arXiv:2601.06148v1 [math.RA] 5 Jan 2026

基于多项式加权主成分分析的实多项式正交等价性证书*

Martin Helmer 数学系，斯旺西大学，斯旺西，SA1 8EN，威尔士，英国
martin.helmer@swansea.ac.uk

David Hong 电气与计算机工程系，特拉华大学，纽瓦克，DE 19716，美国 hong@udel.edu

Hoon Hong 数学系，北卡罗来纳州立大学，罗利，NC 27695，美国 hong@ncsu.edu

2026年1月13日

摘要

假设 $f(x) \in \mathbb{R}[x_1, \ldots, x_n]$ 和 $g(x) \in \mathbb{R}[x_1, \ldots, x_n]$ 是两个具有 d 次、含 n 个变量的实多项式。如果多项式 f 和 g 在正交对称性下相同，那么一个自然的问题是：正交群中的哪个元素诱导了这种正交对称性；也就是说，求出满足 $f(Rx) = g(x)$ 的元素 $R \in O(n)$。一种直接的方法是根据关系 $f(Rx) = g(x)$ 以及正交群的恒等式构造非线性方程组来解决这个问题；然而，当 n 和 d 的值较大时，这种方法的计算开销会变得相当大。为此，我们提出一种替代方案，能够以更强的可扩展性解决这一问题：引入 *Polynomial-Weighted Principal Component Analysis*（PW-PCA）的概念。具体而言，我们将展示如何有效计算 PW-PCA，以及如何利用这些技术获得 *certificate of orthogonalequivalence*，即求出满足 $f(Rx) = g(x)$ 的 $R \in O(n)$。

1 引言

研究多项式在群作用下的行为，是代数中的经典课题。关于这一问题，尤其是如何判断两个多项式是否处于群作用下的同一轨道，不变理论提供了一个成熟且广泛的理论框架，相关内容在许多著作中均有详尽介绍，例如 [9, 20, 25, 8, 22]。考虑一个群 G 作用于由系数取自域 k、由 n 个变量构成的 d 次多项式所组成的向量空间 V。从计算的角度来看，已有算法可用于判断两个多项式 f，g 是否处于同一 G-轨道。这些算法大致通过计算不变量环 $k[V]^G$ 的各种性质来完成这一任务，参见例如 [8, 26, 19], ；它们适用于许多不同的群 G。尽管许多算法方法着重研究 G

*所有作者贡献相同。作者按字母顺序排列。

是有限群，但也会考虑无限群，例如当 G是线性代数群时，参见 [8] 中的详细讨论以及 [11, 10] 中的计算机实现。完备不变量集是指能够生成不变量环 $k[V]^G$的不变量集。在检验两个多项式是否处于同一 G-轨道时，环 $k[V]^G$ 中的变量表示多项式 f 在 V 中的系数 c；由 $r(c) \in k[V]^G$ 指定的给定不变量，在代入同一 G-轨道中任意两个多项式 V 的具体系数后，必须取相同的值。不变理论中还有许多关于特定群所对应不变量的闭式公式的结果。对于正交群，已有多位作者对其进行研究，参见例如 [14, 17] 及其中的参考文献；正交群 $O(n)$情形的最新研究成果似乎是近期的工作 [4]，其中给出了正交群作用于由 n 个变量构成的 d 次多项式向量空间时的有理不变量的简洁闭式表达式，适用于 d 为偶数的情形（并且有可能推广到奇数情形）。

令 $R \bullet f$表示群元素 R对多项式 f的作用。若有两个多项式 f和 g在群作用下等价，那么一个自然的后续问题是：如何获得<嵌入 style id='6'>等价性证书</嵌入>，也就是说，如何找到一个元素 $R \in G$ 使得 $R \bullet f = g$? 不变理论的方法并未直接回答这一问题。本文聚焦于一种专门的方法，它能在 G 为正交群 $O(n)$ 时提供这样的证书。

考虑为正交群 $G = O(n)$的作用获取等价性证书这一问题。给定两个固定的多项式 f和 g，它们在 $\mathbb{R}[x_1, \ldots, x_n]$ 中次数均为 d，可以考虑一个关于 n^2 个变量 $r_{11}, \ldots, r_{nn}$ 的矩阵 R，再加入定义正交群的约束 $R^\top R = I_n$，直接求解关于变量 $r_{11}, \ldots, r_{nn}$ 的方程组（如果有解）；该方程组由关系式 $R \bullet f = f(Rx) = g(x)$ 两边系数相等得到。参见第 5 节中的算法 28。当 $n = 1$ 时，对任意 d，或者当 $d = 1$ 时，对任意 n，该方程组都很容易求解。类似地，若 n 或 d 较小，其他若干特殊情形也相当容易求解。例如，当 $d = 2$时，已知存在简单的解法，可利用 f 和 g的最高次数齐次部分（二次型部分）的特征向量来表示。然而，一般来说，当 n 和 d较大时，求解由此得到的非线性方程组很快就会变得困难。

本文提出一种用于认证多项式正交等价的新颖解决方案，结合代数与统计学的思想，得到一种有效的计算方法，其可扩展性远胜于上文提到的直接求解方法。受广泛使用的数据分析技术主成分分析（PCA）的启发，我们将介绍一种称为多项式加权主成分分析（PW-PCA）的技术。

PW-PCA 使我们能够实质上将任意次数多项式的问题简化为二次齐次多项式的问题。随后，我们运用这种方法设计一种算法，为两个多项式求得正交等价证书。

有趣的是，求取正交等价证书的问题与所谓的正交 *Procrustes* 问题存在天然联系；后者已得到广泛研究，例如在数据点对齐的背景下。然而，这些问题之间存在关键差异，阻碍了将正交 Procrustes 问题的知名技术直接推广到我们的场景中；详见下文备注 5。相较之下，我们提出的方法利用了多项式几何的内在性质。

如上文所述，我们的主要关注点是在已知 f 和 g 正交等价的假设下，获得正交等价证书。然而，作为采用 PW-PCA 获得此类证书的副产品，我们还可以轻松设计出一种算法，原则上可用于在尚未知晓的情况下检验多项式 f 和 g 是否正交等价。但在实践中，这种方法的直接实现需要在代数数域中进行运算，因此会对实际效率造成问题，详见注释 27。因此，我们将其留待未来研究。

论文结构

本文结构安排如下。第2节将精确陈述本文的主要问题。第3节将提出多项式加权形式的主成分分析（PW-PCA），并给出计算该方法的算法。第4节将在一般性假设下，发展

利用PW-PCA解决本文主要问题的数学理论和算法。第5节将讨论该算法初步实现的性能。附录A将证明第4节中假设的一般性。

2 问题

现在，我们对问题作出精确陈述。首先回顾正交群 $O(n)$ 的定义：它是一般线性群 $\mathrm{GL}(n,\mathbb{R})$ 的一个子群，后者是实数项 $n\times n$ 可逆矩阵构成的乘法群。我们还定义了 $O(n)$ 在多项式上的相应群作用，以及多项式在此作用下等价的概念。

定义 1（正交等价）。

1.（群） n 维<形>正交群，记作 $O(n)$，定义如下：

$$O(n)=\left\{R\in\mathbb{R}^{n\times n}:R^{\top}R=I_n\right\}. \tag{1}$$

2.（作用） $R\in O(n)$ 在 $p\in\mathbb{R}[x_1,\ldots,x_n]$ 上的作用 记作 $R\bullet p$，定义如下：

$$(R\bullet p)(x)=p(Rx).$$

3.（等价 *e)* 我们称 f、$g\in\mathbb{R}[x_1,\ldots,x_n]$ 正交等价，并记作 $f\sim_{O(n)}g$，*if*

$$\exists_{R\in O(n)}\ \ g=R\bullet f.$$

有了这一定义，我们便可以精确地表述这样一个问题：求得 $O(n)$ 中的一个元素，以证明两个多项式的等价性。

问题 2（正交等价证书）：求取正交等价证书的一种方法是：对于给定的、以 n 个变量表示且次数为 d 的输入多项式 f 和 g，输出元素 $R\in O(n)$，如下所示。

给定： 次数为 d 的 f、$g\in\mathbb{R}[x_1,\ldots,x_n]$，满足 n、$d\geq 2$ 和 $f\sim_{O(n)}g$。

计算： $R\in O(n)$，使得 $g=R\bullet f$。

现在，我们通过一个简单示例来说明该问题，并将在全文中将其作为贯穿示例。

例3 (简单贯穿示例**)**。 考虑下面给出的两个次数为 $d=3$、属于 $\mathbb{R}[x_1,x_2,x_3]$ 的多项式 f 和 g：

$$\begin{aligned}f=&-27x_1^3+27x_2^2x_3-9x_3,\\ g=&-12x_1^3+12x_1^2x_2-12x_1^2x_3+6x_1x_2^2+36x_1x_2x_3-33x_1x_3^2\\ &+9x_2^3-6x_2^2x_3+6x_2x_3^2-6x_3^3+3x_1-6x_2-6x_3.\end{aligned}$$

这些多项式正交等价，即 $f\sim_{O(n)}g$。验证这种等价性的证书为 $R\in O(n)$

$$R=\begin{bmatrix}\frac{2}{3}&-\frac{1}{3}&\frac{2}{3}\\ \frac{2}{3}&\frac{2}{3}&-\frac{1}{3}\\ -\frac{1}{3}&\frac{2}{3}&\frac{2}{3}\end{bmatrix}.$$

我们可以轻松验证 R 是 $O(n)$ 的一个元素，并且 $f(Rx) = g(x)$。本文的目标是开发一种高效且可扩展的算法（如下文算法 *25* 所示），根据给定的输入多项式 f 和 g，计算上文的矩阵 R 。

备注 4. 在上述例子（例*3*）中，输入多项式系数和输出矩阵元素属于同一个域，即 $\mathbb{Q}$。但情况并非总是如此。例如，考虑以下输入。

$$f = 4x_1^2 - 2x_2^2$$
$$g = x_1^2 - 6x_1x_2 + x_2^2$$

那么，一个输出为

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix},$$

这对应于旋转 $\pi/4$。注意，输入多项式系数属于 $\mathbb{Q}$ ，而输出矩阵元素属于 $\mathbb{Q}(\sqrt{2})$。一般而言，输出矩阵元素属于输入系数域的一个代数扩张。因此，任何试图精确求解主要问题（问题 *2*）的方法，都需要在一个代数扩域中进行计算，这会严重限制该方法的实用性，尤其是在 n 和 d 较大的情况下。因此，人们转而致力于开发同时满足以下两项要求的方法：（*1*）必要时能够精确执行；（*2*）能够安全地近似执行，即该方法具有数值稳定性。

备注 5 (与对称张量和普罗克鲁斯问题的关系**).** 注意，定理 *2* 也可以看作是对称张量上的一种正交 *Procrustes* 问题。为说明这一点，我们将次数为 d 的多项式 f、$g \in \mathbb{R}[x_1, \ldots, x_n]$ 表示为对称系数张量的形式：

$$f(x) = F_0 + \langle F_1, x\rangle + \langle F_2, x \circ x\rangle + \cdots + \langle F_d, x^{\circ d}\rangle$$
$$g(x) = G_0 + \langle G_1, x\rangle + \langle G_2, x \circ x\rangle + \cdots + \langle G_d, x^{\circ d}\rangle,$$

