_{X'} \leq \frac{c_2^A}{c_S} \|w\|_X . \end{aligned}$$

因此伽辽金变分问题(8.341)的唯一解为 $u_M = G_M u$。

4. 由投影算子 G_M 的性质可得

$$G_M \nu_M = \nu_M, \forall \nu_M \in X_M,$$

那么

$$\begin{aligned} \|u - u_M\|_X &= \|u - \nu_M + G_M \nu_M - u_M\|_X \\ &\leq \|u - \nu_M\|_X + \|G_M (u - \nu_M)\|_X \\ &\leq \left(1 + \frac{c_2^A}{c_S}\right) \|u - \nu_M\|_X , \quad \forall \nu_M \in X_M, \end{aligned}$$

由此证得误差项测度(8.344)。

□

现在我们回头来看离散稳定性条件(8.342)的证明。根据前文的介绍可知，若 A 是椭圆算子，则有下式成立

$$\begin{aligned} c_1^A \|\nu_M\|_X &\leq \frac{\langle Aw_M, w_M \rangle}{\|w_M\|_X} \\ &\leq \sup_{\nu_M \in X_M, \|\nu_M\|_X > 0} \frac{\langle Aw_M, \nu_M \rangle}{\|\nu_M\|_X}, \end{aligned}$$

那么若 A 是强制算子，离散稳定性条件的证明见下定理。

Theorem 8.49 (强制算子的离散稳定性条件)**.** *设一个有界线性算子 $A : X \mapsto X'$ 是强制且单射 (injective) 的。设 $X_M \subset X$ 是一个密集 (dense) 协调检测空间序列。那么，存在一个指数 $M_0 \in \mathbb{N}$，使得对于 $M > M_0$ 的情况，离散稳定性条件(8.342)成立。*

证明. 1. 稳定性、误差测度、收敛性。

设任意给定且固定的 $w_M \in X_M$。对于强制算子 $A : X \mapsto X'$，对应存在一个紧凑算子 $C : X \mapsto X'$，使得有界算子 $D = A + C : X \mapsto X'$ 是 X-椭圆的。对应地，构建如下变分问题，其中 $\bar{\nu} := D^{-1} C w_M \in X$ 是唯一解

$$\langle D\bar{\nu}, \nu \rangle = \langle Cw_M, \nu \rangle, \quad \forall \nu \in X.$$

此外，定义伽辽金变分问题，唯一近似解 $\nu_M \in X_M$ 满足

$$\langle D\bar{\nu}_M, \nu_M \rangle = \langle Cw_M, \nu_M \rangle, \quad \forall \nu_M \in X_M.$$

两个变分问题联立，可得伽辽金正交条件

$$\langle D\left(\overline{\nu}-\overline{\nu}_M\right),\nu_M\rangle=0,\quad \forall\nu_M\in X.$$

根据算子 D 的 X-椭圆特性，对伽辽金变分问题应用齐亚引理 Theorem 8.42，可得稳定性测度

$$\|\overline{\nu}_M\|_X\le\frac{1}{c_1^D}\,\|Cw_M\|_{X'}\le\frac{c_2^C}{c_1^D}\,\|w_M\|_X\,.$$

$$\begin{aligned}\hookrightarrow\|w_M-\overline{\nu}_M\|_X&\le\|w_M|\,|_X-\|\overline{\nu}_M\|_X\\&\le\left(1+\frac{c_2^C}{c_1^D}\right)\|w_M\|_X\,.\end{aligned}$$

进而我们有误差测度项

$$\|\overline{\nu}-\overline{\nu}_M\|_X\le\frac{c_2^D}{c_1^D}\inf_{\nu_M\in X_M}\|\overline{\nu}-\nu_M\|_X\,,$$

即检测空间 X_M 的近似属性成立，因此收敛性(8.315)成立。

2. 设一个检测方程 $\nu_M=w_M-\overline{\nu}_M\in X_M$。我们有

$$\begin{aligned}\langle Aw_M,w_M-\overline{\nu}_M\rangle&=\langle Aw_M,w_M-\overline{\nu}\rangle+\langle Aw_M,\overline{\nu}-\overline{\nu}_M\rangle\\&=\underbrace{\langle Aw_M,w_M-D^{-1}Cw_M\rangle}_{=:\mathcal{A}}+\underbrace{\langle Aw_M,\overline{\nu}-\overline{\nu}_M\rangle}_{=:\mathcal{B}},\end{aligned}$$

- 由于 $A:X\mapsto X'$ 有界且可逆

$$\begin{aligned}\mathcal{A}:=\langle Aw_M,w_M-D^{-1}Cw_M\rangle&=\langle Aw_M,D^{-1}\left(D-C\right)w_M\rangle\\&=\langle Aw_M,D^{-1}Aw_M\rangle\\&\ge c_1^D\,\|Aw_M\|_{X'}^2\\&\ge c_1^D\,c_A\,\|w_M\|_X^2\,.\end{aligned}$$

- 由于伽辽金正交条件

$$\begin{aligned}|\mathcal{B}|:=|\langle Aw_M,\overline{\nu}-\overline{\nu}_M\rangle|&=|\langle\left(D-C\right)w_M,\overline{\nu}-\overline{\nu}_M\rangle|\\&=|\langle Dw_M,\overline{\nu}-\overline{\nu}_M\rangle-\langle Cw_M,\overline{\nu}-\overline{\nu}_M\rangle|\\&=\left|\underbrace{\langle w_M,D\left(\overline{\nu}-\overline{\nu}_M\right)\rangle}_{=0}-\langle w_M-C\left(\overline{\nu}-\overline{\nu}_M\right)\rangle\right|\\&=|\langle w_M,C\left(\overline{\nu}-\overline{\nu}_M\right)\rangle|\\&\le\|w\|_X\,\|C\left(\overline{\nu}-\overline{\nu}_M\right)\|_{X'}\,.\end{aligned}$$

3. 由于算子 C 紧凑，那么存在一个子序列 $\{\overline{\nu}_M\}_{M\in\mathbb{N}}$，满足

$$\lim_{M\to\infty}\frac{\|C\left(\overline{\nu}-\overline{\nu}_M\right)\|_{X'}}{\|\overline{\nu}\|_X}=0,$$

因此存在一个指数 $M_0\in\mathbb{N}$，使得满足

$$\langle Aw_M,w_M-\overline{\nu}_M\rangle\ge\frac{1}{2}c_1^Dc_A\left(1+\frac{c_2^D}{c_1^D}\right)^{-1}\|w_M\|\,\|w_M-\overline{\nu}_M\|_X\,,\quad\forall\,M\ge M_0,$$

因此我们有离散稳定性条件(8.342)成立。

□

8.9　有限元

第8.4节介绍了如何将边界积分问题转换为变分问题，本节进一步介绍求得变分问题近似解的方法。首先介绍如何设定合适的有限维检测空间，然后证明索伯列夫空间的一些近似特性。出于简化的考虑，在这里我们只考虑用低阶多项式来构建基方程；更一般形式的有限元近似方法，可参考 Cheng and Cheng (2005); Brenner and Scott (2008)。

8.9.1　参考元

设一个有界域 $\Omega \subset \mathbb{R}^d, d = 1,2,3$，域的边界可能是多边形 (polygonal, $d = 2$) 或多面体 (polyhedra, $d = 3$)。将某一个 $\overline{\Omega}$ 分解为含有有限个网格 (mesh) τ_ℓ 的序列 $\{\mathcal{T}_N\}_{N\in\mathbb{N}}$，分解 (decomposition)可写作

$$\overline{\Omega} = \overline{\mathcal{T}_N} = \bigcup_{\ell=1}^{N} \overline{\tau}_\ell. \tag{8.345}$$

基本概念

在最简单的例子中，每一个网格对应一个区间 (interval, $d = 1$)，三角形 (triangle, $d = 2$)，或多面体 (tetrahedron, $d = 3$)。$\mathcal{T}_N$ 中全部结点 (nodes)的集合用 $\{x_k\}_{k=1}^{M}$ 来表示。对于 $d = 2,3$ 的情况，边 (edge)用 $\{k_j\}_{j=1}^{K}$ 来表示。分解 τ_ℓ 和对应结点 x_k 的关系，可参考图8.1 。

图 8.1: 有限元 τ_ℓ 及其结点 x_k

定义以下三个集合

- $I(k)$ 表示全部网格 τ_ℓ 的集合

$$I(k) := \{\ell \in \mathbb{N} : x_k \in \overline{\tau}_\ell\}, \quad k = 1,\dots,M,$$

- $J(\ell)$ 表示网格 $\overline{\tau}_\ell$ 中全部结点 x_k 的集合

$$J(\ell) := \{k \in \mathbb{N} : x_k \in \overline{\tau}_\ell\}, \quad \ell = 1,\dots,N,$$

  可见 $\dim J(\ell) = d + 1$,

- $K(j)$ 表示网格 $\overline{\tau}_\ell$ 中全部边 k_j 的集合

$$K(j) := \{\ell \in \mathbb{N} : k_j \in \overline{\tau}_\ell\}, \quad j = 1,\dots,K.$$

如果有限元分解(8.345)中，所有两个相邻元之间都有共享的结点 (d=1,2,3)，边 (d=2,3)，三角形 (d=3)，那么我们称这种分解为容许分解 (admissible decomposition)。以 $d = 2$ 为例，容许和不容许分解的例子见图8.2。

在这里我们只考虑可计算域 Ω 中容许分解的情况。对于某一有限元 τ_ℓ，作如下定义

图 8.2: 容许分解，以及不容许分解的三角分解 ($d = 2$)

- Δ_ℓ 表示体积 (volume)

$$\Delta_\ell := \int_{\tau_\ell} dx,$$

- h_ℓ 表示局部网格尺寸 (local mesh size)

$$h_\ell := \Delta_\ell^{\frac{1}{d}},$$

- d_ℓ 表示 (最大) 直径，即有限元 τ_ℓ 中最长的一条边的长度

$$d_\ell := \sup_{x,y\in\tau_\ell} |x-y|,$$

  显然对于 $d = 1$，我们有 $\Delta_\ell = h_\ell = d_\ell$,

- 对应 d_ℓ，r_ℓ 表示 (可包住) 有限元 τ_ℓ 的最大圆 ($d = 2$) 或球体 ($d = 3$) 的半径。

经容许分解(8.345)后得到的有限元 τ_ℓ，如果直径 d_ℓ 由半径 r_ℓ 约束而一致有界 (uniformly bounded)，即满足

$$d_\ell \le c_F\, r_\ell, \quad \ell = 1, \dots, N,$$

那么我们称 τ_ℓ 为正则型有限元 (regular finite element)，其中常数 c_F 与 $\mathcal{T}_N$：

- 对于 $d = 2$ 的情况

$$\begin{aligned} &\pi r_\ell^2 \le \Delta_\ell = h_\ell^2 \le d_\ell^2 \le c_F^2 r_\ell^2, \\ \hookrightarrow &\pi^{\frac{1}{2}} r_\ell \le h_\ell \le d_\ell \le c_F r_\ell, \end{aligned}$$

- 对于 $d = 3$ 的情况

$$\begin{aligned} &\frac{4}{3}\pi r_\ell^3 \le \Delta_\ell = h_\ell^3 \le d_\ell^3 = c_F^3 r_\ell^3, \\ \hookrightarrow &\left(\frac{4}{3}\pi\right)^{\frac{1}{3}} r_\ell \le h_\ell \le d_\ell \le c_F r_\ell. \end{aligned}$$

我们将最大、最小的局部网格尺寸分别定义为 $h_{\max}, h_{\min}$

$$h_{\max} := \max_{\ell=1,\dots,N} h_\ell,$$

$$h_{\min} := \min_{\ell=1,\dots,N} h_\ell,$$

对应地将全局网格尺寸 (global mesh size)定义为

$$h = h_{\max}.$$

如果 $h_{\max}$ 和 $h_{\min}$ 的比值小于等于一个全常数 $c_G \geq 0$，即以下比值有界

$$\frac{h_{\max}}{h_{\min}} \leq c_G, \tag{8.346}$$

那么我们称对应的分解 $\mathcal{T}_N$ 是全局拟一致的 (globally quasi-uniform)的；$c_G \geq 1$ 的值与 $N \in \mathbb{N}$ 无关。

对应地，如果对于任何两个临近的有限元 τ_ℓ 和 τj 都有

$$\frac{h_\ell}{h_j} \leq c_L, \ell, j = 1, \ldots, N,$$

那么我们称这样的分解 $\mathcal{T}_N$ 为局部拟一致 (locally quasi-uniform)。对于某一组相邻的有限元 $\overline{\tau}_\ell$ 和 $\overline{\tau}_j$，如果 $\overline{\tau}_\ell \cap \overline{\tau}_j$ 包括不少于 1 个结点、1 条边或 1 个三角形，那么我们称二者为相邻有限元 (neighboring)。

一维空间的局部参数化

$d = 1$ 时，每个有限元 τ_ℓ 都可以用局部参数化 (local parameterization)的形式予以描述。尤其是对于 $x \in \tau_\ell$, $\ell_1, \ell_2 \in J(\ell)$，我们有

$$x = x_{\ell_1} + \xi\left(x_{\ell_2} - x_{\ell_1}\right) = x_{\ell_1} + \xi h_\ell, \quad \xi \in (0, 1),$$

其中我们将 $\tau \coloneqq (0\ 1)$ 称为参考元 (reference element)。

考虑这样一个方程 $\nu(x)$, $x \in \tau_\ell$，由上面的定义可得

$$\nu(x) = \nu(x_{\ell_1} + \xi h_\ell) =: \widetilde{\nu}_\ell(\xi), \quad \forall \xi \in \tau,$$

即对于 $x \in \tau_\ell$，我们可以将方程 $\nu(x)$ 识别为参考元中的方程 $\widetilde{\nu}_\ell(\xi)$, $\xi \in \tau$，对应范数

$$\begin{aligned}\|\nu\|^2_{L^2(\tau_\ell)} &= \int_{\tau_\ell} (\nu(x))^2 \, dx = \int_\tau |\widetilde{\nu}_\ell(\xi)|^2 \, h_\ell \, d\xi \\ &= h_\ell \, \|\widetilde{\nu}_\ell(\xi)\|^2_{L^2(\tau)} .\end{aligned}$$

我们经常需要计算识别方程的导数，根据链式法则(8.60)可得

1. 一阶导数

$$\begin{aligned}&\frac{d}{d\xi}\widetilde{\nu}_\ell(\xi) = h_\ell \frac{d}{dx}\nu(x), \\ \hookrightarrow &\frac{d}{dx}\nu(x) = h_\ell^{-1}\frac{d}{d\xi}\widetilde{\nu}_\ell(\xi), \quad x \in \tau_\ell, \xi \in \tau.\end{aligned}$$

2. 推广至高阶导数如 $m \in \mathbb{N}$ 的情况

$$\frac{d^m}{dx^m}\nu(x) = h^{-m}\frac{d^m}{d\xi^m}\widetilde{\nu}_\ell(\xi), \quad x \in \tau_\ell, \xi \in \tau,$$

对应局部范数

$$\left\|\frac{d^m}{dx^m}\nu\right\|^2_{L^2(\tau_\ell)} = h_\ell^{1-2m}\left\|\frac{d^m}{d\xi^m}\widetilde{\nu}_\ell\right\|^2_{L^2(\tau)}, \quad m \in \mathbb{N}_0. \tag{8.347}$$

图 8.3: 有限元 τ_ℓ 和参考元 $\tau(d=2)$

二维空间的局部参数化

$d=2$ 时，有限元 τ_ℓ 和参考元 τ 的关系，见图8.3。

参考元 τ 由下面三角形给出

$$\tau = \left\{\xi \in \mathbb{R}^2 : 0 \le \xi_1 \le 1,\, 0 \le \xi_2 \le 1-\xi_1\right\}. \tag{8.348}$$

对 $x \in \tau_\ell$ 作局部参数化

$$x = x_{\ell_1} + \sum_{i=1}^{2} \xi_i \left(x_{\ell_{i+1}} - x_{\ell_1}\right) = x_{\ell_1} + J_\ell\, \xi, \quad \xi \in \tau,$$

其中 J_ℓ 是个雅各比矩阵

$$J_\ell := \begin{pmatrix} x_{\ell_2,1} - x_{\ell_1,1} & x_{\ell_3,1} - x_{\ell_1,1} \\ x_{\ell_2,2} - x_{\ell_1,2} & x_{\ell_3,2} - x_{\ell_1,2} \end{pmatrix}.$$

由此我们得有限元 τ_ℓ 的面积，同时也是体积

$$\begin{aligned} \Delta_\ell &= \int_{\tau_\ell} ds_x = \int_\tau |\det J_\ell|\, d\xi \\ &= |\det J_\ell| \int_0^1 \int_0^{1-\xi_1} d\xi_2 d\xi_1 = \frac{1}{2} |\det J_\ell|, \\ &\hookrightarrow |\det J_\ell| = 2\Delta_\ell. \end{aligned} \tag{8.349}$$

$\nu(x)$, $x \in \tau_\ell$ 对应的识别方程可表示为

$$\nu(x) = \nu\left(x_{\ell_1} + J_\ell \xi\right) =: \widetilde{\nu}_\ell(\xi), \quad \xi \in \tau.$$

继续应用链式法则

$$\begin{aligned} \nabla_\xi \widetilde{\nu}_\ell(\xi) &= J_\ell^\top \nabla_x \nu(x), \\ \hookrightarrow \nabla_x \nu(x) &= J_\ell^{-\top} \nabla_\xi \widetilde{\nu}_\ell(\xi). \end{aligned}$$

等价范数的测度由下引理给出。

Lemma 8.38 (识别方程等价范的测度 $(d=2)$)**.** $d=2$, $m \in \mathbb{N}_0$ 的情况下，识别方程的等价范满足不等式关系

$$\frac{1}{c_m} (2\Delta_\ell)^{1-m} \left\|\nabla_\xi^m \widetilde{\nu}_\ell\right\|_{L^2(\tau)}^2 \le \left\|\nabla_x^m \nu\right\|_{L^2(\tau_\ell)}^2 \le c_m\, (2\Delta_\ell)^{1-m} \left\|\nabla_\xi^m \widetilde{\nu}_\ell\right\|_{L^2(\tau)}^2, \tag{8.350}$$

其中常数 $c_m = \left(\frac{c_F^2}{\pi}\right)^m$。

证明. 对于 $m \in \mathbb{N}_0$，分三种情况来分别证明。

1. $m = 0$ 时。

$$\begin{aligned}\|\nu\|_{L^2(\tau_\ell)}^2 &= \int_{\tau_\ell} |\nu(x)|^2 \, dx \\ &= \int_\tau |\widetilde{\nu}_\ell(\xi)|^2 \, |\det J_\ell| \, d\xi \\ &= 2\Delta_\ell \, \|\widetilde{\nu}_\ell(\xi)\|_{L^2(\tau)}^2 ,\end{aligned}$$

可证得(8.350)成立。

2. $m = 1$ 时。

$$\begin{aligned}\|\nabla_x \nu\|_{L^2(\tau_\ell)}^2 &= \int_{\tau_\ell} |\nabla_x \nu(x)|^2 \, dx \\ &= \int_\tau \left|J_\ell^{-\top} \nabla_\xi \widetilde{\nu}_\ell(\xi)\right|^2 \, |\det J_\ell| \, d\xi \\ &= 2\Delta_\ell \int_\tau \left(J_\ell^{-\top} \nabla_\xi \widetilde{\nu}_\ell(\xi), \nabla_\xi \widetilde{\nu}_\ell(\xi)\right) \, d\xi \\ &= 2\Delta_\ell \int_\tau \left(J_\ell^{-1} J_\ell^{-\top} \nabla_\xi \widetilde{\nu}_\ell(\xi), \nabla_\xi \widetilde{\nu}_\ell(\xi)\right) \, d\xi \\ &\le 2\Delta_\ell \lambda_{\max}\left(J_\ell^{-1} J_\ell^{-\top}\right) \int_\tau |\nabla_\xi \widetilde{\nu}_\ell(\xi)| \, d_\xi \\ &= 2\Delta_\ell \lambda_{\max}\left(J_\ell^{-1} J_\ell^{-\top}\right) \, \|\nabla_\xi \widetilde{\nu}_\ell(\xi)\|_{L^2(\tau)}^2 ,\end{aligned}$$

以及类似地

$$\|\nabla_x \nu\|_{L^2(\tau_\ell)}^2 \ge 2\Delta_\ell \, \lambda_{\min}\left(J_\ell^{-1} J_\ell^{-\top}\right) \, \|\nabla_\xi \widetilde{\nu}_\ell(\xi)\|_{L^2(\tau)}^2 .$$

以上两个不等式都需要计算矩阵 $\left(J_\ell^{-1} J_\ell^{-\top}\right)$ 的特征值，这可以从 $\left(J_\ell^\top J_\ell\right)$ 开始

$$J_\ell^\top J_\ell = \begin{pmatrix} a^2 & ab\cos\alpha \\ ab\cos\alpha & b^2 \end{pmatrix},$$

其中

$$\begin{aligned} a &\coloneqq |x_{\ell_2} - x_{\ell_1}| , \\ b &\coloneqq |x_{\ell_3} - x_{\ell_1}| , \\ \alpha &\coloneqq \sphericalangle\left(x_{\ell_3} - x_{\ell_1}, x_{\ell_2} - x_{\ell_1}\right) . \end{aligned}$$

$\left(J_\ell^{-1} J_\ell^{-\top}\right)$ 的特征值为[20]

$$\lambda_{1,2} = \frac{1}{2}\left[a^2 + b^2 \pm \left(\left(a^2 - b^2\right)^2 + 4a^2b^2\left(\cos\alpha\right)^2\right)^{\frac{1}{2}}\right]$$

- 对应的上限特征值 λ_1 有
$$\frac{1}{2}\left(a^2 + b^2\right) \le \lambda_1 \le a^2 + b^2,$$
- 两个特征值的乘积，根据(8.349)有
$$\lambda_1\lambda_2 = \det\left|J_\ell^\top J_\ell\right| = \det|J_\ell|^2 = 4\Delta_\ell^2.$$

[20]善用 Mathematica!

