Chapter VIII Determinant maximization

§1 Convex spectral functions

goals: (a) Show that
$$F : S^n \to \mathbb{R}$$
$$X \mapsto \begin{cases} -(\det X)^{1/n} & \text{if } X \succeq 0 \\ \infty & \text{otherwise} \end{cases}$$
is a convex function

(b) Characterise all functions on S^n which are convex and depend only on the eigenvalues.

Def. 1 Function $F : S^n \to \mathbb{R} \cup \{\infty\}$ is called spectral if $F(X)$ only depends on the eigenvalues $\lambda_1, \dots, \lambda_n$ of X, i.e. $F(X) = F(AXA^T)$ holds for all $A \in O(n)$.

Remark A spectral function F defines a symmetric function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ by $F(X) = f(\lambda_1, \dots, \lambda_n)$, where f is invariant under permutation of coordinates.

Theorem 2 (Davis 1957)

$F : S^n \to \mathbb{R} \cup \{\infty\}$ is convex and spectral $\iff$ $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex and symmetric.

Some preparation is needed for the proof:

<u>Def. 3</u> (a) The <u>Schur-Horn orbitope</u> of $X \in S^n$ is

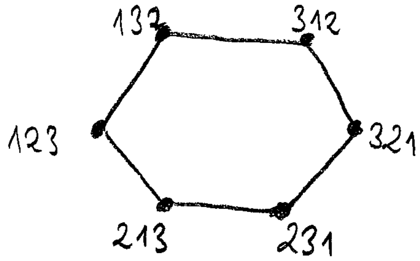
$$SH(X) = \operatorname{conv}\{AXA^T : A \in O(n)\} \subseteq S^n$$

(b) The permutahedron of $(x_1, \dots, x_n) \in \mathbb{R}^n$ is

$$\Pi(x_1, \dots, x_n) = \operatorname{conv}\{(x_{\sigma(1)}, \dots, x_{\sigma(n)}) : \sigma \in S_n\}.$$

<u>Example</u> • $\Pi(1,2,3)$ is a hexagon

• $SH(X)$ is generally not a polytope.



<u>Theorem 4</u> Let $X \in S^n$ be a symmetric matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Then,

$$\operatorname{diag}(SH(X)) = \Pi(\lambda_1, \dots, \lambda_n),$$

where $\operatorname{diag}: S^n \to \mathbb{R}^n$, $\operatorname{diag}(Y) = (Y_{11}, \dots, Y_{nn})$.

<u>Proof</u> <u>"$\supseteq$"</u>: The image $\operatorname{diag}(SH(X))$ is convex since diag is linear. We also have for $\sigma \in S_n$

$$(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)}) \in \operatorname{diag}(SH(X)) \qquad (\text{why?}).$$

Since $\Pi(\lambda_1, \dots, \lambda_n)$ is the smallest convex set containing these vectors, the inclusion follows.

<u>"$\subseteq$"</u>: Perform spectral decomposition of X:

$$X = \sum_{j=1}^{n} \lambda_j u_j u_j^T.$$

For $A \in \mathcal{O}(n)$ define $Y = AXA^T$. Then

$$Y_{ii} = e_i^T Y e_i = e_i^T \left(\sum_{j=1}^{n} \lambda_j A u_j u_j^T A^T \right) e_i$$

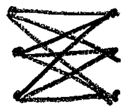
$$= \sum_{j=1}^{n} \lambda_j \underbrace{\left(e_i^T A u_j \right)^2}_{=: S_{ij}}$$

Claim: Matrix $(S_{ij})_{1 \le i,j \le n}$ is doubly stochastic.

Proof: – $S_{ij} \ge 0$: ✓

– row sum: $\sum_{j=1}^{n} S_{ij} = 1$: $\sum_{i=1}^{n} (e_i^T A u_j)^2 = \|A u_j\|^2 = 1$

– column sum: $\sum_{i=1}^{n} S_{ij} = 1$: $\Rightarrow$ same argument.

By the theorem of Birkhoff and von Neumann:
($\hat{=}$ description of the perfect matching polytope of the complete bipartite graph $K_{n,n}$ on $2n$ vertices and n^2 edges ):

$$\{ S \in \mathbb{R}^{n \times n} : S \text{ doubly stochastic} \} = \operatorname{conv} \{ P^\sigma : \sigma \in S_n \},$$

P^σ permutation matrix, $P^\sigma_{ij} = \begin{cases} 1 & \text{if } i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$

So there are coefficients $\alpha_\sigma \ge 0$, $\sum_{\sigma \in S_n} \alpha_\sigma = 1$ with $S = \sum_{\sigma \in S_n} \alpha_\sigma P^\sigma$.

