Answers of Prof. Sturmfels' Problems

Edited by Kenta Nishiyama (Kobe Univ.)

1. Let $\mathcal{S}$ be the set of all points $(x, y) \in \mathbb{R}^2$ such that the matrix

$$\begin{pmatrix} 1 & x & y \\ x & 1 & x \\ y & x & 1 \end{pmatrix} \quad \text{is positive semidefinite.}$$

Draw a picture of the two-dimensional spectrahedron $\mathcal{S}$ and its dual convex figure. Find the maximum of the linear function $x + y$ over $\mathcal{S}$.

Answer. There exist some equivalent definitions for positive semidefinite. Let us adopt the following condition :

$$\text{A matrix } A \text{ is positive semidefinite} \Longleftrightarrow (\text{All principal minors of } A) \geq 0.$$

The spectrahedron $\mathcal{S}$ is written as

$$\begin{aligned} \mathcal{S} &= \{(x, y) \in \mathbb{R}^2 \mid 1 \geq 0, 1 - x^2 \geq 0, 1 - y^2 \geq 0, (y-1)(2x^2 - y - 1) \geq 0\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid y - 1 \leq 0, 2x^2 - y - 1 \leq 0\}. \end{aligned}$$

By the definition, its dual convex $\mathcal{S}^*$ is written as

$$\mathcal{S}^* = \{(u, v) \in \mathbb{R}^2 \mid ux + vy \leq 1 \ (\forall (x, y) \in \mathcal{S})\}.$$

We denote boundaries of $\mathcal{S}$ and $\mathcal{S}^*$ by $g(x, y) := (y-1)(2x^2 - y - 1) = 0$ and $h(u, v) = 0$ respectively. The function h can be obtained by computing a generator of the following elimination ideal. (refer to [1, Chap 3].)

$$\begin{aligned} \langle g, ux + vy - 1, ug_y - vg_x \rangle \cap \mathbb{R}[u, v] &= \langle (u - v + 1)^2 (u + v - 1)^2 (u^2 + 8v^2 + 8v) \rangle \\ &=: \langle h \rangle \end{aligned}$$

The spectrahedron $\mathcal{S}$ and its dual convex $\mathcal{S}^*$ are shown in the following figures.

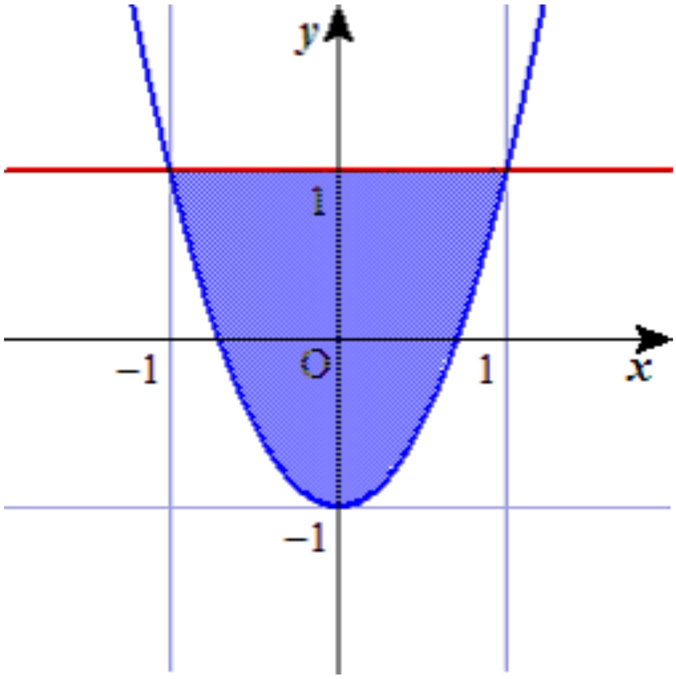


Figure 1: The spectrahedron $\mathcal{S}$

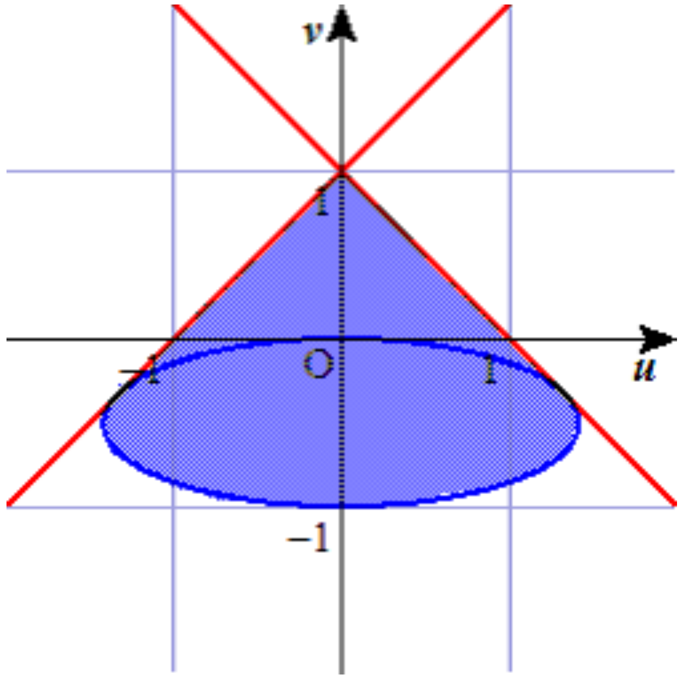


Figure 2: The dual convex of $\mathcal{S}$: $\mathcal{S}^*$

Moreover, by drawing the graph of $y = -x + k$, the function $x + y$ is maximum at the point $(1, 1) \in \mathcal{S}$ and its value is 2. □

References

[1] D. Cox, J. Little and D. O'Shea, *Ideals, Varieties, and Algorithms*, Springer-Verlag, New York, 1992.

2. Let $\mathcal{S}'$ be the set of all points $(x, y) \in \mathbb{R}^2$ such that the matrix

$$\begin{pmatrix} 1 & x & x+y \\ x & 1 & y \\ x+y & y & 1 \end{pmatrix} \quad \text{is positive semidefinite.}$$

Draw a picture of the two-dimensional spectrahedron $\mathcal{S}'$ and its dual convex figure. Find the maximum of the linear function $x + y$ over $\mathcal{S}'$.

Answer. Since the method for solving this problem is the same as the previous problem, I write only the answer.

$$\begin{aligned} \mathcal{S}' &= \{(x, y) \in \mathbb{R}^2 \mid 1 \geq 0, 1 - x^2 \geq 0, 1 - y^2 \geq 0, 1 - (x+y)^2 \geq 0, \\ &\qquad\qquad 1 + 2xy(x+y) - (x+y)^2 - x^2 - y^2 \geq 0\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid 1 - x^2 \geq 0, 1 - y^2 \geq 0, 1 + 2xy(x+y) - (x+y)^2 - x^2 - y^2 \geq 0\}. \end{aligned}$$

The boundary $h(u, v) = 0$ of dual convex is
$h(u, v) = -8u^6 + (24v + 24)u^5 + (-21v^2 - 60v - 12)u^4 + (2v^3 + 24v^2 + 24v - 24)u^3 + (-21v^4 + 24v^3 - 36v^2 + 36v + 24)u^2 + (24v^5 - 60v^4 + 24v^3 + 36v^2 - 24v)u - 8v^6 + 24v^5 - 12v^4 - 24v^3 + 24v^2 - 4.$

These are shown in the following figures. The function $x + y$ is maximum at the point $(1/2, 1/2) \in \mathcal{S}'$ and its value is 1.

