密级 ____________

博士学位论文

多项式系统孤立奇异根的数值精化和可信验证

作者姓名：李楠

指导教师：支丽红 研究员

中国科学院数学与系统科学研究院

学位类别：理学博士

学科专业：应用数学

研究所：数学与系统科学研究院

2013年5月

On Isolated Singular Solutions in Polynomial System Solving

By

Nan Li

A Dissertation Submitted to
The University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Doctor of Applied Mathematics

Academy of Mathematics and Systems Science

May, 2013

摘　要

多项式系统求解问题是数学中一个既古老又经典的问题, 并在科学与工程计算中有着广泛的应用, 例如机器人技术, 编码理论, 最优化理论, 数学生物学, 计算机视觉, 博弈论, 统计学, 机器学习, 控制理论, 密码学等. 而孤立奇异根的识别和处理, 则是代数和几何计算中具有挑战性的课题之一, 并在几何建模问题中经常出现, 例如计算隐式曲线曲面的拓扑和参数曲面的交点等.

实际计算中, 我们经常需要将多项式系统的近似根精化到一个更高的精度, 但数值方法例如 Newton 法, 在孤立奇异根附近收敛速度很慢(有时甚至不收敛). 而验证多项式系统是否有孤立奇异根是一个病态问题, 因为多项式系数任意微小的扰动都可能导致一个孤立奇异根变成一族正则根(有时甚至消失). 因此, 研究多项式系统孤立奇异根的数值精化和可信验证问题, 具有重要的理论意义和很高的应用价值.

本文首先讨论了多项式系统孤立奇异根代数结构的一种恰当表示—局部对偶空间. 利用一些正则化和约化技巧, 我们提出了一种计算宽度为 1 的特殊情形下, 局部对偶空间一组既约基的新算法. 其对于带有近似系数的多项式系统, 或近似奇异根同样适用. 而且数值实验表明, 无论从计算时间还是存储空间上, 算法的效率都是较高的. 更为重要的是, 我们给出了宽度为 1 的特殊情形下, 局部对偶空间一组既约基(重结构)的参数化表示.

其次, 我们研究了多项式系统孤立奇异根的数值精化问题. 基于宽度为 1 的特殊情形下重结构的参数化表示, 我们提出了一种正则化的 Newton 法: 解一个正则化的最小二乘问题作为初始近似根的预处理, 从而得到一个更好的近似根和一个能够达到二次收敛的迭代方向; 通过计算近似局部对偶空间的一组既约基和解一个线性方程组, 从而确定合适的迭代步长. 我们证明了算法在宽度为 1 的孤立奇异根附近是二次收敛的.

最后, 我们研究了多项式系统孤立奇异根的可信验证问题. 利用区间验证方法和宽度为 1 的特殊情形下重结构的参数化表示, 我们提出了一种计算近似奇异根可信误差界的新算法, 其可以验证一个带有微小扰动的多项式系统, 在误差界内有一个宽度为 1 的孤立奇异根. 对于一般的孤立奇异根, 我们提出一种带光滑参数的 deflation 技术, 并且基于这种技术, 我们将验证宽度为 1 的孤立奇异根

的算法推广到了一般情形. 数值实验表明, 算法对于带有近似系数的多项式系统和复的孤立奇异根同样适用.

关键词： 多项式系统, 孤立奇异根, 局部对偶空间, 根的精化, 可信误差界

Abstract

Nowadays, polynomial models are ubiquitous and widely applied across the engineering and sciences, such as in robotics, coding theory, optimization, mathematical biology, computer vision, game theory, statistics, machine learning, control theory, cryptography, and numerous others. A main challenge in algebra and geometry computing is to identify and tackle singular points, which naturally occur when computing the topology of implicit curves or surfaces, the intersection of parametric surfaces in geometric modeling.

A numerical approximation is usually computed to identify an isolated solution of a polynomial system. In practice, we often need to improve the quality of numerical approximations, but numerical methods such like Newton's method converge slowly at singular solutions (or not converge). On the other hand, it is well known that to certify whether a polynomial system has an isolated singular solution is an ill-posed problem, since arbitrary small perturbations of coefficients may transform the singular solution into a cluster of simple roots (or even make it disappear). Therefore, it is hardly possible to verify this problem, if not the entire computation is performed without any rounding error (exact arithmetics).

In this thesis, we first introduce the local dual space for characterizing an isolated singular solution of a polynomial system. By employing some regularization and reduction techniques, we present a novel algorithm for computing a reduced basis of such a space for the special case of breadth one. The algorithm also works for inputs only with limited accuracy, and is efficient both in time and memory use. Moreover, it leads to a parametric representation for a reduce basis (multiplicity structure) of the local dual space.

Based on such a parametric representation and presolving a regularized least squares problem, we propose a regularized Newton's method for refining an approximate singular solution of a given polynomial system. By a careful analysis, we prove the quadratic convergence of the algorithm if the numerical approximation is close to a breadth-one isolated singular solution.

By introducing some well-chosen smoothing parameters to the given system, we develop an improved deflation technique, which derives a square and regular augmented system from an isolated singular solution in a finite number of deflations. Based on this technique, we propose an algorithm for computing verified error bounds such that a slightly perturbed polynomial system is guaranteed to possess an isolated singular solution within the computed bounds.

Keywords: solving polynomial systems, isolated singular solutions, local dual space, root refinement, verified error bounds

目　　录

表　　格

第一章　引言

1.1　问题陈述和研究动机

多项式系统求解问题是数学中一个既古老又经典的问题, 并在科学与工程计算中有着广泛的应用, 例如机器人技术, 编码理论, 最优化理论, 数学生物学, 计算机视觉, 博弈论, 统计学, 机器学习, 控制理论, 密码学等 [39]. 而孤立奇异根的识别和处理, 则是代数和几何计算中一个具有挑战性的课题, 并在几何建模问题中经常出现, 例如计算隐式曲线曲面的拓扑和参数曲面的交点等 [1].

对于给定的多项式系统

$$F(\mathbf{x}) = \begin{cases} f_1(x_1, \ldots, x_n), \\ f_2(x_1, \ldots, x_n), \\ \quad\quad\vdots \\ f_n(x_1, \ldots, x_n), \end{cases}$$

$f_i \in \mathbb{R}[x_1, \ldots, x_n] = \mathbb{R}[\mathbf{x}]$, $i = 1, \ldots, n$, 求解 $F(\mathbf{x}) = \mathbf{0}$ 的符号数值方法主要有: Gröbner 基方法 [2], Ritt-Wu 特征列方法 [31, 42], 实根隔离方法 [32, 45], 乘法矩阵方法 [30, 38], 同伦连续方法 [41], 半正定规划方法 [16] 等. 符号方法能够精确求解, 但速度较慢, 适合求解中小规模的多项式系统; 数值方法速度较快, 适合求解较大规模的多项式系统, 但只能得到有限精度的近似根.

实际问题中, 我们经常需要将已有的近似根精化到一个更高的精度. 当近似根 $\tilde{\mathbf{x}}$ 足够接近 F 的一个正则根 $\hat{\mathbf{x}}$ 时($\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\| = \epsilon \ll 1$), Newton 法

$$\tilde{\mathbf{y}} = -F_{\mathbf{x}}(\tilde{\mathbf{x}})^{-1}F(\tilde{\mathbf{x}})$$

可以达到二次收敛 $\|\tilde{\mathbf{x}} + \tilde{\mathbf{y}} - \hat{\mathbf{x}}\| = \mathcal{O}(\epsilon^2)$ [12]. 然而当近似根 $\tilde{\mathbf{x}}$ 位于 F 的一个奇异根 $\hat{\mathbf{x}}$ 附近时($\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) < n$), Newton 法的收敛速度通常是线性的, 而且有时甚至不收敛 [11]. 这就促使我们研究多项式系统孤立奇异根的数值精化问题. 近年来, 一些数学家致力于利用孤立奇异根的重结构 [20, 22, 23, 27, 38, 43], 来精化多项式系统的近似奇异根 [7, 10, 17, 28, 29, 44]. 本文中, 我们首次给出了宽度为 1 的特殊情形下, 重结构的参数化表示(详情见第二章); 基于参数化的重结

构, 我们提出了一种正则化的 Newton 法, 并且证明了它在宽度为 1 的孤立奇异根附近的二次收敛性(详情见第三章).

大多数基于浮点计算的求根算法, 只能对数值结果的近似程度提供误差“估计”; 近年发展起来的区间验证方法, 能够计算近似根的误差“界”, 从而证明精确根的存在性. 多项式系统正则根的区间验证主要基于如下定理 [15, 25, 33]:

定理 1.1. 假设 $\tilde{\mathbf{x}} \in \mathbb{R}^n$ 是 $F = \{f_1, \ldots, f_n\}$ 的近似根, $f_i \in \mathbb{R}[\mathbf{x}]$. 对于给定的区间向量 $\mathbf{X} \in \mathbb{IR}^n$ 满足 $\mathbf{0} \in \mathbf{X}$ 和区间矩阵 $M \in \mathbb{IR}^{n\times n}$ 满足 $\nabla f_i(\tilde{\mathbf{x}} + \mathbf{X}) \subseteq M_{i,:}$,如果区间判定条件

$$-F_{\mathbf{x}}^{-1}(\tilde{\mathbf{x}})F(\tilde{\mathbf{x}}) + (I_n - F_{\mathbf{x}}^{-1}(\tilde{\mathbf{x}})M)\mathbf{X} \subseteq \mathrm{int}(\mathbf{X})$$

成立, 则存在唯一的 $\hat{\mathbf{x}} \in \tilde{\mathbf{x}} + \mathbf{X}$ 满足 $F(\hat{\mathbf{x}}) = \mathbf{0}$.

而验证多项式系统是否有奇异根是一个病态问题, 因为多项式系数任意微小的扰动都可能导致一个奇异根变成一族正则根. 因此, 利用带有舍入误差的浮点计算, 来验证多项式系统的奇异根的存在性是不可能的 [35]. 针对这个难题, 通过对原多项式系统添加一些光滑参数, 我们提出了一种计算近似奇异根可信误差界的新算法, 其可以验证一个带有微小扰动的多项式系统, 在狭窄的扰动误差界内有一个孤立奇异根(详情见第四章).

1.2 论文结构和主要结果

第二章中, 我们回顾了一些交换代数和代数几何的基本知识, 介绍了多项式系统孤立奇异根的定义和代数结构, 讨论了局部对偶空间的性质和两种计算重结构的经典算法: 基于定义的 Macaulay 方法和基于封闭性的 MMM 方法. 利用一些正则化和约化技巧, 我们提出了一种计算宽度为 1 的特殊情形下, 局部对偶空间一组既约基的新算法 MSB1 , 从而将矩阵的最大规模降低为方程个数 $\times$ 变元个数(与孤立奇异根的重数无关). 更为重要的是, 我们给出了宽度为 1 的特殊情形下, 局部对偶空间一组既约基的参数化表示. 算法 MSB1 对于解析系统同样适用, 而算法 AMSB1 可以用于计算带有近似系数的多项式系统或近似奇异根, 在宽度为 1 的特殊情形下, 近似局部对偶空间的一组既约基. 数值实验表明, 无论从计算时间还是存储空间上, 算法的效率都是较高的.

第三章中, 我们介绍了一种经典的 deflation 技术, 讨论了其关于孤立奇异根的有限终止性, 并且证明了宽度为 1 的特殊情形下的一个猜想: 一次 deflation 会

使多项式系统在孤立奇异根处局部对偶空间的深度严格降 1. 对于宽度为 1 的孤立奇异根附近近似根的数值精化问题, 我们提出了一种新算法 MRRB1: 解一个正则化的最小二乘问题作为初始近似根的预处理, 从而得到一个更好的近似根和一个能够达到二次收敛的迭代方向; 通过计算近似局部对偶空间的一组既约基和解一个线性方程组, 从而确定合适的迭代步长. 我们证明了算法 MRRB1 在宽度为 1 的孤立奇异根附近是二次收敛的, 而且数值实验表明, 算法对于带有近似系数的多项式系统和超定系统(方程个数大于变元个数)同样有效.

第四章中, 我们介绍了一种经典的区间验证方法, 讨论了验证多项式系统孤立奇异根的若干难点和解决思路. 利用第二章中给出的宽度为 1 的特殊情形下, 局部对偶空间一组既约基的参数化表示, 我们提出了一种计算近似奇异根可信误差界的新算法, 其可以验证一个带有微小扰动的多项式系统, 在误差界内有一个宽度为 1 的孤立奇异根. 对于一般的孤立奇异根, 我们提出一种带光滑参数的 deflation 技术, 并且证明了它的有限终止性, 从而将验证宽度为 1 的孤立奇异根的算法推广到一般情形. 在 Matlab 中我们实现了算法 VISS 并且测试了大量的例子, 数值实验表明, 算法对于带有近似系数的多项式系统和复的孤立奇异根同样适用.

最后一章中, 我们总结了已有的工作成果并讨论了今后的努力方向.

第二章　孤立奇异根的代数结构

本章中, 我们回顾了一些交换代数和代数几何的基本知识, 介绍了多项式系统孤立奇异根的定义和代数结构, 讨论了局部对偶空间的性质和重结构的计算, 并且提出了计算一种特殊情形下局部对偶空间一组既约基的新算法.

2.1　根的代数结构

$\mathbb{K}$ 是一个特征为零的域, $\overline{\mathbb{K}}$ 是它的代数闭包, 比如实数域 $\mathbb{K}=\mathbb{R}$, $\overline{\mathbb{K}}=\mathbb{C}$. 如果 $\overline{\mathbb{K}}=\mathbb{K}$, 则称 $\mathbb{K}$ 是代数闭的, 比如复数域 $\overline{\mathbb{K}}=\mathbb{K}=\mathbb{C}$. 事实上, 本文只考虑这两个域 $\mathbb{R}$ 和 $\mathbb{C}$. $\mathbb{K}$ 上定义的 n 元多项式环记为 $\mathbb{K}[\mathbf{x}]:=\mathbb{K}[x_1,\ldots,x_n]$.

定义 2.1 (孤立根)**.** 如果 $\hat{\mathbf{x}}\in\overline{\mathbb{K}}^n$ 和多项式系统 $F=\{f_1,\ldots,f_m\}$, $f_i\in\mathbb{K}[\mathbf{x}]$ 满足

$$\exists\epsilon>0:\{\mathbf{y}\in\overline{\mathbb{K}}^n\mid\|\mathbf{y}-\hat{\mathbf{x}}\|<\epsilon\}\cap F^{-1}(\mathbf{0})=\{\hat{\mathbf{x}}\},$$

则称 $\hat{\mathbf{x}}$ 是 $F(\mathbf{x})=\mathbf{0}$ 的孤立根. $\|\cdot\|$ 是 2-范数, $F^{-1}(\mathbf{0})\subset\overline{\mathbb{K}}^n$ 是 $F(\mathbf{x})=\mathbf{0}$ 的解集.

定义 2.2 (奇异根)**.** 如果 $\hat{\mathbf{x}}\in\overline{\mathbb{K}}^n$ 和多项式系统 $F=\{f_1,\ldots,f_m\}$, $f_i\in\mathbb{K}[\mathbf{x}]$ 满足

$$\hat{\mathbf{x}}\in F^{-1}(\mathbf{0}),\ \mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}}))<n,$$

则称 $\hat{\mathbf{x}}$ 是 $F(\mathbf{x})=\mathbf{0}$ 的奇异根. $F_{\mathbf{x}}$ 是 F 关于变元 $\mathbf{x}$ 的 Jacobian 矩阵.

例 2.1. 考虑如下例子 $F=\{x_1^2x_2-x_1x_2^2,x_1-x_2^2\}$. $(0,0)$ 是 $F(x_1,x_2)=\mathbf{0}$ 的孤立奇异根, 而 $(1,1)$ 只是孤立根. 因为 $F^{-1}(\mathbf{0})=\{(0,0),(1,1)\}$, 且

$$F_{\mathbf{x}}=\begin{pmatrix}2x_1x_2-x_2^2 & x_1^2-2x_1x_2\\ 1 & -2x_2\end{pmatrix},$$

$$F_{\mathbf{x}}(0,0)=\begin{pmatrix}0&0\\1&0\end{pmatrix},\ F_{\mathbf{x}}(1,1)=\begin{pmatrix}1&-1\\1&-2\end{pmatrix}.$$

接下来, 我们简单地回顾一些交换代数和代数几何的基本知识, 进而讨论多项式系统孤立奇异根的代数结构. 这些基本知识都可以在 D. Cox, J. Little 和 D. O'Shea 关于交换代数和代数几何的经典教科书 [3, 4] 中找到.

2.1.1 理想和代数簇

定义 2.3 (多项式理想)**.** $\mathbb{K}[\mathbf{x}]$ 中的一个子集 I 被称为理想, 如果它满足:

1. $0 \in I$.

2. 对任意的 $f, g \in I$, $f + g \in I$.

3. 对任意的 $f \in I$ 和 $h \in \mathbb{K}[\mathbf{x}]$, $hf \in I$.

对于给定的一组多项式 $f_1, \ldots, f_m \in \mathbb{K}[\mathbf{x}]$, 可以定义子集 $I \in \mathbb{K}[\mathbf{x}]$:

$$I = (f_1, \ldots, f_m) := \left\{ \sum_{i=1}^{m} h_i f_i \mid \forall\, h_1, \ldots, h_m \in \mathbb{K}[\mathbf{x}] \right\}.$$

显然, I 是一个$\mathbb{K}[\mathbf{x}]$ 中的理想, 称之为由 $f_1, \ldots, f_m \in \mathbb{K}[\mathbf{x}]$ 生成的理想.

定义 2.4 (仿射代数簇)**.** 理想 I 的(仿射)代数簇定义为

$$V(I) := \{\mathbf{y} \in \overline{\mathbb{K}}^n \mid f(\mathbf{y}) = 0, \forall\, f \in I\}.$$

如果 F 是一个由 $f_1, \ldots, f_m \in \mathbb{K}[\mathbf{x}]$ 构成的多项式系统, I 是由这些多项式生成的理想, 那么 $F^{-1}(\mathbf{0}) = V(I)$. 因此, 可以通过理想的代数簇, 来研究多项式系统的根. 另一方面, 对于给定的代数簇 $V \subset \overline{\mathbb{K}}^n$, 它的消逝理想定义为

$$I(V) := \{f \in \mathbb{K}[\mathbf{x}] \mid f(\mathbf{y}) = 0, \forall\, \mathbf{y} \in V\}.$$

消逝理想 $I(V)$ 是所有以代数簇 V 为解集的多项式构成的理想, 所以它通常要比 I 大. 为了建立起 I 与 $I(V(I))$ 之间的联系, 需要定义根理想:

$$\sqrt{I} := \left\{f \in \mathbb{K}[\mathbf{x}] \mid \exists\, k \in \mathbb{Z}^+ : f^k \in I\right\}.$$

I 与 $I(V(I))$ 的联系就是著名的 Hilbert 零点定理:

定理 2.1. *$\mathbb{K}$ 是一个代数闭域, $I \subseteq \mathbb{K}[\mathbf{x}]$ 是一个理想, 那么 $\sqrt{I} = I(V(I))$.*

例 2.2 (例 2.1 的延续)**.** 令理想 $I = (x_1^2 x_2 - x_1 x_2^2, x_1 - x_2^2)$, 那么 $V(I) = F^{-1}(\mathbf{0}) = \{(0,0),(1,1)\}$, $\sqrt{I} = I(V(I)) = (x_1 - x_2^2, x_1 - x_2)$. 令 $G = \{x_1 - x_2^2, x_1 - x_2\}$, 那么 $V(\sqrt{I}) = G^{-1}(\mathbf{0}) = \{(0,0),(1,1)\}$. 但这时 $(0,0)$ 只是 $G(x_1,x_2) = \mathbf{0}$ 的孤立根而非奇异根. 所以, F 和 G 的解集虽然相同, 但根的代数结构却不同.

由上面的例子可以看出, 根理想只能描述多项式系统根的“位置”, 但并不能保持根的“结构”. 为了刻画根的代数结构, 下面引入不可约代数簇, 素理想和准素理想的概念.

定义 2.5 (不可约代数簇)**.** 代数簇 $V \subset \overline{\mathbb{K}}^n$ 被称为是不可约的, 如果 $V = V_1 \cup V_2$, 那么 $V_1 = V$ 或 $V_2 = V$.

定义 2.6 (素理想)**.** 称理想 $I \subset \mathbb{K}[\mathbf{x}]$ 为素理想, 如果 $fg \in I$, 那么 $f \in I$ 或 $g \in I$.

如果 $\mathbb{K}$ 是代数闭的, 不可约代数簇与素理想之间存在着一一对应的关系. 因为 $V \subset \mathbb{K}^n$ 是不可约代数簇, 当且仅当 $I(V) \subset \mathbb{K}[\mathbf{x}]$ 是素理想. 并且对任意的代数簇 $V \subset \mathbb{K}^n$, 都存在一个有限不可约分解:

定理 2.2. *$V = V_1 \cup \cdots \cup V_k$, V_i 是不可约代数簇.*

对应地, 对任意的根理想 $\sqrt{I} \subset \mathbb{K}[\mathbf{x}]$, 都存在一个有限素理想分解:

定理 2.3. *$\sqrt{I} = P_1 \cap \cdots \cap P_k$, P_i 是素理想.*

例 2.3 (例 2.1 的延续)**.** $I = (x_1^2x_2 - x_1x_2^2, x_1 - x_2^2)$, $V(I) = \{(0,0)\} \cup \{(1,1)\}$, 且

$$\sqrt{I} = (x_1, x_2) \cap (x_1 - 1, x_2 - 1).$$

对任意的理想 $I \subset \mathbb{K}[\mathbf{x}]$, 是否也可以将其分解成有限多个简单理想的交集? 答案是肯定的.

定义 2.7 (准素理想)**.** 理想 $I \subset \mathbb{K}[\mathbf{x}]$ 被称为是准素理想, 如果 $fg \in I$, 那么 $f \in I$ 或 $g^k \in I$.

定理 2.4. *$I = Q_1 \cap \cdots \cap Q_k$, Q_i 是准素理想.*

准素理想 Q 的根理想 $\sqrt{Q}$ 是素理想. 因此, 代数簇 $V(Q) = V(\sqrt{Q})$ 是不可约的, 且 $V(I) = V(Q_1) \cup \cdots \cup V(Q_k)$ 是 $V(I)$ 的不可约分解. 更为重要的是, 理想的准素分解保持根的代数结构不变.

例 2.4 (例 2.1 的延续)**.** $I = (x_1^2x_2 - x_1x_2^2, x_1 - x_2^2)$ 存在准素分解:

$$I = (x_2^4, x_1 - x_2^2) \cap (x_1 - 1, x_2 - 1).$$

准素理想 $(x_2^4, x_1 - x_2^2)$ 的根理想是 (x_1, x_2). 令 $G = \{x_2^4, x_1 - x_2^2\}$, 那么, $(0,0)$ 仍然是 G 的孤立奇异根.

下一小节中, 我们进一步讨论多项式系统孤立奇异根的代数结构.

2.1.2 重数和重结构

首先, 我们介绍单变元多项式根的重数的定义.

定义 2.8 (单变元多项式根的重数)**.** 称 $\hat{x}\in\mathbb{K}$ 是多项式 $p(x)\in\mathbb{K}[x]$ 的 μ 重根, 如果 $\hat{x}$ 满足:

$$p(\hat{x})=p'(\hat{x})=\cdots=p^{(\mu-1)}(\hat{x})=0,\text{ 且 }p^{(\mu)}(\hat{x})\neq 0.$$

即存在多项式 $q(x)\in\mathbb{K}[x]$, 满足 $q(\hat{x})\neq 0$, 使得 $p(x)=(x-\hat{x})^{\mu}q(x)$.

将上述定义翻译成理想与代数簇的语言:

$$\text{理想: } I=(p),\, p(x)=(x-\hat{x}_1)^{\mu_1}\cdots(x-\hat{x}_k)^{\mu_k},$$

$$\text{代数簇和根理想: } V(I)=\{\hat{x}_1,\ldots,\hat{x}_k\},\, \sqrt{I}=\left(\prod_{i=1}^{k}(x-\hat{x}_i)\right),$$

$$\text{不可约分解: } V(I)=\{\hat{x}_1\}\cup\cdots\cup\{\hat{x}_k\},\, \sqrt{I}=(x-\hat{x}_k)\cap\cdots\cap(x-\hat{x}_k),$$

$$\text{准素分解: } I=((x-\hat{x}_k)^{\mu_1})\cap\cdots\cap((x-\hat{x}_k)^{\mu_k}).$$

因此, 对于单变元多项式, 根的代数结构由根的重数决定.

定义 2.9 (多变元多项式根的重数)**.** 假设 $\hat{\mathbf{x}}\in\mathbb{K}^n$ 是多项式系统 $F=\{f_1,\ldots,f_m\}$ 的孤立根, $f_i\in\mathbb{K}[\mathbf{x}]$, $Q_{\hat{\mathbf{x}}}$ 是 $\hat{\mathbf{x}}$ 在理想 $I=(f_1,\ldots,f_m)$ 的准素分解中对应的准素分支, 满足 $\sqrt{Q_{\hat{\mathbf{x}}}}=(x_1-\hat{x}_1,\ldots,x_n-\hat{x}_n)$, 那么, $\hat{\mathbf{x}}$ 的重数 μ 定义为 $\dim(\mathbb{K}[\mathbf{x}]/Q_{\hat{\mathbf{x}}})$.

对于多变元多项式, 根的重数不足以决定根的代数结构, 比如两个不同的准素理想 (x^3,y) 和 (x^2,xy,y^2), 在公共根 $(0,0)$ 处却有相同的重数 $\mu=3$.

例 2.5 (例 2.1 的延续)**.** $\hat{\mathbf{x}}=(0,0)$, $Q_{\hat{\mathbf{x}}}=(x_2^4,x_1-x_2^2)$, $\mu=\dim(\mathbb{R}[\mathbf{x}]/Q_{\hat{\mathbf{x}}})=4$. 事实上, 多项式系统 $F=\{x_1^2x_2-x_1x_2^2,x_1-x_2^2\}$ 和 $G=\{x_2^4,x_1-x_2^2\}$ 都满足:

$$\begin{aligned}
F(\hat{\mathbf{x}})&=\mathbf{0}, & G(\hat{\mathbf{x}})&=\mathbf{0},\\
\tfrac{\partial F}{\partial x_2}(\hat{\mathbf{x}})&=\mathbf{0}, & \tfrac{\partial G}{\partial x_2}(\hat{\mathbf{x}})&=\mathbf{0},\\
\left(\tfrac{1}{2}\tfrac{\partial^2 F}{\partial x_2^2}+\tfrac{\partial F}{\partial x_1}\right)(\hat{\mathbf{x}})&=\mathbf{0}, & \left(\tfrac{1}{2}\tfrac{\partial^2 G}{\partial x_2^2}+\tfrac{\partial G}{\partial x_1}\right)(\hat{\mathbf{x}})&=\mathbf{0},\\
\left(\tfrac{1}{6}\tfrac{\partial^3 F}{\partial x_3^2}+\tfrac{\partial^2 F}{\partial x_1\partial x_2}\right)(\hat{\mathbf{x}})&=\mathbf{0}, & \left(\tfrac{1}{6}\tfrac{\partial^3 G}{\partial x_2^3}+\tfrac{\partial^2 G}{\partial x_1\partial x_2}\right)(\hat{\mathbf{x}})&=\mathbf{0}.
\end{aligned}$$

显然, 这些微分条件是单变元多项式的导数在多变元情形下的推广, 且所有在 $\hat{\mathbf{x}}$ 处满足这些微分条件的多项式构成的集合, 恰为 $Q_{\hat{\mathbf{x}}}$ [27, 定理 8].

由上面的例子可以看出, 对于多变元多项式系统, 理想的准素分解所保持的根的代数结构除了根的重数, 还包括一些在根处满足的微分条件. 它们就是下一节中定义的局部对偶空间的一组基.

2.2　局部对偶空间

本节中, 我们介绍局部对偶空间的定义和性质, 并讨论计算它的一组基的两种著名方法: Macaulay 方法和 MMM 方法.

2.2.1　定义

对于给定的 $\boldsymbol{\alpha} \in \mathbb{N}^n$ 和 $\hat{\mathbf{x}} \in \mathbb{K}^n$, 微分泛函 $\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}} : \mathbb{K}[\mathbf{x}] \to \mathbb{K}$ 定义如下:

$$\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}(f) := \frac{1}{\alpha_1! \cdots \alpha_n!} \cdot \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}(\hat{\mathbf{x}}), \forall f \in \mathbb{K}[\mathbf{x}].$$

微分次数 $\deg(\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}) := |\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_n$. 微分泛函的正规化形式有如下性质:

$$\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}\left(\prod_{i=1}^{n}(x_i - \hat{x}_i)^{\boldsymbol{\beta}_i}\right) = \begin{cases} 1, & \text{如果 } \boldsymbol{\alpha} = \boldsymbol{\beta}, \\ 0, & \text{如若不然.} \end{cases} \tag{2.1}$$

当 $\hat{\mathbf{x}}$ 上下文明确时, 我们将 $\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}$ 简写为 $\mathbf{d}^{\boldsymbol{\alpha}} = d_1^{\alpha_1} \cdots d_n^{\alpha_n}$, $d_i^{\alpha_i} = \frac{1}{\alpha_i!}\frac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}}$. 记 $\mathfrak{D}_{\hat{\mathbf{x}}} := \mathrm{Span}_{\mathbb{K}}\{\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathbb{N}^n\}$, 是所有上述微分泛函, 在 $\mathbb{K}$ 上生成的向量空间. 假设 $\Lambda = \sum c_{\boldsymbol{\alpha}} \mathbf{d}^{\boldsymbol{\alpha}} \in \mathfrak{D}_{\hat{\mathbf{x}}}$, $c_{\boldsymbol{\alpha}} \in \mathbb{K}$, 那么, Λ 的支撑定义为 $\mathrm{supp}(\Lambda) := \{\mathbf{d}^{\boldsymbol{\alpha}} \mid c_{\boldsymbol{\alpha}} \neq 0\}$, 微分阶数定义为 $\deg(\Lambda) := \max\{\deg(\mathbf{d}^{\boldsymbol{\alpha}}) \mid \mathbf{d}^{\boldsymbol{\alpha}} \in \mathrm{supp}(\Lambda)\}$.

定义 2.10 (局部对偶空间)**.** 假设 $\hat{\mathbf{x}} \in \mathbb{K}^n$ 是 $F = \{f_1, \ldots, f_m\}$ 的根, $f_i \in \mathbb{K}[\mathbf{x}]$, 理想 $I = (f_1, \ldots, f_m)$ 在 $\hat{\mathbf{x}}$ 处的局部对偶空间定义为, 对 I 中任意元素都消逝的 $\mathfrak{D}_{\hat{\mathbf{x}}}$ 中微分泛函构成的子空间:

$$\mathcal{D}_{\hat{\mathbf{x}}}(I) := \{\Lambda \in \mathfrak{D}_{\hat{\mathbf{x}}} \mid \Lambda(f) = 0, \forall f \in I\}.$$

局部对偶空间的维数等于根的重数,

$$\dim(\mathcal{D}_{\hat{\mathbf{x}}}(I)) = \dim(\mathbb{K}[\mathbf{x}]/Q_{\hat{\mathbf{x}}}).$$

现在我们可以解释, 为什么对应于 $(0,0)$ 且重数为 3 的准素分支有两种不同的形式 $Q_1 = (x^3, y)$ 和 $Q_2 = (x^2, xy, y^2)$, 原因就是它们的局部对偶空间不同:

$$\mathcal{D}_{\hat{\mathbf{x}}}(Q_1) = \mathrm{Span}_{\mathbb{R}}\{1, d_1, d_1^2\}, \mathcal{D}_{\hat{\mathbf{x}}}(Q_2) = \mathrm{Span}_{\mathbb{R}}\{1, d_1, d_2\}.$$

定义 2.11 (宽度和深度)**.** $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的宽度 κ 和深度 ρ 分别定义为:

$$\kappa := n - \operatorname{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})),\ \rho := \max\{\deg(\Lambda) \mid \Lambda \in \mathcal{D}_{\hat{\mathbf{x}}}(I)\}.$$

显然, $\hat{\mathbf{x}}$ 是 F 的孤立奇异根, 当且仅当 $\kappa > 0$, 且 $0 < \rho < \infty$.

