arXiv:2301.10042v1 [math.AG] 24 Jan 2023

对数稀疏对称矩阵

Dmitrii Pavlov

摘要

如果一个正定矩阵的矩阵对数含有许多零元素，就称它为对数稀疏矩阵。这类矩阵在高维统计学和半定优化中发挥着重要作用。本文从计算代数几何的角度研究对数稀疏矩阵：我们给出了具有给定对数稀疏模式的矩阵集合的 Zariski闭包维数公式，给出了该簇的次数上界，并开发了能够求出其定义方程的隐式化算法。我们通过大量实例展示了这一方法。

1 引言

对数稀疏对称矩阵是矩阵对数具有稀疏性的正定矩阵。这类矩阵出现在高维统计 [2], 中；在此类问题中，协方差矩阵的结构假设对于给出相合估计量是必要的，而稀疏性假设也很自然。此外，一旦稀疏模式固定，相应的对数稀疏矩阵集合便构成一个吉布斯流形 [9]。正如我们在第2节将要回顾的，这是将矩阵指数作用于一个对称矩阵线性系统（LSSM）后得到的流形；此处该系统由稀疏模式定义。吉布斯流形在凸优化[9, 一节 5]中发挥着重要作用。

从实际计算的角度来看，要准确判断给定矩阵是否满足给定的对数稀疏模式，可能颇具挑战。检查给定的多项式方程是否在该矩阵上成立，往往要容易得多。这促使我们研究对数稀疏矩阵族的 Zariski闭包，即在这些矩阵族上消失的多项式的公共零点集。这类Zariski闭包是吉布斯簇的实例。

本文研究由对数稀疏对称矩阵集合的 Zariski 闭包所形成的吉布斯簇。我们说明如何用图对这些簇进行编码，给出其维数公式，并表明实际中可通过简单的线性代数计算该维数。我们提出用于求取其定义方程的数值算法和符号算法。我们还探究图着色如何影响相应的吉布斯簇。此外，我们证明了关于吉布斯簇的一些一般性结果。特别地，对于相应 LSSM 的特征值在有理数域上 $\mathbb{Q}$-线性无关的情形，我们给出吉布斯簇次数的一个上界，并证明置换不变 LSSM 的吉布斯簇具有某种对称性。

本文结构如下。第2节中，我们定义吉布斯流形和吉布斯簇——本研究所需的几何对象；给出吉布斯簇的维数公式和次数上界；研究其定义方程的对称性；并提出一种数值隐式化算法。第3节中，我们正式定义对数稀疏性，说明如何用图对其进行编码，并讨论由对数稀疏性定义的吉布斯簇的特殊性质。第4节中，我们研究由树产生的对数稀疏矩阵族。第5节中，我们研究着色对数稀疏条件。第6节介绍一种用于由对数稀疏性定义的吉布斯簇的符号隐式化算法。最后，第7节讨论对数稀疏性在统计学和优化中的实际意义。

2 吉布斯流形与吉布斯簇

令 $\mathbb{S}^n$ 表示 $n \times n$ 对称矩阵构成的空间。它是一个维数为 $\binom{n+1}{2}$ 的实向量空间。半正定 $n \times n$ 矩阵构成的锥记为 $\mathbb{S}^n_+$。

矩阵指数函数由通常的幂级数定义，该级数对所有实矩阵和复 $n \times n$ 矩阵都收敛。它将对称矩阵映射为正定对称矩阵。零矩阵 0_n 被映射为单位矩阵 id_n。记

$$\exp \,:\, \mathbb{S}^n \to \mathrm{int}(\mathbb{S}^n_+)\,,\; X \,\mapsto\, \sum_{i=0}^{\infty} \frac{1}{i!}\, X^i.$$

此映射可逆，其逆映射为矩阵对数函数，由下列级数给出

$$\log \,:\, \mathrm{int}(\mathbb{S}^n_+) \to \mathbb{S}^n\,,\; Y \,\mapsto\, \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j}\,(Y - \mathrm{id}_n)^j.$$

接下来，我们引入本文中将发挥关键作用的几何对象。固定 d 个线性无关矩阵 A_1、A_2、……、A_d，它们属于 $\mathbb{S}^n$。用 $\mathcal{L}$ 表示 $\mathrm{span}_{\mathbb{R}}(A_1$、……、$A_d)$，这是向量空间 $\mathbb{S}^n \simeq \mathbb{R}^{\binom{n+1}{2}}$ 的一个线性子空间。因此， $\mathcal{L}$ 是一个对称矩阵线性空间（LSSM）。我们关注 $\mathcal{L}$ 在指数映射下的像：

定义 2.1。 $\mathcal{L}$ 的吉布斯流形 $\mathrm{GM}(\mathcal{L})$ 是 d 维流形 $\exp(\mathcal{L}) \subset \mathbb{S}^n_+$。

这确实是凸锥 $\mathbb{S}^n_+$内部的一个 d 维流形。它与 $\mathcal{L} \simeq \mathbb{R}^d$ 微分同胚，该微分同胚由指数映射和对数映射给出。

在某些特殊情况下，吉布斯流形是半代数的，即它是一个代数簇与半正定锥的交集。然而，一般情况下并非如此。我们仍然可以探究，$\mathrm{GM}(\mathcal{L})$ 中任意矩阵的元素之间满足哪些多项式关系。这引出了以下定义。

定义 2.2。 $\mathcal{L}$ 的吉布斯簇 $\mathrm{GV}(\mathcal{L})$ 是 $\mathrm{GM}(\mathcal{L})$ 在 $\mathbb{C}^{\binom{n+1}{2}}$ 中的 Zariski 闭包。