- 因此对应的下限特征值 λ_2 有

$$\lambda_2 = \frac{4\Delta_\ell^2}{\lambda_1} \geq \frac{4\Delta_\ell^2}{a^2+b^2}.$$

- 由此我们有

$$\frac{4\Delta_\ell^2}{a^2+b^2} \leq \lambda_{\min}\left(J_\ell^\top J_\ell\right) \leq \lambda_{\max}\left(J_\ell^\top J_\ell\right) \leq a^2+b^2.$$

- 此外根据定义我们有

$$a^2+b^2 \leq 2d_\ell^2 \leq 2c_F^2 r_\ell^2 \leq \frac{2c_F^2}{\pi}\Delta_\ell,$$

  代入上式，整理得

$$\frac{2\pi}{c_F^2}\Delta_\ell \leq \lambda_{\min}\left(J_\ell^\top J_\ell\right) \leq \lambda_{\max}\left(J_\ell^\top J_\ell\right) \leq \frac{2c_F^2}{\pi}\Delta_\ell.$$

所以，$J_\ell^\top J_\ell$ 的逆矩阵 $J_\ell^{-1}J_\ell^{-\top}$ 的特征值满足

$$\begin{aligned}
&\frac{\pi}{c_F^2}\left(2\Delta_\ell\right)^{-1} \leq \lambda_{\min}\left(J_\ell^{-1}J_\ell^{-\top}\right) \leq \lambda_{\max}\left(J_\ell^{-1}J_\ell^{-\top}\right) \leq \frac{c_F^2}{\pi}\left(2\Delta_\ell\right)^{-1}, \\
\hookrightarrow &\left(\frac{1}{c_m}\right)\left(2\Delta_\ell\right)^{-1}\left(2\Delta_\ell\right)^{-1} \leq \lambda_{\min}\left(J_\ell^{-1}J_\ell^{-\top}\right) \leq \lambda_{\max}\left(J_\ell^{-1}J_\ell^{-\top}\right) \leq c_m\left(2\Delta_\ell\right)^{-1},
\end{aligned}$$

因此证得(8.350)。

3. $m > 1$ 时，可通过递归重复上述步骤证得(8.350)。

□

三维空间的局部参数化

$d = 3$ 时，参考元 τ 表现为四面体 (tetraderon) 形式

$$\tau = \left\{\xi \in \mathbb{R}^3 : 0 \leq \xi_1 \leq 1,\, 0 \leq \xi_2 \leq 1-\xi_1,\, 0 \leq \xi_3 \leq 1-\xi_1-\xi_2\right\}.$$

对 $x \in \tau_\ell$ 作局部参数化

$$x = x_{\ell_1} + \sum_{i=1}^{3}\xi_1\left(x_{\ell_{i+1}} - x_{\ell_1}\right) = x_{\ell_1} + J_\ell\xi, \quad \xi \in \tau,$$

其中雅各比矩阵

$$J_\ell = \begin{pmatrix} x_{\ell_2,1}-x_{\ell_1,1} & x_{\ell_3,1}-x_{\ell_1,1} & x_{\ell_4,1}-x_{\ell_1,1} \\ x_{\ell_2,2}-x_{\ell_1,2} & x_{\ell_3,2}-x_{\ell_1,2} & x_{\ell_4,2}-x_{\ell_1,2} \\ x_{\ell_2,3}-x_{\ell_1,3} & x_{\ell_3,3}-x_{\ell_1,3} & x_{\ell_4,3}-x_{\ell_1,3} \end{pmatrix}.$$

有限元的体积

$$\begin{aligned}
\Delta_\ell &= \int_{\tau_\ell} ds_x = \int_\tau \left|\det J_\ell\right| d\xi \\
&= \left|\det J_\ell\right| \int_0^1 \int_0^{1-\xi_1} \int_0^{1-\xi_1-\xi_2} d\xi_3\, d\xi_2\, d\xi_1 \\
&= \frac{1}{6}\left|\det J_\ell\right|,
\end{aligned}$$

$$\hookrightarrow \left|\det J_\ell\right| = 6\Delta_\ell. \tag{8.351}$$

用识别方程 $\widetilde{\nu}_\ell(\xi)$, $\xi \in \tau$ 来识别 $\nu(x)$, $x \in \tau_\ell$

$$\nu(x) = \nu\left(x_{\ell_1} + J_\ell \xi\right) =: \widetilde{\nu}_\ell(\xi), \quad \xi \in \tau.$$

根据链式法则，可得导数

$$\begin{aligned}
&\nabla_\xi \widetilde{\nu}_\ell(\xi) = J_\ell^\top \nabla_x \nu(x), \\
&\hookrightarrow \nabla_x \nu(x) = J_\ell^\top \nabla_\xi \widetilde{\nu}_\ell(\xi).
\end{aligned}$$

等价范数的测度见下引理。

Lemma 8.39 (识别方程等价范的测度 (d=3))**.** $d = 3$, $m \in \mathbb{N}_0$ 的情况下，识别方程的等价范满足不等式关系

$$c_1 \Delta_\ell h_\ell^{-2m} \left\|\nabla_\xi^m \widetilde{\nu}_\ell\right\|_{L^2(\tau)}^2 \le \left\|\nabla_x^m \nu\right\|_{L^2(\tau_\ell)}^2 \le c_2 \Delta_\ell h_\ell^{-2m} \left\|\nabla_\xi^m \widetilde{\nu}_\ell\right\|_{L^2(\tau)}^2 \tag{8.352}$$

其中常数 $(c_1, c_2) > 0$，可能与 m 和 c_F 有关。

证明. 分三种情况来证明。

1. $m = 0$ 时。

$$\begin{aligned}
\|\nu(x)\|_{L^2(\tau_\ell)}^2 &= \int_{\tau_\ell} |\nu(x)|^2 \, dx \\
&= \int_\tau |\widetilde{\nu}_\ell(\xi)|^2 \, |\det J_\ell| \, d\xi \\
&= 6\Delta_\ell \left\|\widetilde{\nu}_\ell(\xi)\right\|_{L^2(\tau)},
\end{aligned}$$

可证得(8.352)成立。

2. $m = 1$ 时。

$$\begin{aligned}
\|\nabla_x \nu\|_{L^2(\tau_\ell)}^2 &= \int_{\tau_\ell} |\nabla_x \nu(x)|^2 \, dx \\
&= \int_\tau |\det J_\ell| \left|J_\ell^{-\top} \nabla_\xi \widetilde{\nu}_\ell(\xi)\right|^2 d\xi \\
&= |\det J_\ell| \int_\tau \left(J_\ell^{-\top} \nabla_\xi \widetilde{\nu}_\ell(\xi), J_\ell \nabla_\xi \widetilde{\nu}_\ell(\xi)\right) d\xi \\
&= 6\Delta_\ell \int_\tau \left(J_\ell^{-1} J_\ell^{-\top} \nabla_\xi \widetilde{\nu}_\ell(\xi), \nabla_\xi \widetilde{\nu}_\ell(\xi)\right) d\xi
\end{aligned}$$

$$\hookrightarrow \|\nabla_x \nu\|_{L^2(\tau_\ell)}^2 \begin{cases} \ge 6\Delta_\ell \lambda_{\min}\left(J_\ell^{-1} J_\ell^{-\top}\right) \|\nabla_\xi \widetilde{\nu}_\ell\|_{L^2(\tau)}^2 \\ \le 6\Delta_\ell \lambda_{\max}\left(J_\ell^{-1} J_\ell^{-\top}\right) \|\nabla_\xi \widetilde{\nu}_\ell\|_{L^2(\tau)}^2 \end{cases}$$

计算 $\left(J_\ell^\top J_\ell\right)$ 矩阵的特征值

$$J_\ell^\top J_\ell = \begin{pmatrix} a^2 & ab\cos\alpha & ac\cos\beta \\ ab\cos\alpha & b^2 & bc\cos\gamma \\ ac\cos\beta & bc\cos\gamma & c^2 \end{pmatrix},$$

其中

$$
\begin{aligned}
a &:= \left|x_{\ell_2} - x_{\ell_1}\right|, \\
b &:= \left|x_{\ell_3} - x_{\ell_1}\right|, \\
b &:= \left|x_{\ell_4} - x_{\ell_1}\right|, \\
\alpha &:= \sphericalangle\left(x_{\ell_2} - x_{\ell_1}, x_{\ell_3} - x_{\ell_1}\right), \\
\beta &:= \sphericalangle\left(x_{\ell_2} - x_{\ell_1}, x_{\ell_4} - x_{\ell_1}\right), \\
\gamma &:= \sphericalangle\left(x_{\ell_3} - x_{\ell_1}, x_{\ell_4} - x_{\ell_1}\right),
\end{aligned}
$$

对应的三个特征值 $\lambda_i > 0$, $i = 1, 2, 3$，满足

$$
\begin{aligned}
\sum_{i=1}^{3} \lambda_i &= \operatorname{trace}\left(J_\ell^\top J_\ell\right) = a^2 + b^2 + c^2, \\
\prod_{i=1}^{3} \lambda_i &= \det\left(J_\ell^\top J_\ell\right) = \left|\det J_\ell\right|^2 = 36\Delta_\ell
\end{aligned}
$$

- 最大特征值的上限
$$\lambda_{\max}\left(J_\ell^\top J_\ell\right) \le a^2 + b^2 + c^2.$$
- 因此可得最小特征值的下限
$$\lambda_{\min}\left(J_\ell^\top J_\ell\right) \ge \frac{\prod_{i=1}^{3} \lambda_i}{\lambda_{\max}^2\left(J_\ell^\top J_\ell\right)} \ge \frac{36\Delta_\ell^2}{\left(a^2 + b^2 + c^2\right)^2}$$
- 因此我们有
$$\frac{36\Delta_\ell^2}{\left(a^2 + b^2 + c^2\right)^2} \le \lambda_{\min}\left(J_\ell^\top J_\ell\right) \le \lambda_{\max}\left(J_\ell^\top J_\ell\right) \le a^2 + b^2 + c^2.$$

由于假定 τ_ℓ 是正则型有限元 (regular shape)，则所有边的长度对应

$$a^2 + b^2 + c^2 \le 3d_\ell^2 \le 3c_F^2 r_\ell^2 \le 3\left(\frac{4}{3}\pi\right)^{-\frac{2}{3}} c_F^2 h_\ell^2.$$

代入上式可得

$$\frac{4}{c_F^4}\left(\frac{4}{3}\pi\right)^{\frac{4}{3}} h_\ell^2 \le \lambda_{\min}\left(J_\ell^\top J_\ell\right) \le \lambda_{\max}\left(J_\ell^\top J_\ell\right) \le 3\left(\frac{4}{3}\pi\right)^{-\frac{2}{3}} c_F^2 h_\ell^2.$$

可证得(8.352)成立。

3. $m > 1$ 时，可通过递归重复上述步骤证得(8.352)。

□

识别方程等价范的测度不等式

综上，识别方程等价范的测度不等式，$d = 1$ 时为(8.347)，$d = 2$ 时为 Lemma 8.38式(8.350)，$d = 3$ 时为 Lemma 8.39式(8.352)。那么可以得到如下定理

Theorem 8.50 (识别方程等价范的测度不等式)**.** *设一个正则型的容许分解 $\mathcal{T}_N$ 中的有限元 $\tau_\ell \subset \mathbb{R}^d$。如果 ν 充分平滑，对于 $m \in \mathbb{N}_0$ 我们有*

$$c_1 \Delta_\ell h_\ell^{-2m} \left\|\nabla_\xi^m \widetilde{\nu}_\ell\right\|_{L^2(\tau)}^2 \le \left\|\nabla_x^m \nu\right\|_{L^2(\tau_\ell)}^2 \le c_2 \Delta_\ell h^{-2m} \left\|\nabla_\xi^m \widetilde{\nu}_\ell\right\|_{L^2(\tau)}^2,$$

其中常数 $(c_1, c_2) > 0$，并且可能与 m 和 c_F 有关。

8.9.2　形式方程

基于分解元 $\mathcal{T}_N$(8.345)，现在来分析检测空间 (trial spaces)。检测空间由分段多项式构成，这些多项式组成不同形式的基方程 (base functions)。基方程与全局自由度相关，不过更是在局部定义的：是针对有限元 τ_ℓ，通过选取合适的形式方程 (form function)而作局部定义。

考虑这样一个参考元 τ，它可以是一个区间 $(d=1)$，三角形 $(d=2)$ 或是四面体 $(d=3)$。

常数型形式方程

最简单的形式方程可以是个常数

$$\psi_1^0(\xi)=1,\quad \xi\in\tau,$$

对应地，如果有限元 τ_ℓ 中的某个方程 $\nu_h(x)$, $x\in\tau_\ell$ 是个常数，那么 ν 的表现形式可写为

$$\nu_h(x)=\nu_h\left(x_{\ell_1}+J_\ell\xi\right)=\nu_\ell\,\psi_1^0(\xi),\quad x\in\tau_\ell,\,\xi\in\tau,$$

其中系数 ν_ℓ 反映 $\nu_h\in\tau_\ell$ 在参考元 τ 中对应的值。由此可得

$$\|\nu_h\|^2_{L^2(\tau_\ell)}=\Delta_\ell\nu_\ell^2. \tag{8.353}$$

线性形式方程

如果 $\nu_h(x)$, $x\in\tau_\ell$ 是个线性方程，那么 ν_h 可由参考元 τ 中结点的值 $\widetilde{\nu}_k$ 所唯一决定

$$\widetilde{\nu}_h\left(\xi\right)=\sum_{k=1}^{d+1}\widetilde{\nu}_k\psi_k^1\left(\xi\right),\quad \xi\in\tau, \tag{8.354}$$

其中 ψ_k^1 的值如下

- $d=1$ 时结点数 $k=2$

$$\begin{cases}\psi_1^1 & :=1-\xi,\\ \psi_2^1 & :=\xi.\end{cases}$$

- $d=2$ 时结点数 $k=3$

$$\begin{cases}\psi_1^1 & :=1-\xi_1-\xi_2,\\ \psi_2^1 & :=\xi_1,\\ \psi_3^1 & :=\xi_2.\end{cases}$$

- $d=3$ 时结点数 $k=4$

$$\begin{cases}\psi_1^1 & :=1-\xi_1-\xi_2-\xi_3,\\ \psi_2^1 & :=\xi_1,\\ \psi_3^1 & :=\xi_2,\\ \psi_4^1 & :=\xi_3.\end{cases}$$

这样一来我们有：设任意一个有限元 τ_ℓ，对应结点 x_{ℓ_k}, $\ell_k\in J(\ell)$。τ_ℓ 中的线性方程 $\nu_h(x)$, $x\in\tau_\ell$ 可以写为

$$\nu_h\left(x\right)=\nu_h\left(x_{\ell_1}+J_\ell\xi\right)=\sum_{k=1}^{d+1}\nu_{\ell_k}\psi_k^1\left(\xi\right),\quad x\in\tau_\ell,\xi\in\tau. \tag{8.355}$$

与常数形式方程的范数(8.353)类似，线性形式方程 $\nu_h(x)$, $x\in\tau_\ell$ 的范数 $\|\nu_h\|_{L^2}\left(\tau_\ell\right)$ 也可以用结点的值 $\nu_\ell(\xi)$, $\xi\in\tau$ 来表示，见如下引理。

Lemma 8.40 (线性形式方程的范数). 设线性方程 $\nu_h(x), x \in \tau_\ell$ 如(8.355)所示。那么我们有范数不等式关系

$$\frac{\Delta_\ell}{(d+1)(d+2)} \sum_{k=1}^{d+1} \nu_{\ell_k}^2 \leq \|\nu_h\|_{L^2(\tau_\ell)}^2 \leq \frac{\Delta_\ell}{(d+1)} \sum_{k=1}^{d+1} \nu_{\ell_k}^2 \tag{8.356}$$

证明. 线性方程 $\nu_h(x) \in L^2(\tau_\ell)$ 的范数

$$\begin{aligned}
\|\nu_h\|_{L^2(\tau_\ell)}^2 &= \langle \nu_h, \nu_h \rangle_{L^2(\tau_\ell)} \\
&= \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \nu_i \nu_j \int_\tau \psi_i(\xi) \psi_j(\xi) |\det J_\ell| \, d\xi \\
&= \left(G_\ell \underline{\nu}^\ell, \underline{\nu}^\ell \right),
\end{aligned}$$

其中 G_ℓ 是局部质量矩阵 (local mass matrix)

$$G_\ell = \frac{\Delta_\ell}{(d+1)(d+2)} \underbrace{\left(I_{d+1} + \underline{e}_{d+1} \underline{e}_{d+1}^\top \right)}_{=:\mathcal{A}}, \quad \underline{e}_{d+1} = \underline{1} \in \mathbb{R}^{d+1}.$$

提取 G_ℓ 矩阵 $\mathcal{A}$，计算特征值，Mathematica 中代码如下

```
d = 3 (*三维系统d=3。改为2或1，对应二维、一维系统*)
ee = Table[1, d + 1, d + 1]
et = Transpose[ee]
ii = IdentityMatrix[d + 1]
new = ii + ee*et
Eigenvalues[new]
```

可得 $\lambda_1[\mathcal{A}] = 1, \lambda_2[\mathcal{A}] = \ldots = \lambda_{d+1}[\mathcal{A}] = 1$。进而证得(8.356)。 □

许多应用中需要将 $\nu_h(x)$ 的斜率与 $\nu_h(x)$ 自身关联起来，见下引理。

Lemma 8.41 (方程范数与方程斜率的范数). 设线性方程 ν_h 如(8.355)所给定。则以下局部逆不等式关系成立

$$\|\nabla_x \nu_h\|_{L^2(\tau_\ell)} \leq c_I h_\ell^{-1} \|\nu_h\|_{L^2(\tau_\ell)}, \tag{8.357}$$

其中常数 $c_I > 0$。

证明. 分以下几步来证明。

第一步。由 Theorem 8.50 可得 $\nu_h(x), x \in \tau_\ell$ 方程斜率的范，与形式方程 $\widetilde{nu}_\ell(\xi), \xi \in \tau$ 方程斜率范的关系

$$\|\nabla_x \nu_h\|_{L^2(\tau_\ell)}^2 \leq c_2 \Delta_\ell h_\ell^{-2} \|\nabla_\xi \widetilde{\nu}_\ell\|_{L^2(\tau)}^2.$$

那么下面需要计算线性方程

$$\widetilde{\nu}_\ell(\xi) = \sum_{k=1}^{d+1} \nu_{\ell_k} \psi_k^1(\xi)$$

的斜率以及范数。分 $d = 1, 2, 3$ 三种情况。

- $d = 1$ 时，

$$\nabla_\xi \widetilde{\nu}_\ell = \nu_{\ell_2} - \nu_{\ell_1},$$

$$\|\nabla_\xi \widetilde{\nu}\|_{L^2(\tau_\ell)}^2 = (\nu_{\ell_2} - \nu_{\ell_2})^2 \leq 2\left[\nu_{\ell_2}^2 + \nu_{\ell_1}^2\right] \leq 4 \|\nu_h\|_{L^\infty(\tau_\ell)}^2.$$

- $d=2$ 时，

$$\nabla_\xi \widetilde{\nu}_\ell = \begin{pmatrix} \nu_{\ell_2} - \nu_{\ell_1} \\ \nu_{\ell_3} - \nu_{\ell_1} \end{pmatrix},$$

$$\begin{aligned} \|\nabla_\xi \widetilde{\nu}_\ell\|_{L^2(\tau)}^2 &= \frac{1}{2}\left[(\nu_{\ell_2} - \nu_{\ell_1})^2 + (\nu_{\ell_3} - \nu_{\ell_1})^2\right] \\ &\le \frac{1}{2}\left[2\nu_{\ell_2}^2 + 2\nu_{\ell_3}^2 + 4\nu_{\ell_4}^2\right] \\ &\le 4\,\|\nu_h\|_{L^\infty(\tau_\ell)}. \end{aligned}$$

- $d=3$ 时

$$\nabla_\xi \widetilde{\nu}_\ell = \begin{pmatrix} \nu_{\ell_2} - \nu_{\ell_1} \\ \nu_{\ell_3} - \nu_{\ell_1} \\ \nu_{\ell_4} - \nu_{\ell_1} \end{pmatrix},$$

$$\begin{aligned} \|\nabla_\xi \widetilde{\nu}_\ell\|_{L^2(\tau)}^2 &= \frac{1}{6}\left[(\nu_{\ell_2} - \nu_{\ell_1})^2 + (\nu_{\ell_3} - \nu_{\ell_1})^2 + (\nu_{\ell_4} - \nu_{\ell_1})^2\right] \\ &\le \frac{1}{6}\left[2\nu_{\ell_2}^2 + 2\nu_{\ell_3}^2 + 2\nu_{\ell_4}^2 + 6\nu_{\ell_1}^2\right] \\ &\le 4\,\|\nu_h\|_{L^\infty(\tau_\ell)}^2. \end{aligned}$$

汇总可得

$$\begin{aligned} \|\nabla_x \nu_h\|_{L^2(\tau_\ell)}^2 &\le c_2 \Delta_\ell h_\ell^{-2}\,\|\nabla_\xi \widetilde{\nu}_\ell\|_{L^2(\tau)}^2 \\ &\le 4c_2 \Delta_\ell h_\ell^2\,\|\nu_h\|_{L^\infty(\tau_\ell)}^2, \end{aligned}$$

$$\hookrightarrow \|\nabla_x \nu_h\|_{L^2(\tau_\ell)} \le c_I h_\ell^{-1}\,\|\nu_h\|_{L^2(\tau_\ell)}.$$

□

二次形式方程

更高幂次的非线性局部形式方程，可由相应的低幂次形式方程逐级地推定义而得。以二次形式方程为例，定义如下：

- $d=1$ 时，

$$\begin{cases} \psi_1^2(\xi) := 1 - \xi, \\ \psi_2^2(\xi) := \xi, \\ \psi_3^2(\xi) := 4\xi\,(1-\xi). \end{cases}$$

- $d=2$ 时，

$$\begin{cases} \psi_1^2(\xi) := 1 - \xi_1 - \xi_2, \\ \psi_2^2(\xi) := \xi_1, \\ \psi_3^2(\xi) := \xi_2, \\ \psi_4^2(\xi) := 4\xi_1\,(1 - \xi_1 - \xi_2), \\ \psi_5^2(\xi) := 4\xi_1\xi_2, \\ \psi_6^2(\xi) := 4\xi_2\,(1 - \xi_1 - \xi_2). \end{cases}$$

- $d=3$ 时，

$$\begin{cases}\psi_1^2(\xi) := 1-\xi_1-\xi_2-\xi_3,\\ \psi_2^2(\xi) := \xi_1,\\ \psi_3^2(\xi) := \xi_2,\\ \psi_4^2(\xi) := \xi_3,\\ \psi_5^2(\xi) := 4\xi_1\left(1-\xi_1-\xi_2-\xi_3\right),\\ \psi_6^2(\xi) := 4\xi_1\xi_2,\\ \psi_7^2(\xi) := 4\xi_2\left(1-\xi_1-\xi_2-\xi_3\right),\\ \psi_8^2(\xi) := 4\xi_3\left(1-\xi_1-\xi_2-\xi_3\right),\\ \psi_9^2(\xi) := 4\xi_3\xi_1,\\ \psi_{10}^2(\xi) := 4\xi_2\xi_3.\end{cases}$$

注意，线性形式方程与结点 $x_k \in \overline{\tau}_\ell$ 的自由度有关，而非线性形式方程则与边的中值点 $x_{k_j}^*$ 有关。若方程 x_h 在 τ_ℓ 中是二次形式的，那么可以写作

$$\begin{aligned}\nu_h(x) &= \nu_h\left(x_{\ell_1}+J_\ell\xi\right)\\ &= \sum_{k=1}^{\frac{1}{2}(d+1)(d+2)} \nu_{\ell_k}\psi_k^2(\xi), \quad x\in\tau_\ell,\ \xi\in\tau.\end{aligned} \tag{8.358}$$

