



HAL
open science

Variational Principles of Functions In Matrix Analysis

Antoine Mhanna

► **To cite this version:**

| Antoine Mhanna. Variational Principles of Functions In Matrix Analysis. 2015. hal-01186456v1

HAL Id: hal-01186456

<https://inria.hal.science/hal-01186456v1>

Preprint submitted on 25 Aug 2015 (v1), last revised 16 Apr 2016 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Variational Principles of Functions In Matrix Analysis

Antoine Mhanna¹

1 Dept of Mathematics, Lebanese University, Hadath, Beirut, Lebanon.

tmhanat@yahoo.com

Abstract

The main result i present is extending by a very basic lemma the Ky-Fan principle to a larger set of functions satisfying certain conditions.

Keywords : Matrix Analysis, Hermitian matrices, variational characterization, Ky-Fan principle.

1 Introduction

The numerical range of a Hermitian matrix A is called the Rayleigh Quotient $R_A(v)$ which is the application (Rayleigh quotient):

$$\begin{aligned} R_A : \mathbb{C}^n \setminus \{0\} &\rightarrow \mathbb{R} \\ v &\rightarrow \frac{v^* A v}{v^* v}. \end{aligned}$$

This application is well defined because $v^* v = \|v\|^2$ is real and $v^* A v$ is real. The set of $m \times n$ complex matrices is denoted by $\mathbb{M}_{m,n}(\mathbb{C})$.

Proposition 1.1. *Let A be a square matrix, and (v_1, v_2, \dots, v_n) the orthonormal basis of eigenvectors corresponding to the set of eigenvalues $\lambda_i, i = 1, \dots, n$, that is $A v_i = \lambda_i v_i$. Then:*

1. $\lambda_k = R_A(v_k)$ and the range of R_A is $[\lambda_n, \lambda_1]$.
2. $\lambda_1 = \sup_v (R_A(v)) = R_A(B)$, $\lambda_n = \inf_v (R_A(v)) = R_A(C)$
where B is the subspace of eigenvectors corresponding to λ_1 and C is the subspace of eigenvectors corresponding to λ_n .

More specifically:

3. Let V_k be the subspace of \mathbb{C}^n spanned by $\{v_1, v_2, \dots, v_k\}$,
and $W_k = \text{span}\{v_{k+1}, \dots, v_n\} = V_k^\perp$

$$\lambda_k = \inf_{v \in V_k} (R_A(v)) \quad \text{and} \quad \lambda_k = \sup_{v \in W_{k-1}} (R_A(v))$$

Remark 1. Note the difference between (1) and (2) : in (2) the two eigenvalues are uniquely determined by B and C respectively. Obviously the dimension of W_{k-1} is $n - k + 1$, and $\dim(V_k)$ is equal to k .

Proof. Recall that for a Hermitian matrix A there is an orthonormal basis in which A is a diagonal matrix : $\Delta = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ so taking any vector v in \mathbb{C}^n and writing $(\alpha_1, \alpha_2, \dots, \alpha_n)$ as its coordinates in that base:

$$v^* Av = (\alpha_1, \alpha_2, \dots, \alpha_n) \Delta \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \sum_{i=1}^n \lambda_i |\alpha_i|^2. \quad (1)$$

We get this formula:

$$R_A(v) = \frac{\sum_{i=1}^n \lambda_i |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2}$$

As a matter of fact this implies $R_A(\text{span}\{v_i, \dots, v_j\}) = [\lambda_i, \lambda_j]$ and that completes the proof. \square

2 Max principle

In this section let $\lambda_1 \leq \dots \leq \lambda_n$ denote the eigenvalues and (v_1, \dots, v_n) the corresponding eigenvectors of an $n \times n$ Hermitian matrix A , and let $\sigma_1 \leq \dots \leq \sigma_n$ be the singular values of A . Another yet equivalent form of Proposition 1.1 is:

Proposition 2.1. Let A be an $n \times n$ Hermitian matrix then

$$\lambda_1(A) = \max\{R_A(v)\} \text{ so } \lambda_1 = R_A(v_1) \text{ for some } v_1 \in \mathbb{C}^n.$$

$$\vdots$$

$$\lambda_k(A) = \max\{R_A(v), v \perp v_1, v_2, \dots, v_{k-1}\} \text{ and}$$

$$\lambda_k = R_A(v_k) \text{ for some } v_k \perp v_1, v_2, \dots, v_{k-1}.$$

Proof. The characterization of λ_1 is already proved, since the unit ball is a compact set of \mathbb{C}^n and every closed set in a compact is compact, the continuous function x^*Ax will attain its maximum on all described sets, we need to show the proposition for λ_2 and the rest follows easily; take any vector w such that $w \perp v_1$ i.e w orthogonal to v_1 , then $w = \sum_{i=2}^n \alpha_i v_i$ but from Proposition 1.1 $R_A(\text{span}\{v_n, \dots, v_2\}) = [\lambda_n, \lambda_2]$, thus $R_A(w) \leq R_A(v_2)$ since $v_2 \perp v_1$, this completes the proof. \square