任意 LSSM 都可以写成以下形式 $\mathcal{L} = \{y_1A_1 + \ldots + y_dA_d | y_i \in \mathbb{R}\}$，因此可以视为一个元素属于 $\mathbb{R}(y_1, \ldots, y_d)$ 的矩阵。该矩阵的特征值属于代数闭包 $\mathbb{R}(y_1, \ldots, y_d)$，并称为相应 LSSM 的特征值。已知在以下假设下，GV($\mathcal{L}$) 是不可约且单有理的： $\mathcal{L}$ 的特征值 $\mathbb{Q}$-线性无关，且 $\mathcal{L}$ 定义在 $\mathbb{Q}$ [9,定理 3。6]。

在本节中，我们将扩展 [9] 的结果，使其适用于任意 LSSM $\mathcal{L}$。我们首先研究 GV($\mathcal{L}$) 的定义方程的对称性。

考虑变量组 $\mathbf{x} = \{x_{ij} | 1 \leqslant i \leqslant j \leqslant n\}$。对称群 S_n 的一个元素 σ 作用在多项式环 $\mathbb{R}[\mathbf{x}]$ 上，将 x_{ij} 变为 $x_{\sigma(i)\sigma(j)}$，其中 $1 \leqslant i \leqslant j \leqslant n$ （我们将变量 x_{ij} 和 x_{ji} 视为相同）。我们还将考虑 S_n 通过同时置换矩阵的行和列对 $\mathbb{S}^n$ 的作用。

命题 2.3. 设 $\mathcal{L}$ 是一个由 $n \times n$ 个矩阵组成、且在 $\sigma \in S_n$ 的作用下不变的 LSSM。那么，相应吉布斯簇的理想 I(GV($\mathcal{L}$)) 在 σ 的作用下也不变。

证明。为了证明该命题，只需说明：若 $B \in \mathcal{L}$ 由 $A \in \mathcal{L}$ 通过同时置换行和列得到，那么 exp (B) 也能以同样的方式由 exp (A) 得到。由于 exp (B) 是关于 B 的形式幂级数，只需证明对于任意非负整数 k， B^k 能由 A^k 通过同时置换行和列得到。后一个事实可直接由矩阵乘法公式推出。 □

例 2.4。考虑 LSSM

$$\mathcal{L} = \left\{ \begin{pmatrix} y_1+y_2+y_3 & y_1 & y_2 \\ y_1 & y_1+y_2+y_3 & y_3 \\ y_2 & y_3 & y_1+y_2+y_3 \end{pmatrix} \middle| y_1, y_2, y_3 \in \mathbb{R} \right\}.$$

换位 $\sigma = (12) \in S_3$ 作用于 $\mathbb{S}^3$ 的方式如下：

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix} \mapsto \begin{pmatrix} x_{22} & x_{12} & x_{23} \\ x_{12} & x_{11} & x_{13} \\ x_{23} & x_{13} & x_{33} \end{pmatrix}.$$

此作用限制为 $\mathcal{L}$上的线性自同构。 $\mathcal{L}$ 的吉布斯簇是 $\mathbb{C}^6$ 中的一个超曲面，其素理想由一个多项式生成

$$p(x_{11}, x_{12}, x_{13}, x_{22}, x_{23}, x_{33}) = (x_{11} - x_{22})(x_{11} - x_{33})(x_{22} - x_{33}) - \\ - x_{33}(x_{13}^2 - x_{23}^2) - x_{22}(x_{23}^2 - x_{12}^2) - x_{11}(x_{12}^2 - x_{13}^2).$$

σ 在 $\mathbb{C}[x_{11}, x_{12}, x_{13}, x_{22}, x_{23}, x_{33}]$ 上的作用将 p 映射到 $-p$，因此不会改变理想。$\diamond$

定义 2.5. 设 A 是一个 $n \times n$ 矩阵， $\mathcal{L}$ 是由 $n \times n$ 矩阵构成的 LSSM。 A 的中心化子 $C(A)$ 是所有与 A 可交换的矩阵的集合。 A 的 $\mathcal{L}$-中心化子 $C_{\mathcal{L}}(A)$ 是 $C(A) \cap \mathcal{L}$。

以下是 [9, 定理 2。4]。

定理 2.6. 设 $\mathcal{L}$ 是维数为 d 的 $n\times n$ 矩阵的 LSSM。设 k 为 $\mathcal{L}$ 中一般元素的 $\mathcal{L}$-中心化子的维数， m 为由 $\mathcal{L}$ 的特征值张成的 $\mathbb{Q}$-线性空间的维数。则 $\dim \mathrm{GV}(\mathcal{L}) = m+d-k$。

证明. 由 [9, 定理 4.6] 的证明可知，参数化吉布斯簇的映射 ϕ 的一般纤维维数等于该纤维中一般元素的中心化子的维数，即 k。 ϕ 的定义域不可约，且维数为 $m+d$。因此，根据纤维维数定理 [4, 习题 II.3.22],，$\dim \mathrm{GV}(\mathcal{L}) = m+d-k$。 □

注意，当 $m=k$ 时，我们有 $\dim \mathrm{GV}(\mathcal{L}) = d$，因此吉布斯流形是吉布斯簇的正部分，即 $\mathrm{GM}(\mathcal{L}) = \mathrm{GV}(\mathcal{L})\cap\mathbb{S}_n^+$。特别地，在这种情况下，吉布斯流形是一个半代数集。例如，所有对角矩阵的 LSSM 就属于这种情况 [9, 定理 2.7]。

现在给出 LSSM $\mathcal{L}$ 的吉布斯簇的次数上界。下文中， $\mathbb{V}(I)$ 表示由理想 $I\subseteq\mathbb{C}[\mathbf{x}]$ 在 $\mathbb{C}^{\binom{n+1}{2}}$ 中定义的簇。

命题 2.7. 设 $\mathcal{L}$ 是一个由 $n\times n$ 个矩阵组成的 LSSM，其特征值 $\mathbb{Q}$ 线性无关。则 $\deg \mathrm{GV}(\mathcal{L}) \leqslant n^{\binom{n+1}{2}+2n}$。