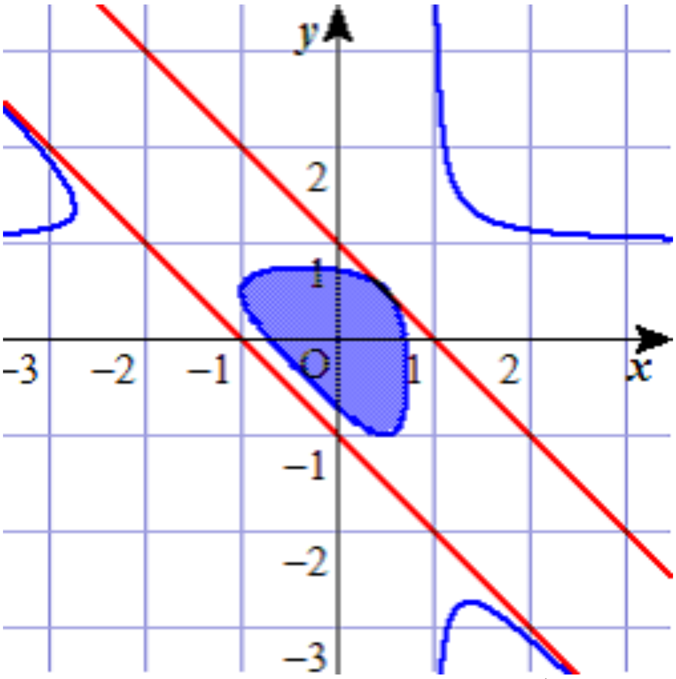


Figure 3: The spectrahedron $\mathcal{S}'$

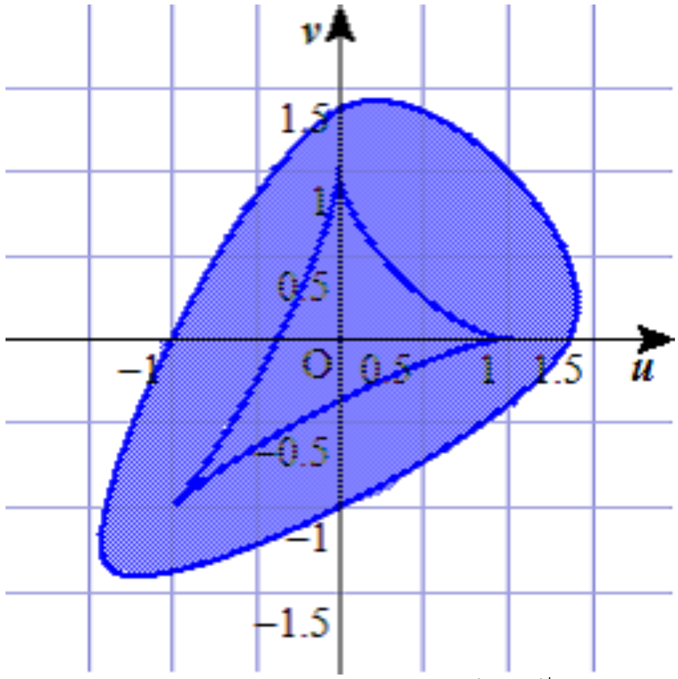


Figure 4: The dual convex of $\mathcal{S}'$: $\mathcal{S}'^*$

3. Fix the four points $(0,0),(0,1),(1,0),(1,2)$ in the plane $\mathbb{R}^2$. Compute their Fermat-Weber point. Also, compute the irreducible polynomial that vanishes on the 4-ellipse with these foci and arbitrary radius $d > 0$.

   Answer. The Fermat-Weber point is the point from where the sum of all Euclidean distances to given points is minimum. In this problem, it is $(1/3, 2/3)$ which is the diagonal intersection point of quadrangle with these 4 points as the vertex.

   The point (x, y) which vanishes on the 4-ellipse with these foci and arbitrary radius $d > 0$ satisfies the following equations :

$$\sum_{i=1}^{4} \sqrt{(x-x_i)^2 + (y-y_i)^2} = d.$$

   Therefore, we can obtain the irreducible polynomial as the only generator of the following ideal.

$$\langle t_1 + t_2 + t_3 + t_4 - d, t_i^2 - \{(x-x_i)^2 + (y-y_i)^2\} \ (i = 1, 2, 3, 4)\rangle \cap \mathbb{R}[x, y, d]$$

   Since this generator is very huge, I do not actually write it down. Its degrees with respect to x, y, d are $10, 10, 16$ respectively. $\square$

4. Let $f(x, y)$ be the polynomial of degree ten computed in the previous exercise, where the radius is now fixed at $d = 17$. Determine all singular points of the complex algebraic curve $\{ (x, y) \in \mathbb{C}^2 : f(x, y) = 0 \}$.

   Answer by Kenta Nishiyama (Kobe Univ.)
   Let us consider the singular locus of $f(x, y)$:

$$\{(x, y) \in \mathbb{C}^2 \mid f(x, y) = f_x(x, y) = f_y(x, y) = 0\}.$$

By computing a primary decomposition of the ideal $\langle f, f_x, f_y\rangle$, we get the following 11 primary components.

$$\langle f, f_x, f_y\rangle = \bigcap_{i=1}^{11} Q_i$$

$$\begin{aligned}
&Q_1 = \langle y-1, 17x-144\rangle, \qquad Q_2 = \langle y-1, 17x+144\rangle,\\
&Q_3 = \langle 2y-1, 289x^2-578x-20231\rangle,\\
&Q_4 = \langle 569y^2-852y-20164, x-y\rangle, \qquad Q_5 = \langle 577y^2-864y-20160, x+y-2\rangle,\\
&Q_6 = \langle 1152y^2-3456y-80351, 2x-1\rangle, \quad Q_7 = \langle 5744y^2-10344y-78055, 2x+4y-5\rangle,\\
&Q_8 = \langle 17y^4-2908y^3+954430y^2-108902572y+3954230201,\\
&\qquad 331776x+17y^3-443y^2+807251y-63680681\rangle,\\
&Q_9 = \langle 41y^4+1556y^3-221524y^2+36102416y+3882021136,\\
&\qquad 336400x-41y^3+4348y^2+15912y-63791944\rangle,\\
&Q_{10} = \langle 65y^4-7640y^3+951016y^2-107793680y+3952590272,\\
&\qquad 169362x-260y^3+11515y^2-2939420y+234068620\rangle,\\
&Q_{11} = \langle 113y^4+28912y^3+3179648y^2+171796144y+3670735216,\\
&\qquad 164738x+452y^3+83443y^2+6917424y+223641480\rangle.
\end{aligned}$$

We note that all primary components are prime ideals. (i.e., the multiplicity of all roots is one.) Hence we can decompose original system of equations into 11 simpler ones. Since the first polynomials of all components are univariate polynomials in y with degree is 4 or less, we can determine y by the well known solution formula. By substituting this y for the second polynomial, x is also determined. We obtain 28 singular points in all by this method. If we want to know only the number of singular points, it is enough to count the number of standard monomials of Gröbner basis with respect to any term-order. □

5. Fix the five points $(0,0), (0,1), (1,0), (1,2), (2,1)$ in the plane $\mathbb{R}^2$. Compute their Fermat-Weber point (x^*, y^*). Show that x^* and y^* are algebraic numbers over $\mathbb{Q}$, and determine their minimal polynomials.

