关于投影到半正定互补集的近似方法

丁超中国科学院数学与系统科学研究院，中国北京
dingchao@amss.ac.cn2026年11月22日

arXiv:2611.29199v1 [math.OC] 22 Nov 2026

摘要

半正定互补约束的增广拉格朗日方法中的近端松弛更新需要进行联合矩阵投影。给定实对称矩阵 A 和 B，我们寻求距离最近的一对 $(X,\ Z)$，使得 X 半正定、Z 负半定，且满足 $XZ=0$。为近似这一投影，我们将投影到 X 允许的子空间上的正交投影算子，替换为一个特征值位于$[0,1]$的对称矩阵 T。根据由此得到的半定规划的一个最优解，我们通过从零子空间开始，按特征值降序依次加入 T 的特征子空间，构造候选子空间。对于每个候选子空间，我们将 A 压缩到该子空间，并将 B 压缩到其正交补。保留第一个压缩矩阵的正特征值和第二个压缩矩阵的负特征值。对于任意输入，最优一对的平方损失至多为最小值的 1.189 倍；当 A半正定且 B 负半定时，至多为最小值的 1.15487 倍。对于任意输入，若 T 在 $(0,\ 1)$ 中至多有两个不同特征值，则该系数可改进为 1.18。证明对均匀谱阈值取平均，并使用计算机辅助界来处理由此得到的标量不等式。

1 投影问题

令 $\mathbb{S}^n$ 表示实对称 $n\times n$ 矩阵，I 表示单位矩阵。对于大小相同的实矩阵，我们使用弗罗贝尼乌斯内积 $\langle H,K\rangle=\mathrm{tr}(H^\top K)$ 及范数 $\|H\|_F=\sqrt{\langle H,H\rangle}$。对于对称矩阵，$H\preceq K$ 表示 $K-H$ 是半正定的。对于 $H\succeq 0$，$H^{1/2}$ 表示其正半定平方根。我们用 $\mathrm{ran}\ H$ 和 $\ker H$ 分别表示 H 的值域和零空间。给定 $A,B\in\mathbb{S}^n$，我们寻求点 (A,B) 到半正定锥法锥映射的图的平方距离，

$$\{(X,Z)\in\mathbb{S}^n\times\mathbb{S}^n : X\succeq 0,\ Z\preceq 0,\ XZ=0\}.$$

矩阵 Z 在 $X\succeq 0$ 处是半正定锥的法向量，当且仅当对每个 $H\succeq 0$，都有 $\langle Z,H-X\rangle\le 0$。取 $H=0$ 和 $H=2X$，得到 $\langle Z,X\rangle=0$。再取 $H=vv^\top$，则对每个 $v\in\mathbb{R}^n$ 都有 $v^\top Zv\le 0$，因此 $Z\preceq 0$。由于 $\langle -Z,X\rangle=\|X^{1/2}(-Z)^{1/2}\|_F^2$，在这些符号条件下，$X$ 与 Z 正交等价于 $XZ=0$。反过来，$Z\preceq 0$ 和 $XZ=0$ 推出对每个 $H\succeq 0$，都有 $\langle Z,H-X\rangle=\langle Z,H\rangle\le 0$。

令 $Y=-Z$，并记 $C=A$、$D=-B$。反射保持弗罗贝尼乌斯距离。对于拟合对 (X,Y)，令 $\mathcal{E}(X,Y)$ 表示其平方损失，v_* 表示互补对上的最小值。因此

$$\mathcal{E}(X,Y)=\|C-X\|_F^2+\|D-Y\|_F^2,\qquad v_*=\min_{X,Y\succeq 0,\ XY=0}\mathcal{E}(X,Y). \tag{1}$$

（X, Y）是（1）中的一个可行对，在原变量中对应（X, $-Y$）。$C,D\succeq 0$ 的情形对应于 $A\succeq 0$ 和 $B\preceq 0$。C 和 D 的值域可能重叠，而且这些矩阵不一定交换。对于 $H\in\mathbb{S}^n$，令 H_+ 表示其正部：在 H 的一组标准正交特征基下，将每个特征值 λ 替换为 $\max\{\lambda,0\}$ 即可得到该正部。

对于线性半定规划中的原始–对偶可行对，对偶间隙等于原始矩阵与对偶松弛矩阵的迹内积。这两个矩阵都是半正定的，因此间隙为零要求它们的值域正交。Wu 等人 [10] 在一个逆问题中利用这一互补条件，对线性半定规划的目标函数和右端数据进行调整。秩界是另一个例子。Ding、Sun 和 Ye [2,的方程（2）–（3）] 使用互补矩阵，对相关矩阵 X施加秩界。对于整数 $0 \le r \le n$，条件 $\operatorname{rank} X \le r$ 等价于存在满足$0 \preceq Y \preceq I$的 Y，使得 $\operatorname{tr} Y = n - r$，且 $XY = 0$。事实上，如果 $\operatorname{rank} X \le r$，我们可以取 Y，使其投影到 $\ker X$ 的一个（$n - r$）维子空间上。反过来，任何这样的 Y 都满足 $\operatorname{rank} Y \ge$、$\operatorname{tr} Y = n - r$，且 $\operatorname{ran} Y \subseteq \ker X$。

该投影本身出现在 Hou、Peng 和 Ding [6, eqs. (33)–(34), (57)] 的增广拉格朗日方法中。该方法引入一对受互补约束的独立松弛矩阵。固定其他变量后，松弛目标函数包含到乘子平移后的约束值的平方距离，以及到前一松弛对的近端平方距离。若它们的权重分别为 $\rho/2$ and $b/2$，且满足 $\rho, b > 0$，则通过配方可得 $(\rho + b)\mathrm{E}(X, Y)/2$，相差一个加性常数。这里，$(C, D)$ 是反射变量中两个中心的加权平均。因此，精确的松弛变量更新即求解 (1)。对于 $c \ge 1$，一个损失至多为 cv_* 的可行对，其松弛目标函数至多比精确块最小值高 $(\rho + b)(c - 1)v_*/2$。这样的可行对可用作非精确方法中的试探更新；该方法还会在接受之前检验下降性和极限法向残差 [6,Algorithm 2, eqs. (39)–(40)]。

当 n 属于输入时，判定 v_* 是否至多为给定的有理数阈值是 NP 难的，即使对于有理正定 C 和 D [6,注 4 及附录 E]也是如此。

注 1.1（不存在全多项式时间近似方案）。对于 (1)，全多项式时间近似方案（FPTAS）将针对任意有理数 $0 < \varepsilon < 1$，在输入比特长度和 $1/\varepsilon$的多项式时间内，返回损失为 $\widehat{v} \le (1+\varepsilon)v_*$的可行对。即使对于有理$C$，$D \succ 0$，这样的方案也将推出 P = NP。为说明这一点，取一个具有 $n \ge 2$ 个顶点的简单无权图，并令 m 为其最大割大小。 [6, 附录 E]中的归约在多项式时间内构造这些矩阵。当 $\delta = (1024n^{16})^{-1}$ 且 $c_0 = \|D\|_F^2$时，其最终估计为 $|v_* - c_0 + 8\delta^2 m| \le \frac{1}{2}\delta^2$。$D$ 的对角元为 1，...，n^2，其非对角部分的弗罗贝尼乌斯范数小于 1。因此 $c_0 \le n^5 + 1$。可行对 $(C, 0)$ 给出 $v_* \le c_0$。选择 $\varepsilon = \delta^2/c_0$，使得 $1/\varepsilon = O(n^{37})$ 且 $0 \le \widehat{v} - v_* \le \delta^2$。由此，$(c_0 - \widehat{v})/(8\delta^2)$ 与 m 的差至多为 $3/16$。将其舍入到最近整数即可在多项式时间内恢复最大割大小，从而解决 NP 完全的简单最大割判定问题 [3]。对于弗罗贝尼乌斯距离本身，同样的结论也成立，因为距离因子为 $1 + \varepsilon/3$ 时，对于 $0 < \varepsilon < 1$，平方损失因子至多为 $1 + \varepsilon$。