例 2.6 (例 2.1 的延续)**.** $F = \{x_1^2x_2 - x_1x_2^2, x_1 - x_2^2\}$ 在 $\hat{\mathbf{x}} = (0,0)$ 处满足的微分条件可以简写为:

$$\mathbf{d^0}(F) = \mathbf{0}, d_2(F) = \mathbf{0}, (d_2^2 + d_1)(F) = \mathbf{0}, (d_2^3 + d_1d_2)(F) = \mathbf{0}.$$

事实上, $\mathcal{D}_{\hat{\mathbf{x}}}(I) = \operatorname{Span}_{\mathbb{R}}\{\mathbf{d^0}, d_2, d_2^2 + d_1, d_2^3 + d_1d_2\}$, 它的宽度为 1, 深度为 3. 且

$$Q_{\hat{\mathbf{x}}} = \{f \in \mathbb{R}[\mathbf{x}] \mid \Lambda(f) = 0, \forall \Lambda \in \mathcal{D}_{\hat{\mathbf{x}}}(I)\} = (x_2^4, x_1 - x_2^2).$$

因此, 对于多变元多项式系统, 孤立奇异根的代数结构由局部对偶空间决定. 下面我们介绍局部对偶空间的一个重要性质—封闭性.

2.2.2 封闭性

定义一个反微分算子 $\Phi_{x_i} : \mathfrak{D}_{\hat{\mathbf{x}}} \to \mathfrak{D}_{\hat{\mathbf{x}}}$:

$$\Phi_{x_i}(\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}) := \begin{cases} \mathbf{d}_{\hat{\mathbf{x}}}^{(\alpha_1,\ldots,\alpha_i-1,\ldots,\alpha_n)}, & \text{如果 } \alpha_i > 0, \\ 0, & \text{如若不然.} \end{cases}$$

记 $\mathcal{D}_{\hat{\mathbf{x}}}^t(I)$, 是 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 中微分阶数小于或等于 t 的微分泛函构成的子空间, $t \in \mathbb{N}$. 事实上, 宽度 $= \dim(\mathcal{D}_{\hat{\mathbf{x}}}^1(I)) - \dim(\mathcal{D}_{\hat{\mathbf{x}}}^0(I))$, 深度 $= \min\{t \in \mathbb{N} \mid \mathcal{D}_{\hat{\mathbf{x}}}^t(I) = \mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)\}$.

命题 2.5. $\forall \Lambda \in \mathcal{D}_{\hat{\mathbf{x}}}^t(I)$, $\Phi_{x_i}(\Lambda) \in \mathcal{D}_{\hat{\mathbf{x}}}^{t-1}(I)$, $i = 1, \ldots, n$.

证明. 显然, $\deg(\Phi_{x_i}(\Lambda)) \leq t - 1$. 多变元微分的乘法法则是著名的 Leibniz 公式:

$$\mathbf{d}^{\boldsymbol{\alpha}}(fg) = \sum_{\boldsymbol{\beta} \leq \boldsymbol{\alpha}} \mathbf{d}^{\boldsymbol{\beta}}(f)\mathbf{d}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}(g),$$

$\boldsymbol{\beta} \leq \boldsymbol{\alpha}$ 意味着 $\beta_i \leq \alpha_i$, $i = 1, \ldots, n$. 对 $\forall f \in I$, $(x_i - \hat{x}_i)f \in I$, 因此

$$0 = \Lambda((x_i - \hat{x}_i)f) = \mathbf{d^0}(x_i - \hat{x}_i)\cdot\Lambda(f) + 1\cdot\Phi_{x_i}(\Lambda)(f) + \Lambda(x_i - \hat{x}_i)\cdot\mathbf{d^0}(f) = \Phi_{x_i}(\Lambda)(f).$$

得到 $\Phi_{x_i}(\Lambda)(f) = 0$ 对 $\forall f \in I$ 都成立. 所以 $\Phi_{x_i}(\Lambda) \in \mathcal{D}_{\hat{\mathbf{x}}}^{t-1}(I)$. □

定义 2.12 (封闭性). 称集合 $\mathbf{\Lambda} = \{\Lambda_1, \ldots, \Lambda_k\}$ 是封闭的, $\Lambda_i \in \mathfrak{D}_{\hat{\mathbf{x}}}$, 如果满足

$$\text{对任意给定} \, 1 \le i \le n, 1 \le j \le k, \exists \, c_{i,j}^1, \ldots, c_{i,j}^k \in \mathbb{K} : \Phi_{x_i}(\Lambda_j) = \sum_{l=1}^{k} c_{i,j}^l \Lambda_l.$$

假设 $\mathbf{\Lambda} = \{\Lambda_1, \ldots, \Lambda_\mu\}$ 是封闭的, 且在 $\mathbb{K}$ 上线性无关, 如果满足 $\Lambda_l(f_j) = 0$, 那么 $\mathbf{\Lambda}$ 是 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组基, $I = (f_1, \ldots, f_m)$, $1 \le l \le \mu$, $1 \le j \le m$.

根据封闭性, 得到 $\Lambda_l((x_i - \hat{x}_i) f_j) = \Phi_{x_i}(\Lambda_l)(f_j) = 0$, 因此, 可以归纳地证明出 $\Lambda_l(f) = 0$, 对 $\forall f \in I$ 都成立. 再由 $\mathbf{\Lambda}$ 在 $\mathbb{K}$ 上线性无关, 和 $\dim(\mathcal{D}_{\hat{\mathbf{x}}}(I)) = \mu$, 推出 $\mathbf{\Lambda}$ 是 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组基.

例 2.7 (例 2.1 的延续). 验证 $\mathbf{\Lambda} = \{\mathbf{d^0}, d_2, d_2^2 + d_1, d_2^3 + d_1 d_2\}$ 是封闭的:

$$\begin{cases} \Phi_{x_1}(d_2) = 0 \\ \Phi_{x_2}(d_2) = \mathbf{d^0} \end{cases}, \begin{cases} \Phi_{x_1}(d_2^2 + d_1) = \mathbf{d^0} \\ \Phi_{x_2}(d_2^2 + d_1) = d_2 \end{cases}, \begin{cases} \Phi_{x_1}(d_2^3 + d_1 d_2) = d_2 \\ \Phi_{x_2}(d_2^3 + d_1 d_2) = d_2^2 + d_1 \end{cases}.$$

且 $\mathbf{d^0}(F) = \mathbf{0}$, $d_2(F) = \mathbf{0}$, $(d_2^2 + d_1)(F) = \mathbf{0}$, $(d_2^3 + d_1 d_2)(F) = \mathbf{0}$. 所以, $\mathbf{\Lambda}$ 是 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组基.

由上面的例子可以看出, 除了定义, 还可以利用封闭性来计算局部对偶空间的一组基, 即根的重结构.

2.2.3 重结构的计算

Macaulay 方法 1916年, 基于矩阵零空间的计算, F.S. Macaulay 提出了一种计算局部对偶空间一组基的算法 [20]. 它的基本思想如下:

假设 $\hat{\mathbf{x}} \in \mathbb{K}^n$ 是多项式系统 $F = \{f_1, \ldots, f_m\}$ 的根, $f_i \in \mathbb{K}[\mathbf{x}]$. 如果 $\Lambda \in \mathfrak{D}_{\hat{\mathbf{x}}}$, $\deg(\Lambda) = t$, 对 $\forall \boldsymbol{\alpha} \in \mathbb{N}^n$, $|\boldsymbol{\alpha}| < t$, 都满足

$$\Lambda\left((\mathbf{x} - \hat{\mathbf{x}})^{\boldsymbol{\alpha}} f_j\right) = 0, \, j = 1, \ldots, m,$$

则 $\Lambda \in \mathcal{D}_{\hat{\mathbf{x}}}^t(I)$. 因此可以逐阶得到 $\mathcal{D}_{\hat{\mathbf{x}}}^t(I)$ 的一组基. 若 $\dim(\mathcal{D}_{\hat{\mathbf{x}}}^t(I)) = \dim(\mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I))$, 对某个 $t \in \mathbb{N}$ 成立, 则 $\mathcal{D}_{\hat{\mathbf{x}}}(I) = \mathcal{D}_{\hat{\mathbf{x}}}^t(I)$. 从这里开始, 将 $\mathbf{d^0}$ 简写为 1.

例 2.8 (例 2.1 的延续). $F = \{x_1^2 x_2 - x_1 x_2^2, x_1 - x_2^2\}$, $\hat{\mathbf{x}} = (0, 0)$. 对 $t = 1$, 构造如下矩阵

$$\begin{array}{c c} & \begin{array}{ccc} 1 & d_1 & d_2 \end{array} \\ \begin{array}{c} f_1 \\ f_2 \end{array} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

矩阵元素为 $\mathbf{d}^{\alpha}\left((\mathbf{x}-\hat{\mathbf{x}})^{\beta} f_j\right)$. 例如, 第二行与第二列的交叉元素, $d_1(f_2)=1$. 通过计算矩阵的零空间, 得到 $\mathcal{D}_{\hat{\mathbf{x}}}^{1}(I)=\mathrm{Span}_{\mathbb{R}}\{1, d_2\}$. 对 $t=2$, 将矩阵扩展为

$$
\begin{array}{cc}
 & \begin{array}{cccccc} 1 & d_1 & d_2 & d_1^2 & d_1 d_2 & d_2^2 \end{array} \\
\begin{array}{c} f_1 \\ f_2 \\ x_1 f_1 \\ x_1 f_2 \\ x_2 f_1 \\ x_2 f_2 \end{array} &
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\end{array}
$$

得到 $\mathcal{D}_{\hat{\mathbf{x}}}^{2}(I)=\mathrm{Span}_{\mathbb{R}}\{1, d_2, d_2^2+d_1\}$. 对 $t=3$, 将矩阵继续扩展为

$$
\begin{array}{cc}
 & \begin{array}{cccccccccc} 1 & d_1 & d_2 & d_1^2 & d_1 d_2 & d_2^2 & d_1^3 & d_1^2 d_2 & d_1 d_2^2 & d_2^3 \end{array} \\
\begin{array}{c} f_1 \\ f_2 \\ x_1 f_1 \\ x_1 f_2 \\ x_2 f_1 \\ x_2 f_2 \\ x_1^2 f_1 \\ x_1^2 f_2 \\ x_1 x_2 f_1 \\ x_1 x_2 f_2 \\ x_2^2 f_1 \\ x_2^2 f_2 \end{array} &
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\end{array}
$$

得到 $\mathcal{D}_{\hat{\mathbf{x}}}^{3}(I)=\mathrm{Span}_{\mathbb{R}}\{1, d_2, d_2^2+d_1, d_2^3+d_1 d_2\}$. 最后发现, 下一阶扩展矩阵的零空间维数仍然为 4. 所以 $\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\{1, d_2, d_2^2+d_1, d_2^3+d_1 d_2\}$.

Macaulay 方法中, 矩阵的最大规模为 $m\binom{\rho+n}{n} \times \binom{\rho+n+1}{n}$.

MMM 方法 90年代, M. Marinari, T. Mora 和 H. Möller 利用局部对偶空间的封闭性提出了一种新算法 [22, 23]. 随后, B. Mourrain 又对算法进行了改进 [27]. 假设 $\{\Lambda_1, \ldots, \Lambda_k\}$ 是 $\mathcal{D}_{\hat{\mathbf{x}}}^{t}(I)$ 的一组基. 由 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的封闭性, $\forall \Lambda \in \mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)$ 满足

$$
\left\{\begin{array}{c}
\Phi_{x_1}(\Lambda)=c_{1,1}\Lambda_1+\cdots+c_{1,k}\Lambda_k, \\
\vdots \\
\Phi_{x_n}(\Lambda)=c_{n,1}\Lambda_1+\cdots+c_{n,k}\Lambda_k.
\end{array}\right. \tag{2.2}
$$

如果这些待定系数确定, 那么就可以得到 $\mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)$ 的一组基. 我们定义一个微分算子 $\Psi_{x_i}: \mathfrak{D}_{\hat{\mathbf{x}}} \rightarrow \mathfrak{D}_{\hat{\mathbf{x}}}$

$$
\Psi_{x_i}(\mathbf{d}_{\hat{\mathbf{x}}}^{\alpha}) := \begin{cases}
\mathbf{d}_{\hat{\mathbf{x}}}^{(\alpha_1,\ldots,\alpha_i+1,\ldots,\alpha_n)}, & \text{如果 } \alpha_1=\cdots=\alpha_{i-1}=0, \\
0, & \text{如若不然.}
\end{cases}
$$

值得注意的是, $\Phi_{x_i}\Psi_{x_i}(\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}) = \mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}$, 但

$$\Psi_{x_i}\Phi_{x_i}(\mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}) := \begin{cases} \mathbf{d}_{\hat{\mathbf{x}}}^{\boldsymbol{\alpha}}, & \text{如果 } \alpha_1 = \cdots = \alpha_{i-1} = 0, \\ 0, & \text{如若不然}. \end{cases}$$

定理 2.6. *[27] 假设 $\hat{\mathbf{x}} \in \mathbb{K}^n$ 是多项式系统 $F = \{f_1, \ldots, f_m\}$ 的根, $f_i \in \mathbb{K}[\mathbf{x}]$, 集合 $\{\Lambda_1, \ldots, \Lambda_k\}$ 是 $\mathcal{D}_{\hat{\mathbf{x}}}^t(I)$ 的一组基, 那么 $\mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)$ 中不含常数项的元素(不含 1)有如下形式*

$$\Lambda = \sum_{j=1}^{k} c_{1,j}\Psi_{x_1}(\Lambda_j) + \sum_{j=1}^{k} c_{2,j}\Psi_{x_2}(\Lambda_j) + \cdots + \sum_{j=1}^{k} c_{n,j}\Psi_{x_n}(\Lambda_j), \tag{2.3}$$

而且满足

(i) $\sum_{j=1}^{k} c_{i,j}\Phi_{x_l}(\Lambda_j) - \sum_{j=1}^{k} c_{l,j}\Phi_{x_i}(\Lambda_j) = 0,\ 1 \le i < l \le n,$

(ii) $\Lambda(F) = \mathbf{0}.$

证明. 令 $\Lambda = \sum_{0<|\boldsymbol{\alpha}|\le t+1} c_{\boldsymbol{\alpha}}\mathbf{d}^{\boldsymbol{\alpha}} \in \mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)$, 定义 Λ 关于 $d_1, \ldots, d_l$ 的截断

$$\Lambda \cap \mathbb{K}[d_1, \ldots, d_l] := \sum_{0<|\boldsymbol{\alpha}|\le t+1} c_{\boldsymbol{\alpha}}\mathbf{d}^{\boldsymbol{\alpha}},\ \mathbf{d}^{\boldsymbol{\alpha}} \in \mathbb{K}[d_1, \ldots, d_l].$$

由(2.2)得到

$$\Lambda \cap \mathbb{K}[d_1] = d_1 \sum_{j=1}^{k} c_{1,j}\Lambda_j = \sum_{j=1}^{k} c_{1,j}\Psi_{x_1}(\Lambda_j),\ \Lambda \cap \mathbb{K}[d_2] = d_2 \sum_{j=1}^{k} c_{2,j}\Lambda_j,$$

它们的相对补集

$$\Lambda \cap \mathbb{K}[d_2] - \Lambda \cap \mathbb{K}[d_1] = \sum_{j=1}^{k} c_{2,j}\Psi_{x_2}(\Lambda_j).$$

所以 $\Lambda \cap \mathbb{K}[d_1, d_2] = (\Lambda \cap \mathbb{K}[d_1]) \cup (\Lambda \cap \mathbb{K}[d_2])$, 满足

$$\Lambda \cap \mathbb{K}[d_1, d_2] = \Lambda \cap \mathbb{K}[d_1] + \sum_{j=1}^{k} c_{2,j}\Psi_{x_2}(\Lambda_j) = \sum_{j=1}^{k} c_{1,j}\Psi_{x_1}(\Lambda_j) + \sum_{j=1}^{k} c_{2,j}\Psi_{x_2}(\Lambda_j).$$

对于 $\Lambda \cap \mathbb{K}[d_1, \ldots, d_l]$ 也采用类似分析, 最后得到 $\Lambda = \Lambda \cap \mathbb{K}[d_1, \ldots, d_n]$ 满足(2.3). 至于条件(i,ii), 它们由封闭性和$\mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)$ 的定义自然满足. □

根据定理 2.6, 当满足条件(i,ii)的线性方程组系数矩阵的零空间维数不变时, 即 $\mathcal{D}_{\hat{\mathbf{x}}}^t(I)=\mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)$, 那么 $\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathcal{D}_{\hat{\mathbf{x}}}^t(I)$. 因此, 可以从 $\mathcal{D}_{\hat{\mathbf{x}}}^0(I)=\mathrm{Span}_{\mathbb{K}}\{1\}$ 出发, 通过定理 2.6 逐阶得到 $\mathcal{D}_{\hat{\mathbf{x}}}^t(I)$ 的一组基, 直至 $\mathcal{D}_{\hat{\mathbf{x}}}^t(I)=\mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)$.

例 2.9 (例 2.1 的延续)**.** $F=\{x_1^2x_2-x_1x_2^2, x_1-x_2^2\}$, $\hat{\mathbf{x}}=(0,0)$. 对 $t=1$, 由(2.3)构造 $\Lambda=c_1d_1+c_2d_2$. 待定系数 c_1, c_2 可以通过计算矩阵

$$\begin{array}{cc} d_1 & d_2 \\ \end{array}$$
$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

的零空间来确定, $\mathcal{D}_{\hat{\mathbf{x}}}^1(I)=\mathrm{Span}_{\mathbb{R}}\{1,d_2\}$. 对 $t=2$, $\Lambda=c_1d_1+c_2d_2+c_3d_1d_2+c_4d_2^2$. 通过计算矩阵

$$\begin{array}{cccc} d_1 & d_2 & d_1d_2 & d_2^2 \end{array}$$
$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

的零空间, 可以得到 $\mathcal{D}_{\hat{\mathbf{x}}}^2(I)=\mathrm{Span}_{\mathbb{R}}\{1,d_2,d_2^2+d_1\}$. 对 $t=3$, $\Lambda=c_1d_1+c_2d_2+c_3d_1d_2+c_4d_2^2+c_5(d_1^2+d_1d_2^2)+c_6d_2^3$. 通过计算矩阵

$$\begin{array}{cccccc} d_1 & d_2 & d_1d_2 & d_2^2 & d_1^2+d_1d_2^2 & d_2^3 \end{array}$$
$$\begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

的零空间, 得到 $\mathcal{D}_{\hat{\mathbf{x}}}^3(I)=\mathrm{Span}_{\mathbb{R}}\{1,d_2,d_2^2+d_1,d_2^3+d_1d_2\}$. 最后发现, 下一阶矩阵的零空间维数仍然为 4. 所以 $\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\{1,d_2,d_2^2+d_1,d_2^3+d_1d_2\}$.

MMM 方法中, 矩阵的最大规模为 $(\frac{1}{2}\mu(n-1)n+m)\times\mu n$.

由上面的例子可以看出, 利用局部对偶空间的封闭性, 我们可以将一些项绑定在一起(比如 $d_1^2+d_1d_2^2$), 同时降低了线性方程的个数(矩阵的行). 下一节中, 利用宽度为 1 的情形下局部对偶空间的特殊性质, 我们给出一种计算重结构的新算法, 从而将矩阵的最大规模降低为 $m\times n$. 并且我们给出这种情形下, 重结构的一种参数化表示.

2.3 宽度为 1 的特殊情形

宽度为 1, 是多项式系统的孤立奇异根中比较特殊的一类情形. 因为在所有的孤立奇异根中, Jacobian 矩阵的列秩亏 1 (宽度为 1)是最为普遍的; 同时这种情形下, 局部对偶空间具有一些类似于单变元多项式的特殊性质.

引理 2.7. 假设 $\hat{\mathbf{x}} \in \mathbb{K}^n$ 是多项式系统 $F = \{f_1, \dots, f_m\}$ 的根, $f_i \in \mathbb{K}[\mathbf{x}]$, 而且满足 $\operatorname{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = n - 1$. 那么

$$\dim(\mathcal{D}_{\hat{\mathbf{x}}}^{t+1}(I)) = \dim(\mathcal{D}_{\hat{\mathbf{x}}}^{t}(I)) + 1,\ t = 0, 1, \dots, \rho - 1,$$

即 $\mu = \rho + 1$.

这个引理的证明可以通过 [37, 定理 2.2] 和 [7, 引理 1] 直接得到. 因此, 对于宽度为 1 的特殊情形, 当我们逐阶计算 $\mathcal{D}_{\hat{\mathbf{x}}}^{t}(I)$ 的一组基时, 每一阶新增的微分泛函至多只有一个, 即存在 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组基 $\{1, \Lambda_1, \dots, \Lambda_{\mu-1}\}$, 满足 $\deg(\Lambda_i) = i$, $i = 1, \dots, \mu - 1$.

下面利用一些正则化和约化技巧, 我们提出了一种计算宽度为 1 的特殊情形下, 局部对偶空间一组既约基的新算法.

2.3.1 假设 $\Lambda_1 = d_1$

令 $\hat{\mathbf{x}} \in \mathbb{K}^n$ 是 $F = \{f_1, \dots, f_m\}$ 的根, $f_i \in \mathbb{K}[\mathbf{x}]$, 满足 $n - \operatorname{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = 1$. 假设 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的某一列为零, 不妨设为第一列, 那么 $\mathcal{D}_{\hat{\mathbf{x}}}^{1}(I) = \operatorname{Span}_{\mathbb{K}}\{1, d_1\}$.

引理 2.8. 假设 $\mathcal{D}_{\hat{\mathbf{x}}}(I) = \operatorname{Span}_{\mathbb{K}}\{1, d_1, \Lambda_2, \dots, \Lambda_{\mu-1}\}$, $\deg(\Lambda_i) = i$, $i = 0, \dots, \mu-1$. 那么 Λ_i 有如下形式

$$\Lambda_i = c_i d_1^i + \sum_{|\boldsymbol{\alpha}|<i} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}},\ c_i \neq 0.$$

证明. 采用归纳法证明. 显然对于 $i = 0, 1$, 引理成立. 假设 Λ_k 有如上形式, 下面证明 Λ_{k+1} 也满足. 由 $\deg(\Lambda_{k+1}) = k + 1$, $\Phi_{x_i}(\Lambda_{k+1}) \in \mathcal{D}_{\hat{\mathbf{x}}}^{k}(I)$, 推出

$$\Lambda_k = \sum_{j=1}^{n} c_{k+1}^{j} d_1^k d_j + \sum_{|\boldsymbol{\alpha}|<k+1} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}},\ 且\ \exists\, l \in \{1, \dots, n\},\ 使得\ c_{k+1}^{l} \neq 0.$$

如果 $l \neq 1$, $\Phi_{x_1}^{k}(\Lambda_{k+1}) \notin \operatorname{Span}_{\mathbb{K}}\{1, d_1\}$, 与封闭性矛盾. 因此, $l = 1$, 引理成立. 这里 $\Phi_{x_i}^{k}(\mathbf{d}^{\boldsymbol{\alpha}}) = \Phi_{x_i} \underbrace{\circ \cdots \circ}_{k-1} \Phi_{x_i}(\mathbf{d}^{\boldsymbol{\alpha}})$. □

将 d_1^i 定义为 Λ_i 的首项, 记为 $\operatorname{lt}(\Lambda_i) = d_1^i$; 将所有的首项系数 c_i 都约化为 1, 记为 $\operatorname{lc}(\Lambda_i) = 1$; 对 $\{1, d_1, \Lambda_2, \dots, \Lambda_{\mu-1}\}$进行约化, 使之满足

$$\operatorname{lt}(\Lambda_i) \notin \operatorname{supp}(\Lambda_j),\ i, j = 0, 1, \dots, \mu - 1,\ i \neq j. \tag{2.4}$$

显然, 这样的一组既约基是存在且唯一的, 而且满足下面的引理.

引理 2.9. 假设 $\Lambda_i = d_1^i + \sum_{1<|\boldsymbol{\alpha}|<i} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}} + a_{i,2}d_2 + \cdots + a_{i,n}d_n$, $i = 2, \ldots, k$. 那么

$$\Phi_{x_1}(\Lambda_{k+1}) = \Lambda_k,\ \Phi_{x_l}(\Lambda_{k+1}) = \sum_{j=0}^{k-1} a_{k+1-j,l}\Lambda_j.$$

证明. 由 $\Phi_{x_i}(\Lambda_{k+1}) \in \mathcal{D}_{\hat{\mathbf{x}}}^k(I)$, 推出

$$\begin{cases} \Phi_{x_1}(\Lambda_{k+1}) = c_{1,0}\Lambda_0 + c_{1,1}\Lambda_1 + \cdots + c_{1,k}\Lambda_k, \\ \Phi_{x_2}(\Lambda_{k+1}) = c_{2,0}\Lambda_0 + c_{2,1}\Lambda_1 + \cdots + c_{2,k}\Lambda_k, \\ \qquad\qquad \vdots \\ \Phi_{x_n}(\Lambda_{k+1}) = c_{n,0}\Lambda_0 + c_{n,1}\Lambda_1 + \cdots + c_{n,k}\Lambda_k. \end{cases} \tag{2.5}$$

根据 $\mathrm{lc}(\Lambda_i) = 1$, 引理 2.8 和(2.4), 得到

$$c_{1,k} = 1,\ c_{l,k} = 0\,(l = 2, \ldots, n),\ c_{1,j} = 0\,(j = 0, \ldots, k-1).$$

将这些系数代入(2.5), 等价于

$$\Phi_{x_1}(\Lambda_{k+1}) = \Lambda_k,\ \Phi_{x_l}(\Lambda_{k+1}) = \sum_{j=0}^{k-1} c_{l,j}\Lambda_j.$$

下面证明 $c_{l,j} = a_{k+1-j,l}$. 利用 $\Phi_{x_1}^j(\Lambda_{k+1}) \in \mathcal{D}_{\hat{\mathbf{x}}}^{k+1-j}(I)$ 和(2.4), 推出

$$\Phi_{x_1}^j(\Lambda_{k+1}) = \Lambda_{k+1-j},\ \Phi_{x_l}\left(\Phi_{x_1}^j(\Lambda_{k+1})\right) = \sum_{0<|\boldsymbol{\alpha}|<k-j} c_{\boldsymbol{\alpha}}\mathbf{d}_{\boldsymbol{\alpha}} + a_{k+1-j,l}\Lambda_0.$$

另一方面, 由(2.5)得到

$$\Phi_{x_1}^j\left(\Phi_{x_l}(\Lambda_{k+1})\right) = \Phi_{x_1}^j\left(\cdots + c_{l,j}\Lambda_j + \cdots\right) = \sum_{0<|\boldsymbol{\alpha}|<k-j} c_{\boldsymbol{\alpha}}\mathbf{d}_{\boldsymbol{\alpha}} + c_{l,j}\Lambda_0.$$

所以 $c_{l,j} = a_{k+1-j,l}$. □

这个引理表明, 除了 $a_{k+1,l}$, (2.5)中所有其它的待定系数都是已确定的, 它们等于 $\mathcal{D}_{\hat{\mathbf{x}}}^k(I)$ 的一组已知既约基 $\{1, d_1, \Lambda_2, \ldots, \Lambda_k\}$ 中一阶项对应的系数. 因此, 计算 $\mathcal{D}_{\hat{\mathbf{x}}}^{k+1}(I)$ 的一组既约基, 所需要的未知参数只是 $a_{k+1,l}$, $l = 2, \ldots, n$. 它们可以通过解线性方程组 $\Lambda_{k+1}(F) = \mathbf{0}$ 唯一确定. 如果方程组无解, 则 $\mathcal{D}_{\hat{\mathbf{x}}}(I) = \mathcal{D}_{\hat{\mathbf{x}}}^k(I)$.

利用下面的定理, 我们可以逐阶得到宽度为 1 的特殊情形下, 局部对偶空间满足(2.4)的一组既约基.

定理 2.10. 假设 $\{1, d_1, \Lambda_2, \ldots, \Lambda_k\}$ 是 $\mathcal{D}_{\hat{\mathbf{x}}}^k(I)$ 的一组既约基，满足

$$\Lambda_i = d_1^i + \sum_{1<|\boldsymbol{\alpha}|<i} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}} + a_{i,2} d_2 + \cdots + a_{i,n} d_n, \ i = 2, \ldots, k.$$

那么，$\{1, d_1, \Lambda_2, \ldots, \Lambda_k, \Lambda_{k+1}\}$ 是 $\mathcal{D}_{\hat{\mathbf{x}}}^{k+1}(I)$ 的一组既约基，其中 Λ_{k+1} 有如下形式

$$\Lambda_{k+1} = \Psi_{x_1}(\Lambda_k) + \sum_{j=0}^{k-1} a_{k+1-j,2} \Psi_{x_2}(\Lambda_j) + \cdots + \sum_{j=0}^{k-1} a_{k+1-j,n} \Psi_{x_n}(\Lambda_j). \tag{2.6}$$

而且满足 $\Lambda_{k+1}(F) = \mathbf{0}$.

证明. 定理可以由引理 2.9 和定理 2.6 直接得到. □

例 2.10 (例 2.1 的延续). $F = \{x_1^2 x_2 - x_1 x_2^2, x_1 - x_2^2\}$, $\hat{\mathbf{x}} = (0, 0)$. 计算 Jacobian 矩阵

$$F_{\mathbf{x}}(\hat{\mathbf{x}}) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

得到 $\mathcal{D}_{\hat{\mathbf{x}}}^1(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_2\}$. 对 $t = 2$, 由(2.6)构造 $\Lambda_2 = d_2^2 + a_{2,1} d_1$. 通过计算矩阵

$$\begin{array}{cc} d_1 & d_2^2 \\ \multicolumn{2}{c}{\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}} \end{array}$$

的零空间，得到 $\mathcal{D}_{\hat{\mathbf{x}}}^2(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_2, d_2^2 + d_1\}$. 对 $t = 3$, 构造 $\Lambda_3 = d_2(d_2^2 + d_1) + a_{3,1} d_1$. 通过计算矩阵

$$\begin{array}{cc} d_1 & d_2^3 + d_1 d_2 \\ \multicolumn{2}{c}{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}} \end{array}$$

的零空间，可以得到 $\mathcal{D}_{\hat{\mathbf{x}}}^3(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_2, d_2^2 + d_1, d_2^3 + d_1 d_2\}$. 对 $t = 4$, 由(2.6)构造 $\Lambda_4 = d_2(d_2^3 + d_1 d_2) + d_1^2 + a_{4,1} d_1$. 矩阵

$$\begin{array}{cc} d_1 & d_2^4 + d_1 d_2^2 + d_1^2 \\ \multicolumn{2}{c}{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}} \end{array}$$

是满秩的，所以 $\mathcal{D}_{\hat{\mathbf{x}}}(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_2, d_2^2 + d_1, d_2^3 + d_1 d_2\}$.

由上面的例子可以看出，为了得到 $\mathcal{D}_{\hat{\mathbf{x}}}^{k+1}(I)$ 中新增的微分泛函 Λ_{k+1}, 只需要计算一个 $m \times n$ 阶矩阵的零空间. 但 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 某一列为零的假设一般情况下并不成立. 当然，可以通过对原系统 F 和它的根 $\hat{\mathbf{x}}$ 进行线性变换，从而使得到的新系统 G 和它的根 $\hat{\mathbf{y}}$ 满足 $G_{\mathbf{y}}(\hat{\mathbf{y}})$ 某一列为零. 但从下面的例子我们会看到，这样做往往会破坏原系统 F 的稀疏性，而且这样得到的 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组基往往会丧失既约性(2.4).