Answer. By symmetry of these points for the line $y = x$, their Fermat-Weber point is on this line. We consider the minimum point of the following function

$$f(x) = \sqrt{2x^2} + 2\sqrt{x^2+(x-1)^2} + 2\sqrt{(x-1)^2+(x-2)^2}.$$

The first and second order derivatives of this are

$$f'(x) = \sqrt{2} + \frac{2(x+(x-1))}{\sqrt{x^2+(x-1)^2}} + \frac{2((x-1)+(x-2))}{\sqrt{(x-1)^2+(x-2)^2}},$$

$$f''(x) = \frac{2}{(x^2+(x-1)^2)\sqrt{x^2+(x-1)^2}} + \frac{2}{((x-1)^2+(x-2)^2)\sqrt{(x-1)^2+(x-2)^2}} > 0.$$

Since $f'(x)$ is strictly increasing function and $f'(1/2) < 0, f'(1) > 0$, $f(x)$ has the only extremum that is in the open interval $(1/2, 1)$ and it is minimal and minimum. If we want a numerical solution, we only have to search for this interval by bisection method.

Let us compute a point x^* which satisfies $f'(x^*) = 0$ algebraically. By squaring both sides twice to eliminate root sign and taking its numerator, we obtain the following polynomial.

$$P(x) := 240x^8 - 1920x^7 + 6656x^6 - 13056x^5 + 15928x^4 - 12512x^3 + 6208x^2 - 1728x + 191$$

$P(x)$ is irreducible and satisfies $P(x^*) = 0$. Hence it is minimal polynomial of x^*. $P(x)$ has 4 real roots, but only the root

$$x^* = \frac{1}{60}\left(60 - \left(3600 + 120\left(-28 + \sqrt{\gamma} - \sqrt{\frac{2752 - 255\alpha - 25\alpha^2 - 255\beta - 25\beta^2 + 1936\sqrt{\gamma}}{\gamma}}\right)\right)^{\frac{1}{2}}\right)$$
$$= 0.6877772297619\cdots$$

is in $(1/2, 1)$, where $\alpha = \sqrt[3]{3528 - 264\sqrt{177}}$, $\beta = \sqrt[3]{3528 + 264\sqrt{177}}$, $\gamma = 5(\alpha + \beta) - 51$. Therefore (x^*, x^*) is the Fermat-Weber point. □

6. Let $\mathcal{T}$ be a spectrahedron obtained by intersecting the cone of positive semidefinite 6×6-matrices with a generic affine-linear space of dimension 6. What are the possible ranks of extreme points on $\mathcal{T}$?

   Answer. For the Pataki inequalities :

$$\binom{r+1}{2} \leq \binom{n+1}{2} - m, \quad \binom{n-r+1}{2} \leq m,$$

   putting $n = m = 6$, we obtain $r = 3$ or 4 or 5. □

7. Consider the unit circle C and the unit square S in the plane $\mathbb{R}^2$. Draw their Minkowski sum $C + S$. Prove that $C + S$ is not spectrahedron. Can you express $C + S$ as the projection of a 3-dimensional spectrahedron?

   Answer. The Minkowski sum $C + S$ is shown in the following figures.

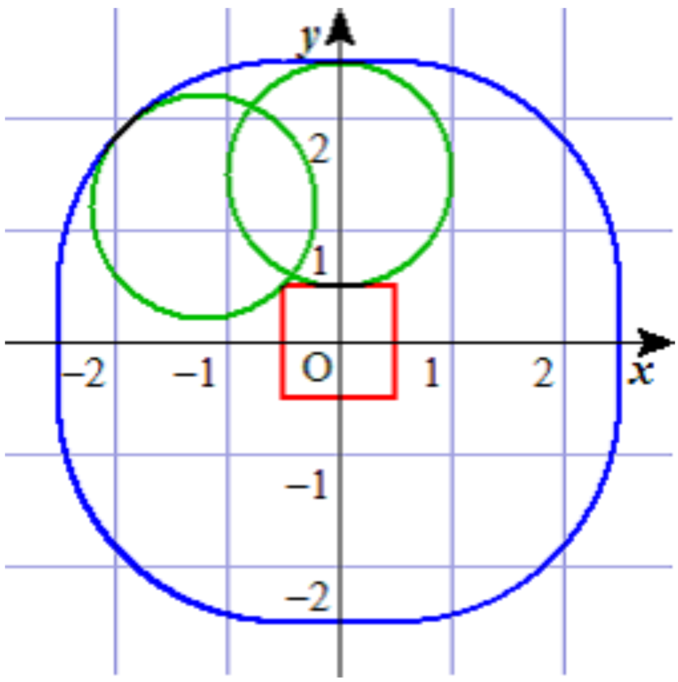


Figure 5: Minkowski sum $C + S$

We recall that a spectrahedron is written as a set

$$\{x \in \mathbb{R}^n \mid f_1(x) \geq 0, \ldots, f_k(x) \geq 0\}$$

by using polynomials $f_i(x)$. The Zariski closure of a part of quarter circle of $C+S$ is a complete circle interpolated to the remaining 3/4 part. It is a proper subset of $C+S$. Hence $C+S$ is not a spectrahedron.

Moreover $C+S$ can be expressed as the projection of a 4-dimensional spectrahedron by direct product of two positive semidefinite matrices of C and S. However, it seems not possible to express is as a projection of a 3-dimensional spectrahedron. $\square$

8. How many straight lines (over $\mathbb{C}$) lie on Cayley's cubic surface ?

Answer. The Cayley's cubic is a cubic surface in 3-dimensional projective space which has 4 conical points. It is given by the following equation.

$$x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_1 + x_4x_1x_2 = 0$$

It contains 9 straight lines :

(a) $x_i = x_j = 0$ (6 lines)

(b) $x_i + x_j = x_k + x_l = 0$ (3 lines)

In general, a smooth cubic surface contains 27 straight lines. In this case, multiplicity of straight lines of (a) is 4, and that of (b) is 1. Therefore if we count with multiplicity, it includes $6 \times 4 + 3 \times 1 = 27$ straight lines. $\square$

9. Let $\mathcal{U}$ be any spectrahedron obtained by intersecting the cone of positive semidefinite 4×4-matrices with an affine-linear space of dimension 3. Can you draw $\mathcal{U}$? How many matrices of rank 2 can $\mathcal{U}$ contain?