其中，标量 F_0、$G_0 \in \mathbb{R}$ 分别表示 f 和 g 的常数项，向量 F_1、$G_1 \in \mathbb{R}^n$ 表示它们的线性项，对称矩阵 F_2、$G_2 \in \mathbb{R}^{n\times n}$ 表示它们的二次项，依此类推；$\circ$ 表示通常的张量外积，$\langle \cdot, \cdot \rangle$ 表示通常的张量内积。通过标准的张量代数运算（参见例如 *[1,* 第 *3]* 章），可得对于任意 $R \in \mathbb{R}^{n\times n}$，有

$$\begin{aligned} f(Rx) &= F_0 + \langle F_1, Rx\rangle + \langle F_2, (Rx) \circ (Rx)\rangle + \cdots + \langle F_d, (Rx)^{\circ d}\rangle \\ &= F_0 + \langle F_1 \times_1 R^\top, x\rangle + \langle F_2 \times_1 R^\top \times_2 R^\top, x \circ x\rangle + \cdots + \langle F_d \times_1 R^\top \cdots \times_d R^\top, x^{\circ d}\rangle, \end{aligned}$$

其中，$\times_i$ 表示沿模式 i 的张量与矩阵乘积。因此，我们寻求满足以下条件的 $R \in O(n)$：

$$G_i = F_i \times_1 R^\top \cdots \times_i R^\top, \quad \textit{for each } i \in \{1, \ldots, d\},$$

这可以表述为如下最小二乘优化问题

$$\operatorname*{argmin}_{R\in O(n)} \sum_{i=1}^{d} \|F_i \times_1 R^\top \cdots \times_i R^\top - G_i\|_{\mathsf{F}}^2, \tag{2}$$

其中，$\|\cdot\|_{\mathsf{F}}$ 表示张量的通常弗罗贝尼乌斯范数。

这与广为研究的正交 *Procrustes* 问题十分相似，该问题旨在为给定矩阵 A 和 B，寻找使 $\|RA - B\|_{\mathsf{F}}^2$ 最小化的 *[24]*，$R \in O(n)$。然而，两者存在关键差异。首先，问题 (2) 涉及所有阶数不超过 d 的对称张量，而不只是矩阵（即二阶张量）。此外，正交矩阵 R 会沿着 F_i 的所有模态相乘。这意味着，即使对于 (2) 中的二阶项——其目标是将矩阵 $F_2 \in \mathbb{R}^{n\times n}$ 变换为矩阵 $G_2 \in \mathbb{R}^{n\times n}$——该变换也对应于 $F_2 \times_1 R^\top \times_2 R^\top = R^\top F_2 R$，而非正交 *Procrustes* 问题中的 $R^\top F_2$。因此，通过对$BA^\top$ 进行极分解来求解正交 *Procrustes* 问题的常规方法，并不能直接推广来求解 (2)。本文提出了一种不同的方法，转而利用多项式几何的内在性质。

3 多项式加权主成分分析 s

受统计多变量分析的启发，本节将构建一种主成分分析的多项式加权形式，用于解决上一节提出的问题：如何认证多项式的正交等价，即问题2。

3.1 主成分分析简述

我们首先简要回顾主成分分析（PCA） [16, 21]，为介绍我们提出的多项式加权主成分分析作铺垫；全面综述请参见专著 [18]。给定一组 N数据点 $x^{(1)}, \dots, x^{(N)} \in \mathbb{R}^n$，PCA 会得到一组主轴，使其最大程度地捕捉数据中的变异性。由于许多数据集具有潜在的低维结构，通常只需相对较少的主轴，就能捕捉到大部分有意义的变异性。这些主轴通过寻找使数据点方差最大的方向来确定。具体而言，第一个主轴通过最大化[1]得到

$$\mathrm{Var}_u(x^{(1)}, \dots, x^{(N)}) = \frac{1}{N} \sum_{i=1}^{N} (u^\top x^{(i)})^2, \tag{3}$$

相对于单位方向向量 $u \in \mathbb{R}^n$。随后，第二个主轴通过最大化 $\mathrm{Var}_u(x^{(1)}, \dots, x^{(N)})$得到，其中单位方向向量 $u \in \mathbb{R}^n$ 与第一条主轴正交；依此类推。由此得到一组按顺序排列的正交归一主轴 $v_1, \dots, v_n \in \mathbb{R}^n$，如下所示：

$$\begin{aligned}
v_1 &\in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\|=1} \mathrm{Var}_u(x^{(1)}, \dots, x^{(N)}), \\
v_2 &\in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\|=1} \mathrm{Var}_u(x^{(1)}, \dots, x^{(N)}) \quad \text{s.t.} \quad u \perp v_1, \\
&\vdots \\
v_n &\in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\|=1} \mathrm{Var}_u(x^{(1)}, \dots, x^{(N)}) \quad \text{s.t.} \quad u \perp \{v_1, \dots, v_{n-1}\}.
\end{aligned} \tag{4}$$

值得注意的是，沿方向 $u \in \mathbb{R}^n$ 的方差 $\mathrm{Var}_u(x^{(1)}、\text{.....}、x^{(N)})$ 可以表示为一个二次型

$$\mathrm{Var}_u(x^{(1)}, \dots, x^{(N)}) = \frac{1}{N} \sum_{i=1}^{N} (u^\top x^{(i)})({x^{(i)}}^\top u) = u^\top \left(\frac{1}{N} \sum_{i=1}^{N} x^{(i)} {x^{(i)}}^\top \right) u = u^\top \mathrm{Cov}(x^{(1)}, \dots, x^{(N)}) u,$$

由协方差矩阵定义

$$\mathrm{Cov}(x^{(1)}, \dots, x^{(N)}) = \frac{1}{N} \sum_{i=1}^{N} x^{(i)} {x^{(i)}}^\top \in \mathbb{R}^{n \times n}. \tag{5}$$

因此，式（4）中的一系列优化问题的解由协方差矩阵的特征向量给出，也就是说，$v_1, \dots, v_n$ 是 $\mathrm{Cov}(x^{(1)}, \dots, x^{(N)})$ 的特征向量，按特征值递减的顺序排列 $\lambda_1 \geq \cdots \geq \lambda_n$。因此，主轴有时也等价地定义为协方差矩阵的特征向量。这里需要注意的是，主轴和特征向量一样，只有在特征值互不相同时（即当 $\lambda_1 > \cdots > \lambda_n$ 时）才是唯一的；即便如此，它们也只在符号意义下唯一。最后还需注意，尽管主成分分析通常首先针对有限数据点集合进行介绍，但它可以立即

[1]为简单起见，我们假设数据点已进行中心化处理，其平均值为 $(x^{(1)} + \cdots + x^{(N)})/N = 0$；只需从所有数据点中减去平均值，即可轻松完成这一步。

推广到服从任意概率测度 μ 的数据分布，只需将上述有限平均值替换为关于该测度的积分。具体而言，上述方差和协方差变为

$$\mathrm{Var}_u(\mu) = \int (u^\top x)^2 \, d\mu(x), \qquad \mathrm{Cov}(\mu) = \int xx^\top \, d\mu(x), \tag{6}$$

which 通常简洁地写作期望 $\mathbb{E}[(u^\top x)^2]$和 $\mathbb{E}[xx^\top]$，分别表示。

3.2 多项式加权方差与协方差的定义 e

现在，我们为多项式 $f \in \mathbb{R}[x_1, \ldots, x_n]$定义一种多项式加权的方差与协方差概念，它将构成所提出的多项式加权主成分分析的基础。类似于式 (3) 和 (5) 如何刻画数据在不同方向上的"分散程度"，这里的总体思路是刻画 f 沿任意给定方向 $u \in \mathbb{R}^n$的"大小"。我们将研究范围限定为齐次多项式 f，并照常将研究限制在单位球面上

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\}, \tag{7}$$

不失一 般性。这自然引出了以下多项式加权方差概念。

定义 6 (齐次多项式的多项式加权方差)。 齐次多项式 $f \in \mathbb{R}[x_1, \ldots, x_n]$ 沿方向 $u \in \mathbb{R}^n$ 的多项式加权方差定义为

$$\mathrm{Var}_u(f) = \int_{\mathbb{S}^{n-1}} f(x)^2 \, (u^\top x)^2 \, d\mu(x), \tag{8}$$

其中， μ表示 $\mathbb{S}^{n-1}$上的球面测度。

换言之，Var$_u$(f)表示沿方向 u 的方差，其中点 x 在球面$\mathbb{S}^{n-1}$上均匀分布，并以其对应的多项式值的平方 $f(x)^2$为权重。正如我们将在下文的定理 22 中所示，这一概念刻画了 f的"取向"。采用加权方差概念的加权主成分分析变体也已用于统计场景，例如用于考虑不同点的相对噪声水平 [15]。这里，多项式 f 决定赋予每个点的权重。

与式 (3) 中定义的数据沿不同方向的方差一样，多项式加权方差 Var$_u$(f)也可以表示为如下二次型:

$$\mathrm{Var}_u(f) = \int_{\mathbb{S}^{n-1}} f(x)^2 \, (u^\top x)(x^\top u) \, d\mu(x) = u^\top \left(\int_{\mathbb{S}^{n-1}} f(x)^2 \, xx^\top \, d\mu(x) \right) u.$$

由此得到如下多项式加权协方差矩阵的定义。

定义 7 (齐次多项式的多项式加权协方差)。 齐次多项式 $f \in \mathbb{R}[x_1, \ldots, x_n]$ 的多项式加权协方差矩阵定义为

$$\mathrm{Cov}(f) = \int_{\mathbb{S}^{n-1}} f(x)^2 \, xx^\top \, d\mu(x) \in \mathbb{R}^{n \times n}, \tag{9}$$

其中 μ 表示 $\mathbb{S}^{n-1}$ 上的球面测度。根据此定义，我们有 方差$_u(f) = u^\top$ 协方差(f)u。

为了进一步直观地理解这些概念，请注意，多项式加权方差和协方差也可以等价地写作

$$\mathrm{Var}_u(f) = \rho_f \cdot \int_{\mathbb{S}^{n-1}} (u^\top x)^2 \left(\frac{f(x)^2}{\rho_f} \right) d\mu(x), \qquad \mathrm{Cov}(f) = \rho_f \cdot \int_{\mathbb{S}^{n-1}} xx^\top \left(\frac{f(x)^2}{\rho_f} \right) d\mu(x),$$

其中 $\rho_f = \int_{\mathbb{S}^{n-1}} f(x)^2 \, d\mu(x)$ 对 $f(x)^2$ 进行归一化，使其成为概率密度（即 $f(x)^2/\rho_f$ 的积分为 1）。在这种解释下，多项式加权方差和协方差可以视为球面 $\mathbb{S}^{n-1}$ 上各点的方差和协方差，这些点的概率密度由 $f(x)^2/\rho_f$ 给出。

3.3 Main 定义：多项式加权主成分分析

现在我们可以介绍所提出的多项式加权主成分分析（PW-PCA）。其总体思路是为多项式定义主轴，类似于式（4）中为数据定义主轴的方式，具体做法是最大化多项式加权方差（定理 6）。由于多项式加权方差仅对齐次多项式有定义，我们采用其齐次化形式 $f \in \mathbb{R}[x_1, \ldots, x_{n+1}]$，定义如下

$$\overline{f}(x_1, \ldots, x_{n+1}) = x_{n+1}^{\deg(f)} f(x_1/x_{n+1}, \ldots, x_n/x_{n+1}). \tag{10}$$

这引出了 PW-PCA 的如下定义。它基本上是公式 (4) 的直接类比，只是在适当处理齐次化变量 x_{n+1}时存在一些细微之处。

定义 8（多项式加权主成分分析）。对于多项式 $f \in \mathbb{R}[x_1, \ldots, x_n]$，如果满足以下条件，我们称 v_1、……、$v_n \in \mathbb{R}^n$ 为 f 的一组主轴：

$$\begin{aligned} v_1 &\in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\|=1} \operatorname{Var}_{[u,0]}(\overline{f}), \\ v_2 &\in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\|=1} \operatorname{Var}_{[u,0]}(\overline{f}) \quad s.t. \quad u \perp v_1, \\ &\vdots \\ v_n &\in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\|=1} \operatorname{Var}_{[u,0]}(\overline{f}) \quad s.t. \quad u \perp \{v_1, \ldots, v_{n-1}\}. \end{aligned} \tag{11}$$