例 2.11. 考虑如下例子

$$F=\left\{x_1^2+x_2-3, x_1+\frac{1}{8}x_2^2-\frac{3}{2}\right\}.$$

$\hat{\mathbf{x}}=(1,2)$ 是 $F(x_1,x_2)=\mathbf{0}$ 的孤立奇异根. 计算 Jacobian 矩阵

$$F_{\mathbf{x}}(\hat{\mathbf{x}})=\begin{pmatrix} 2 & 1 \\ 1 & \frac{1}{2} \end{pmatrix},$$

它的任意一列都不为零. 因此, 对变元进行线性变换

$$x_1=-\frac{1}{2}y_1+2y_2, x_2=y_1+y_2,$$

从而得到新系统

$$G=\left\{\frac{1}{4}y_1^2-2y_1y_2+4y_2^2+y_1+y_2-3, \frac{1}{8}y_1^2+\frac{1}{4}y_1y_2+\frac{1}{8}y_2^2-\frac{1}{2}y_1+2y_2-\frac{3}{2}\right\},$$

和它的根 $\hat{\mathbf{y}}=\left(\frac{6}{5},\frac{4}{5}\right)$. 此时, 新系统的 Jacobian 矩阵

$$G_{\mathbf{y}}(\hat{\mathbf{y}})=\begin{pmatrix} 0 & 5 \\ 0 & \frac{5}{2} \end{pmatrix}$$

的第一列为零, 但 G 也变得十分稠密. 通过定理 2.10, 得到一组既约基

$$\mathcal{D}_{\hat{\mathbf{y}}}(J)=\mathrm{Span}_{\mathbb{R}}\left\{1, d_1, d_1^2-\frac{1}{20}d_2\right\}, \ J=(g_1,g_2).$$

最后通过线性变换得到

$$\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\left\{1,-\frac{1}{2}d_1+d_2, \frac{1}{4}d_1^2-\frac{1}{2}d_1d_2+d_2^2-\frac{1}{10}d_1-\frac{1}{20}d_2\right\}, \ I=(f_1,f_2).$$

但这组基不再满足既约性(2.4).

为了避免线性变换, 我们对定理 2.10 直接进行推广.

2.3.2 一组既约基

令 $\hat{\mathbf{x}}\in\mathbb{K}^n$ 是 $F=\{f_1,\ldots,f_m\}$ 的根, $f_i\in\mathbb{K}[\mathbf{x}]$, 满足 $n-\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}}))=1$. 那么, 必然存在 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的某一列能够被其他列线性表示, 不妨设为第一列. 因此,

$$\mathcal{D}_{\hat{\mathbf{x}}}^1(I)=\mathrm{Span}_{\mathbb{K}}\{1, d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n\}.$$

下面我们在 $\mathbb{K}[d_1,\ldots,d_n]$ 中引入项序 $d_1>\cdots>d_n$, 并对引理 2.8, 2.9 和定理 2.10 进行推广.

引理 2.11. 假设 $\mathcal{D}_{\hat{\mathbf{x}}}(I) = \mathrm{Span}_{\mathbb{K}}\{1, d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n, \Lambda_2, \ldots, \Lambda_{\mu-1}\}$, 满足 $\deg(\Lambda_i) = i$, $i = 0, \ldots, \mu - 1$. 那么 Λ_i 有如下形式

$$\Lambda_i = c_i d_1^i + \sum_{\mathbf{d}_{\boldsymbol{\alpha}} < d_1^i} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}},\ c_i \neq 0.$$

证明. 同样采用归纳法证明. 显然对于 $i = 0, 1$, 引理成立. 假设 Λ_k 有如上形式, 下面证明 Λ_{k+1} 也满足. 由 $\deg(\Lambda_{k+1}) = k + 1$, $\Phi_{x_i}(\Lambda_{k+1}) \in \mathcal{D}^k_{\hat{\mathbf{x}}}(I)$, 推出

$$\exists l \in \{1, \ldots, n\},\ 使得\ \Lambda_k = c_{k+1} d_1^k d_l + \sum_{\mathbf{d}_{\boldsymbol{\alpha}} < d_1^k d_l} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}},\ c_{k+1} \neq 0.$$

如果 $l = 1$, 引理成立. 如果 $l \neq 1$, $\Phi^k_{x_1}(\Lambda_{k+1}) \notin \mathcal{D}^1_{\hat{\mathbf{x}}}(I)$, 与封闭性矛盾. □

所以, 将 d_1^i 定义为 Λ_i 的首项, 记为 $\mathrm{lt}(\Lambda_i) = d_1^i$; 将所有的首项系数 c_i 都约化为 1, 记为 $\mathrm{lc}(\Lambda_i) = 1$; 对 $\{1, d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n, \Lambda_2, \ldots, \Lambda_{\mu-1}\}$ 进行约化, 使之满足(2.4). 显然, 这样的一组既约基也是存在且唯一的, 而且满足下面的引理.

引理 2.12. 假设 $\Lambda_i = d_1^i + \sum_{d_1 < \mathbf{d}^{\boldsymbol{\alpha}} < d_1^i} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}} + a_{i,2}d_2 + \cdots + a_{i,n}d_n$, $i = 1, \ldots, k$. 那么

$$\Phi_{x_1}(\Lambda_{k+1}) = \Lambda_k,\ \Phi_{x_l}(\Lambda_{k+1}) = \sum_{j=0}^{k} a_{k+1-j,l}\Lambda_j.$$

证明. 由 $\Phi_{x_i}(\Lambda_{k+1}) \in \mathcal{D}^k_{\hat{\mathbf{x}}}(I)$ 可知, (2.5)依然成立, 而且证明也是类似的, 只不过

$$c_{l,k} = a_{1,l},\ l = 2, \ldots, n.$$

所以, 求和运算的上界由 $k - 1$ 变为 k. □

定理 2.13. 假设 $\{1, \Lambda_1, \Lambda_2, \ldots, \Lambda_k\}$ 是 $\mathcal{D}^k_{\hat{\mathbf{x}}}(I)$ 的一组既约基, 满足

$$\Lambda_i = d_1^i + \sum_{d_1 < \mathbf{d}^{\boldsymbol{\alpha}} < d_1^i} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}} + a_{i,2}d_2 + \cdots + a_{i,n}d_n,\ i = 1, \ldots, k.$$

那么, $\{1, \Lambda_1, \Lambda_2, \ldots, \Lambda_k, \Lambda_{k+1}\}$ 是 $\mathcal{D}^{k+1}_{\hat{\mathbf{x}}}(I)$ 的一组既约基, 其中 Λ_{k+1} 有如下形式

$$\Lambda_{k+1} = \Psi_{x_1}(\Lambda_k) + \sum_{j=0}^{k} a_{k+1-j,2}\Psi_{x_2}(\Lambda_j) + \cdots + \sum_{j=0}^{k} a_{k+1-j,n}\Psi_{x_n}(\Lambda_j). \tag{2.7}$$

而且满足 $\Lambda_{k+1}(F) = \mathbf{0}$.

证明. 定理可以由引理 2.12 和定理 2.6 直接得到. □

例 2.12 (例 2.11 的延续). $F = \{x_1^2 + x_2 - 3, x_1 + \frac{1}{8}x_2^2 - \frac{3}{2}\}$, $\hat{\mathbf{x}} = (1, 2)$. 通过计算 Jacobian 矩阵的零空间, 得到 $\mathcal{D}_{\hat{\mathbf{x}}}^1(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_1 - 2d_2\}$. 由(2.7)构造

$$\Lambda_2 = d_1(d_1 - 2d_2) + (-2)d_2(-2d_2) + a_{2,2}d_2.$$

通过计算矩阵

$$\begin{array}{c} \begin{array}{cc} d_2 & d_1^2 - 2d_1d_2 + 4d_2^2 \end{array} \\ \begin{bmatrix} 1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{array}$$

的零空间, 得到 $\mathcal{D}_{\hat{\mathbf{x}}}^2(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_1 - 2d_2, d_1^2 - 2d_1d_2 + 4d_2^2 - d_2\}$. 由(2.7)构造

$$\Lambda_3 = d_1(d_1^2 - 2d_1d_2 + 4d_2^2) + (-2)d_2(4d_2^2) + (-1)d_2(d_1 - 2d_2) + a_{3,2}d_2.$$

最后发现, 矩阵

$$\begin{array}{c} \begin{array}{cc} d_2 & d_1^3 - 2d_1^2d_2 + 4d_1d_2^2 - 8d_2^2 - d_1d_2 + 2d_2^2 \end{array} \\ \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{4} \end{bmatrix} \end{array}$$

是满秩的, 所以 $\mathcal{D}_{\hat{\mathbf{x}}}(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_1 - 2d_2, d_1^2 - 2d_1d_2 + 4d_2^2 - d_2\}$.

定理 2.13 更重要的意义在于, 我们给出了宽度为 1 的特殊情形下, 局部对偶空间一组既约基的参数化表示. 对于给定的一个多项式系统 $F = \{f_1, \ldots, f_m\}$, $f_i \in \mathbb{K}[\mathbf{x}]$, 它的一个根 $\hat{\mathbf{x}} \in \mathbb{K}^n$, 满足 $n - \mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = 1$, 和 $\mathbb{K}[d_1, \ldots, d_n]$ 的某个项序 $>$, 我们可以得到 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组既约基 $\mathbf{\Lambda} = \{1, \Lambda_1, \ldots, \Lambda_{\mu-1}\}$. 而且, 这样一组既约基由 $(\mu - 1)(n - 1)$ 个参数 $a_{i,j}$ 和(2.7)唯一确定.

2.3.3 算法及实现

根据定理 2.13, 我们下面给出计算宽度为 1 的特殊情形下, 局部对偶空间的一组既约基的算法. 首先讨论一些算法的细节.

调整变元次序 对于给定的一个多项式系统 $F = \{f_1, \ldots, f_m\}$, $f_i \in \mathbb{K}[\mathbf{x}]$ 和它的一个根 $\hat{\mathbf{x}} \in \mathbb{K}^n$, 满足 $n - \mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = 1$. 我们需要调整变元次序, 使

$$\mathcal{D}_{\hat{\mathbf{x}}}^1(I) = \mathrm{Span}_{\mathbb{K}}\{1, d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n\}$$

成立. 计算 Jacobian 矩阵 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的一个非零零向量 $(v_1,\ldots,v_n)^T$, $\exists t\in\mathbb{N}$, 满足

$$|v_t|\geq |v_i|,\ i=1,\ldots,n.$$

交换 F 和 $\hat{\mathbf{x}}$ 中变元次序 $x_1\leftrightarrow x_t$. 另外,

$$a_{1,2}=\frac{v_2}{v_t},\ldots,a_{1,t}=\frac{v_1}{v_t},\ldots,a_{1,n}=\frac{v_n}{v_t}.$$

记 $\mathbf{a}_1:=(1,a_{1,2},\ldots,a_{1,n})^T$.

矩阵的 LU 分解　根据定理 2.13, 得到 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组既约基 $\{1,\Lambda_1,\ldots,\Lambda_{\mu-1}\}$ 有如下形式

$$\Lambda_i=\Delta_i+a_{i,2}d_2+\cdots+a_{i,n}d_n,\ \Delta_i=\sum_{1<|\boldsymbol{\alpha}|\leq i}c_{\boldsymbol{\alpha}}\mathbf{d}_{\boldsymbol{\alpha}}.$$

当计算每一阶新增的微分泛函时, 需要解线性方程组

$$\begin{pmatrix} d_2(f_1) & \cdots & d_n(f_1)\\ \vdots & & \vdots\\ d_2(f_m) & \cdots & d_n(f_m)\end{pmatrix}\begin{pmatrix} a_{i,2}\\ \vdots\\ a_{i,n}\end{pmatrix}=-\begin{pmatrix}\Delta_i(f_1)\\ \vdots\\ \Delta_i(f_m)\end{pmatrix}.$$

记等式左边的系数矩阵为 M, $\mathbf{a}_i:=(0,a_{i,2},\ldots,a_{i,n})^T$, $i=2,\ldots,\mu-1$. 那么每次解线性方程组时, 更新的只有等式右边的向量, 而系数矩阵 M 保持不变. 因此, 计算 M 的 LU 分解, $M=P\cdot L\cdot U$, 使得当计算每一阶新增的微分泛函时, 只需解两个三角化的线性方程组.

微分泛函的赋值　在实验中我们有时会遇到微分泛函 Δ_i 的形式十分稠密, 从而导致 $\Delta_i(f_j)$ 的赋值非常耗时.

例 2.13. 考虑如下例子 $F=\{f_1,\ldots,f_s\}$

$$\begin{aligned} f_i &= x_i^3+x_i^2-x_{i+1},\ \text{如果}\ i<s,\\ f_s &= x_s^2,\end{aligned}$$

$\hat{\mathbf{x}}=(0,\ldots,0)$ 是 $F(\mathbf{x})=\mathbf{0}$ 重数为 2^s 的孤立奇异根.

对于 $s=6$, 大约需要 $17MB$ 的存储和 2 个小时的时间, 得到 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组既约基; 对于 $s=7$, 在 3 天内无法得到结果; 对于 $s=9$, 估计的存储空间已经大于 $1GB$. 为此, 我们递归地构造了两组多项式系统

$$F_i := G_i + F_{\mathbf{x}} \cdot \mathbf{a}_i,\ G_i := \sum_{j=1}^{i-1} \frac{j}{i} \cdot F_{i-j,\mathbf{x}} \cdot \mathbf{a}_j,$$

$F_1 = F_{\mathbf{x}} \cdot \mathbf{a}_1$, 满足

$$F_i(\hat{\mathbf{x}}) = \Lambda_i(F),\ G_i(\hat{\mathbf{x}}) = \Delta_i(F). \tag{2.8}$$

因此, 我们用这两组多项式系统在 $\hat{\mathbf{x}}$ 处的计值来取代 $\Delta_i(f_j)$ 的赋值, 从而达到降低计算时间和减少存储空间的目的. (2.8)的证明请见附录中命题 5.2.

算法 下面我们就给出计算宽度为 1 的特殊情形下, 局部对偶空间的一组既约基的算法 MSB1 (Multiplicity Structure for Breadth One).

算法：*MSB1*

输入：▸ 一个多项式系统 $F=\{f_1,\ldots,f_m\}$, $f_i \in \mathbb{K}[\mathbf{x}]$ 和它的一个根 $\hat{\mathbf{x}} \in \mathbb{K}^n$.

输出：▸ $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组既约基.

第 1 步 计算 Jacobian 矩阵 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的零空间 $V=\mathrm{null}(F_{\mathbf{x}}(\hat{\mathbf{x}}))$.

- 如果 $\dim(V)=1$, 调整变元次序, 计算 M 的 LU 分解和 $G_2(\hat{\mathbf{x}})$. 令 $i:=2$.
- 否则, 退出.

第 2 步 解两个三角化的线性方程组

- 如果有解 $\mathbf{a}_i$, 则令 $i:=i+1$, 计算 $G_i(\hat{\mathbf{x}})$, 重复第 2 步.
- 否则, 令 $\mu:=i$, 继续.

第 3 步 利用 $\mathbf{a}_i$ 和(2.7)构造 $\mathbf{\Lambda}=\{1,\Lambda_1,\ldots,\Lambda_{\mu-1}\}$. 返回 $\mathbf{\Lambda}$.

另外, 如果给定的多项式系统系数是近似的或给定的根是近似的, 我们用矩阵的奇异值分解来代替算法 MSB1 中 Jacobian 矩阵零空间的计算和线性方程组的求解, 从而给出计算宽度为 1 的特殊情形下, 近似局部对偶空间的一组既约基的算法 AMSB1 (Approximate Multiplicity Structure for Breadth One).

算法：*AMSB1*

输入：▸ 多项式系统 $F=\{f_1,\ldots,f_n\}$, $f_i\in\mathbb{K}[\mathbf{x}]$, 近似根 $\tilde{\mathbf{x}}\in\mathbb{K}^n$ 和容忍度 τ.

输出：▸ 近似局部对偶空间的一组既约基.

第 1 步 计算 Jacobian 矩阵 $F_{\mathbf{x}}(\tilde{\mathbf{x}})$ 的奇异值分解 $F_{\mathbf{x}}(\tilde{\mathbf{x}})=U\cdot\Sigma\cdot V^{\star}$.

- 如果 $\sigma_n<\tau\ll\sigma_{n-1}$, 调整变元次序,

$$\mathbf{a}_1 := \left(1,\frac{v_2}{v_t},\ldots,\frac{v_1}{v_t},\ldots,\frac{v_n}{v_t}\right)^T,$$

  令 $i:=2$, 计算 $G_2(\tilde{\mathbf{x}})$.

- 否则, 退出.

第 2 步 计算矩阵 $(G_i(\tilde{\mathbf{x}}),M)$ 的奇异值分解 $(G_i(\tilde{\mathbf{x}}),M)=U\cdot\Sigma\cdot V^{\star}$.

- 如果 $\sigma_n<\tau$,

$$\mathbf{a}_i := \left(1,\frac{v_2}{v_1},\ldots,\frac{v_n}{v_1}\right)^T,$$

  令 $i:=i+1$, 计算 $G_i(\tilde{\mathbf{x}})$, 重复第 2 步.

- 否则, 令 $\mu:=i$, 继续.

第 3 步 利用 $\mathbf{a}_i$ 和(2.7)构造 $\mathbf{\Lambda}=\{1,\Lambda_1,\ldots,\Lambda_{\mu-1}\}$. 返回 $\mathbf{\Lambda}$.

注 1. 在矩阵的奇异值分解 $U\cdot\Sigma\cdot V^{\star}$ 中, $\Sigma=\mathrm{diag}\{\sigma_1,\ldots,\sigma_n\}$, σ_n 是矩阵的最小奇异值, $(v_1,\ldots,v_n)^T$ 是 V 的最后一列, 即对应于 σ_n 的右奇异向量.

在 Maple 中我们实现了算法 MSB1 和算法 AMSB1, 程序代码和一些实验结果可以在 http://www.mmrc.iss.ac.cn/~lzhi/Research/hybrid/breadthone/ 找到.

实验结果　从下面的例子可以看到, 算法 MSB1 对于解析系统同样适用.

例 2.14. [6, 例 6] 考虑如下例子

$$F=\{x^2\sin(y),y-z^2,z+\sin(x^n)\}.$$

$\hat{\mathbf{x}}=(0,0,0)$ 是 $F(x,y,z)=\mathbf{0}$ 重数为 $2(n+1)$ 的孤立奇异根.

表 2.1: 例子 2.14 的实验结果

n	5	50	100	200	300	400	500
根的重数	12	102	202	402	602	802	1002
时间(秒)	0.056	0.608	2.077	9.596	35.415	105.060	232.490

表 2.2: 例子 2.13 的实验结果

变量个数	6	7	8	9	10
根的重数	64	128	256	512	1024
时间(秒)	0.593	1.377	3.445	10.913	44.659

表 2.2 展示了, 利用算法 MSB1 计算出例 2.13 中所有参数 $\mathbf{a}_i$ 需要的时间.

例 2.15. [7] 考虑如下例子

$$\begin{aligned} F \;=\; & \Big\{14\,x + 33\,y - 3\,\sqrt{5}\left(x^2 + 4\,xy + 4\,y^2 + 2\right) + \sqrt{7} + x^3 + 6\,x^2y \\ & +12\,xy^2 + 8\,y^3, 41\,x - 18\,y - \sqrt{5} + 8\,x^3 - 12\,x^2y + 6\,xy^2 - y^3 \\ & +3\,\sqrt{7}\left(4\,xy - 4\,x^2 - y^2 - 2\right)\Big\}. \end{aligned}$$

$\hat{\mathbf{x}} = (\frac{2\sqrt{7}}{5} + \frac{\sqrt{5}}{5}, -\frac{\sqrt{7}}{5} + \frac{2\sqrt{5}}{5})$ 是 $F(x, y) = \mathbf{0}$ 的 5 重根.

首先需要指出的是, 不同于那些基于 Gröbner 基的算法, 尽管多项式系统系数中含有无理数 $\sqrt{5}$ 和 $\sqrt{7}$, 算法 MSB1 依然可以得到 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组既约基.

$$\Big\{1, d_1 + \frac{1}{3}d_2, d_1^2 + \frac{1}{3}d_1d_2 + \frac{1}{9}d_2^2, d_1^3 + \frac{1}{3}d_1^2d_2 + \frac{1}{9}d_1d_2^2 + \frac{1}{27}d_2^3 - \frac{125}{81}d_2,$$

$$d_1^4 + \frac{1}{3}d_1^3d_2 + \frac{1}{9}d_1^2d_2^2 + \frac{1}{27}d_1d_2^3 + \frac{1}{81}d_2^4 - \frac{125}{81}d_1d_2 - \frac{125}{81}d_2^2.\Big\}$$

下面我们将 F 中的所有多项式系数和 $\hat{\mathbf{x}}$ 都截断到小数点后 5 位, 然后利用

算法 AMSB1 得到近似局部对偶空间的一组既约基

$$
\begin{aligned}
&\{1, d_1 + 0.33341\, d_2, d_1^2 + 0.33343\, d_1 d_2 + 0.11116\, d_2^2,\\
&d_1^3 + 0.33343\, d_1^2 d_2 + 0.11117\, d_1 d_2^2 + 0.03706\, d_2^3 - 1.54321\, d_2,\\
&d_1^4 + 0.33343\, d_1^3 d_2 + 0.11117\, d_1^2 d_2^2 + 0.03706\, d_1 d_2^3 + 0.01235\, d_2^4\\
&-1.54321\, d_1 d_2 - 0.51441\, d_2^2\},
\end{aligned}
$$

满足

$$
\begin{aligned}
\Lambda_0(F) &= (-0.00066377, -0.00039331)^T,\\
\Lambda_1(F) &= (-0.00023342, 0.00023341)^T,\ \Lambda_2(F) = (-0.00000997, 0.00000997)^T,\\
\Lambda_3(F) &= (-0.00060593, 0.00060608)^T,\ \Lambda_4(F) = (0.00080432, -0.00080428)^T.
\end{aligned}
$$

第三章 近似奇异根的精化

经典的求根问题(root-finding)通常由两部分组成: 第一, 全局地对互不相同的根进行隔离(root isolation), 第二, 局部地对每一个近似根分别进行精化(root refinement). 本章主要讨论多项式系统近似根的精化问题:

问题 3.1. 假设 $F(\hat{\mathbf{x}}) = \mathbf{0}$. 对于给定的多项式系统 $F = \{f_1, \ldots, f_n\}$, $f_i \in \mathbb{C}[\mathbf{x}]$, 和 $\hat{\mathbf{x}}$ 附近的近似根 $\tilde{\mathbf{x}} \in \mathbb{C}^n$, 满足 $\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\| = \epsilon \ll 1$, 如何将 $\tilde{\mathbf{x}}$ 精化到更高的精度?

值得注意的是, 本章只考虑方程个数等于变元个数的多项式系统($m = n$). 除非特别声明, $\hat{\mathbf{x}}$ 表示 F 的一个精确根, $\tilde{\mathbf{x}}$ 表示 $\hat{\mathbf{x}}$ 附近的一个近似根.

众所周知, 如果 F 和 $\hat{\mathbf{x}}$ 满足 $\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = n$, Newton 法

$$\tilde{\mathbf{y}} = -F_{\mathbf{x}}(\tilde{\mathbf{x}})^{-1}F(\tilde{\mathbf{x}}), \tag{3.1}$$

当初始近似 $\tilde{\mathbf{x}}$ 足够接近精确根 $\hat{\mathbf{x}}$ 时, 可以达到二次收敛 $\|\tilde{\mathbf{x}} + \tilde{\mathbf{y}} - \hat{\mathbf{x}}\| = \mathcal{O}(\epsilon^2)$ [12]. 然而当 $\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) < n$ 时, Newton 法的收敛速度经常是线性的, 而且有时甚至不收敛 [11]. 我们看下面的例子:

例 3.1 (例 2.11 的延续)**.** 多项式系统 $F = \{x_1^2 + x_2 - 3, x_1 + 0.125x_2^2 - 1.5\}$ 有一个孤立奇异根 $\hat{\mathbf{x}} = (1, 2)$. 对近似根 $\tilde{\mathbf{x}} = (1.01, 2.01)$ 采用 Newton 法(3.1):

- 10 次迭代后, $\tilde{\mathbf{x}} \approx (1.00007008380087, 1.99985982936138)$;
- 100 次迭代后, $\tilde{\mathbf{x}} \approx (1.00000033405564, 1.99999933188861)$;
- 1000 次迭代后, $\tilde{\mathbf{x}} \approx (1.00000033405564, 1.99999933188861)$.

本章中, 我们介绍了处理奇异根的经典技术 deflation, 讨论了它对于孤立奇异根的有限终止性, 基于宽度为 1 的特殊情形下重结构的参数化表示, 提出了一种正则化的 Newton 法, 并且证明了算法在宽度为 1 的孤立奇异根附近是二次收敛的.

3.1 Deflation 技术

80年代, T. Ojika 等人提出了一种处理奇异根的 deflation 方法 [28, 29]. 随后, A. Leykin 等人又对方法进行了改进 [17]. 它的基本思想是: 通过添加一些新方程, 使原系统的一个孤立奇异根变成新系统的一个正则根或者正则根的一部分. 下面我们列举两种添加新方程的方法: 子式法和零向量法.

假设多项式系统 $F = \{f_1, \ldots, f_n\}$, $f_i \in \mathbb{C}[\mathbf{x}]$ 和它的根 $\hat{\mathbf{x}} \in \mathbb{C}^n$, 满足

$$r = \mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) < n.$$

记 $F_{\mathbf{x}}^{r+1}$ 是 Jacobian 矩阵 $F_{\mathbf{x}}$ 所有 $r+1$ 阶子式所构成的集合. 那么对 $\forall\, A \in F_{\mathbf{x}}^{r+1}$, 由 $r = \mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}}))$, 得到 $\det(A(\hat{\mathbf{x}})) = 0$. 因此, $\hat{\mathbf{x}}$ 是增广系统

$$G(\mathbf{x}) = \begin{cases} F, \\ \det(A), \forall\, A \in F_{\mathbf{x}}^{r+1}, \end{cases} \tag{3.2}$$

的根.

例 3.2 (例 2.11 的延续). $F = \{x_1^2 + x_2 - 3, x_1 + 0.125x_2^2 - 1.5\}$ 有一个孤立奇异根 $\hat{\mathbf{x}} = (1, 2)$. 计算 Jacobian 矩阵

$$F_{\mathbf{x}} = \begin{pmatrix} 2x_1 & 1 \\ 1 & 0.25x_2 \end{pmatrix},$$

所以 $G = \{f_1, f_2, 0.5x_1x_2 - 1\}$. 此时 G 的 Jacobian 矩阵

$$G_{\mathbf{x}} = \begin{pmatrix} 2x_1 & 1 \\ 1 & 0.25x_2 \\ 0.5x_2 & 0.5x_1 \end{pmatrix}, \text{ 且 } G_{\mathbf{x}}(\hat{\mathbf{x}}) = \begin{pmatrix} 2 & 1 \\ 1 & 0.5 \\ 1 & 0.5 \end{pmatrix}$$

仍然是列秩亏的, 所以继续添加 $G_{\mathbf{x}}$ 的二阶子式作为新方程, 得到 $H = \{g_1, g_2, g_3, x_1^2 - 0.5x_2, 0.5x_1 - 0.125x_2^2\}$. 此时 H 的 Jacobian 矩阵

$$H_{\mathbf{x}}(\hat{\mathbf{x}}) = \begin{pmatrix} 2 & 1 \\ 1 & 0.5 \\ 1 & 0.5 \\ 2 & -0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

是列满秩的, 因此 $\hat{\mathbf{x}} = (1, 2)$ 现在是增广系统 $H(x_1, x_2) = \mathbf{0}$ 的正则根.

除了 Jacobian 矩阵的 $r+1$ 阶子式, 还可以利用 $F_{\mathbf{x}}$ 的非零零向量来添加新方程. 存在唯一的 $\hat{\boldsymbol{\lambda}} \in \mathbb{C}^{r+1}$, 使得 $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})$ 是增广系统

$$G(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} F, \\ F_{\mathbf{x}} B \boldsymbol{\lambda}, \\ \mathbf{h}^{\star} \boldsymbol{\lambda} - 1, \end{cases} \tag{3.3}$$

的根, 其中随机矩阵 $B \in \mathbb{C}^{n \times (r+1)}$ 和随机向量 $\mathbf{h} \in \mathbb{C}^{r+1}$ 中的元素来自复单位圆, 概率为 1 地保证了 $\hat{\boldsymbol{\lambda}}$ 的唯一性($\operatorname{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = r \Leftrightarrow \operatorname{corank}(F_{\mathbf{x}}(\hat{\mathbf{x}})B) = 1$).

例 3.3 (例 2.11 的延续)**.** 根据 F 的 Jacobian 矩阵, 可以得到 $G = \{f_1, f_2, 2x_1x_3 + x_4, x_3 + 0.25x_2x_4, x_3 - 1\}$. 此时 G 的 Jacobian 矩阵

$$G_{\mathbf{x}} = \begin{pmatrix} 2x_1 & 1 & 0 & 0 \\ 1 & 0.25x_2 & 0 & 0 \\ 2x_3 & 0 & 2x_1 & 1 \\ 0 & 0.25x_4 & 1 & 0.25x_2 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \text{ 且 } G_{\mathbf{x}}(\hat{\mathbf{y}}) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 0.5 & 0 & 0 \\ 2 & 0 & 2 & 1 \\ 0 & -0.5 & 1 & 0.5 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

仍然是列秩亏的, 其中 $\hat{\mathbf{y}} = (1, 2, 1, -2)$, 所以我们继续添加 $G_{\mathbf{x}}$ 的非零零向量, 得到 $H = \{g_1, \ldots, g_5, 2x_1x_5 + x_6, x_5 + 0.25x_2x_6, 2x_3x_5 + 2x_1x_7 + x_8, 0.25x_4x_6 + x_7 + 0.25x_2x_8, x_7, x_5 - 1\}$. 此时 H 的 Jacobian 矩阵

$$H_{\mathbf{x}}(\hat{\mathbf{z}}) = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -0.5 & 1 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & -0.5 & 0 & 0 & 1 & 0.5 & 0 & 0 \\ 0 & 0 & 2 & 0 & 2 & 0 & 2 & 1 \\ 0 & -0.5 & 0 & -0.5 & 0 & -0.5 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

是列满秩的, 其中 $\hat{\mathbf{z}} = (1, 2, 1, -2, 1, -2, 0, -2)$, 因此原系统 $F(x_1, x_2) = \mathbf{0}$ 的孤立奇异根 $\hat{\mathbf{x}} = (1, 2)$ 现在是增广系统 $H(x_1, \ldots, x_8) = \mathbf{0}$ 正则根 $\hat{\mathbf{z}}$ 的一部分.

称(3.2)或(3.3)为 F 的一次 deflation. 如果 $\hat{\mathbf{x}}$ 或 $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})$, 仍然是(3.2)或(3.3)的奇异根, 那么我们对增广系统(3.2)或(3.3)再做一次 deflation. 如果 $\hat{\mathbf{x}}$ 是 F 的一个孤立奇异根, 那么在重复有限次 deflation 后, 我们总可以得到一个增广系统和它的一个正则根, 即 deflation 方法是有限终止的.

3.1.1 有限终止性

A. Leykin 等人证明了, 一次 deflation (3.3) 会使 $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})$ 的重数严格小于 $\hat{\mathbf{x}}$ 的重数 [17, 定理 3.1]. 随后, B. Dayton 和 Z. Zeng 进一步证明了, 一次 deflation 后严格下降的不只是根的重数, 还有局部对偶空间的深度.