Answer. Sorry, it is under some more investigation....

10. Maximize the determinant over the spectrahedron $\mathcal{S}$ and $\mathcal{S}'$ in Ex. 1 and 2.

Answer by Yasuhide NUMATA (Univ. of Tokyo)
1. By symmetry, the maximum point is on the line $x = 0$. Hence the determinant

$$\det \begin{pmatrix} 1 & 0 & y \\ 0 & 1 & 0 \\ y & 0 & 1 \end{pmatrix} = 1 - y^2$$

is the maximum at the point $(x, y) = (0, 0) \in \mathcal{S}$.

2. The determinant

$$\det \begin{pmatrix} 1 & x & x+y \\ x & 1 & y \\ x+y & y & 1 \end{pmatrix} = 1 + xy(x+y) + xy(x+y) - (x+y)^2 - x^2 - y^2$$

is fixed by mapping $(x, y) \mapsto (y, x)$. Mapping $(x, y) \mapsto (-x, -y)$ also fixes the determinant. Hence the maximum point of the determinant is fixed by them. Such a point is only $(0, 0) \in \mathcal{S}'$, so, the determinant is the maximum at the point. $\square$

11. Fix the Gaussian graphical model defined by the 4-cycle. Give an example of a sample covariance matrix for which the MLE does not exist. What is the maximal rank of a psd 4×4-matrix with this property?

Answer by Tomonari SEI (Univ. of Tokyo)
The 4-cycle model is defined by

$$\begin{aligned}
\mathcal{L}_{\succ 0}^{-1} &= \left\{\Sigma \in \mathbb{S}_{\succ 0}^{4} \mid \Sigma^{-1} \in \mathcal{L}\right\}, \\
\mathcal{L} &= \left\{K = (K_{ij}) \in \mathbb{R}^{4\times 4} \mid K_{ij} = 0 \text{ if } \{i,j\} \notin E\right\}, \\
E &= \{\{1,2\},\{2,3\},\{3,4\},\{4,1\}\}.
\end{aligned}$$

In general, a sample covariance matrix S has the (unique) MLE if and only if the fiber

$$\text{fiber}_{\mathcal{L}}(S) = \{\Sigma \in \mathbb{S}_{\succ 0}^{m} : \langle \Sigma, K\rangle = \langle S, K\rangle \text{ for } \forall K \in \mathcal{L}\}$$

is non-empty (Barndorff-Nielsen 1978, Theorem 9.13). For the 4-cycle, the following sample covariance matrix has the rank 3 and satisfies $\text{fiber}_{\mathcal{L}}(S) = \emptyset$:[1]

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1}$$

Indeed, if a psd matrix Σ satisfies $\langle \Sigma, K\rangle = \langle S, K\rangle$ ($\forall K \in \mathcal{L}$), then Σ is not positive definite because $\Sigma_{44} = 0$. On the other hand, if S has the rank 4, then $\text{fiber}_{\mathcal{L}}(S)$ is non-empty because $S \in \text{fiber}_{\mathrm{L}}(S)$. Hence, the maximal rank of a psd matrix S with $\text{fiber}_{\mathcal{L}}(S) = \emptyset$ is 3. □

Remark. The equation (1) has the empty fiber for any other graphical model. Therefore no information of the 4-cycle model was used. We modify the problem as follows. We first define a terminology. For a given S, *the effective size* of S is defined to be the maximum number n such that any $n \times n$-principal submatrix of S has the rank n. Then we consider the problem

For the 4-cycle model, determine the maximal effective size of S with $\text{fiber}_{\mathcal{L}}(S) = \emptyset$.

The solution of this problem was given by Buhl (1993). For the m-cycle, if the effective size of S is greater than or equal to 3, then $\text{fiber}_{\mathcal{L}}(S) \neq \emptyset$. On the other hand, there exists a S with the effective size 2 such that $\text{fiber}_{\mathrm{L}}(S) = \emptyset$. We prove this fact for 4-cycle.

First, we prove that the fiber is non-empty if the effective size of S is greater than or equal to 3. Any edge of the 4-cycle graph is a subset of at least one of the two subsets $\{1,2,3\}$ or $\{1,3,4\}$. Therefore if a positive definite matrix Σ satisfies $\Sigma(1,2,3) = S(1,2,3)$ and $\Sigma(1,3,4) = S(1,3,4)$, then Σ is an element of $\text{fiber}_{\mathcal{L}}(S)$. Here $\Sigma(1,2,3)$ denotes the $(1,2,3)$-principal submatrix of Σ. Such a Σ can be constructed explicitly. Since the effective size of S is greater than or equal to 3, $S(1,3,4)$ is positive definite. Therefore it is possible to write

$$S(1,3,4) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ v_1 & v_3 & 1 \end{pmatrix} \begin{pmatrix} S(1,3) & 0 \\ 0 & \tilde{S}(4) \end{pmatrix} \begin{pmatrix} 1 & 0 & v_1 \\ 0 & 1 & v_3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then if we set

$$\Sigma := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ v_1 & 0 & v_3 & 1 \end{pmatrix} \begin{pmatrix} S(1,2,3) & 0 \\ 0 & \tilde{S}(4) \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & v_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & v_3 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

[1] This is pointed out by Dr. Kei Kobayashi.

Σ is positive definite and satisfies $\Sigma(1,2,3) = S(1,2,3)$ and $\Sigma(1,3,4) = S(1,3,4)$. This means $\mathrm{fiber}_{\mathcal{L}}(S)$. Note that $\Sigma_{24} \neq S_{24}$ in general.

Next we construct an example of S with the effective size 2 and the empty fiber. For three distinct numbers $\alpha, \beta, \gamma \in (0, \pi)$, we put

$$X = \sqrt{2} \begin{pmatrix} 1 & \cos\alpha & \cos\beta & \cos\gamma \\ 0 & \sin\alpha & \sin\beta & \sin\gamma \end{pmatrix}.$$

Then any two columns of X are linearly independent. Therefore

$$S := \frac{1}{2} X^\top X = \begin{pmatrix} 1 & \cos\alpha & \cos\beta & \cos\gamma \\ \cos\alpha & 1 & \cos(\beta-\alpha) & \cos(\gamma-\alpha) \\ \cos\beta & \cos(\beta-\alpha) & 1 & \cos(\gamma-\beta) \\ \cos\gamma & \cos(\gamma-\alpha) & \cos(\gamma-\beta) & 1 \end{pmatrix}$$

has the effective size 2. Now assume $\mathrm{fiber}_{\mathcal{L}}(S) \neq \emptyset$. Then there must exist x and y such that

$$\Sigma = \begin{pmatrix} 1 & \cos\alpha & x & \cos\gamma \\ \cos\alpha & 1 & \cos(\beta-\alpha) & y \\ x & \cos(\beta-\alpha) & 1 & \cos(\gamma-\beta) \\ \cos\gamma & y & \cos(\gamma-\beta) & 1 \end{pmatrix} \succ 0$$