类似于 Lemma 8.40，二次形式方程的范数可写作

$$\begin{aligned}\|\nu_h\|_{L^2(\tau_\ell)}^2 &= \sum_{i,j=1}^{\frac{1}{2}(d+1)(d+2)} \nu_i\nu_j\int_\tau \psi_i^2(\xi)\psi_j^2(\xi)\ \left|\det J_\ell\right|\ d\xi\\ &= \left(G_\ell\underline{\nu}^\ell,\underline{\nu}^\ell\right),\end{aligned}$$

- $d=1$ 时，局部质量矩阵 G_ℓ 为

$$G_\ell = \begin{pmatrix}\frac{1}{3} & \frac{1}{6} & \frac{1}{3}\\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{3} & \frac{8}{15}\end{pmatrix},$$

对应特征值计算得

$$\begin{aligned}\lambda_1 &= \frac{\Delta_\ell}{6},\\ \lambda_2 &= \Delta_\ell\left[\frac{31}{60}+\frac{\sqrt{89}}{20}\right]\\ \lambda_2 &= \Delta_\ell\left[\frac{31}{60}+\frac{\sqrt{89}}{20}\right].\end{aligned}$$

- $d=2,3$ 时，
  - 等价范数的估计与线性形式方程情况下的 Lemma 8.40式(8.356)一致

$$c_1\Delta_\ell\sum_{k=1}^{\frac{1}{2}(d+1)(d+2)}\nu_{\ell_k}^2 \le \|\nu_h\|_{L^2(\tau_\ell)}^2 \le c_2\Delta_\ell\sum_{k=1}^{\frac{1}{2}(d+1)(d+2)}\nu_{\ell_k}^2, \tag{8.359}$$

    其中二次形式方程 ν_h 的定义如(8.358)。
  - 逆不等式也与线性形式下的情况 Lemma 8.41式(8.357)相一致。

8.9.3　检测空间

对于二阶偏微分方程的边界值问题，为了计算近似解，需要我们构建检测空间 $\mathcal{S}_h^1(\mathcal{T}_N)$，空间中包括局部分段线性方程和全局连续方程。

对于容许分解(8.345)来说，在分解产生的结点 x_k 处，$\mathcal{S}_h^1(\mathcal{T}_N)$ 便由结块方程 (nodal function) 所唯一确定，$\nu_k = \nu_h(x_k)$。通过这样的方式，可以将有限元 τ_ℓ 用局部形式方程予以表现。

全局检测空间 $\mathcal{S}_h^1(\mathcal{T}_N)$ 的维数 $\dim \mathcal{S}_h^1(\mathcal{T}_N)$ 也因此等于分解后的结点数。对应地，$\mathcal{S}_h^1(\mathcal{T}_N)$ 的基 $\psi_k^1(x)$ 如图所示，可以写为

$$\psi_k^1(x) := \begin{cases} 1 & x = x_k, \\ 0 & x = x_\ell \neq x_k, \\ \text{线性} & \text{其他情况}. \end{cases}$$

图 8.4: 检测空间 $\mathcal{S}_h^1(\mathcal{T}_N)$ 的基 $\psi_k^1(x)$

分段线性方程 $\nu_h(x) \in \mathcal{S}_h^1(\mathcal{T}_N)$ 可以表示为

$$\nu_h(x) = \sum_{k=1}^{M} \nu_k \psi_k^1(x).$$

进而检测空间中的范数见下引理

Lemma 8.42 (光谱等价范数不等式)**.** *$\nu_h(x) \in \mathcal{S}_h^1(\mathcal{T}_N)$ 在分解中的光谱等价范数不等式为*

$$\frac{1}{(d+1)(d+2)} \sum_{k=1}^{M} \left(\sum_{\ell \in I(k)} \Delta_\ell \right) \nu_k^2 \leq \|\nu_h\|_{L^2(\mathcal{T}_N)} \leq \frac{1}{(d+1)} \sum_{k=1}^{M} \left(\sum_{\ell \in I(k)} \Delta_\ell \right) \nu_k^2. \tag{8.360}$$

证明. 由 Lemma 8.40可得

$$\begin{aligned} \|\nu_h\|_{L^2(\mathcal{T}_N)}^2 &= \sum_{\ell}^{N} \|\nu_h\} \|\nu_h\|_{L^2(\mathcal{T}_\ell)}^2 \\ &\leq \sum_{\ell=1}^{N} \frac{\Delta_\ell}{d+1} \sum_{k=1}^{d+1} \nu_{\ell_k}^2 \\ &= \frac{1}{d+1} \sum_{k=1}^{M} \left(\sum_{\ell \in I(k)} \Delta_\ell \right) \nu_k^2 \end{aligned}$$

为上限界。采用类似的方法可得下限界。证得(8.360)。 □

Lemma 8.43 (逆等价范数不等式)**.** $\nu_h(x) \in \mathcal{S}_h^1(\mathcal{T}_N)$ 在分解中的逆等价范数不等式为

$$\|\nabla_x \nu_h\|_{L^2(\mathcal{T}_N)}^2 \leq c_I \sum_{\ell=1}^{N} h_\ell^{-2} \|\nu_h\|_{L^2(\tau_\ell)}^2 . \tag{8.361}$$

如果分解是全局拟一致的，见(8.346)，那么我们有

$$\|\nabla_x \nu_h\|_{L^2(\mathcal{T}_N)} \leq c\, h^{-1} \|\nu_h\|_{(\mathcal{T}_N)} . \tag{8.362}$$

证明. 可由 Lemma 8.41直接求得。 □

为了求得检测空间 $\mathcal{S}_h^1(\mathcal{T}_N)$ 的近似特性，首先要定义相应的插值算子和投影算子，然后计算误差测度项。

检测空间的插值算子

设 $\nu \in C(\mathcal{T}_N)$ 是一个全局连续的方程。在含有分段线性方程的空间中定义插值如下

$$I_h \nu(x) := \sum_{k=1}^{M} \nu(x_k)\varphi_k(x) \in \mathcal{S}_h^1(\mathcal{T}_N) . \tag{8.363}$$

Lemma 8.44 (分段线性插值的局部误差测度项)**.** 设给定的 $\nu_{|\tau_\ell} \in H^2(\tau_\ell)$。则我们有局部误差测度项

$$\|\nu - I_h \nu\|_{L^2(\tau_\ell)} \leq c\, h_\ell^2 \, |\nu|_{h^2(\tau_\ell)} . \tag{8.364}$$

证明. 1. 根据识别方程等价范的测度不等式 Theorem 8.50可得

$$\|\nu - I_h \nu\|_{L^2(\tau_\ell)} \leq c\, \Delta_\ell \, \|\widetilde{\nu}_\ell - I_\tau \widetilde{\nu}_\ell\|_{L^2(\tau)} ,$$

其中 $I_\tau : H^2(\tau) \mapsto L^2(\tau)$ 是相应参考元 τ 的线性插值算子，满足

$$\|I_\tau \widetilde{\nu}_\ell\|_{L^2(\tau)} \leq \mathrm{meas}(\tau) \, \|\widetilde{\nu}_\ell\|_{L^\infty(\tau)} ,$$

$\mathrm{meas}(\tau)$ 表示 τ 的测度 (measure)。由索伯列夫嵌入定理 (第225页8.2.4节) 可得

$$\|\widetilde{\nu}_\ell\|_{L^\infty(\tau)} \leq c \, \|\widetilde{\nu}_\ell\|_{L^2(\tau)} ,$$

可见 $I_\tau : H^2(\tau) \mapsto L^2(\tau)$ 有界。

2. 对于任意且固定的方程 $\omega \in L^2(\tau)$，可定义一个线性泛函如下

$$f(u) := \int_\tau [(I - I_\tau)\, u(\xi)]\, \omega(\xi) d\xi.$$

如果给定方程 $u \in H^2(\tau)$，那么我们有

$$\begin{aligned} |f(u)| &= \left| \int_\tau [(I - I_\tau)\, u(\xi)]\, \omega(\xi)\, d\xi \right| \\ &\leq \|(I - I_\tau)\, u\|_{L^2(\tau)} \, \|\omega\|_{L^2(\tau)} \\ &\leq c \, \|u\|_{H^2(\tau)} \, \|\omega\|_{L^2(\tau)} , \end{aligned}$$

即线性泛函 f 有界。

对于任一线性方程 $q \in \mathcal{P}_1(\tau)$，我们有 $I_\tau q = q$，因此

$$f(q) = 0, \quad \forall\, q \in \mathcal{P}_1(\tau).$$

可见 Bramble-Hilbert 引理 Lemma 8.1的全部前提条件均得到满足。根据该引理可得

$$|f(u)| \le \widetilde{c}\, \|\omega\|_{L^2(\tau)}\, |u|_{H^2(\tau)}\,.$$

3. 定义 $u(\xi)$ 和 $\omega(\xi)$ 的值如下

$$\begin{cases} u := \widetilde{\nu}_\ell, \\ \omega := (I - I_\ell)\, \widetilde{\nu}_\ell, \end{cases}$$

进而

$$\begin{aligned} \|\omega\|^2_{L^2(\tau)} &= \int_\tau (\omega(\xi))^2\, d\xi = \int_\tau [(I - I_\tau)\, \widetilde{\nu}_\ell(\xi)]\, \omega(\xi)\, d\xi = \int_\tau [(I - I_\tau)\, u(\xi)]\, \omega(\xi)\, d\xi = |f\,(\widetilde{\nu}_\ell)| \\ &\le \widetilde{c}\, \|\omega\|_{L^2(\tau)}\, |\widetilde{\nu}_\ell|_{H^2(\tau)} = \widetilde{c}\, \|(I - I_\tau)\, \widetilde{\nu}_\ell\|_{L^2(\tau)}\, |\widetilde{\nu}_\ell|_{H^2(\tau)}\,, \\ &\hookrightarrow \|(I - I_\tau)\, \widetilde{\nu}_\ell\|_{L^2(\tau)} \le \widetilde{c}\, |\widetilde{\nu}_\ell|_{H^2(\tau)}\,. \end{aligned}$$

4. 代入 Theorem 8.50可得

$$\begin{aligned} \|\nu - I_h \nu\|_{L^2(\tau)} &\le c\, \Delta_\ell\, \|\widetilde{\nu}_\ell\|_{H^2(\tau)} \\ &\le \hat{c}\, h_\ell^2\, |\nu|_{H^2(\tau_\ell)}\,. \end{aligned}$$

□

由 Lemma 8.44可得分段线性插值的全局误差项

$$\|\nu - I_h \nu\|^2_{L^2(\mathcal{T}_N)} \le c \sum_{\ell=1}^{\ell} h_\ell^4\, |\nu|^2_{H^2(\tau_\ell)}\,, \tag{8.365}$$

以及

$$\|\nu - I_h \nu\|^2_{H^1(\mathcal{T}_N)} \le c \sum_{\ell=1}^{\ell} h_\ell^2\, |\nu|^2_{H^2(\tau_\ell)}\,. \tag{8.366}$$

检测空间的投影算子

为了应用插值算子执行上述运算，需要假定待差值的方程 $\nu \in C\,(\mathcal{T}_N)$ 是全局连续的，这个假设过于强硬，有时较难满足。一个替代方案是以下若假设：

第二部分

模型估计

第三部分

待分类

第九章　最大似然估计

> The maximum-likelihood procedure in any problem is what you are most likely to do if you don’t know any statistics.
>
> *Harrison H. Barrett*
> *Foundations of Image Science*

9.1　简介

设一个 $M \times 1$ 的向量 g，描述一组随机数据。对应地，设一个 $K \times 1$ 的参数向量 θ，用来描述这组随机数据 g。那么，概率密度方程 (probability density function, PDF) $pr(g|\theta)$ 用于表示，在给定 θ 的情况下，随机数据 g 的抽样分布 (sampling distribution)。

反之，若在现实中已经观测到了 g，则我们可将 $pr(g|\theta)$ 看成是一个关于参数 θ 的方程，称为似然方程 (likelihood function)，定义为 $L(\theta|g)$

$$L(\theta|g) = pr(g|\theta),$$

需要注意的是，$L(\theta|g)$ 并不是 θ 的概率密度方程。

我们将对 θ 的估计命名为 $\hat{\theta}$。$\hat{\theta}$ 是个关于 g 的方程：由于 g 是随机数据向量，导致 $\hat{\theta}(g)$ 也是个关于 g 的随机变量向量。本质上来说，参数估计反映的是一个映射过程

$$\begin{cases} \theta \mapsto g \mapsto \hat{\theta}, \\ pr(\theta) \mapsto pr(g|\theta) \mapsto pr(\hat{\theta}|\theta). \end{cases}$$

从这个角度来说，$\hat{\theta}$ 是一个关于 θ 的随机变量。

9.2　近似的测度

估计值 $\hat{\theta}$ 相比于真实值 θ 往往存在误差。误差常常来自于两个方面：

- 随机误差，又称 precision：如估计偏差，噪音等。常用方差 (viariance) 与以量化描述。
- 系统误差，又称 accuracy：如估计偏误，校准误差，模型误设定等。常用偏误 (bias) 或均方误差 (MSE, EMSE) 予以量化描述。

两种误差均可由估计值的条件分布 $pr(\hat{\theta}|\theta)$ 予以量化。

9.2.1 估计的偏误：标量参数

以标量参数 θ 及其 (随机) 估计值 $\hat{\theta}$ 为例，估计的条件均值 $\overline{\hat{\theta}}$ 可由似然方程计算而得

$$\overline{\hat{\theta}} = \left\langle \hat{\theta}(g) \right\rangle_{g|\theta} = \int d^{M_g}\, pr(g|\theta)\hat{\theta}(g), \tag{9.1}$$

其中 $\langle\cdot\rangle_{g|\theta}$ 表示在给定的 θ 下，各个随机估计值 $\hat{\theta}(g)$ 的平均值。

如果已知估计 $\hat{\theta}$ 的条件概率密度 $pr\left(\hat{\theta}|\theta\right)$，那么(9.1)可进一步改写为

$$\overline{\hat{\theta}} = \int d\hat{\theta}\, pr\left(\hat{\theta}|\theta\right)\, \hat{\theta}. \tag{9.2}$$

偏误

我们将 $\overline{\hat{\theta}}$ 相对于 θ 的偏差称为关于真实值的条件偏误 (conditional bias on true parameter)，定义为 $b(\theta)$

$$b(\theta) = \overline{\hat{\theta}} - \theta. \tag{9.3}$$

如果某个估计 $\overline{\hat{\theta}}$，对于任意 θ 都满足 $b(\theta) = 0$，那我们称之为无偏估计 (unbiased estimate)。

可测度性

可测度性 (estimability) 或可识别性 (identifiability)。如果无论随机数据集 g 的取值如何，都存在关于参数 θ 的无偏估计 $\hat{\theta}$，满足 $b(\theta) = 0 \quad \forall\theta$，那我们称参数 θ 是可估计的 (estimable) 或可识别的 (identifiable)，表示为 $\overline{\overline{\hat{\theta}}}$

$$\overline{\overline{\hat{\theta}}} = \int d\theta\, pr(\theta)\, \overline{\hat{\theta}} = \int d\theta\, pr(\theta) \int d^{M_g}\, pr(g|\theta)\, \hat{\theta}(g). \tag{9.4}$$

测度的方差和均方误差

某一测度 $\hat{\theta}$ 相对于测度均值 $\overline{\theta}$ 的波动，我们用方差 $var\left(\hat{\theta}\right)$ 来表示

$$var\left(\hat{\theta}\right) = \sigma_{\hat{\theta}}^2 = \left\langle \left|\hat{\theta}(g) - \overline{\hat{\theta}}\right|^2 \right\rangle_{g|\theta}. \tag{9.5}$$

某一测度 $\hat{\theta}$ 相对于真实值 θ 的波动，我们用均方误差 (mean square error, MSE)来表示

$$MSE(\theta) = \left\langle \left|\hat{\theta} - \theta\right|^2 \right\rangle_{g|\theta}. \tag{9.6}$$

考虑到 θ 随 g 而具有的随机性特征，将均方误差再严 θ 取平均值，得整体均方误差 (ensemble mean square error, EMSE)

$$EMSE(\theta) = \left\langle \left\langle \left|\hat{\theta} - \theta\right|^2 \right\rangle_{g|\theta} \right\rangle_{\theta}. \tag{9.7}$$

9.2.2 估计的偏误：向量参数

设一个 P 维的参数向量 θ，对应估计值 $\hat{\theta}$。我们将平均值 $\overline{\hat{\theta}}$(有时也写作 $\langle\theta\rangle$) 定义如下

$$\overline{\hat{\theta}}\,(g) = \int d^{M_g}\, pr\,(g|\theta)\, \hat{\theta}\,(g) = \int d^{P_{\hat{\theta}}}\, pr\left(\hat{\theta}|\theta\right)\, \hat{\theta}\,(g)\,. \tag{9.8}$$

平均偏误

平均偏误 $\overline{b}(\theta)$ 定义为

$$\overline{b}(\theta) = \langle b(\theta) \rangle_\theta \tag{9.9}$$

测度的方差和均方误差

我们将 P 维随机估计向量 $\hat{\theta}$ 中，第 p 个元素的均值表示为 $\langle \hat{\theta} \rangle_p = \overline{\hat{\theta}}_p$。对应的方差为

$$\begin{aligned} var\left(\hat{\theta}_p\right) &= \left\langle \left[\hat{\theta}_p - \langle \hat{\theta} \rangle\right] \left[\hat{\theta}_p - \langle \hat{\theta} \rangle\right]^\top \right\rangle_{g|\theta} \\ &= \int_\infty d^{M_g} \left|\hat{\theta}_p(g) - \langle \hat{\theta}_p(g) \rangle\right|^2 pr(g|\theta) \\ &= \int_\infty d^{P_\theta} \left|\hat{\theta}_p - \langle \hat{\theta}_p \rangle\right|^2 pr\left(\hat{\theta}|\theta\right). \end{aligned} \tag{9.10}$$

将所有 $p \in P$ 个方差汇总在一起，得到整体的方差协方差矩阵，定义为 $K_{\hat{\theta}}$

$$K_{\hat{\theta}} = \left\langle \left(\hat{\theta} - \overline{\theta}\right) \left(\hat{\theta} - \overline{\theta}\right)^\dagger \right\rangle = \left\langle \Delta\hat{\theta}\, \Delta\hat{\theta}^\dagger \right\rangle. \tag{9.11}$$

参数向量的均方误差

$$\begin{aligned} MSE &= \left\langle \left\|\hat{\theta} - \theta\right\|^2 \right\rangle_{g|\theta} \\ &= \int_\infty d^{M_g} \left\|\hat{\theta}(g) - \theta\right\|^2 pr(g|\theta) \\ &= \mathrm{trace}\left(K_{\hat{\theta}}\right) + \mathrm{trace}\left(b(\theta) b(\theta)^\dagger\right) \end{aligned} \tag{9.12}$$

9.3　费雪信息矩阵和克拉美罗下界

参数估计的方差协方差矩阵 K，存在一些限定条件。

9.3.1　标量参数

任意标量参数的无偏估计,其方差都必须满足克拉美罗下界 (Cramér-Rao lower bound, CRLB)

$$var\left(\hat{\theta}\right) \geq \underbrace{\frac{1}{\left\langle \left[\frac{\partial}{\partial\theta} \log pr(g|\theta)\right]^2 \right\rangle_{g|\theta}}}_{=:F}, \tag{9.13}$$

其中我们将分母的部分称为费雪信息矩阵 (Fischer information matrix)。

若估计是有偏的，克拉美罗下界为

$$var\left(\hat{\theta}\right) \geq \frac{\left(\frac{db(\theta)}{d\theta} + 1\right)^2}{\left\langle \left[\frac{\partial}{\partial\theta} \log pr(g|\theta)\right]^2 \right\rangle_{g|\theta}} \tag{9.14}$$

9.3.2 向量参数

对于 P 维向量参数 θ，费雪信息矩阵 F 成为 $P \times P$ 维的埃米特矩阵 (Hermitian matrix)，其中第 jk 个元素 F_{jk} 表示为

$$\begin{aligned} F_{jk} &= \left\langle \left[\frac{\partial}{\partial\theta_j}\log pr\left(g|\theta\right)\right]\left[\frac{\partial}{\partial\theta_k}\log pr\left(g|\theta\right)\right]\right\rangle_{g|\theta} \\ &= \int_{\infty} d^{M_g}\, pr\left(g|\theta\right)\left[\frac{1}{pr\left(g|\theta\right)}\frac{\partial}{\partial\theta_j}pr\left(g|\theta\right)\right]\left[\frac{1}{pr\left(g|\theta\right)}\frac{\partial}{\partial\theta_k}pr\left(g|\theta\right)\right]. \end{aligned} \tag{9.15}$$

在得到费雪信息矩阵 F 的基础上，来看 (参数向量 θ 的无偏估计向量 $\hat{\theta}$ 的) 方差协方差矩阵 $K_{\hat{\theta}}$。$n \times n$ 矩阵 $K_{\hat{\theta}}$ 中的第 nn 个元素，等于 $\hat{\theta}$ 向量中第 nn 个元素的方差，并且也满足克拉美洛下界

$$\left[K_{\hat{\theta}}\right]_{nn} = var\left(\hat{\theta}_{nn}\right) \geq \left[F^{-1}\right]_{nn}, \tag{9.16}$$

这就是说，为了计算无偏估计向量 $\hat{\theta}$ 中第 n 个元素方差的下界 (或方差协方差矩阵 $K_{\hat{\theta}}$ 中第 nn 元素的下界)，我们需要首先计算费雪信息矩阵的逆矩阵，提取其中的第 nn 个元素。

一个满足克拉美罗下界条件的无偏估计是有效估计 (efficient)。

9.4 最大似然估计

前面讨论了一个参数估计 $\hat{\theta}$ 所应当具有的特性。随后的问题就是，如何找到这个估计值? 常见的方法之一是最大似然估计 (maximum likelihood estimation)

$$\hat{\theta}_{ML} \equiv \mathrm{argmax}_{\theta}\, pr\left(g|\theta\right), \tag{9.17}$$

有时也用对数似然方程来表示

$$\hat{\theta}_{ML} \equiv \underset{\theta}{\mathrm{argmax}} \log pr\left(g|\theta\right), \tag{9.18}$$

即是说，选择合适的参数值 $\hat{\theta}_{ML}$，使得在 $\hat{\theta}_{ML}$ 下可能实际观测到数据集合 g 的概率最大。

没有其他额外限定条件的情况下，$\hat{\theta}_{ML}$ 常取值于 (对数) 似然方程对 θ 的梯度 (gradient)，即一阶导数等于 0 的点

$$\underbrace{\nabla_{\theta}\log pr\left(g|\theta\right)}_{=:s(g)} = 0, \quad \leftrightarrow \theta = \hat{\theta}_{ML}, \tag{9.19}$$

我们将这个表示斜率的随机向量 $s(g)$ 称为 score，用于描述 (对数) 似然方程对参数变化的敏感程度，因此有时也称敏感方程 (sensitivity function)。

$s(g)$ 的方差协方差矩阵就等于费雪信息矩阵的逆。

9.4.1 最大似然估计：to be or not to be?