Plugging in yields:

$$Y_{ii} = \sum_{j=1}^{n} \lambda_j \Big(\sum_{\sigma\in S_n} \alpha_\sigma P^\sigma\Big)_{ij}$$

$$= \sum_{\sigma\in S_n} \alpha_\sigma \sum_{j=1}^{n} \lambda_j P^\sigma_{ij}$$

$$= \sum_{\sigma\in S_n} \alpha_\sigma \lambda_{\sigma^{-1}(i)}.$$

So $(Y_{11}, \dots, Y_{nn}) \in \Pi(\lambda_1, \dots, \lambda_n)$. ⊠

Now we can prove Theorem 2

<u>Proof</u> (Thm. 2)

<u>„$\Rightarrow$"</u>: f symmetric: ✓

f convex: Let $x, y \in \mathbb{R}^n$, $\alpha \in [0,1]$. Define $X = \mathrm{Diag}(x_1, \dots, x_n)$, $Y = \mathrm{Diag}(y_1, \dots, y_n)$. Then

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &= F(\alpha X + (1-\alpha)Y) \\ &\le \alpha F(X) + (1-\alpha) F(Y) \\ &= \alpha f(x) + (1-\alpha) f(y). \end{aligned}$$

<u>„$\Leftarrow$"</u>: We shall show that

$$F(X) = \max_{A\in O(n)} f(\mathrm{diag}(AXA^T)) \qquad (*)$$

holds.

From this representation it follows that F is convex since it is a maximum of a set of convex functions

$$g_A : S^n \to \mathbb{R}, \quad g_A(X) = f(\operatorname{diag}(AXA^T)), \quad A \in O(n).$$

"$\leq$" in $(*)$: Spectral decomposition of $X = A^T \operatorname{Diag}(\lambda_1, \dots, \lambda_n) A$, with $A \in O(n)$. Then

$$F(X) = f(\lambda_1, \dots, \lambda_n) = f(\operatorname{diag}(AXA^T)).$$

"$\geq$" in $(*)$: $\max\limits_{A \in O(n)} f(\operatorname{diag}(AXA^T)) \leq \max\limits_{Y \in SH(X)} f(\operatorname{diag}(Y))$

$$\overset{\text{Thm. 4}}{=} \max_{x \in \Pi(\lambda_1, \dots, \lambda_n)} f(x)$$

$$\overset{\substack{f \text{ convex} \\ \text{and symmetric}}}{=} f(\lambda_1, \dots, \lambda_n)$$

Maximum of a convex function over convex polytope is attained at a vertex.

$$= F(X). \qquad \boxtimes$$

Corollary 5 $F(X) = \begin{cases} -(\det X)^{1/n} & \text{if } X \succeq 0 \\ \infty & \text{otherwise} \end{cases}$

is convex and spectral. (Minkowski's determinant inequality).

<u>Proof</u> to show: $f(x_1, \dots, x_n) = -\left(\prod_{i=1}^{n} x_i\right)^{1/n}$

is symmetric and convex.

symmetry: obvious

convexity: follows from the inequality between the arithmetic mean and the geometric mean (AM-GM inequality).

<u>Lemma 6</u> AM-GM inequality

(a) For $x_1, \dots, x_n \geq 0$:

$$\left(\prod_{i=1}^{n} x_i\right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^{n} x_i$$

with equality iff $x_1 = \dots = x_n$

(geometric interpretation: Volume of parallelepiped with side length $S = \sum_{i=1}^{n} x_i$ is maximal iff $x_1 = \dots = x_n = \frac{S}{n}$.)

(b) For $x_1, \dots, x_n > 0$, $y_1, \dots, y_n > 0$:

$$\left(\prod_{i=1}^{n} x_i\right)^{1/n} + \left(\prod_{i=1}^{n} y_i\right)^{1/n} \leq \left(\prod_{i=1}^{n} (x_i + y_i)\right)^{1/n}.$$

with equality iff the vectors $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ are linearly dependent.

<u>Proof</u> (a) George Pólya's (1887 - 1985) best mathematical dream: We have $1+y \le e^y$ for all $y \in \mathbb{R}$.

So $y \le e^{y-1}$ and $y^{1/n} \le \left(e^{y-1}\right)^{1/n} = e^{y/n - 1/n}$.

Define $y_i = \dfrac{x_i}{\frac{1}{n}\sum_{j=1}^{n} x_j}$. Then

$$\left(\prod_{i=1}^{n} y_i\right)^{1/n} \le \left(\prod_{i=1}^{n} e^{y_i - 1}\right)^{1/n} = e^{\frac{1}{n}\sum_{i=1}^{n} y_i - 1} = e^0 = 1.$$

So the inequality follows. Note that $1+y = e^y$ iff $y=0$.

(b)
$$\frac{\left(\prod_{i=1}^{n} x_i\right)^{1/n} + \left(\prod_{i=1}^{n} y_i\right)^{1/n}}{\left(\prod_{i=1}^{n} (x_i + y_i)\right)^{1/n}} = \left(\prod_{i=1}^{n} \frac{x_i}{x_i + y_i}\right)^{1/n} + \left(\prod_{i=1}^{n} \frac{y_i}{x_i + y_i}\right)^{1/n}$$

$$\overset{(a)}{\le} \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{x_i+y_i} + \frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{x_i+y_i}$$

$$= 1.$$

Equality holds iff (by (a)):

$$\frac{x_1}{x_1+y_1} = \frac{x_2}{x_2+y_2} = \dots = \frac{x_n}{x_n+y_n} \quad \text{and} \quad \frac{y_1}{x_1+y_1} = \dots = \frac{y_n}{x_n+y_n}.$$

Hence, $\dfrac{x_i}{x_j} = \dfrac{x_i+y_i}{x_j+y_j} = \dfrac{y_i}{y_j}$ for all $i \ne j$. So $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ are lin. dep.