约束 $XY = 0$ 要求 X 和 Y 的值域彼此正交。将 C和 D 分别投影到半正定锥上，可能会使它们拟合后的值域发生重叠。一旦为 X 选择了一个子空间，并为 Y 选择了其正交补，两个半正定投影便能给出最佳拟合。我们将投影算子放宽为作用于 X的矩阵 $0 \preceq T \preceq I$。将其特征值进行阈值化即可得到候选子空间。表 1给出了损失界。阈值舍入的界是否尖锐仍是一个开放问题（注 2.2）。

2 半定松弛与谱阈值舍入

设 P为正交投影算子，并令 $Q = I - P$。用 $v(P)$ 表示满足 X，$Y \succeq 0$, $X = PXP$ 和 $Y = QYQ$ 的条件下 $\mathcal{E}(X, Y)$ 的最小值。最后两个条件使值域位于正交子空间中，因此每一对这样的对象都是互补的。将输入分解为

表1：由（4）的最优解进行谱阈值舍入时的平方损失因子。对于 $v_* > 0$，每个因子都给出了平方损失除以 v_*的上界。其平方根给出了相应距离比的上界。

假设	平方损失因子
$C, D \in \mathbb{S}^n$	1.189
$C, D \succeq 0$	1.15487
T 在 (0,1) 中至多有两个不同特征值	1.18

相对于 $\operatorname{ran} P$ 和 $\operatorname{ran} Q$ 的分块，可得

$$(X_P, Y_P) = ((PCP)_+, (QDQ)_+), \quad v(P) = \|C\|_F^2 + \|D\|_F^2 - \|(PCP)_+\|_F^2 - \|(QDQ)_+\|_F^2. \tag{2}$$

对于第一个矩阵，只有 P 块可以变化。根据半正定锥的 Frobenius投影公式 [4, 定理 2.1], ，截断PCP的负特征值可使该块的损失最小。在 $\operatorname{ran} Q$ 上对 QDQ 进行相同论证即可。这两个压缩投影也出现在 [6, 第 6.1] 节的固定子空间松弛拟合中。反过来，对于任意互补对，可以令 P 为投影到 $\operatorname{ran} X$ 上的投影算子，因为 $\operatorname{ran} Y \subseteq \ker X$。由此可得 $v_* = \min_{P=P^\top=P^2} v(P)$。由于$v(P)$ 在正交投影算子的紧集上连续，因此最小值可以取得。

2.1 松弛

在投影算子 P 处，映射 $Z \mapsto PZP$恰好保留 $\operatorname{ran} P$ 上的块。要将此映射扩展到 $0 \preceq T \preceq I$，需在带有弗罗贝尼乌斯内积的整个空间 $\mathbb{R}^{n\times n}$ 上进行。通过定义三个线性算子对矩阵 Z 的作用如下：

$$S_T(Z) = \tfrac{1}{2}(TZ + ZT), \qquad J_T(Z) = TZ - ZT, \qquad H_T(Z) = 2Z - TZ - ZT.$$

令 $\dagger$表示 Moore–Penrose逆，并令 $\mathcal{K}_T = S_T - \frac{1}{2} J_T H_T^\dagger J_T$。选取 T 的一组标准正交特征基 (e_i)。若 e_i, e_j的特征值分别为 x, y，则矩阵单位 $e_i e_j^\top$是 S_T, J_T, H_T 的公共特征向量，其特征值依次为 $(x+y)/2$、$x-y$ 和 $2-x-y$。这些矩阵单位构成 $\mathbb{R}^{n\times n}$ 的一组标准正交基，因此这三个算子均为自伴且彼此交换。$\mathcal{K}_T$ 在 $e_i e_j^\top$上的特征值为

$$k(x,y) = \frac{x(1-x) + y(1-y)}{2 - x - y} \quad ((x,y) \neq (1,1)), \qquad k(1,1) = 1. \tag{3}$$

由于 $1 - k(x, y) = ((1-x)^2 + (1-y)^2)/(2-x-y)$ 不在 $(1, 1)$ 处，而 $k(x, y)$ 是 x 和 y的加权平均，权重为 $1-x$，$1-y$，因此我们在 $[0,1]^2$ 上有 $0 \le k \le 1$。加权平均公式还保证了在 $(1, 1)$ 处的连续性。因此，$\mathcal{K}_T$ 是半正定的，且其所有特征值都属于 $[0,1]$。在投影算子 P 处，当两个特征值都为 1 时，乘子为 1；否则为 0，因此 $\mathcal{K}_P(Z) = PZP$。对于 $U \in \mathbb{R}^{n\times n}$，定义扩展实值二次型

$$q_T(U) = \begin{cases} \langle U, \mathcal{K}_T^\dagger U\rangle, & U \in \operatorname{ran} \mathcal{K}_T, \\ +\infty, & \text{otherwise.} \end{cases}$$

对于对称输入 H，令 $\Phi_H(T, U)$ 表示 U 对松弛损失的贡献。我们用 $L(T, U, W)$ 表示总损失，并用 ℓ_* 表示其最小值，

$$\begin{aligned} \Phi_H(T,U) &= \|H\|_F^2 - 2\langle H, U\rangle + q_T(U), \\ L(T,U,W) &= \Phi_C(T,U) + \Phi_D(I-T, W), \\ \ell_* &= \min_{0\preceq T\preceq I,\ U,W\succeq 0} L(T,U,W). \end{aligned} \tag{4}$$

若可行三元组 $(T,\ U,\ W)$ 满足 $L(T,\ U,\ W) < \infty$，则称其为有限的；等价地， $q_T(U)$ 和 $q_{I-T}(W)$ 均为有限值。令 x_i 表示 T 的特征值，并使用相应的特征基表示各项 U_{ij}。 q_T 定义中的值域条件要求：每当 $k(x_i,\ x_j) = 0$ 时， $U_{ij} = 0$ 。在此定义域上，

$$q_T(U) = \sum_{k(x_i,x_j)>0} U_{ij}^2/k(x_i,x_j).$$

对于线性矩阵不等式，使用 $\mathbb{R}^{n\times n}$ 的固定标准矩阵单位基，且该基的选取与 T 无关。令 $z \in \mathbb{R}^{n^2}$ 为按列堆叠 U 各列所得的向量，基的顺序也采用相同排列。对于 $\tau \in \mathbb{R}$，上图集条件 $q_T(U) \le \tau$等价于

$$\begin{pmatrix} \tau & z^\top & 0 \\ z & S_T & J_T \\ 0 & J_T & 2H_T \end{pmatrix} \succeq 0. \tag{5}$$