1. 选择最大似然估计法的理由

   (a) 有效。如果的确存在一个有效的最大似然估计值。

   (b) 渐进有效。随着观测数据的数量增加，估计效果逐渐提升。

   (c) 渐进无偏。

   (d) 渐进一致。

   (e) 便于计算。相对来说。

(f) 有助于严谨地强制使得估计结果与实际观测数据相一致。

(g) 无需先验信息 $pr(\theta)$ 设定。与之相反地，参考贝叶斯估计。

2. 不选择最大似然估计法的理由

(a) (严谨地强制使得估计结果与实际观测数据相一致)。
现实中观测到的数据常常是充满噪音的。强行严谨一致容易导致计算出的最大似然估计值，哪怕是无偏的，也随之充满噪音。

(b) (无需先验信息 $pr(\theta)$ 设定)。
现实中总是存在着一些先验信息 $pr(\theta)$，我们应当充分利用这些信息，哪怕他们有时并非完全无偏[1]。

9.5　伯努利实验：最大似然估计法

现在我们从伯努利实验 (Bernoulli trials)入手，介绍如何应用最大似然估计法求解相关问题。基于本节的求解思路，我们随后在 python 环境中编写程序求解，见第9.6节。

手头一枚一元硬币，现在我们连续掷 10 次，有 7 次头朝上。我们想要利用最大似然法，估计每一次投掷得到头朝上的概率。

用 $p \in (0,1)$ 来表示每次投掷得到头朝上的概率 (参数的实际值)。那么连续投掷 N 次，在这 N 次中得到 $n \in \mathbb{N}$ 次头朝上的概率可以表示为

$$pr\left(n|p\right) = \binom{N}{n} p^n \left(1-p\right)^{N-n}, \tag{9.20}$$

其中二项式系数满足

$$\binom{N}{n} = \frac{N!}{n!\,(N-n)!}. \tag{9.21}$$

随机变量 n 的均值和方差分别为

$$\langle n \rangle = Np, \tag{9.22}$$

$$\sigma^2 = Np(1-p). \tag{9.23}$$

在上一组实验中我们观测到的数据是投掷 N 次得到 n 次头 (如 10 次得到 7 次头)。我们想要知道如果反复重复这组实验，出现 N/n 的最大概率 p 是多少？可用最大似然法进行估计：改变 p 的值，使得 $pr(n|p)$ 达到最大。

对 $pr(n|p)$ 取对数，然后对 p 求导，得 score 方程 $s(n)$

$$\begin{aligned} \log pr(n|p) &= constant + n\log p + (N-n)\log(1-p), \\ s(n) &= \frac{d}{dp}\log pr(n|p) = \frac{n}{p} + \frac{N-n}{1-p}. \end{aligned} \tag{9.24}$$

设 $s(n)=0$，对应的 $p=\hat{p}_{ML}$ 即为 p 的最大似然估计

$$\hat{p}_{ML} = \frac{n}{N}, \tag{9.25}$$

[1]一种将先验 $pr(\theta)$ 与最大似然估计相结合的方法称为加权似然法 (weighted likeihood)，如

$$\hat{\theta}_{WL} \equiv \underset{\theta}{\operatorname{argmax}}\, pr\left(g|\theta\right) pr\left(\theta\right) = \underset{\theta}{\operatorname{argmax}}\, pr\left(\theta|g\right).$$

分子表示观测到头的次数，分母表示总的投掷次数。

最大似然估计 $\hat{p}_{ML}$ 的均值和方差，由(9.22)-(9.23)得

$$\langle \hat{p}_{ML} \rangle = \frac{\langle n \rangle}{N}, \tag{9.26}$$

$$var\left(\hat{p}_{ML}\right) = \frac{var(n)}{N^2} = \frac{p(1-p)}{N} \tag{9.27}$$

据此可得估计的费雪信息矩阵

$$F = \left\langle \left[\frac{n}{p} - \frac{N-n}{1-p}\right]^2 \right\rangle = \frac{\langle [n-Np]^2 \rangle}{p^2\left(1-p\right)^2} = \frac{var(n)}{p^2\left(1-p\right)^2} = \frac{N}{p\left(1-p\right)}, \tag{9.28}$$

因此我们有

$$F^{-1} = var\left(\hat{p}_{ML}\right), \tag{9.29}$$

并且最大似然估计 $\hat{p}_{ML}$ 是有效的。

9.6 伯努利实验：最大似然估计的 Python 程序实现

伯努利实验 (频率法) 在 Python 环境下的实现

伯努利实验(Bernoulli trials) 是指重复若干次 (设为 T 次) 某一项实验 (或称迭代 iteration)，每次实验 (iteration) 的结果可能是二者之一：成功 (0) 或者失败 (1)。假定我们现在来检测成功的概率，可以基于频率学派的思路，使用最大似然估计法来实现 (maximum likelihood, MLB)。

```
In [2]: # 初始设定
    %matplotlib inline
    from __future__ import division
    import numpy as np
    import pandas as pd
    import statsmodels.api as sm
    import sympy as sp
    import pymc3
    import matplotlib.pyplot as plt
    import matplotlib.gridspec as gridspec
    from mpl_toolkits.mplot3d import Axes3D
    from scipy import stats
    from scipy.special import gamma
    from sympy.interactive import printing
    printing.init_printing()
```

9.6.1 模型构建

设一个伯努利实验

$$y \sim Bernoulli(\theta) = Binomial(1, \theta). \tag{9.30}$$

我们将概率密度方程，或称边际似然方程，写为

$$p(y|\theta) = \theta^y \ (1-\theta)^{1-y} = \begin{cases} \theta & y = 1 \\ 1-\theta & y = 0. \end{cases} \tag{9.31}$$

设 $\theta = 0.3$，随机进行 100 次试验。

In [3]:
```
# 模拟数据
np.random.seed(123)

nobs = 100
theta = 0.3
# numpy.random.binomial(n, p, size=None)
# https://docs.scipy.org/doc/numpy-1.13.0/reference/generated/numpy.random.binomial.html
Y = np.random.binomial(1, theta, nobs)
```

In [4]:
```
# 画图

fig = plt.figure(figsize=(10,5))
gs = gridspec.GridSpec(1, 2, width_ratios = [5,1])
ax1 = fig.add_subplot(gs[0])
ax2 = fig.add_subplot(gs[1])

ax1.plot(range(nobs), Y, "x")
ax2.hist(-Y, bins=2)

ax1.yaxis.set(ticks=(0,1), ticklabels=("Failure", "Success"))
ax2.xaxis.set(ticks=(-1,0), ticklabels=("Success", "Failure"))

ax1.set(title=r"Bernoulli Trial Outcomes $(\theta = 0.3)$", xlabel="Trial", ylim=(-0.2, 1.
ax2.set(ylabel="Frequency")

fig.tight_layout()
```

9.6.2 似然方程

现在假定从随机变量 100 个 y 中，随机提取 T 个值，构成一个样本

$$Y = (y_1, \ldots y_T)^\top \tag{9.32}$$

对应的联合似然方程为

$$p(Y|\theta) = \prod_{i=1}^{T} \theta^{y_i}\ (1-\theta)^{1-y_i} = \theta^s\ (1-\theta)^{T-s}, \quad s = \sum_i y_i, \tag{9.33}$$

s 就是观测到的成功的次数 (对应 $y=1$)；$T-s$ 是观测到的失败的次数 ($y=0$)。
随后的目标就是，如何将似然方程最大化?

9.6.3 最大似然估计：解析法

一种较为直观的方法是解析法。

- 似然方程 $L(\theta; Y)$

$$L(\theta; Y) = \theta^s\ (1-\theta)^{T-s}. \tag{9.34}$$

- 对数似然方程 $\log(\theta; Y)$

$$\log L(\theta; Y) = s \log\theta + (T-s)\log(1-\theta). \tag{9.35}$$

- score 矩阵 (一阶导数)

$$\frac{\partial \log L(\theta; Y)}{\partial\theta} = \frac{s}{\theta} - \frac{T-s}{1-\theta} = s\,\theta^{-1} - (T-s)\ (1-\theta)^{-1}. \tag{9.36}$$

- 黑塞矩阵 (二阶导数)

$$H(\theta) = \frac{\partial^2 \log L(\theta; Y)}{\partial\theta^2} = -s\,\theta^{-2} - (T-s)\ (1-\theta)^{-2}. \tag{9.37}$$

- 信息矩阵

$$\begin{aligned} I[\theta] &= -E[H(\theta)] = -E\left[-s\,\theta^{-2} - (T-s)\ (1-\theta)^{-2}\right] \\ &= (\theta\,T)\ \theta^{-2} + (T-\theta\,T)\ (1-\theta)^{-2} = \frac{T}{\theta(1-\theta)}. \end{aligned} \tag{9.38}$$

- 方差协方差矩阵 (逆信息矩阵)

$$I[\theta]^{-1} = \frac{\theta(1-\theta)}{T}. \tag{9.39}$$

```
In [5]: t, T, s = sp.symbols("theta, T, s")

        # 构建符号方程
```

```
likelihood = (t ** s) * ((1 - t)**(T - s))
loglike = s * sp.log(t) + (T - s) * sp.log(1-t)
score = (s/t) - ((T-s)/(1-t))
hessian = - s/(t**2) - (T - s)/((1-t)**2)
information = T/(t * (1-t))

# 利用 sympy.lambdify 将符号方程转换为 numpy.lambda
# sympy.utilities.lambdify.implemented_function(symfunc, implementation)
# http://docs.sympy.org/latest/modules/utilities/lambdify.html
_likelihood = sp.lambdify((t,T,s), likelihood, modules='numpy')
_loglike = sp.lambdify((t,T,s), loglike, modules='numpy')
_score = sp.lambdify((t,T,s), score, modules='numpy')
_hessian = sp.lambdify((t,T,s), hessian, modules='numpy')
```

对数似然方程 $\log L(\theta;Y)$ 对 θ 求偏导，值为 0 处对应的 θ 值定义为 $\widehat{\theta}_{MLE}$

$$\begin{aligned} \left.\frac{\partial \log L(\theta;Y)}{\partial\theta}\right|_{\widehat{\theta}_{MLE}} &= 0, \\ \hookrightarrow \widehat{\theta}_{MLE} &= \frac{s}{T}. \end{aligned} \tag{9.40}$$

对应地

$$H\left(\widehat{\theta}_{MLE}\right) = -\frac{T^2}{s} - \frac{T^2}{T-s} < 0, \quad \Leftarrow T > s, \tag{9.41}$$

二阶导数小于 0 确保最大似然条件成立。

In [6]:
```
# 计算极大似然估计系数 $\theta$ 的解析值

theta_hat_analytic = Y.sum() / nobs
var_analytic = (theta_hat_analytic * (1 - theta_hat_analytic)) / nobs

print('Analytic MLE Results: %.2f (%.10f)' % (theta_hat_analytic, var_analytic ** 0.5))
```

```
Analytic MLE Results: 0.22 (0.0414246304)
```

9.6.4　最大似然估计：数值法

在很多情况下，我们无法通过解析法求得最大化的解。这时需要数值近似求解，利用scipy.optimize最优算法。以下几点需要注意。

- 对似然方程的负数求最小化，达到最大化的效果。
- 最优算法常常是在不受约束的参数空间 $\mathbb{R}^k$ 中进行的 (k 表示参数数量即维度)。如果实际研究对象的参数值是受限的，如本例 $\theta \in (0,1)$，则需要将受限参数首先转换为无约束空间中的参数 $(0,1) \ni \theta \to \varphi \in \mathbb{R}$，然后使用最优算法求解。对应的转换矩阵

$$\theta = g(\varphi) = \frac{\exp(\varphi)}{1+\exp(\varphi)}, \quad \varphi \in \mathbb{R}. \tag{9.42}$$

- 只有得到了黑塞矩阵，根据逆黑塞矩阵的负数才能计算方差协方差矩阵。而黑塞矩阵的解析形往往无法求得——可以利用scipy.optimize.fmin_bfgs数值近似黑塞矩阵。
- 利用数值近似黑塞矩阵算得的方差协方差矩阵 $Cov(\varphi)$，需要转换为 $Cov(\theta)$

$$Cov(\theta) = \left.\frac{\partial g(\varphi)}{\partial \varphi}\right|_{\varphi_{MLE}} \times Cov(\varphi) \times \left.\frac{\partial g(\varphi)}{\partial \varphi}^{\top}\right|_{\varphi_{MLE}} \tag{9.43}$$

转换矩阵的偏导数 (斜率) 为

$$\frac{\partial g(\varphi)}{\partial \varphi} = \frac{\partial \left[\exp(\varphi)\ (1+\exp(\varphi))^{-1}\right]}{\partial \varphi} = \frac{\exp(\varphi)}{(1+\exp(\varphi))^2} \tag{9.44}$$

9.6.5 无约束参数 $g(\varphi)$ 构成的模型系统

- 逆转换方程

$$\begin{aligned} \theta = g(\varphi) &= \frac{\exp(\varphi)}{1+\exp(\varphi)}, \\ 1+\exp(\varphi)\,\theta &= \exp(\varphi), \\ \theta &= \exp(\varphi)\,(1-\theta), \\ \varphi &= \log\left(\frac{\theta}{1-\theta}\right) = g^{-1}(\theta) \end{aligned} \tag{9.45}$$

- 似然方程 $L(\theta; Y)$

$$L(g(\varphi); Y) = g(\varphi)^s\ (1-g(\varphi))^{T-s} \tag{9.46}$$

- 对数似然方程 $\log(g(\varphi); Y)$

$$\log L(g(\varphi); Y) = s \log g(\varphi) + (T-s) \log\left(1-g(\varphi)\right) \tag{9.47}$$

- score 矩阵 (一阶导数)

$$\begin{aligned} \frac{\partial \log L(g(\varphi); Y)}{g(\varphi)} &= \frac{\partial \log L(g(\varphi); Y)}{g(\varphi)}\,\frac{\partial g(\varphi)}{\partial \varphi} \\ &= \left(\frac{s}{g(\varphi)} - \frac{T-s}{1-g(\varphi)}\right)\frac{\exp(\varphi)}{(1+\exp(\varphi))^2} \\ &= \left(\frac{s}{\frac{\exp(\varphi)}{1+\exp(\varphi)}} - \frac{T-s}{1-\frac{\exp(\varphi)}{1+\exp(\varphi)}}\right)\frac{\exp(\varphi)}{(1+\exp(\varphi))^2} \\ &= s - T\,\frac{\exp(\varphi)}{1+\exp(\varphi)} = s - T\,g(\varphi). \end{aligned} \tag{9.48}$$

- 黑塞矩阵 (二阶导数)

$$H(\varphi) = \frac{\partial^2 \log L(g(\varphi); Y)}{\partial g(\varphi)^2} = -T. \tag{9.49}$$

- 信息矩阵

$$I[\varphi] = -E[H(\varphi)] = T. \tag{9.50}$$

- 方差协方差矩阵

$$I[\varphi]^{-1} = \frac{1}{T}. \tag{9.51}$$

Delta 近似法

进而我们有 Delta 近似法

$$\begin{aligned}
I_{Delta}[\widehat{\theta}_{MLE}]^{-1} &= \left.\frac{\partial g(\varphi)}{\partial \varphi}\right|_{\varphi_{MLE}} \times I[\varphi]^{-1} \times \left.\frac{\partial g(\varphi)}{\partial \varphi}^{\top}\right|_{\varphi_{MLE}} \\
&= \left\{ \frac{\exp\left[\log\left(\frac{\widehat{\theta}_{MLE}}{1-\widehat{\theta}_{MLE}}\right)\right]}{\left(1+\exp\left[\log\left(\frac{\widehat{\theta}_{MLE}}{1-\widehat{\theta}_{MLE}}\right)\right]\right)^2} \right\}^2 \frac{1}{T} = \left[\left(\frac{\widehat{\theta}_{MLE}}{1-\widehat{\theta}_{MLE}}\right)\left(1+\frac{\widehat{\theta}_{MLE}}{1-\widehat{\theta}_{MLE}}\right)^{-2}\right]^2 \frac{1}{T} \\
&= \left[\frac{\widehat{\theta}_{MLE}}{1-\widehat{\theta}_{MLE}}\left(1-\widehat{\theta}_{MLE}\right)^2\right]^2 \frac{1}{T} = \left[\widehat{\theta}_{MLE}\left(1-\widehat{\theta}_{MLE}\right)\right]^2 \frac{1}{T} \\
&= I\left[\widehat{\theta}_{MLE}\right]^{-1}\left[\widehat{\theta}_{MLE}\left(1-\widehat{\theta}_{MLE}\right)\right]
\end{aligned} \tag{9.52}$$

Delta 近似容易导致方差协方差矩阵被低估。可以写出误差项

$$I[\theta]^{-1} - I_{Delta}[\theta]^{-1} = \frac{\theta(1-\theta) - [\theta(1-\theta)]^2}{T}, \tag{9.53}$$

或者误差项的偏移 (相对误差)

$$\frac{\left|I[\theta]^{-1} - I_{Delta}[\theta]^{-1}\right|}{I[\theta]^{-1}} = \frac{\frac{\theta(1-\theta)-[\theta(1-\theta)]^2}{T}}{\frac{\theta(1-\theta)}{T}} = 1-\theta(1-\theta) \tag{9.54}$$

```
In [7]: fig = plt.figure(figsize = (12,4))

        # 误差项
        ax = fig.add_subplot(121, projection="3d")
        delta = 0.025
        T_range = np.arange(0,100,10)
        # numpy.linspace(start, stop, num=50, endpoint=True, retstep=False, dtype=None)[source]
        # https://docs.scipy.org/doc/numpy-1.13.0/reference/generated/numpy.linspace.html
        TH_range = np.linspace(0,1)
        # numpy.meshgrid(*xi, **kwargs)[source]
        # https://docs.scipy.org/doc/numpy-1.13.0/reference/generated/numpy.meshgrid.html
        # [X,Y] = meshgrid(x,y) 将向量 x 和 y 定义的区域转换成矩阵 X 和 Y,
        # 这两个矩阵可以用来表示 mesh 和 surf 的三维空间点以及两个变量的赋值。
        # 其中矩阵 X 的行向量是向量 x 的简单复制，而矩阵 Y 的列向量是向量 y 的简单复制。
        T_grid, TH_grid = np.meshgrid(T_range, TH_range)
        # 误差项的测算
        E_grid = (TH_grid * (1 - TH_grid) - (TH_grid * (1-TH_grid)) ** 2)/T_grid
        ax.plot_wireframe(T_grid, TH_grid, E_grid, color="#348ABD")
```

```
    ax.set(
        title = "Absolute Error from the Delta Method",
        xlabel= "T",
        ylabel= r"$\theta$"
    )
    ax.xaxis.set(pane_color = (1,1,1))
    ax.yaxis.set(pane_color = (1,1,1))

    # 误差项的偏移（相对误差）
    ax = fig.add_subplot(122)
    X = np.linspace(0,1)
    ax.plot(X, 1 - X * (1-X))
    ax.set(
        title = "Relative Error from the Delta Method",
        xlabel = r"$\theta$",
        ylim = (0.7, 1)
    )
    ax.vlines(theta_hat_analytic, 0.7, 1, "r")
```

Out[7]: <matplotlib.collections.LineCollection at 0x125ab96a0>

- 左图可以看出，绝对误差项随着样本数量 T 的增加而逐渐增大，相对误差的变化则与 T 无关。
- 右图可以看出，在最大似然估计的解析解 ($\theta = 0.22$) 处，相对误差的值约为 82%。

最大似然估计：算法举例

In [8]:
```
# 载入
from scipy import optimize

# 构建转换方程
transform = lambda phi: np.exp(phi) / (1 + np.exp(phi))
```

```
## 转换方程的斜率（一阶导数）
transform_grad = lambda phi: np.exp(phi) / ((1 + np.exp(phi)) **2)

# 构建优化所调用的方程。为负，是因为我们使用最小化的算法。
f = lambda params: -_loglike(
    transform(params[0]), nobs, Y.sum()
)
fprime = lambda params: np.matrix(
    -_score(
        transform(params[0]), nobs, Y.sum()
    )
)
fhess = lambda params: np.matrix(
    -_hessian(
        transform(params[0]), nobs, Y.sum()
    )
)
```

scipy.optimize 可以使用一系列算法，这里以两种方法为例作简要介绍。

第一种方法叫做 牛顿共轭梯度法 (Newton Conjugate Gradient Method, Newton-CG method)。

```
In [9]: # 利用牛顿共轭梯度法作最大似然估计

#%timeit res_ncg = optimize.fmin_ncg(f, np.array([0.5]), fprime, fhess=fhess, disp=False)
print ("Timing: 10 loops, best of 3: 4.89 ms per loop")

res_ncg = optimize.fmin_ncg (f, np.array([0.5]), fprime, fhess = fhess)[0]
var_ncg = np.linalg.inv(
    -_hessian(transform(res_ncg), nobs, Y.sum())
)[0,0]
#var_ncg = np.linalg.inv(-_hessian(transform(res_ncg), nobs, Y.sum()))[0,0]
print ("Newton-CG MLE estimate: %.2f (%.10f)" % (transform(res_ncg), var_ncg**0.5))
```

```
Timing: 10 loops, best of 3: 4.89 ms per loop
Optimization terminated successfully.
     Current function value: 52.690796
     Iterations: 35
     Function evaluations: 51
     Gradient evaluations: 85
     Hessian evaluations: 35
Newton-CG MLE estimate: 0.22 (0.0414252425)
```

第二种方法叫做 BFMG 法(Broyden–Fletcher–Goldfarb–Shanno algorithm, BFMG method)。

```
In [10]: # 利用 BFMG 法作最大似然估计
     #%timeit res_bfgs = optimize.fmin_bfgs(f, np.array([0.5]), fprime=fprime, disp=False)
```

```
print("Timing: 1000 loops, best of 3: 1.85 ms per loop")

res_bfgs_full = optimize.fmin_bfgs(f, np.array([0.5]), fprime = fprime, full_output = True)
res_bfgs = res_bfgs_full[0]
est_var_bfgs = res_bfgs_full[3]

# 计算信息矩阵；计算信息矩阵的逆矩阵，得到的方差协方差矩阵
var_bfgs = np.linalg.inv(
    -_hessian(transform(res_bfgs), nobs, Y.sum())
)[0,0]
# 根据 (解析)delta 方法得到的方差协方差矩阵
var_bfgs_delta = (transform_grad(res_bfgs)**2 * est_var_bfgs)[0,0]

print(
    "BFGS MLE estimate: %.2f (%.10f)" % (transform(res_bfgs), var_bfgs**0.5)
)
print(
    "BFGS MLE estimate (delta) : %.2f (%.10f)" % (transform(res_bfgs), var_bfgs_delta ** 0.5)
)
```

```
Timing: 1000 loops, best of 3: 1.85 ms per loop
Optimization terminated successfully.
     Current function value: 52.690796
     Iterations: 6
     Function evaluations: 7
     Gradient evaluations: 7
BFGS MLE estimate: 0.22 (0.0414246303)
BFGS MLE estimate (delta) : 0.22 (0.0171611895)
```

进而我们有误差项和误差项的偏移

```
In [11]: absolute_error = var_bfgs - var_bfgs_delta
         relative_error = (var_bfgs - var_bfgs_delta) / var_bfgs

         print(absolute_error)
         print(relative_error)
```

```
0.00142149357495
0.828376209448
```