Corollary 2.1. *Let A be an $m \times n$ matrix, and let $\sigma_1 \geq \dots \geq \sigma_n$ denotes the singular values of A , we have*

$$\begin{aligned} \sigma_1(A) &= \max\{\|Ax\|, \|x\| = 1\} \text{ so } \sigma_1 = \|Ax_1\| \text{ for some } x_1 \in \mathbb{C}^n. \\ &\vdots \\ \sigma_k(A) &= \max\{\|Ax\|, x \perp x_1, x_2, \dots, x_{k-1}, \|x\| = 1\} \text{ and} \\ \sigma_k &= \|Ax_k\| \text{ for some } x_k \perp x_1, x_2, \dots, x_{k-1}. \end{aligned}$$

Proof. Since $\sigma_i^2(A)$ are the eigenvalues of A^*A , and $\|Ax\|^2 = \langle Ax, Ax \rangle = \langle A^*Ax, x \rangle$ for every $x \in \mathbb{C}^n$, we get the desired representation from Proposition 2.1. \square

Theorem 2.1. *Let $A \in \mathbb{M}_{m,n}(\mathbb{C})$ If $A = W\Sigma V$ is the singular value decomposition, $W = \begin{pmatrix} t_1 & \dots & t_m \\ \vdots & \dots & \vdots \end{pmatrix}$ and $V^* = \begin{pmatrix} s_1 & \dots & s_n \\ \vdots & \dots & \vdots \end{pmatrix}$ we have:*

$$\sigma_1(A) = \max \frac{|x^*Ay|}{\|x\|\|y\|} \quad \text{and} \quad \sigma_k(A) = \max_{\substack{x \in \text{span}\{t_1, \dots, t_{k-1}\}^\perp \\ y \in \text{span}\{s_1, \dots, s_{k-1}\}^\perp}} \frac{|x^*Ay|}{\|x\|\|y\|}$$

for all $k \leq \min(m, n)$.

Proof. Let x and y both be unit vectors, $x \in \mathbb{C}^m$ and $y \in \mathbb{C}^n$, W (respectively V) is an $m \times m$ (respectively $n \times n$) unitary matrix, we first expand the quadratic form in terms of the S.V.D of A :

$$|x^*Ay| = |x^*W\Sigma Vy| = \left| \sum_{i=1}^n \sigma_i (x^*t_i)(s_i^*y) \right| \leq \sigma_1.$$

where the last inequality holds because x and y are both unit vectors. This bound is attained by $x = t_1$ and $y = s_1$ so σ_1 is the optimal value of the optimization problem: $\max_{x,y} \frac{|x^*Ay|}{\|x\|\|y\|}$. To characterize the smaller singular values, we constrain x and y to lie

in $\text{span}\{t_1, \dots, t_{k-1}\}^\perp$ and $\text{span}\{s_1, \dots, s_{k-1}\}^\perp$ respectively. Then we can ignore the $k - 1$ largest singular values and

$$|x^*Ay| = \left| \sum_{i=k}^n \sigma_i(x^*t_i)(s_i^*y) \right| \leq \sigma_k.$$

this bound is attained by $x = t_k$ and $y = s_k$ so σ_k is the optimal value of the optimization problem:

$$\max_{\substack{x \in \text{span}\{t_1, \dots, t_{k-1}\}^\perp \\ y \in \text{span}\{s_1, \dots, s_{k-1}\}^\perp}} \frac{|x^*Ay|}{\|x\|\|y\|}$$

□

3 Variational characterizations

Lemma 3.1. *Let E be a \mathbb{K} -Hilbert space ($\mathbb{K} \equiv \mathbb{R}$ or \mathbb{C}) of dimension n and let f be a continuous function from E to \mathbb{R} . If $r > 0$ is any fixed real number, Set:*

$$h_1 := \max\{f(x), \|x\|_s = r\} \text{ so } h_1 = f(v_1) \text{ for some } v_1 \in \mathbb{C}^n.$$

$$h_2 := \max\{f(x), \|x\|_s = r, x \perp v_1\} \text{ so } h_2 = f(v_2) \text{ for some } v_2 \in \mathbb{C}^n.$$

⋮

$$h_n := \max\{f(x), \|x\|_s = r, x \perp v_1, v_2, \dots, v_{n-1}\}$$

$$\text{so } h_n = f(v_n) \text{ for some } v_n \perp v_1, \dots, v_{n-1}.$$

and set:

$$q_1 := \min\{f(x), \|x\|_s = r\} \text{ so } q_1 = f(w_1) \text{ for some } w_1 \in \mathbb{C}^n.$$

$$q_2 := \min\{f(x), \|x\|_s = r, x \perp w_1\} \text{ so } q_2 = f(w_2) \text{ for some } w_2 \in \mathbb{C}^n.$$

⋮

$$q_n := \min\{f(x), \|x\|_s = r, x \perp w_1, w_2, \dots, w_{n-1}\}$$

$$\text{so } q_n = f(w_n) \text{ for some } w_n \perp w_1, w_2, \dots, w_{n-1}.$$

$\mathcal{A} \equiv (v_1, \dots, v_n)$ and $\mathcal{B} \equiv (w_1, \dots, w_n)$ are two orthonormal basis of E . Let $k \leq n$, and $m < z$ since we study f on the ball of radius r denoted B_r we will suppose that our function f is only defined from B_r to \mathbb{R} .