$\boxtimes$

end of proof of Corollary 5:

We have for $x, y \in \mathbb{R}^n_{\geq 0}$ $\alpha \in [0,1]$ that $f(\alpha x + (1-\alpha) y)$ equals

$$-\left(\prod_{i=1}^{n} (\alpha x_i + (1-\alpha) y_i)\right)^{1/n} \overset{L.6(b)}{\leq} -\left(\prod_{i=1} \alpha x_i\right)^{1/n} - \left(\prod (1-\alpha) y_i\right)^{1/n}$$

$$= \alpha f(x) + (1-\alpha) f(y). \qquad \boxtimes$$

§2 MAXDET optimization

Theorem 1 The set

$$\mathcal{D}^n = \{ (X, s) \in \mathcal{S}^n \times \mathbb{R} : X \succeq 0,\ s \geq 0,\ (\det X)^{1/n} \geq s \}$$

is a proper convex cone.

Proof • $\mathcal{D}^n$ is a convex cone.

Let $\alpha \geq 0$, $(X, s), (Y, t) \in \mathcal{D}^n$ be given:

Then $(\alpha X, \alpha s) \in \mathcal{D}^n$ because $\alpha X \succeq 0$, $\alpha s \geq 0$ and $(\det \alpha X)^{1/n} = \alpha (\det X)^{1/n} \geq \alpha s$.

Then $(X+Y, s+t) \in \mathcal{D}^n$ because $X+Y \succeq 0$, $s+t \geq 0$ and $(\det (X+Y))^{1/n} \geq (\det X)^{1/n} + (\det Y)^{1/n} \geq s+t$.

Corollary 1.5.

- $\mathcal{D}^n$ is pointed:

  Suppose $(X, s), (-X, -s) \in \mathcal{D}^n$. We have $X = 0$ since S_+^m is pointed. We have $s = 0$ since $\mathbb{R}_{\geq 0}$ is pointed.

- $\operatorname{int} \mathcal{D}^n \neq \emptyset$:

  Consider an open neighborhood of $(I_n, 1)$.

- $\mathcal{D}^n$ is closed:

  Clear because det is continuous.

<u>Def. 2</u> MAXDET problem, primal standard form

$$
\begin{aligned}
p^* = \sup \ & \langle (C, c), (X, s) \rangle \\
& (X, s) \in \mathcal{D}^n \\
& \langle (A_j, a_j), (X, s) \rangle = b_j, \quad j \in [m].
\end{aligned}
$$

If $(C, c) = (0, 1)$ and $a_j = 0$, then the above problem simplifies

$$
\begin{aligned}
p^* = \sup \ & s \\
& X \succeq 0, \ (\det X)^{1/n} \geq s \\
& \langle A_j, X \rangle = b_j, \ j \in [m]
\end{aligned}
\quad \Longleftrightarrow \quad
\begin{aligned}
p^* = \sup \ & (\det X)^{1/n} \\
& X \succeq 0 \\
& \langle A_j, X \rangle = b_j, \\
& j \in [m].
\end{aligned}
$$

Lemma 3 Let $X \in S_+^n$ be a positive semidefinite matrix. Then
$$\mathrm{Tr}(X) - n(\det X)^{1/n} \geq 0, \qquad (*)$$
where equality holds iff X is a multiple of the identity matrix.

Proof We may assume that X is a diagonal matrix $X = \mathrm{Diag}(x_1, \dots, x_n)$. Then
$$(*) \iff \sum_{i=1}^{n} x_i - n\left(\prod_{i=1}^{n} x_i\right)^{1/n} \geq 0 \quad \text{(AM-GM inequality)}.$$
$\boxtimes$

Theorem 4 The dual cone of $\mathcal{D}^n$ is

$$(\mathcal{D}^n)^* = \{ (Y,t) \in \mathcal{S}^n \times \mathbb{R} : \ Y \succeq 0 , \ (\det Y)^{1/n} \geq -\frac{t}{n} \}.$$

Proof (Thm. 4)

"$\supseteq$": Let $(Y,t) \in S^n \times \mathbb{R}$ with $Y \succeq 0$ and $(\det Y)^{1/n} \geq -\frac{t}{n}$, and let $(X,s) \in \mathcal{D}^n$. Then

$$\langle (X,s), (Y,t) \rangle = \mathrm{Tr}(XY) + st.$$

Would like to apply Lemma 3 but XY is generally not even symmetric. Solution: Use Cholesky factorization: $X = LL^T$. Then

$$\mathrm{Tr}(XY) = \mathrm{Tr}(LL^TY) = \mathrm{Tr}(L^TYL) \text{ and } L^TYL \succeq 0.$$