第三种方法是牛顿拉夫森法 (Newton-Raphson method)。比起前两种方法来，牛顿拉夫森法的数值计算速度更快，但稳健性较低。python 的 statsmodels包提供了这种算法的程序。

```
In [20]: # 利用 BFMG 法作最大似然估计
         # 导入 Statsmodels 包
```

```
from statsmodels.base.model import LikelihoodModel

class Bernoulli(LikelihoodModel):
    _loglike = lambda self, theta, T, s: _loglike(theta, T, s)
    _score = lambda self, theta, T, s: _score(theta, T, s)
    _hessian = lambda self, theta, T, s: _hessian(theta, T, s)

    def __init__(self, endog, exog=None, **kwargs):
        # numpy.asarray(a, dtype=None, order=None)[source]
        # https://docs.scipy.org/doc/numpy-1.13.0/reference/generated/numpy.asarray.html
        endog = np.asarray(endog)
        super(Bernoulli, self).__init__(endog, exog, **kwargs)
        self.T = self.endog.shape[0]
        self.s = self.endog.sum()

    def loglike(self, params):
        """
        Joint log-likelihood for Bernoulli trials
        """
        return self._loglike(params[0], self.T, self.s)

    def score(self, params):
        """
        Gradient of the joint log-likelihood for Bernoulli trials
        """
        return self._score(params[0], self.T, self.s)

    def hessian(self, params):
        """
        Hessian of the joint log-likelihood for Bernoulli trials
        """
        return np.matrix(self._hessian(params[0], self.T, self.s))
```

```
In [26]: mod = Bernoulli(Y)
     #%timeit res = mod.fit(start_params=[0.5], disp=False)
     print("1000 loops, best of 3: 277 µs per loop")
     # 最大似然估计
     res = mod.fit(start_params=[0.5])
     print ("Statsmodels MLE Estimate: %.2f (%.10f)" % (res.params[0,0], res.bse[0]))
```

```
1000 loops, best of 3: 277 µs per loop
Optimization terminated successfully.
     Current function value: 0.526908
     Iterations 2
Statsmodels MLE Estimate: 0.22 (0.0414246304)
```

第十章　工作论文

10.1　问题的提出

10.1.1　研究目标

- 构建一个基于 DSGE 框架的中心——边缘经济模型，解释造成耐用消费品相对于费耐用消费品价格的通货膨胀、产出等实际经济变量波动的机制。
- 模型的参数估计，部分采用校准法，部分采用贝叶斯估计。

  结合中国当前经济运行的实际可观测数据做估计。
- 利用估计的 DSGE 模型，量化分析当经济体受到外生冲击，导致市场信贷信号发生变化时，经济政策应当如何设定和响应，来达到整体福利最大化 (加权效用函数最大化) 的目标。

  设定一系列可供选择的经济政策工具，如货币政策工具、宏观审慎政策工具等，对于实现福利最大化目标的经济绩效的比较与评价。

10.1.2　模型设定

基于 DSGE 框架的中心——边缘经济模型，主要考虑这样一个国家，拥有统一的货币、货币政策和（可以是不同的）宏观审慎政策。国家内部按地域划分为中心和边缘两个经济体，每个经济体中都有

- 两个生产部门，两种最终消费品：非耐用消费品，耐用消费品（房地产）。

  产品均在本地生产，最终产品的生产条件均为完全竞争。中间产品的生产条件为垄断竞争，存在价格粘性。

  最终产品中，非耐用消费品可以跨地区进行贸易和消费；耐用消费品只能在本地消费。
- 每个地区的家庭部门都分为两类：储蓄者，借款者。都是追求效用最大化的理性行为人，效用获得的手段主要有三种：消费非耐用品、购买耐用品 (住宅投资) 增加本家庭的房屋以及休闲。储蓄者和借款者的不同主要表现在
  - 时间偏好、进而消费习惯不同。借款者比储蓄者更缺乏耐心，因此更偏好提前消费。这使得在均衡条件下出现信贷市场：储蓄者将家庭收入的一部分存入金融中介机构，赚取存款利息。借款者利用现有住宅存量作抵押，向金融机构贷款用于住宅投资，通过增加住宅存量提高效用水平。
  - 引入Bernanke et al. (1999) 的金融加速器机制，描述家庭部门中不同类型个体受同一外部冲击的影响不同。假定存在一个针对住宅存量的特定质量冲击，该冲击会影响到借款者现有住宅存量 (抵押品) 的价值，借款者的经济决策（如期还本付息，或是违约），银

行的存贷款利差，进而整体经济运行状况。特定质量冲击对储蓄者不产生总量层面上的实质影响：储蓄者无需向银行借款，自有住房存量也不必用于抵押。

- 金融部门：每个经济体分别有一个境内金融中介机构（如商业银行），从储蓄者那里获取存款，向借款者提供贷款，当经济体内部出现资金供需不匹配情况时，通过发行债券向其他经济体中的金融机构进行融资。

  对应地，存在一个国家级跨境金融中介机构，单一功能提供跨经济体的资金流动渠道：为利用债券交易平衡两个经济体中金融中介机构之间的资金供需关系。通过这样的设定，一个经济体内部的储蓄和（住宅）投资不必每期都相等：以中心地区为例，若出现额外的信贷需求，可在跨境金融机构的协助下，通过边缘地区的资金供应予以满足，融资成本表现为跨境金融机构收取的风险溢价，风险溢价的大小取决于中心地区在边缘地区持有的资产头寸。

- 中央银行：执行货币政策，通过观测过去、预测未来的全国经济数据，如 CPI 通胀率、实际经济产出等，制定执行相关利率政策，以平稳物价、产出增长。

  - 一系列外部冲击，如前文提到的对住宅存量的质量冲击等。若无特别指明，假定部门的外部冲击表现为 AR(1) 过程；假定一个单位根形式的中性级数冲击对各部门均产生影响。

完全部模型后，进一步汇总、理。

  - 引入Calvo (1983) 的定价机制，在模型中一系列名义和实际粘性，参考Smets and Wouters (2003); Iacoviello and Neri (2010); Christiano et al. (2005)。

下面以中心地区经济体为主，分部门对模型作进一步阐述。

10.2 境内金融中介部门

境内金融机构如商业银行的资金来源是从家庭部门储蓄者中吸收存款 S_t，在下一期以存款利率 R_t 返还。资金使用方面表现为向符合资质的家庭部门借款者发放贷款 S_t^B，收取贷款利率 R_t^L。期间形成的贷存利差 (spread rate) 用 R_t^L/R_t 表示，反映金融机构中介服务品的价格，与信贷市场的实际供需关系等因素有关。

借款者获得银行贷款的“资质”是自有住宅存量的价值 $P_t^D D_t^B$，P_t 是市场上的名义住宅价格，D_t^B 表示借款者住宅存量。我们将贷款相对于抵押品价值的比率 $S_t^B/\left(P_t^D D_t^B\right)$ 定义为贷款价值比 (Loan To Value ratio, LTV)，反映银行放贷的风险程度。

借款人获得的贷款主要用于住宅投资 I_t^B，以增加 $t+1$ 期住宅存量 D_t^B，提升家庭效用与福利水平。

根据金融加速器机制，授信过程中的风险主要体现在房地产市场的违约风险，这是由贷款行为的性质所决定的。在期初，银行拟定利率 R_t^L，对有资质的借款申请者授信 S_t^B 用于住宅投资，并要求借款者在下一期返还 $R_t^L S_t^B$。在期初，特定质量冲击还未实际发生，银行基于对过去信息的不完全掌握，以及对未来状况的模糊预测，事先设定 R_t^L 值。而只有到了期末，特定质量冲击才会真正出现并被识别，从而影响借款人在期末的 (事后) 决策。

面方程等式。

我们设质量冲击是异质的，对不同借款人的冲击程度各异，$t-1$ 期末编号为 j 的借款人承受质量冲击水平 ω_t^j，影响当期住宅存量价值。如果冲击水平较低，ω_{t-1}^j 值较小，j 的当期住宅存量价值仍高于应还贷款数 $\omega_{t-1}^j P_t^D D_t^B > R_{t-1}^L S_{t-1}^B$，那么 j 选择如约还本付息；反之，如果冲击水平较高，ω_{t-1}^j 值较大，$\omega_{t-1}^j P_t^D D_t^B > R_{t-1}^L S_{t-1}^B$，$j$ 的当期住宅存量价值无力偿还贷款，会选择违约。若违约发生，银行会引入债务清偿机构，强制收回 j 的抵押品残值 $\omega_{t-1}^j P_t^D D_t^B$，转手在住宅市场

上出售。债务清偿机构收取 $0<\mu<1$ 比例的佣金，银行实际回收 $(1-\mu)\,\omega_{t-1}^{j}P_t^D D_t^B$。假定债务清偿机构由储蓄者所有，其利润最终流转进入家庭部门的储蓄者手中[1]。

本文此处设定与Bernanke et al. (1999) 的原始假设不同。在原始假设中，为了评估因违约而收回住宅的实际价值 $\omega_{t-1}P_t^D D_t^B$，银行需要支付 μ 比例的监督成本，从而若违约发生，房屋存量的一部分价值被摧毁。本文对这一假设的调整，主要基于两点考虑。第一，从模型设定上来讲，Bernanke et al. (1999) 假设会使随机质量冲击对家庭部门的住宅投资、住宅存量、进而整体经济运行状态产生很大幅度的波动，这与观测到的现实情况不符 (Forlati and Lambertini, 2011)。第二，本文研究的核心目标之一，是探讨一系列宏观审慎政策工具对维护宏观经济运行稳定性的效果评价，主要基于全体福利效果的分析；若是在前期模型构建中出现社会财富的“蒸发”，在随后经济政策的福利效果评价中，我们将很难区分福利损耗中有多少是由不恰当的宏观审慎政策工具选取所导致的，有多少是由外生风险冲击所导致的。

现在将风险冲击 ω_t^j 对储蓄者 j 个体的影响，扩展到全部储蓄者。假设 ω_t^j 呈对数正态分布[2]

$$\log\left(\omega_t^j\right)\sim\mathcal{N}\left(-\frac{\sigma_{\omega,t}^2}{2},\sigma_{\omega,t}^2\right),\quad \omega_t^j\in(0,\infty)\ \forall\, t,j. \tag{10.1}$$

其中 t 期质量冲击在 j 个储蓄者中的分布，用标准差 $\sigma_{\omega,t}^2$ 表示。对于标准化的冲击值 ω_t^j，通过选取合适的 $\sigma_{\omega,t}^2$，我们使得 $E\left[\omega_t^j\right]\equiv 1\leftrightarrow E\left[\log\omega_t^j\right]\equiv 0$，即从同期来看尽管住宅市场中存在异质性特定质量冲击，但总量层面上来看，质量冲击的风险仍为零。

此外跨期来看，我们设标准差 $\sigma_{\omega,t}^2$ 是个时变对数 $AR(1)$ 过程

$$\log\left(\sigma_{\omega,t}\right)=\left(1-\rho_{\sigma_\omega}\right)\log\left(\overline{\sigma}_\omega\right)+\rho_{\sigma_\omega}\log\left(\sigma_{\omega,t-1}\right)+u_{\omega,t},\quad u_{\omega,t}\sim\mathcal{N}\left(0,\sigma_{\omega,t}\right). \tag{10.2}$$

这样的设定，使得 $\sigma_{\omega,t}$ 的增加，在保持 $E\left[\omega_t^j\right]$ 不变的情况下，提高了 ω_t^j 的分布的偏度：原有分布更多向左侧几种，从而出现更多 ω_t^j 取较低值的情况。

从事后观察来看，存在一个门槛值 $\overline{\omega}_t^p$，使得个体 j 受到的冲击若是低于门槛值 $\omega_t^j<\overline{\omega}_t^p$ 只能违约；若是高于门槛值 $\omega_t^j\geq\overline{\omega}_t^p$ 选择如约还本付息。$\overline{\omega}_t^p$ 显然是一个影响经济系统的重要指标。然而 t 期初，银行在制定贷款利率 R_t^L 时，并不清楚 $\overline{\omega}_t^p$ 这一关键信息，换句话说，不知道会有多少住宅贷款违约，多少如期还款；假定银行基于对 $t+1$ 期房价的预期 $E_tP_{t+1}^D$，只能设一个事先门槛值 $\overline{\omega}_t^a$ 来作近似，依据的计算标准为

$$\begin{aligned}\overline{\omega}_t^a E_t\left[P_{t+1}^D D_{t+1}^D\right]&=R_t^L S_t^B,\\ \hookrightarrow\overline{\omega}_t^a&=R_t^L\frac{S_t^B}{E_t\left[P_{t+1}^D D_{t+1}^B\right]}=R_t^L\left(LTV_t\right),\end{aligned} \tag{10.3}$$

不难看出两点。第一，事先和事后门槛值不一定相等，这主要取决于 t 期实际发生的质量冲击在多大程度上偏离了最初的预期，用 $\sigma_{\omega,t}$ 表示。第二，$\overline{\omega}_t^a$ 是个关于贷款价值比 LTV 的增函数，也即同等价值的抵押品，获得的贷款越多，风险越大，未来出现违约问题的可能性越高。

基于设定好的事先门槛值 $\overline{\omega}_t^a$，可以用特定质量冲击的累积密度方程 (Cummulative Density Function, CDF) $F\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)$ 来描述 t 期初银行在做贷款利率的决策时，预计会有多少比例的借款人会选择在 t 期末 ($t+1$ 期初) 违约

$$F\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)=\int_0^{\overline{\omega}_t^a}dF\left(\omega;\sigma_{\omega,t}\right)d\omega, \tag{10.4}$$

[1]储蓄者仍保有其余并未拿去做抵押的房屋存量价值 (Suh, 2012)。

[2]膏按：关于对数正态分布，一个 (不太恰当的类比) 介绍，可见 (朱彦元, 2018, Sec 10.5.4)，即本笔记第6.5.4节。

以及多少比例的借款人会如约还本付息

$$1-F\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)=\int_{\overline{\omega}_t^a}^{\infty}dF\left(\omega;\sigma_{\omega,t}\right)d\omega. \tag{10.5}$$

在此基础上，所有低于门槛值 (即存在违约风险) 的质量冲击的平均值用 $G\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)$ 来表示

$$G\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)\equiv\int_0^{\overline{\omega}_t^a}\omega\,dF\left(\omega,\sigma_{\omega,t}\right), \tag{10.6}$$

那么银行在 t 期初预计 $t+1$ 期的违约住房贷款数额为

$$G\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)E_t\left[P_{t+1}^D D_{t+1}^B\right]=\int_0^{\overline{\omega}_t^a}\omega dF\left(\omega,\sigma_{\omega,t}\right)\left[P_{t+1}^D D_{t+1}^B\right]. \tag{10.7}$$

在本模型中，银行对贷款利率的设定不只取决于自身的成本——收益考量，还受到宏观审慎经济政策的影响。具体来说，假定存在某种政策工具 η，通过影响银行的资产负债表，即能贷出多少款，来影响信贷市场出清条件

$$n\lambda\frac{S_t-B_t}{\eta_t}=n\left(1-\lambda\right)S_t^B, \tag{10.8}$$

等式左侧和右侧分别表示总负债和总资产。系数 $0\le n\le 1$ 表示中心地区人口数占全国总人口的比重；$1-n$ 为边缘地区人口数站全国总人口的比重。$0\le\lambda\le 1$ 表示中心地区全部人口中，储蓄者所占比重；$1-\lambda$ 表示借款者的比重。变量 B_t 表示中心地区在边缘地区的债权。出于简化模型的考虑，我们假定宏观审慎政策工具 η_t 被置于其他现有监管规定之上，从而使其与境内净存款、贷款等呈线性关系。具体来说，η_t 可以表示为资本附加 (capital surchage)、贷款损失准备 (loss-loan provision)、存款准备金 (reserve requirement) 等具体形式，通过直接限制可贷出资金量来影响存贷款利差，对应 $\eta>1$。此外也可设 $\eta<1$，作为一种“非传统”货币工具，被用于中央银行实现控制存贷利差等目标 (Gertler and Karadi, 2011)。在本研究中，一方面在模型估计过程中，设 $\eta_t=\eta\equiv 1\,\forall t$，即假定宏观审慎政策是“中性”的。另一方面在随后宏观审慎经济政策的福利效果分析中，设 η 围绕稳态值 1 作反周期变化。

文作详细说明，一个 ref。

假定商业银行面对完全竞争的市场条件。此时银行授信的期望收益率应该与其融资成本相互抵消，后者在本模型中等价于储蓄利率 R_t(我们假定商业银行从其他金融机构的融资成本也等于 R_t)，由此有银行的参与约束条件

$$n\lambda R_t\left(S_t-B_t\right)=n\left(1-\lambda\right)E_t\left\{\left(1-\mu\right)G\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)P_{t+1}^D D_{t+1}^B+\left[1-F\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)\right]R_t^L S_t^B\right\},$$

该条件确保了金融机构对债权人的义务（式左侧）与对债务人的期望回报（式右侧）相等。期望回报由两部分构成，违约资产处置后回收的部分，和如约还本付息的部分。将(10.8)代入上式左侧有

$$n\lambda R_t\left(S_t-B_t\right)=n\left(1-\lambda\right)\eta_t R_t S_t^B,$$

进而参与约束条件进一步改写为

$$\eta_t R_t=E_t\left\{\left(1-\mu\right)\frac{G\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)}{LTV_t}+\left[1-F\left(\overline{\omega}_t^a,\sigma_{\omega,t}\right)\right]R_t^L\right\}. \tag{10.9}$$

第四部分

附录

Appendices

第一章　Post-Keynesian Economics 的一个小综述

主要参考自 Hart and Kriesler (2015)。

A.1　Introduction

A.2　PK 的主要方法论

A.3　PK 下的市场结构与定价

A.3.1　PK 下的市场结构

A.3.2　PK 下的定价策略

markup pricing principle 的一系列变体，如

- Kalecki (1937) 考虑到企业往往面临内部/外部的融资需求，(其他企业所定的)price markup 会影响到企业的现金流，进而使得金融部门，企业融资行为和企业 (家庭) 投资决策这三者之间存在重大关联，可见货币和金融的确会对实体经济产生影响 (Ball, 1964; Eichner, 1973; Wood, 1975; Harcourt and Kenyon, 1976)。
- (企业自身所定的)price markup，也会被理解为企业自己在多变、激烈和难以预测的市场竞争中寻求立足、发展的关键决策之一。

因此价格更多反映企业的利益决策而非其所在产业或市场的条件；价格是基于企业战略而非基于成本制定的；这一价格反映市场的非均衡状态，对于整个经济体动态发展起到重要作用。

PK 视野下，企业不是追求短期利润最大化的行为人，而是在面临巨大不确定性条件下的长期决策制定者。企业的产量和定价决策反映了寡头垄断的市场条件，企业自身及其竞争对手的行为，以及融资情况都是该条件的重要组成部分。

A.4　PK 下的宏观经济

MS 更多遵循萨伊定律，即供应会产生它所对应的需求，导致就业和产出都不受需求的限制。PK 经济体更多强调有效需求，货币非中性，不确定情况下的决策，以及主观并且变化迅速的期望，这使得 PK 对萨伊定律提出挑战。值得指出的是，该挑战并不依赖于对市场完全竞争程度或者“粘性”的假定。由此 PK 坚决反对任何将凯恩斯经济学与一般均衡框架融合在一起的尝试，即反对新古典主义综合。PK 也不同于 New Keynesian，后者致力于将一系列摩擦 (frictions) 引入到 New

Classical 体系当中Akerlof (2007)。在 PK 看来，这两种思路忽略了发达资本主义经济体中最本质的不稳定因素，从而仍然使市场运作暴露在重大风险中。

延续 Keynes 和 Kalecki 的思路，PK 强调真实产出和就业的根本决定因素是有效需求的水平，真实产出和就业的波动主要是由投资支出的变化所导致的，投资支出受到对周边情形的期望（“动物精神”）的影响。对于“储蓄增加导致投资增加”的传统观点，PK 持相反意见：只要能够寻找到合适的融资渠道，投资通过乘数效应可以增加收入，进而提高储蓄。

PK 在几个方面修正了 Keynes 的理论体系，尤其是在对金融市场、制度的描述上。在 Keynes 看来，不确定性是人们持有货币以追求保值的最关键原因，持有货币的行为将人们过去已经发生了的经济行动和不确定性的未来连接在一起，进一步扩大了不确定性的风险。如Kaldor (1985) 对货币主义者的批判，货币供应的内生性被过分夸大了：随着货币供应增加，金融机构可以放出更多贷款，导致存款增加和/或对金融财产的购买增加。这一系列借入和借出的决策取决于融资成本以及对未来的期望——金融财产的规模和组成结构直接取决于 balance sheet 中放贷者和潜在借款人的立场。他们的乐观或悲观心态，会影响到金融市场整体的行为，进而放大、波及对实体经济运转情况的看法。从这个意义上来讲，金融不稳定性假设认为资本主义经济体中，金融部门和实体经济部门的不稳定性是彼此相关和不可避免的 (Minsky, 2015)，导致中央银行和财政稳定政策的努力注定无效，“有效市场”及其相关资产定价模型也需要被否定。

PK 不再关注于供应-需求互动决定的均衡价格和产品数量，而是关注产业结构和 markup 定价原则，这使得对产品相对价格的分析不再有意义：在宏观经济层面上，产出就业和总物价均发生波动。excess capacity 的存在意味着即便需求增加了，产出也会跟着调整以满足需求，从而不会出现通货膨胀的压力。当逼近 full capacity utilisation 的情况下，通货膨胀的压力也主要来自 mark-up pricing 机制，而不是原材料、工资等要素成本。就业水平主要取决于有效需求水平，而不是实际工资。实际工资不取决于市场，而是取决于以下两方面因素，第一如上文所说的企业定价策略，第二是围绕全部国民收入，在劳动收入和资本收入之间的讨价还价，这就涉及到随时间变化的制度性要素。在 PK 看来，通货膨胀是国民收入中劳动收入和资本收入比重分配不恰当所导致的，也只有重新调整比重分配才能缓解通货膨胀——这呼吁一种 permanent income policy 的出台。

PK 对经济增长和经济发展尤为关注，在 PK 看来它们是资本主义经济发展动态的最核心要素。然而在 PK 看来，它们都不是平稳和连续的。

A.5　PK 视野下的经济政策

与主流经济学衍生出的经济政策相比，根据 PK 所得的经济增侧往往有些反直觉。总的说来，PK 认为市场运作在短期和长期都存在失灵和低效率情况，持干预市场运作的立场。

第二章　Keynesian, New Keynesian and New Classical Economics

主要参考自Greenwald and Stiglitz (1987)。

简要说来。New Classical Economics 学派认为，macro-theory 缺乏 micro 基础，需要从一系列微观行为准则（理性、最大化的企业和个人行为）推导出经济总体层面上的动态；承认动态对于了解宏观行为的重要性；在分析动态行为时，期望扮演核心角色，如理性预期的形成。“理性预期学派”的问题并不在于其假定之一，理性的预期是不是可能，而在于另一个古典主义假定，市场出清。根据这样的假定，理性预期模型会得到如下结果：没有失业，政府宏观政策是不必要的。