定理 3.1. *[7] 假设 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的深度 $0 < \rho_I < \infty$, $I = (f_1, \ldots, f_n)$, 那么 $\mathcal{D}_{(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})}(J)$ 的深度 $\rho_J < \rho_I$. 其中 $J = (g_1, \ldots, g_{2n+1})$ 是由(3.3)定义的.*

证明. 令 $\Lambda_J \in \mathcal{D}_{(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})}(J)$, $\deg(\Lambda_J) = k$, 有如下形式

$$\Lambda_J = \Lambda_I + \sum_{j=1, \Lambda_j \in R_1}^{r+1} \Lambda_j d_{n+j} + \sum_{\mathbf{d}_{\boldsymbol{\alpha}} \in R_1, \mathbf{d}_{\boldsymbol{\beta}} \in R_2, |\boldsymbol{\beta}| > 1} c_{\boldsymbol{\alpha}+\boldsymbol{\beta}} \mathbf{d}_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\beta}},$$

$R_1 = \mathbb{C}[d_1, \ldots, d_n]$, $R_2 = \mathbb{C}[d_{n+1}, \ldots, d_{n+r+1}]$, $\Lambda_I = \sum_{\mathbf{d}_{\boldsymbol{\alpha}} \in R_1} c_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\alpha}}$.

首先证明 $\Lambda_I \in \mathcal{D}_{\hat{\mathbf{x}}}(I)$ 且 $\deg(\Lambda_I) = k$.

$$\begin{aligned} \Lambda_J \in \mathcal{D}_{(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})}(J) \ &\Rightarrow \ \Lambda_J((\mathbf{x} - \hat{\mathbf{x}})^{\boldsymbol{\alpha}} g_i) = 0 \Rightarrow \Lambda_J((\mathbf{x} - \hat{\mathbf{x}})^{\boldsymbol{\alpha}} f_i) = 0 \\ &\Rightarrow \ \Lambda_I((\mathbf{x} - \hat{\mathbf{x}})^{\boldsymbol{\alpha}} f_i) = 0 \Rightarrow \Lambda_I \in \mathcal{D}_{\hat{\mathbf{x}}}(I), \end{aligned}$$

$i = 1, \ldots, n$. 当 $\deg(\Lambda_J) = 1$ 时, 如果 $\deg(\Lambda_I) = 0$, 那么

$$\Lambda_J = c_{n+1} d_{n+1} + \cdots + c_{n+r+1} d_{n+r+1},$$

满足 $\Lambda_J(G) = \mathbf{0}$, 即满足

$$\begin{pmatrix} F_{\mathbf{x}}(\hat{\mathbf{x}}) B \\ \mathbf{h}^{\star} \end{pmatrix} \begin{pmatrix} c_{n+1} \\ \vdots \\ c_{n+r+1} \end{pmatrix} = \mathbf{0}.$$

由 $\operatorname{corank}(F_{\mathbf{x}}(\hat{\mathbf{x}}) B) = 1$ 和 $\mathbf{h}$ 中元素的随机性, 推出等式左边的矩阵满秩. 因此

$$c_{n+1} = \cdots = c_{n+r+1} = 0,$$

与 $\deg(\Lambda_J) = 1$ 矛盾. 所以 $\deg(\Lambda_I) = 1$. 运用引理 2.8 和 2.11 中类似的技巧, 可以归纳地证明出 $\deg(\Lambda_I) = \deg(\Lambda_J)$. 由封闭性得到

$$\Phi_{\lambda_j}(\Lambda_J) = \Lambda_j + \sum_{\mathbf{d}_{\boldsymbol{\alpha}} \in R_1, \mathbf{d}_{\boldsymbol{\beta}} \in R_2, |\boldsymbol{\beta}|>0} c_{\boldsymbol{\alpha}+\boldsymbol{\beta}} \mathbf{d}_{\boldsymbol{\alpha}} \mathbf{d}_{\boldsymbol{\beta}} \in \mathcal{D}_{(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})}(J),$$

$j = 1, \ldots, r+1$, 所以 $\Lambda_j \in \mathcal{D}_{\hat{\mathbf{x}}}(I)$.

假设 $\Lambda_J \in \mathcal{D}_{(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})}(J)$, $\deg(\Lambda_J) = \rho_I$, 则 $\deg(\Lambda_I) = \rho_I$, 且

$$\begin{aligned}
\Lambda_J(F_{\mathbf{x}} B \boldsymbol{\lambda}) = \mathbf{0} \quad &\Leftrightarrow \quad \Lambda_I(F_{\mathbf{x}} B \boldsymbol{\lambda}) + \sum_{j=1}^{r+1} \Lambda_j d_{n+j}(F_{\mathbf{x}} B \boldsymbol{\lambda}) = \mathbf{0} \\
&\Leftrightarrow \quad \Lambda_I\Big(\sum_{i=1}^{n} p_i(\boldsymbol{\lambda}) \frac{\partial F}{\partial x_i}\Big) + \sum_{j=1}^{r+1} \Lambda_j d_{n+j}\Big(\sum_{i=1}^{n} p_i(\boldsymbol{\lambda}) \frac{\partial F}{\partial x_i}\Big) = \mathbf{0} \\
&\Leftrightarrow \quad \Big(\sum_{i=1}^{n} p_i(\hat{\boldsymbol{\lambda}}) \Lambda_I \frac{\partial}{\partial x_i}\Big)(F) + \Big(\sum_{j=1}^{r+1} \sum_{i=1}^{n} b_{i,j} \Lambda_j \frac{\partial}{\partial x_i}\Big)(F) = \mathbf{0} \\
&\Leftrightarrow \quad \Lambda(F) = \mathbf{0},
\end{aligned}$$

$p_i(\boldsymbol{\lambda}) = b_{i,1}\lambda_1 + \cdots + b_{i,r+1}\lambda_{r+1}$, $b_{i,j}$ 是随机矩阵 B 中第 i 行第 j 列的元素.

最后通过归纳法可以证明

$$\Lambda = \sum_{i=1}^{n} p_i(\hat{\boldsymbol{\lambda}}) \Lambda_I \frac{\partial}{\partial x_i} + \sum_{j=1}^{r+1} \sum_{i=1}^{n} b_{i,j} \Lambda_j \frac{\partial}{\partial x_i} \in \mathcal{D}_{\hat{\mathbf{x}}}(I),$$

且 $\deg(\Lambda) = \rho_I + 1$, 这与 $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的深度为 ρ_I 矛盾. 所以 $\rho_J < \rho_I$. □

根据定理 3.1, 如果 $\hat{\mathbf{x}}$ 是 F 的一个孤立奇异根, 那么重复至多 ρ_I 次 deflation 后可以得到一个增广系统和它的一个正则根. 而且对于宽度为 1 的特殊情形, B. Dayton 和 Z. Zeng 提出了一个猜想: $\rho_J = \rho_I - 1$. 接下来利用第二章中给出的, 宽度为 1 的特殊情形下, 局部对偶空间一组既约基的参数化表示, 我们来证明这个猜想.

3.1.2 一个猜想

猜想 3.1. [7] 假设 $\hat{\mathbf{x}} \in \mathbb{C}^n$ 是多项式系统 $F = \{f_1, \ldots, f_n\}$ 的根, $f_i \in \mathbb{C}[\mathbf{x}]$, 而且满足 $\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = n - 1$, $\rho_I < \infty$, $I = (f_1, \ldots, f_n)$. 那么, 存在唯一的 $\hat{\boldsymbol{\lambda}} \in \mathbb{C}^n$,

使得 $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})$ 是增广系统

$$G(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} F, \\ F_{\mathbf{x}}\boldsymbol{\lambda}, \\ \mathbf{h}^{\star}\boldsymbol{\lambda} - 1, \end{cases}$$

的根, 并且满足 $\rho_J = \rho_I - 1$. 随机向量 $\mathbf{h} \in \mathbb{C}^n$, $J = (g_1, \ldots, g_{2n+1})$.

事实上, 由 $\operatorname{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = n - 1$, 必然存在 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的某一列能够被其他列线性表示, 不妨设为第一列. 否则我们可以通过调整变元的次序, 使 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的第一列可以被其他列线性表示.

定理 3.2. 假设 $\operatorname{rank}(M) = n - 1$, M 是 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 去掉第一列剩下的子矩阵. 那么存在唯一的 $\hat{\boldsymbol{\lambda}} \in \mathbb{C}^n$, 使得 $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})$ 是增广系统

$$G(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} F, \\ F_{\mathbf{x}}\boldsymbol{\lambda}, \\ \lambda_1 - 1, \end{cases}$$

的根. 而且如果 $\rho_I = 1$, $\operatorname{rank}(G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})) = 2n$; 如果 $\rho_I > 1$, $\operatorname{rank}(G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})) = 2n - 1$, $\rho_J = \rho_I - 1$. $J = (g_1, \ldots, g_{2n+1})$, $G = \{g_1, \ldots, g_{2n+1}\}$.

证明. 令 $d_1 > \cdots > d_n$, 根据定理 2.13, $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 存在唯一的一组既约基

$$\mathcal{D}_{\hat{\mathbf{x}}}(I) = \operatorname{Span}_{\mathbb{C}}\{1, d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n, \Lambda_2, \ldots, \Lambda_{\rho_I}\}$$

满足 $\deg(\Lambda_i) = i$, $i = 2, \ldots, \mu - 1$. 由 $F_{\mathbf{x}}(\hat{\mathbf{x}})\hat{\boldsymbol{\lambda}} = \mathbf{0}$, 显然

$$\hat{\lambda}_1 = 1, \hat{\lambda}_2 = a_{1,2}, \ldots, \hat{\lambda}_n = a_{1,n}.$$

计算 G 的 Jacobian 矩阵

$$G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}) = \begin{pmatrix} \frac{\partial F(\hat{\mathbf{x}})}{\partial x_1} \quad M & \mathbf{0} \\ F_{\mathbf{xx}}(\hat{\mathbf{x}})\hat{\boldsymbol{\lambda}} & \frac{\partial F(\hat{\mathbf{x}})}{\partial x_1} \quad M \\ \mathbf{0} & 1 \quad \mathbf{0} \end{pmatrix},$$

其中

$$F_{\mathbf{xx}}(\hat{\mathbf{x}})\hat{\boldsymbol{\lambda}} = \left(\sum_{i=1}^{n} \hat{\lambda}_i \frac{\partial^2 F(\hat{\mathbf{x}})}{\partial x_1 x_i}, \ldots, \sum_{i=1}^{n} \hat{\lambda}_i \frac{\partial^2 F(\hat{\mathbf{x}})}{\partial x_n x_i} \right).$$

显然 $\mathrm{rank}(G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}}))\geq 2n-1$. 如果 $\rho_I=1$, 假设 $\mathbf{v}=(v_1,\ldots,v_{2n})^T$ 是 $G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})$ 的一个非零零向量. 那么 $v_1\neq 0$, $v_1d_1+\cdots+v_nd_n\in\mathcal{D}_{\hat{\mathbf{x}}}(I)$, 且

$$\sum_{i=1}^{n}v_i^2d_i^2+\sum_{i\neq j}v_iv_jd_id_j\in\mathcal{D}_{\hat{\mathbf{x}}}(I),$$

与 $\rho_I=1$ 矛盾. 所以, $G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})$ 不存在非零零向量, 即 $\mathrm{rank}(G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}}))=2n$. 如果 $\rho_I>1$, $\mathbf{v}=(1,a_{1,2},\ldots,a_{1,n},1,2a_{2,2},\ldots,2a_{2,n})^T$ 是 $G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})$ 的一个非零零向量, 所以 $\mathrm{rank}(G_{\mathbf{x},\boldsymbol{\lambda}}(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}}))=2n-1$.

下面证明 $\rho_J=\rho_I-1$. 如果 $\rho_I>1$, 由

$$\mathcal{D}^2_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{C}}\{1,d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n,\Delta_2+a_{2,2}d_2+\cdots+a_{2,n}d_n\}$$

可以推出

$$\mathcal{D}^1_{(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})}(J)=\mathrm{Span}_{\mathbb{C}}\{1,d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n+2a_{2,2}d_{n+2}+\cdots+2a_{2,n}d_{2n}\}.$$

事实上, 可以归纳地证明出

$$\mathcal{D}^k_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{C}}\{1,d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n,\ldots,\Delta_k+a_{k,2}d_2+\cdots+a_{k,n}d_n\}$$

可以诱导出

$$\begin{aligned}\mathcal{D}^{k-1}_{(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})}(J) &= \mathrm{Span}_{\mathbb{C}}\{1,d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n+2a_{2,2}d_{n+2}+\cdots+2a_{2,n}d_{2n},\\ &\quad \ldots,\Delta'_{k-1}+ka_{k,2}d_{n+2}+\cdots+ka_{k,n}d_{2n}\},\end{aligned}$$

Δ'_i 的表示形式和具体证明请见附录中的命题 5.3. 因此我们得到 $\rho_J\geq\rho_I-1$, 再根据定理 3.1, 推出 $\rho_J=\rho_I-1$. □

根据定理 3.2, 如果 $\hat{\mathbf{x}}$ 是 F 的一个宽度为 1 的孤立奇异根, 直至 ρ_I 次 deflation 后我们才可以得到一个增广系统和它的一个正则根. 值得注意的是, (3.3)定义的一次 deflation 会使新系统的变元个数和方程个数都翻倍, 最坏的情形(例如宽度为 1)会使最终增广系统的规模达到

$$(2^{\rho_I}n)\times(2^{\rho_I}n).$$

事实上, 对于宽度为 1 的特殊情形, 还可以添加那些满足(2.8)的方程作为一次 deflation [7], 重复 ρ_I 步后得到增广系统

$$G(\mathbf{x}, a_{1,2}, \ldots, a_{1,n}, \ldots, a_{\rho_I,2}, \ldots, a_{\rho_I,n}) = \begin{cases} F, \\ F_1 = F_{\mathbf{x}} \cdot \mathbf{a}_1, \\ \quad\vdots \\ F_{\rho_I} = \sum_{j=1}^{\rho_I - 1} \frac{j}{\rho_I} \cdot F_{\rho_I - j, \mathbf{x}} \cdot \mathbf{a}_j + F_{\mathbf{x}} \cdot \mathbf{a}_{\rho_I}, \end{cases}$$

$\mathbf{a}_1 = (1, a_{1,2}, \ldots, a_{1,n})^T$, $\mathbf{a}_i = (0, a_{i,2}, \ldots, a_{i,n})^T$, $i = 2, \ldots, \rho_I$, 和它的一个正则根. 从而将系统的规模降低为

$$(n + \rho_I n) \times (n + \rho_I n) = (\mu n) \times (\mu n).$$

第四章中, 我们会利用上述增广系统来验证宽度为 1 的孤立奇异根.

例 3.4 (例 2.11 的延续). 由定理 2.13, $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 存在唯一的一组既约基

$$\mathcal{D}_{\hat{\mathbf{x}}}(I) = \{1, d_1 + x_3 d_2, d_1^2 + x_3 d_1 d_2 + x_3^2 d_2^2 + x_4 d_2\},$$

$I = (x_1^2 + x_2 - 3, x_1 + 0.125 x_2^2 - 1.5)$. 因此我们得到增广系统

$$G(x_1, x_2, x_3, x_4) = \{f_1, f_2, 2x_1 + x_3, 0.25 x_2 x_3 + 1, x_4 + 1, 0.125 x_3^2 + 0.25 x_2 x_4 + 1\}$$

和它的一个正则根 $\hat{\mathbf{y}} = (1, 2, -2, -1)$.

下一节中, 对于宽度为 1 的孤立奇异根附近近似根的数值精化问题, 我们提出了一种正则化的 Newton 法: 解一个正则化的最小二乘问题作为初始近似根的预处理, 从而得到一个更好的近似根和一个能够达到二次收敛的迭代方向; 通过计算近似局部对偶空间的一组既约基和解一个线性方程组, 从而确定合适的迭代步长. 算法中涉及的矩阵的最大规模仅为 $n \times n$, 并且我们证明了算法在宽度为 1 的孤立奇异根附近是二次收敛的.

3.2 正则化的 Newton 法

本节主要考虑如下问题:

问题 3.2. 假设 $F(\hat{\mathbf{x}}) = \mathbf{0}$ 且 $\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = n - 1$. 对于给定的多项式系统 $F = \{f_1, \ldots, f_n\}$, $f_i \in \mathbb{C}[\mathbf{x}]$ 和 $\hat{\mathbf{x}}$ 附近的近似根 $\tilde{\mathbf{x}} \in \mathbb{C}^n$, 满足 $\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\| = \epsilon \ll 1$, 如何将 $\tilde{\mathbf{x}}$ 精化到更高的精度?

记 $A = F_{\mathbf{x}}(\tilde{\mathbf{x}})$, $\mathbf{b} = -F(\tilde{\mathbf{x}})$, $\sigma_1 \geq \ldots \geq \sigma_n$ 是 A 的奇异值, 那么

$$\|\mathbf{b}\| = \mathcal{O}(\epsilon),\ \sigma_n = \mathcal{O}(\epsilon),\ \sigma_i = \Theta(1),\ i = 1, \ldots, n-1. \tag{3.4}$$

注 2. $\mathcal{O}(\epsilon)$ 表示 ϵ 的高阶无穷小, $\Theta(\epsilon)$ 表示 ϵ 的同阶无穷小.

3.2.1　正则化的最小二乘问题

对于宽度为 1 的近似奇异根 $\tilde{\mathbf{x}}$, 我们解一个带阻尼的最小二乘问题(damped least-squares problem)来代替 Newton 法(3.1):

$$\min\ \|A\mathbf{y} - \mathbf{b}\|^2 + \lambda\|\mathbf{y}\|^2,$$

实数 λ 被称为正则化参数(regularization parameter) [40].

定理 3.3. 如果选择 A 的最小奇异值 σ_n 作为正则化参数, 正则化最小二乘问题

$$(A^\star A + \sigma_n I_n)\mathbf{y} = A^\star \mathbf{b} \tag{3.5}$$

的解 $\tilde{\mathbf{y}}$ 满足

$$\|\tilde{\mathbf{y}}\| = \mathcal{O}(\epsilon),\ \|F(\tilde{\mathbf{x}} + \tilde{\mathbf{y}})\| = \mathcal{O}(\epsilon^2).$$

证明. 令 $A = U \cdot \Sigma \cdot V^\star$ 是 A 的奇异值分解, $\Sigma = \mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$, 那么(3.5)的解

$$\tilde{\mathbf{y}} = V \cdot (\Sigma^2 + \sigma_n I_n)^{-1} \cdot \Sigma \cdot U^\star \mathbf{b}.$$

由(3.4)和酉变换保持2-范数不变, 推出

$$\|\tilde{\mathbf{y}}\|^2 = \sum_{i=1}^{n} \left(\frac{\sigma_i |b_i'|}{\sigma_i^2 + \sigma_n} \right)^2 = \mathcal{O}(\epsilon^2),$$

$\mathbf{b}' = (b_1', \ldots, b_n')^T = U^\star \mathbf{b}$, $\|\mathbf{b}'\| = \|\mathbf{b}\| = \mathcal{O}(\epsilon)$. 因此 $\|\tilde{\mathbf{y}}\| = \mathcal{O}(\epsilon)$.

F 在 $\tilde{\mathbf{x}}$ 处的 Taylor 展开满足

$$F(\hat{\mathbf{x}}) = -\mathbf{b} + A(\hat{\mathbf{x}} - \tilde{\mathbf{x}}) + \mathcal{O}(\epsilon^2).$$

所以

$$\| -\mathbf{b} + A(\hat{\mathbf{x}} - \tilde{\mathbf{x}})\| = \mathcal{O}(\epsilon^2).$$

再由酉变换保持2-范数不变得到

$$\| - U^{\star}\mathbf{b} + \Sigma \cdot V^{\star}(\hat{\mathbf{x}} - \tilde{\mathbf{x}})\| = \mathcal{O}(\epsilon^2).$$

因为 $\sigma_n = \mathcal{O}(\epsilon)$ 和 $\|V^{\star}(\hat{\mathbf{x}} - \tilde{\mathbf{x}})\| = \|\hat{\mathbf{x}} - \tilde{\mathbf{x}}\| = \epsilon$, 所以

$$|b_n'| = \mathcal{O}(\epsilon^2). \tag{3.6}$$

另一方面,

$$A\tilde{\mathbf{y}} - \mathbf{b} = U \cdot \mathrm{diag}\left\{\frac{-\sigma_n}{\sigma_1^2 + \sigma_n}, \ldots, \frac{-\sigma_n}{\sigma_n^2 + \sigma_n}\right\} \cdot \mathbf{b}',$$

因此

$$\|A\tilde{\mathbf{y}} - \mathbf{b}\|^2 = \sum_{i=1}^{n}\left(\frac{\sigma_n|b_i'|}{\sigma_i^2 + \sigma_n}\right)^2,$$

$$\frac{\sigma_n}{\sigma_n^2 + \sigma_n} = \Theta(1), \ \frac{\sigma_n}{\sigma_i^2 + \sigma_n} = \mathcal{O}(\epsilon), \ i = 1, \ldots, n-1,$$

由(3.6)推出

$$\|A\tilde{\mathbf{y}} - \mathbf{b}\| = \mathcal{O}(\epsilon^2). \tag{3.7}$$

最后由 F 在 $\tilde{\mathbf{x}}$ 处的 Taylor 展开, 我们得到

$$\|F(\tilde{\mathbf{x}} + \tilde{\mathbf{y}})\| \leq \| - \mathbf{b} + A\tilde{\mathbf{y}}\| + \mathcal{O}(\epsilon^2) = \mathcal{O}(\epsilon^2).$$

□

根据定理3.3, 通过解一个正则化的最小二乘问题(3.5), 可以得到一个“更好的”近似根 $\tilde{\mathbf{x}} + \tilde{\mathbf{y}}$, 满足

$$\|\tilde{\mathbf{x}} + \tilde{\mathbf{y}} - \hat{\mathbf{x}}\| \leq \|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\| + \|\tilde{\mathbf{y}}\| = \epsilon + \mathcal{O}(\epsilon).$$

如果 $\|\tilde{\mathbf{x}} + \tilde{\mathbf{y}} - \hat{\mathbf{x}}\| = \mathcal{O}(\epsilon^2)$, 则已经达到二次收敛, 但这通常是不可能的. 所以大多数情况下, 我们得到

$$\|\tilde{\mathbf{x}} + \tilde{\mathbf{y}} - \hat{\mathbf{x}}\| = \Theta(\epsilon).$$

虽然如此, 下面我们证明 $F_{\mathbf{x}}(\tilde{\mathbf{x}} + \tilde{\mathbf{y}})$ 对应于 σ_n 的右奇异向量 $\mathbf{v}_n$ 是一个能够达到二次收敛的迭代方向.

定理 3.4. 假设 $A = U{\cdot}\Sigma{\cdot}V^\star$ 是 $A = F_{\mathbf{x}}(\tilde{\mathbf{x}}+\tilde{\mathbf{y}})$ 的奇异值分解, $\Sigma = \operatorname{diag}\{\sigma_1, \ldots, \sigma_n\}$, $U = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$, $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. 那么

$$|\mathbf{v}_n^\star(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}})| = \Theta(\epsilon),\ |\mathbf{v}_i^\star(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}})| = \mathcal{O}(\epsilon^2),\ i = 1, \ldots, n-1.$$

证明. F 在 $\tilde{\mathbf{x}} + \tilde{\mathbf{y}}$ 处的 Taylor 展开满足

$$F(\hat{\mathbf{x}}) = F(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}) + A(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}}) + \mathcal{O}(\epsilon^2).$$

等式两边分别乘以 $\mathbf{u}_i^\star$, 得到

$$-\mathbf{u}_i^\star F(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}) = \mathbf{u}_i^\star A(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}}) + \mathcal{O}(\epsilon^2).$$

因为 $\|F(\tilde{\mathbf{x}} + \tilde{\mathbf{y}})\| = \mathcal{O}(\epsilon^2)$ 和 $\mathbf{u}_i^\star A = \sigma_i \mathbf{v}_i^\star$, 所以

$$\sigma_i|\mathbf{v}_i^\star(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}})| = \mathcal{O}(\epsilon^2).$$

再由(3.4)推出

$$|\mathbf{v}_i^\star(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}})| = \mathcal{O}(\epsilon^2),\ i = 1, \ldots, n-1.$$

最后因为 $\|\tilde{\mathbf{x}} + \tilde{\mathbf{y}} - \hat{\mathbf{x}}\| = \Theta(\epsilon)$, 所以 $|\mathbf{v}_n^\star(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}})| = \Theta(\epsilon)$. □

因此 $\mathbf{v}_n$ 是一个能够达到二次收敛的迭代方向, 步长可以取为

$$\delta = \mathbf{v}_n^\star(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}}).$$

那么 $|\mathbf{v}_i^\star(\tilde{\mathbf{x}} + \tilde{\mathbf{y}} + \delta\mathbf{v}_n - \hat{\mathbf{x}})| = \mathcal{O}(\epsilon^2)$, $i = 1, \ldots, n$, 即 $\|\tilde{\mathbf{x}} + \tilde{\mathbf{y}} + \delta\mathbf{v}_n - \hat{\mathbf{x}}\| = \mathcal{O}(\epsilon^2)$.

例 3.5 (例 2.11 的延续). $F = \{x_1^2 + x_2 - 3, x_1 + 0.125x_2^2 - 1.5\}$ 和 $\hat{\mathbf{x}} = (1, 2)$. 对于近似根 $\tilde{\mathbf{x}} = (1.01, 2.01)$, $\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\| = \epsilon \approx 0.01414$, $F(\tilde{\mathbf{x}}) \approx (0.0301, 0.01501)^T$, $\sigma_1 = 2.517$, $\sigma_2 = 0.00598$, 通过解正则化的最小二乘问题(3.5), 得到

$$\tilde{\mathbf{y}} \approx (-0.01193, -0.005992),\ \|\tilde{\mathbf{y}}\| \approx 0.01135,$$

$$\tilde{\mathbf{x}} + \tilde{\mathbf{y}} \approx (0.9981, 2.004),\ F(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}) \approx (0.0001442, 0.00007224)^T.$$

计算 Jacobian 矩阵 $F_{\mathbf{x}}(\tilde{\mathbf{x}} + \tilde{\mathbf{y}})$ 的奇异值分解 $F_{\mathbf{x}}(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}) = U \cdot \Sigma \cdot V^\star$, 满足

$$\sigma_1 = 2.497,\ \sigma_2 = 0.00004657,\ V^\star(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}) = (-0.00006629, 0.004449)^T.$$

如果取步长为 $\delta = \mathbf{v}_2^\star(\hat{\mathbf{x}} - \tilde{\mathbf{x}} - \tilde{\mathbf{y}}) \approx -0.004449$, 那么

$$\tilde{\mathbf{x}} + \tilde{\mathbf{y}} + \delta\mathbf{v}_2 = (1.00005927, 2.00002969).$$

由上面的例子可以看出, 对于更好的近似根 $\tilde{\mathbf{x}}+\tilde{\mathbf{y}}$, Jacobian 矩阵的右奇异向量 $\mathbf{v}_n$ 确实是一个能够达到二次收敛的迭代方向. 下面通过计算近似局部对偶空间的一组既约基和解一个线性方程组, 我们来确定合适的步长 δ.

3.2.2 近似重结构

首先, 假设 $\tilde{\mathbf{x}}\in\mathbb{C}^n$ 是多项式系统 $F=\{f_1,\ldots,f_n\}$ 的近似根, $f_i\in\mathbb{C}[\mathbf{x}]$, 而且满足 $F(\hat{\mathbf{x}})=\mathbf{0}$, $\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}}))=n-1$ 和 $\|\tilde{\mathbf{x}}-\hat{\mathbf{x}}\|=\epsilon\ll 1$. 如果 $\{1,\Lambda_1,\ldots,\Lambda_{\mu-1}\}$ 是理想 $I=(f_1,\ldots,f_n)$ 在 $\tilde{\mathbf{x}}$ 处, 近似局部对偶空间的一组既约基, 那么

$$\|\Lambda_i(F)\|=\mathcal{O}(\epsilon),\ i=1,\ldots,\mu-1. \tag{3.8}$$

例 3.6 (例 2.1 的延续)**.** 多项式系统 $F=\{x_1^2x_2-x_1x_2^2,x_1-x_2^2\}$ 有一个孤立奇异根 $\hat{\mathbf{x}}=(0,0)$. 根据定理 2.13, $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 存在唯一的一组既约基

$$\begin{aligned}\mathcal{D}_{\hat{\mathbf{x}}}(I) &= \mathrm{Span}_{\mathbb{R}}\{1,d_2+a_1d_1,d_2^2+a_1d_1d_2+a_1^2d_1^2+a_2d_1,\\ &\quad d_2^3+a_1d_1d_2^2+a_1^2d_1^2d_2+a_1^3d_1^3+a_2d_1d_2+2a_1a_2d_1^2+a_3d_1\},\end{aligned}$$

$I=(x_1^2x_2-x_1x_2^2,x_1-x_2^2)$. 对近似根 $\tilde{\mathbf{x}}=(0.01,0.01)$, $\|\tilde{\mathbf{x}}-\hat{\mathbf{x}}\|\approx 0.01414$, 通过算法 AMSB1, 我们得到近似局部对偶空间的一组既约基

$$\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\{1,\Lambda_1,\Lambda_2,\Lambda_3\},\ a_1\approx 0.02,\ a_2\approx 1,\ a_3\approx -0.00001825,$$

满足

$$\Lambda_0(F)\approx(0.2\times 10^{-5},0.0099)^T,\ \Lambda_1(F)\approx(0.000306,-0.8993\times 10^{-7})^T,$$

$$\Lambda_2(F)\approx(0.0111,0.00005832)^T,\ \Lambda_3(F)\approx(0.0608,-0.00001831)^T.$$

通过解正则化的最小二乘问题(3.5), 我们可以得到更好的近似根 $\tilde{\mathbf{x}}+\tilde{\mathbf{y}}$. 假设奇异值分解 $F_{\mathbf{x}}(\tilde{\mathbf{x}}+\tilde{\mathbf{y}})=U\cdot\Sigma\cdot V^\star$, $\Sigma=\mathrm{diag}\{\sigma_1,\ldots,\sigma_n\}$, $U=(\mathbf{u}_1,\ldots,\mathbf{u}_n)$, $V=(\mathbf{v}_1,\ldots,\mathbf{v}_n)$. 下面证明, 如果 $\hat{\mathbf{x}}$ 的重数 $\mu>2$, 那么 $\sigma_n=\mathcal{O}(\epsilon^2)$. 为了简化证明过程, 对变元进行线性变换

$$\mathbf{x}=W\mathbf{z},$$

酉矩阵 $W=(\mathbf{v}_n,\mathbf{v}_1,\ldots,\mathbf{v}_{n-1})$. 显然 $\hat{\mathbf{z}}=W^\star\hat{\mathbf{x}}$ 是 $H(\mathbf{z}):=F(W\mathbf{z})$ 的根, 满足

$$\mathrm{rank}(H_{\mathbf{z}}(\hat{\mathbf{z}}))=\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})W)=n-1.$$

事实上, $\hat{\mathbf{z}}$ 的重数也等于 μ [19, 定理 3.3]. $\tilde{\mathbf{z}} = W^\star(\tilde{\mathbf{x}}+\tilde{\mathbf{y}})$ 是 H 的近似根, 满足

$$\|\tilde{\mathbf{z}}-\hat{\mathbf{z}}\| = \|W^\star(\tilde{\mathbf{x}}+\tilde{\mathbf{y}}-\hat{\mathbf{x}})\| = \Theta(\epsilon),\ H(\tilde{\mathbf{z}}) = F(\tilde{\mathbf{x}}+\tilde{\mathbf{y}}) = \mathcal{O}(\epsilon^2).$$

根据定理 3.4, 得到

$$|\tilde{z}_1-\hat{z}_1| = |\mathbf{v}_n^\star(\tilde{\mathbf{x}}+\tilde{\mathbf{y}}-\hat{\mathbf{x}})| = \Theta(\epsilon),$$

$$|\tilde{z}_i-\hat{z}_i| = |\mathbf{v}_{i-1}^\star(\tilde{\mathbf{x}}+\tilde{\mathbf{y}}-\hat{\mathbf{x}})| = \mathcal{O}(\epsilon^2),\ i=2,\ldots,n.$$

Jacobian 矩阵的第一列满足

$$\left\|\frac{\partial H(\tilde{\mathbf{z}})}{\partial z_1}\right\| = \|F_{\mathbf{x}}(\tilde{\mathbf{x}}+\tilde{\mathbf{y}})\mathbf{v}_n\| = \sigma_n.$$

由(3.4)和定理 2.10, 理想 $J=(h_1,\ldots,h_n)$ 在 $\tilde{\mathbf{z}}$ 处的近似局部对偶空间, 存在一组既约基 $\{1,d_1,\Lambda_2,\ldots,\Lambda_{\mu-1}\}$ 满足(3.8).