So

$$\begin{aligned} \mathrm{Tr}(XY) + st &= \mathrm{Tr}(L^TYL) + st \\ &\overset{L.3}{\geq} n \det(L^TYL)^{1/n} + st \\ &= n \det(LL^T)^{1/n} (\det Y)^{1/n} + st \\ &= n (\det X)^{1/n} (\det Y)^{1/n} + st \\ &\geq n \cdot s \cdot \left(-\frac{t}{n}\right) + st \\ &\geq 0. \end{aligned}$$

"$\subseteq$": Consider $(Y,t) \in (\mathcal{D}^n)^*$. For $(X,0) \in \mathcal{D}^n$ we have

$$0 \leq \langle (X,0), (Y,t) \rangle = \langle X, Y \rangle + 0 \cdot t = \langle X, Y \rangle.$$

Since $(S^m_{\geq 0})^* = S^m_{\geq 0}$, it follows that $Y \geq 0$.

Now consider $(X, (\det X)^{1/n}) \in \mathcal{D}^n$ with $X > 0$. Then

$$0 \leq \mathrm{Tr}(XY) + (\det X)^{1/n}\, t.$$

Therefore,

$$-t \leq \frac{\mathrm{Tr}(XY)}{(\det X)^{1/n}}.$$

Minimise the function, depending on X, of the RHS.

<u>1st case</u>: Y not positive definite, only positive semidefinite

Then the infimum is zero: Let $u_1 \in \mathbb{R}^n$ be a unit vector s.t. $u_1^T Y u_1 = 0$. Complete u_1 to an ONB $u_1, \dots, u_n$ and define $X = u_1 u_1^T + \varepsilon \sum_{i=2}^n u_i u_i^T$ for $\varepsilon > 0$. Then,

$$0 \leq \frac{\mathrm{Tr}(XY)}{(\det X)^{1/n}} = \frac{u_1^T Y u_1 + \varepsilon \sum_{i=2}^n u_i^T Y u_i}{(\varepsilon^{n-1})^{1/n}}$$

$$\leq \frac{0 + \varepsilon (n-1)\, \lambda_{max}(Y)}{\varepsilon^{(n-1)/n}} = \varepsilon^{\frac{1}{n}} (n-1)\, \lambda_{max}(Y) \xrightarrow{\varepsilon \to 0} 0.$$

<u>2nd case</u>: Y is positive definite

Then the minimum is attained at $X = Y^{-1}$ and it is equal to $(\det Y)^{1/n} \cdot n$. This follows from the AM-GM inequality (Lemma 3) as follows:

By Lemma 3 (see also proof of "$\supseteq$"):

$$\mathrm{Tr}(XY) \geq n(\det X)^{1/n}(\det Y)^{1/n}$$

with equality iff XY is a multiple of the identity.

Hence,

$$-t \leq n(\det Y)^{1/n} \Rightarrow (\det Y)^{1/n} \geq -\frac{t}{n}. \qquad \boxtimes$$

<u>Def.</u> 5 MAXDET problem, primal standard form

$$p^* = \sup \langle (C,c), (X,s) \rangle$$
$$(X,s) \in \mathcal{D}^n$$
$$\langle (A_j, a_j), (X,s) \rangle = b_j, \quad j \in [m].$$

If $(C,c) = (0,1)$ and $a_j = 0$, then the above problem simplifies

$$p^* = \sup s \qquad \Longleftrightarrow \qquad p^* = \sup (\det X)^{1/n}$$
$$X \succeq 0, \ (\det X)^{1/n} \geq s \qquad\qquad X \succeq 0$$
$$\langle A_j, X \rangle = b_j, \ j \in [m] \qquad\qquad \langle A_j, X \rangle = b_j, \ j \in [m].$$

MAXDET problem, dual standard form

$$d^* = \inf \sum_{j=1}^m y_j b_j$$
$$\sum_{j=1}^m y_j (A_j, a_j) - (C,c) \in (\mathcal{D}^n)^*.$$

This problem simplifies, when setting $(C, c) = (0, 1)$, $a_j = 0$, as follows:

$$d^* = \inf \sum_{j=1}^{m} y_j b_j$$
$$\sum_{j=1}^{m} y_j A_j - C \succeq 0$$
$$\det\Big(\sum_{j=1}^{m} y_j A_j - C\Big)^{1/n} \geq \frac{1}{m}.$$

Of course, the duality theory developed in Chapter III.5 also holds here.

Especially useful for MAXDET: optimality condition.

<u>Theorem 6</u> Suppose $p^* = d^*$ and suppose (X, s) is feasible for the primal MAXDET problem and $(y_1, \dots, y_m)$ is feasible for the dual MAXDET problem. Then, (X, s), $(y_1, \dots, y_m)$ are both optimal if and only if the following three condition hold:

(i) $\cancel{\Big(\sum_{j=1}^{m} y_j A_j - C\Big)^{-1} = X}$ $\quad X\Big(\sum_{j=1}^{m} y_j A_j - C\Big) = \alpha I$ for some $\alpha > 0$

(ii) $s = (\det X)^{1/n}$

(iii) $-\dfrac{\sum y_j a_j - c}{n} = (\det Y)^{1/n}$.