另一大流派 New Keynesian Economics 学派，认为现实世界中真实存在的一系列经济现象，如失业、信贷配给、经济周期等，都是不能由标准 micro-theory 所解决的。因此致力于开发 micro 理论，来解释 macro 现象。NK 学派的核心工作就是从 micro 和 macro 层面了解不充分信息和不完全市场及其后果。市场失灵中的表现之一是失业。与 New Classical 相比，NK 同样追求建立一个 single 解释框架，但 NK 的框架致力于解释失业，而不像 New Classical 学派那样忽视甚至否定失业的存在。

B.1　Keynesian 的四个核心

NK 从 traditional Keynesian 经济学发展而来。Keynes 对经济运行有着显著不同于 standard neo-classical 理论的看法，比如企业行为决策并不是基于理性的计算，而是”动物精神”。他用图形的形式表示他的观点，但当他试图构建理论模型去描述这一观点的时候，他更多回到了传统的 neo classical 框架中——这很有可能是因为他受过严格的 neo-classical 训练，束缚了他表达自己全新思想的能力。Hicks 等后人将 Keynes 的观点进一步模型化系统化理论化，成为理解失业、经济周期等问题的强有力工具。

Keynes 的观点中，如下四点对于我们构建一个解释失业和经济周期波动的模型来说至关重要。

B.1.1　失业与有效工资理论

第一点。该通用模型必须能够解释失业的持续存在，以及一系列关键经济变量的周期性变化。为了解释失业，就需要发展出一个关于劳动力市场的理论，Keynes 从传统经济学思想 (如 fixed price 学派) 中借鉴了 (名义，实际) 工资粘性的假定，但 Keynesian 从经验和理论两个层面批评了 Fixed price 学派的工资粘性假定是不完善的。经验层面，大萧条中工资下跌了 1/3，经历通货膨胀的国家即使调整了实际工资，仍然出现失业情况。理论层面，Fixed Price 学派没能对工资粘性做出解释，没能提供制度上的设计。

根据 Keynes 的理论，工资并不需要是完全粘性的，而仅仅需要确保工资不会跌落到完全市场出清的水平，如有效工资理论 (efficiency wage theories(Stiglitz, 1984, 1987))。有效工资理论的假

定如下：

- 对异质劳动者属性特征的信息掌握是不充分的，
- 劳动者的工作表现是无法充分监控的，
- 无法与劳动者签订可以完全反映他劳动成果的工作合同。

根据这样的假定，有效工资理论认为：

- 劳动力的质量及生产率，企业的利润都可能随着工资的增加而增加，
- 工资增加可能导致劳动者周转率降低，
- 由于企业需要承担部分的劳动者周转成本，工资上升在某种程度上可以导致企业利润的上升。

然而当失业存在，工资仍可能不会降低，因为企业认识到如果他们降低员工工资，会导致生产率降低，周转率上升，利润降低。这是指充分竞争市场下，企业只在一定范围内是工资制定者的情况。如果瓦尔拉斯工资 (Walrasian wage)——劳动力需求等于劳动力供应情况下对应的工资水平——相当低，企业有可能有动机提高自己员工的工资以提高利润，对应的有效工资 (efficiency wage) 就是企业利润最大化条件下的工资水平。显然外部经济情况的不同会导致有效工资的不同，使得工资并不是完全刚性的；实际工资并不会跌落到市场出清那么低的水平上。这就肯定了政策干预的重要性，如失业救济等会影响到市场上的均衡工资水平。

B.1.2 价格变化与经济波动

第二点。该通用模型不仅能够解释失业的存在，还要能够解释失业等核心变量的波动。分为两个问题。

1. 导致经济体出现大幅度波动的冲击，是怎么产生的?
   - 战争、油价波动等外生原因。
   - 内部机制。
     - 对投资、尤其是存货的需求变动。然而在理论上，对于 concave 生产函数、低工资、低利率的情况，应当存在生产平滑效果：存货应当降低而不是扩大经济波动。
     - 储蓄的变化。储蓄有助于稳定消费。在储蓄不足的情况下，implicit contracts 提供的保障机制起到稳定收入、进而稳定消费的作用。因此 implicit contracts 应当是降低而不是扩大经济波动。

   Keynes 正确地强调了投资变化对理解经济波动的重要性。但他用动物精神 (unexplained changes in expectations) 来解释投资的变化。这种解释不能让人完全满意：会导致经济周期理论具有无法解释甚至是非理性的特征。

2. 根据 typical micro 模型，利率、工资等价格的变化有助于缓解需求或供应的波动，进而如存货一样起到稳定经济波动的效果。因此大的外部冲击 (影响需求或供应) 只可能导致经济体中均衡价格的小幅度变化。然而为什么工资、利率等价格往往是刚性的?

   Keynes 不仅需要解释为什么对投资的需求曲线 shifted，还需要解释为什么利率调整无法有效缓解投资需求变化所产生的波动。Keynes 的回答不能让人完全满意：他认为对投资需求的降低会导致利率的降低，然而真实世界中当出现经济衰退，利率有时上升；甚至有时利率的变化不会影响到公司的经济决策。

Keynes 只是指出工资可以保持不变，但未能解释工资为什么会保持不变。有效工资理论解释了为什么工资不会跌落到市场出清条件那么低的水平。Akerlof and Yellen (1985) 的 near-rational 行为模型解释了为什么企业可以调整工资而往往倾向于不这么做。类似地，一个资本市场理论解释了为什么利率不会降低到那么低的水平 (Stiglitz and Weiss, 1981, 1983, 2009)，甚至常常是变动很小的。同时，有效工资理论也强调了一个企业制定的工资往往和其他企业制定的工资之间存在互动关系。

总之，上述理论对工资，利率，和价格刚性做出了解释，它们互相作用，放大了对经济系统的冲击，导致在波动往往很剧烈而不是很缓和 (multiplier effects)。

B.1.3 储蓄与投资，信贷配给

Keynes 敏锐注意到了家庭部门的储蓄和生产部门的投资是不同的，要加以区分。然而在构建模型时，Keynes 又回到 standard neo-classical 模型的老路上，储蓄如何转化为投资，不做进一步分析；投资由利率决定，因此不存在信贷配给 (credit rationing)。这使得尽管 Keynes 强调资本市场的不充分性特征，却也无法做深入的探讨。

B.1.4 供应与技术进步

在 Marshallian 的框架中，经济的波动是由供应波动和需求波动两方面共同作用的结果。Keynes 基于对大萧条的分析，将精力投入到分析导致需求波动的诸因素。然而对于供应端的波动，以及技术进步，他并没有做深入解释。

B.2 New Keynesian 经济学的四个核心

NK 是在 Keynes 思想的基础上发展起来的，但也注意到了 Keynes 过度依赖于 neoclassical 理论分析框架，以及虽然认识到了资本市场是不完善的，但没能充分认识到不完善所带来的信息成本损耗及其经济效果。NK 的核心有以下几点。

B.2.1 有效工资

有效工资模型，用于解释非刚性工资为什么不会完全跌落到市场出清条件下的低水平。

B.2.2 资产配给

资本市场是不充分竞争的，表现在企业经理人与潜在投资者之间的信息不对称，信息不对称导致资产配给问题。资产配给问题的重要性在于，如果企业想要获得更多资本用于扩大生产，就需要从市场上借入更多资金，承担更大的投资风险。风险还表现在，在缺乏期货市场的情况下，企业当期筹资用于扩大生产，在未来时间里用增加的产出换取利润抵付，这就存在风险。在分析企业经营决策时，就必须考虑企业的风险承受意愿。

B.2.3 信贷配给

有意愿承担风险扩大生产的企业，还面临借不到钱的问题，即信贷配给问题。与 neoclassical 的分析框架不同，资本品提供者在面临供不应求的情况下，可能会选择不提高利率，其逻辑类似于失业 (劳动力供大于求) 情况下企业的工资定价：提高利率可能会使资本品所有者的期望回报率降低，其原因之一是选择效应 (selection effects，申请借贷人的集合向不利于借出者的方向变化)，其

原因之二是激励效应 (incentive effects，借入者为了偿还更高的利率，会倾向于采取更激进、风险更高的生产形式)。

B.2.4 货币政策

NK 中货币政策的重要性，比起表现在个人持有货币维持账目平衡来，更多体现在信贷的可获得性上。信息不对称导致若商业银行决定缩紧发放贷款，需要资金的生产者很难找到其他合适的放贷人。此外商业银行的放贷决策也类似于企业生产者的生产决策，充满不确定性和风险。中央银行可以通过货币政策影响商业银行的放贷意愿。

B.2.5 小结

总之，NK 提供一套解释经济运行的通用理论，该理论基于微观经济学原理而建立。与 traditional Keynesian 理论相比，NK

1. 弥补了传统 K 理论的不一致性，如
   (a) 内部不一致，如期望的形成机制，
   (b) 预测和实际观察情况之间的不一致，
2. 填补了传统 K 理论的空白，如传统 K 仅仅假定工资是刚性的，NK 则至少在一定程度上解释了工资为什么是刚性的。

NK 进一步探讨了

1. 失业问题，如有效工资理论
2. 经济周期波动问题，其基本逻辑是：对经济系统的冲击影响到企业实际利用资本的存量。
   (a) 信贷市场问题，企业面对不充分的信贷市场，信息不对称，融资有风险；即使信贷市场是完全的 (企业能以合理的利率得到自己所需的全部融资)，企业融资数额仍然受自己风险承受能力的限制。
   (b) 根据借款契约的束缚，企业借贷从事生产活动需要承诺的义务是定死的，随着 working capital 的降低，借贷风险 (破产的概率) 升高，生产活动的规模降低，并且需要过了一段时间才能使 working capital 缓慢回升到正常水平。

   这使得 NK 不仅可以解释总量层面上的冲击 (如货币冲击导致的物价暴跌) 如何影响总量层面上的经济系统，还可以解释部门层面上的冲击 (如未提前预知的 demand shift，或操作价格的石油联盟等) 如何影响总量层面上的经济系统。从而生产意愿成为关于 working capital 的 concave 函数，并且 working capital 的再分配也会产生总量效果。

B.3 凯恩斯的不足

NK 能够解释分析一些现象，这些现象都是 traditional Keynesian 理论所很难解释的，如

1. 为什么经济衰退中的企业不降低产品价格，即 markup 的周期运动模式，
2. 投资和存货的周期性变化特征，
3. 失业者在愿意降低工资重新求职的情况下，为什么仍然很难再就业，

4. 为什么一个 unanticipated 工资-物价下降，通过降低企业的 working capital 存量，会导致经济衰退恶化而非好转，

下面做详细说明。

Keynes 最重要的错误在于他的企业理论，以及对货币如何影响经济活动的解释上。这两点不足都是由于他未能充分理解资本市场的本质。

B.3.1 债券和股票的区别

Keynes 未能正确区分长期债券和长期股票，而将二者统称为长期资产。然而区别的确存在：

1. 风险性。即便不考虑破产清算问题，二者在风险性上有本质区别：在经济衰退中，债券的价值上升，股票的价值下降。因此在个人理财账户中，二者是互补品而非替代品。

2. 企业承诺。借入债券/借贷的企业，需要承诺在指定日期偿还；而对于股票，企业则不需要受此承诺约束。从这个角度来说，对于企业和投资者而言，二者远非完全替代品。尤其在经济衰退中，需要融资的企业很少借助股票渠道，见Greenwald et al. (1984) 的 adverse selection model。

B.3.2 需求和供应

Keynes 从需求端对经济波动的解释，有助于回答“为什么企业不改变定价策略，降价促销”，但该解释进一步导致下个问题的出现：为什么一个小的开放经济体会存在 Keynesian 失业的情况？只需要它改变汇率，就可以产生对该经济体商品近乎无限的需求。

NK 不对供应和需求作明确的区分。在有明确需求的情况下，企业倾向于扩大生产 (供应)，此时需求引导生产；在需求不够明确的情况下，企业未必会扩大生产，而会基于供应曲线从事经济活动。这样，NK 可以解释为什么在经济周期中，企业愿意生产的产品数量是波动的。此外 NK 还可以结识为什么在经济衰退时期，(可以调整自己产品价格的) 企业倾向于基于生产成本，做一个更高的 price markup。在不充分竞争和不完全信息的市场上，企业在制定产品价格时，面临现在和未来利润的权衡：今天降价导致今天利润下降，以及明天销量上升、利润上升。衰退时期企业产品定价更高的原因在于，生产企业面对更高的资本成本 (并不是指市场利率)，以及更严重的资产配给问题。

B.3.3 投资的决定因素

Keynes 认为在给定期望水平的情况下，投资的最主要决定因素是利率。不论这里的“利率”应当是指实际利率还是名义利率[1]，通过对现实世界的观察能够发现，实际利率相对于其他经济变量，波动相对较小。NK 模型中在解释投资时，作为解释变量的 (实际) 利率不应当是个常数或近似常数。

此外 NK 认为，投资还受到特定时段信贷可获得程度的影响。所谓特定时段，是指货币政策会对经济活动产生影响的时期。在现实世界中，当面临经济衰退的时候，就算银行仍然愿意以现行利率将款项借出给有良好资质的企业，但很可能愿意借入的企业是不足的，这导致货币政策失效。

Keynesian-neoclassical 理论无法解释存货波动的问题，即为什么存货往往激化而不是缓和经济波动。NK 可以解释这个问题，认为资本的有效成本增加意味着企业倾向于在经济衰退时期减少库存。资本有效成本的增加是由于 equity rationing 和/或 working capital 供应量的减少所导致的。

[1] 目前经济学界通常认为，实际利率更为恰当。

B.3.4 货币政策

在 Keynes 的分析框架中，是不可能通过货币政策影响经济运行的。从最简单的思路出发，可行的方案分三步走：

1. 政府采取改变货币供应量的举措；

   NK：至少对完成交易这一目的而言，存在着货币的替代物，可能使得政府改变货币供应的举动无法完成既定目标。相当多的交易可以不依赖于货币完成，仅靠信用足矣，这使得许多基于 cash-in-advance 的模型无法较好解释现实。此外交易和收入之间的关系较弱：大多数交易体现为财产的交换，而经济周期波动往往伴随着财富的变化，进而财产分配的变化。

2. 在给定个人货币需求函数 (假定个人对货币的需求取决于利率和收入) 的前提下，利率发生变化；

   NK: 既然对货币的需求往往是基于财产的，那么显然财富而非收入成为个人经济行为的关注要素。在交易目的考虑之外，短期债券可以充当货币的完全替代品，此时持有货币的机会成本表现为短期货币利率。但影响投资行为的利率必须是实际利率[2]。

3. 利率变化导致投资变化。

B.4 小结

方法论。

一系列 facts: 市场是很复杂的。经济模型应当致力于描述市场的核心特征，而非完美复制市场中的一切。企业和个人今天的行为决策受到昨天决策的影响，基于对未来的期望。人们对未来的期望是非理性的。市场的确存在但是不完全的。价格的确调整。失业率较高时，工资会下降。

这些 facts 给经济模型构建者带来挑战。无法建立这样一个动态模型来完美反映未来的一切 facts，因为未来是不可完全预知的。构建模型应当致力于解决其中一些核心问题 (而非全部)，比如失业。

政策。长久以来关于在出现失业时政府应当做些什么，存在较大争议：a) 自由放任，什么都不做，b) 致力于降低工资以提升就业，c) 使用货币政策，d) 扩大政府支出。Keynes 理论的成功在于它为持有 d) 立场的人提供了理论依据。New Classical 理论为持有 a) 立场的人提供依据

市场效率。NK 认同 K 的观点，不充分竞争、不完全信息导致市场失灵，从而导致失业的确是资本主义经济面临的核心难题之一。

[2]此外，Cash Management Accounts 方面的最新研究表明，完全可以通过提供 interest bearing "money" 来替代直接持有货币的方案，此时经济个体的最优决策就变成了考虑他所拟持有债务的 maturity structure。

参考文献

Adda J, Cooper R. Dynamic Economics: Quantitative Methods and Applications. London: MIT Press, 2003.

Adolfson M, Las en S, Lind J, Villani M. Evaluating an estimated new Keynesian small open economy model. Journal of Economic Dynamics and Control 2008a;32(8):2690–2721.

Adolfson M, Laséen S, Lindé J, Villani M. Bayesian estimation of an open economy DSGE model with incomplete pass-through. Journal of International Economics 2007;72(2):481–511.

Adolfson M, Laséen S, Lindé J, Villani M. Evaluating an estimated new Keynesian small open economy model. Journal of Economic Dynamics and Control 2008b;32(8):2690–2721.

Agranovich MS. Sobolev Spaces, Their Generalizations and Elliptic Problems in Smooth and Lipschitz Domains. Springer Monographs in Mathematics. Cham: Springer International Publishing, 2015.

Akerlof GA. The Missing Motivation in Macroeconomics. The American Economic Review 2007;97(1):3–36.

Akerlof GA, Yellen JL. A Near-Rational Model of the Business Cycle, With Wage and Price Inertia. The Quarterly Journal of Economics 1985;100:823.

Aldrich EM, Kung H. Computational Methods for Production-Based Asset Pricing Models with Recursive Utility. Technical Report 87; 2011.

Alvarez F, Jermann UJ. Using Asset Prices to Measure the Cost of Business Cycles. Journal of Political Economy 2004;112(6):1223–1256.

An S, Schorfheide F. Bayesian analysis of DSGE models - Rejoinder. Econometric Reviews 2007;26(2-4):211–219.

Ancona A. Elliptic operators, conormal derivatives and positive parts of functions (with an appendix by Haïm Brezis). Journal of Functional Analysis 2009;257(7):2124–2158.

Anderson EW, Hansen LP, McGrattan ER, Sargent TJ. On the Mechanics of Forming and Estimating Dynamic Linear Economies. In: Amman HM, Rust J, editors. Handbook of Computational Economics. Amsterdam: North-Holland; 1996. p. 171–252.

Anderson GS. A Reliable and Computationally Efficient Algorithm for Imposing the Saddle Point Property in Dynamic Models. IFAC Proceedings Volumes 1998;31(16):355–366.

Anderson GS, Moore G. A linear algebraic procedure for solving linear perfect foresight models. Economics Letters 1985;17(3):247–252.

Andreasen MM, Fernández-Villaverde J, Rubio-Ramírez JF. The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications. Technical Report 2016-07; 2016.

Ang WT. Hypersingular Integral Equations in Fracture Analysis . Cambridge: Woodhead Publishing, 2013.

Angrist JD, Pischke JS. Mostly Harmless Econometrics. An Empiricist's Companion. Princeton University Press, 2008.

Aruoba SB, Fernández-Villaverde J, Rubio-Ramírez JF. Comparing Solution Methods for Dynamic Equilibrium economies. Journal of Economic Dynamics and Control 2006;30(12):2477–2508.

Atkinson KE. The Numerical Solution of Boundary Integral Equations. University of Iowa, 1996.

Atkinson KE. The Numerical Solution of Integral Equations of the Second Kind. Cambridge University Bridge, 1997.

Babuška I. Error-Bounds for Finite Element Method. Numerische Mathematik 1971;16(4):322–333.

Babuška I. The Finite Element Method with Lagrangian Multipliers. Numerische Mathematik 1973;20(3):179–192.

Babuška I, Suri M. The P and H-P Versions of the Finite Element Method, Basic Principles and Properties. SIAM Review 1994;36(4):578–632.

Balke NS, Canova F, Milani F, Wynne MA. DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments. Emerald Group Publishing Limited, 2012.

Ball RJ. Inflation and the Theory of Money 1964;.

Bapat RB. Linear Algebra and Linear Models. Springer Science & Business Media, 2012.

Bastani H, Guerrieri L. Advances in Automatic Differentiation. In: Bischof CH, Bücker HM, Hovland P, Naumann U, Utke J, editors. Advances in Automatic Differentiation. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 303–313.

Bayoumi T, Faruqee H, Laxton D, Karam PD, Rebucci A, Lee J, Hunt B, Tchakarov I. GEM: A New International Macroeconomic Model. International Monetary Fund, 2004.

Behrndt J, ter Elst AFM. Dirichlet-to-Neumann Maps on Bounded Lipschitz Domains. Journal of Differential Equations 2015;259(11):5903–5926.

Bellman R. Dynamic Programming. Princeton University Press, 1957.

Bender CM, Orszag SA. Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory. New York: Springer, 2013.

Benhabib J, Farmer REA. Indeterminacy and Sunspots in Macroeconomics. In: Handbook of Macroeconomics. Elsevier; 1999. p. 387–448.

Benigno P, Woodford M. Optimal Monetary and Fiscal Policy: A Linear-Quadratic Approach. NBER Macroeconomics Annual 2003;18:271–333.

Bernanke BS. **Inflation Expectations and Inflation Forecasting**. 2007.

Bernanke BS, Gertler M, Gilchrist S. The financial accelerator in a quantitative business cycle framework. In: Handbook of Macroeconomics. Elsevier; 1999. p. 1341–1393.

Binder M, Pesaran MH. Multivariate Rational Expectations Models and Macroeconomic Modelling: A Review and Some New Results. In: Pesaran MH, Wickens MR, editors. Handbook of Applied Econometrics: Macroeconomics. Oxford: Faculty of Economics, University of Cambridge; 1995. p. 139–187.

Blanchard OJ, Kahn CM. The Solution of Linear Difference Models under Rational Expectations. Econometrica 1980;48(5):1305–1311.

Bloom N. The Impact of Uncertainty Shocks. Econometrica 2009;77:623–685.

Bocola L. The Pass-Through of Sovereign Risk. Journal of Political Economy 2016;124(4):879–926.

Bogachev VI. Measure Theory Volume I. Springer, 2007a.

Bogachev VI. Measure Theory Volume II. Springer, 2007b.

Boyd JP. Chebyshev and Fourier Spectral Methods. 2nd ed. Dover Publications, 2001.

Boyd JP, Petschek R. The Relationships Between Chebyshev, Legendre and Jacobi Polynomials: The Generic Superiority of Chebyshev Polynomials and Three Important Exceptions. Journal of Scientific Computing 2013;59(1):1–27.

Boyd S, Vandenberghe L. Convex Optimization. Cambridge University Press, 2004.

Bramble JH. The Lagrange Multiplier Method for Dirichlet's Problem. Mathematics of Computation 1981;37(155):1–11.

Brenner SC, Scott RL. The Mathematical Theory of Finite Element Methods. volume 15 of *Texts in Applied Mathematics*. New York, NY: Springer, 2008.

Brubakk L, Husebø TA, Maih J, Olsen K, Østnor M. Finding NEMO: Documentation of the Norwegian economy model. Technical Report; 2006.

Buiter WH. The unfortunate uselessness of most'state of the art' academic monetary economics. MPRA paper 2009;.

Caldara D, Fernández-Villaverde J, Rubio-Ramírez JF, Yao W. Computing DSGE Models with Recursive Preferences and Stochastic Volatility. Review of Economic Dynamics 2012;15(2):188–206.

Callahan JJ. Advanced Calculus: A Geometric View. Undergraduate Texts in Mathematics. New York, NY: Springer New York, 2010.

Calvo GA. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 1983;12(3):383–398.

Canova F. Bridging cyclical DSGE models and the raw data. Manuscript 2009;.

Canova F. Methods for Applied Macroeconomic Research. Princeton University Press, 2011.