The condition $\det(\Sigma(1,2,3)) > 0$ implies

$$\begin{aligned} \det(\Sigma(1,2,3)) &= 1 + 2\cos\alpha\cos(\beta-\alpha)x - \cos^2\alpha - \cos^2(\beta-\alpha) - x^2 \\ &= -(x - \cos\alpha\cos(\beta-\alpha))^2 + \sin^2\alpha\sin^2(\beta-\alpha) \\ &= -(x-\cos\beta)(x - \cos(2\alpha-\beta)) > 0. \end{aligned}$$

This is equivalent to $x \in (\min\{\cos\beta, \cos(2\alpha-\beta)\}, \max\{\cos\beta, \cos(2\alpha-\beta)\})$. Similarly, the condition $\det(\Sigma(1,3,4)) > 0$ implies

$$-(x-\cos\beta)(x-\cos(2\gamma-\beta)) > 0.$$

This is equivalent to $x \in (\min\{\cos\beta, \cos(2\gamma-\beta)\}, \max\{\cos\beta, \cos(2\gamma-\beta)\})$. However, if $\alpha < \beta < \gamma$ or $\gamma < \beta < \alpha$, then there is no x satisfying the above two inequalities. Hence $\mathrm{fiber}_{\mathcal{L}}(S) = \emptyset$.

Buhl (1993) proved the following more general result. Let S be the sample covariance matrix $S = (1/n)X^\top X$ of n samples of m-dimensional standard normal random vectors $X = (X_{ti})_{1\le t\le n, 1\le i\le m}$. Then S has the rank $\min(n,m)$ and the effective size $\min(n,m)$ almost surely.

Corollary 3.3 of Buhl (1993) For an undirected graph (V,E), let (V,E^+) be a decomposable graph that includes (V,E). Let the degree of maximum clique of E (resp. E^+) be q (resp. q^+). Define the graphical model corresponding to (V,E) by $\mathcal{L} = \{K \mid K_{ij} = 0 \text{ if } \{i,j\} \notin E\}$. Then $\mathrm{fiber}_{\mathcal{L}}(S)$ is almost surely non-empty if $n \ge q^+$, and almost surely empty if $n < q$. □

Corollary 4.2 of Buhl (1993) Let $m \ge 4$ and $n \ge 1$. Consider the m-cycle model. Then

$$\mathrm{P}[\mathrm{fiber}_{\mathcal{L}}(S) = \emptyset] = \begin{cases} 0 & \text{if } n \ge 3\ (= q^+), \\ 2/(m-1)! & \text{if } n = 2, \\ 1 & \text{if } n = 1\ (= q-1). \end{cases}$$

□

References

[1] Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*, John Wiley & Sons.

[2] Buhl, S. L. (1993). On the existence of maximum likelihood estimators for graphical Gaussian models, *Scand. J. Statist.*, **20**, 263–270.

12. Let $\mathcal{L}$ be the row space of the matrix

$$A = \begin{pmatrix} 4 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix}$$

The variety $\mathcal{L}^{-1}$ is a curve in $\mathbb{P}^4$. Find its prime ideal $P_{\mathcal{L}}$ and its degree. Explain the geometry and statistical meaning of MLE for this model.

Answer by Kenta Nishiyama (Kobe Univ.)
For the matrix A, let us consider the concentration matrix

$$K = \begin{pmatrix} 4\lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 3\lambda_1+\lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 2\lambda_1+2\lambda_2 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1+3\lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 4\lambda_2 \end{pmatrix}.$$

The variety $\mathcal{L}^{-1}$ is a curve in $\mathbb{P}^4$ parametrized by the inverse matrix which we write as $K^{-1} = \mathrm{diag}(x_1, x_2, x_3, x_4, x_5)$. K^{-1} is a diagonal matrix and its (i,i)-elements are

$$K_{ii}^{-1} = \frac{\prod_{j\neq i} K_{jj}}{\det(K)}.$$

Hence the prime ideal $P_{\mathcal{L}}$ is computed as the following elimination ideal.

$$\begin{aligned} P_{\mathcal{L}} &= \langle x_i - \prod_{j\neq i} K_{jj}\ (i=1,\dots,5), 1-\det(K)\rangle \cap \mathbb{R}[x_1,x_2,x_3,x_4,x_5] \\ &= \langle x_2x_3 - 3x_2x_5 + 2x_3x_5, 2x_2x_4 - 3x_2x_5 + x_4x_5, x_3x_4 - 2x_3x_5 + x_4x_5, \\ &\qquad 3x_1x_4 - 4x_1x_5 + x_4x_5, x_1x_2 - 4x_1x_5 + 3x_2x_5, x_1x_3 - 2x_1x_5 + x_3x_5\rangle \end{aligned}$$

The generator $1-\det(K)$ is a non-zero condition for $\det(K)$. However, since $\prod_{j\neq i} K_{jj}$ are homogeneous, we can obtain the same prime ideal from the ideal except this condition. Furthermore the Hilbert polynomial of $P_{\mathcal{L}}$ is

$$HP_{P_{\mathcal{L}}}(t) = 2t^2 + 3t + 1.$$

A head term $\frac{D}{d!}t^d$ of the Hilbert polynomial is expressed by using a degree D of a projection variety. (refer to [1, Chap 9], [2].) In this case, the head term is $\frac{4}{2!}t^2$. Therefore degree of $\mathcal{L}^{-1}$ is 4. □

In statistical terminology, the reciprocal of the variance is called the precision. This problem corresponds to a linear regression model for precisions.

References

[1] D. Cox, J. Little and D. O'Shea, *Ideals,Varieties, and Algorithms*, Springer-Verlag, New York, 1992.

[2] D. Cox, J. Little and D. O'Shea, *Using Algebraic Geometry*, Springer-Verlag, New York, 1998.

13. Let $\mathcal{L}$ be the space of Hankel matrices of format $m \times m$ and let $\mathcal{C}_{\mathcal{L}}$ be the spectrahedral cone consisting of all positive definite Hankel matrices. Determine the dual cone $\mathcal{K}_{\mathcal{L}}$ and its boundary hypersurface $H_{\mathcal{L}}$. What is the projective variety $\mathcal{L}^{-1}$? What does MLE mean for this model?

Answer by Akimichi Takemura (Univ. of Tokyo)
We give a sketch of a solution to the dual cone $\mathcal{K}_{\mathcal{L}}$ and its boundary $H_{\mathcal{L}}$. Projective variety $\mathcal{L}^{-1}$ can be described by Cramer's formula, but I am not sure about their algebraic properties. We also give a statistical model of multivariate normal distribution with Hankel type covariance matrix. However the model is not very natural and it is somewhat hard to discuss interpretations of the maximum likelihood estimation. I am not sure of the model whose concentration matrix is of Hankel type.