证明. 根据 [9, 算法 1],，$\mathrm{GV}(\mathcal{L})$ 的素理想 J 可通过从理想 I 消去得到，该理想由次数至多为 n 的多项式生成。因此，$\deg \mathrm{GV}(\mathcal{L}) = \deg \mathbb{V}(J) \leqslant \deg \mathbb{V}(I)$。簇 $\mathbb{V}(I)$ 位于维数为 $\binom{n+1}{2}+2n+d$ 的仿射空间中，其中 $d = \dim \mathcal{L}$。注意，$\dim \mathbb{V}(I) \geqslant \dim \mathcal{L}$，因此 $\operatorname{codim}\mathbb{V}(I) \leqslant \binom{n+1}{2}+2n$。于是，根据贝祖定理，我们有 $\deg \mathbb{V}(I) \leqslant n^{\binom{n+1}{2}+2n}$，由此证明了该命题。 □

正如我们将在下文看到的，命题 2.7 给出的上界通常过于保守。

已知吉布斯簇的次数后，便可使用数值方法求出其定义方程。一般而言，这使得我们能够计算出无法通过符号算法求得的吉布斯簇理想。

现在我们介绍用于数值求解吉布斯簇方程的算法 1 。我们将 $\langle P\rangle$ 记作由 $P\subseteq\mathbb{C}[\mathbf{x}]$生成的理想。

遗憾的是，命题 2.7 中的次数上界限制了该算法在实际中适用于 $n\leqslant 3$。不过，如果吉布斯簇是一个超曲面，那么找到一个代数方程后，算法便可立即终止。该方程的次数通常远低于命题 2.7 中的次数上界（例如，例 2.4 中的吉布斯簇由一个三次方程定义，而命题2.7 给出的上界等于 3^{12}），因此对于更大的 n，也可以用该算法找到定义方程。

尽管算法 1 使用浮点计算，但对于定义在 $\mathbb{Q}$ 上的 LSSM，可以对其进行调整以得到精确方程。可以使用计算机代数系统中的内置命令来完成这一点，例如 `rationalize` 中的 Julia。将吉布斯簇的参数化代入所得方程，即可检查有理化过程的正确性。

算法 1 已知次数的吉布斯簇数值隐式化 e

输入：LSSM $\mathcal{L}$，给定为 d 个线性无关矩阵 A_1、.....、A_d 的 $\mathbb{R}$-张成空间；GV($\mathcal{L}$) 的次数为 k；
输出：一组方程，作为集合定义 GV($\mathcal{L}$)。

(S1) 要求 $\mathcal{L}$ 具有 $\mathbb{Q}$-线性无关的特征值。
(S2) 设 $I := \{0\}$, $l = 1$, $N = \binom{n+1}{2}$
(S3) 对 $l = 1$ 到 k 执行
(a) 在 $\mathcal{L}$ 中选取 $M > \binom{N+l-1}{l}$ 个随机样本 (b) 令 E 为所选 M 个样本的矩阵指数集合 (c) 通过计算 E 中各元素上的所有 l 次单项式，构造范德蒙德矩阵 A
(d) 令 I_l 为 ker A 的基 (e) $I := \langle I \cup I_l \rangle$

(S4) 返回 I 的一组生成元。

3 对数稀疏模式

每个集合 $S \subseteq \{(i,j) | 1 \leqslant i \leqslant j \leqslant n\}$ 都以下述方式定义对称矩阵上的一个稀疏模式。

定义 3.1. 我们称 $A = \{a_{ij}\} \in \mathbb{S}^n$ 满足由 S 给出的稀疏条件，若对于所有 $(i,j) \in S$，都有 $a_{ij} = 0$。所有满足由 S 给出的稀疏条件的对称矩阵构成一个 LSSM，其基为 $\left\{ \frac{E_{ij}+E_{ji}}{2} \middle| (i,j) \notin S \right\}$，其中 E_{ij} 是一个矩阵单位，即只有位置 (i,j) 上的一个非零元素为 1 的矩阵。

我们将这个 LSSM 记作 $\mathcal{L}_S$，并写作 $\mathcal{L}_S = \left\{ \sum_{(i,j) \notin S} y_{ij} \left(\frac{E_{ij}+E_{ji}}{2} \right) \middle| y_{ij} \in \mathbb{R} \right\}$.

例 3.2。令 $n = 4$ 和 $S = \{(1,2), (1,3), (2,4)\}$。对应的 LSSM

$$\mathcal{L}_S = \begin{pmatrix} y_{11} & 0 & 0 & y_{14} \\ 0 & y_{22} & y_{23} & 0 \\ 0 & y_{23} & y_{33} & y_{34} \\ y_{14} & 0 & y_{34} & y_{44} \end{pmatrix}$$

由方程 $y_{12} = y_{13} = y_{24} = 0$ 定义。 ◇

定义 3.3. 我们称 $A \in \text{int}(\mathbb{S}^n_+)$ 满足由 S 给出的对数稀疏条件，当且仅当 $\log A \in \mathcal{L}_S$。

稀疏模式可以用图来编码，从而可以从组合学的角度对其进行研究。具体而言，对于任意一个以 n 个节点为顶点的简单无向图 G，我们按如下方式定义集合 $S_G \subseteq \{(i,j)|1 \leqslant i \leqslant j \leqslant n\}$：当且仅当 G 中的节点 i 和 j 之间没有边时，$(i,j) \in S_G$。在这种情况下，我们也将相应的 LSSM 记为 $\mathcal{L}_G$。注意，如果 G 有 n 个节点和 e 条边，那么 $\dim \mathcal{L}_G = n + e$。

我们关注满足由 G 给定的对数稀疏模式的矩阵集合的代数描述。这个矩阵集合恰好是 $\mathcal{L}_G$ 的吉布斯流形。由于非连通图对应具有分块对角结构的 LSSM，且分块对角矩阵按块进行指数运算，因此我们只考虑连通 G 的情形。

由图给出的 LSSM 有一个优点：求其吉布斯簇的维数可以归结为一个简单的线性代数过程，即计算矩阵中心化子。以下结果说明了这一点。

命题 3.4. 设 $\mathcal{L}_G$ 是由简单连通图 G 给出的 LSSM，且该图有 n 个节点。那么，它的特征值在 $\mathbb{Q}$ 上线性无关。