Cheng AHD, Cheng DT. Heritage and early history of the boundary element method. Engineering Analysis with Boundary Elements 2005;29(3):268–302.

Cho JO, Cooley TF, Kim HSE. Business cycle uncertainty and economic welfare. Review of Economic Dynamics 2015;18(2):185–200.

Christiano LJ. Solving Dynamic Equilibrium Models by a Method of Undetermined Coefficients. Computational Economics 2002;20(1):21–55.

Christiano LJ, Fisher JDM. Algorithms for solving dynamic models with occasionally binding constraints. Journal of Economic Dynamics and Control 2000;24(8):1179–1232.

Christiano LJ, Martin Eichenbaum , Evans CL. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. Journal of Political Economy 2005;113(1):1–45.

Christiano LJ, Trabandt M, Walentin K. Chapter 7 - DSGE Models for Monetary Policy Analysis. In: Friedman BM, editor. Handbook of Monetary Economics. Elsevier; 2010. p. 285–367 T2 –.

Christoffel K, Coenen G, Warne A. The New Area-wide Model of the Euro Area. European Central Bank Working Paper Series 2008;:122.

Ciarlet PG. The Finite Element Method for Elliptic Problems. 2nd ed. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2002.

Clough RW. The Finite Element Method in Plane Stress Analysis. In: Proceedings of the nd ASCE Conference on Electronic Computation. American Society of Civil Engineers; 1960. p. 1–35.

Clough RW, Wilson EL. Early finite element research at Berkeley. In: Fifth US National Conference on Computational Mechanics. Berkeley; 1999. .

Cochrane JH. Asset Pricing. (Revised Edition). Princeton University Press, 2000.

Cooley TF. Frontiers of Business Cycle Research. Princeton University Press, 1995.

Costabel M. Boundary Integral Operators on Lipschitz Domains: Elementary Results. SIAM Journal on Mathematical Analysis 1988;19(3):613–626.

Costabel M, Stephan E. Boundary integral equations for mixed boundary value problems in polygonal domains and Galerkin approximation. Banach Center Publications 1985;15(1):175–251.

Dautray R, Lions JL. Mathematical Analysis and Numerical Methods for Science and Technology:Volume 4 Integral Equations and Numerical Methods. Berlin, Heidelberg: Springer, 1990.

Del Negro M, Schorfheide F. Bayesian macroeconometrics. The Oxford handbook of Bayesian … 2011;.

Demkowicz L. Computing with HP-Adaptive Finite Elements: Volume I One and Two Dimensional Elliptic and Maxwell Problems, 2006.

Demkowicz L, Kurtz J, Pardo D, Paszenski M, Rachowicz W, Zdunek A. Computing with hp-Adaptive Finite Elements: Volume II Frontiers Three Dimensional Elliptic and Maxwell Problems with Applications. CRC Press, 2007.

Den Haan WJ, De Wind J. Nonlinear and stable perturbation-based approximations. Journal of Economic Dynamics and Control 2012;36(10):1477–1497.

Den Haan WJ, Marcet A. Solving the Stochastic Growth Model by Parameterizing Expectations. Journal of Business & Economic Statistics 1990;8.

Den Haan WJ, Marcet A. Accuracy in Simulations. The Review of Economic Studies 1994;61(1):3–17.

Di Nezza E, Palatucci G, Valdinoci E. Hitchhiker's guide to the fractional Sobolev spaces. Bulletin des Sciences Mathématiques 2012;136(5):521–573.

Díaz-Giménez J. Linear-Quadratic Approximations: An Introduction. In: Computational Methods for the Study of Dynamic Economies. OUP Oxford; 1999. p. 13–29.

Edge RM, Kiley MT, Laforte JP. A comparison of forecast performance between federal reserve staff forecasts, simple reduced-form models, and a DSGE model. Journal of Applied Econometrics 2010;25(4):720–754.

Eichner AS. A Theory of the Determination of the Mark-up Under Oligopoly. The Economic Journal 1973;83(332):1184–1200.

Erceg CJ, Guerrieri L, Gust C. SIGMA: A New Open Economy Model for Policy Analysis. International Journal of Central Banking 2006;.

Fagan G, Henry Jrm, Mestre R. An area-wide model for the euro area. Economic Modelling 2005;22(1):39–59.

Fernández-Villaverde J. The econometrics of DSGE models. SERIEs 2010;1(1-2):3–49.

Fernández-Villaverde J, Gurerrón-Quintana P, Rubio-Ramírez JF, Uribe M. Risk Matters: The Real Effects of Volatility Shocks. American Economic Review 2011;101:2530–2561.

Fernández-Villaverde J, Rubio-Ramírez JF. Comparing Dynamic Equilibrium Models to Data: A Bayesian Approach. Journal of Econometrics 2004;123(1):153–187.

Fernández-Villaverde J, Rubio-Ramírez JF. Solving DSGE models with perturbation methods and a change of variables. Journal of Economic Dynamics and Control 2006;30(12):2509–2531.

Fernández-Villaverde J, Rubio-Ramírez JF. Estimating Macroeconomic Models: A Likelihood Approach. The Review of Economic Studies 2007;74(4):1059–1087.

Fernández-Villaverde J, Rubio-Ramírez JF, Schorfheide F. Solution and Estimation Methods for DSGE Models. In: Taylor JB, Uhlig H, editors. Handbook of Macroeconomics Vol. 2. Elsevier; 2016. p. 527–724.

Forlati C, Lambertini L. Risky Mortgages in a DSGE Model. International Journal of Central Banking 2011;7(1):285–336.

Fornberg B. A Practical Guide to Pseudospectral Methods 1996;.

Fortin M. An Analysis of the Convergence of Mixed Finite Element Methods. RAIRO Analyse numérique 1977;4:341–354.

Fries TP, Belytschko T. The extended/generalized finite element method: An overview of the method and its applications. International Journal for Numerical Methods in Engineering 2010;84(3):253–304.

Galí J. Monetary Policy, inflation, and the Business Cycle: An introduction to the new Keynesian Framework. Princeton University Press, 2008.

Galor O. Discrete Dynamical Systems. Springer Science & Business Media, 2007.

Gaspar J, Judd KL. Solving Large-Scale Rational-Expectations Models. Macroeconomic Dynamics 1997;1(1):45–75.

Gertler M, Karadi P. A Model of Unconventional Monetary Policy. Journal of Monetary Economics 2011;58(1):17–34.

Glaser A, Liu X, Rokhlin V. A Fast Algorithm for the Calculation of the Roots of Special Functions. SIAM Journal on Scientific Computing 2007;29(4):1420–1438.

Gradshteyn IS, Ryzhik IM. Table of Integrals, Series, and Products. Academic Press, 2014.

Greenwald B, Stiglitz JE. Keynesian, New Keynesian and New Classical Economics. Oxford Economic Papers 1987;39(1):119–133.

Greenwald B, Stiglitz JE, Weiss A. Informational Imperfections in the Capital Market and Macroeconomic Fluctuations. The American Economic Review 1984;74(2):194–199.

Guiggiani M, Krishnasamy G, Rudolphi TJ, Rizzo FJ. A General Algorithm for the Numerical Solution of Hypersingular Boundary Integral Equations. Journal of Applied Mechanics 1992;59(3):604–614.

Hansen LP, Heaton JC, Li N. Consumption Strikes Back? Measuring Long-Run Risk. Journal of Political Economy 2008;116(2):260–302.

Hansen LP, Sargent TJ. Misspecification in Recursive Macroeconomic Theory. Book Manuscript, 2004.

Hansen LP, Sargent TJ. Recursive Models of Dynamic Linear Economies. Princeton University Press, 2013.

Harcourt GC, Kenyon P. PRICING AND THE INVESTMENT DECISION. Kyklos 1976;29(3):449–477.

Harrison R, Nikolov K, Quinn M, Ramsay G, Scott A, Thomas R. The Bank of England Quarterly Model. Technical Report; 2005.

Hart N, Kriesler P. Post-Keynesian Economics . Technical Report 2015 ECON 12; UNSW Business School; 2015.

Heinonen J. Lectures on Lipschitz Analysis. University of Jyväskylä, 2005.

Herbst EP, Schorfheide F. Bayesian Estimation of DSGE Models. Princeton University Press, 2015.

Hörmander L. The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis. Vienna: Springer, 1983a.

Hörmander L. The Analysis of Linear Partial Differential Operators II: Differential Operators with Constant Coefficients. Classics in Mathematics. Berlin, Heidelberg: Springer, 1983b.

Hörmander L. The Analysis of Linear Partial Differential Operators III: Pseudo-Differential Operators. Classics in Mathematics. Berlin, Heidelberg: Springer, 1994a.

Hörmander L. The Analysis of Linear Partial Differential Operators IV: Fourier Integral Operators. Classics in Mathematics. Berlin, Heidelberg: Springer, 1994b.

Hsiao GC, Wendland WL. A Finite Element Method for Some Integral Equations of the First Kind. Journal of Mathematical Analysis and Applications 1977;58:449–481.

Hughes TJ. The Finite Element Method: Linear Static and Dynamic Finite Element Analysis. Dover Publications, 2000.

Iacoviello M, Neri S. Housing Market Spillovers: Evidence from an Estimated DSGE Model. American Economic Journal: Macroeconomics 2010;2(2):125–164.

Inada KI. On a Two-Sector Model of Economic Growth: Comments and a Generalization. Review of Economic Studies 1963;30(2):119–127.

Jarrow R, Protter P. **Structural** Versus Reduced Form Models: a New Information Based Perspective. Journal of Investment Management 2004;2(2):1–10.

Jin HH, Judd KL. Perturbation methods for general dynamic stochastic models . Bank of England, 2002.

Jovanović BS, Süli E. Analysis of Finite Difference Schemes: For Linear Partial Differential Equations with Generalized Solutions. volume 46. London: Springer, 2014.

Judd KL. Projection Methods for Solving Aggregate Growth Models. Journal of Economic Theory 1992a;:410–452.

Judd KL. Projection methods for solving aggregate growth models. Journal of Economic Theory 1992b;58(2):410–452.

Judd KL. Numerical Methods in Economics. MIT Press, 1998.

Judd KL, Guu SM. Perturbation Solution Methods for Economic Growth Models. In: Economic and Financial Modeling with Mathematica®. New York, NY: Springer, New York, NY; 1993. p. 80–103.

Judd KL, Guu SM. Asymptotic methods for asset market equilibrium analysis. Economic Theory 2001;18(1):127–157.

Judd KL, Guu SM. Asymptotic Methods for Asset Market Equilibrium Analysis. Economic Theory 2003;18(1):127–157.

Judd KL, Maliar L, Maliar S. How to Solve Dynamic Stochastic Models Computing Expectations Just Once. NBER Working Paper 2011;.

Jung M, Steinbach O. A Finite Element-Boundary Element Algorithm for Inhomogeneous Boundary Value Problems. Computing 2002;68(1):1–17.

Kaldor N. The Scourge of Monetarism 1985;.

Kalecki M. The Principle of Increasing Risk. Economica 1937;4(16):440–447.

Kim J, Kim S, Schaumburg E, Sims CA. Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. Journal of Economic Dynamics and Control 2008;32(11):3397–3414.

Kim J, Kim SH. Spurious welfare reversals in international business cycle models. Journal of International Economics 2003;60(2):471–500.

Kimball MS. Precautionary Saving in the Small and in the Large. Econometrica 1990;58(1):53–73.

King RG, Plosser CI, Rebelo ST. Production, growth and business cycles I. The basic neoclassical model. Journal of Monetary Economics 1988a;21(2):195–232.

King RG, Plosser CI, Rebelo ST. Production, Growth and Business Cycles: II. New Directions. Journal of Monetary Economics 1988b;21(2):309–341.

King RG, Plosser CI, Rebelo ST. Production, Growth and Business Cycles: Technical Appendix. Computational Economics 2002;20(1-2):87–116.

King RG, Rebelo ST. Resuscitating Real Business Cycles. In: Handbook of Macroeconomics. Elsevier; 1999. p. 85.

King RG, Watson MW. The Solution of Singular Linear Difference Systems under Rational Expectations. International Economic Review 1998;39(4):1015–1026.

King RG, Watson MW. System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations. Computational Economics 2002;20(1-2):57–86.

Klein P. Using the generalized Schur form to solve a multivariate linear rational expectations model. Journal of Economic Dynamics and Control 2000;24(10):1405–1423.

Koop G, Pesaran MH, Potter SM. Impulse Response Analysis in Nonlinear Multivariate Models. Journal of Econometrics 1996;74(1):119–147.

Kopecky KA, Suen RM. Finite State Markov-Chain Approximations to Highly Persistent Processes. Review of Economic Dynamics 2010;13(3):701–714.

Kydland FE, Prescott EC. Time to build and aggregate fluctuations. Econometrica: Journal of the Econometric Society 1982;50(6):1345.

Ladyzhenskaya . Linear and Quasilinear Elliptic Equations. Academic Press, 1968.

Lanczos C. Trigonometric Interpolation of Empirical and Analytical Functions. Studies in Applied Mathematics 1938;17(1-4):123–199.

Leeper EM. Equilibria under 'active' and 'passive' monetary and fiscal policies. Journal of Monetary Economics 1991;27(1):129–147.

Leland HE. Saving and Uncertainty: The Precautionary Demand for Saving. The Quarterly Journal of Economics 1968;82(3):465–473.

Levintal O. Fifth-Order Perturbation Solution to DSGE Models. Journal of Economic Dynamics and Control 2017;80:1–16.

Lifanov IK, Poltavskii LN, Vainikko MGM. Hypersingular Integral Equations and Their Applications. CRC Press, 2003.

Ljungqvist L, Sargent TJ. Recursive Macroeconomic Theory. MIT Press, 2004.

Lubik TA, Schorfheide F. Computing Sunspot Equilibria in Linear Rational Expectations Models. Journal of Economic Dynamics and Control 2003;28(2):273–285.

Lucas RE. Econometric policy evaluation: A critique. Carnegie-Rochester Conference Series on Public Policy 1976;1:19–46.

Marcet A, Marshall DA. Solving Nonlinear Rational Expectations Models by Parameterized Expectations: Convergence to Stationary Solutions. Federal Reserve Bank of Minneapolis Working Paper 1994;.

Marcet A, Sargent TJ. Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information. Journal of Political Economy 1989;97(6):1306–1322.

Mas-Colell A, Whinston MD, Green JR. Microeconomic Theory. Oxford University Press, 1995.

Mason JC, Handscomb DC. Chebyshev Polynomials. CRC, 2003.

Mazya V. Sobolev Spaces in Mathematics I: Sobolev Type Inequalities. Springer, 2009a.

Mazya V. Sobolev Spaces in Mathematics II: Applicatins in Analysis and Partial Differential Equations, 2009b.

McCallum BT. On non-uniqueness in rational expectations models. Journal of Monetary Economics 1983;11(2):139–168.

McGrattan ER. Solving the Stochastic Growth Model with a Finite Element Method. Journal of Economic Dynamics and Control 1996;20(1-3):19–42.

McLean W. Strongly Eliptic Systems and Boundary Integral Equations. Cambridge University Press, 2000.

Medina JP, Soto C. The Chilean business cycles through the lens of a stochastic general equilibrium model. Central Bank of Chile Working Papers 2007;457.

Mikhailov SE. About traces, extensions and co–normal derivative operators on Lipschitz domains. In: Integral Methods in Science and Engineering: Theoretical and Practical Aspects. Birkhäuser; 2006. p. 149–160.

Mikhailov SE. Traces, extensions, co-normal derivatives and solution regularity of elliptic systems with smooth and non-smooth coefficients. Preprint in Brunel University Research Archive BURA 2009;:1–28.

Milani F. The Modeling of Expectations in Empirical DSGE Models: A Survey. Advances in Econometrics 2012;28:3–38.

Minsky HP. Can "It" Happen Again?: Essays on Instability and Finance 2015;.

Moll VH. Special Integrals of Gradshteyn and Ryzhik: the Proofs - Volume I. Taylor & Francis Group, 2015.

Moll VH. Special Integrals of Gradshteyn and Ryzhik: the Proofs - Volume II, 2016.

Murchison S, Rennison A. ToTEM: the Bank of Canada's New Quarterly Projection Model . Technical Report 97; 2006.

Muscat J. Functional Analysis: An Introduction to Metric Spaces, Hilbert Spaces, and Banach Algebras. Cham: Springer, 2014.

Neri U. Singular Integrals. volume 200 of *Lecture Notes in Mathematics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1971.

Obstfeld M, Rogoff K. Exchange Rate Dynamics Redux. Journal of Political Economy 1995;103(3):624–660.

Of G, Steinbach O, Urthaler P. Fast Evaluation of Volume Potentials in Boundary Element Methods. SIAM Journal on Scientific Computing 2010;32(2):585–602.

Palais RS. A simple proof of the Banach contraction principle. Journal of Fixed Point Theory and Applications 2007;2(2):221–223.

Privault N. Understanding Markov Chains: Examples and Applications. Springer, 2013.

Rabanal P, Rubio-Ramírez JF. Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach. Journal of Monetary Economics 2005;52(6):1151–1166.

Renardy M, Rogers RC. An Introduction to Partial Differential Equations. 2nd ed. Springer, 2004.

Rouwenhorst KG. Asset Pricing Implications of Equilibrium Business Cycle Models. In: Frontiers of Business Cycle Research. Princeton University Press; 1995. p. 294–330.

Rubio-Ramírez JF, Fernández-Villaverde J. How Structural Are Structural Parameters? In: NBER Macroeconomics Annual 2007, Volume 22. 2008. p. 83–137.

Ruge-Murcia FJ. Methods to Estimate Dynamic Stochastic General Equilibrium Models. Journal of Economic Dynamics and Control 2007;31(8):2599–2636.

Samuelson PA. The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments. Review of Economic Studies 1970;37(4):537–542.

Sandmo A. The Effect of Uncertainty on Saving Decisions. The Review of Economic Studies 1970;37(3):353.

Santos MS. Differentiability and Comparative Analysis in Discrete-time Infinite-horizon Optimization. Journal of Economic Theory 1992;57(1):222–229.

Schmitt-Grohé S, Uribe M. Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function. Journal of Economic Dynamics and Control 2004;28(4):755–775.

Schorfheide F. Estimation and Evaluation of DSGE Models: Progress and Challenges. Federal Researve Bank of Philadelphia Working Paper 2011;.

Seber GAF, Lee AJ. Linear Regression Analysis. John Wiley & Sons, 2003.

Shen J, Tang T, Wang LL. Spectral Methods: Algorithms, Analysis and Applications. Springer Science & Business Media, 2011.

Sikorski K. Optimal solution of nonlinear equations. Journal of Complexity 1985;1(2):197–209.

Simmonds JG, Mann JE. A First Look at Perturbation Theory. Dover Publications, 2013.

Simon HA. Dynamic Programming Under Uncertainty with a Quadratic Criterion Function. Econometrica 1956;24(1):74.

Sims CA. Solving Linear Rational Expectations Models. Computational Economics 2002;20(1-2):1–20.

Smets F, Wouters R. An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. Journal of the European Economic Association 2003;1(5):1123–1175.

Söderlind P. Solution and estimation of RE macromodels with optimal policy. European Economic Review 1999;43(4-6):813–823.

Šolín P, Segeth K, Dolezel I. Higher-Order Finite Element Methods. CRC Press, 2003.

Stachurski J. Economic Dynamics 2008;:1–392.

Steinbach O, Wendland WL. On C. Neumann's Method for Second-Order Elliptic Systems in Domains with Non-smooth Boundaries. Journal of Mathematical Analysis and Applications 2001;262(2):733–748.

Stiglitz J, Weiss A. Credit rationing and collateral. In: Edwards J, Franks J, Mayer C, Schaefer S, editors. Recent developments in corporate finance. Cambridge: Cambridge University Press; 2009. p. 101–135.

Stiglitz JE. Theories of Wage Rigidity. NBER Working Papers 1984;.

Stiglitz JE. The Causes and Consequences of The Dependence of Quality on Price. Journal of Economic Literature 1987;25(1):1–48.

Stiglitz JE, Weiss A. Credit Rationing in Markets with Imperfect Information. The American Economic Review 1981;71(3):393–410.

Stiglitz JE, Weiss A. Incentive Effects of Terminations: Applications to the Credit and Labor Markets. The American Economic Review 1983;73(5):912–927.

Stokey N, Lucas RE, Prescott EC. Recursive Methods in Economic Dynamics, 1989.

Suh H. Macroprudential Policy: Its Effects and Relationship to Monetary Policy. Federal Researve Bank of Philadelphia Working Paper 2012;12(28).

Swanson ET, Anderson GS, Levin AT. Higher-Order Perturbation Solutions to Dynamic, Discrete-Time Rational Expectations Models. SSRN Electronic Journal 2006;:1–31.

Tartar L. An Introduction to Sobolev Spaces and Interpolation Spaces. Springer, 2007.

Tauchen G. Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions. Economics Letters 1986;20(2):177–181.

Tauchen G, Hussey R. Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models. Econometrica 1991;59(2):371.

Theil H. A Note on Certainty Equivalence in Dynamic Planning. Econometrica 1957;25(2):346.

Thompson JF, Warsi ZU, Mastin CW. Numerical Grid Generation: Foundations and Applications. North-Holland, 1985.

Tovar Mora CE. DSGE Models and Central Banks. Economics: The Open-Access, Open-Assessment E-Journal 2009;3(2009-16):1–31.

Uhlig H. A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily. In: Computational Methods for the Study of Dynamic Economies. OUP Oxford; 1999. p. 30–61.

Vainikko G. Multidimensional Weakly Singular Integral Equations. volume 1549 of *Lecture Notes in Mathematics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993.

Verchota G. Layer potentials and regularity for the Dirichlet problem for Laplace's equation in Lipschitz domains. Journal of Functional Analysis 1984;59(3):572–611.

Vladimirov VS. Equations of Mathematical Physics. New York: Marcel Dekker, 1971.

Wood A. A Theory of Profits 1975;.

Woodford M. Interest and Prices. Foundations of a Theory of Monetary Policy. Princeton University Press, 2011.

Wu HH, Wu S. Various Proofs of the Cauchy-Schwarz Inequality. OCTOGON MATHEMATICAL MAGAZINE 2011;17(1):221–229.

Yosida K. Functional Analysis. Spinger, 1978.

朱彦元 . DSGE **模型笔记** . 同济大学中德工程学院, 2018.