同时对角化可得 $\ker H_T \subseteq \ker J_T$。根据广义舒尔补判据 [1, 附录 A.5.5], 消去 $2H_T$ 后，得到一个对角块为 τ、K_T 且非对角块为 z的分块矩阵。半正定性要求 z 与 $\ker \mathrm{K}_T$正交。在 $\operatorname{ran} \mathrm{K}_T$上配方后，得到 $q_T(U) \le \tau$。(5) 中的所有项都是 (T,U,τ) 的仿射函数。再引入一个标量变量来界定 $q_{I-T}(W)$，并在 (5) 中以 $(I-T,W)$ 替换 (T,U)。将两个标量上界与 (4) 中的仿射项之和最小化，得到一个含有两个阶数为 $2n^2+1$ 的线性矩阵不等式的半定规划。当 $T = P$是投影算子时，K_P投影到矩阵 $U = PUP$ 上。在该子空间上， $q_P(U) = \|U\|_F^2$。同样， $q_{I-P}(W)$ 有限当且仅当 $W = (I-P)W(I-P)$，此时其值为 $\|W\|_F^2$。因此，在 $T = P$ 处最小化松弛后的目标函数可得 $v(P)$，故 $\ell_* \le v_*$。对于任意可行的 T，界 $q_T(U) \ge \|U\|_F^2$给出 $L(T,U,W) \ge \|C-U\|_F^2 + \|D-W\|_F^2$。因此，$0 \le \ell_*$ 以及 U,W 在每个有限下水平集上都有界。(5) 中的上图集是闭集，且 T 的可行集是紧集，因此该松弛问题能够达到其最小值。

2.2 阈值舍入

令 $\lambda_1 > \cdots > \lambda_m$ 为 T 的不同特征值，其特征子空间为 E_1，……， E_m。从零子空间开始，依次加入 E_1，……， E_m，最终得到 $\mathbb{R}^n$。在步骤 j 之后，该子空间恰好包含特征值至少为 λ_j 的特征子空间。每个重复特征值都连同其完整的特征子空间一起加入。令 $\mathcal{P}(T)$ 表示这些 $m+1 \le n+1$ 子空间上正交投影算子的集合。该族中的最小损失为

$$V(T) = \min_{P\in\mathcal{P}(T)} v(P). \tag{6}$$

算法 1 半定松弛与谱阈值舍入

要求： 实对称输入 C， D。

1: 为（4）选择一个最优三元组 $(T,\ U,\ W)$。2: 对 T 的不同特征值进行排序，并构造嵌套投影算子 $\mathcal{P}(T)$。3: **对于** $P \in \mathcal{P}(T)$ **执行**4: 设置 $Q = I - P$，计算 $(PCP)_+$ 和 $(QDQ)_+$，并根据（2）计算 $v(P)$。5: **结束循环**6: **返回** 使 $v(P)$ 取最小值的一对 $(X_P,\ Y_P)$。

对于固定的 T，算法 1 返回一对损失 $V(T)$，即使有多个投影算子达到最小值。对于任意阈值 $s \in [0,1]$，记 $P_s(T)$ 为投影到 T 的特征子空间之和的投影算子，其中这些特征子空间对应的特征值至少为 s。对于固定的 T，将其简写为 P_s。在证明中， s 在 $[0,1]$，上均匀分布，而 $\mathbb{E}$ 表示关于随机阈值的期望。每个 P_s 都属于 $\mathrm{P}(T)$，因此在每个阈值处， $V(T) \le v\,P_s$，从而 $V(T) \le \mathbb{E}\,v\,P_s$。

定理 2.1。 对于任意有限可行三元组 $(T,\ U,\ W)$，其满足 (4)，

$$V(T) \le 1.189\, L(T,U,W) \quad (C,D \in \mathbb{S}^n), \qquad V(T) \le 1.15487\, L(T,U,W) \quad (C,D \succeq 0).$$

令 c 为适用不等式中的常数。算法 *1* 返回一对平方损失$V(T) \le c v_*$。更一般地，对于绝对容差 $\eta \ge 0$，如果一个有限可行三元组满足$L(T,\ U,\ W) \le \ell_* + \eta$，则阈值舍入给出 $V(T) \le c(v_* + \eta)$。

对于一个最优三元组， $L(T,\ U,\ W) = \ell_* \le v_*$。对于最后一个结论，使用 $L(T,\ U,\ W) \le \ell_* + \eta \le v_* + \eta$。这里 η 表示松弛目标函数中的绝对误差。两个结论都假设所选子空间具有精确特征子空间和精确拟合。若要得到相对误差的比特复杂度界，还需要估计将松弛问题求解到该精度所需的代价。如果$v_* = 0$，则最优松弛解给出 $V(T) = 0$，因此返回的数对等于 (1) 中的输入数对。

注 2.2。 定理 1.189 和1.15487 中的界，以及命题 2.1和 59/50 命题 5.1中的界，来自对均匀阈值下期望损失的上界估计。舍入规则返回所有候选子空间中损失最小的结果。对于 $v_* > 0$，在最优松弛解处， $V(T)/v_*$ 能达到多接近这些界，目前仍未解决。

3 任意对称输入的界

固定一个有限可行三元组 (T_0, U, W)。我们首先在固定基下改变 T_0 的特征值，同时保持 U, W 不变。若特征值仍按顺序排列，且每个原有特征子空间保持完整，则新矩阵的每个阈值投影算子都已属于$\mathcal{P}(T_0)$。我们利用这一自由度，在可容许特征值上最小化 L 。极小化点处的导数给出一个恒等式，可抵消期望损失估计中的两个项。

3.1 关于特征值的驻点性

选取 T_0 的特征基，并按特征值递减排列。在保持 U, W 不变的条件下，在此基下为对角的矩阵 T 中最小化 $L(T,U,W)$。要求 T 的对角元按非增顺序属于 $[0,1]$，并且对应于 T_0 相等特征值的对角元相等。这个紧集包含 T_0，因此由下半连续性可知，存在一个极小化点 T，其损失有限。相邻的不同特征值在 T 处可能变为相等。此时，某个阈值要么同时包含两个原有特征子空间，要么都不包含。因此，$\mathcal{P}(T) \subseteq \mathcal{P}(T_0)$ 并且

$$V(T_0) \le V(\bar{T}), \qquad L(\bar{T},U,W) \le L(T_0,U,W). \tag{7}$$

本小节中始终使用所选特征基，并将 T 的特征值记为 x_i。为沿可容许扰动求导，我们使用多项式

$$\lambda(x) = x(1-x)(2x-1)\big[-0.001447 + 0.213196(2x-1)^2 - 0.166343(2x-1)^4\big].$$

在全文中，用作系数的有限小数均表示精确的有理数。扰动 $x \mapsto x + \epsilon\lambda(x)$ 固定两个端点。由于 λ' 有界，该映射在所有充分小的 $|\epsilon|$ 下严格递增。因此，内部特征值仍处于 $(0,1)$ 中，且特征值保持其顺序和相等关系。由 (3) 可知，零乘子仅出现在端点对处，因此所有零乘子 $k(x_i, x_j)$ 的位置均保持不变。将关于第一和第二个自变量的偏导数分别记为 $\partial_1 k$、$\partial_2 k$。对于 $k(x,y) > 0$ 和 $(x,y) \ne (1,1)$，定义

$$J_\lambda(x,y) = \frac{\lambda(x)\partial_1 k(x,y) + \lambda(y)\partial_2 k(x,y)}{k(x,y)^2}.$$

令 $J_\lambda(1,\ 1) = 0$，因为该扰动固定了 1。还令 $J_\lambda(x,\ y) = 0$，即使 $k(x,\ y) = 0$，只是为了将这些指标包含在下面的求和中。相应的矩阵元素由