我们将相应的方差 $\lambda_i = \operatorname{Var}_{[v_i,0]}(\overline{f})$ 称为主方差，并将二元组 (λ, V) 称为 f 的 *PW-PCA*，其中 $\lambda =$（λ_1、……、λ_n）$\in \mathbb{R}^n$，且 $V =$（v_1、……、v_n）$\in \mathbb{R}^{n \times n}$。

换言之，主轴 v_1、……、v_n 为 $\mathbb{R}^n$ 提供了一个标准正交基，捕捉齐次化 f 较大的方向；而主方差 λ_1、……、λ_n 则给出了相应的多项式加权方差。

3.4 Alg多项式加权主成分分析算法

现在，我们将发展一些理论，为计算定理 8 中引入的多项式加权主成分分析（PW-PCA）提供一种高效算法。第一个理论结果将式（11）中的一系列优化问题归结为对多项式加权协方差矩阵的前导主 $n \times n$ 子矩阵进行特征分解（定理 7）。

命题 9(PW-PCA 通过特征分解)。对于任意 $f \in \mathbb{R}[x_1, \ldots, x_n]$，以下各项等价：

1. (λ, V) 是 f 的 *PW-PCA*
2. $V \operatorname{diag}(\lambda) V^\top$ 是 $\operatorname{Cov}(\overline{f})_{1:n,1:n}$ 的正交特征分解

其中 *ei* 征分解按非递增顺序排列特征值，即 $\lambda_1 \geq \cdots \geq \lambda_n$.

该定理类似于以下事实：数据点的主轴由协方差矩阵的特征向量给出（如第3.1节所述），并且可以用基本相同的方式加以证明。一种典型方法是通过拉格朗日乘子刻画（11）中的约束优化问题的解，例如 [18, 第1节1]中所述。这里我们提供另一种采用常见方法的初等证明（用特征值和特征向量表示协方差），这种方法有时能带来更深入的理解。

定理 *9* 的证明。利用定理 7 中所述的多项式加权方差与协方差之间的关系，我们首先有

$$\operatorname{Var}_{[u,0]}(\overline{f}) = [u,0]^\top \operatorname{Cov}(\overline{f})[u,0] = \begin{bmatrix} u \\ 0 \end{bmatrix}^\top \begin{bmatrix} \operatorname{Cov}(\overline{f})_{1:n,1:n} & \operatorname{Cov}(\overline{f})_{1:n,n+1} \\ \operatorname{Cov}(\overline{f})_{n+1,1:n} & \operatorname{Cov}(\overline{f})_{n+1,n+1} \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = u^\top \operatorname{Cov}(\overline{f})_{1:n,1:n} u.$$

由于对于任意 $u \in \mathbb{R}^n$，$\mathrm{Var}_{[u,0]}(f) \geq 0$，因此 $\mathrm{Cov}(f)_{1:n,\,1:n}$ 是半正定的，其特征值均为非负值。现在我们逐一证明两个方向：

- (2) $\Longrightarrow$ (1)。

  假设 $V\mathrm{diag}(\lambda)V^\top$ 是 $\mathrm{Cov}(f)_{1:n,1:n}$ 的正交特征分解，即 $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ 是其特征值，对应的正交归一特征向量为 $v_1, \ldots, v_n$，它们由 V 的各列给出。代入后可得

$$\mathrm{Var}_{[u,0]}(\overline{f}) = u^\top \mathrm{Cov}(\overline{f})_{1:n,1:n} u = u^\top V \,\mathrm{diag}(\lambda) V^\top u = \lambda_1 (v_1^\top u)^2 + \cdots + \lambda_n (v_n^\top u)^2.$$

  因此，对于任意满足 $\|u\| = 1$ 的 u，都有如下上界

$$\mathrm{Var}_{[u,0]}(\overline{f}) \leq \lambda_1 (v_1^\top u)^2 + \cdots + \lambda_1 (v_n^\top u)^2 = \lambda_1 \|V^\top u\|^2 = \lambda_1 \|u\|^2 = \lambda_1,$$

  由于 $\mathrm{Var}_{[v_1,0]}(f) = \lambda_1 (v_1^\top v_1)^2 + \cdots + \lambda_n (v_n^\top v_1)^2 = \lambda_1$，该上界在 v_1 处取到，因此

$$v_1 \in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\| = 1} \mathrm{Var}_{[u,0]}(\overline{f}).$$

  类似地，对于任意满足 $\|u\| = 1$ 且满足 $u \perp v_1$ 的 u，都有如下上界

$$\mathrm{Var}_{[u,0]}(\overline{f}) = \lambda_2 (v_2^\top u)^2 + \cdots + \lambda_n (v_n^\top u)^2 \leq \lambda_2 (v_2^\top u)^2 + \cdots + \lambda_2 (v_n^\top u)^2 = \lambda_2 \|V^\top u\|^2 = \lambda_2 \|u\|^2 = \lambda_2,$$

  由于 $\mathrm{Var}_{[v_2,0]}(f) = \lambda_1 (v_1^\top v_2)^2 + \cdots + \lambda_n (v_n^\top v_2)^2 = \lambda_2$，该上界在 v_2 处取到，因此

$$v_2 \in \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\| = 1} \mathrm{Var}_{[u,0]}(\overline{f}) \quad \text{s.t.} \quad u \perp v_1.$$

  以类似的方式可知，$v_3, \ldots, v_n$ 可解出（11）中的相应优化问题，因此，$v_1, \ldots, v_n$ 构成 f 的一组主轴，其对应的主方差为 $\lambda_1, \ldots, \lambda_n$。至此，这一方向的证明完成。

- (1) $\Longrightarrow$ (2)。

  假设 (λ, V) 是 f 的 PW-PCA，即，V 的列 $v_1, \ldots, v_n$ 是 f 的一组主轴，其对应的主方差为 $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$。接下来，回顾一下，$\mathrm{Cov}(f)_{1:n,1:n}$ 是一个对称矩阵，因此总存在正交特征分解 $\tilde{V}\,\mathrm{diag}(\tilde{\lambda})\tilde{V}^\top$。那么，根据与上文相同的论证，对于任意满足 $\|u\| = 1$ 的 u，我们有如下上界

$$\mathrm{Var}_{[u,0]}(\overline{f}) = u^\top \mathrm{Cov}(\overline{f})_{1:n,1:n} u = u^\top \tilde{V} \,\mathrm{diag}(\tilde{\lambda}) \tilde{V}^\top u = \tilde{\lambda}_1 (\tilde{v}_1^\top u)^2 + \cdots + \tilde{\lambda}_n (\tilde{v}_n^\top u)^2 \leq \tilde{\lambda}_1,$$

  该上界由 $\tilde{v}_1$ 达到。现在注意，与 λ_1 相关联的 $\tilde{}$ 特征子空间中的任意单位向量（即具有相同特征值的特征向量所张成的空间）也能达到该上界，而且除此之外的任何其他单位向量都无法达到该上界。因此，我们有

$$\mathrm{span}(\tilde{v}_i : \tilde{\lambda}_i = \tilde{\lambda}_1) \cap \mathbb{S}^{n-1} = \operatorname*{argmax}_{u \in \mathbb{R}^n : \|u\| = 1} \mathrm{Var}_{[u,0]}(\overline{f}),$$

  因此可得，$v_1 \in \mathrm{span}(\tilde{v}_i : \lambda_i = \lambda_1)$ 且 $\lambda_1 = \lambda_1$。因此，v_1 是 $\mathrm{Cov}(f)_{1:n,1:n}$ 的一个特征向量，其对应特征值为 λ_1。

  通过类似的论证可知，$v_2, \ldots, v_n$ 都是 $\mathrm{Cov}(f)_{1:n,1:n}$ 的特征向量，其对应特征值为 $\lambda_2 \geq \cdots \geq \lambda_n$。换言之，对于每个 $i \in \{1, \ldots, n\}$，$\mathrm{Cov}(f)_{1:n,1:n} v_i = \lambda_i v_i$，简洁地写成矩阵形式为

$$\mathrm{Cov}(\overline{f})_{1:n,1:n} V = [\mathrm{Cov}(\overline{f})_{1:n,1:n} v_1, \ldots, \mathrm{Cov}(\overline{f})_{1:n,1:n} v_n] = [\lambda_1 v_1, \ldots, \lambda_n v_n] = V \,\mathrm{diag}(\lambda).$$

  现在回想一下 根据构造，$v_1, \ldots, v_n$ 是正交归一的，因此 V 是一个正交矩阵，于是

$$V \,\mathrm{diag}(\lambda) V^\top = \mathrm{Cov}(\overline{f})_{1:n,1:n} V V^\top = \mathrm{Cov}(\overline{f})_{1:n,1:n},$$

  即，$V\mathrm{diag}(\lambda)V^\top$ 是 $\mathrm{Cov}(f)_{1:n,1:n}$ 的正交特征分解。至此，这一方向的证明完成。

□

接下来需要找到一种高效计算多项式加权协方差矩阵的方法。值得注意的是，其定义涉及多元曲面积分，直接实现并不容易。这里我们给出一个易于实现且高效的显式代数表达式。为此，我们先回顾并引入一些标准记号。

记号 10。 我们使用以下记号：

- Γ 表示伽马函数。
- !! 表示双阶乘。
- e_i 表示$\mathbb{R}^n$中的第 i个标准基向量。
- $\eta(s) = \begin{cases} s & \text{；如果 } s \text{ 为奇数} \\ s-1 & \text{，如果为偶数，则将“舍入” } s \text{ 为不} \end{cases}$

大于它的最近奇数。

现在给出我们将用于获得计算 PW-PCA 的高效算法的多项式加权协方差的代数表达式。该表达式表明，多项式的每一项只会对多项式加权协方差的少数几个条目产生贡献。

定理 11 （多项式加权协方差的表达式）。 设 $f \in \mathbb{R}[x_1, \ldots, x_n]$ 是一个次数为 d 的齐次多项式。设 $a \cdot x^\mu$ 表示 f^2 中的每一项，即 $f^2 = \sum a \cdot x^\mu$。则有

$$\mathrm{Cov}(f) \quad = \quad \frac{\pi^{n/2}}{2^d\,\Gamma(n/2+d+1)} \cdot \sum a \cdot \Psi(x^\mu) \tag{12}$$

其中 $\Psi(x^\mu) \in \mathbb{N}_{\geq 0}^{n \times n}$ 由下式给出

$$\Psi(x^\mu) = \left(\prod_{k=1}^{n} \eta(\mu_k)!!\right) \begin{cases} \mathrm{diag}(\mu) + I_n, & \textit{if } \mu \textit{ has no odd entries,} \\ e_u e_v^\top + e_v e_u^\top, & \textit{if } \mu \textit{ has two odd entries at indices } u \textit{ and } v, \\ 0_{n\times n}, & \textit{otherwise.} \end{cases} \tag{13}$$

定理 11 的证明结合了在球面上对单项式积分的标准公式 [13]（即计算球面上均匀分布的统计矩），并仔细分析多项式加权协方差中出现的单项式指数。具体而言，我们的分析揭示并利用了每个多项式项对多项式加权协方差的贡献所具有的稀疏性。

定理 11 的证明。 在定义 7 中应用 (9) 并进行因式分解，可得

$$\mathrm{Cov}(f) = \int_{\mathbb{S}^{n-1}} \left(\sum a \cdot x^\mu\right) xx^\top \, d\mu(x) = \sum a \cdot \int_{\mathbb{S}^{n-1}} x^\mu \, xx^\top \, d\mu(x) = \sum a \cdot \mathrm{Cov}(x^\mu). \tag{14}$$