证明. 将变量 y_{ij} 中的 $i \neq j$ 取零，并将变量 y_{ii} 取为在 n $\mathbb{Q}$ 上线性无关的代数数，可得到 $\mathcal{L}$ 的一个对角元素，其特征值在 $\mathbb{Q}$ 上线性无关。这立即说明 $\mathcal{L}$ 的特征值在 $\mathbb{Q}$ 上线性无关。

□

现在我们来讨论如何计算一般元素 $A \in \mathcal{L}_G$ 的 $\mathcal{L}_G$-中心化子。一种方法是直接求解方程组 $\binom{n}{2}$ equations $XA - AX = 0$，其中变量 x_{ij} 取自域 $\mathbb{Q}(a_{ij})$，x_{ij} 是 $X \in \mathcal{L}_G$ 的元素，而 a_{ij} 是 A 的元素。不过，还有一种方法可以更明确地描述 $\mathcal{L}_G$-中心化子。

注意，根据命题 3.4，$\mathcal{L}_G$ 的特征值在 $\mathbb{Q}$ 上线性无关。特别地，这意味着 $A \in \mathcal{L}$ 的特征值一般而言互异，且 A 一般而言是非亏格的（[6，定义 1.4.4]）。因此，根据 [7，定理 4.4.17、推论 4.4.18]，我们有 $C(A) = \mathrm{span}_{\mathbb{R}}(\mathrm{id}_n, A, \ldots, A^{n-1})$，其中 id_n 是单位矩阵 $n \times n$。因此，求 $C_{\mathcal{L}}(A)$ 可归结为求 $\mathrm{span}_{\mathbb{R}}(\mathrm{id}_n, A, \ldots, A^{n-1})$ 与 $\mathcal{L}$ 的交集。这个交集可通过求解线性方程组 $p_0\mathrm{id}_n + p_1 A + \ldots + p_{n-1}A^{n-1} = \sum c_{ij}E_{ij}{}_{(i,j)\notin S_G}$ 来找到。

关于变量 $p_0, \ldots, p_{n-1}, c_{ij}$。由于 id_n 和 A 都属于 $\mathcal{L}$，交集至少是二维的，因此得到以下命题。

命题 3.5. 设 G 是一个有 n 个节点和 e 条边的简单连通图。则 $\mathrm{dimGV}(\mathcal{L}_G) \leqslant 2n + e - 2$。

我们猜想，$\dim \mathrm{GV}(\mathcal{L}_G) = \min\left(2n + e - 2, \binom{n+1}{2}\right)$。当 $2n+e-2 \leqslant \binom{n+1}{2}$时，该猜想等价于断言 $\{A^2, \ldots, A^{n-1}\} \cup \{E_{ij} | (i,j) \in E(G)\} \cup \{E_{ii} | i = 1, \ldots, n\}$ 是一个线性无关集。这里，$E(G)$表示 G的边集。当 G 是一棵树时，该猜想成立，正如下节所示。

本节最后，我们刻画与 $n \leqslant 4$ 个顶点上的简单连通图相对应的 LSSM 的吉布斯簇。对于 $n \leqslant 3$，我们总有 $\mathrm{dimGV}(\mathcal{L}_G) = \binom{n+1}{2}$

因此，GV($\mathcal{L}_G$) 是整个环境空间 $\mathbb{C}^{\binom{n+1}{2}}$。对于 $n = 4$，有 6 个两两不同构的简单连通图，其中 2 个是树。如果 G 不是树，我们再次有 dim GV($\mathcal{L}_G$) $= 6 = \binom{n+1}{2}$ 且 GV($\mathcal{L}_G$) $= \mathbb{C}^{\binom{n+1}{2}}$。如果 G 是树，则 GV($\mathcal{L}_G$) 是一个超曲面。下一节将讨论这 2 个超曲面的定义方程。

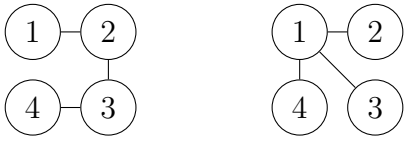
4 树所确定的稀疏模式

树是一类重要的图，对于给定数量的节点，它们所对应的 LSSM 具有可能的最小维数。值得注意的是，对于这类 LSSM，吉布斯簇的维数只取决于图中的节点数（或者等价地，矩阵的大小），且这种依赖关系是线性的。

定理 4.1. 设 $\mathcal{L}_G$ 是由具有 n 个节点的树 G 所确定的 LSSM。那么 dim GV($\mathcal{L}_G$) $= 3n - 3$。