<u>Proof</u> follows immediately from Chapter II and the equality condition in Lemma 3.

$$0 = \left\langle (X,s), \left(\sum_{j=1}^{m} y_j A_j - C, \sum_{j=1}^{m} y_j a_j - c\right)\right\rangle$$

$$= \mathrm{Tr}\left(X\left(\sum_{j=1}^{m} y_j A_j - C\right)\right) + s\left(\sum_{j=1}^{m} y_j a_j - c\right)$$

$$\geq n\,(\det X)^{1/n}\left(\det\left(\sum y_j A_j - C\right)\right)^{1/n} + s\left(\sum_{j=1}^{m} y_j a_j - c\right)$$

$$\geq n\ s\ \left(-\frac{\sum y_j a_j - c}{n}\right) + s\left(\sum y_j a_j - c\right)$$

$$= 0.$$

The equality case for the first inequality yields (i).
The equality case for the second inequality yields (ii) and (iii).

§ 3 Approximation of polytopes by ellipsoids

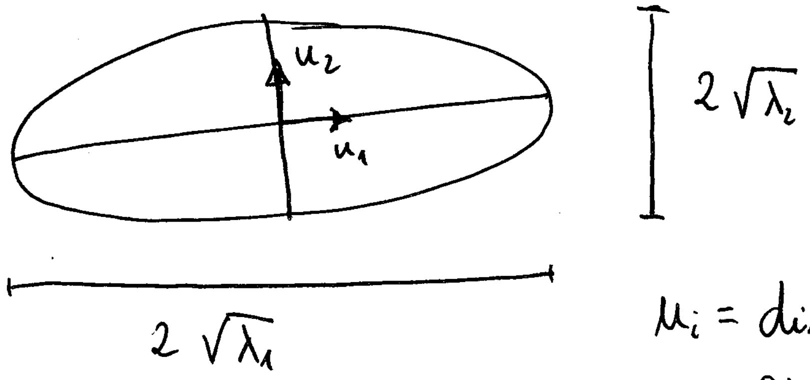
Recall:

Def. 1 Matrix $A \in S^n_{\succ 0}$, vector $x \in \mathbb{R}^n$ define the ellipsoid

$$\mathcal{E}(A, x) = \{ y \in \mathbb{R}^n : (y-x)^T A^{-1} (y-x) \leq 1 \}.$$

Ex.: $\mathcal{E}(r^2 I_n, 0) = r B_n$, where $B_n = \{ y \in \mathbb{R}^n : \|y\| \leq 1 \}$
n-dimensional unit ball.

Geometric properties can be read off from the spectral decomposition of $A = \sum_{i=1}^{n} \lambda_i u_i u_i^T$.



u_i = direction of axis.

$$\operatorname{vol} \mathcal{E}(A, x) = \sqrt{\det A} \; \underbrace{\operatorname{vol} B_n}_{\frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}}$$

Ellipsoid is an affine image of the unit ball:

$$\mathcal{E}(A^2, x) = \{ Ay + x : y \in B_n \}$$

Theorem 2 (inner approximation)

Let $\mathcal{E}(A^2, x)$ be an ellipsoid and let

$$P = \{ y \in \mathbb{R}^n : a_1^T y \leq b_1, \dots, a_m^T y \leq b_m \}$$

be a polytope. We have

$$\mathcal{E}(A^2, x) \subseteq P \iff \| A a_j \| \leq b_j - a_j^T x \,, \quad j \in [m]$$
$$\iff (A a_j, b_j - a_j^T x) \in \mathcal{L}^{n+1}.$$

Proof

$$\forall y \in \mathbb{R}^n : \|y\| \leq 1 : \quad a_j^T (Ay + x) \leq b_j \,, \quad j \in [m]$$

$$\iff \max_{\substack{y \in \mathbb{R}^n \\ \|y\| \leq 1}} a_j^T A y \leq b_j - a_j^T x \,, \quad j \in [m]$$

Cauchy-Schwarz gives the maximum:

$$\max_{\|y\| \leq 1} (A a_j)^T y = (A a_j)^T \frac{A a_j}{\|A a_j\|} = \| A a_j \| \,,$$

and the claim follows. $\boxtimes$

Goal: Find best inner approximation for P, i.e.