定理 3.5. 如果 $\mu>2$, 那么

$$\sigma_n = \left\|\frac{\partial H(\tilde{\mathbf{z}})}{\partial z_1}\right\| = \mathcal{O}(\epsilon^2).$$

证明. 如果 $\mu>2$, 根据定理 2.10 和(3.8), 存在二阶微分泛函

$$\|\Lambda_2(H)\| = \left\|\frac{1}{2}\frac{\partial^2 H(\tilde{\mathbf{z}})}{\partial z_1^2} + a_{2,2}\frac{\partial H(\tilde{\mathbf{z}})}{\partial z_2} + \cdots + a_{2,n}\frac{\partial H(\tilde{\mathbf{z}})}{\partial z_n}\right\| = \mathcal{O}(\epsilon).$$

令 $\mathbf{u}_n$ 是 $F_{\mathbf{x}}(\tilde{\mathbf{x}}+\tilde{\mathbf{y}})$ 的奇异值分解中对应于 σ_n 的左奇异向量. 由

$$\mathbf{u}_n^\star H_{\mathbf{z}}(\tilde{\mathbf{z}}) = \mathbf{u}_n^\star F_{\mathbf{x}}(\tilde{\mathbf{x}}+\tilde{\mathbf{y}})W = \sigma_n\mathbf{v}_n^\star W = \sigma_n(1,0,\ldots,0),$$

推出

$$\mathbf{u}_n^\star\frac{\partial H(\tilde{\mathbf{z}})}{\partial z_i} = 0,\ 2\le i\le n.$$

因此

$$\left|\mathbf{u}_n^\star\frac{\partial^2 H(\tilde{\mathbf{z}})}{\partial z_1^2}\right| = \mathcal{O}(\epsilon).$$

H 在 $\tilde{\mathbf{x}}$ 处的 Taylor 展开满足

$$H(\hat{\mathbf{z}}) = H(\tilde{\mathbf{z}}) + H_{\mathbf{z}}(\tilde{\mathbf{z}})(\hat{\mathbf{z}}-\tilde{\mathbf{z}}) + H_{\mathbf{zz}}(\tilde{\mathbf{z}})(\hat{\mathbf{z}}-\tilde{\mathbf{z}})^2 + \mathcal{O}(\epsilon^3),$$

其中 $(\hat{\mathbf{z}}-\tilde{\mathbf{z}})^2$ 表示由所有二阶单项式 $(\hat{z}_i-\tilde{z}_i)(\hat{z}_j-\tilde{z}_j)$ 构成的向量, $H_{\mathbf{zz}}(\tilde{\mathbf{z}})$ 表示由所有二阶偏导数 $\frac{\partial^2 H(\tilde{\mathbf{z}})}{\partial z_i \partial z_j}$ 构成的矩阵. 基于

$$|\tilde{z}_1-\hat{z}_1|=\Theta(\epsilon),\ |\tilde{z}_i-\hat{z}_i|=\mathcal{O}(\epsilon^2),\ i=2,\ldots,n,$$

除了 $(\hat{z}_1-\tilde{z}_1)^2$ 外, $(\hat{\mathbf{z}}-\tilde{\mathbf{z}})^2$ 中其他元素都是 $\mathcal{O}(\epsilon^3)$ 的. 因为

$$\left|\mathbf{u}_n^\star \frac{\partial^2 H(\tilde{\mathbf{z}})}{\partial z_1^2}(\hat{z}_1-\tilde{z}_1)^2\right|=\mathcal{O}(\epsilon^3),$$

所以

$$\left|\mathbf{u}_n^\star\, H_{\mathbf{zz}}(\tilde{\mathbf{z}})(\hat{\mathbf{z}}-\tilde{\mathbf{z}})^2\right|=\mathcal{O}(\epsilon^3).$$

另一方面, 令 $\mathbf{u}_0$ 是 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的左零向量. 利用类似的分析, 可以得到

$$|\mathbf{u}_0^\star H_{\mathbf{zz}}(\hat{\mathbf{z}})(\tilde{\mathbf{z}}-\hat{\mathbf{z}})^2|=\mathcal{O}(\epsilon^3).$$

结合 H 在 $\hat{\mathbf{x}}$ 处的 Taylor 展开

$$H(\tilde{\mathbf{z}})=H(\hat{\mathbf{z}})+H_{\mathbf{z}}(\hat{\mathbf{z}})(\tilde{\mathbf{z}}-\hat{\mathbf{z}})+H_{\mathbf{zz}}(\hat{\mathbf{z}})(\tilde{\mathbf{z}}-\hat{\mathbf{z}})^2+\mathcal{O}(\epsilon^3),$$

推出

$$\mathbf{u}_0^\star H(\tilde{\mathbf{z}})=\mathcal{O}(\epsilon^3).$$

再由 $\|\mathbf{u}_n-\mathbf{u}_0\|=\mathcal{O}(\epsilon)$ 和 $H(\tilde{\mathbf{z}})=\mathcal{O}(\epsilon^2)$, 得到

$$|\mathbf{u}_n^\star H(\tilde{\mathbf{z}})|\le|(\mathbf{u}_n-\mathbf{u}_0)^\star H(\tilde{\mathbf{z}})|+|\mathbf{u}_0^\star H(\tilde{\mathbf{z}})|=\mathcal{O}(\epsilon^3).$$

因此

$$|\mathbf{u}_n^\star H_{\mathbf{z}}(\tilde{\mathbf{z}})(\hat{\mathbf{z}}-\tilde{\mathbf{z}})|=|\mathbf{u}_n^\star H(\tilde{\mathbf{z}})|+|\mathbf{u}_n^\star H_{\mathbf{zz}}(\tilde{\mathbf{z}})(\hat{\mathbf{z}}-\tilde{\mathbf{z}})^2|+\mathcal{O}(\epsilon^3)=\mathcal{O}(\epsilon^3).$$

等价于

$$|\mathbf{u}_n^\star H_{\mathbf{z}}(\tilde{\mathbf{z}})(\hat{\mathbf{z}}-\tilde{\mathbf{z}})|=|\sigma_n(1,0,\ldots,0)(\hat{\mathbf{z}}-\tilde{\mathbf{z}})|=\sigma_n|\tilde{z}_1-\hat{z}_1|=\mathcal{O}(\epsilon^3).$$

所以 $\sigma_n=\mathcal{O}(\epsilon^2)$. □

事实上, 对于 $\mathbf{v}_n = (v_{n,1}, \ldots, v_{n,n})^T$, $\exists t \in \mathbb{N}$, 满足

$$|v_{n,t}| \geq |v_{n,i}|,\ i = 1, \ldots, n.$$

交换变元次序 $x_1 \leftrightarrow x_t$, 可以得到理想 $I = (f_1, \ldots, f_n)$ 在 $\tilde{\mathbf{x}} + \tilde{\mathbf{y}}$ 处, 近似局部对偶空间的一组既约基 $\{1, d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n, \Lambda_2, \ldots, \Lambda_{\mu-1}\}$,

$$a_{1,2} = \frac{v_2}{v_t}, \ldots, a_{1,t} = \frac{v_1}{v_t}, \ldots, a_{1,n} = \frac{v_n}{v_t},\ \mathbf{a}_1 = (1, a_{1,2}, \ldots, a_{1,n})^T.$$

如果 $\mu > 2$,

$$\|\Lambda_1(F)\| = \|F_{\mathbf{x}}(\tilde{\mathbf{x}} + \tilde{\mathbf{y}})\mathbf{a}_1\| = \frac{\sigma_n}{|v_{n,t}|} = \mathcal{O}(\epsilon^2).$$

不仅如此, 我们还发现

$$\|\Lambda_i(F)\| = \mathcal{O}(\epsilon^2),\ i = 1, \ldots, \mu - 2.$$

即通过解正则化的最小二乘问题(3.5), 我们不仅得到了更好的近似根 $\tilde{\mathbf{x}} + \tilde{\mathbf{y}}$, 而且近似重结构也得到了精化.

定理 3.6. 如果 $\mu > 2$,

$$\|\Lambda_i(F)\| = \mathcal{O}(\epsilon^2),\ i = 1, \ldots, \mu - 2.$$

证明. 首先根据定理 3.5, 对于 $\mu = 3$ 定理成立. 如果 $\mu = 4$, 根据定理 3.2, 存在唯一的 $\hat{\boldsymbol{\lambda}} \in \mathbb{C}^n$, 使得 $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})$ 是增广系统

$$G(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} F, \\ F_{\mathbf{x}}\boldsymbol{\lambda}, \\ \lambda_1 - 1, \end{cases}$$

的根, 并且重数为 3. 显然 $(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}, \mathbf{a}_1)$ 是 $G(\mathbf{x}, \boldsymbol{\lambda})$ 的近似根, 满足

$$\|(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}, \mathbf{a}_1) - (\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})\| = \Theta(\epsilon),\ \|G(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}, \mathbf{a}_1)\| = \mathcal{O}(\epsilon^2).$$

由定理 3.4 和 3.5 的证明, 存在 $(\tilde{\mathbf{x}} + \tilde{\mathbf{y}}, \mathbf{a}_1)$ 处的一阶微分泛函

$$\|\Lambda_1'(G)\| = \|(d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n + 2a_{2,2}d_{n+2} + \cdots + 2a_{2,n}d_{2n})(G)\| = \mathcal{O}(\epsilon^2).$$

由定理 3.2 的证明, Λ_1' 可以诱导出 $\tilde{\mathbf{x}} + \tilde{\mathbf{y}}$ 处的一阶微分泛函 Λ_2,

$$\Lambda_2 = \Delta_2 + a_{2,2}d_2 + \cdots + a_{2,n}d_n,\ \|\Lambda_2(F)\| = \|\Lambda_1'(G)\| = \mathcal{O}(\epsilon^2).$$

因此对于 $\mu = 4$ 定理成立. 余下的证明可以通过归纳法得到. □

例 3.7 (例 2.1 的延续). $F=\{x_1^2x_2-x_1x_2^2, x_1-x_2^2\}$ 有一个孤立奇异根 $\hat{\mathbf{x}}=(0,0)$. 对于近似根 $\tilde{\mathbf{x}}=(0.01,0.01)$, 通过解正则化的最小二乘问题(3.5), 得到

$$\tilde{\mathbf{x}}+\tilde{\mathbf{y}}\approx(0.000107,0.0102).$$

对更好的近似根 $\tilde{\mathbf{x}}+\tilde{\mathbf{y}}$, $\|\tilde{\mathbf{x}}+\tilde{\mathbf{y}}-\hat{\mathbf{x}}\|\approx 0.0102$, 通过算法 AMSB1, 得到近似局部对偶空间的一组既约基

$$\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\{1,\Lambda_1,\Lambda_2,\Lambda_3\},\ a_1\approx 0.0204,\ a_2\approx 1,\ a_3\approx -0.4429\times 10^{-5},$$

满足

$$\Lambda_0(F)\approx(0.1125\times 10^{-7},0.2987\times 10^{-5})^T,\ \Lambda_1(F)\approx(0.436\times 10^{-5},-0.4627\times 10^{-9})^T,$$

$$\Lambda_2(F)\approx(0.0006379,0.1357\times 10^{-6})^T,\ \Lambda_3(F)\approx(0.04163,-0.4429\times 10^{-5})^T.$$

事实上, 基于定理 3.5 和 3.6, 通过解线性方程组

$$\begin{pmatrix}\Delta_\mu(f_1) & d_2(f_1) & \cdots & d_n(f_1)\\ \vdots & \vdots & & \vdots\\ \Delta_\mu(f_n) & d_2(f_n) & \cdots & d_n(f_n)\end{pmatrix}\begin{pmatrix}a_{\mu,1}\\ a_{\mu,2}\\ \vdots\\ a_{\mu,n}\end{pmatrix}=-\begin{pmatrix}\Lambda_{\mu-1}(f_1)\\ \vdots\\ \Lambda_{\mu-1}(f_n)\end{pmatrix}, \tag{3.9}$$

可以确定步长 $\delta:=\frac{a_{\mu,1}}{\mu}$. 我们看下面两个例子:

例 3.8 (例 2.1 的延续). 对于近似根 $\tilde{\mathbf{x}}=(0.01,0.01)$, 通过解正则化的最小二乘问题(3.5), 得到

$$\tilde{\mathbf{x}}+\tilde{\mathbf{y}}\approx(0.000107,0.01019877),$$

再通过算法 AMSB1, 得到近似局部对偶空间的一组既约基

$$\begin{aligned}\mathcal{D}_{\hat{\mathbf{x}}}(I) &= \mathrm{Span}_{\mathbb{R}}\{1, d_2+a_1d_1, d_2^2+a_1d_1d_2+a_1^2d_1^2+a_2d_1\\ &\quad d_2^3+a_1d_1d_2^2+a_1^2d_1^2d_2+a_1^3d_1^3+a_2d_1d_2+2a_1a_2d_1^2+a_3d_1\},\\ \text{和 } \Delta_4 &= d_2^4+a_1d_1d_2^3+a_1^2d_1^2d_2^2+a_1^3d_1^3d_2+a_1^4d_1^4\\ &\quad +a_2d_1d_2^2+2a_1a_2d_1^2d_2+3a_1^2a_2d_1^3+a_3d_1d_2+(2a_1a_3+a_2^2)d_1^2,\end{aligned}$$

满足

$$a_1\approx 0.02039754,\ a_2\approx 1.00000014,\ a_3\approx -0.00000443.$$

最后通过解线性方程组

$$\begin{pmatrix} \Delta_4(f_1) & d_1(f_1) \\ \Delta_4(f_2) & d_1(f_2) \end{pmatrix} \begin{pmatrix} a_{4,1} \\ a_{4,2} \end{pmatrix} = - \begin{pmatrix} \Lambda_3(f_1) \\ \Lambda_3(f_2) \end{pmatrix},$$

得到步长

$$\delta := \frac{a_{4,1}}{4} \approx -0.01009929.$$

从而

$$\tilde{\mathbf{x}} + \tilde{\mathbf{y}} + \delta(1, a_1) \approx (-0.000099, 0.00009948).$$

例 3.9 (例 2.11 的延续). $F = \{x_1^2 + x_2 - 3, x_1 + 0.125x_2^2 - 1.5\}$ 有一个孤立奇异根 $\hat{\mathbf{x}} = (1, 2)$. 对于近似根 $\tilde{\mathbf{x}} = (1.01, 2.01)$, 通过解正则化的最小二乘问题(3.5), 得到

$$\tilde{\mathbf{x}} + \tilde{\mathbf{y}} \approx (0.99806632, 2.00400783),$$

再通过算法 AMSB1, 得到近似局部对偶空间的一组既约基

$$\mathcal{D}_{\hat{\mathbf{x}}}(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_1 + a_1 d_2, d_1^2 + a_1 d_1 d_2 + a_1^2 d_2^2 + a_2 d_2\},$$

和 $\Delta_3 = d_1^3 + a_1 d_1^2 d_2 + a_1^2 d_1 d_2^2 + a_1^3 d_2^3 + a_2 d_1 d_2 + 2a_1 a_2 d_2^2$, 满足

$$a_1 \approx -1.99610606, \ a_2 \approx -0.99882255.$$

最后通过解线性方程组

$$\begin{pmatrix} \Delta_3(f_1) & d_2(f_1) \\ \Delta_3(f_2) & d_2(f_2) \end{pmatrix} \begin{pmatrix} a_{3,1} \\ a_{3,2} \end{pmatrix} = - \begin{pmatrix} \Lambda_2(f_1) \\ \Lambda_2(f_2) \end{pmatrix},$$

得到步长

$$\delta := \frac{a_{3,1}}{3} \approx 0.00197084.$$

从而

$$\tilde{\mathbf{x}} + \tilde{\mathbf{y}} + \delta(1, a_1) \approx (1.00003716, 2.00007382).$$

由上面两个例子可以看出, 通过解线性方程组(3.9)的确可以得到合适的步长 δ.

3.2.3 算法和二次收敛性

根据前面的分析, 我们接下来给出算法 MRRB1(Multiple Roots Refiner for Breadth One), 并证明算法在宽度为 1 的孤立奇异根附近是二次收敛的.

算法：*MRRB1*

输入：▸ 多项式系统 $F=\{f_1,\ldots,f_n\}$, $f_i\in\mathbb{C}[\mathbf{x}]$, 近似根 $\tilde{\mathbf{x}}\in\mathbb{C}^n$ 和容忍度 τ.

输出：▸ 精化的近似根 $\tilde{\mathbf{x}}$.

第 0 步 解正则化的最小二乘问题(3.5), 令 $\tilde{\mathbf{x}}:=\tilde{\mathbf{x}}+\tilde{\mathbf{y}}$.

第 1 步 计算 Jacobian 矩阵 $F_{\mathbf{x}}(\tilde{\mathbf{x}})$ 的奇异值分解 $F_{\mathbf{x}}(\tilde{\mathbf{x}})=U\cdot\Sigma\cdot V^{\star}$.

- 如果 $\sigma_n<\tau\ll\sigma_{n-1}$, 利用算法 AMSB1 计算 $\mathbf{a}_1$, $F_{\mu-1}(\tilde{\mathbf{x}})$ 和 $G_\mu(\tilde{\mathbf{x}})$.
- 否则, 退出.

第 2 步 解线性方程组(3.9)

$$(G_\mu(\tilde{\mathbf{x}}),M)\,\mathbf{a}_\mu=-F_{\mu-1}(\tilde{\mathbf{x}}),$$

令 $\delta:=\frac{a_{\mu,1}}{\mu}$.

第 3 步 返回 $\tilde{\mathbf{x}}:=\tilde{\mathbf{x}}+\delta\cdot\mathbf{a}_1$.

定理 3.7. *假设 $\tilde{\mathbf{x}}\in\mathbb{C}^n$ 是 $F=\{f_1,\ldots,f_n\}$ 的近似根, $f_i\in\mathbb{C}[\mathbf{x}]$, 满足 $F(\hat{\mathbf{x}})=\mathbf{0}$, $\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}}))=n-1$ 和 $\|\tilde{\mathbf{x}}-\hat{\mathbf{x}}\|=\epsilon\ll 1$. 如果重数 μ 的计算正确无误, 那么算法 MRRB1 返回的精化根 $\tilde{\mathbf{x}}$ 满足*

$$\|\tilde{\mathbf{x}}-\hat{\mathbf{x}}\|=\mathcal{O}(\epsilon^2).$$

证明. 根据算法 MRRB1, 通过解正则化的最小二乘问题(3.5), 得到更好的近似根 $\tilde{\mathbf{x}}:=\tilde{\mathbf{x}}+\tilde{\mathbf{y}}$. 调整变元次序, 根据定理 3.6 和(3.8),

$$\|\Lambda_i(F)\|=\mathcal{O}(\epsilon^2),\ 0\le i\le\mu-2,\ \|\Lambda_{\mu-1}(F)\|=\mathcal{O}(\epsilon),$$

$\{1,d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n,\Lambda_2,\ldots,\Lambda_{\mu-1}\}$ 是理想 $I=(f_1,\ldots,f_n)$ 在 $\tilde{\mathbf{x}}$ 处, 近似局部对偶空间的一组既约基, 满足

$$\Lambda_i(F)=F_i(\tilde{\mathbf{x}})=G_i(\tilde{\mathbf{x}})+a_{i,2}\frac{\partial F(\tilde{\mathbf{x}})}{\partial x_2}+\cdots+a_{i,n}\frac{\partial F(\tilde{\mathbf{x}})}{\partial x_n}.$$

由局部对偶空间的封闭性推出

$$\|\Lambda_i((\mathbf{x}-\tilde{\mathbf{x}})^{\boldsymbol{\alpha}}F)\| = \mathcal{O}(\epsilon^2),\ 0 \le i \le \mu-2,\ \forall \boldsymbol{\alpha} \in \mathbb{N}^n.$$

通过解线性方程组(3.9), 其中 $\Delta_\mu(F) = G_\mu(\tilde{\mathbf{x}})$, 得到

$$\Lambda_\mu := \Lambda_{\mu-1} + a_{\mu,1} \cdot \Delta_\mu + a_{\mu,2} \cdot d_2 + \cdots + a_{\mu,n} \cdot d_n,$$

满足 $\Lambda_\mu(F) = \mathbf{0}$. 由 $\|\Lambda_{\mu-1}(F)\| = \mathcal{O}(\epsilon)$ 和(3.9)中的系数矩阵满秩($\sigma_n > \tau$)得到

$$\|\mathbf{a}_\mu\| = \mathcal{O}(\epsilon).$$

再由 $\Phi_{x_i}(\Delta_\mu) \in \mathrm{Span}_{\mathbb{C}}\{1, \Lambda_1, \ldots, \Lambda_{\mu-1}\}$ 和 $|a_{\mu,1}| = \mathcal{O}(\epsilon)$, 推出

$$\|\Lambda_\mu((\mathbf{x}-\tilde{\mathbf{x}})^{\boldsymbol{\alpha}}F)\| = \mathcal{O}(\epsilon^2).$$

因此

$$\|M_{\mu+1} \cdot \Lambda_i(\mathbf{v}(\mathbf{x})_\mu)\| = \mathcal{O}(\epsilon^2),\ i = 0, 1, \ldots, \mu-2, \mu,$$

$M_{\mu+1}$ 是多项式系统 F 和它所有直至 μ 阶的延拓在 $\tilde{\mathbf{x}}$ 处的 Taylor 展开对应的系数矩阵, 即 Macaulay 方法中系数矩阵的 μ 阶截断(详情见 [44]),

$$\mathbf{v}(\mathbf{x})_\mu = ((x_1-\tilde{x}_1)^\mu, \ldots, (\mathbf{x}-\tilde{\mathbf{x}})^{\boldsymbol{\alpha}}, \ldots, x_n-\tilde{x}_n, 1)^T.$$

值得注意的是, 我们通过封闭性条件直接构造出 $M_{\mu+1}$ 的近似零空间而无需计算它本身(矩阵规模很大). 类似于 [44, 注 18] 中的分析, 基于近似零向量 $\Lambda_i(\mathbf{v}(\mathbf{x})_\mu)$ 得到的乘法矩阵 $\widetilde{M}_{x_i}$ 的迹满足

$$\frac{1}{\mu}\mathrm{Trace}(\widetilde{M}_{x_i}) = \frac{1}{\mu}\mathrm{Trace}(M_{x_i}) + \mathcal{O}(\epsilon^2) = \hat{x}_i - \tilde{x}_i + \mathcal{O}(\epsilon^2).$$

事实上, 乘法矩阵 $\widetilde{M}_{x_1}$ 满足

$$\widetilde{M}_{x_1} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & a_{\mu,1} \end{pmatrix}.$$

因此
$$\left|\tilde{x}_1+\frac{a_{\mu,1}}{\mu}-\hat{x}_1\right|=\mathcal{O}(\epsilon^2).$$
最后根据定理 3.4, 我们得到
$$\|\tilde{\mathbf{x}}-\hat{\mathbf{x}}\|=\mathcal{O}(\epsilon^2),\ \tilde{\mathbf{x}}:=\tilde{\mathbf{x}}+\frac{a_{\mu,1}}{\mu}\cdot\begin{pmatrix}1\\a_{1,2}\\\vdots\\a_{1,n}\end{pmatrix}.$$
□

3.3 数值实验

在 Maple 中我们实现了算法 MRRB1, 程序代码和一些例子的实验结果可以在 http://www.mmrc.iss.ac.cn/~lzhi/Research/hybrid/rootrefinerbreadthone/ 找到.

表 3.1 展示了算法 MRRB1 对下述例子的测试结果:

1. Ojika1 [28]: $x^2+y-3,\ x+\frac{1}{8}y^2-\frac{3}{2}$
2. Ojika2 [28]: $x^2+y+z-1,\ x+y^2+z-1,\ x+y+z^2-1$
3. Ojika3 [28]: $x+y+z-1,\ 2x^3+5y^2-10z+5z^3+5,\ 2x+2y+z^2-1$
4. Ojika4 [28]: $x+x^3z+xy^2z-xz,\ 10y-2x^2yz-y^3z-yz,$
   $6x^4z^2-3x^2y^2z^2-x^2z^2+28x^2z-3y^4z^3+2y^2z^2+7y^2z+z^2-11z+10$
5. Decker2 [8]: $x+y^3,\ x^2y-y^4$
6. DZ3 [7]: $14x+33y-3\sqrt{5}(x^2+4xy+4y^2+2)+\sqrt{7}+x^3+6x^2y+12xy^2+8y^3,$
   $41x-18y-\sqrt{5}+8x^3-12x^2y+6xy^2-y^3+3\sqrt{7}(4xy-4x^2-y^2-2)$
7. Dayton2 [5]: $2x^2-x-x^3+z^3,\ x-y-x^2+xy+z^2,\ xy^2z-x^2z-y^2z+x^3z$
8. RG [35]: $x^2x_2-x_1x_2^2,\ x_1-x_2^2$
9. SY5 [36]: $x_1+x_2-2, x_1^2+x_2^2-2$

10. Menzel [24]: x^2-xy+y^2+x-2, $3x^2+2xy+2y-7$, $2x^2-8x+y^2-4y+9$

表 3.1 中最后一列表示近似奇异根的精度变化: 第一个数字表示初始近似的精度, 一个箭头表示迭代一次算法 MRRB1.

表 3.1:　算法 MRRB1 的实验结果

系统	孤立奇异根	方程	变量	重数	精度
Ojika1	$(1,2)$	2	2	3	$2\to3\to8\to16$
Ojika2	$(1,0,0)$	3	3	2	$2\to3\to7\to14$
Ojika3	$(-\frac{5}{2},\frac{5}{2},1)$	3	3	2	$2\to4\to7\to16$
Ojika4	$(0,0,10)$	3	3	3	$3\to4\to8\to17$
Decker2	$(0,0)$	2	2	4	$3\to8\to25$
DZ3	$(\frac{2\sqrt{7}}{5}+\frac{\sqrt{5}}{5},-\frac{\sqrt{7}}{5}+\frac{2\sqrt{5}}{5})$	2	2	5	$3\to8\to12$
Dayton2	$(0,0,0)$	3	3	5	$3\to5\to10\to20$
RG	$(0,0)$	2	2	4	$2\to4\to7\to16$
SY5	$(1,1)$	2	2	2	$2\to4\to9\to19$
Menzel1	$(1,1)$	3	2	2	$2\to4\to9\to20$

注 3. 例子 DZ3 和 Menzel1 表明, 算法 MRRB1 对于带有近似系数的多项式系统和超定系统(方程个数大于变元个数)同样有效.

第四章　可信误差界的计算

本章中, 我们介绍了经典的多项式系统正则根的区间验证方法, 讨论了验证孤立奇异根的若干难点和解决思路. 利用宽度为 1 的特殊情形下重结构的参数化表示, 我们提出了一种计算近似奇异根可信误差界的新算法, 从而保证了一个带有微小扰动的多项式系统, 在误差界内有一个宽度为 1 的孤立奇异根. 对于一般的孤立奇异根, 我们提出一种带光滑参数的 deflation 技术, 并且基于这种技术, 我们将验证宽度为 1 孤立奇异根的算法推广到了一般情形.

4.1　根的区间验证

大多数基于浮点计算的求根算法, 只能对数值结果的近似程度提供误差“估计”. 我们希望能够计算近似根的误差“界”, 从而证明精确根的存在性.

问题 4.1. 对于给定的一个多项式系统 $F=\{f_1,\ldots,f_n\}$, $f_i\in\mathbb{R}[\mathbf{x}]$ 和它的一个近似根 $\tilde{\mathbf{x}}\in\mathbb{R}^n$, 如何得到一个区间向量 $\mathbf{X}\in\mathbb{IR}^n$, 使得存在唯一的 $\hat{\mathbf{x}}\in\tilde{\mathbf{x}}+\mathbf{X}$ 满足 $F(\hat{\mathbf{x}})=\mathbf{0}$?

记 $\mathbb{IR}$ 是实数域上的区间集合, $\mathbb{IR}^n$ 和 $\mathbb{IR}^{n\times n}$ 分别是实数域上的区间向量集合和区间矩阵集合. 除非特殊声明, 本章只考虑实系数多项式系统的实根.

下面我们介绍多项式系统正则根的区间验证方法, 并且讨论验证奇异根的若干难点和解决思路.

4.1.1　正则根

多项式系统正则根的区间验证主要基于如下定理 [15, 25, 33]:

定理 4.1. *假设 $\tilde{\mathbf{x}}\in\mathbb{R}^n$ 是 $F=\{f_1,\ldots,f_n\}$ 的近似根, $f_i\in\mathbb{R}[\mathbf{x}]$. 对于给定的区间向量 $\mathbf{X}\in\mathbb{IR}^n$ 满足 $\mathbf{0}\in\mathbf{X}$ 和区间矩阵 $\mathbf{M}\in\mathbb{IR}^{n\times n}$ 满足 $\nabla f_i(\tilde{\mathbf{x}}+\mathbf{X})\subseteq\mathbf{M}_{i,:}$,如果区间判定条件*

$$-F_{\mathbf{x}}^{-1}(\tilde{\mathbf{x}})F(\tilde{\mathbf{x}})+(I_n-F_{\mathbf{x}}^{-1}(\tilde{\mathbf{x}})M)\mathbf{X}\subseteq\mathrm{int}(\mathbf{X})$$

成立, 则存在唯一的 $\hat{\mathbf{x}}\in\tilde{\mathbf{x}}+\mathbf{X}$ 满足 $F(\hat{\mathbf{x}})=\mathbf{0}$. 而且所有矩阵 $\widetilde{M}\in\mathbf{M}$ 都是非奇异的, 特别地, $Jacobian$ 矩阵 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 非奇异.

定理 4.1 的证明主要基于 Brouwer 的区间不动点定理和多变元系统的中值定理 [33]: 对 $\forall \tilde{\mathbf{y}} \in \tilde{\mathbf{x}} + \mathbf{X}$, 存在 $\widetilde{M} \in M$ 满足 $F(\tilde{\mathbf{y}}) = F(\tilde{\mathbf{x}}) + \widetilde{M}(\tilde{\mathbf{y}} - \tilde{\mathbf{x}})$. 基于定理 4.1, S. Rump 在 INTLAB [34] 中已经实现了算法 verifynlss. 值得注意的是, 算法 verifynlss 由两部分组成: 首先对近似根 $\tilde{\mathbf{x}}$ 进行精化, 然后利用定理 4.1 进行验证, 从而得到 $\hat{\mathbf{x}} \in \tilde{\mathbf{x}} + \mathbf{X}$.

例 4.1 (例 2.1 的延续)**.** 对于 $F = \{x_1^2 x_2 - x_1 x_2^2, x_1 - x_2^2\}$ 和近似根 $(1.01, 1.01)$,

$$\begin{aligned}
&\gg\ f = \text{inline}('[x(1)^2 * x(2) - x(1) * x(2)^2; x(1) - x(2)^2]')\\
&f =\\
&\text{Inline function :}\\
&f(x) = [x(1)^2 * x(2) - x(1) * x(2)^2; x(1) - x(2)^2]\\
&\gg\ \text{verifynlss}(f, [1.01; 1.01])\\
&\text{intval ans} =\\
&[0.99999999999999, 1.00000000000001]\\
&[0.99999999999999, 1.00000000000001]
\end{aligned}$$

而对于近似根 $(0.01, 0.01)$,

$$\begin{aligned}
&\gg\ \text{verifynlss}(f, [0.01; 0.01])\\
&\text{intval ans} =\\
&[\qquad\qquad \text{NaN}, \qquad\qquad \text{NaN}]\\
&[\qquad\qquad \text{NaN}, \qquad\qquad \text{NaN}]
\end{aligned}$$

这是因为 $(1, 1)$ 是 $F(x_1, x_2) = \mathbf{0}$ 的正则根, 而 $(0, 0)$ 是奇异根.

由上面的例子可以看出, 算法 verifynlss 只适用于多项式系统正则根的区间验证(Jacobian 矩阵 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 非奇异).