证明. 根据命题 3.4，由 $\mathcal{L}_G$ 的特征值张成的 $\mathbb{Q}$-线性空间的维数等于 n。$\mathcal{L}_G$ 的维数等于 $2n-1$，因为 G 是一棵树，因此有 $n-1$ 条边。接下来需要计算 $\mathcal{L}_G$ 中一般元素的 $\mathcal{L}_G$-中心化子的维数。设 $A \in \mathcal{L}_G$。我们要寻找方程 $AY - YA = 0$,$Y \in \mathcal{L}_G$ 的解。这是关于未知数 y_{ij} 的齐次线性方程组。我们有 $(AY - YA)_{ik} = \sum a_{ij}y_{jk} - \sum y_{ij}a_{jk}$。注意，由于 $Y \in \mathcal{L}_G$，当且仅当 (i, j) 是 G 的一条边或 $i = j$ 时，$y_{ij} \neq 0$。对于 a_{ij}，一般情形下也如此。因此，$(AY - YA)_{ik}$ 不恒为零，当且仅当存在 j，使得 (i, j) 和 (j, k) 是 G 的边；或者 (i, k) 本身就是 G 的一条边。用图 G 的术语来说，这意味着 $(AY - YA)_{ik}$ 不恒为零，当且仅当从 i 到 k 存在一条边长至多为 2 的路径。由于 G 是一棵树，这样的路径至多只有一条。因此，若 i 和 k 之间存在一条经过节点 j、边长为 2 的路径，那么 $AY - YA$ 中对应的元素等于 $a_{ij}y_{jk} - a_{jk}y_{ij}$。当 y_{jk} 与 y_{ij} 成比例，比例系数为 a_{ij}/a_{jk} 时，该元素等于零（注意，a_{jk} 一般不为零）。由于 G 是连通的，我们得出，所有满足 $i \neq j$ 的 y_{ij} 都成比例。若 i 和 k 之间有一条边，则 $AY - YA$ 中对应的元素等于 $y_{ii}a_{ik} + y_{kk}a_{ik} - (a_{ii} - a_{kk})y_{ik}$。若其等于零，则 y_{kk} 是 y_{ii} 和 y_{ik} 的线性组合。我们得出，由于 G 是连通的，且所有 y_{ik} 都成比例，因此所有 y_{ii} 都可以表示为 y_{11} 和一个满足 $j \neq k$ 的 y_{jk} 的线性组合。因此，中心化子（即所考虑线性系统的解空间）的维数至多为 2。由于它包含 id_n 和 A，其维数恰为 2。由此，定理的结论根据 $m = n$、$d = 2n - 1$ 和 $k = 2$ 的定理 2.6 即得。 □

例 4.2. 对于 $n = 4$，恰有两棵非同构树，如下所示。根据定理 4.1，它们的吉布斯簇的维数为 9。因此，这些吉布斯簇是 $\mathbb{C}^{\binom{n+1}{2}} = \mathbb{C}^{10}$中的超曲面。



对应的线性状态空间模型为

$$\begin{pmatrix} y_{11} & y_{12} & 0 & 0 \\ y_{12} & y_{22} & y_{23} & 0 \\ 0 & y_{23} & y_{33} & y_{34} \\ 0 & 0 & y_{34} & y_{44} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} y_{11} & y_{12} & y_{13} & y_{14} \\ y_{12} & y_{22} & 0 & 0 \\ y_{13} & 0 & y_{33} & 0 \\ y_{14} & 0 & 0 & y_{44} \end{pmatrix},$$

分别。

对于 4-链，左侧图的吉布斯簇由一个次数为 6、含有 96 项的齐次方程定义。对于右侧的图，定义方程同样是次数为 6 的齐次方程，含有 60 项。这两个方程是使用算法1找到的。$\diamond$

5 彩色图的对数稀疏性

由彩色图定义的稀疏 LSSM，出现在代数统计学中对彩色高斯图模型的研究中 [5], [10]。本节研究这类 LSSM 的吉布斯簇的性质。

考虑图 G ，并假设其顶点用 p 种颜色标记，边用 q 种颜色标记。相应的 LSSM $\mathcal{L}$ 由以下三组方程定义。

1. $x_{ij} = 0$ 若 (i, j) 不是 G的一条边
2. $x_{ii} = x_{jj}$ 如果顶点 i 和 j 具有相同的颜色。
3. $x_{ij} = x_{kl}$ 如果 (i, j) 和 (k, l) 是 G 中颜色相同的边 .

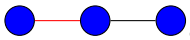
显然，$\dim \mathcal{L} = p + q$。

我们将用 $\mathcal{G}$ 表示彩色图，并用 $\mathcal{L}_{\mathcal{G}}$表示对应的线性状态空间模型。相应的无色图通常记作 G。注意，由于 $\mathcal{L}_{\mathcal{G}} \subseteq \mathcal{L}_G$，吉布斯簇的包含关系也成立：$\mathrm{GV}(\mathcal{L}_{\mathcal{G}}) \subseteq \mathrm{GV}(\mathcal{L}_G)$。由于对于任意 $\mathcal{G}$，单位矩阵都属于 $\mathcal{L}_{\mathcal{G}}$ ，命题 3.5 中的维数界对彩色图也成立。

定义 5.1. 我们称$X \in \mathbb{S}^n_+$ 满足由 $\mathcal{G}$给出的彩色稀疏模式，当且仅当 $X \in \mathcal{L}_{\mathcal{G}}$。

命题 5.2. 设 $\mathcal{G}$ 是一个具有 n 个节点的彩色图，其中顶点标记为 p种颜色，边标记为 q 种颜色。那么 $\dim \mathrm{GV}(\mathcal{L}_{\mathcal{G}}) \leqslant n + p + q - 2$。

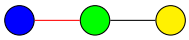
请注意，如果 $\mathcal{G}$ 是一个彩色图，那么 $\mathcal{L}_{\mathcal{G}}$ 的特征值不一定 $\mathbb{Q}$线性无关。因此，命题5.2 给出的上界并不总能取到。

例5.3。考虑图 。对应的 LSSM 为

$$\begin{pmatrix} y_1 & y_2 & 0 \\ y_2 & y_1 & y_3 \\ 0 & y_3 & y_1 \end{pmatrix}.$$

这个 LSSM 的特征值 $\mathbb{Q}$线性相关：它们满足方程 $2\lambda_1 = \lambda_2 + \lambda_3$。我们有 dim GV($\mathcal{L}$) $= 3 < n + p + q - 2 = 3 + 1 + 2 - 2 = 4$。请注意，在这种情况下，dim GV($\mathcal{L}_\mathcal{G}$) $=$ dim GM($\mathcal{L}_\mathcal{G}$)，而 $\mathcal{L}_\mathcal{G}$ 的吉布斯流形，即满足由 $\mathcal{G}$ 给出的着色对数稀疏条件的矩阵集合，是其吉布斯簇的正部分，即 GM($\mathcal{L}_\mathcal{G}$) $=$ GV($\mathcal{L}_\mathcal{G}$) $\cap \mathbb{S}^n_+$。这意味着，可以用代数方式描述具有该图所给出的着色对数稀疏模式的矩阵集合。$\diamond$

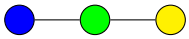
为了说明同一图的不同着色如何影响吉布斯簇，我们在本节最后分析底层图为 3-链的彩色图。我们使用 [9,算法 1]。

1. 

