Chapter VI Determinant maximization

§ 1 Convex spectral functions

goals: (a) Show that $F : S^n \to \mathbb{R}$

$$X \mapsto \begin{cases} -(\det X)^{1/n} & \text{if } X \succeq 0 \\ 0 & \text{otherwise} \end{cases}$$

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

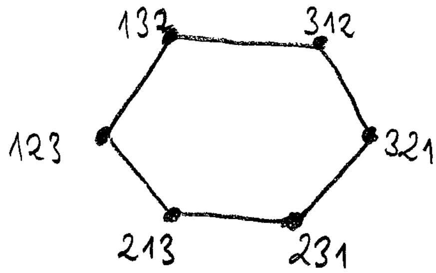
$$SH(X) = \operatorname{conv}\{AXA^T : A \in O(n)\} \subseteq S^n$$

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

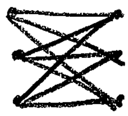
$$Y_{ii} = e_i^T Y e_i = e_i^T \left(\sum_{j=1}^{n} \lambda_j A u_j u_j^T A^T \right) e_i$$

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

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

<u>Proof:</u> - $S_{ij} \ge 0$: ✓

- row sums: $\sum_{j=1}^{n} S_{ij} = 1$: $\sum_{i=1}^{n} (e_i^T A u_j)^2 = \|A u_j\|^2 = 1$
- column sums: $\sum_{i=1}^{n} S_{ij} = 1$: 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 $(Y_{11}, \dots, Y_{nn}) \in \Pi(\lambda_1, \dots, \lambda_n)$. $\boxtimes$

Now we can prove Theorem 2

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

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

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

holds.