4.1.2 奇异根

利用第三章中介绍的 deflation 技术, 可以将多项式系统奇异根的可信验证问题转化为一个增广系统正则根的区间验证. 但通过 deflation 技术得到的增广系统总是超定的(over-determined), 从而导致定理 4.1 依然无法适用.

为了克服这个困难, 我们试图对原多项式系统添加一些光滑参数, 构造一个带扰动的多项式系统, 然后验证它在狭窄的扰动误差界内有一个孤立奇异根. 这

种解决思路是合理的, 因为验证多项式系统是否有奇异根是一个病态问题—多项式系数任意微小的扰动都可能导致一个奇异根变成一族正则根. 因此, 利用带有舍入误差的浮点计算, 来验证多项式系统奇异根的存在性是不可能的 [35].

目前已经有一些工作 [13, 21, 35] 沿着这种解决思路进行了尝试.

Y. Kanzawa 和 S. Oishi 首先注意到光滑参数的添加位置和 Jacobian 矩阵行的线性无关性有关 [13]. 他们提出了一种带光滑参数的 deflation 技术, 使增广系统的变元个数等于方程个数. 然后利用定理 4.1, 证明了增广系统正则根的存在唯一性, 从而给出了“不完美”奇异根(imperfect singular solution)的区间验证.

例 4.2 (例 2.11 的延续)**.** 对于 $F = \{x_1^2 + x_2 - 3, x_1 + 0.125x_2^2 - 1.5\}$ 和它的一个近似根 $\tilde{\mathbf{x}} = (1.01, 2.01)$, F 的 Jacobian 矩阵

$$F_{\mathbf{x}}(\tilde{\mathbf{x}}) = \begin{pmatrix} 2.02 & 1 \\ 1 & 0.5025 \end{pmatrix}.$$

由于 $F_{\mathbf{x}}(\tilde{\mathbf{x}})$ 的第一行能被第二行近似线性表示, 我们将第一个光滑参数 x_3 添加到第一个方程上, 得到 $G = \{f_1 + x_3, f_2, 2x_1 + x_4, 0.25x_2x_4 + 1\}$ 和它的一个近似根 $\tilde{\mathbf{y}} = (1.01, 2.01, 0, -2.014)$. G 的 Jacobian 矩阵

$$G_{\mathbf{x}}(\tilde{\mathbf{y}}) = \begin{pmatrix} 2.02 & 1 & 1 & 0 \\ 1 & 0.5025 & 0 & 0 \\ 2 & 0 & 0 & 1 \\ 0 & -0.5035 & 0 & 0.5025 \end{pmatrix}.$$

由于 $G_{\mathbf{x}}(\tilde{\mathbf{y}})$ 的第三行能被另三行近似线性表示, 我们将第二个光滑参数 x_5 添加到第三个方程上, 得到 $H = \{g_1, g_2, g_3 + x_5, g_4, 2x_1 + x_6 + x_7, 0.25x_2x_6 + 1, x_8 + 2, 0.25x_4x_6 + 0.25x_2x_8\}$ 和近似根 $\tilde{\mathbf{z}} = (1.01, 2.01, 0, -2.014, 0, -1.99, 0.03, -2)$.

$$\gg\ f = \mathrm{inline}('[x(1)^2 + x(2) - 3 + x(3); x(1) + 0.125 * x(2)^2 - 1.5;$$
$$2 * x(1) + x(4) + x(5); 1 + 0.25 * x(2) * x(4); 2 * x(1) + x(6) + x(7);$$
$$1 + 0.25 * x(2) * x(6); x(8) + 2; 0.25 * x(4) * x(6) + 0.25 * x(2) * x(8)]')$$
$$f =$$
$$\text{Inline function :}$$
$$f(x) = [x(1)^2 + x(2) - 3 + x(3); x(1) + 0.125 * x(2)^2 - 1.5;$$
$$2 * x(1) + x(4) + x(5); 1 + 0.25 * x(2) * x(4); 2 * x(1) + x(6) + x(7);$$
$$1 + 0.25 * x(2) * x(6); x(8) + 2; 0.25 * x(4) * x(6) + 0.25 * x(2) * x(8)]$$

$\gg$ verifynlss(f, [1.01; 2.01; 0; −2.014; 0; −1.99; 0.03; −2])

intval ans =

[0.99999999999999, 1.00000000000001]

[1.99999999999999, 2.00000000000001]

[−0.00000000000001, 0.00000000000001]

[−2.00000000000001, −1.99999999999999]

[−0.00000000000001, 0.00000000000001]

[−2.00000000000001, −1.99999999999999]

[−0.00000000000001, 0.00000000000001]

[−2.00000000000000, −2.00000000000000]

上述区间向量内, 存在增广系统 H 唯一的一个正则根, 从而验证了 F 的一个 $(\hat{x}_3, \hat{x}_5)$-"不完美"奇异根(详情见 [13]).

S. Rump 和 S. Graillat 进一步分析了二重根(double roots), 验证了一个带扰动的多项式系统在狭窄的扰动误差界内有一个二重根 [35].

例 4.3. 对例 4.2 中的 $G = \{x_1^2+x_2+x_3-3, x_1+0.125x_2^2-1.5, 2x_1+x_4, 0.25x_2x_4+1\}$ 和近似根 $\tilde{\mathbf{y}} = (1.01, 2.01, 0, -2.014)$, 扰动多项式系统 $\widetilde{G} = \{g_1, g_2, g_3 + \hat{x}_5, g_4\}$ 在扰动误差界

$$-0.00000000000001 \leq \hat{x}_5 \leq 0.00000000000001$$

内存在唯一的二重根 $(\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)$, 且 $\mathrm{rank}(\widetilde{G}_{\mathbf{x}}(\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4)) = 3$.

A. Mantzaflaris 和 B. Mourrain 利用近似局部对偶空间一组给定的既约基, 验证了一个原系统附近的多项式系统(另一个系统)有一个孤立奇异根 [21]. 这种验证方法很大程度上依赖于给定重结构的近似程度.

沿着同样的思路, 下面利用第二章中给出的, 宽度为 1 的特殊情形下重结构的参数化表示, 我们提出了一种计算近似奇异根可信误差界的新算法, 从而保证了一个带有微小扰动的多项式系统, 在误差界内有一个宽度为 1 的孤立奇异根.

4.2 宽度为 1 的孤立奇异根的验证

我们注意到, [13] 中的方法之所以只能验证"不完美"奇异根, 是因为他们将光滑参数不仅添加到原系统上, 同样也添加到其他 deflation 系统上. 这也是为什么 [35] 中的方法可以验证二重根的原因(唯一的光滑参数只添加到原系统上). 同时我们注意到, [21] 中的方法虽然将所有光滑参数都添加到原系统上, 但由于他们利用"近似的"重结构生成 deflation 系统, 所以只能验证一个原系统附近的多项式系统(另一个系统)有一个孤立奇异根. 因此我们考虑利用第二章中给出的, 宽度为 1 的特殊情形下重结构的参数化表示, 结合一种恰当的光滑参数添加方法, 来验证宽度为 1 的孤立奇异根的存在性.

假设 $\hat{\mathbf{x}} \in \mathbb{R}^n$ 是 $F = \{f_1, \ldots, f_n\}$ 的根, $f_i \in \mathbb{R}[\mathbf{x}]$, 满足 $n - \mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = 1$. 那么, 必然存在 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的某一列能够被其他列线性表示, 不妨设为第一列; 必然存在 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的某一行能够被其他行线性表示, 不妨设为第一行. 如若不然, 我们计算 Jacobian 矩阵 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的一个非零零向量 $(v_1, \ldots, v_n)^T$ 和 $F_{\mathbf{x}}^T(\hat{\mathbf{x}})$ 一个非零零向量 $(u_1, \ldots, u_n)^T$. 那么 $\exists t_1, t_2 \in \mathbb{N}$, 满足

$$|v_{t_1}| \geq |v_i|,\ |u_{t_2}| \geq |u_i|,\ i = 1, \ldots, n.$$

交换变元次序 $x_1 \leftrightarrow x_{t_1}$ 和方程次序 $f_1 \leftrightarrow f_{t_2}$.

假设理想 $I = (f_1, \ldots, f_n)$ 在 $\tilde{\mathbf{x}}$ 处, 近似局部对偶空间的一组既约基有参数化表示 $\mathrm{Span}_{\mathbb{R}}\{1, d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n, \Lambda_2, \ldots, \Lambda_{\mu-1}\}$. 我们引入 $\mu - 1$ 个光滑参数 $b_0, b_1, \ldots, b_{\mu-2}$, 并考虑增广系统:

$$G(\mathbf{x}, \mathbf{b}, \mathbf{a}) = \begin{pmatrix} F_0(\mathbf{x}, \mathbf{b}) = F(\mathbf{x}) + \left(\sum_{\nu=0}^{\mu-2} \frac{b_\nu x_1^\nu}{\nu!}\right) \mathbf{e}_1 \\ F_1(\mathbf{x}, \mathbf{b}, a_{1,2}, \ldots, a_{1,n}) \\ \vdots \\ F_{\mu-1}(\mathbf{x}, \mathbf{b}, a_{1,2}, \ldots, a_{1,n}, \ldots, a_{\mu-1,2}, \ldots, a_{\mu-1,n}) \end{pmatrix}, \tag{4.1}$$

$\mathbf{b} = (b_0, b_1, \ldots, b_{\mu-2})$, $\mathbf{a} = (a_{1,2}, \ldots, a_{1,n}, \ldots, a_{\mu-1,2}, \ldots, a_{\mu-1,n})$, $\mathbf{e}_1 = (1, 0, \ldots, 0)^T$,

$$F_k(\mathbf{x}, \mathbf{b}, a_{1,2}, \ldots, a_{1,n}, \ldots, a_{k,2}, \ldots, a_{k,n}) = \Lambda_k(F_0),\ k = 1, \ldots, \mu - 1.$$

值得注意的是, 我们将光滑参数 b_ν 乘以第一个变元的某个幂次 x_1^ν 添加到原系统的第一个方程上, 而选择第一个变元和第一个方程正是因为 Jacobian 矩阵的第一列和第一行能够被其他列行线性表示. (4.1)中, 增广系统 $G(\mathbf{x}, \mathbf{b}, \mathbf{a})$ 的方程个数 μn 等于变元个数 $n + (\mu - 1) + (n - 1)(\mu - 1)$.

定理 4.2. 如果 $(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 是增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})=\mathbf{0}$ 的正则根, 那么 $\hat{\mathbf{x}}$ 是扰动多项式系统 $\widetilde{F}(\mathbf{x}):=F_0(\mathbf{x},\hat{\mathbf{b}})=\mathbf{0}$ 的 μ 重根, 且满足 $n-\operatorname{rank}(\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}}))=1$.

证明. 由 $G(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})=\mathbf{0}$, 得到 $\widetilde{F}(\hat{\mathbf{x}})=\mathbf{0}$, 且

$$F_1(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{a}_{1,2},\ldots,\hat{a}_{1,n})=\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}})\cdot\hat{\mathbf{a}}_1=\mathbf{0},$$

$\hat{\mathbf{a}}_1=(1,\hat{a}_{1,2},\ldots,\hat{a}_{1,n})^T$. 因为 $\hat{\mathbf{a}}_1\neq\mathbf{0}$, 所以

$$\operatorname{rank}(\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}}))\leq n-1.$$

而且 $\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}})$ 的第一列能够被其他列线性表示. 因此

$$\operatorname{rank}(M)=\operatorname{rank}(\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}}))\leq n-1,$$

M 是 $\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}})$ 去掉第一列的子矩阵.

同样地, 由 $F_k(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{a}_{1,2},\ldots,\hat{a}_{1,n},\ldots,\hat{a}_{k,2},\ldots,\hat{a}_{k,n})=\mathbf{0}$ 和(2.8), 得到

$$\operatorname{rank}(\Delta_k(\widetilde{F}),M)=\operatorname{rank}(M)\leq n-1,\ k=2,\ldots,\mu-1.$$

下面证明

$$\operatorname{rank}(M)=n-1\text{ 且 }\operatorname{rank}(\Delta_\mu(\widetilde{F}),M)=n.$$

事实上, 利用等价形式

$$\frac{\partial F_k}{\partial a_{j,i}}=\frac{\partial F_{k-j}}{\partial x_i},\ i=2,\ldots,n,\ j=1,\ldots,k-1,$$

可以得到增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})$ 的 Jacobian 矩阵 $G_{\mathbf{x},\mathbf{b},\mathbf{a}}$ 的简化形式

$$G_{\mathbf{x},\mathbf{b},\mathbf{a}}=\begin{pmatrix}
F_{0,\mathbf{x}} & \mathbf{e}_1 & x_1\mathbf{e}_1 & \cdots & \frac{x_1^{\mu-3}}{(\mu-3)!}\mathbf{e}_1 & \frac{x_1^{\mu-2}}{(\mu-2)!}\mathbf{e}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\
F_{1,\mathbf{x}} & \mathbf{0} & \mathbf{e}_1 & \cdots & \frac{x_1^{\mu-4}}{(\mu-4)!}\mathbf{e}_1 & \frac{x_1^{\mu-3}}{(\mu-3)!}\mathbf{e}_1 & F_{0,\mathbf{y}} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\
F_{2,\mathbf{x}} & \mathbf{0} & \mathbf{0} & \cdots & \frac{x_1^{\mu-5}}{(\mu-5)!}\mathbf{e}_1 & \frac{x_1^{\mu-4}}{(\mu-4)!}\mathbf{e}_1 & F_{1,\mathbf{y}} & F_{0,\mathbf{y}} & \cdots & \mathbf{0} & \mathbf{0}\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\
F_{\mu-2,\mathbf{x}} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{e}_1 & F_{\mu-3,\mathbf{y}} & F_{\mu-4,\mathbf{y}} & \cdots & F_{0,\mathbf{y}} & \mathbf{0}\\
F_{\mu-1,\mathbf{x}} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & F_{\mu-2,\mathbf{y}} & F_{\mu-3,\mathbf{y}} & \cdots & F_{1,\mathbf{y}} & F_{0,\mathbf{y}}
\end{pmatrix},\quad(4.2)$$

这里 $\mathbf{y}=(x_2,\ldots,x_n)$, $F_{k,\mathbf{y}}$ 是 $F_{k,\mathbf{x}}$ 去掉第一列的子矩阵, $F_{k,\mathbf{x}}$ 是 F_k 关于变元 $\mathbf{x}$ 的 Jacobian 矩阵, $k=0,1,\ldots,\mu-1$.

假设 $\mathrm{rank}(M) \le n-2$, 则存在非零向量 $\mathbf{v}$ 满足 $M\mathbf{v}=\mathbf{0}$. 那么

$$(0,\ldots,0,\mathbf{v}^T)^T$$

是矩阵 $G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 的零向量($F_{0,\mathbf{y}}(\hat{\mathbf{x}},\hat{\mathbf{b}})=M$), 与 $(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 是 $G(\mathbf{x},\mathbf{b},\mathbf{a})=\mathbf{0}$ 的正则根矛盾. 因此

$$\mathrm{rank}(\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}}))=\mathrm{rank}(\Delta_{k-1}(\widetilde{F}),M)=\mathrm{rank}(M)=n-1.$$

另一方面, 假设 $\mathrm{rank}(\Delta_\mu(\widetilde{F}),M)=n-1$, 则存在向量 $\mathbf{v}$ 满足 $M\mathbf{v}=-\Delta_\mu(\widetilde{F})$. 由(2.8)推出

$$(1,\hat{a}_{1,2},\ldots,\hat{a}_{1,n},0,\ldots,0,2\hat{a}_{2,2},\ldots,2\hat{a}_{2,n},\ldots,(\mu-1)\hat{a}_{\mu,2},\ldots,(\mu-1)\hat{a}_{\mu,n},\mathbf{v}^T)^T.$$

是矩阵 $G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 的零向量, 同样与 $(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 是 $G(\mathbf{x},\mathbf{b},\mathbf{a})=\mathbf{0}$ 的正则根矛盾. 因此

$$\mathrm{rank}(\Delta_\mu(\widetilde{F}),M)=n.$$

最后由定理 2.13, 我们得到 $\hat{\mathbf{x}}$ 是扰动多项式系统 $\widetilde{F}(\mathbf{x})=\mathbf{0}$ 的 μ 重根. □

根据定理 4.2, 如果能够验证增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})=\mathbf{0}$ 正则根 $(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 的存在性, 那么就可以保证一个带有微小扰动的多项式系统 $\widetilde{F}(\mathbf{x})=\mathbf{0}$ 有一个宽度为 1 的孤立奇异根 $\hat{\mathbf{x}}$. 下面证明, 如果多项式系统 $F(\mathbf{x})=\mathbf{0}$ 有一个宽度为 1 的孤立奇异根 $\hat{\mathbf{x}}$, 那么满足 $G(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})=\mathbf{0}$ 和 $G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$ 非奇异的增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})$ 必存在.

定理 4.3. *假设 $\hat{\mathbf{x}}$ 是多项式系统 $F(\mathbf{x})=\mathbf{0}$ 的一个宽度为 1 的孤立奇异根, 而且满足 Jacobian 矩阵 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的第一列和第一行能够分别被其他列和其他行线性表示. 那么存在形如(4.1)的增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})$ 和它的一个正则根 $(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$.*

证明. 由 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的第一列和第一行能够分别被其他列和其他行线性表示, 推出

$$\mathrm{rank}(M,\mathbf{e}_1)=n,$$

M 是 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 去掉第一列的子矩阵, $\mathbf{e}_1=(1,0,\ldots,0)^T$.

假设理想 $I=(f_1,\ldots,f_n)$ 在 $\hat{\mathbf{x}}$ 处局部对偶空间的一组既约基有参数化表示

$$\mathrm{Span}_{\mathbb{R}}\{1,d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n,\Lambda_2,\ldots,\Lambda_{\mu-1}\}.$$

构造形如(4.1)的增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})$. 显然 $\exists\,\hat{\mathbf{a}}\in\mathbb{R}^{(n-1)(\mu-1)}$ 满足

$$G(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})=\mathbf{0},$$

$\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\{1,d_1+\hat{a}_{1,2}d_2+\cdots+\hat{a}_{1,n}d_n,\Lambda_2,\ldots,\Lambda_{\mu-1}\}$.

下面证明 $G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$ 非奇异.

假设 $\mathrm{rank}(G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}))<\mu n$, 存在非零向量 $\mathbf{v}$, 满足 $G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})\mathbf{v}=\mathbf{0}$. 由 $\mathrm{rank}(M,\mathbf{e}_1)=n$ 和(4.2)的第一行, 得到 $v_1\neq 0$, 不妨设 $v_1=1$.

类似于定理 4.2 的证明, 我们推出 $M(v_{\mu n-n+2},\ldots,v_{\mu n})^T=-\Delta_\mu(F)$. 因此

$$\mathrm{rank}(\Delta_\mu(F),M)=n-1.$$

根据定理 2.13, 这与 $\dim(\mathcal{D}_{\hat{\mathbf{x}}}(I))=\mu$ 矛盾. 所以, $G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$ 不存在非零零向量, 即 $G_{\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$ 非奇异. □

对于给定的多项式系统 F 和近似奇异根 $\tilde{\mathbf{x}}$, 我们首先利用宽度为 1 的特殊情形下重结构的参数化表示计算

$$\tilde{\mathbf{a}}=(\tilde{a}_{1,2},\ldots,\tilde{a}_{1,n},\ldots,\tilde{a}_{\mu-1,2},\ldots,\tilde{a}_{\mu-1,n})$$

和构造形如(4.1)的增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})$, 然后利用定理 4.1 验证近似根 $(\tilde{\mathbf{x}},\mathbf{0},\tilde{\mathbf{a}})$ 附近的正则根 $(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$, 满足 $G(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})=\mathbf{0}$.

定理 4.4. *假设定理 4.1 适用于增广系统 $G(\mathbf{x},\mathbf{b},\mathbf{a})$ 和近似根 $(\tilde{\mathbf{x}},\mathbf{0},\tilde{\mathbf{a}})$, 并且得到正则根 $(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 的可信误差界. 那么 $\hat{\mathbf{x}}$ 是扰动多项式系统 $\widetilde{F}(\mathbf{x}):=F_0(\mathbf{x},\hat{\mathbf{b}})$ 的孤立奇异根, 满足*

$$\mathrm{rank}(\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}}))=n-1,$$

且 $\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\{1,d_1+\hat{a}_{1,2}d_2+\cdots+\hat{a}_{1,n}d_n,\Lambda_2,\ldots,\Lambda_{\mu-1}\}$.

证明. 定理可以由定理 4.2 和定理 4.3 直接得到. □

例 4.4 (例 2.11 的延续). 对于 $F=\{x_1^2+x_2-3,x_1+0.125x_2^2-1.5\}$ 和它的一个近似根 $\tilde{\mathbf{x}}=(1.01,2.01)$, 利用宽度为 1 的特殊情形下重结构的参数化表示

$$\mathcal{D}_{\hat{\mathbf{x}}}(I)=\mathrm{Span}_{\mathbb{R}}\{1,d_1+a_1d_2,d_1^2+a_1d_1d_2+a_1^2d_2^2+a_2d_2\},$$

得到 $\tilde{a}_1 = -2.014$, $\tilde{a}_2 = -1.002$ 和构造

$$G(x_1, x_2, b_0, b_1, a_1, a_2) = \begin{pmatrix} x_1^2 + x_2 - 3 + b_0 + b_1 x_1 \\ x_1 + 0.125 x_2^2 - 1.5 \\ 2x_1 + a_1 + b_1 \\ 1 + 0.25 a_1 x_2 \\ 1 + a_2 \\ 0.125 a_1^2 + 0.25 a_2 x_2 \end{pmatrix}.$$

利用 INTLAB 中的算法 verifynlss, 得到可信误差界

$\gg f = \text{inline}('[x(1)^2 + x(2) - 3 + x(3) + x(1) * x(4); x(1) + 0.125 * x(2)^2 - 1.5;$

$2 * x(1) + x(4) + x(5); 1 + 0.25 * x(2) * x(5);$

$1 + x(6); 0.125 * x(5)^2 + 0.25 * x(2) * x(6)]')$

$f =$

Inline function :

$f(x) = [x(1)^2 + x(2) - 3 + x(3) + x(1) * x(4); x(1) + 0.125 * x(2)^2 - 1.5;$

$2 * x(1) + x(4) + x(5); 1 + 0.25 * x(2) * x(5);$

$1 + x(6); 0.125 * x(5)^2 + 0.25 * x(2) * x(6)]$

$\gg$ verifynlss(f, [1.01; 2.01; 0; 0; −2.014; −1.002])

intval ans =

[0.99999999999999, 1.00000000000001]

[1.99999999999999, 2.00000000000001]

[−0.00000000000001, 0.00000000000001]

[−0.00000000000001, 0.00000000000001]

[−2.00000000000001, −1.99999999999999]

[−1.00000000000000, −1.00000000000000]

上述区间向量内, 存在增广系统 G 唯一的一个正则根 $(\hat{\mathbf{x}}, \hat{\mathbf{b}}, \hat{\mathbf{a}})$. 最后根据定理 4.4, 扰动多项式系统 $\widetilde{F} = \{x_1^2 + x_2 - 3 + \hat{b}_0 + \hat{b}_1 x_1, x_1 + 0.125 x_2^2 - 1.5\}$ 在扰动误差界

$$-0.00000000000001 \le \hat{b}_0 \le 0.00000000000001$$
$$-0.00000000000001 \le \hat{b}_1 \le 0.00000000000001$$

内存在唯一的孤立奇异根 $(\hat{x}_1, \hat{x}_2)$, 而且

$$\mathcal{D}_{\hat{\mathbf{x}}}(I) = \mathrm{Span}_{\mathbb{R}}\{1, d_1 + \hat{a}_1 d_2, d_1^2 + \hat{a}_1 d_1 d_2 + \hat{a}_1^2 d_2^2 + \hat{a}_2 d_2\}.$$

注意到宽度为 1 的特殊情形下, 光滑参数添加的形式(乘以某个变元的幂次)和位置(原系统的某个方程), 是由 Jacobian 矩阵中能够被其他列和其他行线性表示的“某列”和“某行”决定的. 受此启发, 下面对于一般的孤立奇异根, 我们提出一种带光滑参数的 deflation 技术, 并且基于这种技术, 我们将验证宽度为 1 孤立奇异根的方法推广到一般情形.

4.3 带光滑参数的 deflation 技术

假设 $\hat{\mathbf{x}} \in \mathbb{R}^n$ 是 $F = \{f_1, \ldots, f_n\}$ 的孤立奇异根, $f_i \in \mathbb{R}[\mathbf{x}]$, 满足

$$\mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = n - d, \ (1 < d \leq n).$$

第一, 必然存在 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的某 d 列能够被其他列线性表示, 记 $\mathbf{c} = \{c_1, \ldots, c_d\}$ $(1 \leq c_1 < \ldots < c_d \leq n)$ 是它们的指标集合, $F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}})$ 是 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 去掉第 $c_1, \ldots, c_d$ 列的子矩阵, 那么

$$\mathrm{rank}(F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}})) = n - d. \tag{4.3}$$

第二, 必然存在 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的某 d 行能够被其他行线性表示, 记 $\mathbf{k} = \{k_1, \ldots, k_d\}$ $(1 \leq k_1 < \ldots < k_d \leq n)$ 是它们的指标集合, 那么

$$\mathrm{rank}(F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}}), I_{\mathbf{k}}) = n, \tag{4.4}$$

$I_{\mathbf{k}} = (\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_d})$, $\mathbf{e}_{k_i} = (0, \ldots, 0, \overset{k_i}{1}, 0, \ldots, 0)^T$.

基于(4.3)和(4.4)的假设, 下面我们讨论一种带光滑参数的 deflation 技术.

4.3.1 一阶 deflation

我们引入 d 个光滑参数 $b_{0,1}, \ldots, b_{0,d}$, 并考虑增广系统:

$$G(\mathbf{x}, \mathbf{b}_0, \mathbf{a}_1) = \begin{pmatrix} F(\mathbf{x}) + \sum_{i=1}^{d} b_{0,i} \mathbf{e}_{k_i} \\ F_{\mathbf{x}}(\mathbf{x}) \mathbf{v}_1 \end{pmatrix}, \tag{4.5}$$

$\mathbf{b}_0 = (b_{0,1}, \ldots, b_{0,d})$, $\mathbf{a}_1 = (a_{1,1}, \ldots, a_{1,n-d})$, $\mathbf{v}_1 = (a_{1,1}, \ldots, \overset{c_1}{1}, \ldots, \overset{c_d}{1}, \ldots, a_{1,n-d})^T$. 显然增广系统 $G(\mathbf{x}, \mathbf{b}_0, \mathbf{a}_1)$ 的方程个数 $2n$ 等于变元个数 $n + d + (n - d)$.

事实上, 如果我们不考虑(4.5)中的光滑参数, 增广系统

$$\widetilde{G}(\mathbf{x},\mathbf{a}_1)=\begin{pmatrix}F(\mathbf{x})\\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1\end{pmatrix} \tag{4.6}$$

等价于(3.3)中的一次 deflation. 由(4.3), 存在唯一的 $\hat{\mathbf{a}}_1\in\mathbb{R}^{n-d}$, 使得 $(\hat{\mathbf{x}},\hat{\mathbf{a}}_1)$ 是增广系统 $\widetilde{G}(\mathbf{x},\mathbf{a}_1)=\mathbf{0}$ 的根. 显然, $(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}_1)$ 是 $G(\mathbf{x},\mathbf{b}_0,\mathbf{a}_1)=\mathbf{0}$ 的根.

定理 4.5. *比较(4.5)和(4.6)中两个增广系统的 Jacobian 矩阵,*

$$\mathrm{null}\left(G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}_1)\right)=\left\{\begin{pmatrix}\mathbf{y}_1\\ \mathbf{0}\\ \mathbf{y}_2\end{pmatrix}\in\mathbb{R}^{2n}\mid\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}\in\mathrm{null}\left(\widetilde{G}_{\mathbf{x},\mathbf{a}_1}(\hat{\mathbf{x}},\hat{\mathbf{a}}_1)\right)\right\}.$$

证明. 计算(4.5)和(4.6)中两个增广系统的 Jacobian 矩阵

$$G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}_1)=\begin{pmatrix}F_{\mathbf{x}}(\hat{\mathbf{x}}) & I_{\mathbf{k}} & \mathcal{O}_{n,n-d}\\ F_{\mathbf{xx}}(\hat{\mathbf{x}})\hat{\mathbf{v}}_1 & \mathcal{O}_{n,d} & F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}})\end{pmatrix},$$

$$\widetilde{G}_{\mathbf{x},\mathbf{a}_1}(\hat{\mathbf{x}},\hat{\mathbf{a}}_1)=\begin{pmatrix}F_{\mathbf{x}}(\hat{\mathbf{x}}) & \mathcal{O}_{n,n-d}\\ F_{\mathbf{xx}}(\hat{\mathbf{x}})\hat{\mathbf{v}}_1 & F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}})\end{pmatrix},$$

$F_{\mathbf{xx}}$ 是 F 的 Hessian 矩阵, $\mathcal{O}_{i,j}$ 是 $i\times j$ 阶零矩阵. 若 $\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}\in\mathrm{null}\left(\widetilde{G}_{\mathbf{x},\mathbf{a}_1}(\hat{\mathbf{x}},\hat{\mathbf{a}}_1)\right)$, 则 $\begin{pmatrix}\mathbf{y}_1\\ \mathbf{0}\\ \mathbf{y}_2\end{pmatrix}\in\mathrm{null}\left(G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}_1)\right)$.

另一方面, 假设 $\begin{pmatrix}\mathbf{y}_1\\ \mathbf{z}\\ \mathbf{y}_2\end{pmatrix}\in\mathrm{null}\left(G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}_1)\right)$, 那么

$$F_{\mathbf{x}}(\hat{\mathbf{x}})\mathbf{y}_1+I_{\mathbf{k}}\mathbf{z}=\mathbf{0}.$$

由 $\mathrm{rank}(F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}}),I_{\mathbf{k}})=n$, 显然 $\mathbf{z}=\mathbf{0}$. □

根据定理 4.5, 与(4.6)中的一次 deflation 相比, (4.5)中带光滑参数的 deflation 得到的增广系统不仅变元个数等于方程个数, 而且保留了 Jacobian 矩阵零空间的所有信息. 因此根据定理 3.1, 重复至多 ρ_I 次(4.5)中带光滑参数的 deflation 后, 可以得到一个增广系统(变元个数等于方程个数)和它的一个正则根.

4.3.2 二阶 deflation

假设 $(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)$ 仍然是 $G(\mathbf{x}, \mathbf{b}_0, \mathbf{a}_1)$ 的孤立奇异根, 满足

$$\operatorname{rank}(G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)) = 2n - d', \ (1 < d' \leq 2n).$$

记 $\mathbf{c}' = \{c'_1, \ldots, c'_{d'}\}$ 和 $\mathbf{k}' = \{k'_1, \ldots, k'_{d'}\}$ 是满足

$$\operatorname{rank}(G^{\mathbf{c}'}_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)) = 2n - d'. \tag{4.7}$$

$$\operatorname{rank}(G^{\mathbf{c}'}_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1), I_{\mathbf{k}'+n}) = 2n, \tag{4.8}$$

的指标集合, $G^{\mathbf{c}'}_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)$ 是 $G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)$ 去掉第 $c'_1, \ldots, c'_{d'}$ 列的子矩阵,

$$I_{\mathbf{k}'+n} = \begin{pmatrix} \mathcal{O}_{n,d'} \\ I_{\mathbf{k}'} \end{pmatrix}, \ I_{\mathbf{k}'} = (\mathbf{e}_{k'_1}, \mathbf{e}_{k'_2}, \ldots, \mathbf{e}_{k'_{d'}}).$$

定理 4.6. *存在指标集合 $\mathbf{c}' \subseteq \mathbf{c}$ 和 $\mathbf{k}' \subseteq \mathbf{k}$ 分别满足(4.7)和(4.8).*

证明. 假设

$$G^{\mathbf{c}}_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1) = \begin{pmatrix} F^{\mathbf{c}}_{\mathbf{x}}(\hat{\mathbf{x}}) & I_{\mathbf{k}} & \mathcal{O}_{n,n-d} \\ \star & \mathcal{O}_{n,d} & F^{\mathbf{c}}_{\mathbf{x}}(\hat{\mathbf{x}}) \end{pmatrix},$$

是 $G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)$ 去掉第 $c_1, \ldots, c_d$ 列的子矩阵. 由(4.3)和(4.4)推出

$$\operatorname{rank}(G^{\mathbf{c}}_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)) = 2n - d.$$

因此 $d' \leq d$, 且存在指标集合 $\mathbf{c}' \subseteq \mathbf{c}$ 满足(4.7).