   对应的 LSSM 为

   $$\mathcal{L}_\mathcal{G} = \begin{pmatrix} y_1 & y_4 & 0 \\ y_4 & y_2 & y_5 \\ 0 & y_5 & y_3 \end{pmatrix}.$$

   dim GV($\mathcal{L}_\mathcal{G}$) $= 6$ 且不存在在吉布斯簇上成立的多项式方程 .

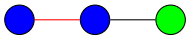
2. 

   对应的 LSSM 为

   $$\mathcal{L}_\mathcal{G} = \begin{pmatrix} y_1 & y_4 & 0 \\ y_4 & y_2 & y_4 \\ 0 & y_4 & y_3 \end{pmatrix}.$$

   dimGV($\mathcal{L}_\mathcal{G}$) $= 5$ 且吉布斯簇是一个三次超曲面，其素理想由多项式生成

   $$x_{11}x_{13}x_{23} - x_{12}^2x_{23} + x_{12}x_{22}x_{13} - x_{12}x_{13}^2 - \\ - x_{12}x_{13}x_{33} + x_{12}x_{23}^2 - x_{22}x_{13}x_{23} + x_{13}^2x_{23}.$$

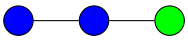
3. 

   对应的 LSSM 为

   $$\mathcal{L}_\mathcal{G} = \begin{pmatrix} y_1 & y_3 & 0 \\ y_3 & y_1 & y_4 \\ 0 & y_4 & y_2 \end{pmatrix}.$$

   dimGV($\mathcal{L}_\mathcal{G}$) $= 5$ 且吉布斯簇是一个三次超曲面。其素理想由多项式生成

   $$-x_{11}x_{12}x_{23} + x_{11}x_{22}x_{13} - x_{11}x_{13}x_{33} + x_{12}x_{22}x_{23} - \\ - x_{22}^2x_{13} + x_{22}x_{13}x_{33} + x_{13}^3 - x_{13}x_{23}^2.$$

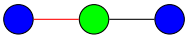
4. 

相应的 LSSM 为

$$\mathcal{L}_{\mathcal{G}} = \begin{pmatrix} y_1 & y_3 & 0 \\ y_3 & y_1 & y_3 \\ 0 & y_3 & y_2 \end{pmatrix}.$$

dimGV($\mathcal{L}_{\mathcal{G}}$) $= 4$。吉布斯簇是一个完全交，其素理想由以下多项式生成

$$x_{11} - x_{22} + x_{33},$$
$$-x_{12}x_{23} + x_{22}x_{13} - x_{13}^2 - x_{13}x_{33} + x_{23}^2.$$

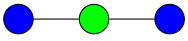
5. 

相应的 LSSM 为

$$\mathcal{L}_{\mathcal{G}} = \begin{pmatrix} y_1 & y_3 & 0 \\ y_3 & y_2 & y_4 \\ 0 & y_4 & y_1 \end{pmatrix}.$$

dimGV($\mathcal{L}_{\mathcal{G}}$) $= 5$，且吉布斯簇是一个三次超曲面。其素理想由以下多项式生成

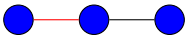
$$-x_{11}x_{12}x_{23} + x_{12}^2x_{13} + x_{12}x_{23}x_{33} - x_{13}x_{23}^2.$$

6. 

相应的 LSSM 为

$$\mathcal{L}_{\mathcal{G}} = \begin{pmatrix} y_1 & y_3 & 0 \\ y_3 & y_2 & y_3 \\ 0 & y_3 & y_1 \end{pmatrix}.$$

dim GV($\mathcal{L}_{\mathcal{G}}$) $= 4$，且吉布斯簇是一个仿射子空间，其素理想由 $x_{12} - x_{23}$ 和 $x_{11} - x_{33}$ 生成。

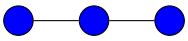
7. 

相应的 LSSM 为

$$\mathcal{L}_{\mathcal{G}} = \begin{pmatrix} y_1 & y_2 & 0 \\ y_2 & y_1 & y_3 \\ 0 & y_3 & y_1 \end{pmatrix}.$$

dim GV($\mathcal{L}_{\mathcal{G}}$) $= 3$。吉布斯簇的素理想由 7 个多项式生成:

$$x_{12}x_{13} - x_{22}x_{23} + x_{23}x_{33},$$
$$x_{11}x_{13} - x_{12}x_{23} + x_{13}x_{33},$$
$$x_{11}x_{22} - x_{11}x_{33} - x_{22}^2 + x_{22}x_{33} + x_{13}^2,$$
$$x_{12}^2 - x_{22}^2 + x_{13}^2 + x_{33}^2,$$
$$x_{11}x_{12} - x_{12}x_{22} + x_{13}x_{23},$$
$$x_{11}^2 - x_{22}^2 + x_{13}^2 + x_{23}^2,$$
$$-x_{12}x_{22}x_{23} + x_{12}x_{23}x_{33} + x_{22}^2x_{13} - x_{13}^3 - x_{13}x_{33}^2.$$

8. 