$\mathcal{E}(A^2, x)$ with $\mathcal{E}(A^2, x) \subseteq P$ and $\operatorname{vol} \mathcal{E}(A^2, x)$ maximal.

$$\operatorname{vol} \mathcal{E}(A^2, x) = \det A \operatorname{vol} B_n.$$

Conic programming formulation:

$$\max s$$
$$(A,s) \in \mathcal{D}^n, \quad x \in \mathbb{R}^n$$
$$(Aa_j, b_j - a_j^T x) \in \mathcal{L}^{n+1}, \quad j \in [m]$$

$$= \max s$$
$$(A,s) \in \mathcal{D}^n, \; x \in \mathbb{R}^n, \; (y_j, t_j) \in \mathcal{L}^{n+1}, \; j \in [m]$$
$$y_j = Aa_j, \quad t_j = b_j - a_j^T x, \quad j \in [m].$$

<u>Remarks</u>: (i) We work with the cone $\mathbb{R}^n$ here which is not pointed. This is more convenient and does not cause problems. In principle one could write $x = x^+ - x^-$, $x^+ \in \mathbb{R}^n_{\geq 0}$, $x^- \in \mathbb{R}^n_{\geq 0}$.

(ii) The conic program has a <u>unique</u> solution because the function $F(X) = -(\det X)^{1/n}$ is strictly convex on line segments $[X, Y]$ with $X \neq \alpha Y$. (follows from the equality case of the AM-GM inequality, Lemma 1.6 (b)). The optimal ellipsoid is uniquely determined.

<u>Def. 3</u> Let P be a polytope. The ellipsoid E with $E \subseteq P$ having largest volume is called <u>Loewner-John ellipsoid</u> $\mathcal{E}_{in}(P)$ of P.

Theorem 4 (outer approximation)

Let $\mathcal{E}(A,x)$ be an ellipsoid and let $P = \operatorname{conv}\{x_1, \dots, x_N\} \subseteq \mathbb{R}^n$ be a polytope. We have

$$P \subseteq \mathcal{E}(A,x) \iff \exists s \in \mathbb{R} : \begin{pmatrix} s & d^T \\ d & A^{-1} \end{pmatrix} \in \mathcal{S}^{n+1}_{\succeq 0}, \quad d = A^{-1}x$$

$$\text{and} \quad x_i^T A^{-1} x_i - 2 x_i^T d + s \leq 1, \quad i \in [N].$$

Proof For verifying the condition $P \subseteq \mathcal{E}(A,x)$ it suffices to consider only the points x_i, $i \in [N]$.

$$\begin{aligned}
x_i \in \mathcal{E}(A,x) &\iff (x_i - x)^T A^{-1} (x_i - x) \leq 1 \\
&\iff x_i^T A^{-1} x_i - 2 x_i^T A^{-1} x + x^T A^{-1} x \leq 1 \\
(\text{use } d = A^{-1}x) &\iff x_i^T A^{-1} x_i - 2 x_i^T d + d^T A d \leq 1 \\
&\iff x_i^T A^{-1} x_i - 2 x_i^T d + s \leq 1
\end{aligned}$$

with $s \geq d^T A d$.

Condition $s \geq d^T A d$ is equivalent to $\begin{bmatrix} s & d^T \\ d & A^{-1} \end{bmatrix} \succeq 0$ by the Schur complement because A is positive definite. $\boxtimes$

<u>Goal:</u> Find best outer approximation of P, i.e. $\mathcal{E}(A,x)$ with $P \subseteq \mathcal{E}(A,x)$ and $\operatorname{vol}(\mathcal{E}(A,x)) = \sqrt{\det A}\,\operatorname{vol} B_n$ minimal.

<u>Conic programming formulation:</u>

$$\max \left(\det A^{-1}\right)^{1/n}$$
$$\begin{pmatrix} s & d^T \\ d & A^{-1} \end{pmatrix} \in S^{m+1}_{\succeq 0}$$
$$x_i^T A^{-1} x_i - 2 x_i^T d + s \leq 1, \quad i \in [N]$$

$$= \max t$$
$$(B,t) \in \mathcal{D}^n, \quad \begin{pmatrix} s & d^T \\ d & B \end{pmatrix} \in S^{m+1}_{\succeq 0}$$
$$s_1, \ldots, s_N \geq 0$$
$$x_i^T B x_i - 2 x_i^T d + s + s_i = 1, \quad i \in [N]$$

$$= \max t$$
$$(B,t) \in \mathcal{D}^n, \quad Y \in S^{m+1}_{\succeq 0}, \quad s \in \mathbb{R}^N$$
$$\langle E_{ij}, B \rangle + \langle -E_{i+1,j+1}, Y \rangle = 0 \qquad 1 \leq i \leq j \leq n$$
$$\left\langle \begin{bmatrix} 1 & -x_i^T \\ -x_i & x_i x_i^T \end{bmatrix}, Y \right\rangle + s_i = 1, \qquad i \in [N].$$

<u>Def. 5</u> Let P be a polytope. The ellipsoid E with $P \subseteq E$ having smallest volume is called <u>Loewner–John ellipsoid</u> $\mathcal{E}_{out}(P)$ of P.

Consider the dual MAXDET program

$$\min \sum_{i=1}^{N} y_i$$

$$\left(\sum_{i,j} z_{ij} E_{ij}, -1\right) \in (\mathcal{D}^n)^*$$

$$-\sum_{i,j} z_{ij} E_{i+1,j+1} + \sum_i y_i \begin{bmatrix} 1 & -x_i^T \\ -x_i & x_i x_i^T \end{bmatrix} \succeq 0$$

$$y_i \geq 0, \qquad i \in [N].$$

Lemma 6 Both programs are strictly feasible, therefore strong duality holds, if $\dim P = n$.