二次型的有限性而消失，并且零乘子在该扰动过程中保持为零。对于 $k > 0$， $1/k$ 沿扰动在 $\epsilon = 0$ 处的导数为 $-J_\lambda(x, y)$。此外，$1-(x+\epsilon\lambda(x)) = (1-x) + \epsilon\lambda(1-x)$，因为 $\lambda(1-x) = -\lambda(x)$。在极小化点 T处， ϵ 的两个方向都是可行的¯。因此， L 沿该曲线在 $\epsilon = 0$ 处的导数为零。只有两个二次项发生变化，从而得到

$$\sum_{ij} J_\lambda(x_i, x_j) U_{ij}^2 + \sum_{ij} J_\lambda(1-x_i, 1-x_j) W_{ij}^2 = 0. \tag{8}$$

¯ ¯在本小节的其余部分令 $c = 1.189$。由（7），只需证明 $V(T) \le cL(T, U, W)$。

3.2 阈值化和对角缩放后的期望损失

¯在 $[0,1]$ 上均匀分布地选择 s，并在本小节中令 $P_s = P_s(T)$。我们通过在 ran P_s 上对 U 进行对角缩放，为 C 构造一个拟合。对于 x， $s \in [0,1]$，定义

$$\begin{aligned} p(x) &= 0.743787 - 0.486537x + 0.627925x^2 - 0.392282x^3, \\ g(x) &= (0.426605 - 2.315288x + 2.696582x^2)/(0.12 + x), \\ \rho_s(x) &= \frac{1 + (1-x)g(x)(x-2s)}{1 - (1-x)p(x)}. \end{aligned}$$

令 R_s 在所选特征基下为对角矩阵。其在特征值 x 处的元素为：当 $x \ge s$ 时为 $\rho_s(x)$，否则为零。 $\rho_s(x)$ 的分母为正，这是因为 p 在 $[0,1]$ 上递减，且 $p(0) < 1$。于是 $R_s U R_s$ 为半正定矩阵，且其值域包含于 ran P_s 中。对于 D，使用 $I - T$ 和阈值 $1 - s$ 对 W 应用相同规则。除非 s 是 T 的特征值，否则条件 $1 - x \ge 1 - s$ 等价于 $x < s$。因此，为 W 选取的子空间几乎处处为 ran$(I - P_s)$。

对于条目 (i, j)，其中 $x_i, x_j > 0$，令 $x = \max(x_i, x_j)$，并令 $y = \min(x_i, x_j)$。对于这对值，令 $h_s = \rho_s(x)\rho_s(y)$。 $R_s U R_s$ 的 (i, j) 项为 $h_s U_{ij}$（当 $s \le y$ 时），否则为零。它对 $\mathbb{E}\,\|C - R_s U R_s\|_F^2$ 的贡献为

$$C_{ij}^2 - 2C_{ij}U_{ij}\int_0^y h_s\,ds + U_{ij}^2 \int_0^y h_s^2\,ds.$$

从缩放后的松弛损失 $c(C_{ij}^2 - 2C_{ij}U_{ij} + U_{ij}^2/k(x,y))$ 中减去该期望损失。对于固定的 U_{ij}，该差值关于 $C_{ij} \in \mathbb{R}$ 的最小值为 $M(x,y)U_{ij}^2$，其中

$$M(x,y) = \frac{c}{k(x,y)} - \int_0^y h_s^2\,ds - \frac{(c - \int_0^y h_s\,ds)^2}{c-1}. \tag{9}$$

对于几乎所有的 s，这两个缩放后的矩阵在 $\operatorname{ran} P_s$ 和 $\operatorname{ran}(I - P_s)$上均可行。因此，它们的损失从上方界定 $v(P_s)$，其上界为 (2)。

Lem引理 3.1（ $M + J_\lambda$ 的逐点下界）。 在 $[0,1]$ 上定义 f，并在 $[0,1$上定义 θ) b y

$$\begin{aligned} f(x) &= (-0.398243 + 0.543957x - 0.533234x^2)/(0.057958 + x), \\ \theta(x) &= 2.191072\log(1-x) - 1.031425x - 0.481436x^2. \end{aligned}$$

对于 $x, y \in [0,1)$，令

$$G(x,y) = \sqrt{(1-x)(1-y)}\, f(x) f(y) \cos(\theta(x) - \theta(y)).$$

设 G（x、 y） $= 0$，若 $x = 1$ 或 $y = 1$。则对于 $0 < y \le x \le 1$,

$$M(x,y) + J_\lambda(x,y) \ge G(x,y). \tag{10}$$

证明。对于 $y<1$，令 $b=1-y$ 且令 $t=\sqrt{(1-x)/(1-y)}$。由于 b，$k(x,\ y)>0$，所需不等式等价于 $(k(x,\ y)/b)(M+J_\lambda-G)\geq 0$。附录 A.1 给出了该缩放差值在 $[0,1]^2$ 的覆盖中的每个矩形上的正下界。在 $x=y=1$ 处，我们有$k=h_s=1$，因此 (9) 给出 $M=0$。根据定义，$J_\lambda=G=0$ 在该处同样成立。 □

任意输入下定理 2.1 的证明。矩阵 $[G(x_i,x_j)]$ 是一个格拉姆矩阵。令 $x=1$ 处的向量为零向量，令 $x<1$处的向量为 $\sqrt{1-x}\,f(x)(\cos\theta(x),\sin\theta(x))$。$x_i$、$x_j$ 处向量的内积等于 $G(x_i,x_j)$。对于大小相同的矩阵，记 $\circ$ 为逐元素乘积。根据舒尔积定理 [5, 第 7.5] 节，$U\circ U=[U_{ij}^2]$ 也是半正定的。直接验证如下：令 $u_i^\top$为 $U^{1/2}$的第 i 行。那么 $U_{ij}=u_i^\top u_j$ 且 $U_{ij}^2=u_iu_i^\top,u_ju_j^\top$。两个半正定矩阵 A、B的迹内积非负，因为 $\operatorname{tr}(AB)=\operatorname{tr}(A^{1/2}BA^{1/2})\geq 0$。因此 $\sum_{ij}G(x_i,x_j)U_{ij}^2\geq 0$。

若 $x_i=0$，则 $q_{\bar{T}}(U)$ 的有限性迫使 $U_{ii}=0$。不等式 $U_{ij}^2\leq U_{ii}U_{jj}$ 因而给出：对每个 j，都有 $U_{ij}=0$。该行或该列中的每个元素都会对 $c\Phi_C(\bar{T},U)-\mathbb{E}\,\|C-R_sUR_s\|_F^2$ 贡献 $(c-1)C_{ij}^2\geq 0$。对于两个特征值均为正的元素，将 (10) 乘以U_{ij}^2 后求和。由于涉及 G 的和非负，C 的损失差至少为 $-\sum_{ij}J_\lambda(x_i,x_j)U_{ij}^2$。$D$ 对应的下界使用 W 和特征值 $1-x_i$。涉及 J_λ 的两个和根据 (8) 相互抵消。因此，构造出的配对 $\bar{\ }$ $\bar{\ }$ $\bar{\ }$ $\bar{\ }$ 的期望损失至多为 $cL(T,U,W)$。由于 $V(T)\leq\mathbb{E}\,v(P_s)$，由 (7) 得 $V(T_0)\leq V(T)\leq cL(T,U,W)\leq cL(T_0,U,W)$。 □

4 反射后的半正定输入

固定一个有限可行三元组 $(T,\ U,\ W)$ 并假设 $C,\ D\succeq 0$。取 s 在 $[0,1]$ 上均匀分布，并记 $P_s=P_s(T)$。压缩 P_sCP_s 和 $(I-P_s)D(I-P_s)$ 均为半正定，因此是（2）中的极小化对。