根据 (4.4), 显然

$$\operatorname{rank}(G^{\mathbf{c}}_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1), I_{\mathbf{k}+n}) = 2n,$$

$I_{\mathbf{k}+n} = \begin{pmatrix} \mathcal{O}_{n,d} \\ I_{\mathbf{k}} \end{pmatrix}$. 因此存在指标集合 $\mathbf{k}' \subseteq \mathbf{k}$ 满足(4.8). □

如果 $(\hat{\mathbf{x}}, \mathbf{0}, \hat{\mathbf{a}}_1)$ 依然是 $G(\mathbf{x}, \mathbf{b}_0, \mathbf{a}_1)$ 的孤立奇异根, 我们再做一步(4.5)中带光滑参数的 deflation. 根据定理 4.6, 得到新的增广系统:

$$H(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1, \mathbf{a}_1, \mathbf{a}_2) = \begin{pmatrix} F(\mathbf{x}) + I_{\mathbf{k}}\mathbf{b}_0 \\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1 + I_{\mathbf{k}'}\mathbf{b}_1 \\ G_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\mathbf{x}, \mathbf{b}_0, \mathbf{a}_1)\mathbf{v}_2 \end{pmatrix} \tag{4.9}$$

$\mathbf{b}_1=(b_{1,1},\ldots,b_{1,d'})$, $\mathbf{a}_2=(a_{2,1},\ldots,a_{2,2n-d'})$,

$$\mathbf{v}_2=(a_{2,1},\ldots,\overset{c'_1}{1},\ldots,\overset{c'_{d'}}{1},\ldots,a_{2,2n-d'})^T.$$

值得注意的是, 增广系统 $H(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2)$ 中新的光滑参数 $\mathbf{b}_1$ 并没有添加到原系统 F 上, 而是添加到 deflation 系统 $F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1$ 上. 下面我们考虑改进的增广系统:

$$\widetilde{H}(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2)=\begin{pmatrix} F(\mathbf{x})+X_0\mathbf{b}_0+X_1\mathbf{b}_1 \\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1+I_{\mathbf{k}'}\mathbf{b}_1 \\ \widetilde{G}_{\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1)\mathbf{v}_2 \end{pmatrix}, \tag{4.10}$$

$X_0=I_{\mathbf{k}}$, $X_1=(x_{c'_1}\mathbf{e}_{k'_1},\ldots,x_{c'_{d'}}\mathbf{e}_{k'_{d'}})$,

$$\widetilde{G}(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1)=\begin{pmatrix} F(\mathbf{x})+X_0\mathbf{b}_0+X_1\mathbf{b}_1 \\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1+I_{\mathbf{k}'}\mathbf{b}_1 \end{pmatrix}.$$

根据(4.7), 存在唯一的 $\hat{\mathbf{a}}_2\in\mathbb{R}^{2n-d'}$, 使得 $(\hat{\mathbf{x}},\mathbf{0},\mathbf{0},\hat{\mathbf{a}}_1,\hat{\mathbf{a}}_2)$ 是 $H(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2)$ 的根, 也是 $\widetilde{H}(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2)$ 的根.

定理 4.7. *比较(4.9)和(4.10)中两个增广系统的 Jacobian 矩阵,*

$$\operatorname{null}\left(H_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}},\mathbf{0},\mathbf{0},\hat{\mathbf{a}}_1,\hat{\mathbf{a}}_2)\right)=\operatorname{null}\left(\widetilde{H}_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}},\mathbf{0},\mathbf{0},\hat{\mathbf{a}}_1,\hat{\mathbf{a}}_2)\right).$$

证明. 计算(4.9)中增广系统的 Jacobian 矩阵 $H_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}},\mathbf{0},\mathbf{0},\hat{\mathbf{a}}_1,\hat{\mathbf{a}}_2)$

$$\begin{pmatrix} F_{\mathbf{x}}(\hat{\mathbf{x}}) & I_{\mathbf{k}} & \mathcal{O}_{n,d'} & \mathcal{O}_{n,n-d} & \multicolumn{3}{c}{\mathcal{O}_{2n,2n-d'}} \\ F_{\mathbf{xx}}(\hat{\mathbf{x}})\hat{\mathbf{v}}_1 & \mathcal{O}_{n,d} & I_{\mathbf{k}'} & F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}}) & & & \\ \star & \mathcal{O}_{n,d} & \mathcal{O}_{n,d'} & \star & F_{\mathbf{x}}^{\mathbf{c}'}(\hat{\mathbf{x}}) & I_{\mathbf{k}} & \mathcal{O}_{n,n-d} \\ \star & \mathcal{O}_{n,d} & \mathcal{O}_{n,d'} & \star & \star & \mathcal{O}_{n,d} & F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}}) \end{pmatrix},$$

和(4.10)中增广系统的 Jacobian 矩阵 $\widetilde{H}_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}},\mathbf{0},\mathbf{0},\hat{\mathbf{a}}_1,\hat{\mathbf{a}}_2)$

$$\begin{pmatrix} F_{\mathbf{x}}(\hat{\mathbf{x}}) & I_{\mathbf{k}} & \hat{X}_1 & \mathcal{O}_{n,n-d} & \multicolumn{3}{c}{\mathcal{O}_{2n,2n-d'}} \\ F_{\mathbf{xx}}(\hat{\mathbf{x}})\hat{\mathbf{v}}_1 & \mathcal{O}_{n,d} & I_{\mathbf{k}'} & F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}}) & & & \\ \star & \mathcal{O}_{n,d} & I_{\mathbf{k}'} & \star & F_{\mathbf{x}}^{\mathbf{c}'}(\hat{\mathbf{x}}) & I_{\mathbf{k}} & \mathcal{O}_{n,n-d} \\ \star & \mathcal{O}_{n,d} & \mathcal{O}_{n,d'} & \star & \star & \mathcal{O}_{n,d} & F_{\mathbf{x}}^{\mathbf{c}}(\hat{\mathbf{x}}) \end{pmatrix}.$$

$\hat{X}_1 = (\hat{x}_{c'_1}\mathbf{e}_{k'_1}, \ldots, \hat{x}_{c'_{d'}}\mathbf{e}_{k'_{d'}})$.

类似于定理 4.5, 对 $\forall\,(\mathbf{y}_1, \mathbf{z}_1, \mathbf{z}_2, \mathbf{y}_2, \mathbf{y}_3)^T \in \mathrm{null}\,(H_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2))$, 必然满足 $\mathbf{z}_1 = \mathbf{0}$, $\mathbf{z}_2 = \mathbf{0}$. 显然

$$\mathrm{null}\,(H_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2)) \subseteq \mathrm{null}\left(\widetilde{H}_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2)\right).$$

由于$\mathbf{k}' \subseteq \mathbf{k}$, 通过若干次初等列变换我们可以将 $\widetilde{H}_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2)$ 变成 $H_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2)$, 因此它们的秩相等. 所以,

$$\mathrm{null}\,(H_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2)) = \mathrm{null}\left(\widetilde{H}_{\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\hat{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2)\right).$$

□

根据定理 4.7, 与(4.9)中的 deflation 相比, (4.10)中改进的 deflation 不仅将新的光滑参数 $\mathbf{b}_1$ 也添加到原系统 F 上, 而且保持了增广系统 Jacobian 矩阵零空间不变. 事实上, 如果记 $\widetilde{F}(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1) := F(\mathbf{x}) + X_0\mathbf{b}_0 + X_1\mathbf{b}_1$, 那么

$$\begin{aligned}
\widetilde{F}_{\mathbf{x}}(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1)\mathbf{v}_1 &= F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1 + (\mathbf{0}, \cdots, \overset{c'_1}{b_{1,1}\mathbf{e}_{k'_1}}, \cdots, \overset{c'_{d'}}{b_{1,d'}\mathbf{e}_{k'_{d'}}}, \cdots, \mathbf{0})\,\mathbf{v}_1 \\
&= F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1 + (\mathbf{e}_{k'_1}, \ldots, \mathbf{e}_{k'_{d'}})\,\mathbf{b}_1\;(\mathbf{c}' \subseteq \mathbf{c}) \\
&= F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1 + I_{\mathbf{k}'}\mathbf{b}_1
\end{aligned}$$

对于给定的多项式系统 F 和近似奇异根 $\tilde{\mathbf{x}}$, 如果我们能够利用(4.10)中改进的 deflation 得到增广系统 $\widetilde{H}(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1, \mathbf{a}_1, \mathbf{a}_2)$ 和近似根 $(\tilde{\mathbf{x}}, \mathbf{0}, \mathbf{0}, \tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2)$, 然后利用定理 4.1 得到 $\widetilde{H}(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1, \mathbf{a}_1, \mathbf{a}_2)$ 正则根 $(\hat{\mathbf{x}}, \hat{\mathbf{b}}_0, \hat{\mathbf{b}}_1, \hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2)$ 的可信误差界. 那么, $\hat{\mathbf{x}}$ 是扰动多项式系统 $\widetilde{F}(\mathbf{x}, \hat{\mathbf{b}}_0, \hat{\mathbf{b}}_1)$ 的孤立奇异根, 因为

$$\widetilde{F}(\hat{\mathbf{x}}, \hat{\mathbf{b}}_0, \hat{\mathbf{b}}_1) = \widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}}, \hat{\mathbf{b}}_0, \hat{\mathbf{b}}_1)\hat{\mathbf{v}}_1 = \mathbf{0}.$$

例 4.5. [7, DZ1] 考虑如下例子:

$$F = \{x_1^4 - x_2x_3x_4, x_2^4 - x_1x_3x_4, x_3^4 - x_1x_2x_4, x_4^4 - x_1x_2x_3\}.$$

有一个 131 重根 $(0, 0, 0, 0)$.

对于近似根 $\tilde{\mathbf{x}} = (0.01, 0.01, 0.01, 0.01)$, F 的 Jacobian 矩阵

$$F_{\mathbf{x}}(\tilde{\mathbf{x}}) = \begin{pmatrix} 0.4\times10^{-5} & -0.0001 & -0.0001 & -0.0001 \\ -0.0001 & 0.4\times10^{-5} & -0.0001 & -0.0001 \\ -0.0001 & -0.0001 & 0.4\times10^{-5} & -0.0001 \\ -0.0001 & -0.0001 & -0.0001 & 0.4\times10^{-5} \end{pmatrix}.$$

近似满足 $d=4$, $\mathbf{c}=\mathbf{k}=\{1,2,3,4\}$. 因此得到增广系统

$$G(\mathbf{x},\mathbf{b}_0)=\begin{pmatrix} F(\mathbf{x})+I_{\mathbf{k}}\mathbf{b}_0 \\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1 \end{pmatrix}$$

和 $\tilde{\mathbf{b}}_0=(0,0,0,0)$, $\mathbf{v}_1=(1,1,1,1)^T$.

G 的 Jacobian 矩阵

$$G_{\mathbf{x},\mathbf{b}_0}(\tilde{\mathbf{x}},\tilde{\mathbf{b}}_0)=\begin{pmatrix} F_{\mathbf{x}}(\tilde{\mathbf{x}}) & I_{\mathbf{k}} \\ F_{\mathbf{xx}}(\tilde{\mathbf{x}})\mathbf{v}_1 & \mathcal{O}_{4,4} \end{pmatrix},$$

$$F_{\mathbf{xx}}(\tilde{\mathbf{x}})\mathbf{v}_1=\begin{pmatrix} 0.0012 & -0.02 & -0.02 & -0.02 \\ -0.02 & 0.0012 & -0.02 & -0.02 \\ -0.02 & -0.02 & 0.0012 & -0.02 \\ -0.02 & -0.02 & -0.02 & 0.0012 \end{pmatrix}$$

近似满足 $d'=4$, $\mathbf{c}'=\mathbf{k}'=\{1,2,3,4\}$. 因此根据(4.10), 得到增广系统

$$H(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_2)=\begin{pmatrix} F(\mathbf{x})+X_0\mathbf{b}_0+X_1\mathbf{b}_1 \\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1+I_{\mathbf{k}'}\mathbf{b}_1 \\ \widetilde{G}_{\mathbf{x},\mathbf{b}_0}(\mathbf{x},\mathbf{b}_0,\mathbf{b}_1)\mathbf{v}_2 \end{pmatrix}$$

和近似根 $\tilde{\mathbf{b}}_1=(0,0,0,0)$, $\tilde{\mathbf{a}}_2=(0.0003,0.0003,0.0003,0.0003)$,

$$\mathbf{v}_2=(1,1,1,1,a_{2,1},a_{2,2},a_{2,3},a_{2,4})^T.$$

利用 INTLAB 中的算法 verifynlss, 得到可信误差界

$$-10^{-321}\le \hat{x}_i,\ \hat{b}_{0,i},\ \hat{b}_{1,i},\ \hat{a}_{2,i}\le 10^{-321},$$

$i=1,2,3,4$. 因此 $(\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{x}_4)$ 是扰动多项式系统

$$\widetilde{F}(\mathbf{x},\hat{\mathbf{b}}_0,\hat{\mathbf{b}}_1)=\begin{pmatrix} x_1^4-x_2x_3x_4+\hat{b}_{0,1}+\hat{b}_{1,1}x_1 \\ x_2^4-x_1x_3x_4+\hat{b}_{0,2}+\hat{b}_{1,2}x_2 \\ x_3^4-x_1x_2x_4+\hat{b}_{0,3}+\hat{b}_{1,3}x_3 \\ x_4^4-x_1x_2x_3+\hat{b}_{0,4}+\hat{b}_{1,4}x_4 \end{pmatrix}$$

的孤立奇异根.

由上面的例子可以看出, 正因为(4.10)中改进的 deflation 将原本应该添加到 deflation 系统 $F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1$ 上的光滑参数 $\mathbf{b}_1$, 以 $X_1\mathbf{b}_1$ 的形式添加到原系统 F 上, 所以成功验证了扰动多项式系统 $\widetilde{F}(\mathbf{x},\hat{\mathbf{b}}_0,\hat{\mathbf{b}}_1)$ 的孤立奇异根. 下面我们将这种改进的 deflation 推广到更高阶情形.

4.3.3 高阶 deflation

我们现在递归地给出改进的高阶 deflation:

假设 $(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$ 仍然是 $G_t(\mathbf{x},\mathbf{b},\mathbf{a})$ 的孤立奇异根,

$$\mathbf{b}=(\mathbf{b}_0,\ldots,\mathbf{b}_{t-1}),\ \mathbf{a}=(\mathbf{a}_1,\ldots,\mathbf{a}_t),\ t>0,$$

满足

$$\operatorname{rank}(G_{t,\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})=2^t n-d^{(t)},\ (1<d^{(t)}\le 2^t n).$$

记 $\mathbf{c}^{(t)}=\{c_1^{(t)},\ldots,c_{d^{(t)}}^{(t)}\}$ 和 $\mathbf{k}^{(t)}=\{k_1^{(t)},\ldots,k_{d^{(t)}}^{(t)}\}$ 是满足

$$\operatorname{rank}(G^{\mathbf{c}^{(t)}}_{t,\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}))=2^t n-d^{(t)}. \tag{4.11}$$

$$\operatorname{rank}(G^{\mathbf{c}^{(t)}}_{t,\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}}),I_{\mathbf{k}^{(t)}+2^t n-n})=2^t n, \tag{4.12}$$

的指标集合, $G^{\mathbf{c}^{(t)}}_{t,\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$ 是 $G_{t,\mathbf{x},\mathbf{b},\mathbf{a}}(\hat{\mathbf{x}},\mathbf{0},\hat{\mathbf{a}})$ 去掉第 $c_1^{(t)},\ldots,c_{d^{(t)}}^{(t)}$ 列的余子式,

$$I_{\mathbf{k}^{(t)}+2^t n-n}=\begin{pmatrix}\mathcal{O}_{2^t n-n,d^{(t)}}\\ I_{\mathbf{k}^{(t)}}\end{pmatrix},\ I_{\mathbf{k}^{(t)}}=(\mathbf{e}_{k_1^{(t)}},\mathbf{e}_{k_2^{(t)}},\ldots,\mathbf{e}_{k_{d^{(t)}}^{(t)}}).$$

我们考虑增广系统:

$$G_{t+1}(\mathbf{x},\mathbf{b},\mathbf{b}_t,\mathbf{a},\mathbf{a}_{t+1})=\begin{pmatrix}F(\mathbf{x})+\sum_{i=0}^t X_i\mathbf{b}_i\\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1+\sum_{i=1}^t X_i'\mathbf{b}_i\\ G_{1,\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\mathbf{x},\mathbf{b},\mathbf{b}_{t+1},\mathbf{a}_1)\mathbf{v}_2\\ G_{2,\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\mathbf{x},\mathbf{b},\mathbf{b}_{t+1},\mathbf{a}_1,\mathbf{a}_2)\mathbf{v}_3\\ \vdots\\ G_{t,\mathbf{x},\mathbf{b},\mathbf{a}}(\mathbf{x},\mathbf{b},\mathbf{b}_{t+1},\mathbf{a})\mathbf{v}_{t+1}\end{pmatrix}, \tag{4.13}$$

$X_i=\frac{1}{i!}\left(x^i_{c_1^{(i)}}\mathbf{e}_{k_1^{(i)}},\ldots,x^i_{c_{d^{(i)}}^{(i)}}\mathbf{e}_{k_{d^{(i)}}^{(i)}}\right)$, $X_i'=\frac{1}{(i-1)!}\left(x^{(i-1)}_{c_1^{(i)}}\mathbf{e}_{k_1^{(i)}},\ldots,x^{(i-1)}_{c_{d^{(i)}}^{(i)}}\mathbf{e}_{k_{d^{(i)}}^{(i)}}\right)$,

$$G_j(\mathbf{x},\mathbf{b},\mathbf{b}_{t+1},\mathbf{a}_1,\ldots,\mathbf{a}_j)=\begin{pmatrix}F(\mathbf{x})+\sum_{i=0}^t X_i\mathbf{b}_i\\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1+\sum_{i=1}^t X_i'\mathbf{b}_i\\ G_{1,\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\mathbf{x},\mathbf{b},\mathbf{b}_{t+1},\mathbf{a}_1)\mathbf{v}_2\\ \vdots\\ G_{j-1,\mathbf{x},\mathbf{b}_0,\ldots,\mathbf{b}_{j-2},\mathbf{a}_1,\ldots,\mathbf{a}_{j-1}}(\mathbf{x},\mathbf{b},\mathbf{b}_{t+1},\mathbf{a}_1,\ldots,\mathbf{a}_{j-1})\mathbf{v}_j\end{pmatrix}.$$

4.3.4 有限终止性

如果 $\hat{\mathbf{x}}$ 是 F 的孤立奇异根, 那么(4.13)中改进的 deflation 是有限终止的.

定理 4.8. *假设 $\hat{\mathbf{x}} \in \mathbb{R}^n$ 是 $F=\{f_1,\ldots,f_n\}$ 的孤立奇异根, $f_i \in \mathbb{R}[\mathbf{x}]$, 那么重复至多 ρ_I 次(4.13)中改进的 deflation 后, 可以得到一个增广系统和它的一个正则根.*

证明. 首先根据定理 4.7, 可以归纳地证明出, 利用(4.5)和(4.13)中的 deflation 得到的增广系统 Jacobian 矩阵零空间总保持不变; 然后根据定理 4.5, 可以归纳地证明出, 利用(4.5)和(4.6)中的 deflation 得到的增广系统 Jacobian 矩阵零空间的维数总相等; 最后根据定理 3.1, 重复至多 ρ_I 次(4.6)中的 deflation 后, 可以得到一个增广系统和它的一个正则根, 所以定理成立. □

4.3.5 一般的孤立奇异根的验证

定理 4.9. *存在指标集合 $\mathbf{c}^{(t)} \subseteq \mathbf{c}^{(t-1)}$ 和 $\mathbf{k}^{(t)} \subseteq \mathbf{k}^{(t-1)}$ 分别满足(4.11)和(4.12).*

证明. 根据定理 4.6, 4.7 可以归纳地得到. □

对于给定的多项式系统 F 和近似奇异根 $\tilde{\mathbf{x}}$, 重复 t 次(4.13)中改进的 deflation 后, 我们得到增广系统 $G_t(\mathbf{x},\mathbf{b},\mathbf{a})$ 和近似根 $(\tilde{\mathbf{x}},\mathbf{0},\tilde{\mathbf{a}})$.

定理 4.10. *假设定理 4.1 适用于增广系统 $G_t(\mathbf{x},\mathbf{b},\mathbf{a})$ 和近似根 $(\tilde{\mathbf{x}},\mathbf{0},\tilde{\mathbf{a}})$, 并且得到正则根 $(\hat{\mathbf{x}},\hat{\mathbf{b}},\hat{\mathbf{a}})$ 的可信误差界. 那么 $\hat{\mathbf{x}}$ 是扰动多项式系统*

$$\widetilde{F}(\mathbf{x},\hat{\mathbf{b}}) := F(\mathbf{x}) + \sum_{i=0}^{t} X_i \hat{\mathbf{b}}_i$$

的孤立奇异根, 满足

$$\operatorname{rank}(\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}})) < n.$$

证明. 显然 $\widetilde{F}(\hat{\mathbf{x}},\hat{\mathbf{b}})=\mathbf{0}$. 根据定理 4.9, 得到

$$\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}},\hat{\mathbf{b}})\hat{\mathbf{v}}_1 = F_{\mathbf{x}}(\mathbf{x})\hat{\mathbf{v}}_1 + \sum_{i=1}^{t} \hat{X}_i' \hat{\mathbf{b}}_i = \mathbf{0}.$$

再由 $\hat{\mathbf{v}}_1 \neq \mathbf{0}$, 推出 $\operatorname{rank}(\widetilde{F}_{\mathbf{x}}(\hat{\mathbf{x}})) < n$. □

例 4.6. [7, DZ2] 考虑如下例子:

$$F = \{x_1^4, x_1^2x_2 + x_2^4, x_3 + x_3^2 - 7x_1^3 - 8x_1^2\}.$$

有一个 16 重根 $(0, 0, -1)$.

对于近似根 $\tilde{\mathbf{x}} = (0.01, 0.01, -1.01)$, 应用二阶 deflation 得到

$$G_2(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1, \mathbf{a}_1, \mathbf{a}_2) = \begin{pmatrix} F(\mathbf{x}) + X_0\mathbf{b}_0 + X_1\mathbf{b}_1 \\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1 + X_1'\mathbf{b}_1 \\ G_{1,\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1, \mathbf{a}_1)\mathbf{v}_2 \end{pmatrix},$$

$$X_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, X_1 = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \\ 0 & 0 \end{pmatrix}, X_1' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},$$

此时 $d^{(2)} = 1$, $\mathbf{c}^{(2)} = \mathbf{k}^{(2)} = \{1\}$. 因此根据(4.13), 得到增广系统

$$G_3(\mathbf{x}, \mathbf{b}, \mathbf{a}) = \begin{pmatrix} F(\mathbf{x}) + X_0\mathbf{b}_0 + X_1\mathbf{b}_1 + X_2\mathbf{b}_2 \\ F_{\mathbf{x}}(\mathbf{x})\mathbf{v}_1 + X_1'\mathbf{b}_1 + X_2'\mathbf{b}_2 \\ G_{1,\mathbf{x},\mathbf{b}_0,\mathbf{a}_1}(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{a}_1)\mathbf{v}_2 \\ G_{2,\mathbf{x},\mathbf{b}_0,\mathbf{b}_1,\mathbf{a}_1,\mathbf{a}_2}(\mathbf{x}, \mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{a}_1, \mathbf{a}_2)\mathbf{v}_3, \end{pmatrix},$$

$$X_2 = \frac{1}{2}\begin{pmatrix} x_1^2 \\ 0 \\ 0 \end{pmatrix}, X_2' = \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix}.$$

利用 INTLAB 中的算法 verifynlss, 我们得到扰动多项式系统

$$\widetilde{F}(\mathbf{x}, \hat{\mathbf{b}}) = \{f_1 + \hat{b}_{0,1} + \hat{b}_{1,1}x_1 + \frac{1}{2}\hat{b}_{2,1}x_1^2, f_2 + \hat{b}_{0,2} + \hat{b}_{1,2}x_2, f_3\}.$$

在扰动误差界

$$-0.00000000000001 \leq \hat{b}_{0,1},\ \hat{b}_{0,2},\ \hat{b}_{1,1},\ \hat{b}_{1,2},\ \hat{b}_{2,1} \leq 0.00000000000001,$$

内有孤立奇异根 $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$, 满足

$$-0.00000000000001 \leq \hat{x}_1,\ \hat{x}_2 \leq 0.00000000000001,$$
$$-1.00000000000001 \leq \hat{x}_3 \leq -0.99999999999999.$$

4.4 数值实验

根据上述分析, 我们接下来给出算法 VISS (Verification for Isolated Singular Solution).

算法：*VISS*

输入：▸ 多项式系统 $F=\{f_1,\ldots,f_n\}$, $f_i\in\mathbb{R}[\mathbf{x}]$, 近似根 $\tilde{\mathbf{x}}\in\mathbb{C}^n$ 和容忍度 τ.

输出：▸ 扰动多项式系统 $\widetilde{F}(\mathbf{x},\mathbf{b})$ 和可信误差界 $\mathbf{X}$, $\mathbf{B}$.

第 1 步 计算 Jacobian 矩阵 $F_{\mathbf{x}}(\tilde{\mathbf{x}})$ 的奇异值分解 $F_{\mathbf{x}}(\tilde{\mathbf{x}})=U\cdot\Sigma\cdot V^{\star}$, $\sigma_0:=+\infty$.

- 如果 $\sigma_n<\tau\ll\sigma_{n-1}$, 根据(4.1)构造增广系统 G 和计算 $(\tilde{\mathbf{x}},\mathbf{0},\tilde{\mathbf{a}})$.
- 如果 $\sigma_{n-d+1}<\tau\ll\sigma_{n-d}$, 根据(4.13)构造增广系统 G 和计算 $(\tilde{\mathbf{x}},\mathbf{0},\tilde{\mathbf{a}})$.

第 2 步 利用 INTLAB 中的算法 verifynlss 计算 $\mathbf{X}$, $\mathbf{B}$, 同时返回 $\widetilde{F}(\mathbf{x},\mathbf{b})$.

诸一骏帮我们在 matlab 中实现了算法 VISS, 程序代码和一些例子的实验结果可以在 `http://www.mmrc.iss.ac.cn/~lzhi/Research/hybrid/VISS/` 找到.

表 4.1 展示了算法 VISS 对下述例子和表 3.1 中部分例子的测试结果:

1. DZ1 [7]: $x_1^4-x_2x_3x_4$, $x_2^4-x_1x_3x_4$, $x_3^4-x_1x_2x_4$, $x_4^4-x_1x_2x_3$, $(0,0,0,0)$
2. DZ2 [7]: x^4, x^2y+y^4, $z+z^2-7x^3-8x^2$, $(0,0,-1)$
3. cbms1 [39]: x^3-yz, y^3-xz, z^3-xy, $(0,0,0)$
4. cbms2 [39]: $x^3-3x^2y+3xy^2-y^3-z^2$, $z^3-3z^2x+3zx^2-x^3-y^2$, $y^3-3y^2z+3yz^2-z^3-x^2$, $(0,0,0)$
5. mth191 [17]: $x^3+y^2+z^2-1$, $x^2+y^3+z^2-1$, $x^2+y^2+z^3-1$, $(0,1,0)$
6. KSS [14]: $f_i=x_i^2+\sum_{j=1}^{10}x_j-2x_i-9\ (i=1,\ldots,10)$, $(1,\ldots,1)$
7. Caprasse [26]: $x_2^2x_3+2x_1x_2x_4-2x_1-x_3$, $2x_2x_3x_4+x_1x_4^2-x_1-2x_3$, $-x_1x_3^3+4x_2x_3^2x_4+4x_1x_3x_4^2+2x_2x_4^3+4x_1x_3+4x_3^2-10x_2x_4-10x_4^2+2$, $-x_1^3x_3+4x_1x_2^2x_3+4x_1^2x_2x_4+2x_2^3x_4+4x_1^2-10x_2^2+4x_1x_3-10x_2x_4+2$, $(2,-i\sqrt{3},2,i\sqrt{3})$

8. cyclic9 [9]: $f_i = \sum_{j=0}^{8} \prod_{k=j}^{i+j} x_k \, (i = 0, \ldots, 7)$, $f_8 = 1 - \prod_{j=0}^{8} x_j$,

   $(z_0, z_1, z_2, z_0, -z_2, -z_1, z_0, -z_2, -z_1)$, $z_0 = -.9396926 - .3420201i$,

   $z_1 = -2.4601472 - .8954204i$, $z_2 = -.3589306 - .1306401i$

9. LZ [18]: $f_i = x_i^2 + x_i - x_{i+1} \, (i = 1, \ldots, 999)$, $f_{1000} = x_{1000}^3$, $(0, \ldots, 0)$

表 4.1 中第四列表示 deflation 过程中增广系统的 Jacobian 矩阵近似零空间的维数变化, 最后两列表示孤立奇异根和光滑参数的可信误差界规模.

表 4.1: 算法 VISS 的实验结果

系统	变元	重数	$d^{(t)}$	$\|\mathbf{X}\|$	$\|\mathbf{B}\|$
DZ1	4	131	$4 \to 4 \to 0$	e-321	e-321
DZ2	3	16	$2 \to 2 \to 1 \to 0$	e-14	e-14
cbms1	3	11	$3 \to 0$	e-321	e-321
cbms2	3	8	$3 \to 0$	e-321	e-321
mth191	3	4	$2 \to 0$	e-14	e-14
KSS	10	638	$9 \to 0$	e-14	e-14
Caprasse	4	4	$2 \to 0$	e-14	e-14
cyclic9	9	4	$2 \to 0$	e-13	e-13
RG	2	4	$1 \to 1 \to 1 \to 0$	e-14	e-14
LZ	1000	3	$1 \to 1 \to 0$	e-12	e-12
Ojika1	2	3	$1 \to 1 \to 0$	e-14	e-14
Ojika2	3	2	$1 \to 0$	e-14	e-14
Ojika3	3	2	$1 \to 0$	e-14	e-14
Ojika4	3	3	$1 \to 1 \to 0$	e-14	e-14
Decker2	3	4	$1 \to 1 \to 1 \to 0$	e-14	e-14
DZ3	2	4	$1 \to 1 \to 1 \to 0$	e-7	e-7

注 4. 例子 KSS 表明, 对于一般的孤立奇异根, 实际的 deflation 次数有时要远远小于理论的界; 例子 Caprasse 和 cyclic9 表明, 算法 VISS 对于复的孤立奇异根同

样适用. 值得注意的是例子 DZ3, 由于我们将多项式系数截断成浮点小数, 因此算法 VISS 得到的扰动误差界 $\mathbf{B}$ 不包含零.

第五章　结论与展望

本文主要研究多项式系统孤立奇异根的数值精化和可信验证问题.

第二章中, 我们讨论了多项式系统在孤立奇异根处, 局部对偶空间的性质和重结构的计算. 利用一些正则化和约化技巧, 我们提出了一种计算宽度为 1 的特殊情形下, 局部对偶空间一组既约基的新算法 MSB1 , 从而将算法涉及的矩阵的最大规模降低为方程个数 $\times$ 变元个数(首次与孤立奇异根的重数无关). 更为重要的是, 我们首次给出了宽度为 1 的特殊情形下, 局部对偶空间一组既约基的参数化表示. 算法 MSB1 对于解析系统同样适用, 而算法 AMSB1 可以用于计算带有近似系数的多项式系统或近似奇异根, 在宽度为 1 的特殊情形下, 近似局部对偶空间的一组既约基. 数值实验表明, 无论从计算时间还是存储空间上, 算法都是十分高效的.