Hankel matrix is a symmetric matrix with equal "anti-diagonal" elements. For $(k+1) \times (k+1)$ it is of the following form.

$$A = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \cdots & \gamma_k \\ \gamma_1 & \gamma_2 & \gamma_3 & & & \gamma_{k+1} \\ \gamma_2 & \gamma_3 & & & & \vdots \\ \gamma_3 & & & & & \\ \vdots & & & & & \gamma_{2k-1} \\ \gamma_k & \gamma_{k+1} & \cdots & & \gamma_{2k-1} & \gamma_{2k} \end{pmatrix} \tag{2}$$

In this problem we consider non-negative definite A. We refer to [1] for Hankel matrices. In particular the equivalence of (iv) and (v) in Theorem 3.9 (p.619) of [1] seems to be of basic importance and I explain the equivalence here.

For a real s,

$$A = \begin{pmatrix} 1 \\ s \\ s^2 \\ \vdots \\ s^k \end{pmatrix} (1, s, , s^2, \ldots, s^k) = \begin{pmatrix} 1 & s & s^2 & \ldots & s^k \\ s & s^2 & & & s^{k+1} \\ s^2 & & & & \vdots \\ \vdots & & & & s^{2k-1} \\ s^k & s^{k+1} & \ldots & s^{2k-1} & s^{2k} \end{pmatrix} \tag{3}$$

is of rank 1 non-negative definite and of Hankel type. Let $A(s) = (1, s, \ldots, s_k)^T (1, s, \ldots, s^k)$ denote this matrix. Let $s_1, \ldots, s_r$ ($r \leq k+1$) be distinct real numbers and let $p_1, \ldots, p_r$ be positive. In view of Vandermonde determinant, $(1, s_i, \ldots, s_i^k)$, $i = 1, \ldots, r$ are linearly independent and

$$A = p_1 A(s_1) + \cdots + p_r A(s_r) \tag{4}$$

is a non-negative definite Hankel matrix of rank r. The important fact is that the converse is also true by (iv) and (v) of Theorem 3.9 (p.619) of [1] [2]. Namely, a $(k+1) \times (k+1)$ Hankel

[2] As I understood [1]

matrix A is of rank r and non-negative definite if and only if A is written as (4) with distinct real values $s_1, \dots, s_r$ and positive values $p_1, \dots, p_r$.

(4) corresponds to a multivariate normal distribution in a natural way, if A is regarded as a covariance matrix[3]. Let $Z_1, \dots, Z_r$ be independent standard normal random variables. Then the random vector (with mean 0)

$$Z_1\sqrt{p_1}(1, s_1, \dots, s_1^k) + \cdots + Z_r\sqrt{p_r}(1, s_r, \dots, s_r^k)$$

has the covariance matrix A in (4). However this random vector does not appear often in statistical literature.

In any case, the important fact is that non-negative definite Hankel matrices are non-negative combinations of matrices of the form $A(s)$. Therefore the problem of obtaining its dual reduces to obtaining the dual cone to the "moment cone", which is generated by the "moment curve" $(1, s, \dots, s^k)$, $s \in \mathbb{R}$. This in turn reduces to obtaining conditions on the coefficients of a polynomial which is non-negative on the whole real line. Let us check this more closely.

By (8) of p.4 of the distributed material [2], the dual cone is written as

$$\mathcal{K}_{\mathcal{L}} = \{(t_0, \dots, t_{2k}) : t_0\gamma_0 + \cdots + t_{2k}\gamma_{2k} \ge 0, \quad \forall(\gamma_0, \dots, \gamma_{2k}) : (2) \text{ n.n.d.}\}. \tag{5}$$

Comparing (2) and (3), we see that $\gamma_0 = 1$, $\gamma_1 = s$, $\gamma_2 = s^2$, $\dots$, $\gamma_{2k} = s^{2k}$. Hence (5) is rewritten as follows:

$$\mathcal{K}_{\mathcal{L}} = \{(t_0, \dots, t_{2k}) : t_0 + t_1 s + t_2 s^2 + \cdots + t_{2k}s^{2k} \ge 0, \quad \forall s \in \mathbb{R}\}. \tag{6}$$

Therefore the interior of this cone is the cone with positive constant term and no real roots. I did not have time to check known results on this problem. A simple-minded approach for evaluating the boundary $H_{\mathcal{L}}$ of $\mathcal{K}_{\mathcal{L}}$ is as follows.

The boundary with $t_0 = 0$ or $t_{2k} = 0$ seem to trivial. Therefore assume $t_0 > 0$, $t_{2k} > 0$. Then we only need to consider the boundary of set of polynomials with no real roots. Polynomials on the boundary have multiple real roots. Therefore we can use the discriminant, which is the square of Vandermonde determinant. The discriminant is a polynomial in the coefficients $t_0, \dots, t_{2k}$. Therefore $H_{\mathcal{L}}$ is a subset of the set of zeros of the discriminant. However it should be noted that the set of zeros of the discriminant contains positive polynomials with multiple complex roots, such as $(x^2+1)^2$. Therefore the set intersects the interior of $\mathcal{K}_{\mathcal{L}}$[4]. This is unsatisfactory.

Now we briefly consider $\mathcal{L}^{-1}$, which is a variety consisting of inverses of matrices of $\mathcal{C}_{\mathcal{L}}$. $\Sigma \in \mathcal{L}^{-1}_{\succ 0}$ is equivalent to the condition that anti-diagonal elements of $K = \Sigma^{-1}$ are equal. By Cramer's formula it is equivalent to the conditions that anti-diagonal cofactors of Σ are equal. I am not sure about algebraic property of this condition. □

References

[1] Raúl E. Curto and Lawrence A. Fialkow. (1991). Recursiveness, positivity, and truncated moment problems. *Houston Journal of Mathematics*, **17**, 603–635.

[2] Bernd Sturmfels and Caroline Uhler. (2009). Multivariate Gaussians, Semidefinite Matrix Completion, and Convex Algebraic Geometry. `arXiv:0906.3529v1`

[3] Due to Satoshi Kuriki

[4] Due to Kei Kobayashi

14. The *trigonometric moment curve* in $\mathbb{R}^4$ is given by the parametrization

$$\theta \mapsto \big(\cos(\theta), \sin(\theta), \cos(2\theta), \sin(2\theta)\big).$$

Show that this curve is algebraic and compute a Gröbner basis for its prime ideal. Prove that the convex hull of this curve is a spectrahedron. Classify all faces (of all dimensions 1, 2 and 3) of this spectrahedron.