因此，问题归结为计算次数为 $2d$ 的单项式 x^μ 的协方差矩阵。也就是说，我们需要计算 $\mathrm{Cov}(x^\mu)$，其条目由下式给出

$$\mathrm{Cov}(x^\mu)_{ij} = \left(\int_{\mathbb{S}^{n-1}} x^\mu \, xx^\top \, d\mu(x)\right)_{ij} = \int_{\mathbb{S}^{n-1}} x^\mu \, x_i x_j \, d\mu(x) = \int_{\mathbb{S}^{n-1}} x^{\mu+e_i+e_j} \, d\mu(x).$$

关于这些单项式在单位球面上积分的公式见 [13]。具体而言，应用 [13], 中的主定理，我们得到以下两种情况

$$\begin{aligned} \text{if } \mu + e_i + e_j \text{ has no odd entries}: \quad & \mathrm{Cov}(x^\mu)_{ij} = 2\frac{\prod_{k=1}^{n} \Gamma(1/2 + (\mu + e_i + e_j)_k/2)}{\Gamma(\sum_{k=1}^{n}(1/2 + (\mu + e_i + e_j)_k/2))}, \\ \text{otherwise}: \quad & \mathrm{Cov}(x^\mu)_{ij} = 0. \end{aligned} \tag{15}$$

由于.....的和 $\mu+e_i+e_j$ 的条目始终为 $2d+2$，因此第一种情况下的分母可简化为 t o

$$\Gamma\left(\sum_{k=1}^{n}(1/2+(\mu+e_i+e_j)_k/2)\right)=\Gamma(n/2+(2d+2)/2)=\Gamma(n/2+d+1). \tag{16}$$

此外，对于这种“仅含偶数条目”的情形，$(\mu+e_i+e_j)/2$ 的所有条目都是整数，因此分子也可以改写为

$$\begin{aligned}\prod_{k=1}^{n}\Gamma(1/2+(\mu+e_i+e_j)_k/2)&=\prod_{k=1}^{n}\left(\frac{(2(\mu+e_i+e_j)_k/2-1)!!}{2^{(\mu+e_i+e_j)_k/2}}\sqrt{\pi}\right)\\&=\pi^{n/2}\frac{\prod_{k=1}^{n}(2(\mu+e_i+e_j)_k/2-1)!!}{2^{\sum_{k=1}^{n}(\mu+e_i+e_j)_k/2}}=\frac{\pi^{n/2}}{2^{d+1}}\prod_{k=1}^{n}((\mu+e_i+e_j)_k-1)!!.\end{aligned} \tag{17}$$

将 (16) 和 (17) 代回 (15) 并进行因式分解，得到

$$\mathrm{Cov}(x^\mu)=\frac{\pi^{n/2}}{2^d\,\Gamma(n/2+d+1)}\cdot\Psi(x^\mu), \tag{18}$$

其中，$\Psi(x^\mu)\in\mathbb{N}_{\geq 0}^{n\times n}$ 的条目定义为

$$\Psi(x^\mu)_{ij}=\begin{cases}\prod_{k=1}^{n}((\mu+e_i+e_j)_k-1)!!, & \text{if } \mu+e_i+e_j \text{ has no odd entries,}\\ 0, & \text{otherwise.}\end{cases} \tag{19}$$

请注意，$\mu+e_i+e_j$ 最终不含奇数条目只有两种情形：

1. μ 不含奇数条目。在这种情形下，当 $i=j$ 时，$\mu+e_i+e_j$ 不含奇数条目，因此 $\Psi(x^\mu)$ 是一个对角矩阵，其条目为

$$\Psi(x^\mu)_{ii}=\prod_{k=1}^{n}((\mu+e_i+e_i)_k-1)!!=\underbrace{((\mu_i+2)-1)!!}_{(\mu_i+1)(\mu_i-1)!!}\prod_{k\neq i}(\mu_k-1)!!=(\mu_i+1)\prod_{k=1}^{n}(\mu_k-1)!!.$$

   因此，对于这种情形，我们有

$$\Psi(x^\mu)=\left(\prod_{k=1}^{n}(\mu_k-1)!!\right)\cdot(\mathrm{diag}(\mu)+I_n)=\left(\prod_{k=1}^{n}\eta(\mu_k)!!\right)\cdot(\mathrm{diag}(\mu)+I_n). \tag{20}$$

2. μ 在索引 u 和 v处有两个奇数条目。在这种情形下，当 $(i,j)=(u,v)$ 或 $(i,j)=(v,u)$ 时，$\mu+e_i+e_j$ 没有奇数条目，因此 $\Psi(x^\mu)$ 只有以下两个非零条目

$$\Psi(x^\mu)_{uv}=\Psi(x^\mu)_{vu}=\prod_{k=1}^{n}((\mu+e_u+e_v)_k-1)!!=((\mu_u+1)-1)!!((\mu_v+1)-1)!!\prod_{k\neq u,v}(\mu_k-1)!!.$$

   因此，对于这种情形，我们有

$$\Psi(x^\mu)=((\mu_u+1)-1)!!((\mu_v+1)-1)!!\prod_{k\neq u,v}(\mu_k-1)!!\cdot(e_ue_v^\top+e_ve_u^\top)=\left(\prod_{k=1}^{n}\eta(\mu_k)!!\right)\cdot(e_ue_v^\top+e_ve_u^\top). \tag{21}$$

对于其他情形，$\mu+e_i+e_j$ 始终至少有一个奇数项，因此在这些情况下有 $\Psi(x^\mu)=0_{n\times n}$。证明通过合并方程 (14)、(18)、(19)、(20) 和 (21)，并进行因式分解，得以完成。 □

现在，我们有了高效计算 PW-PCA 的方法。具体而言，根据定理 9，我们只需对多项式加权协方差矩阵的主导主 $n \times n$ 子矩阵进行正交特征分解；该分解可通过定理 11 中的表达式高效计算。由此得到以下算法。

算法 12（PW-PCA）。

输入： $f \in \mathbb{R}[x_1, \ldots, x_n]$ **输出：** (λ, V)，f 的 *PW-PCA 1.* $f \leftarrow x_{n+1}^{\deg f} f\ (x_1/x_{n+1}, \ldots, x_n/x_{n+1})$，$f$ 的齐次化 *2.* $C \leftarrow$ 协方差(f)，通过对 f 应用定理 *11* 得到 *3.* $C \leftarrow C_{1:n,\ 1:n}$ *4.* $\lambda \leftarrow$ n 的特征值列表，按非递增顺序排列 *5.* $V \leftarrow$ 矩阵，其列为 C 对应的特征向量

—

我们通过贯穿示例来演示上述算法。我们分别以精确方式（精确算术）和近似方式（双精度浮点运算）执行了该算法；对于近似情况，我们仅展示 3 位十进制位数。

例 13（贯穿示例续）。令 $f = -27x_1^3 + 27x_2^2x_3 - 9x_3$ 为 *PW-PCA* 算法（算法 *12*）的输入。下面我们执行该算法的各个步骤。

1. $\overline{f} \leftarrow -27x_1^3 + 27x_2^2x_3 - 9x_3x_4^2$。

2. $\overline{C} \leftarrow \frac{\pi^2}{2^3 \cdot 5!} \begin{bmatrix} 78489 & 0 & -2916 & 0 \\ 0 & 20655 & 0 & 0 \\ -2916 & 0 & 16767 & 0 \\ 0 & 0 & 0 & 12879 \end{bmatrix}$

$\approx \begin{bmatrix} 806.933 & 0.000 & -29.979 & 0.000 \\ 0.000 & 212.351 & 0.000 & 0.000 \\ -29.979 & 0.000 & 172.379 & 0.000 \\ 0.000 & 0.000 & 0.000 & 132.407 \end{bmatrix}$

3. $C \leftarrow \frac{\pi^2}{2^3 \cdot 5!} \begin{bmatrix} 78489 & 0 & -2916 \\ 0 & 20655 & 0 \\ -2916 & 0 & 16767 \end{bmatrix}$

$\approx \begin{bmatrix} 806.933 & 0.000 & -29.979 \\ 0.000 & 212.351 & 0.000 \\ -29.979 & 0.000 & 172.379 \end{bmatrix}$

4. $\lambda \leftarrow \frac{\pi^2}{2^3 \cdot 5!} \begin{bmatrix} 47628 + 243\sqrt{16273} & 20655 & 47628 - 243\sqrt{16273} \end{bmatrix}$

$\approx [808.346,\ 212.351,\ 170.966]$

5. $V \leftarrow \begin{bmatrix} \frac{\sqrt{32546 - 254\sqrt{16273}}}{24} + \frac{127\sqrt{32546 - 254\sqrt{16273}}\sqrt{16273}}{390552} & 0 & \frac{\sqrt{32546 + 254\sqrt{16273}}}{24} - \frac{127\sqrt{32546 + 254\sqrt{16273}}\sqrt{16273}}{390552} \\ 0 & 1 & 0 \\ -\frac{\sqrt{32546 - 254\sqrt{16273}}\sqrt{16273}}{32546} & 0 & \frac{\sqrt{32546 + 254\sqrt{16273}}\sqrt{16273}}{32546} \end{bmatrix}$

$\approx \begin{bmatrix} 0.999 & 0.000 & 0.047 \\ 0.000 & -1.000 & 0.000 \\ -0.047 & 0.000 & 0.999 \end{bmatrix}$

结果　算法 *12* 的输出随后为（λ、V），即以近似形式 b 给出的 *PW-PCA* y

$$(\lambda, V) \approx \left([808.346,\ 212.351,\ 170.966], \begin{bmatrix} 0.999 & 0.000 & 0.047 \\ 0.000 & -1.000 & 0.000 \\ -0.047 & 0.000 & 0.999 \end{bmatrix} \right).$$

请注意，为了在计算机上以算法方式表示（λ, V）的精确版本，我们需要在有理数的某个代数域扩张中进行运算。

4 PW-PCA 正交等价证书

本节将建立所需的数学理论，并据此提出相应算法，以解决本文的主要问题（问题 2）。为此，我们将探讨正交等价与上一节介绍的多项式加权主成分分析（PW-PCA）之间的联系。核心结果见定理 17。我们将先准确陈述这些核心结果，再通过若干引理给出详细证明。最后，我们将把这些核心结果改写为一个算法。

按照主成分分析理论中的惯例，核心结果将针对主轴唯一确定（仅相差符号）的多项式陈述（见定义 16）；等价地，也将针对主方差各不相同的多项式陈述。正如预期，几乎所有多项式的主方差都各不相同；换言之，一般的多项式具有互不相同的主方差。本文将始终假设所有多项式均满足这一条件，具体表述见下文的假设 14。

假设 14 (全局假设：主方差互不相同)。 从现在起，我们假设每个多项式都满足以下条件：其主方差 $\lambda_1, \ldots, \lambda_n$ 互不相同，即 $\lambda_1 > \cdots > \lambda_n$。

如上所述，主方差互不相同这一性质对于一个扎里斯基稠密的多项式集合成立，也就是说，一般的多项式都具有这一性质。下文的命题 15 将对此作出精确表述。

命题 15 (主方差互不相同的一般性)。 设 $n \geq 2$ 和 $d \geq 2$。假设 *14* 是一般的，其含义是，对于次数至多为 d的几乎所有 $f \in \mathbb{R}[x_1, \ldots, x_n]$，该假设都成立。

证明。 证明在概念上很直接，基于一个标准论证，该论证将表明，不具有互不相同主方差的多项式集合在扎里斯基拓扑下是闭集。然而，由于证明涉及若干冗长且繁琐的代数运算，且这些运算并不能带来显著的理论见解，我们选择将证明留至附录A。 □