固定 T 的一个特征基，将其特征值记为 x_i，并令 $a_i=1-x_i$。压缩P_sCP_s 恰好在 $s\leq\min(x_i,\ x_j)$ 时保留元素 C_{ij}。因此，该元素产生的期望平方损失为 $\max(a_i,\ a_j)C_{ij}^2$。对于 $k=k(x_i,\ x_j)>0$，配方后，$\Phi_C(T,\ U)$ 中的对应项为

$$C_{ij}^2-2C_{ij}U_{ij}+U_{ij}^2/k=(1-k)C_{ij}^2+(U_{ij}-kC_{ij})^2/k.$$

若 $k=0$，则 $q_T(U)$ 的有限性要求 $U_{ij}=0$。在这两种情况下，对松弛损失的贡献都至少为 $(1-k)C_{ij}^2$。式（3）给出 $1-k=\beta(a_i,\ a_j)$，其中

$$\beta(a,b)=\frac{a^2+b^2}{a+b}\quad(a+b>0),\qquad\beta(0,0)=0.$$

对于 $a,\ b>0$，令 $t=|\log(a/b)|$ 和 $c_+=1.15487$。将 $c_+\beta(a,\ b)-\max(a,\ b)$ 除以 $\sqrt{ab}$，得到一个仅关于 t 的函数，

$$\frac{c_+\beta(a,b)-\max(a,b)}{\sqrt{ab}}=F(t):=c_+\frac{\cosh t}{\cosh(t/2)}-e^{t/2}.\tag{11}$$

引理 4.1（余弦和的下界）。设 $(\omega_m,\ d_m)$，$m=1,\ \ldots\ldots,\ 98$，为表 2 中的非负有理频率和权重对。则

$$F(t)\geq\sum_{m=1}^{98}d_m\cos(\omega_mt)\qquad(t\geq 0).\tag{12}$$

证明。令 $\delta(t)=F(t)-\sum_m d_m\cos(\omega_mt)$。附录 A.2 给出 $\delta(j/4000)\geq 2535846/10^{12}$其中 $j=0,\ldots,12000$ 以及 $|\delta''(t)|\leq 54$ 在 $[0,3]$ 上。相邻节点之间 δ 的线性插值函数

节点处的值至少为两个端点值中的较小者。在节点区间 $[a,b]$ 上，线性插值余项为 $\delta''(\xi)(t-a)(t-b)/2$，其中某个 $\xi\in(a,b)$ [8, 见公式 (3.3.5)]。其绝对值至多为$54/(8\cdot 4000^2)$，小于 δ 在节点处的下界。因此 $\delta\geq 0$ 在 $[0,3]$ 上成立。对于 $t\geq 3$，附录 A.2 给出 $F(t)>\sum_m d_m\geq\sum_m d_m\cos(\omega_m t)$。

□

定理 *2.1* 在 C时，$D\succeq 0$。对于每个 $a_i>0$，通过拼接坐标对构造一个向量

$$\sqrt{a_i d_m}\,(\cos(\omega_m\log a_i),\sin(\omega_m\log a_i))$$

对应于频率 m。对于 $a_i=0$，取零向量。令 K 为这些向量的格拉姆矩阵。如果 a_i、$a_j>0$，则由余弦差公式和 (11)–(12) 得到 $K_{ij}\leq c_+\beta(a_i,a_j)-\max(a_i,a_j)$。如果 a_i 或 a_j 为零，则 $K_{ij}=0$，右侧等于 $(c_+-1)\max(a_i,a_j)\geq 0$。

将关于 K_{ij} 的下界乘以 C_{ij}^2 并求和。由于 $C\succeq 0$，关于 $U\circ U$ 在小节 3 中的论证给出 $C\circ C\succeq 0$。它与 K 的迹内积为非负数。因此$\sum_{ij}\max(a_i,a_j)C_{ij}^2\leq c_+\sum_{ij}\beta(a_i,a_j)C_{ij}^2$。现在，关于 $\Phi_C(T,U)$的下界给出

$$\mathbb{E}\,\|C-P_sCP_s\|_F^2\leq c_+\Phi_C(T,U).$$

将相同的估计应用于 D、W 使用 $I-T$和阈值 $1-s$。相应的投影算子 $I-P_s$几乎必然成立。将两个估计相加得到 $\mathbb{E}\,v(P_s)\leq c_+L(T,U,W)$，以及 $V(T)\leq\mathbb{E}\,v(P_s)$ 由（6）可得。

□

5 至多具有两个不同内部特征值时的界

对于 (0，1) 中的两个不同特征值，我们在每个内部特征子空间上分别进行缩放。随后，半正定性将两个特征子空间之间的块的弗罗贝尼乌斯范数界定为两个对角块的范数之和的函数。

命题 5.1。假设 T 在 $(0,1)$ 中至多具有两个不同特征值。对于每个有限可行三元组 $(T、U、W)$，有 $V(T)\leq(59/50)L(T、U、W)$。特征值 0 和 1 可以同时存在，其重数以及内部特征值的重数均不受限制。

证明。令 $c=59/50$ 和 $\alpha=c-1$。首先，假设存在两个不同的内部特征值$0<y<x<1$，并令 $r=y/x$。在 $[0,1]$ 上均匀分布地选取 σ，并记作 $P_\sigma=P_\sigma(T)$。对于参数 $\kappa\in\{0,\ 1/2,\ 1\}$，令

$$u=1+\kappa(1-r)(1-x),\qquad b=1-r\kappa(1-x)[2+\kappa(1-r)(1-x)].$$

对于 $b\geq 0$，在 T的特征基下定义对角矩阵 R。在 y-特征子空间上，若 $\sigma\leq y$，其对角元素为 $1/\sqrt{y}$；否则为零。在 x-特征子空间上，对于 $\sigma\leq y$，对角元素为 $u/\sqrt{x}$；对于 $y<\sigma\leq x$，为 $\sqrt{b}/x$；对于 $\sigma>x$，为零。在特征值 0 和 1 对应的特征子空间上，对每个 σ 分别取零和一。于是 RUR 为半正定矩阵，且其值域包含于 $\operatorname{ran}P_\sigma$。$q_T(U)$ 的有限性和 U 的半正定性迫使特征值 0 对应的行和列均为零。

固定 $\xi,\eta\in\{x,y,1\}$，其中 $k(\xi,\eta)>0$。令 h_σ为在阈值 σ处，分配给 ξ 和 η 的 R 的对角元的乘积。定义

$$m_{\xi\eta}=\frac{c}{k(\xi,\eta)}-\mathbb{E}h_\sigma^2-\frac{(c-\mathbb{E}h_\sigma)^2}{\alpha}.$$

对于一个元素 U_{ij}，其行和列分别对应特征值 ξ 和 η，按照（9）对 C_{ij}进行最小化，可得到 $m_{\xi\eta}U_{ij}^2$，作为该元素对 $c\Phi_C(T,U)$ –的贡献的下界

$\mathbb{E}\,\|C-RUR\|_F^2$。附录 A.3 利用伯恩斯坦系数表明，对于每一对 $(x,\ y)$，都可以选择 κ，使得 $b\geq 0$ 且

$$m_{xx},m_{yy},m_{1x},m_{1y}\geq 0,\qquad m_{xy}+\sqrt{m_{xx}m_{yy}}\geq 0. \tag{13}$$