插图

2.1 DSGE 模型的工作流程 18

3.1 不同 m 取值下的包络定理 31

7.1 单项式方程 105
7.2 切比雪夫多项式 $T_n(x)$ 107
7.3 切比雪夫多项式的根 108
7.4 切比雪夫多项式的根 108
7.1 张量模拟 112
7.1 kinks 113
7.2 三角方程形式 114
7.C.1 正交多项式满足的关系 133
7.C.2 正态分布 134
7.C.3 埃米特多项式 135
7.C.4 伽玛分布 137
7.C.5 贝塔分布 142
7.C.6 切比雪夫插值近似 153
7.F.1 矩形法则 $(N=4)$ 168
7.F.2 梯形法则 $(N=4)$ 168
7.F.3 3 阶牛顿——寇特斯法则，对应子区间中的 4 个配点 170
7.F.4 龙格现象 $(N=6)$ 171
7.F.5 比较矩形法则和梯形法则下，近似误差随 N 值增大的收敛情况 172
7.F.6 适应性求积 175
7.F.7 非周期方程的 T-周期扩展 $t \in [0,1]$ 178
7.F.8 $f(t)$ 和 $g(\theta)$ 180
7.F.9 $f(x), x \in [-1,1]$ 185
7.F.10 求积法则误差的衰减 187
7.G.1 傅里叶变换 188
7.G.2 洛伦兹方程 $(\alpha=2)$ 191
7.G.3 $f(t)=\cos^2(3t), t \in (-2,2)$ 196
7.G.4 锯齿波 197
7.G.5 吉布斯现象 198

8.1 有限元 τ_ℓ 及其结点 x_k 339
8.2 容许分解，以及不容许分解的三角分解 $(d=2)$ 340

8.3 有限元 τ_ℓ 和参考元 $\tau(d=2)$ 342
8.4 检测空间 $\mathcal{S}_h^1(\mathcal{T}_N)$ 的基 $\psi_k^1(x)$ 351

表格

3.1 偏好冲击的转移矩阵 29

7.C.1 经典正交多项式 132
7.G.1 傅里叶分析的常见概念界定 187
7.G.2 原方程域和傅里叶变换方程域 189
7.G.3 雅各比方程的泊松求和 202

Todo list

curse of dimensionality，可由映射法得到一定程度上的缓解，做一个 reference。 20
加入一个 reference。 22
Karl Whelan 对 forward-looking 和 backward-looking 的描述 22
to be finished... 51
补一个 reference 64
补一个 reference。 68
补一个 ref 69
补一个 ref，见下文的范例。 72
reference 81
reference 84
下文有注释，对应距离方程其实是三角方程的一种。到时把讲义内容补充到这里来。 102
reference 102
reference 和 reference 103
reference 106
reference 112
见下文，补 reference。 113
reference 115
完备 (completeness) 的脚注 218 117
这部分还需要做进一步的说明。 138
弱条件补一个 reference 160
补一个 reference 161
解的存在性和唯一性 161
解的存在性和唯一性 162
补一个描述：Hilbert space 162
reference 177
补一个 measurable function 的词条 214
section 6.1.1 299
更多自然密度的介绍见后 303
作 reference 326
写完全部模型后，待进一步汇总、整理。 376
后面方程等式。 376
下文作详细说明，做一个 ref。 378

索引

adaptive quadrature ..... 适应性求积, 174
adjoint operator ..... 伴随算子, 250
angular frequency ..... 角频率, 186, 188, 189
axiomatic set theory ..... 公理集合论, 210

Babuška-Brezzi-Ladyshenskaya condition ..... BBL 条件, 333
Banach space ..... 巴拿赫空间, 217, 249
base function ..... 基方程, 347
basis
  Chebyshev polynomial ..... 切比雪夫多项式 (基), 105
  global ..... 全局基, 104
  Jacobi polynomial ..... 雅各比多项式 (基), 105
  local ..... 局部基, 104
  monomial ..... 单项式 (基), 104
  trigonometric series ..... 三角序列 (基), 105
  unidimensional ..... 单维基, 104
basis function ..... 基方程, 101
Bellman equation ..... 贝尔曼等式, 25
Bernoulli trials ..... 伯努利实验, 363
Bessel function ..... 贝塞尔方程, 276
Beta distribution ..... 贝塔分布, 141
Beta integral ..... 贝塔积分, 121
BFMG algorithm ..... BFMG 算法, 371
bijection ..... 双射, 164
binomial theorem ..... 二项式定理, 120
boundary integral equations ..... 边界积分等式, 272
boundary value problem ..... 边界值问题, 205
Bramble-Hilbert Lemma ..... Bramble-Hilbert 引理, 233

Calderón projection ..... 卡尔德隆投影, 298, 316
capacity ..... 容度, 303
cardinality ..... 势, 211
Cauchy principal value ..... 柯西主值, 291
Cauchy sequnce ..... 柯西序列, 164
Cauchy-Schwarz inequality ..... 柯西——施瓦茨不等式, 220
chain rule ..... 链式法则, 240

Chebyshev distance ..... 切比雪夫距离, 213
Christoffel numbers ..... 克里斯托费尔数, 131
Christoffel-Darboux relation ..... 克里斯托费尔——达布关系, 128
Chu-Vandermonde summation ..... 朱世杰——范德蒙德求和公式, 122, 143
Clenshaw-Curtis Rule ..... 克伦肖——柯蒂斯求积法则, 174, 179
closed curve ..... 闭曲线, 293
closed range theorem ..... 闭值域定理, 250
coercive form ..... 强制形式, 162
collocation
..... 配点, 106, 117
orthogonal ..... 正交配点, 106, 117
compact operator ..... 紧凑算子, 258, 336
complement ..... 补集, 210
completeness
Euclidean space ..... 完备性 (欧几里得空间), 218
complex modulus ..... 复数绝对值, 219
complex number
imaginary part ..... 虚部 (复数), 164
real part ..... 实部 (复数), 164
complex plane ..... 复数平面, 219
conditional bias ..... 条件偏误, 360
conditional density ..... 条件密度, 155
conforming trial spaces ..... 协调检测空间, 324
conjugate ..... 共轭复数, 219
continuous inverse ..... 连续逆, 164
contraction mapping ..... 收缩映射, 248
convergence
absolute ..... 绝对收敛, 218
convolution ..... 卷积, 201
countable additivity ..... 可数可加性, 211
cover(topology) ..... 覆盖 (拓扑学), 210
Cramér-Rao lower Bound ..... 克拉美罗下界, 361
curse of dimensionality ..... 维数灾难, 109, 112
Céa lemma, ... 齐亚引理, 165, 326

decomposition
..... 分解 (有限元), 339
admissible ..... 容许分解, 339
differentiability classification ..... 方程微分的类, 221
dirac delta function ..... 狄拉克方程, 117, 192
Dirac distribution ..... 狄拉克分布, 235
Dirichlet boundary value condition ..... 狄利克雷边界值条件, 207, 261
Dirichlet condition
homogeneous 齐次狄利克雷条件, 163

Dirichlet to Neumann map ..... 狄利克雷到诺依曼的映射, 308
distance function ..... 距离方程, 102
distribution
    regular ..... 正则分布, 235
    singular ..... 奇异分布, 235
    stationary ..... 平稳分布, 110
distribution ..... 分布, 235
divergence theorem ..... 高斯散度定理, 205
dual space ..... 对偶空间, 219, 225, 241
duality pairing ..... 对偶配对, 241, 280
dynamic programming ..... 动态规划, 23

edge ..... 边, 339
ellipticity
    semi ..... 半椭圆性, 305, 307
embedding theorem of Sobolev space ..... 索伯列夫空间的嵌入定理, 226
embeddings ..... 嵌入, 226
ensemble mean suqare error (EMSE) ..... 整体均方误差, 360
Envelope Theorem ..... 包络定理, 23
equivalence of norms ..... 范数等价, 213, 229, 241
Euclidean norm ..... 欧几里得范数, 213
Euler integral representation ..... 欧拉积分表达式, 121
Euler transformation formula ..... 欧拉转换公式, 122
extended real number line ..... 扩展实数线, 211
exterior algebra ..... 外代数, 205
exterior normal vector ..... 外部法向量, 292

family of sets ..... 集族, 210
finite element
    neighboring ..... 相邻有限元, 341
    regular ..... 正则型有限元, 340, 346
finite elements method ..... 有限元法, 112
Fischer information matrix ..... 费雪信息矩阵, 361
form function ..... 形式方程, 347
Fourier cosine series ..... 傅里叶余弦级数, 196
Fourier series ..... 傅里叶, 194
Fourier series ..... 傅里叶级数, 187, 223
Fourier sine series ..... 傅里叶正弦级数, 196
Fourier synthesis ..... 傅里叶综合, 198
Fourier transform
    ..... 傅里叶变换, 187, 236
    Discrete ..... 离散傅里叶变换, 187
    inverse ..... 逆傅里叶变换, 236
    rotational symmetries ..... 傅里叶变换的旋转对称, 236

Semidiscrete ..... 半离散傅里叶变换, 187
Fredholm alternative theorem ..... 弗雷德霍姆二择一定理, 259
Fredholm boundary integral function
first kind ..... 第一类弗雷德霍姆边界积分方程, 316
second kind ..... 第二类弗雷德霍姆边界积分方程, 316
full width at half maximum(FWHM) ..... 半峰全宽, 191
fundamental solution ..... 基本解, 272

Galerkin method ..... 伽辽金近似法, 160
Galerkin Orthogonality ..... 伽辽金正交, 325
Galerkin-Bubnov method ..... 伽辽金——布博诺夫法, 324
Galerkin-Petrov method ..... 伽辽金——佩特洛夫法, 329
gamma distribution ..... 伽玛分布, 137
Gauss quadrature formula ..... 高斯求积公式, 131
Gauss summation formula ..... 高斯求和公式, 121
Gauss theorem ..... 高斯散度定理, 205
Gauss-Hermite quadrature ..... 高斯——埃米特求积, 158
Gaussian function ..... 高斯方程, 191
generalized partial derivative ..... 广义偏导数, 222
generating function
Hermite polynomial ..... 母方程 (埃米特多项式), 136
Jacobi polynomial ..... 母方程 (雅各比多项式), 145
Laguerre polynomial ..... 母方程 (拉盖尔多项式), 139
Legendre polynomial ..... 母方程 (勒让德多项式), 147
the first kind Chebyshev polynomial ..... 母方程 (第一类切比雪夫多项式), 150
the second kind Chebyshev polynomial ..... 母方程 (第二类切比雪夫多项式), 150
Gibbs phenomenon ..... 吉布斯现象, 197
Gram-Schmidt orthogonality process ..... 格拉姆-施密特正交化过程, 125
Green identities
first 格林第一恒等式, 206, 260
second 格林第二恒等式, 206, 272
grid ..... 网格, 113
grid distance ..... 网格距离, 213
Gårding inequality ..... Gårding 不等式, 258

h-refinement finite elements ..... h——改进 (有限元分析), 115
Hankel moment matrix ..... 汉克尔矩阵, 126
Hermitian symmetric ..... 埃米特对称, 219
Hilbert space ..... 希尔伯特空间, 162, 217
hp-adaptive finite elements ..... hp——有限元法, 115
hypergeometric
function ..... 超几何方程, 119
function, generalized ..... 广义超几何方程, 120
series ..... 超几何序列, 119

series, monic ..... 首一超几何序列, 119
hypergeometric function
Jacobi polynomial ..... 超几何方程 (雅各比多项式), 142
Laguerre polynomial ..... 超几何方程 (拉盖尔多项式), 138
Hölder condition ..... 霍德尔条件, 224
Hölder continuous ..... 霍德尔连续, 210
Hölder inequality
generalization ..... 广义霍德尔不等式, 225
Hölder inequality ..... 霍德尔不等式, 225
Hölder space ..... 霍德尔空间, 209
Hölder's inequality ..... 霍德尔不等式, 216

iamge ..... 像, 250
Inada condition ..... 稻田条件, 24
index family ..... 索引族, 209
inf-sup condition ..... 上极限——下极限条件, 163
injection ..... 内射, 259
injection ..... 单射, 164
inner product
duality ..... 双内积, 162
inner product ..... 内积, 125, 218
inner product space ..... 内积空间, 218
integrable singular points ..... 可积奇异点, 173
integration by parts ..... 分部积分, 206, 222
integration by parts ..... 分部积分公式, 221
interior boundary trace ..... 内界迹, 243
interpolation
Chebyshev ..... 切比雪夫插值, 106, 151
Lagrange ..... 拉格朗日插值法, 131
interpolation theorem ..... 插值定理, 239
intersection ..... 交集, 211

Jacobian function ..... 雅各比方程, 201

kernel ..... 核, 249
Korn inequality ..... 科恩不等式, 163
Kronecker product ..... 克罗内克乘积, 125

L-infinity norm, 213
Lagrange interpolation ..... 拉格朗日插值, 152
Lamé elasticity parameters ..... 拉梅弹性参数, 163
Laplace function ..... 拉普拉斯方程, 277
Laplace operator ..... 拉普拉斯算子, 272
Lax-Milgram theorem ..... 拉克斯一密格拉蒙定理, 163, 248, 256
Lebesgue integrable function ..... 勒贝格可积方程, 210

Lebesgue space
weighted ..... 加权勒贝格空间, 217
Leibniz rule ..... 莱布尼兹法则, 133
Lesbegue space ..... 勒贝格空间, 210
Levenberg-Marquardt algorithm ..... 莱文贝格——马夸特方法, 116
likelihood function ..... 似然方程, 359
linear frequency ..... 线频率, 189
Lipschitz continuous function ..... 利普希茨连续方程, 209
Lipschitz domain ..... 利普希茨域, 209
Lipschitz hypograph ..... 利普希茨亚图, 209
Lorentzian function ..... 洛伦兹方程, 190
LP-norm ..... LP 范数, 213

Manhattan distance ..... 曼哈顿距离, 213
Markov chain
finite ..... 有限马尔科夫链, 110
mass matrix
local ..... 局部质量矩阵, 348
matrix
identity ..... 单位矩阵, 129
infinite Jacobi ..... 无限雅各比矩阵, 128
leading principle submatrix ..... 领先主子矩阵, 128
tridiagonal ..... 三角对角矩阵, 128
maximum norm ..... 最大范数, 213
mean square error (MSE) ..... 均方误差, 360
measurable sets ..... 可测集合, 211
measurable space ..... 可测空间, 211
measure
counting ..... 计数测度, 212
measure ..... 测度, 211
mesh ..... 网格, 339
mesh size
global ..... 全局网格尺寸, 340
local ..... 局部网格尺寸, 340
metric function ..... 度量方程, 102
Minkowski inequality
integral ..... 闵可夫斯基积分不等式, 216
Minkowski inequality ..... 闵可夫斯基不等式, 215
moment ..... 矩, 125

natural density ..... 自然密度, 303
Neumann boundary value condition ..... 诺依曼边界值条件, 207, 267
Neumann series ..... 诺依曼级数, 316, 317
Newton algorithm ..... 牛顿算法, 116

Newton Conjugate Gradient Method (Newton-CG method) ..... 牛顿共轭梯度法, 371
Newton-Raphson Algorithm ..... 牛顿拉夫森算法, 372
Newton—Cotes Rule ..... 牛顿——寇特斯法则, 167
node ..... 结点, 113, 339
norm ..... 范数, 125, 163, 209, 212
norm equivalence theorem of Sobolev space, 228
normal distribution ..... 正态分布, 133
normed vector space ..... 赋范向量空间, 213
null space ..... 零空间, 249
numerical integration ..... 数值积分, 167
numerical quadrature ..... 数值求积, 167

orthogonal complement ..... 正交补, 249
orthogonality condition
  Hermite polynomial ..... 正交条件 (埃米特多项式), 134
  Jacobi polynomial ..... 正交条件 (雅各比多项式), 144
  Laguerre polynomial ..... 正交条件 (拉盖尔多项式), 138
  Legendre polynomial ..... 正交关系 (勒让德多项式), 146
  the first kind Chebishev polynomial ..... 正交条件 (第一类切比雪夫多项式), 149
  the second kind Chebishev polynomial ..... 正交条件 (第二类切比雪夫多项式), 149
Ostrogradsky-Gauss theorem ..... 奥斯特罗格拉德斯基——高斯定理, 205

p-norm ..... p 范数, 213
p-refinement finite elements ..... p——改进 (有限元分析), 115
Paley-Wiener theorems ..... 佩利——维纳诸定理, 193
parameterization
  local ..... 局部参数化, 341
Parseval theorem ..... 帕塞瓦尔定理, 223
Pascal triangle identity ..... 帕斯卡三角, 133
periodic function ..... 周期方程, 193
Pfaff's transformation formula ..... 法夫转换公式, 122
Pfaff-Saalschütz summation formula ..... 法夫——萨尔舒茨求和公式, 123
piecewise linear approximation ..... 分段线性近似, 115
piecewise smooth boundary ..... 分段平滑边界, 210
Plancherel's theorem ..... 普朗歇尔定理, 199
Pochhammer symbol ..... 伯赫哈默尔符号, 119, 121
Poincaré inequality ..... 庞加莱不等式, 232
Poisson equation ..... 泊松方程, 277
Poisson summation formula ..... 泊松求和式, 199
polar coordinate system ..... 极坐标系, 236
polynomial
  Chebyshev, first kind ..... 第一类切比雪夫多项式, 148
  Chebyshev, second kind ..... 第二类切比雪夫多项式, 148
  complete ..... 完全多项式, 112

Hermite ..... 埃米特多项式 , 133
Jacobi ..... 雅各比多项式, 141
Lagrange ..... 拉格朗日多项式, 161
Laguerre ..... 拉盖尔多项式, 137
monic orthogonal ..... 首一正交多项式, 125
orthonormal ..... 标准正交多项式, 125
potential
adjoint double layer ..... 伴随双层位势, 287
adjoint double layer ..... 伴随双层位势算子, 284
Bessel ..... 贝塞尔位势, 238
double layer ..... 双层位势算子, 287
Newton ..... 牛顿位势, 277
single layer ..... 单层位势, 282
powerset ..... 幂集, 210
probability density function(PDF) ..... 概率密度方程, 359
projection method ..... 映射法, 101
pseudospectral ..... 伪光谱, 117

quadrature points ..... 求积点, 167
quadrature rule ..... 求积法则, 167
quadrature weights ..... 求积权重, 167
quansi-uniform
globally ..... 全局拟一致, 341, 352
quasi-uniform
locally ..... 局部拟一致, 341
quotient space ..... 商空间, 213

r-refinement finite elements ..... r——改进 (有限元分析), 115
range ..... 值域, 250
rapidly decreasing function ..... 急减方程, 236
rapidly decreasing space ..... 急减空间, 236
rectangular (quadrature) rule ..... 矩形 (求积) 法则, 167
reference element ..... 参考元, 295, 341
residual function (projection method) , 101
Riemann-Stieltjes integral ..... 黎曼——斯蒂尔杰斯积分, 124
Riesz map ..... 里兹映射, 247
Riesz representation theorem ..... 里兹表现定理, 245, 272
Riesz-Schauder theorem ..... Riesz-Schauder 定理, 259
Robin boundary value condition ..... 罗宾边界值条件, 207, 271
Rodrigues formula
Hermite polynomial ..... 罗德里格斯公式 (埃米特多项式), 133
Jacoby polynomial ..... 罗德里格斯公式 (雅各比多项式), 141
Laguerre polynomial ..... 罗德里格斯公式 (拉盖尔多项式), 137
Legendre polynomial ..... 罗德里格斯共识 (勒让德多项式), 145

the first kind Chebyshev polynomial ..... 罗德里格斯公式 (第一类切比雪夫多项式), 149
the second kind Chebyshev polynomial ..... 罗德里格斯公式 (第二类切比雪夫多项式), 149
rotation ..... 旋转, 292, 295
Runge phenomenon ..... 龙格现象, 170

Schwarz inequality ..... 施瓦茨不等式, 232
score (maximum likelihood) ..... 评分 (最大似然估计), 362
second order linear differential equation
Hermite polynomial ..... 二阶线性微分方程 (埃米特多项式), 136
Jacobi Polynomial ..... 二阶线性微分方程 (雅各比多项式), 145
Laguerre polynomial ..... 二阶线性微分方程 (拉盖尔多项式), 141
selected points ..... 选点, 117
self-adjoint operator, 205
seminorm
..... 半范数, 212
trivial ..... 平凡半范数, 213
seminormed vector space ..... 半赋范向量空间, 216
sensitivity function (maximum likelihood) ..... 敏感方程 (最大似然估计), 362
sequence space
countably infinite dimensional ..... 可数无限维数列空间, 214
sigma
additivity ..... sigma 可加, 211
algebra ..... sigma 代数, 210
finite ..... sigma 有限, 212
Simpson's (quadrature) rule ..... 辛普森 (求积) 法则, 170
sinc finction
normalized ..... 标准化 sinc 方程, 274
unnormalized ..... 非标准化 sinc 方程, 274
sine integral ..... 正弦积分, 274
singular integral
weak ..... 弱奇异积分, 281, 283
Sobolev space
..... 索伯列夫空间, 160
Sobolev-Slobodeckij
norm ..... Sobolev-Slobodeckij 范数, 224
seminorm ..... Sobolev-Slobodeckij 半范数, 224
space ..... Sobolev–Slobodeckij 空间, 224
Sobolev-Slobodeckij norm ..... Sobolev-Slobodeckij 范数, 241
spectral elements ..... 光谱元, 115
spectral method ..... 光谱法, 104
spectral method ..... 有限元法, 104
spherical coordinate system ..... 三维坐标系, 237, 278
Steklov-Poincaré operator ..... Steklov-Poincaré 算子, 308
stiffness matrix ..... 刚度矩阵, 325

Stokes equations ..... 斯托克斯公式, 163
Stone-Weierstrass Theorem ..... 斯通-魏尔斯特拉斯定理, 104
Strang's First Lemma ..... 斯特朗第一引理, 327
Strang's Second Lemma ..... 斯特朗第二引理, 328
subadditivity ..... 次可加, 212
support ..... 支撑, 209
surface integral ..... 表面积分, 281, 283

taxicab geometry ..... 出租车距离, 213
telescoping sum ..... 裂项和, 129
tempered distribution space ..... 缓增分布空间, 236, 273
tensor ..... 张量, 111
test space ..... 试探空间, 235
three-term recurrence relation
    Hermite polynomial ..... 三项递推关系 (埃米特多项式), 135
    Laguerre polynomial ..... 三项递推关系 (拉盖尔多项式), 139
    Legendre polynomial ..... 三项递推关系 (勒让德多项式), 147
    monic polynomial ..... 三项递推关系 (首一正交多项式), 127
    orthonormal polynomial ..... 三项递推关系 (标准正交多项式), 127
    the first kind Chebishev polynomial ..... 三项递推关系 (第一类切比雪夫多项式), 149
    the second kind Chebishev polynomial ..... 三项递推关系 (第二类切比雪夫多项式), 149
trace ..... 迹, 206, 260
transition matrix ..... 转移矩阵, 110
transition probability ..... 转移概率, 110
trapezoidal (quadrature) rule ..... 梯形 (求积) 法则, 168
trial space ..... 检测空间, 347
triangle inequality ..... 三角不等式, 215
trigonometric integral ..... 三角积分, 274
triple 三元组, 211
truncation
    Chebyshev ..... 切比雪夫截断, 106, 154
truncation ..... 截断, 182

unbiased estimate ..... 无偏估计, 360
uniform cone condition ..... 均匀法锥条件, 239
uniform convergence ..... 均匀收敛, 117
uniform distribution ..... 均匀分布, 145
uniform norm ..... 均匀范数, 213
union ..... 并集, 210

value function ..... 价值方程, 25
vector space
    normed ..... 赋范向量空间, 215
vector space ..... 向量空间, 212

wavenumber ..... 波数, 189
weight function ..... 权重方程, 124
weighted residual methods ..... 加权残差法, 101

zero vector ..... 零向量, 212
ZFC axioms ..... ZFC 集合论, 210