Proof Strict feasibility of the primal: Since P is a polytope, P is bounded. So there is a ball $B(x,r)$ with center x and radius r so that $P \subseteq B(x,r)$. Make r so big that there are no vertices of P lying on the boundary of $B(x,r)$. This defines a strictly feasible solution of the primal: For $\varepsilon > 0$ sufficiently small define

$$B = \frac{1}{r^2} I_n, \quad t = \varepsilon$$

$$Y = \begin{bmatrix} s & d^T \\ d & B \end{bmatrix}, \quad d = Bx, \quad s = d^T B^{-1} d + \varepsilon$$

$$s_i = 1 - x_i^T B x_i - 2x_i^T d + s.$$

Strict feasibility of the dual:

Since $\dim P = n$ we may assume, after performing an affine tranformation that $x_1 = e_1, \ldots, x_n = e_n,\ x_{n+1} = 0$.

Set z_{ij} so that $\sum_{i,j} z_{ij} E_{ij} = I_n$. Then $(I_n, -1) \in \operatorname{int}(\mathcal{D}^n)^*$.

For $\varepsilon > 0$ set $y_1 = \ldots = y_n = 2$, $y_{n+1} = 2n + \varepsilon$.

$$y_{n+2} = \ldots = y_N = \varepsilon.$$

Then

$$-I_n + \sum_{i=1}^{N} y_i \begin{bmatrix} 1 & -x_i \\ -x_i & x_i x_i^T \end{bmatrix}$$

$$= \underbrace{\varepsilon \sum_{i=n+2}^{N} \begin{bmatrix} 1 \\ -x_i \end{bmatrix} \begin{bmatrix} 1 \\ -x_i \end{bmatrix}^T}_{\geq 0} + \underbrace{\sum_{i=1}^{n} \left(2 \begin{bmatrix} 1 \\ -e_i \end{bmatrix} \begin{bmatrix} 1 \\ -e_i \end{bmatrix}^T - \begin{bmatrix} 0 \\ e_i \end{bmatrix} \begin{bmatrix} 0 \\ e_i \end{bmatrix}^T \right)}_{\begin{bmatrix} 2n & -2 & \cdots & -2 \\ -2 & 1 & & \\ \vdots & & \ddots & \\ -2 & & & 1 \end{bmatrix}}$$

$$+ \begin{bmatrix} 2n+\varepsilon & 0 \\ 0 & 0 \end{bmatrix}$$

This matrix is positive definite: Use Schur complement:

$$(-2e)^T I_n (-2e) = 4n < 4n + \varepsilon. \qquad \boxtimes$$

<u>Theorem 7</u> (John's optimality condition, 1948)

Let $P \subseteq \mathbb{R}^n$ be a polytope with $\dim P = n$. We have that the following statements are equivalent:

(i) $\mathcal{E}_{out}(P) = B_n$ (B_n, the n-dimensional unit ball, $B_n = \{x \in \mathbb{R}^n : \|x\| \le 1\}$)

$\iff$ $P \subseteq B_n$, $\exists\, \lambda_1, \dots, \lambda_M > 0$, $x_1, \dots, x_M$ vertices of P

so that

(a) $\|x_i\| = 1$, $i \in [M]$

(b) $\sum_{i=1}^{M} \lambda_i x_i = 0$

(c) $\sum_{i=1}^{M} \lambda_i x_i x_i^T = I$ Text Text

(ii) $\mathcal{E}_{in}(P) = B_n \iff B_n \subseteq P$, $\exists\, \lambda_1, \dots, \lambda_M > 0$, $x_1, \dots, x_M \in \partial P$ so that (a) - (c) holds.

<u>Remarks</u> * By applying an affine transformation one can always assume that the Loewner-John ellipsoid is the unit ball.

* Historically, the John optimality conditions were among the first optimality conditions of a nonlinear program.

* Condition (b) says that not all points $x_1, \dots, x_M$ lie on one side of the unit sphere S^{n-1}, i.e.

is not possible

* Condition (c) says that $x_1, \dots, x_M$ behave similar to an orthonormal basis: For $x, y \in \mathbb{R}^n$

$$x^T y = x^T I_n y = \sum_{i=1}^{M} \lambda_i (x^T x_i)(x_i^T y).$$

* Condition (b) and (c) imply: $M \geq n+1$.

<u>Proof</u> (ii) follows from (i) by considering the polar polytope

$$P^* = \{ y \in \mathbb{R}^n : x^T y \leq 1 \text{ for all } x \in P \}.$$

$\rightarrow$ exercise.

(i) "$\Longleftarrow$" Application of weak duality:

Consider the dual program . Define

$$(y_1, \dots, y_M, y_{M+1}, \dots, y_N) = \left(\frac{\lambda_1}{n}, \dots, \frac{\lambda_M}{n}, 0, \dots, 0 \right).$$

$$\sum z_{ij} E_{ij} = \frac{1}{n} I_n.$$

This is a feasible solution for the dual (check it!).