Answer by Tomonari SEI (Univ. of Tokyo)
Denote the curve by

$$C = \{(x,y,z,w) = (\cos(\theta), \sin(\theta), \cos(2\theta), \sin(2\theta)) \mid \theta \in [0, 2\pi)\}.$$

By the double-angle formula, the following ideal determines C:

$$I = \langle x^2 + y^2 - 1, z - (x^2 - y^2), w - 2xy\rangle.$$

We check I is prime. Since x and y (and hence z and w) are written in a rational function

$$x = \frac{1-t^2}{1+t^2} \quad \text{and} \quad y = \frac{2t}{1+t^2},$$

the variety C determined by I is irreducible[5]. We prove $I = I(C)$, which implies I is prime. Let f be a polynomial such that $f(x,y,z,w) = 0$ over C. By substituting $z = x^2 - y^2$, $w = 2xy$ and $y = \sqrt{1-x^2}$, we obtain a univariate equation

$$\sum_{j=0}^{J} a_j x^j + \sum_{j=0}^{J'} b_j x^j \sqrt{1-x^2} = 0, \quad -1 \le x \le 1.$$

Since $\sqrt{1-x^2}$ is not rational function, we have $a_j = b_j = 0$ for all j and therefore f belongs to I. Hence I is prime.

The Gröbner basis of I with respect to the graded reverse lexicographic order is, by computation on Risa/Asir,

$$G = \{1 - z^2 - w^2, -wx + (z+1)y, (-z+1)x - wy, -2y^2 - z + 1, -2xy + w, 2x^2 - z - 1\}.$$

Now we prove the convex hull of C is the spectrahedron

$$S = \left\{ (x,y,z,w) \in \mathbb{R}^4 \;\middle|\; \begin{pmatrix} 1 & x & y \\ x & (z+1)/2 & w/2 \\ y & w/2 & (-z+1)/2 \end{pmatrix} \succ 0 \right\} \tag{7}$$

A method how to obtain this formula will be described later. We first prove $C \subset S$, and hence $\mathrm{Conv}(C) \subset S$. In fact, if $z = x^2 - y^2$, $w = 2xy$ and $x^2 + y^2 = 1$, then

$$\begin{aligned}
\begin{pmatrix} 1 & x & y \\ x & (z+1)/2 & w/2 \\ y & w/2 & (-z+1)/2 \end{pmatrix} &= \begin{pmatrix} 1 & x & y \\ x & x^2 & xy \\ y & xy & y^2 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} \begin{pmatrix} 1 & x & y \end{pmatrix} \\
&\succeq 0.
\end{aligned}$$

[5] Cox, Little & O'Shea (1992). *Ideals, Varieties, and Algorithms*, Springer., §4-5, Prop.6.

Conversely, we prove $S \subset \mathrm{Conv}(C)$. Let (x, y, z, w) be an element of S. It is sufficient to prove that there exists a random vector (X, Y) with $X^2 + Y^2 = 1$ such that

$$\mathrm{E}[(1, X, Y)^\top (1, X, Y)] = \begin{pmatrix} 1 & x & y \\ x & (1+z)/2 & w/2 \\ y & w/2 & (1-z)/2 \end{pmatrix} \tag{8}$$

This problem is a variant of so-called moment problems. We first remark that the spectrahedron S is invariant under the orthogonal transformation on the $(2, 3)$-principal submatrix. Hence we can assume $y = 0$ in (8) without loss of generality. Since the right hand side of (8) is positive semi-definite, it is easily deduced that $x, z, w \in [-1, 1]$. If $x = \pm 1$, then we have $z = 1$ and $w = 0$, and therefore a (deterministic) random variable $(X, Y) \equiv (\pm 1, 0)$ satisfies (8). Now we assume $-1 < x < 1$. The Schur complement of (8) is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -x^2 + (1+z)/2 & w/2 \\ 0 & w/2 & (1-z)/2 \end{pmatrix}.$$

Since the $(2, 3) \times (2, 3)$-submatrix of this matrix should be positive semi-definite and its trace is $1 - x^2$, there exists a $\phi \in [0, \pi/2]$ and $\rho \in [-1, 1]$ such that

$$\begin{pmatrix} -x^2 + (1+z)/2 & w/2 \\ w/2 & (1-z)/2 \end{pmatrix} = \begin{pmatrix} (1-x^2)\cos^2\phi & \rho(1-x^2)\cos\phi\sin\phi \\ \rho(1-x^2)\cos\phi\sin\phi & (1-x^2)\sin^2\phi \end{pmatrix}.$$

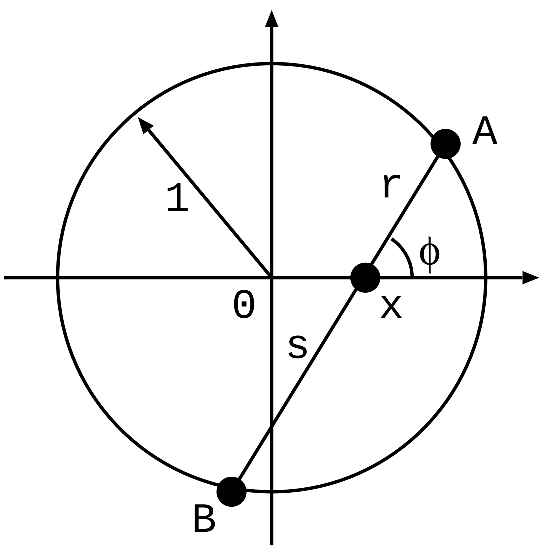
Then the problem is to find (X, Y) satisfying

$$\mathrm{E}[(1, X, Y)^\top (1, X, Y)] = \begin{pmatrix} 1 & x & 0 \\ x & x^2 + (1-x^2)\cos^2\phi & \rho(1-x^2)\cos\phi\sin\phi \\ 0 & \rho(1-x^2)\cos\phi\sin\phi & (1-x^2)\sin^2\phi \end{pmatrix}. \tag{9}$$

The parameter ρ is interpreted as the correlation coefficient between X and Y. We first consider the case $\rho = 1$. Define two points A and B in the unit circle by

$$\begin{aligned} A &= (x + r\cos\phi, r\sin\phi), \\ B &= (x - s\cos\phi, s\sin\phi) \end{aligned}$$

as is shown below.



Define a random variable (X_1, Y_1) by

$$(X_1, Y_1) = \begin{cases} A & \text{with probability } Cr^{-1}, \\ B & \text{with probability } Cs^{-1}, \end{cases}$$

with $C = 1/(r^{-1} + s^{-1})$. Then we can prove

$$\mathrm{E}[(1, X_1, Y_1)^\top (1, X_1, Y_1)] = \begin{pmatrix} 1 & x & 0 \\ x & x^2 + C(r+s)\cos^2\phi & C(r+s)\cos\phi\sin\phi \\ 0 & C(r+s)\cos\phi\sin\phi & C(r+s)\sin^2\phi \end{pmatrix}.$$

However, since $X_1^2 + Y_1^2 \equiv 1$, we obtain $C(r+s) = 1 - x^2$. Hence (X_1, Y_1) satisfies (9) with $\rho = 1$. Now we consider the general case $\rho \in [-1, 1]$. Let $\bar{A}$ and $\bar{B}$ be the reflected point of A and B with respect to the x-axis. Define a random variable (X, Y) by