为了精确且系统地处理固有的符号歧义，我们引入一个有限群，称为“符号翻转群”，并仿照正交群（见定义 1），为相关的作用和等价关系给出记号。

定义 16 （符号翻转群、作用和等价关系）。

1.（群） n 维的符号翻转群，记作 $\pm(n)$，定义为

$$\pm(n) = \{-1, 1\}^n. \tag{22}$$

2.（作用） $\sigma \in \pm(n)$ 对 $p \in \mathbb{R}[x_1, \ldots, x_n]$ 的作用 记作 $\sigma \bullet p$，定义为 y

$$(\sigma \bullet p)(x) = p(\mathrm{diag}(\sigma)x).$$

3.（等价 等价）我们称 f、$g \in \mathbb{R}[x_1, \ldots, x_n]$ 是符号翻转等价的，并记作 $f \sim_{\pm(n)} g$，i f

$$\exists_{\sigma \in \pm(n)}\ g = \sigma \bullet f.$$

现在我们可以陈述本节的关键理论结果了。

定理 17（关键结果：通过 PW-PCA 实现正交等价）。假设 f、$g \in \mathbb{R}[x_1, \ldots, x_n]$ 均满足一般情形下的定理 ***14***，并且 (λ_f, V_f) 和 (λ_g, V_g) 分别是 f 和 g 的 ***PW-PCA***。则有

C1: $f \sim_{O(n)} g \implies \lambda_f = \lambda_g$,ˆ
C2: $f \sim_{O(n)} g \iff f \sim_{\pm(n)}$ ˆ
g,ˆ其中 $f = V_f \bullet f$ 且 ˆ$g = V_g \bullet g$。

此外，我们还有

C3: $\sigma \bullet \hat{f} = \hat{g} \implies R \bullet f = g \quad where\ R = V_f \operatorname{diag}(\sigma) V_g^\top$.

在证明定理 17 之前，我们先在下文的备注 18 中汇总几条重要观察。

备注 18。

C1：这一论断对 f 与 g之间的正交等价性提供了部分检验。取其逆否命题可知，若 $\lambda_f \neq \lambda_g$，则 f 与 g 不正交等价。

由于 λ 属于 f 和 g 的系数域的实代数扩张，因此检验涉及比较实代数数。通常，可以不必直接处理实代数数，而改为比较相应协变矩阵的特征多项式的系数。

C2：该结论提供了一种完全检验，用于检验 f 与 g 是否正交等价。若仅根据定义检验$f \sim_{O(n)} g$，就需要在无限集合 $O(n)$ 中搜索满足 $R \bullet f = g$ 的 R。显然，这在算法上是不可能的。该结论告诉我们，可以改为检验ˆ $\hat{f} \sim_{\pm(n)}$ ˆg，这需要在有限集合 $\pm(n)$ 中搜索满足 $\sigma \bullet f =$ ˆg 的 σ。显然，这在算法上是可行的。因此，该结论可视为将无限搜索空间 $O(n)$缩减为有限搜索空间 $\pm(n)$。

ˆ这一论断也可以表述为：对于每个 $h \in \mathbb{R}[x_1, \ldots, x_n]$，多项式 h是 h 的轨道在忽略符号差异下的规范形式，而 V_h 是规范化简器，也就是说，是 $O(n)$ 中的一个ˆ元素，它作用于 h，将其变换为其规范形式 h。

ˆ**C3**：这一论断展示了一种简单的方法，用于找到一个 $R \in O(n)$，使得当$f \sim_{\pm(n)}$ ˆg时，$R \bullet f = g$：找到 σ 使得 $\sigma \bullet$ ˆ$f =$ ˆg。于是我们有 $R = V_f \operatorname{diag}(\sigma) V_g^\top$。

本节的其余部分将专门用于证明关键定理（定理 17）。该定理的证明被分解为若干个较小的引理，即引理 19、20、22、23 和 24；读者也可以单独理解这些引理。在证明这些引理之前，我们将概述这些引理的使用方式，展示它们的自顶向下的依赖关系。

- 定理 17 的证明使用了引理 19 和引理 24。
- 引理 24 的证明使用了引理 23。
- 引理 23 的证明使用了引理 20 和引理 22。

现在让我们从引理 19 开始。它指出，对于几乎所有的 f，其主方差是唯一的，主轴除符号翻转外也是唯一的。

引理 19（PW-PCA 的唯一性，除符号翻转外）。 假设 $f \in \mathbb{R}[x_1, \ldots, x_n]$ 满足一般性定理 *14*。若 (λ_1, V_1) 和 (λ_2, V_2) 是 f 的两个 *PW-PCA*，则

$$\lambda_1 = \lambda_2 \qquad and \qquad \exists_{\sigma \in \pm(n)} \;\; V_2 = V_1 \operatorname{diag} \sigma.$$

定理 *19* 的证明。根据 PW-PCA 的定义，显然有 $\lambda_1 = \lambda_2$。由于 f 满足一般性假设， C_f 的特征值互不相同。因此，每个特征值对应的特征子空间维数均为一。于是，特征向量在符号意义下唯一。因此，存在一些符号翻转$\sigma \in \{-1, 1\}^n$，使得 $V_2 = V_1 \operatorname{diag} \sigma$.

□

以下四个引理探讨一个自然的问题：以下对象在正交群作用下如何变化?

- 齐次化（引理 20）
- 齐次多项式的 PW 协方差（引理 22）
- 齐次化多项式的 PW 协方差（引理 23）
- PW-PCA（引理 24）

这些引理表明，在正交群作用下，它们都具有等变性。

引理 20（齐次化的正交等变性）。 设 $f \in \mathbb{R}[x_1, \ldots, x_n]$ 且 $R \in O(n)$。则有

$$\overline{R \bullet f} \;=\; \overline{R} \bullet \overline{f} \quad where \quad \overline{R} = \begin{bmatrix} R & \\ & 1 \end{bmatrix} \in O(n+1).$$

证明。 设 $d = \deg f$。则 $d = \deg g$。注意

$$\begin{aligned}
\left(\overline{R \bullet f}\right)(\overline{x}) &= \left(\overline{R \bullet f}\right)(x, x_{n+1}) && \text{where } \overline{x} = (x, x_{n+1}) \\
&= (R \bullet f)(x/x_{n+1}) x_{n+1}^d && \text{from the definition of homogenization (10)} \\
&= f(R(x/x_{n+1})) x_{n+1}^d && \text{since } g = R \bullet f \\
&= f((Rx)/x_{n+1}) x_{n+1}^d && \text{since } R(x/x_{n+1}) = (Rx)/x_n \\
&= \overline{f}(Rx, x_{n+1}) && \text{from the definition of homogenization (10)} \\
&= \overline{f}(\overline{R}\overline{x}) && \text{from the definition of } \overline{R} \\
&= \overline{R} \bullet \overline{f}(\overline{x}) && \text{since } \overline{R} \in O(n+1).
\end{aligned}$$

注意，上述等式对每个 x 都成立。因此有

$$\overline{R \bullet f} \;=\; \overline{R} \bullet \overline{f}.$$

引理得证。 □

现在定义 R 属于 $O(n)$ 的元素在对称矩阵上的作用。

定义21（对称矩阵上的正交作用）。 作用 $R \in O(n)$ 在对称矩阵$M \in \mathbb{R}^{n \times n}$上的作用记作 $R \bullet M$，定义为

$$R \bullet M \;=\; R^\top M R.$$

当将 M 视为二次型的一个表示时，这是标准定义。

我们将在下文引理的陈述和证明中采用上述定义。回顾一下，对于齐次多项式 f，其多项式加权协方差 $\mathrm{Cov}(f)$ 是一个实对称矩阵。

引理 22 （齐次多项式的 PW 协方差正交等变性）。 设 $f \in \mathbb{R}[x_1, \ldots, x_n]$是一个齐次多项式，且设 $R \in O(n)$。则有

$$\mathrm{Cov}(R \bullet f) \quad = \quad R \bullet \mathrm{Cov}(f).$$

定理 *22* 的证明。注

$$\begin{aligned}
\mathrm{Cov}(R \bullet f) &= \int_{\mathbb{S}^{n-1}} f(Rx)^2 \, xx^\top \, d\mu(x) && \text{from Definition 7} \\
&= \int_{\mathbb{S}^{n-1}} f(Rx)^2 \, (R^\top R) xx^\top (R^\top R) \, d\mu(x) && \text{since } R^\top R = I \\
&= \int_{\mathbb{S}^{n-1}} f(Rx)^2 \, R^\top (Rx)(Rx)^\top R \, d\mu(x) && \text{by re-associating} \\
&= R^\top \left(\int_{\mathbb{S}^{n-1}} f(Rx)^2 \, (Rx)(Rx)^\top \, d\mu(x) \right) R && \text{since } R \text{ does not depend on } x \\
&= R^\top \left(\int_{R\mathbb{S}^{n-1}} f(y)^2 \, yy^\top \, d\mu(R^\top y) \right) R && \text{by the change of variable } y = Rx \\
&= R^\top \left(\int_{\mathbb{S}^{n-1}} f(y)^2 \, yy^\top \, d\mu(y) \right) R && \text{since } d\mu(R^\top y) = d\mu(y) \\
&= R^\top \, \mathrm{Cov}(f) R. && \text{from Definition 7} \\
&= R \bullet \mathrm{Cov}(f) && \text{from Definition 21.}
\end{aligned}$$

我们已经证明了该引理。

□

下一个引理与上面的引理 22 类似，但有一种情况例外：我们从一个非齐次多项式开始并将其齐次化。

引理 23 (齐次化多项式的 PW 协方差正交等变性)。 设 $f \in \mathbb{R}[x_1, \ldots, x_n]$和 $R \in O(n)$。那么我们有

$$\mathrm{Cov}(\overline{R \bullet f})_{1:n,1:n} \quad = \quad R \bullet \mathrm{Cov}(\overline{f})_{1:n,1:n}$$

证明。注

$$\begin{aligned}
\mathrm{Cov}(\overline{R \bullet f}) &= \mathrm{Cov}(\overline{R} \bullet \overline{f}) && \text{from Lemma 20} \\
&= \overline{R}^\top \, \mathrm{Cov}(\overline{f}) \overline{R} && \text{from Lemma 22} \\
&= \begin{bmatrix} R & \\ & 1 \end{bmatrix}^\top \begin{bmatrix} \mathrm{Cov}(\overline{f})_{1:n,1:n} & \mathrm{Cov}(\overline{f})_{1:n,n+1} \\ \mathrm{Cov}(\overline{f})_{n+1,1:n} & \mathrm{Cov}(\overline{f})_{n+1,n+1} \end{bmatrix} \begin{bmatrix} R & \\ & 1 \end{bmatrix} && \text{by showing the parts} \\
&= \begin{bmatrix} R^\top \mathrm{Cov}(\overline{f})_{1:n,1:n} \, R & R^\top \mathrm{Cov}(\overline{f})_{1:n,n+1} \\ \mathrm{Cov}(\overline{f})_{n+1,1:n} R & \mathrm{Cov}(\overline{f})_{n+1,n+1} \end{bmatrix} && \text{by multiplying.}
\end{aligned}$$

通过令右上部分相等，我们得到

$$\mathrm{Cov}(\overline{R \bullet f})_{1:n,1:n} \quad = \quad R^\top \, \mathrm{Cov}(\overline{f})_{1:n,1:n} \, R \quad = \quad R \bullet \mathrm{Cov}(\overline{f})_{1:n,1:n}.$$