Consider the primal program . Define

$$(B,t) = (I_n, 1), \quad Y = \begin{bmatrix} 0 & 0 \\ 0 & I_n \end{bmatrix}.$$

This is a feasible solution for the primal because $P \subseteq B_n$. The value of the primal is equal to 1. The value of the dual is equal to $\frac{1}{n} \sum_{i=1}^{M} \lambda_i = 1$, too, because

$$\sum \lambda_i \overset{(i)}{=} \sum \lambda_i x_i^T x_i = \mathrm{Tr}\left(\sum \lambda_i x_i x_i^T\right) \overset{(iii)}{=} \mathrm{Tr}(I) = n.$$

"$\Longrightarrow$" Application of strong duality:

$\cdot\ P \subseteq B_n$: ✓

Let $y_1^*, \dots, y_N^*$ be an optimal solution of the dual program. After reordering we may assume

$$y_1^* > 0, \dots, y_M^* > 0, \quad y_{M+1}^* = \dots = y_N^* = 0.$$

By assumption, $(B,t) = (I_n, 1)$, $Y = \begin{bmatrix} 0 & 0 \\ 0 & I_n \end{bmatrix}$ is an optimal solution of the primal. By complementary slackness we know $s_i = 0$, $i = 1, \dots, M$.

Hence $\left\langle \begin{bmatrix} 1 & -x_i^T \\ -x_i^T & x_i x_i^T \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & I_n \end{bmatrix} \right\rangle = 1$ for $i \in [M]$.

This shows $\|x_i\| = 1$, condition (a).

Let $Z^* = \sum z_{ij}^* E_{ij}$ be an optimal solution for the dual. Then (Theorem 2.6): $I_n Z^* = \alpha I_n$ for $\alpha > 0$, and $(\det Z^*)^{1/n} = \frac{1}{n}$. Hence, $Z^* = \frac{1}{n} I_n$.

Let Y^* be optimal for the primal. Then

$$Y^* \left(- \sum_{i,j} z_{ij}^* E_{i+1,j+1} + \sum_i y_i^* \begin{bmatrix} 1 & -x_i^T \\ -x_i & x_i x_i^T \end{bmatrix} \right) = 0$$

$$\Longleftrightarrow \sum_i y_i^* \begin{bmatrix} 1 & -x_i^T \\ -x_i & x_i x_i^T \end{bmatrix} = \begin{bmatrix} * & 0 \\ 0 & \frac{1}{n} I_n \end{bmatrix}.$$

(because $\begin{bmatrix} 0 & 0 \\ 0 & I_n \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ C & D \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$)

Define $\lambda_i = n y_i^*$, then (b), (c) follow from the equality above. $\boxtimes$

<u>Corollary 8</u> Let $P \subseteq \mathbb{R}^n$ be a polytope with $\dim P = n$.

(i) If $\mathcal{E}_{out}(P) = B_n$, then $\frac{1}{n} B_n \subseteq P \subseteq B_n$.

(ii) If $\mathcal{E}_{in}(P) = B_n$, then $B_n \subseteq P \subseteq n B_n$.

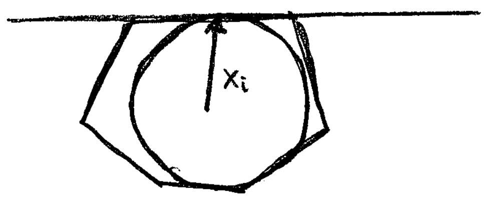
Both inclusions are best possible $\to$ Exercises.

<u>Proof</u> (i) follows from (ii) by polarity.

(ii) Apply Theorem 7: We have

$$\sum_{i=1}^{M} \lambda_i x_i = 0, \quad \sum_{i=1}^{M} \lambda_i x_i x_i^T = I_n, \quad \lambda_i > 0.$$

Consider the supporting hyperplane through $x_i \in B_n \cap \partial P$.



It is orthogonal to x_i.

Hence,

$$B_n \subseteq P \subseteq Q = \{x \in \mathbb{R}^n : x_i^T x \le 1, \ i \in [M]\}.$$

If $x \in Q$, then $x^T x_i \in [-\|x\|, 1]$. Hence,

$$\begin{aligned}
0 &\le \sum_{i=1}^{M} \lambda_i (1 - x^T x_i)(\|x\| + x^T x_i) \\
&= \|x\| \sum_{i=1}^{M} \lambda_i + (1 - \|x\|) \sum_{i=1}^{M} \lambda_i x^T x_i - \sum_{i=1}^{M} \lambda_i (x^T x_i)^2 \\
&= \|x\| \, n + 0 - \|x\|^2
\end{aligned}$$

$\Rightarrow \|x\| \le n.$

because $\sum \lambda_i = \sum \lambda_i x_i^T x_i$

$= \mathrm{Tr}\left(\sum \lambda_i x_i x_i^T\right)$

$= \mathrm{Tr}(I_n) = n.$

☒