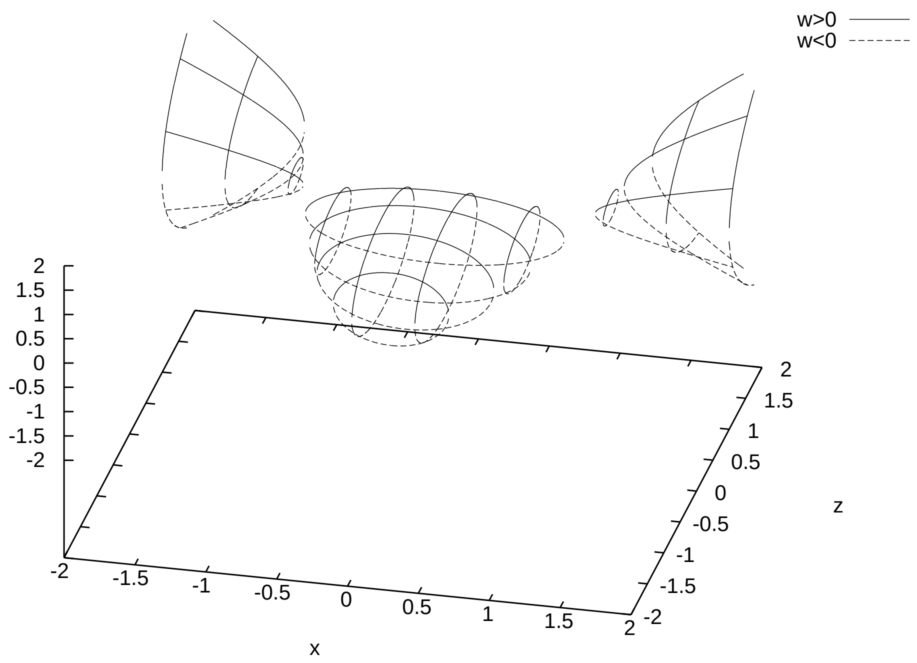
$$(X, Y) = \begin{cases} A & \text{with probability } Cr^{-1}(1+\rho)/2, \\ B & \text{with probability } Cs^{-1}(1+\rho)/2, \\ \bar{A} & \text{with probability } Cr^{-1}(1-\rho)/2, \\ \bar{B} & \text{with probability } Cs^{-1}(1-\rho)/2. \end{cases}$$

Then the moment of (X, Y) satisfies (9). We have proved $S \subset \mathrm{Conv}(C)$.

Finally, we classify the faces of S. The boundary of S is contained in a 3-dimensional variety $V = \{(x, y, z, w) \mid \det M(x, y, z, w) = 0\}$, where

$$M(x, y, z, w) := \begin{pmatrix} 1 & x & y \\ x & (1+z)/2 & w/2 \\ y & w/2 & (1-z)/2 \end{pmatrix}.$$

The variety V partitions $\mathbb{R}^4$ into three regions in that the three eigenvalues of $M(x, y, z, w)$ are $(+, +, +)$, $(+, +, -)$ and $(+, -, -)$, respectively. Here we remark that at least one of the eigenvalues of $M(x, y, z, w)$ must be positive because its trace is 2. Our boundary ∂S is identified that the eigenvalues of $M(x, y, z, w)$ are $(+, +, 0)$ or $(+, 0, 0)$. The following figure shows ∂S (the middle surface).



Next the 2-dimensional face is contained in the points $(x,y,z,w) \in V$ where the derivative of $\det M$ at (x,y,z,w) vanishes. This condition is described as

$$-x(1-z) + yw = xw - y(1+z) = -z - y^2 + x^2 = -w + 2xy = 0$$

with $\det M(x,y,z,w) = 0$ and $x^2+y^2 \leq 1$. The Gröbner bases of this system with $\det M(x,y,z,w) = 0$ is given by

$$G' = \{z^2+w^2-1, wx+(-z-1)y, (z-1)x+wy, 2y^2+z-1, -2xy+w, -2x^2+z+1\}.$$

This is equal to the Gröbner bases G for the original curve C. Hence the 2-dimensional face is empty and the 1-dimensional face is C. Since C is closed smooth curve, the 0-dimensional face is also empty.

Appendix: How to obtain (7) According to Gouveia et al., we construct a candidate of the convex hull of a real variety $\mathcal{V}_{\mathbb{R}}(I) \subset \mathbb{R}^n$ as follows. First, if a $y \in (\mathbb{R}[x]/I)^*$, the dual space of $\mathbb{R}[x]/I$, is given, then we define a bilinear form H_y as follows

$$\begin{array}{rccc} H_y: & (\mathbb{R}[x]/I) \times (\mathbb{R}[x]/I) & \to & \mathbb{R} \\ & (f+I, g+I) & \mapsto & y(fg+I) \end{array}$$

Here H_y is in general an infinite-dimensional matrix. Let $H_{y,t}$ be the truncated matrix of H_y up to t of the degree of $\mathbb{R}[x]/I$ (the degree of each class in $\mathbb{R}[x]/I$ is defined by the minimum degree of representatives of the class). Then we obtain

$$\mathrm{Conv}(\mathcal{V}_{\mathbb{R}}(I)) \subset \{(y(x_1+I)), \ldots, y(x_n+I)) \mid H_{y,t} \succeq 0,\ y(1+I) = 1\}$$

by Definition 2.3 and Theorem 2.4 of Gouveia et al. (2008). Now denote the coordinate of $\mathbb{R}^4$ by x_1, x_2, x_3, x_4 and consider the ideal

$$I = \langle 1-x_3^2-x_4^2, -x_4x_1+(x_3+1)x_2, (-x_3+1)x_1-x_4x_2, -2x_2^2-x_3+1, -2x_1x_2+x_4, 2x_1^2-x_3-1\rangle.$$

Select a $y \in (\mathbb{R}[x]/I)^*$ and denote $c_0 = y(1+I) = 1, c_i = y(x_i+I), c_{ij} = y(x_ix_j+I), \cdots$ and so on. Then the matrix $H_{y,1}$ is written as, with respect to the representative $1, x_1, x_2, x_3, x_4$,

$$H_{y,1} = \begin{pmatrix} 1 & c_1 & c_2 & c_3 & c_4 \\ c_1 & (c_3+1)/2 & c_4/2 & c_{13} & c_{23}+c_2 \\ c_2 & c_4/2 & (-c_3+1)/2 & c_{23} & -c_{13}+c_1 \\ c_3 & c_{13} & c_{23} & c_{33} & c_{34} \\ c_4 & c_{23}+c_2 & -c_{13}+c_1 & c_{34} & -c_{33}+1 \end{pmatrix}.$$

In particular, the condition where the first 3×3-submatrix is positive semi-definite is given by

$$\begin{pmatrix} 1 & c_1 & c_2 \\ c_1 & (c_3+1)/2 & c_4/2 \\ c_2 & c_4/2 & (-c_3+1)/2 \end{pmatrix} \geq 0.$$

This implies (7). If we use the first 5×5-submatrix instead of 3×3-submatrix, then we will not obtain any candidate of a spectrahedron because the latent variables c_{ij} are contained in the expression. □

References

[1] Gouveia, J. Parrilo, P. A. and Thomas, R. R. (2008). Theta bodies for polynomial ideals, arXiv:0809.3480v2