我们已经证明了该引理。 □

现在我们证明最后一个引理，该引理将确立获得定理17证明所需的 PW-PCA 的正交等变性。

引理 24 (PW-PCA 的正交等变性)。 设 $f \in \mathbb{R}[x_1, \ldots, x_n]$和 $R \in O(n)$。则以下条件等价。

1. (λ, V) 是 f 的 *PW-PCA*。*2.* $(\lambda, R^\top V)$ 是 $R \bullet f$ 的 *PW-PCA*。

证明。 假设 (λ, V) 是 f 的一个 PW-PCA。那么以下条件等价:

$$
\begin{aligned}
& (\lambda, V) \text{ is a PW-PCA of } f && \\
\iff\ & V \operatorname{diag}(\lambda) V^\top = \operatorname{Cov}(\overline{f})_{1:n,1:n} && \text{from Theorem 9} \\
\iff\ & R^\top V \operatorname{diag}(\lambda) V^\top R = R^\top \operatorname{Cov}(\overline{f})_{1:n,1:n} R && \text{since } R \text{ is invertible} \\
\iff\ & \left(R^\top V\right) \operatorname{diag}(\lambda) \left(R^\top V\right)^\top = R^\top \operatorname{Cov}(\overline{f})_{1:n,1:n} R && \text{since } V^\top R = \left(R^\top V\right)^\top \\
\iff\ & \left(R^\top V\right) \operatorname{diag}(\lambda) \left(R^\top V\right)^\top = \operatorname{Cov}(\overline{R \bullet f})_{1:n,1:n} && \text{from Lemma 23} \\
\iff\ & (\lambda, R^\top V) \text{ is a PW-PCA of } R \bullet f && \text{from Theorem 9.}
\end{aligned}
$$

由此我们证明了该引理。 □

现在我们准备证明定理 17，这是本节的关键定理。证明将关键地用到 PW-PCA 在符号翻转意义下的唯一性（引理 19）以及 PW-PCA 的正交等变性（引理 24）。

定理 *17* 的证明。 设 $f, g \in \mathbb{R}[x_1, \ldots, x_n]$ 分别具有 PW-PCA (λ_f, V_f) 和 (λ_g, V_g)，其中 f和 g 均满足定理 14。我们逐一证明各项结论:

C1. $f \sim_{O(n)} g \implies \lambda_f = \lambda_g$。假设 $f \sim_{O(n)} g$。那么存在 $R \in O(n)$ 使得 $R \bullet f = g$。由定理 24 可知，λ_f 因而是 g 的主方差；最后由定理 19 可得 $\lambda_f = \lambda_g$，从而完成此项结论的证明。

C2. $f \sim_{O(n)} g \iff f \sim_{\pm(n)} \hat{g}$.

对于 $\implies$ 方向，假设 $f \sim_{O(n)} g$。那么存在 $R \in O(n)$，使得 $R \bullet f = g$。根据定理24， $R^\top V_f$ 是 g 的一组主轴；接着根据定理 19，存在 $\sigma \in \pm(n)$，使得 $V_g = R^\top V_f \operatorname{diag}(\sigma)$。因此， $RV_g = V_f \operatorname{diag}(\sigma)$，所以

$$
\begin{aligned}
(\sigma \bullet \hat{f})(x) &= \hat{f}(\operatorname{diag}(\sigma) x) && \text{根据 } \sigma \bullet \text{的定义} \quad \hat{f} \\
&= f(V_f \operatorname{diag}(\sigma) x) && \text{根据定义} \quad \hat{f} \\
&= f(R V_g x) && \text{因为 } RV_g = V_f \operatorname{diag}(\sigma) \\
&= g(V_g x) && \text{因为 } R \bullet f = g \\
&= \hat{g}(x) && \text{根据} \hat{g} \text{的定义}
\end{aligned}
$$

因此，我们有 $f \sim_{\pm(n)} \hat{g}$，这就完成了这一方向的证明。

对于 $\impliedby$ 方向，假设 $f \sim_{\pm(n)} \hat{g}$。那么存在 $\sigma \in \pm(n)$，使得 $\sigma \bullet f = \hat{g}$，因此

$$
\begin{aligned}
g(x) &= g(V_g V_g^\top x) && \text{因为 } V_g V_g^\top = I_n \\
&= \hat{g}(V_g^\top x) && \text{根据} \hat{g} \text{的定义} \\
&= \hat{f}(\operatorname{diag}(\sigma) V_g^\top x) && \hat{} \text{因为 } \sigma \bullet f = \hat{g} \\
&= f(V_f \operatorname{diag}(\sigma) V_g^\top x) && \hat{} \text{ 根据 } f \text{ 的定义} \\
&= f(Rx) && \text{通过定义 } R = V_f \operatorname{diag}(\sigma) V_g^\top \\
&= (R \bullet f)(x) && \text{根据 } R \bullet f \text{ 的定义,}
\end{aligned}
$$

其中，由于它是正交矩阵的乘积， $R \in O(n)$。因此，我们得到 $f \sim_{O(n)} g$，从而完成了这一方向的证明。

C3. $\sigma \bullet \hat{f} = \hat{g} \implies R \bullet f = g$ 其中 $R = V_f \,\mathrm{diag}(\sigma) V_g^\top$。

假设 $\hat{}$设 $\sigma \bullet f = \hat{g}$。那么，根据 **C2** 的 $\Longleftarrow$ 方向中的类似论证，我们有

$$(R \bullet f)(x) = f(Rx) = f(V_f \,\mathrm{diag}(\sigma) V_g^\top x) = \hat{f}(\mathrm{diag}(\sigma) V_g^\top x) = \hat{g}(V_g^\top x) = g(V_g V_g^\top x) = g(x),$$

因此 $R \bullet f = g$，从而完成了这一断言的证明。

我们已经证明了关键定理。 □

迄今发展的理论，尤其是定理 17，直接导出了以下算法，用于解决本文的主要问题（问题 2）。

算法 25 （正交等价性证书）。

输入： $f, g \in \mathbb{R}[x_1, \ldots, x_n]$，次数为 d，使得 $n, d \geq 2$ 且 $f \sim_{O(n)} g$

输出： $R \in O(n)$，使得 $R \bullet f = g$

1. $(\lambda_f, V_f) \leftarrow$加权多项式主成分分析$(f)$

$(\lambda_{\hat{g}}, V_g) \leftarrow$加权多项式主成分分析$(g)$ *2.*

$f \leftarrow V_f \bullet f$ $\hat{g} \leftarrow V_g \bullet g$ *3.* 求

$\sigma \in \{-1, 1\}^n$，使得 $\sigma \bullet f = \hat{g}$ *4.*

$R \leftarrow V_f$ 对角 $\sigma V_g^\top$

我们通过贯穿示例说明上述算法。我们使用双精度（~ 16 位有效十进制数字）浮点运算执行了该算法，仅显示 3 位十进制数字。

例26 (贯穿示例续**)**。再次考虑下面给出的两个次数为 $d = 3$ 的多项式 f 和 g，它们属于 $\mathbb{R}[x_1, x_2, x_3]$：

$$f = -27x_1^3 + 27x_2^2 x_3 - 9x_3,$$
$$g = -12x_1^3 + 12x_1^2 x_2 - 12x_1^2 x_3 + 6x_1 x_2^2 + 36x_1 x_2 x_3 - 33x_1 x_3^2 + 9x_2^3 - 6x_2^2 x_3 + 6x_2 x_3^2 - 6x_3^3 + 3x_1 - 6x_2 - 6x_3.$$

现在，我们使用上述输入 f 和 g 运行算法 *25*。通过此过程，我们得到元素 $R \in O(n)$，以此证明 f 和 g正交等价。

1. $\lambda_f, V_f \leftarrow [808.346,\ 212.351,\ 170.966], \begin{bmatrix} 0.999 & 0.000 & 0.047 \\ 0.000 & -1.000 & 0.000 \\ -0.047 & 0.000 & 0.999 \end{bmatrix}$

$\lambda_g, V_g \leftarrow [808.345,\ 212.351,\ 170.966], \begin{bmatrix} 0.682 & 0.667 & -0.302 \\ -0.364 & 0.667 & 0.650 \\ 0.635 & -0.333 & 0.697 \end{bmatrix}$

2. $\hat{}$ $f \leftarrow -26.910x_1^3 - 3.806x_1^2 x_3 - 1.271x_1 x_2^2 - 0.179x_1 x_3^2 + 26.970x_2^2 x_3 - 0.003x_3^3 + 0.424x_1 - 8.990x_3$ +项的系数小于 10^{-3}

$\hat{g} \leftarrow -26.910x_1^3 - 3.806x_1^2 x_3 - 1.271x_1 x_2^2 - 0.179x_1 x_3^2 + 26.970x_2^2 x_3 - 0.003x_3^3 + 0.424x_1 - 8.990x_3$ +项的系数小于 10^{-3}

3. $\sigma \leftarrow [1, -1, 1]$

$$4.\ R \leftarrow \begin{bmatrix} 0.667 & -0.333 & 0.667 \\ 0.667 & 0.667 & -0.333 \\ -0.333 & 0.667 & 0.667 \end{bmatrix}$$

我们可以轻松验证， R 是 $O(n)$ 的一个元素，并且 $f(Rx) = g(x)$，在浮点运算误差范围内成立，因此 R 确实是所需的证书。

我们以对上述方法的一系列评论作为本节的结尾 .

备注 27. 以下是关于算法 *25* 的几点观察。

a) 请注意，在使用算法 *25* 寻找证书时，并未用到来自PW-PCA 的 λ_f 和 λ_g 。我们仍将它们列出，因为它们或许有助于解决其他问题。注意，由于f 和 g 等价，定理 *17* 的**C1**保证了 $\lambda_f = \lambda_g$，这一点可从上述示例第 *1* 步中它们的数值看出。

b) 虽然没有必要，但由于我们使用浮点运算执行了该算法，还是对上述示例的输出进行合理性检查。简单计算可得

$$||R \bullet f - g|| = 2.035 \times 10^{-13}.$$

回想一下，我们使用的是具有 *16* 位有效十进制位数的浮点数，可见误差小得可以忽略。事实上，如果使用精度更高的浮点运算（$\sim$ *30*位有效十进制位数），误差会更小： 3.090×10^{-27}。

c) 如果合理性检查未通过（也就是说， $||R \bullet f - g||$ 并非可忽略），那么根据关键定理（定理 *17*），原因是以下情况之一（或多种）： f 和 g 中至少有一个不满足一般性假设（假设 *14*）；或者用户误输入了 f 和 g，而它们并不正交等价。由于 f 和 g 不满足一般性条件的可能性极低（可通过检查 λ_f 和 λ_g来确认），因此可以得出结论：用户极有可能误输入了 f 和 g，而它们并不正交等价。

d) 因此，如果所有计算都能精确完成，上述算法也可用于在一般性条件下检验任意两个多项式 f 和 g 是否正交等价。实际上，要完成此类精确计算，我们需要在代数数域中进行运算，而这在实践中可能会带来较高的计算开销。

5 所提算法的性能

本节通过计算机实验考察所提算法的性能。在讨论结果之前，我们先介绍所采用的整体实验设置。

实验设置

- 程序：我们使用 Julia编程语言 [3]，基于标准库和以下程序包实现了所提算法：Nemo、Combinatorics 和 GenericLinearAlgebra。由于我们的主要目标是考察算法的性能方面，因此采用了较为直接的初步算法实现。对于基线算法（下文将予以介绍），我们使用了 AlgebraicSolving 程序包，该程序包采用了高度优化的 msolve 库 [2]。

  我们还使用了以下 Julia 程序包来运行实验，并将结果导出为表格和图形：CacheVariables、CairoMakie [7], Dictionaries、LaTeXStrings、LaTeXTabulars、Pluto、Progress- Logging、StableRNGs 和 TerminalLoggers。