相应的 LSSM 为

$$\mathcal{L}_{\mathcal{G}} = \begin{pmatrix} y_1 & y_2 & 0 \\ y_2 & y_1 & y_2 \\ 0 & y_2 & y_1 \end{pmatrix}.$$

这是一个交换族，因此根据 [9, 定理 2。7], $\dim \mathrm{GV}(\mathcal{L}_{\mathcal{G}}) = 2$。吉布斯簇的素理想由 3 个线性形式和 1 个二次曲面生成： $x_{22} - x_{13} - x_{33}$， $x_{12} - x_{23}$， $x_{11} - x_{33}$ 以及 $-2x_{13}x_{33} + x_{23}^2$。

6 从解析方程到代数方程

由于对数在 $\mathbb{R}_{>0}$ 上是解析函数，因此满足由图 G 给出的对数稀疏模式的矩阵集合可以通过形式幂级数方程来定义。写出这些方程的一种紧凑方式是使用西尔维斯特公式。

定理 6.1（Sylvester [11]）。设 $f\colon D \to \mathbb{R}$ 是开集 $D \subset \mathbb{R}$上的解析函数， $M \in \mathbb{R}^{n\times n}$ 是一个具有 n 个互异特征值 $\lambda_1, \ldots, \lambda_n$ 且这些特征值属于 D的矩阵。那么

$$f(M) = \sum_{i=1}^{n} f(\lambda_i) M_i, \quad with \quad M_i = \prod_{j\neq i} \frac{1}{\lambda_i - \lambda_j}(M - \lambda_j \cdot \mathrm{id}_n).$$

我们注意到， 右侧的乘积是在交换环 $\mathbb{R}[M]$ 中进行的 .

令 f 为对数函数，可得到 $\log X$ 的参数化，其元素 x_{ij} 由 X的特征值 λ_i 及 X 的特征值及其对数 $\log \lambda_i$诱导出的对数稀疏条件施加于 X ，要求该参数化的某些分量为零，因此得到关于 x_{ij}、 λ_i和 $\log$ λ_i的多项式方程组。利用 λ_i 和 $\log \lambda_i$ 之间由特征多项式的系数给出的多项式关系，对该方程组中的变量进行消去后，可得到 $\mathrm{GV}(\mathcal{L}_G)$ 的一组定义方程。该过程由算法2描述。

定理 6.2。算法*2* 是正确的。步骤 J 中计算得到的理想***(S9)*** 是$\mathrm{GV}(\mathcal{L}_G)$ 的素理想。

证明。由于 $\mathcal{L}_G$ 的特征值 $\mathbb{Q}$在 E_2 上线性无关，由其生成的理想是素理想。此外，对于任意正定的 X ，不存在 $\mathbb{C}$-代数关系将 X 的特征值及其对数联系起来 X (这是 Ax-Schanuel 定理[1, (SP)]的推论)。这两项事实确保了 X、 λ和 $\log \lambda$ 之间的所有代数关系都已纳入考虑，因此该算法是正确的。由 E_1 和 E_2 生成的理想经过饱和后也因此是素理想，而步骤**(S9)** 中的消去保持素性。

□

注意，素性 J 意味着 $\mathrm{GV}(\mathcal{L}_G)$ 不可约，正如 [9, 定理 3。6]。

与 [9, 算法 1] 相比，该算法的优势在于使用了更小的多项式环，而且消去的变量更少。

A1算法 2 由图 G 给出的 $\mathcal{L}_G$ 吉布斯簇的隐式化

输入：一个简单无向连通图 G;
输出 GV($\mathcal{L}_G$) 的一组定义方程。

(S1) 令 $S := \{(i,j)|1 \leqslant i \leqslant j \leqslant n$ 和 $(i,j) \notin E(G)\}$。 (S2) 令 $\{a_{ij}\} = A := \sum_{i=1}^{n} \log(\lambda_i) X_i$, 其中 $X_i = \prod_{j \neq i} \frac{1}{\lambda_i - \lambda_j}(X - \lambda_j \cdot \mathrm{id}_n)$, 且 $X = (x_{ij})$ 是由变量组成的对称矩阵。 (S3) 令 $E_1 := \{a_{ij}|(i,j) \in S\}$。 (S4) 消去 E_1 中的分母，并记下最小公分母 D。 (S5) 计算特征多项式 $P_X(x_{ij}; \lambda) = \det(X - \lambda \mathrm{id}_n) = c_0(x_{ij}) + c_1(x_{ij})\lambda + \ldots + c_n(x_{ij})\lambda^n$。 (S6) 令 $E_2 := \{$多项式 $n\ (-1)^i \sigma_{n-i}(\lambda) - c_i(y)\}$，其中 $\sigma_{n-i}(\lambda)$ 是变量 λ_1、.....、λ_n 中的第 $(n-i)$ 个初等对称多项式。 (S7) 令 I 为 $\mathbb{Q}[x_{ij}, \lambda, \log$ 中由 E_1 和 E_2 生成的理想。 (S8) $I := I : D^{\infty}$。 (S9) 令 $J = I \cap \mathbb{Q}[x_{ij}]$。 (S10) 返回 J 的一组生成元。

7 应用相关性

本节将展示对数稀疏矩阵在两个应用领域中发挥的作用：高维统计和半定规划。

高维统计中的一个典型问题是估计长度为 n 的随机向量的协方差矩阵 $l \ll n$ 个样本。众所周知，在这种情况下，如果不对协方差矩阵的结构作出额外假设，就无法得到相合估计量。在某些情况下，可以假设协方差矩阵具有固定的对数稀疏模式 [2], [3]，从而解决这个问题。该假设的一个优点是，一旦协方差矩阵 C具有对数稀疏模式，集中矩阵 $K = C^{-1}$也会自动具有该模式，因为（$\exp\ L$）$^{-1} = \exp$（$-L$）。原则上，可以放宽对对数稀疏性的结构假设，改为假设协方差矩阵属于吉布斯簇。这样放宽假设的好处是，检查矩阵是否满足给定的一组多项式方程，通常比计算矩阵对数并检查其是否满足稀疏条件更简单。

现在我们说明，对半定规划进行熵正则化时如何产生对数稀疏矩阵 [9]。首先给出基本定义。

我们固定一个任意的线性映射 π： $\mathbb{S}^n \to \mathbb{R}^d$。它可以写成如下形式

$$\pi(X) = \big(\langle A_1, X\rangle, \langle A_2, X\rangle, \ldots, \langle A_d, X\rangle\big).$$