涉及零特征值的项对损失差的贡献为 $(c-1)C_{ij}^2\geq 0$。对于特征子空间中对应特征值 1 的部分，两个乘子均为 1，因此 $m_{11}=0$。式 (13) 中关于 m_{1x} 和 m_{1y} 的界涵盖了其他涉及特征值 1 的项。对于 $\xi,\eta\in\{x,y\}$，令 $U_{\xi\eta}$ 表示这些特征子空间之间的块。两个非对角块互为转置，因此它们的总贡献为$2m_{xy}\|U_{xy}\|_F^2$。当 $m_{xy}<0$时，提取因子 $U=\Gamma\Gamma^\top$，并令 Γ_x,Γ_y 为相应的行块。此时 $\|U_{xy}\|_F^2=\Gamma_x^\top\Gamma_x,\Gamma_y^\top\Gamma_y$，故由柯西–施瓦茨不等式可得 $\|U_{xy}\|_F^2\leq\|U_{xx}\|_F\|U_{yy}\|_F$。由式 (13) 可知，$x$和 y特征子空间上的四个块的总贡献至少为

$$\left(\sqrt{m_{xx}}\,\|U_{xx}\|_F-\sqrt{m_{yy}}\,\|U_{yy}\|_F\right)^2.$$

若 $m_{xy}\geq 0$，则四个区块的贡献均非负。若 x 是唯一的内部特征值，则任取 $y\in(0,x)$，并令辅助 y 子空间为零维。所有涉及该子空间的项均为零。若 $T=P$ 是投影算子，则 $P\in\mathcal{P}(P)$；而 $q_P(U)$ 和 $q_{I-P}(W)$ 的有限性意味着 $U=PUP$ 且 $W=(I-P)W(I-P)$。因此，$V(P)\leq v(P)\leq L(P,U,W)$。

对于 D,W，将此构造应用于阈值为 $1-\sigma$ 的 $I-T$。两种缩放可以采用不同的 κ 值，因为 κ 只影响所选子空间内的乘子。这些子空间几乎必然互补。将两个损失界相加，得到 $V(T)\leq\mathbb{E}\,v(P_\sigma)\leq(59/50)L(T,U,W)$。

□

对于 $n\leq 2$，T 在(0,1)中自动至多有两个不同特征值。在更高维情形下，谱约束 $0\preceq T\preceq I$ 允许出现更多特征值。对于任意对称输入，$V(T_*)\leq 1.18v_*$ 是否对最优解中出现的每个 T_* 都成立，仍是一个未解决的问题（4）。

致谢

在作者于2026年11月21日举行的研究报告会上，香港理工大学应用数学系的齐厚铎询问是否有求解该投影问题的近似算法。作者感谢他提出这个问题，这促使作者在当晚稍后撰写了这篇短文。作者在撰写本文时使用了生成式人工智能，并对其内容承担全部责任。

A 标量不等式的验证

A.1 逐点不等式的区间界

固定 $c=1.189$，其中 p、g、λ如第 3 节所述，而 f、θ如引理 3.1 所述。对于 $0<y<1$和 $y\leq x\leq 1$，令 $b=1-y$、$t=\sqrt{(1-x)/(1-y)}$，以及 $u=t^2$。于是 $y=1-b$、$x=1-bu$、$0<b<1$和 $0\leq t\leq 1$。我们在 (b,t) 中的闭正方形 $[0,1]^2$ 上计算区间界。在此计算中，不带参数的 k、M、J_λ、G 表示它们在 (x,y) 处的值。令 $v=y+ux$，则有

$$k=\frac{v}{1+u}=1-b\,\frac{1+u^2}{1+u}.$$

对于乘积 h_s（在9）中，令 $h_s=1+bH_s$。令 $a_s=ug(x)(x-2s)$、$b_s=g(y)(y-2s)$以及 $D_0=(1-bup(x))(1-bp(y))$，可得

$$H_s=\frac{up(x)+p(y)(1-bup(x))}{D_0}+\frac{a_s+b_s(1+ba_s)}{D_0}.$$

多项式 p 在 $[0,1]$ 上递减，因为它的导数是一个首项系数为负且判别式为负的二次多项式。由于 $p(0) < 1$， D_0 的两个因子均为正。在 $[0,1]$ [8, 第 3.5 节(v)]，上使用三点高斯–勒让德求积公式，节点为 $(r_1, r_2, r_3) = (\tfrac{1}{2} - \sqrt{3/20}, 1/2, \tfrac{1}{2} + \sqrt{3/20})$，权重为 $(w_1, w_2, w_3) = (5/18, 4/9, 5/18)$。对于 $i = 1, 2$ 令 $m_i = \sum_{j=1}^{3} w_j (H_{yr_j})^i$。矩恒等式 $\sum_j w_j r_j^{\ell} = 1/(\ell+1)$ 对于 $\ell = 0, \dots, 4$ 给出 $\int_0^y H_s^i \, ds = y m_i$，因为 H_s 是关于 s的二次多项式。

对于 $t > 0$，取 $\omega = 2.191072$ 并使用相位

$$-2\omega \log t + (y - x)\big[-1.031425 - 0.481436(x + y)\big].$$

定义 G_0为 $t f(x) f(y)$ 乘以此相位的余弦，其中 $t > 0$，并令 $G_0 = 0$在 $t = 0$处取值。当 b时，$t > 0$，相位等于 $\theta(y) - \theta(x)$，因为 $(1-x)/(1-y) = t^2$。因此 $G = bG_0$。这个显式相位也定义了 G_0 在 $b = 0$ 上的值，其中 $t > 0$，无需计算 $\theta(1)$。

令 $C_0 = y m_2 + (y m_1 - 1)^2/(c - 1)$，并且对于 b， $k > 0$，令 $D_\lambda = (k/b) J_\lambda$。将 $h_s = 1 + b H_s$ 代入（9），并合并 b 的各次幂，得到

$$\frac{k}{b}(M + J_\lambda - G) = c\,\frac{1 + u^2}{1 + u} - k - k(bC_0 + G_0) + D_\lambda. \tag{14}$$

令 ℓ 为从 $\lambda(x)$ 中去掉因子 $x(1-x)$ 后所得的多项式，这样 ℓ 在 0 和 1 处也有定义。对于导数项，令

$$d_x = \ell(x)\frac{u^2 + 2u - 1}{1 + u}, \qquad d_y = \ell(y)\frac{1 + 2u - u^2}{1 + u}.$$

对 k 求导，得到 $D_\lambda = (ux\,d_x + y\,d_y)/v$，该式在 $b, k > 0$处成立。凡是 $v > 0$，均使用此表达式。此时，其值是 d_x 和 d_y的凸组合。因此，在每个矩形上， d_x, d_y 的下界中较小者，都是满足 $v > 0$的每个点处 D_λ 的下界。即使 v 的区间包含零、无法除以该区间，此结论仍然成立。方程 (14) 仅在 $b, k > 0$时才需要。 $b = 0$ 的情形给出 $x = y = 1$，引理 3.1 对此作了直接处理。

令 $[t_0, t_1]$ 为矩形的 t区间。若 $t_0 < 0.28$，则用 $[-1, 1]$给出 G_0 中的余弦的界。否则，$t_0 > 0$，我们按如下方式对显式相位中的对数进行界定。在 $[t_0, t_1]$的中点 m 处，令 $z = (1-m)/(1+m) \le 9/16$。$\log((1+z)/(1-z))$ 的级数 $= -\log m$ 给出 $-2\omega \log m$的近似 $4\omega \sum_{j=0}^{55} z^{2j+1}/(2j+1)$。截断余项的上界为 $4\omega z^{113}/[113(1 - z^2)] < 10^{-16}$。对于 $t \in [t_0, t_1]$，由中值定理可得$|2\omega(\log t - \log m)| \le \omega(t_1 - t_0)/t_0$。将此界与截断误差相加，即可得到整个区间上 $-2\omega \log t$ 的界。