- 基准测试：对于每一对 (n, d)，我们按如下方式生成十对 (f, g)：

1. 生成次数为 d 的 $f \in \mathbb{Z}[x_1, \ldots, x_n]$ 齐次多项式，其系数随机取自 -100 到 100 的范围（不含零）。这会生成一个稠密多项式。2. 使用凯莱变换 [5] 按如下方式生成一个正交 $R \in \mathbb{Q}^{n \times n}$：生成一个条目介于 -10 和 10 之间的随机$n \times n$反对称矩阵 $S \in \mathbb{Z}^{n \times n}$，以及一个随机符号向量$\sigma \in \{-1, 1\}^n$。然后计算 $R = (S - I_n)^{-1}(S + I_n)\,\mathrm{diag}(\sigma)$。3. 计算 $g = R \bullet f \in \mathbb{Q}[x_1, \ldots, x_n]$。

我们生成了 f 和 g，其系数为精确有理系数，因为基线算法（如下所述）涉及求解一个高度超定的多项式方程组，即使系数中存在微小的数值误差，也几乎总会使该方程组变得不相容。值得注意的是，所提出的算法（定理 25）不涉及求解超定方程组，因此不受这一限制。使用浮点数安全运行该算法。

- 计时：所有计时均按墙钟时间测量，我们报告十组（f, g）数据对的中位运行时间。
- 计算机：实验在一台配备 M1 Max CPU（64GB 内存）的 2021年 Macbook Pro 上运行。
- 可复现研究：可在网上 https://gitlab.com/ dahong/ortho-equiv-pwpca 获取用于复现实验的代码。

与基线方法比较

据我们所知，目前没有公开可用的软件能够寻找正交等价的证书。因此，我们实现了以下简单直接的算法，作为比较用的基线方法。

算法28（基线）。

输入： $f, g \in \mathbb{R}[x_1, \ldots, x_n]$ 次数为 d，且 $n, d \geq 2$ 和 $f \sim_{O(n)} g$

输出： $R \in O(n)$，使得 $R \bullet f = g$

1. $M \leftarrow \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & & \vdots \\ r_{n1} & & r_{nn} \end{bmatrix}$ *where* r_{ij} *are indeterminates*
2. $h \leftarrow M \bullet f - g \quad \in \mathbb{R}[r_{11}, \ldots, r_{nn}][x_1, \ldots, x_n]$
3. $H \leftarrow h$ 的所有系数以及 $M^\top M - I_n \subset \mathbb{R}[r_{11}, \ldots, r_{nn}]$ 的所有条目
4. $H = 0$ 的一个实数解 $\in \mathbb{R}^{n^2}$
5. $R \leftarrow$ 通过将 M 实例化到 S 上而得到的矩阵 $\in \mathbb{R}^{n \times n}$

下表展示了所提方法和基线方法的中位运行时间。我们对基线方法进行的 $n = 5$, $d = 10$ 次运行均未能完成；该方法在第一次试验运行约 154 分钟（9240 秒）后崩溃。我们在表中以“>9240”表示这一情况。

n	d	# 术语	基线方法（秒）	所提方法（秒）	近似加速比
3	7	36	0.024	0.004	6.0 倍
	8	45	0.043	0.005	8.6 倍
	9	55	0.075	0.008	9.4 x
	10	66	0.139	0.015	9.3 x
4	7	120	0.944	0.039	24.2 x
	8	165	3.048	0.072	42。3 ×
	9	220	8.933	0.141	63。4 x
	10	286	26.166	0.256	102。2 x
5	7	330	133.690	0.277	482。6 x
	8	495	261.556	0.706	370。5 x
	9	715	1334.248	1.746	764.2 x
	10	1001	>9240	4.005	>2307.1 x

我们有以下观察：

- 所提方法比基线方法快得多。事实上，对于 n的三个取值，所提方法在 $d = 10$ 上都比基线方法在 $d = 7$上快，尽管它需要处理项数多得多的多项式。

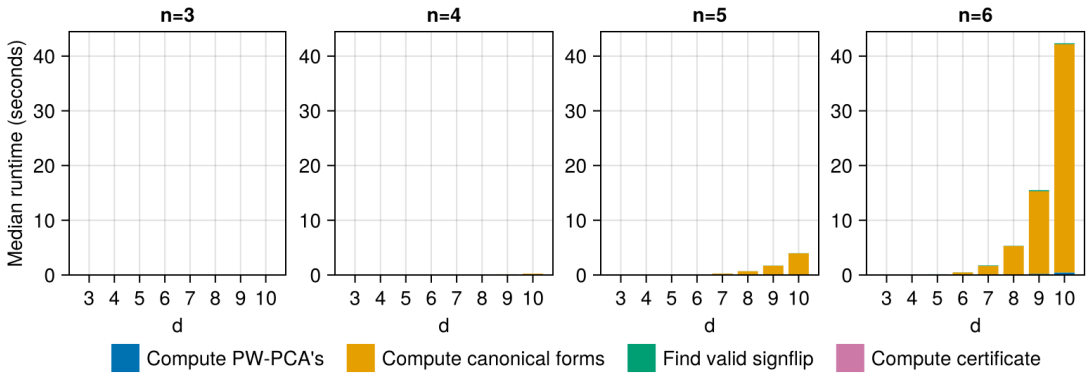
- 与基线方法相比，所提方法在变量数量 n 和次数 d方面的增长都平缓得多。因此，速度提升通常会随着 n 和 d的增加而增大。

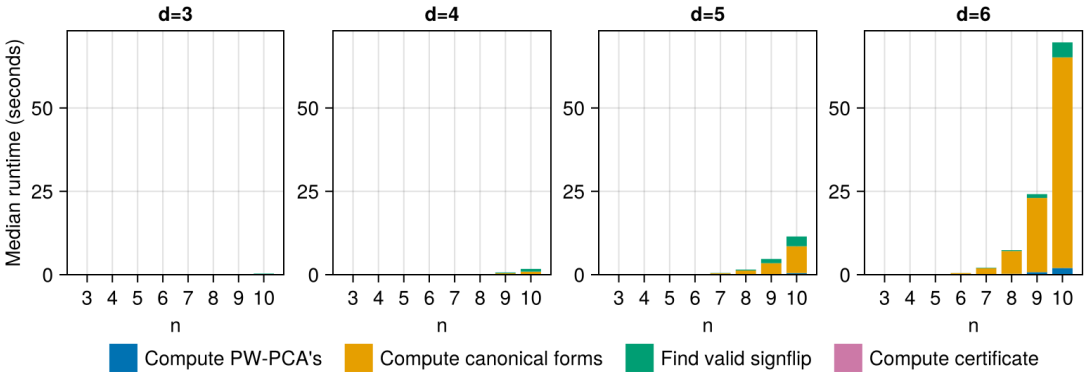
所提算法的性能详情

本文最后将详细考察所提算法的性能。具体而言，我们评估其可扩展性、各步骤所耗费的时间，并指出未来可能的改进方向。为此，回顾一下，所提算法（定理 25）由以下步骤组成：

1. 计算 f 和 g 的 PW-PCA。ˆ 2. 计算 f 和 ˆg 的规范形式。3. 找到有效的符号翻转 σ。4. 计算证书 R。

以下两组图表展示了各步骤耗时的中位数，并采用坐标轴范围一致的坐标轴，以便清楚地观察该方法如何随变量数量n 和次数 d变化。请记住，这些都是稠密多项式。



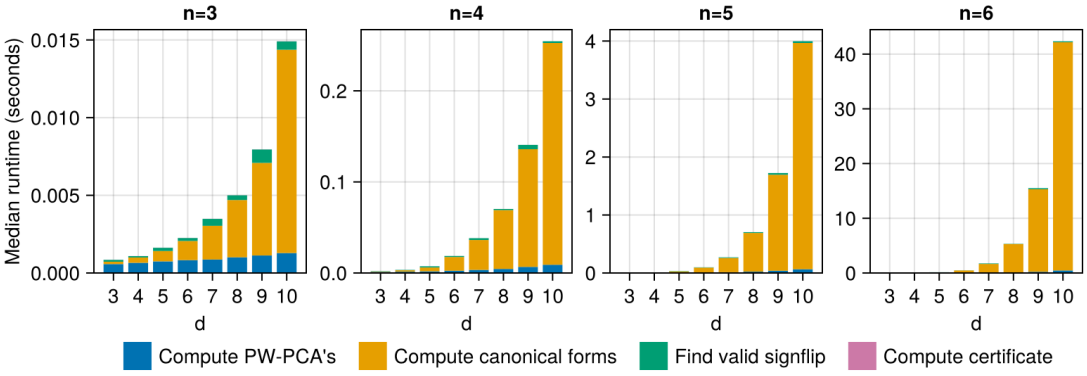


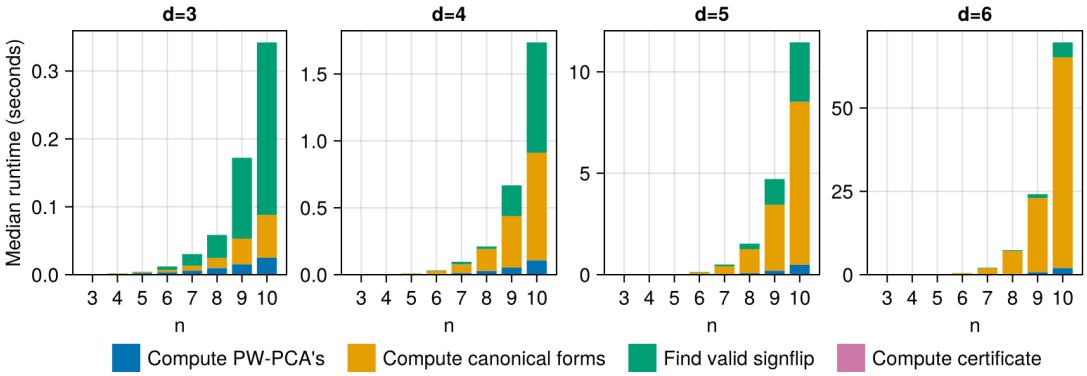
我们作出以下观察：

- 当 n 较小或 d较小时，图表看起来几乎是空的。这是因为这些情况下的耗时相对于较大的情况非常短，因此柱形不可见。
- 顶部一行的图表表明，对于固定的较小 n，所提方法能够很好地扩展到 d 的中等取值。同样，底部一行的图表表明，对于固定的较小 d，所提方法能够很好地扩展到 n 的中等取值。在这两种情况下，运行时间的增长似乎主要是由计算规范形式的成本增加所驱动的。
- 对于较大的 n 和 d，图表实际上表明，运行时间基本上由计算规范形式所主导。

总体而言，进一步提升该方法可扩展性的主要障碍似乎在于提高计算规范形式的速度。

为了更仔细地检查各步骤耗时的详细分布，我们再次绘制相同的计时结果，但这次针对每张图分别调整坐标轴范围。





我们作出如下观察：

- 尽管 PW-PCA 的定义涉及全局优化和多元积分，但其运行时间短得出奇。这可能归功于定理 11 的有效性。

- 从图表的底行可以看出，对于固定的 d，随着 n 增大，寻找有效符号翻转所占的比重越来越大。这可能是因为当前实现通过在指数级庞大的 2^n 可能符号翻转 σ空间上进行朴素循环，来寻找有效的符号翻转。只有找到一个符号翻转 σ，使得 $\sigma \bullet f$ 与$\hat{}g$ 之间产生的误差处于容差范围内时，它才会退出这一$\hat{}$循环。$\hat{}$利用 f 和$\hat{}g$ 中的结构来加速这一步骤，是未来研究的一个有希望的方向。