第三章中, 我们研究了多项式系统孤立奇异根的数值精化问题. 我们首次证明了经典的 deflation 方法, 在宽度为 1 的特殊情形下的一个猜想: 一次 deflation 会使多项式系统在孤立奇异根处局部对偶空间的深度严格降 1. 对于精化宽度为 1 的孤立奇异根附近的近似根, 我们提出了一种新算法 MRRB1: 解一个正则化的最小二乘问题作为初始近似根的预处理, 从而得到一个更好的近似根和一个能够达到二次收敛的迭代方向; 通过计算近似局部对偶空间的一组既约基和解一个线性方程组, 从而确定合适的迭代步长. 我们证明了算法 MRRB1 在宽度为 1 的孤立奇异根附近是二次收敛的, 而且数值实验表明, 算法对于带有近似系数的多项式系统和超定系统(方程个数大于变元个数)同样有效.

第四章中, 我们研究了多项式系统孤立奇异根的可信验证问题. 利用区间验证方法和第二章中给出的宽度为 1 的特殊情形下, 局部对偶空间一组既约基的参数化表示, 我们提出了一种计算近似奇异根可信误差界的新算法, 其可以验证一个带有微小扰动的多项式系统, 在误差界内有一个宽度为 1 的孤立奇异根. 对于一般的孤立奇异根, 我们提出一种带光滑参数的 deflation 技术, 并证明了它的有限终止性, 从而将验证宽度为 1 的孤立奇异根的方法推广到一般情形. 在 Matlab 中我们实现了算法 VISS 并且测试了大量的例子, 数值实验表明, 算法对于带有近似系数的多项式系统和复的孤立奇异根同样适用.

今后的工作主要包含以下几个方向:

1. 将多项式系统在孤立奇异根处的局部对偶空间的概念, 推广到不可约代数簇上, 并且以此为工具研究多项式系统正维解的奇异性.

2. 将已有的工作应用到几何建模问题中去, 例如计算隐式曲线曲面的拓扑和参数曲面的交点等.

3. 研究更多非线性系统(例如微分方程)解的可信验证问题.

附　　录

假设 $\hat{\mathbf{x}} \in \mathbb{K}^n$ 是 $F = \{f_1, \ldots, f_m\}$ 的根, $f_i \in \mathbb{K}[\mathbf{x}]$, 满足 $n - \mathrm{rank}(F_{\mathbf{x}}(\hat{\mathbf{x}})) = 1$, 且 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 的第一列能够被其他列线性表示. 因此, 由定理 2.13, $\mathcal{D}_{\hat{\mathbf{x}}}(I)$ 的一组既约基有参数化表示 $\{1, d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n, \ldots, \Lambda_{\mu-1}\}$, 满足

$$\Lambda_k = \Delta_k + a_{k,2}d_2 + \cdots + a_{k,n}d_n,\ k = 2, \ldots, \mu - 1,$$

$$\Delta_k = \sum_{1<|\boldsymbol{\alpha}|\leq k} c_{\boldsymbol{\alpha}}\mathbf{d}_{\boldsymbol{\alpha}} = \Psi_{x_1}(\Lambda_{k-1}) + \sum_{j=1}^{k-1} a_{k-j,2}\Psi_{x_2}(\Lambda_j) + \cdots + \sum_{j=1}^{k-1} a_{k-j,n}\Psi_{x_n}(\Lambda_j).$$

这里 $\mathbf{d}^{\boldsymbol{\alpha}} = d_1^{\alpha_1}\cdots d_n^{\alpha_n}$, $d_i^{\alpha_i} = \frac{1}{\alpha_i!}\frac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}}$. 下面证明 Δ_k 满足的另一个等价形式.

命题 5.1.

$$k\Delta_k = \Lambda_{k-1}\frac{\partial}{\partial x_1} + \sum_{j=2}^{n}(a_{1,j}\Lambda_{k-1} + \cdots + (k-1)a_{k-1,j}\Lambda_1)\frac{\partial}{\partial x_j}.$$

证明. 采用归纳法证明. 首先根据定理 2.13 得到

$$\begin{aligned}
\Delta_2 &= (d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n)d_1 + \sum_{j=2}^{n} a_{1,j}(a_{1,j}d_j + \cdots + a_{1,n}d_n)d_j \\
&= \frac{1}{2}\Bigg[(\frac{\partial}{\partial x_1} + a_{1,2}\frac{\partial}{\partial x_2} + \cdots + a_{1,n}\frac{\partial}{\partial x_n})\frac{\partial}{\partial x_1} \\
&\quad + \sum_{j=2}^{n} a_{1,j}(\frac{\partial}{\partial x_1} + a_{1,2}\frac{\partial}{\partial x_2} + \cdots + a_{1,n}\frac{\partial}{\partial x_n})\frac{\partial}{\partial x_j}\Bigg].
\end{aligned}$$

因此对于 $k = 2$, 命题成立. 假设 Δ_k 满足如上等式, 下面证明 Δ_{k+1} 也满足. 记

$$\Gamma_k := \Lambda_{k-1}\frac{\partial}{\partial x_1} + \sum_{j=2}^{n}(a_{1,j}\Lambda_{k-1} + \cdots + (k-1)a_{k-1,j}\Lambda_1)\frac{\partial}{\partial x_j},$$

$C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Lambda_k}$, $C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_k}$ 和 $C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Gamma_k}$ 分别是 $\mathbf{d}^{\boldsymbol{\alpha}}$ 在 Λ_k, Δ_k 和 Γ_k 中的系数. 下面证明

$$C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Gamma_{k+1}} = (k+1)C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}},\ \forall\, \boldsymbol{\alpha} \in \mathbb{N}^n,\ |\boldsymbol{\alpha}| \leq k+1.$$

对于 $\mathbf{d}^{\boldsymbol{\alpha}} = d_1^{\alpha_1}\cdots d_n^{\alpha_n}$, 记录 $\boldsymbol{\alpha}$ 中非零元素的顺序和位置. 如果 $\alpha_i \neq 0$,

$$j_{\alpha_1+\cdots+\alpha_{i-1}+l} := i,\ l = 1,\ldots,\alpha_i.$$

那么 $\mathbf{d}^{\boldsymbol{\alpha}} = \prod_{l=1}^{|\boldsymbol{\alpha}|} d_{j_l}$. 例如对于 $\mathbf{d}^{\boldsymbol{\alpha}} = d_1^2 d_3 d_4^2$, 记 $j_1 = j_2 = 1, j_3 = 3, j_4 = j_5 = 4$.

根据引理 2.12,

$$\begin{cases} C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}} = C_{\Phi_{x_1}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_k}, & \alpha_1 \neq 0, \\ C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}} = a_{1,j} C_{\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_k} + \cdots + a_{k,j} C_{\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_1}, & \alpha_j \neq 0. \end{cases} \tag{5.1}$$

接下来利用上述公式分次分类地证明 $C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Gamma_{k+1}} = (k+1) C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}}$.

1. 如果 $|\boldsymbol{\alpha}| = 2$ 且 $\alpha_1 \neq 0$, 即 $\mathbf{d}^{\boldsymbol{\alpha}} = d_{j_1} d_{j_2}$, $j_1 = 1, 1 < j_2 \leq n$, 那么

$$C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Gamma_{k+1}} = a_{k,j_2} + k\, a_{k,j_2} = (k+1) a_{k,j_2} = (k+1) C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}}.$$

2. 如果 $|\boldsymbol{\alpha}| = 2$ 且 $\alpha_1 = 0$, 即 $\mathbf{d}^{\boldsymbol{\alpha}} = d_{j_1} d_{j_2}$, $1 < j_1 \leq j_2 \leq n$, 那么

$$\begin{aligned} C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Gamma_{k+1}} &= a_{1,j_1} a_{k,j_2} + \cdots + k a_{k,j_1} a_{1,j_2} \\ &\quad + a_{1,j_2} a_{k,j_1} + \cdots + k a_{k,j_2} a_{1,j_1} \\ &= (k+1)(a_{1,j_1} a_{k,j_2} + \cdots + a_{k,j_1} a_{1,j_2}) \\ &= (k+1) C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}}. \end{aligned}$$

3. 如果 $|\boldsymbol{\alpha}| > 2$ 且 $\alpha_1 \neq 0$, 利用归纳法,

$$\begin{aligned} & C_{\Phi_{x_1}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Gamma_k} = k C_{\Phi_{x_1}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Delta_k} \\ & = (\alpha_1 - 1) C_{\Phi_{x_1}^2(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_{k-1}} + \sum_{j=2}^{n} \alpha_j \left(a_{1,j} C_{\Phi_{x_1}\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_{k-1}} + \cdots + (k-2) a_{k-2,j} C_{\Phi_{x_1}\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_1} \right). \end{aligned}$$

再由(5.1), 得到 $C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}} = C_{\Phi_{x_1}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_k} = C_{\Phi_{x_1}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Delta_k}$ 且

$$\begin{aligned} & C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Gamma_{k+1}} = \alpha_1 C_{\Phi_{x_1}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_k} + \sum_{j=2}^{n} \alpha_j \left(a_{1,j} C_{\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_k} + \cdots + (k-1) a_{k-1,j} C_{\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_1} \right) \\ & = \alpha_1 C_{\Phi_{x_1}^2(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_{k-1}} + \sum_{j=2}^{n} \alpha_j \left(a_{1,j} C_{\Phi_{x_1}\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_{k-1}} + \cdots + (k-2) a_{k-2,j} C_{\Phi_{x_1}\Phi_{x_j}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Lambda_1} \right) \\ & = (k+1) C_{\Phi_{x_1}(\mathbf{d}^{\boldsymbol{\alpha}})}^{\Delta_k} = (k+1) C_{\mathbf{d}^{\boldsymbol{\alpha}}}^{\Delta_{k+1}}. \end{aligned}$$

4. 如果 $|\boldsymbol{\alpha}| > 2$ 且 $\alpha_1 = 0$,

$$\begin{aligned}C^{\Gamma_{k+1}}_{\mathbf{d}^{\boldsymbol{\alpha}}} =& a_{1,j_1} C^{\Delta_k}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})} + \cdots + k a_{k,j_1} C^{\Delta_1}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})} \\ &+ \cdots + a_{1,j_{|\alpha|}} C^{\Delta_k}_{\Phi_{x_{j_{|\alpha|}}}(\mathbf{d}^{\boldsymbol{\alpha}})} + \cdots + k a_{k,j_{|\alpha|}} C^{\Delta_1}_{\Phi_{x_{j_{|\alpha|}}}(\mathbf{d}^{\boldsymbol{\alpha}})}.\end{aligned}$$

再由(5.1)推出 $C^{\Delta_k}_{\mathbf{d}^{\boldsymbol{\alpha}}} = a_{1,j_1} C^{\Lambda_{k-1}}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})} + \cdots + a_{k-1,j_1} C^{\Lambda_1}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})}$.

考虑 $C^{\Gamma_{k+1}}_{\mathbf{d}^{\boldsymbol{\alpha}}}$ 中关于 a_{p,j_1} 的部分, $1 \le p \le k$,

$$\begin{aligned}& p a_{p,j_1} C^{\Lambda_{k-p+1}}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})} \\ &+ a_{p,j_1}\left(a_{1,j_2} C^{\Lambda_{k-p}}_{\Phi_{x_{j_1}}\Phi_{x_{j_2}}(\mathbf{d}^{\boldsymbol{\alpha}})} + \cdots + (k-p) a_{k-p,j_2} C^{\Lambda_1}_{\Phi_{x_{j_1}}\Phi_{x_{j_2}}(\mathbf{d}^{\boldsymbol{\alpha}})}\right) \\ &+ \cdots + a_{p,j_1}\left(a_{1,j_{|\alpha|}} C^{\Lambda_{k-p}}_{\Phi_{x_{j_1}}\Phi_{x_{j_{|\alpha|}}}(\mathbf{d}^{\boldsymbol{\alpha}})} + \cdots + (k-p)\cdot a_{k-p,j_{|\alpha|}} C^{\Lambda_1}_{\Phi_{x_{j_1}}\Phi_{x_{j_{|\alpha|}}}(\mathbf{d}^{\boldsymbol{\alpha}})}\right) \\ &= p\cdot a_{p,j_1} C^{\Lambda_{k-p+1}}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})} + (k-p+1) a_{p,j_1} C^{\Lambda_{k-p+1}}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})} \\ &= (k+1) a_{p,j_1} C^{\Lambda_{k-p+1}}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})}.\end{aligned}$$

因此, $C^{\Gamma_{k+1}}_{\mathbf{d}^{\boldsymbol{\alpha}}} = (k+1)\sum_{p=1}^{k} a_{p,j_1} C^{\Lambda_{k-p+1}}_{\Phi_{x_{j_1}}(\mathbf{d}^{\boldsymbol{\alpha}})} = (k+1) C^{\Delta_{k+1}}_{\mathbf{d}^{\boldsymbol{\alpha}}}$.

根据上述分析我们得到, 对 $\forall \boldsymbol{\alpha} \in \mathbb{N}^n$, $|\boldsymbol{\alpha}| \le k+1$, $C^{\Gamma_{k+1}}_{\mathbf{d}^{\boldsymbol{\alpha}}} = (k+1) C^{\Delta_{k+1}}_{\mathbf{d}^{\boldsymbol{\alpha}}}$. 所以命题成立. □

命题 5.2. 假设两组多项式系统

$$F_i := G_i + F_{\mathbf{x}} \cdot \mathbf{a}_i,\ G_i := \sum_{j=1}^{i-1} \frac{j}{i} \cdot F_{i-j,\mathbf{x}} \cdot \mathbf{a}_j,\ i = 2, \ldots, \mu - 1,$$

$F_1 = F_{\mathbf{x}} \cdot \mathbf{a}_1$. $\mathbf{a}_1 := (1, a_{1,2}, \ldots, a_{1,n})^T$, $\mathbf{a}_i := (0, a_{i,2}, \ldots, a_{i,n})^T$. 那么

$$F_i(\hat{\mathbf{x}}) = \Lambda_i(F),\ G_i(\hat{\mathbf{x}}) = \Delta_i(F).$$

证明. 首先由 $F_1 = F_{\mathbf{x}} \cdot \mathbf{a}_1$, 得到

$$\begin{aligned}\Lambda_1(F) &= (d_1 + a_{k,2} d_2 + \cdots + a_{k,n} d_n)(F) \\ &= F_{\mathbf{x}}(\hat{\mathbf{x}}) \cdot \mathbf{a}_1 \\ &= F_1(\hat{\mathbf{x}}).\end{aligned}$$

根据命题5.1, 可以归纳地证明出

$$
\begin{aligned}
\Delta_i(F) &= \frac{1}{i}\left[\Lambda_{i-1}\frac{\partial}{\partial x_1}+\sum_{j=2}^{n}(a_{1,j}\Lambda_{i-1}+\cdots+(i-1)a_{i-1,j}\Lambda_1)\frac{\partial}{\partial x_j}\right](F)\\
&= \frac{1}{i}\left[\Lambda_{i-1}\left(\frac{\partial}{\partial x_1}+a_{1,2}\frac{\partial}{\partial x_2}+\cdots+a_{1,n}\frac{\partial}{\partial x_n}\right)\right.\\
&\quad\left.+\sum_{j=2}^{i-1}(i-j)\Lambda_{i-j}\left(a_{j,2}\frac{\partial}{\partial x_2}+\cdots+a_{j,n}\frac{\partial}{\partial x_n}\right)\right](F)\\
&= \sum_{j=1}^{i-1}\frac{j}{i}\cdot F_{i-j,\mathbf{x}}(\hat{\mathbf{x}})\cdot\mathbf{a}_j\\
&= G_i(\hat{\mathbf{x}}),
\end{aligned}
$$

还有

$$
\begin{aligned}
\Lambda_i(F) &= (\Delta_i+a_{k,2}d_2+\cdots+a_{k,n}d_n)(F)\\
&= \Delta_i(F)+F_{\mathbf{x}}(\hat{\mathbf{x}})\cdot\mathbf{a}_i\\
&= G_i(\hat{\mathbf{x}})+F_{\mathbf{x}}(\hat{\mathbf{x}})\cdot\mathbf{a}_i\\
&= F_i(\hat{\mathbf{x}}).
\end{aligned}
$$

所以命题成立. □

命题 5.3. 假设 $\mathrm{rank}(M)=n-1$, M 是 $F_{\mathbf{x}}(\hat{\mathbf{x}})$ 去掉第一列剩下的子矩阵, 那么增广系统

$$
G(\mathbf{x},\boldsymbol{\lambda})=\begin{cases}F,\\ F_{\mathbf{x}}\boldsymbol{\lambda},\\ \lambda_1-1,\end{cases}
$$

在 $(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})$ 处局部对偶空间的一组既约基有如下形式:

$$
\begin{aligned}
\mathcal{D}_{(\hat{\mathbf{x}},\hat{\boldsymbol{\lambda}})}(J) &= \mathrm{Span}_{\mathbb{C}}\{1,d_1+a_{1,2}d_2+\cdots+a_{1,n}d_n+2a_{2,2}d_{n+2}+\cdots+2a_{2,n}d_{2n},\\
&\quad\ldots,\Delta'_{\mu-2}+(\mu-1)a_{\mu-1,2}d_{n+2}+\cdots+(\mu-1)a_{\mu-1,n}d_{2n}\},
\end{aligned}
$$

而且满足

$$
\Delta'_k(F_{\mathbf{x}}\boldsymbol{\lambda})=(k+1)\Delta_{k+1}(F).
$$

证明. 类似于定理 3.1 的证明, 可以归纳地得到 Δ_k' 满足

$$\Delta_k' = \Lambda_k + \sum_{j=2}^{n}(2a_{2,j}\Lambda_{k-1} + \cdots + ka_{k-1,j}\Lambda_1)d_{n+j} + \sum_{\mathbf{d}_{\boldsymbol{\alpha}}\in R_1, \mathbf{d}_{\boldsymbol{\beta}}\in R_2, |\boldsymbol{\beta}|>1} c_{\boldsymbol{\alpha}+\boldsymbol{\beta}}\mathbf{d}_{\boldsymbol{\alpha}}\mathbf{d}_{\boldsymbol{\beta}},$$

$R_1 = \mathbb{C}[d_1, \ldots, d_n]$, $R_2 = \mathbb{C}[d_{n+2}, \ldots, d_{n+r+1}]$. 因此

$$\begin{aligned} \Delta_k'(F_{\mathbf{x}}\boldsymbol{\lambda}) &= \Lambda_k(F_{\mathbf{x}}\mathbf{a}_1) + \sum_{j=2}^{n}(2a_{2,j}\Lambda_{k-1} + \cdots + ka_{k-1,j}\Lambda_1)\left(\frac{\partial F}{\partial x_j}\right) \\ &= \Lambda_k\left(\frac{\partial F}{\partial x_1}\right) + \sum_{j=2}^{n}(a_{1,j}\Lambda_k + \cdots + ka_{k-1,j}\Lambda_1)\left(\frac{\partial F}{\partial x_j}\right) \end{aligned}$$

根据命题 5.1得到

$$\Delta_k'(F_{\mathbf{x}}\boldsymbol{\lambda}) = (k+1)\Delta_{k+1}(F).$$

所以增广系统 $G(\mathbf{x}, \boldsymbol{\lambda})$ 在 $(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}})$ 处局部对偶空间的一组既约基满足

$$\begin{aligned} \Lambda_k'(F_{\mathbf{x}}\boldsymbol{\lambda}) &= [\Delta_k' + (k+1)a_{k+1,2}d_{n+2} + \cdots + (k+1)a_{k+1,n}d_{2n}]\,(F_{\mathbf{x}}\boldsymbol{\lambda}) \\ &= [(k+1)\Delta_{k+1} + (k+1)a_{k+1,2}d_2 + \cdots + (k+1)a_{k+1,n}d_n]\,(F) \\ &= (k+1)\Lambda_{k+1}(F) = \mathbf{0}. \end{aligned}$$

□

参考文献

[1] L. Alberti, B. Mourrain, and J. Wintz. Topology and arrangement computation of semi-algebraic planar curves. *Computer Aided Geometric Design*, 25:631–651, 2008.

[2] B. Buchberger. Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. *Journal of Symbolic Computation*, 41:475–511, 2006.

[3] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra.* Undergraduate Texts in Mathematics. Springer, 1996.

[4] D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry.* Graduate Texts in Mathematics. Springer, 2005.

[5] B. Dayton. Numerical local rings and local solutions of nonlinear systems. In *Proceedings of the 2007 International Workshop on Symbolic-Numeric Computation*, SNC '07, pages 79–86, New York, NY, USA, 2007. ACM.

[6] B. Dayton, T. Li, and Z. Zeng. Multiple zeros of nonlinear systems. *Mathematics of Computation*, 80:2143–2168, 2011.

[7] B. Dayton and Z. Zeng. Computing the multiplicity structure in solving polynomial systems. In *Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation*, ISSAC'05, pages 116–123, New York, NY, USA, 2005. ACM.

[8] D. Decker, H. Keller, and C. Kelly. Convergence rate for Newton's method at singular points. *SIAM Journal on Numerical Analysis*, 20:296–314, 1983.

[9] J. Faugére. A new efficient algorithm for computing Gröbner bases. *Journal of Pure and Applied Algebra*, 139:61–88, 1998.

[10] M. Giusti, G. Lecerf, B. Salvy, and J. Yakoubsohn. On location and approximation of clusters of zeros: case of embedding dimension one. *Foundations of Computational Mathematics*, 7(1):1–58, 2007.

[11] A. Griewank and M. Osborne. Analysis of Newton's method for irregular singularities. *SIAM Journal on Numerical Analysis*, 20(4):747–773, 1983.

[12] L. Kantorovich and G. Akilov. *Functional Analysis in Normed Spaces*. Pergamon, New York, 1964.

[13] Y. Kanzawa and S. Oishi. Imperfect singular solutions of nonlinear equations and a numerical method of proving their existence. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 82(6):1062–1069, 1999.

[14] H. Kobayashi, H. Suzuki, and Y. Sakai. Numerical calculation of the multiplicity of a solution to algebraic equations. *Mathematics of Computation*, 67:257–270, 1998.

[15] R. Krawczyk. Newton-algorithmen zur bestimmung von nullstellen mit fehlerschranken. *Computing*, 4:187–201, 1969.

[16] J. Lasserre, M. Laurent, and P. Rostalski. Semidefinite characterization and computation of zero-dimensional real radical ideals. *Foundations of Computational Mathematics*, 8:607–647, 2008.

[17] A. Leykin, J. Verschelde, and A. Zhao. Newton's method with deflation for isolated singularities of polynomial systems. *Theoretical Computer Science*, 359(1):111–122, 2006.

[18] N. Li and L. Zhi. Compute the multiplicity structure of an isolated singular solution: case of breadth one. *Journal of Symbolic Computation*, 47(6):700–710, 2012.

[19] N. Li and L. Zhi. Computing isolated singular solutions of polynomial systems: case of breadth one. *SIAM Journal on Numerical Analysis*, 50(1):354–372, 2012.

[20] F. Macaulay. *The Algebraic Theory of Modular Systems*. Cambridge University Press, 1916.

[21] A. Mantzaflaris and B. Mourrain. Deflation and certified isolation of singular zeros of polynomial systems. In *Proceedings of the 2011 International Symposium on Symbolic and Algebraic Computation*, ISSAC'11, pages 249–256, New York, NY, USA, 2011. ACM.

[22] M. Marinari, T. Mora, and H. Möller. Gröbner duality and multiplicities in polynomial solving. In *Proceedings of the 1995 International Symposium on Symbolic and Algebraic Computation*, ISSAC'95, pages 167–179, 1995.

[23] M. Marinari, T. Mora, and H. Möller. On multiplicities in polynomial system solving. *Transactions of the American Mathematical Society*, 348:3283–3321, 1996.

[24] R. Menzel. On solving nonlinear least-squares problems in case of rankdeficient jacobians. *Computing*, 34:63–72, 1985.

[25] R. Moore. A test for existence of solutions to nonlinear systems. *SIAM Journal on Numerical Analysis*, 14(4):611–615, 1977.

[26] S. Moritsugu and K. Kuriyama. On multiple zeros of systems of algebraic equations. In *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation*, ISSAC'99, pages 23–30, New York, NY, USA, 1999. ACM.

[27] B. Mourrain. Isolated points, duality and residues. *Journal of Pure and Applied Algebra*, 117 & 118:469–493, 1997.

[28] T. Ojika. Modified deflation algorithm for the solution of singular problems. *Journal of Mathematical Analysis and Applications*, 123(1):199–237, 1987.

[29] T. Ojika, S. Watanabe, and T. Mitsui. Deflation algorithm for the multiple roots of a system of nonlinear equations. *Journal of Mathematical Analysis and Applications*, 96:463–479, 1983.

[30] G. Reid and L. Zhi. Solving polynomial systems via symbolic-numeric reduction to geometric involutive form. *Journal of Symbolic Computation*, 44:280–291, 2009.

[31] J. Ritt. *Differential Algebra*. AMS, 1950.

[32] F. Rouillier. Solving zero-dimensional systems through the rational univariate representation. *Applicable Algebra in Engineering, Communication and Computing*, 9:433–461, 1999.

[33] S. Rump. Solving algebraic problems with high accuracy. In *Proceedings of the symposium on A new approach to scientific computation*, pages 51–120, San Diego, CA, USA, 1983. Academic Press Professional, Inc.

[34] S. Rump. INTLAB - INTerval LABoratory. *Developments in Reliable Computing*, pages 77–104, 1999.

[35] S. Rump and S. Graillat. Verified error bounds for multiple roots of systems of nonlinear equations. *Numerical Algorithms*, 54(3):359–377, 2010.

[36] Y. Shen and T. Ypma. Newton's method for singular nonlinear equations using approximate left and right nullspaces of the jacobian. *Applied Numerical Mathematics*, 54:256–265, 2005.

[37] R. Stanley. Hilbert function of graded algebras. *Advances in Mathematics*, 28:57–83, 1973.

[38] H. Stetter. *Numerical Polynomial Algebra*. SIAM, Philadelphia, 2004.

[39] B. Sturmfels. *Solving Systems of Polynomial Equations*. Number 97 in CBMS Reginal Conference Series in Mathematics. AMS, 2002.

[40] A. Tikhonov and V. Arsenin. *Solution of Ill Posed Problems*. V. H. Winston and Sons, Washington, DC, 1977.

[41] J. Verschelde. Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software*, 25(2):251–276, 1999.

[42] W. Wu. *Mathematics Mechanization*. Science Press and Kluwer Academic Publishers, 2000.

[43] X. Wu and L. Zhi. Computing the multiplicity structure from geometric involutive form. In *Proceedings of the 2008 International Symposium on Symbolic and Algebraic Computation*, ISSAC'08, pages 325–332, New York, NY, USA, 2008. ACM.

[44] X. Wu and L. Zhi. Determining singular solutions of polynomial systems via symbolic-numeric reduction to geometric involutive forms. *Journal of Symbolic Computation*, 47:227–238, 2012.

[45] B. Xia and T. Zhang. Real solution isolation using interval arithmetic. *Computers & Mathematics with Applications*, 52(6-7):853–860, 2006.

发表文章目录

[1] *An improved method for evaluating Max Noether conditions: case of breadth one*. N. Li. In SNC '11: Proceedings of the 2011 International Workshop on Symbolic-Numeric Computation, ACM, New York, NY, USA, pp. 102-103, 2011.

[2] *Computing the multiplicity structure of an isolated singular solution: case of breadth one*. N. Li and L. Zhi. Journal of Symbolic Computation, 47(6): pp. 700-710, 2012.

[3] *Computing isolated singular solutions of polynomial systems: case of breadth one*. N. Li and L. Zhi. SIAM Journal on Numerical Analysis, 50(1): pp. 354-372, 2012.

[4] *Verified error bounds for isolated singular solutions of polynomial systems: case of breadth one*. N. Li and L. Zhi. Theoretical Computer Science, 479: pp. 163-173, 2013.

[5] *Verified error bounds for isolated singular solutions of polynomial systems*. N. Li and L. Zhi. Submitted.

简　　历

李楠, 男, 1985年10月9日出生于河北唐山.

教育经历

2008.9–2013.6　硕博研究生, 应用数学专业, 中国科学院数学与系统科学研究院, 数学机械化重点实验室, 导师: 支丽红 研究员.

2004.9–2008.6　理学学士, 信息与计算科学专业, 吉林大学数学学院.

学术经历

1. 第10届亚洲计算机数学研讨会(ASCM2012), 中国北京, 2012年10月.
2. IMA 研究生暑期项目"Algebraic Geometry for Applications", 佐治亚理工学院, 美国亚特兰大, 2012年6月至7月.
3. 国家自然科学基金项目重大研究计划可信软件基础研究重点支持项目: 基于符号数值混合计算的误差可控算法及其应用, 2012年1月至今.
4. 2011年国际符号数值计算研讨会(SNC2011), 美国圣何西, 2011年6月.
5. 法国国立信息与自动化研究院(INRIA)和中国科学院联合培养博士生项目, Paris-Rocquencourt 研究中心, 法国巴黎, 2010年12月至2011年12月.

获奖经历

2013.1	博士研究生国家奖学金
2012.9	数学与系统科学研究院院长奖学金
2012.6	数学与系统科学研究院博时奖学金
2012.4	中国科学院研究生院三好学生
2010.9	中国科学院中欧联合培养基金
2010.5	数学与系统科学研究院金融界奖学金
2008.6	Ky and Yu-Fen Fan Fund Travel Grant from the AMS

致　谢

五年的博士生生活一晃而过, 回首走过的岁月, 心中倍感充实, 论文即将完成之日, 感慨良多. 首先诚挚地感谢我的导师支丽红研究员, 从论文选题, 中期修改到最终定稿的过程中, 自始至终都倾注着导师的心血. 导师以严谨的治学之道, 宽厚仁慈的胸怀, 积极乐观的生活态度, 为我树立了一辈子学习的典范, 她的教诲与鞭策, 将激励我在科学和教育的道路上励精图治, 开拓创新.

衷心地感谢 Fabrice Rouillier 研究员和 Anton Leykin 教授, 在我参加 INRIA-CAS 联合培养博士生项目和 IMA 研究生暑期项目时给予我的支持和帮助. 而且在同他们讨论问题的过程中, 我学到了很多知识也得到了很多启发. 感谢华东师范大学研究生诸一骏同学, 感谢他帮助我编写最后一章中的程序代码.

感谢数学机械化中心的全体同仁, 没有他们就没有我这篇论文. 感谢吴文俊院士, 高小山研究员, 李洪波研究员, 李子明研究员, 刘卓军研究员, 王定康研究员, 闫振亚研究员, 冯如勇副研究员, 袁春明副研究员, 程进三副研究员, 黄雷助理研究员, 张立先助理研究员, 李伟助理研究员, 感谢他们五年来给予我的关心和鼓励. 感谢周代珍老师, 丁建敏老师和李佳老师, 感谢她们帮我处理诸多日常事务. 感谢杨争峰师兄, 李冰玉师姐, 吴晓丽师姐, 李斌师兄, 陈绍示师兄, 郭峰师兄, 马玥师姐, 梁野师兄, 李喆师姐, 李子佳, 马晓栋, 姚守彬, 吴宝峰, 张可, 郭庆东, 刘琦, 以及实验室其他同学. 他们给我单调的科研生活增添了绚丽的色彩, 在我心情低落时为我分忧, 在我信心不足时给我打气, 与他们在一起的日子开心而充实, 他们将是我记忆里最美的风景.

感谢我的妻子唐浩女士, 没有她的体贴, 包容和支持, 相信这五年的博士生生活将是很不一样的光景. 这篇论文的完成同样离不开了她的心血和奉献!

最后谨以此文献给我挚爱的双亲, 他们在背后的默默支持是我前进的动力. 在此祝愿他们身体健康, 心情愉快!