对于余弦，令 a 为与其自变量区间中点最近的 2^{-15} 的整数倍。这些值满足 $|a| < 10$。从麦克劳林级数的常数项开始，在第 j步将前一项乘以 $-a^2/[(2j)(2j-1)]$，初始值为 $j = 1$。求和至 46 次项，并用 $10^{48}/48! < 10^{-12}$ 的绝对值包住余项。要对整个自变量区间上的余弦进行界定，根据 1-Lipschitz性质，将 $\cos a$ 的包围区间扩大，扩大幅度为从 a 到较远端点的距离。再将结果与 $[-1, 1]$取交集。

对和、差及乘积使用区间算术，使每个计算得到的区间都包含所有可能值 [7, Chs. 2–4]。对包含零的区间进行平方时，将零作为下端点。除以正区间时，乘以其倒数。使用霍纳法则计算系数多项式和截断的对数级数。将每个计算得到的 binary64 端点 a 向外调整 $(1 + |a|)\,2^{-49}$，以涵盖端点及调整本身的舍入误差。对输入常数和高斯节点也使用相同的误差余量。通过平方可知，用 0.3872983346207417 近似 $\sqrt{3/20}$ 的误差小于 10^{-16}。

从正方形 $[0,1]^2$ 开始。在每个矩形上计算 $k = v/(1+u)$。如果D_0 的区间下端点不为正，则将该矩形二分。否则，对 H_s 的两项各计算一次 $1/D_0$ 。对于 D_λ，先取 d_x, d_y 的下端点中较小者。如果 v 的下端点为正，则还计算 $(ux\,d_x + y\,d_y)/v$，并保留两个下界中较大的一个。如果 (14) 右侧的下界严格为正，则接受该矩形；否则将其二分。每次二分都在中点处进行，交替沿 b 和 t 方向切分，并从b开始。11 839 610 个接受的矩形覆盖整个正方形。其最大深度为 31，最小下界大于 1.7×10^{-10}。

A.2 余弦下界的验证

令 $\delta(t) = F(t) - \sum_m d_m \cos(\omega_m t)$，其中 F由（11）给出，且 $c_+ = 1.15487$。表 2 中的频率和权重给出 $\mu := \sum_m d_m < 4/25$ 以及 $\sum_m d_m \omega_m^2 < 50$。总权重 μ为余弦和提供上界，适用于 $t \ge 3$。第二个和为区间 $[0,3]$上余弦和的二阶导数绝对值提供上界。我们用它将节点处的 δ 下界推广到节点之间的区间。

表 2：余弦界的数据（12）。每对相邻列给出标题中所示的缩放频率和权重。每行中的各对数据从左向右读取。

$100\omega_m$	$10^{10}d_m$	$100\omega_m$	$10^{10}d_m$	$100\omega_m$	$10^{10}d_m$	$100\omega_m$	$10^{10}d_m$	$100\omega_m$	$10^{10}d_m$
297	132530224	312	727610379	578	53987891	593	204854308	866	32656035
881	84909950	896	659902	1156	45136331	1186	22232023	1447	29884142
1477	13113541	1737	20568717	1767	9213989	2028	15637737	2058	6156926
2320	12938571	2350	3655984	2611	10571697	2641	2470659	2902	8812243
2932	1688686	3193	7439563	3223	1184365	3485	6592322	3515	596429
3776	5682429	3806	399829	4067	4904453	4097	301871	4359	4391955
4389	100622	4650	3810937	4680	104577	4942	3422173	4972	10216
5233	2983613	5263	50448	5525	2690875	5816	2349846	5831	51934
6107	2051889	6122	102072	6399	1928833	6429	8925	6690	1719182
6720	32956	6982	1585608	7012	1771	7273	1418523	7303	26625
7565	1317605	7856	1183139	7886	21876	8148	1101216	8439	995000
8469	15422	8731	925759	9022	837907	9052	11290	9314	778576
9605	700447	9635	15101	9897	653642	9927	3180	10188	588757
10218	15274	10480	553920	10771	505787	10801	3121	11063	466742
11354	411623	11369	14789	11646	389134	11937	343807	11952	11385
12228	314534	12258	8956	12520	291648	12811	258416	12841	4092
13103	235801	13394	200668	13409	9619	13686	185786	13977	157653
14007	3685	14269	138679	14560	109888	14575	6376	14852	95887
15143	74052	15173	2461	15435	54110	15465	3257	15726	27727
15756	10203	16018	11849	16048	7260				

在 $[0,3]$, 上取 $N = 4000$，并使用节点 $t_j = j/N$（ $j = 0$、.....、$3N$）。用 i 表示虚数单位，用 Re 表示实部。令 $\zeta_m = \omega_m/N$。在相邻节点处， $e^{\mathrm{i}\omega_m t_j}$ 乘以 $e^{\mathrm{i}\zeta_m}$。为近似该乘子，固定 $S = 10^{24}$，并令 R_S 将实部和虚部分别向下舍入到 $1/S$的整数倍。定义

$$q_m = R_S\left(\sum_{\ell=0}^{17} \frac{(\mathrm{i}\zeta_m)^\ell}{\ell!}\right), \qquad z_{m,0} = 1, \qquad z_{m,j+1} = R_S(q_m z_{m,j}).$$

指数余项至多为 $\zeta_m^{18}/18! \ < 1/S$，两个坐标舍入造成的误差小于 $2/S$。因此 $|q_m - e^{\mathrm{i}\zeta_m}| < 3/S$。假设$|z_{m,\,j} - e^{\mathrm{i}j\zeta_m}| \le 8j/S$。由于$j \le 3N$ 且 8(3N)$/S < 1$，我们有$|z_{m,\,j}| \le 2$。在下一步中，乘子误差至多贡献 2(3$/S$)，舍入至多贡献 $2/S$。从 $z_{m,\,0} = 1$ 开始，这证明了对所有 $j \le 3N$ 均有$|z_{m,\,j} - e^{\mathrm{i}j\zeta_m}| \le 8j/S$。取实部并按权重 d_m 求和，得到节点 j 处余弦和的上界 $\sum_m d_m$Re$(z_{m,\,j}) + 8j\mu/S$。

令 L_j 为 $\sum_{\ell=0}^{32}(t_j/2)^\ell/\ell!$ 向下舍入到 $1/S$的整数倍，并令 $U_j = L_j + 2/S$。在此区间上，连续被略去项的比值小于二分之一，因此正余项有界于 $2(3/2)^{33}/33! < 1/S$。于是 $L_j \le e^{t_j/2} \le U_j$。将 $z = e^{t/2}$代入（11）得

$$F(t) = (c_+ - 1)z + \frac{c_+}{z} - \frac{2c_+ z}{z^2+1} \tag{15}$$

对于 $z \ge 1$，分式 $z/(z^2+1)$ 递减。因此，在（15）的最后一项中，将 z 替换为 L_j，即可得到该项的下界。减去余弦和的上界后，可得 $\delta(t_j)$的有理数下界。对于 $j = 0, \dots, 3N$，该下界满足

$$(c_+ - 1)L_j + \frac{c_+}{U_j} - \frac{2c_+ L_j}{L_j^2+1} - \sum_m d_m \operatorname{Re} z_{m,j} - \frac{8j\mu}{S} \;\ge\; \frac{2535846}{10^{12}}.$$