- 从两行图表可以看出，对于固定的 n，随着 d增大，计算规范形式所占的比重越来越大。这可能是因为当前实现以朴素方式展开 $f(V_f x)$和$g(V_g x)$，而当次数较大时，这一过程可能会变得非常耗时。

  值得注意的是，这些展开中的每一个都对应于一次“对称张量乘矩阵”计算，其中涉及相应的对称系数张量和变换矩阵。张量领域已经对这类运算进行了大量研究，例如 [23, 6]中的研究。或许可以利用类似技术来加速这里的运算。

  另请注意，规范形式仅用于检查：对于给定的 σ，$\sigma \bullet V_f \bullet f = V_g \bullet g$ 是否成立，即 $h(x) = \sigma \bullet V_f \bullet f - V_g \bullet g$ 关于 x 是否恒等于零。因此，另一个有前景的方向是避免完全展开这些多项式的开销：可以使用 Schwartz-Zippel 引理进行随机点测试，也可以进行“通用”点测试，即根据 h 的某些结构属性，选取经过精心挑选的 x 值来计算 $h(x)$。

总体而言，这些图表表明，若要进一步提升可扩展性，很可能需要先提高计算规范形式的速度，再提高找到有效符号翻转的速度。这两个方向都很有希望，可供未来研究探索。

致谢：感谢 Alexander Demin 就使用 Julia 软件包实现所提算法提出的宝贵建议。我们也感谢 Irina Kogan 就计算不变量理论的当前前沿进展进行了有益讨论。

Martin Helmer 获英国皇家学会资助，资助编号为 RSWF\R2\242006 ；同时还获美国空军科学研究办公室（AFOSR）资助，项目编号为 FA9550-22-1-0462，由 Frederick Leve 博士负责管理，并谨此致谢。David Hong 的工作部分获得美国国家科学基金会数学科学博士后研究奖学金（DMS 2103353）资助。Hoon Hong 的工作部分获得美国国家科学基金会 CCF 2212461 和 CCF 2331401 项目资助。

参考文献

[1] G. Ballard 和 T. G. Kolda。《数据科学中的张量分解》。剑桥大学出版社，第 1 版，2025年6月。[2] J. Berthomieu、C. Eder 和 M. Safey El Din。msolve：求解多项式方程组的程序库。载于《2021年国际符号与代数计算研讨会》，第 46 届国际符号与代数计算研讨会，页 51–58，俄罗斯圣彼得堡，2021年7月。ACM。[3] J. Bezanson、A. Edelman、S. Karpinski 和 V. B. Shah。Julia：数值计算的新方法。《SIAM评论》，59(1):65–98，2017年。[4] H. Breloer。正交群作用下偶次多项式的有理不变量。arXiv 预印本 arXiv:2501.17504，2025年。[5] A. Cayley。关于特征为零或素特征的域上正交群及其他经典群的代数结构。《J. Reine Angew.》，32，1846年。[6] H. A. Daas、G. Ballard、L. Grigori、S. Kumar、K. Rouse 和 M. V´erit´e。《并行对称张量与同一向量相乘计算中的通信量最小化》，2025年6月。[7] S. Danisch 和 J. Krumbiegel。Makie.jl：适用于 Julia 的灵活高性能数据可视化工具。《开放源代码软件期刊》，6(65):3349，2021年。[8] H. Derksen 和 G. Kemper。《计算不变量理论》。Springer，2015年。[9] I. Dolgachev。《不变理论讲义》。第 296 号。剑桥大学出版社，2003年。[10] L. Ferraro、F. Galetto、F. Gandini、H. Huang、T. Hawes、M. Mastroeni 和 X. Ni。InvariantRing：群作用的不变量。2.0 版。Macaulay2 程序包，网址：https://github.com/Macaulay2/M2/tree/stable/M2/Macaulay2/packages。[11] L. Ferraro、F. Galetto、F. Gandini、H. Huang、T. Hawes、M. Mastroeni 和 X. Ni。Macaulay2 的 InvariantRing 程序包。《代数与几何软件期刊》，14，2023年。[12] C. Fieker、W. Hart、T. Hofmann 和 F. Johansson。Nemo/hecke：用于 Julia 编程语言的计算机代数和数论程序包。载于《2017年 ACM 国际符号与代数计算研讨会论文集》，ISSAC '17，页 157–164，美国纽约州纽约市，2017年。ACM。[13] G. B. Folland。如何计算多项式在球面上的积分。《美国数学月刊》，108(5):446–448，2001年。[14] P. G¨orlach、E. Hubert 和 T. Papadopoulo。正交群作用下偶次三元形式的有理不变量。《计算数学基础》，19(6):1315–1361，2019年。[15] D. Hong、F. Yang、J. A. Fessler 和 L. Balzano。适用于高维异方差数据的最优加权主成分分析。《SIAM数据科学数学期刊》，5(1):222–250，2023年3月。[16] H. Hotelling。将一组统计变量分解为主成分的分析。《教育心理学期刊》，24(6):417–441，1933年9月。[17] E. Hubert 和 M. Jalard。 $SO_3(\mathbb{R})$ 与 $O_3(\mathbb{R})$ 表示的不变量域的代数独立生成元 $\mathbb{R}^3 \oplus \mathcal{H}$。2025年。[18] I. T. Jolliffe。《主成分分析》。施普林格出版社，纽约州纽约市，第 2 版，2002年。[19] I. A. Kogan。不变量：计算与应用。载于《2023年国际符号与代数计算研讨会论文集》，页 31–40，2023年。

[20] M. D. Neusel. 不变理论，第 36 卷。美国数学学会，200
7.[21] K. Pearson。LIII. 关于空间点系的最佳拟合直线和平面。伦敦、爱丁堡和都柏林哲学杂志与科学杂志，2(11):559–572，1901 年 11 月。[22] V. L. Popov 和 E. B. Vinberg。不变理论。载于代数几何 *IV*：线性代数群与不变理论，第 123–278 页。Springer，1994 年。[23] M. D. Schatz、T. M. Low、R. A. van de Geijn 和 T. G. Kolda。利用张量中的对称性实现高性能：与对称张量的乘法。*SIAM* 科学计算杂志，36(5):C453–C479，2014 年 1 月。[24] P. H. Sch¨onemann。正交 Procrustes 问题的广义解。*Psychometrika*，31(1):1– 10，1966 年 3 月。

[25] T. A. Springer. 不变理论，第 585 卷。Springer，2006 年。

[26] B. Sturmfels. 不变理论中的算法。Springer，2008 年。

附录

定理 15 的证明

我们将证明假设 14 的一般性。令 $n \geq 2$ 和 $d \geq 2$ 为任意但固定的元素。$x_1, \ldots, x_n$ 中次数至多为 d 的所有多项式构成一个 $\mathbb{R}$-向量空间，它与所有可能系数的集合一一对应，即 $\mathbb{R}^m$，其中 m 是 $x_1, \ldots, x_n$ 中次数为 d 的所有单项式的数量。令 a 表示 f 的系数向量，并令

$$E = \{a \in \mathbb{R}^m : C_f \text{ has a multiple eigenvalue}\}.$$

只需证明 E 包含于某个 Zariski 闭集中。为此，只需找到 $h \in \mathbb{R}[a] \setminus \{0\}$，使得 $E \subseteq \{a \in \mathbb{R}^m : h(a) = 0\}$。我们提出

$$h = \text{discriminant}_\lambda(p) \in \mathbb{R}[a] \quad \text{where} \quad p = |\lambda I - C_f| \in \mathbb{R}[a][\lambda].$$

根据线性代数和判别式的基础理论，我们有 $h = \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)^2$，其中 λ_i 是 C_f 的特征值。因此显然有 $E \subseteq \{a \in \mathbb{R}^m : h(a) = 0\}$。现在只需证明 $h \neq 0$ （不恒为零）。为此，只需找到 $a^* \in \mathbb{R}^m$，等价地，找到 f^*，使得 C_{f^*}具有互不相同的特征值。

经过数 次试验错误后，我们提出了与以下特定多项式 f^* 对应的 a^*：

$$f^* = \sum_{1 \leq s \leq n} g_s x_1^{d-2} x_s^2 \quad \text{where } g_s = n + 1 - s$$

注意，$(f^*)^2$ 中的每个指数都是偶数。因此，由定理 11 可立即得出，C_{f^*} 是一个对角矩阵。因此，C_{f^*} 的特征值就是它的对角元素。因此，只需证明 C_{f^*} 的对角元素互不相同。事实上，我们将证明 $C_{f^*,11} > C_{f^*,22} > \cdots > C_{f^*,nn}$。

将定理 11 应用于 f^*，并进行初等但繁琐的计算和化简，可得到 C_{f^*} 的对角项的如下表达式：

$$C_{f^*,ii} = \begin{cases} (2d+1)(2d-1)(2d-3)g_1^2 + 2(2d-1)(2d-3)g_1 u + 2(2d-3)v + (2d-3)u^2 & \text{如果 } i = 1 \\ (2d-1)(2d-3)g_1^2 + 4(2d-3)g_1 g_i + 8g_i^2 + 2(2d-3)g_1 u + 4g_i u + 2v + u^2 & \text{如果 } i \geq 2 \end{cases}$$

其中

$$W=\frac{\pi^{(n+1)/2}(2d-5)!!}{2^d\Gamma(\frac{n+1}{2}+d+1)},\qquad u=g_2+\cdots+g_n,\qquad \text{and}\quad v=g_2^2+\cdots+g_n^2.$$

通过检查上面 $C_{f^*,ii}$ 的表达式，并回想 $g_2>g_3>\cdots>g_n$，可以立即看出 $C_{f^*,22}>C_{f^*,33}>\cdots>C_{f^*,nn}$ 对于 $i\geq 2$ 成立。因此，剩下要证明的是 $C_{f^*,11}>C_{f^*,22}$。为此，注意

$$\begin{aligned}&\left(C_{f^*,11}-C_{f^*,22}\right)/W\\ =&\underbrace{2d\left(2d-1\right)\left(2d-3\right)}_{>0}g_1^2+\underbrace{2\left(2d-2\right)\left(2d-3\right)}_{>0}g_1u+\underbrace{2\left(2d-4\right)}_{\geq 0}v+\left(2d-4\right)u^2-4\left(2d-3\right)g_1g_2-8g_2^2-4g_2u\end{aligned}$$

$>2d\left(2d-1\right)\left(2d-3\right)g_1g_2+2\left(2d-2\right)\left(2d-3\right)g_2u+2\left(2d-4\right)g_2^2+\left(2d-4\right)u^2-4\left(2d-3\right)g_1g_2-8g_2^2-4g_2u$因为 $g_1^2>g_1g_2$、 $g_1u>g_2u$，且 $v\geq g_2^2=\left(2d\left(2d-1\right)\left(2d-3\right)-4\left(2d-3\right)\right)g_1g_2+\underbrace{\left(2\left(2d-2\right)\left(2d-3\right)-4\right)g_2u+\left(2d-4\right)u^2}_{>0}+\underbrace{\left(2\left(2d-4\right)-8\right)g_2^2}_{\geq 0}$ $>\left(2d\left(2d-1\right)\left(2d-3\right)-4\left(2d-3\right)\right)g_2^2+\left(2\left(2d-2\right)\left(2d-3\right)-4\right)g_2^2+\left(2d-4\right)g_2^2+\left(2\left(2d-4\right)-8\right)g_2^2$因为 $g_1g_2>g_2^2$、 $g_2u\geq g_2^2$，且 $u^2\geq g_2^2=8(d+1)d(d-2)g_2^2$通过因式分解

≥ 0

因此 $C_{f^*,\ 11}>C_{f^*,\ 22}$。最后，我们证明了假设 14 的一般性。 □