这里， $A_i \in \mathbb{S}^n$，而 $\langle A_i$、 $X\rangle$ ： $=$ trace(A_iX)。半正定锥 $\mathbb{S}^n_+$在该线性映射 π 下的像 π($\mathbb{S}^n_+$)是一个谱半代数影。在我们的设定中，它是 $\mathbb{R}^d$中的一个满维半代数凸锥。

Sem定规划（SDP）是如下凸优化问题 ：

$$\text{Minimize} \quad \langle C, X\rangle \quad \text{subject to} \quad X \in \mathbb{S}^n_+ \ \text{ and } \ \pi(X) = b.$$

例如，参见[8, 第 12]章。一个 SDP 问题实例由代价矩阵 $C \in\mathbb{S}^n$ 和右端向量 $b \in \mathbb{R}^d$指定。可行域 $\mathbb{S}^n_+ \cap \pi^{-1}$($b$)是一个谱半代数体。当且仅当 b 属于 π($\mathbb{S}^n_+$)时，SDP 问题才有可行解。

考虑 LSSM $\mathcal{L} = \text{span}_{\mathbb{R}}(A_1, \ldots, A_d)$。我们通常假设 $\mathcal{L}$ 包含一个正定矩阵。该假设确保每个谱半代数体 π^{-1}(b)都是紧的。

SDP 的熵正则化[9, 第 5] 节定义如下：

$$\text{Minimize} \quad \langle C, X\rangle \ - \ \epsilon \cdot h(X) \quad \text{subject to} \quad X \in \mathbb{S}^n_+ \ \text{ and } \ \pi(X) = b.$$

这里 $\epsilon > 0$ 是一个参数， h 表示冯·诺伊曼熵

$$h : \mathbb{S}^n_+ \to \mathbb{R}, \ X \mapsto \text{trace}\big(X - X \cdot \log(X)\big).$$

下一个结果以定理 5.1 的形式出现在 [9] 中，展示了吉布斯流形在半定规划中的作用。将 ϵ 和代价矩阵 C 纳入 LSSM 后，可得到以下对称矩阵仿射空间（ASSM）：

$$\mathcal{L}_\epsilon \ := \ \mathcal{L} - \frac{1}{\epsilon} C \quad \text{for any } \epsilon > 0.$$

这里允许 $\epsilon = \infty$ 这一情形，此时对 C 的依赖消失，ASSM 就是 LSSM，即 $\mathcal{L}_\infty = \mathcal{L}$。注意，ASSM 的吉布斯流形的定义与 LSSM 的情形类似。

定理7.1. 对于 $b \in \pi(\mathbb{S}_+^n)$， $\pi^{-1}(b)$与吉布斯流形 GM($\mathcal{L}_\epsilon$)的交集仅包含一个点 X_ϵ^*。该点是正则化半定规划的最优解。对于$\epsilon = \infty$，它是谱半代数体 $\pi^{-1}(b)$上冯·诺伊曼熵的唯一最大化点。

满足给定对数稀疏模式的矩阵集合是吉布斯流形，对应于一类特定的半定规划约束。如果稀疏模式由图 G给出，则谱半代数体由某些元素固定的半正定矩阵构成。更确切地说，当且仅当 $i = j$ 或(i, j)是 G的一条边时，元素 x_{ij} 固定。此外，如果图还带有颜色，则当节点 i 和 j 颜色相同时，加入约束 $x_{ii} = x_{jj}$ ；当边(i, j)和(k, l)颜色相同时，加入约束 $x_{ij} = x_{kl}$。

例 7.2. 设 G 为 4-链。相应的 LSSM 为

$$\mathcal{L}_G = \begin{pmatrix} y_{11} & y_{12} & 0 & 0 \\ y_{12} & y_{22} & y_{23} & 0 \\ 0 & y_{23} & y_{33} & y_{34} \\ 0 & 0 & y_{34} & y_{44} \end{pmatrix}.$$

该谱半代数体由以下矩阵组成

$$\begin{pmatrix} b_{11} & b_{12} & x_{13} & x_{14} \\ b_{12} & b_{22} & b_{23} & x_{24} \\ x_{13} & b_{23} & b_{33} & b_{34} \\ x_{14} & x_{24} & b_{34} & b_{44} \end{pmatrix} \in \mathbb{S}_+^n,$$

其中 b_{ij} 处的元素固定，而 x_{ij} 处的元素可以任意取值，只要矩阵为半正定即可。$\diamond$

致谢

作者谨此感谢 Bernd Sturmfels、Simon Telen 和 Piotr Zwernik 提供的宝贵讨论与建议。

参考文献

[1] J. Ax. 关于 Schanuel猜想。数学年刊，93(2):252–268，1971。

[2] H. S. Battey. 一类新型结构化协方差及其逆协方差矩阵的特征结构。Bernoulli, 23:3166–3177，2017。

[3] H. S. Battey。稀疏性诱导。《加拿大统计学杂志》，2022年 .

[4] R. Hartshorne。《代数几何》。数学研究生教材。Springer，纽约，纽约州，1977年。[5] S. Højsgaard 和 S. Lauritzen。具有边和顶点对称性的图形高斯模型。《皇家统计学会杂志》，70:1005–1027，2008年。[6] R. A. Horn 和 C. R. Johnson。《矩阵分析》。剑桥大学出版社，1985年。[7] R. A. Horn 和 C. R. Johnson。《矩阵分析专题》。剑桥大学出版社，1991年。[8] M. Micha lek 和 B. Sturmfels。《非线性代数导论》，数学研究生学习丛书第211卷。美国数学学会，普罗维登斯，2021年。[9] D. Pavlov、B. Sturmfels 和 S. Telen。吉布斯流形，2022年。[10] B. Sturmfels 和 C. Uhler。多元高斯分布、半正定矩阵补全与凸代数几何。《统计数学研究所年刊》，62:603–638，2010年。[11] J. Sylvester。论行星理论中的长期不等式方程。《哲学杂志》第五辑，16:100:267–269，1883年。

作者地址：Dmitrii Pavlov，莱比锡马克斯·普朗克数学研究所
dmitrii.pavlov@mis.mpg.de