记 $\operatorname{sech} u = 1/\cosh u$。在 $[0,3]$ 上，我们有 $\cosh(t/2) < 3$, $e^{t/2} < 5$，且 $|\operatorname{sech}''(t/2)| \le 1$，其中撇号表示 sech 关于其自变量的导数。对 $\cosh t/\cosh(t/2) = 2\cosh(t/2) - \operatorname{sech}(t/2)$ 求两次导数，可得该比值的二阶导数的绝对值上界为 $7/4$。$e^{t/2}$ 的二阶导数至多为 $5/4$。结合$\sum_m d_m\omega_m^2 < 50$，这些界给出 $|\delta''| < 54$，这正是引理 4.1 中插值估计所需的结果。

对于 $t \ge 3$，$z = e^{t/2} > 4$。方程（15）给出 $F(t) \ge (c_+ - 1)\, z - c_+/z \ge 4(c_+ - 1) - c_+/4 > 4/25 > \mu$。由于 $\sum_m d_m \cos(\omega_m t) \le \mu$，这就证明了12）在 $[3,\infty)$ 上成立。

A.3 两个内部特征值的多项式界

固定 $0 < y < x < 1$ 和 $\kappa \in \{0, \pm 2, 1\}$。令 $c = 59/50$，$\alpha = c - 1$，$r = y/x$，并令 $\bar{s} = \sqrt{r}$，取 u，b 如命题5.1的证明中所述。系数$m_{\xi\eta}$ 涉及 $\sqrt{b}$ 并要求 $b \ge 0$。下面的多项式检验针对每个候选 κ 定义，并将 $b \ge 0$ 列为待检验的条件之一。

对于 $b \ge 0$，将三个内部系数乘以 x 得到

$$\begin{aligned} d = x m_{xx} &= \alpha(1-x) - r(1-r)\kappa^2(1-x)^2[2 + \kappa(1-r)(1-x)]^2, \\ x m_{yy} &= \alpha(1-y)/r, \\ x m_{xy} &= cx/k(x,y) - u^2 - x(c - us)^2/\alpha. \end{aligned}$$

令 $n_0 = 1 - x + r(1-y) > 0$。则 $x/k(x,y) = (2 - x - y)/n_0$。要检验混合系数，构造多项式

$$\chi = \alpha c(2 - x - y) - n_0[\alpha u^2 + x(c - us)^2], \qquad \Delta = \alpha(1-y)d(\alpha n_0)^2 - r\chi^2.$$

由于 $\chi = \alpha n_0 x m_{xy}$，它与 m_{xy}同号。一旦 $d \ge 0$，m_{xx} 和 m_{yy} 均为非负数。如果 $\chi \ge 0$，则 $m_{xy} + \sqrt{m_{xx}m_{yy}} \ge 0$ 立即成立。如果 $\chi < 0$，不等式 $\Delta \ge 0$ 已足够，因为 $\Delta = r(\alpha n_0 x)^2(m_{xx}m_{yy} - m_{xy}^2)$。对于系数 m_{1x}，令

$$m_0 = 1 - \tfrac{1}{2}\{r(u-1)^2 + (1-r)(b-1)^2\}.$$

给定 $\sigma \le x$，事件 $\sigma \le y$ 发生的概率为 $r = y/x$。因此，$\sqrt{x}\,R$ 在 x-特征子空间上的对角项以概率 r 和 $1 - r$分别取值 u 或 $\sqrt{b}$。直接展开可得 $ru^2 + (1-r)b = 1$。将偏离 1 的平方展开，并利用 $(\sqrt{b} - 1)^2 \le (b-1)^2$，即可得到 $m_0 \le ru + (1-r)\sqrt{b} \le 1$中的下界。上界由柯西–施瓦茨不等式得出。在 m_{1x} 的定义中，我们有 $k(1,x) = x$, $\mathbb{E}h_\sigma^2 = 1$，以及$\mathbb{E}h_\sigma = \sqrt{x}\,[ru + (1-r)\sqrt{b}]$。定义 $E = \alpha(c - x) - x(c - \sqrt{x}\,m_0)^2$。那么，对于$m_{1x} \ge 0$，只需满足 $E \ge 0$。对于 m_{1y}，将 x 替换为 y，并在 $r = 1$ 处计算 E。于是 $m_0 = 1$，且此界与 κ无关。

将 $x = t^2$ 替换为 $t = \sqrt{x}$ 并代入 $r = s^2$，可使 b，d，χ，Δ，E 成为关于 $(t,\ s)$ 的多项式。用 $t = 1 - (1 - s)z$ 参数化三角形 $t \ge s$，并用 $s = 1 - (1 - t)z$参数化三角形 $s \ge t$，两种情况下均令 $z \in [0, 1]$。在第一种参数化中约去所有因子 $1 - s$，在第二种中约去所有因子 $1 - t$。由于这些因子在正方形上非负，商非负即可推出原多项式非负。

将每个二进矩形子区域重新缩放到 $[0, 1]^2$ 坐标为 (ξ, η)的区域。对于一个双次数为 $\sum_{p=0}^{m}\sum_{q=0}^{n} a_{pq}\xi^p\eta^q$ 的多项式 (m, n)，幂基到伯恩斯坦基转换 [9, 第 3 节，式 (9)]给出索引 $0 \le i \le m$，$0 \le j \le n$处的系数，

$$\sum_{p=0}^{i}\sum_{q=0}^{j} a_{pq}\frac{\binom{i}{p}\binom{j}{q}}{\binom{m}{p}\binom{n}{q}}.$$

根据二项式定理，单变量伯恩斯坦基函数 $\binom{m}{i}\xi^i(1 - \xi)^{m-i}$ 非负且和为 1。它们的张量积也具有同样的性质，因此每个多项式都不小于其最小伯恩斯坦系数 [9, 式 (10)]。对于给定的 κ，检查约化多项式 b, d, E 的所有伯恩斯坦系数是否均非负。对于 χ 和 Δ 的约化多项式，也必须至少有一个满足这一条件。若某个 $\kappa \in \{0, 1\,2, 1\}$满足这两个条件，则接受该矩形。将其余矩形在中点处二等分；对两种参数化，分别按 (z, s) 和 (t, z) 的顺序交替划分坐标。使用精确的有理数运算，得到 62 个深度不超过 11 的接受矩形。

参考文献

[1] S. Boyd 和 L. Vandenberghe。《凸优化》。剑桥大学出版社，2004 年。[2] C. Ding、D. Sun 和 J. J. Ye。带半正定锥互补约束的数学规划的一阶最优性条件。《数学规划》147 (2014)，539–579。[3] M. R. Garey、D. S. Johnson 和 L. Stockmeyer。一些简化的 NP 完全图问题。《理论计算机科学》1 (1976)，237–267。[4] N. J. Higham。计算最近的对称半正定矩阵。《线性代数及其应用》103 (1988)，103–118。[5] R. A. Horn 和 C. R. Johnson。《矩阵分析》，第 2 版。剑桥大学出版社，2013 年。[6] W. Hou、Y. Peng 和 C. Ding。一种用于半正定互补优化的非精确增广拉格朗日法。arXiv:2611.14885v1，2026 年。[7] R. E. Moore、R. B. Kearfott 和 M. J. Cloud。《区间分析导论》。工业与应用数学学会，2009 年。[8] N. M. Temme。数值方法。载于《美国国家标准与技术研究院数学函数数字图书馆》第 3 章，美国国家标准与技术研究院。访问日期：2026年11月22日。[9] J. Titi 和 J. Garloff。张量伯恩斯坦形式的矩阵方法。《应用数学与计算》346 (2019)，254–271。

[10] J. Wu、Y. Zhang、L. Zhang 和 Y. Lu。求解逆线性半定规划问题的序贯凸规划方法。亚太运筹学杂志 **33**（2016），1650025。