- If $f(\text{span}(v_z, \dots, v_m)) \leq rf(v_m)$ and $f(\alpha_i v_i + \alpha_{i+1} v_{i+1}) \leq \sum_{j=i}^{i+1} |\alpha_j|^2 f(v_j)$ (max condition) for all $i = 1, \dots, n$ then $\sum_{i=1}^k rh_i = \max_{B_k} \sum_{i=1}^k f(x_i)$, if f is positive valued we have also $\prod_{i=1}^k rh_i = \max_{B_k} \prod_{i=1}^k f(x_i)$.
- If $f(\text{span}(w_z, \dots, w_m)) \geq rf(w_z)$ and $f(\alpha_i w_i + \alpha_{i+1} w_{i+1}) \geq \sum_{j=i}^{i+1} |\alpha_j|^2 f(w_j)$ (min condition) for all $i = 1, \dots, n$ then $\sum_{i=1}^k rq_i = \min_{B_k} \sum_{i=1}^k f(x_i)$, if f is positive valued we have also $\prod_{i=1}^k rq_i = \min_{B_k} \prod_{i=1}^k f(x_i)$.

where $B_k = (x_1, \dots, x_k)$ denotes an orthogonal basis of dimension k such that $\|x_i\|_s = r$ for $1 \leq i \leq k$ and $\alpha_i \in \mathbb{C}$, $\beta_i \in \mathbb{C}$ for all $i = 1, \dots, k$.

Proof. It suffices to prove the maximum version as the other has a similar proof, we will take r to be 1, as the same proof holds if $r \neq 1$. Let us suppose first that E is of dimension $n = 2$, take two orthonormal vectors x_1 and x_2 and since (v_1, v_2) is a basis, necessarily we have

$$\left\{ \begin{array}{l} x_1 = \alpha v_1 + \beta v_2 \\ x_2 = i v_1 + j v_2 \\ |\alpha|^2 + |\beta|^2 = 1 \quad (\|x_1\|_s = 1) \\ |i|^2 + |j|^2 = 1 \quad (\|x_2\|_s = 1) \\ \alpha \bar{i} = -\beta \bar{j} \quad (x_1 \perp x_2) \\ |\alpha|^2 |i|^2 = |\beta|^2 |j|^2 \end{array} \right. \implies \left\{ \begin{array}{l} |\alpha|^2 |i|^2 + |\beta|^2 |j|^2 = i^2 \\ |\beta|^2 (|i|^2 + |j|^2) = |i|^2 \\ |\beta|^2 = |i|^2 \\ |\alpha|^2 = |j|^2 \end{array} \right.$$

The max condition let us write

$$\begin{aligned} f(x_1) &\leq |\alpha|^2 f(v_1) + |\beta|^2 f(v_2) \\ f(x_2) &\leq |i|^2 f(v_1) + |j|^2 f(v_2) = |\beta|^2 f(v_1) + |\alpha|^2 f(v_2) \end{aligned}$$

adding both inequalities yields to $f(x_1) + f(x_2) \leq f(v_1) + f(v_2)$ and we have equality if $v_1 = x_1$ and $x_2 = v_2$. For any n , $n > 2$ we can write:

$$\left\{ \begin{array}{l} x_1 = \alpha_{1,1}v_1 + \alpha_{2,1}v_2 + \cdots + \alpha_{n,1}v_n \\ x_2 = \alpha_{1,2}v_1 + \alpha_{2,2}v_2 + \cdots + \alpha_{n,2}v_n \\ \vdots \\ x_k = \alpha_{1,k}v_1 + \alpha_{2,k}v_2 + \cdots + \alpha_{n,k}v_n \\ |\alpha_{1,1}|^2 + \cdots + |\alpha_{n,1}|^2 = 1 \quad (\|x_1\|_s = 1) \\ |\alpha_{1,2}|^2 + \cdots + |\alpha_{n,2}|^2 = 1 \quad (\|x_2\|_s = 1) \\ \vdots \\ |\alpha_{1,k}|^2 + \cdots + |\alpha_{n,k}|^2 = 1 \quad (\|x_k\|_s = 1) \end{array} \right.$$

Here we assume that $k < n$, let us take any two vectors from the orthonormal basis (x_1, \dots, x_k) , we take the first x to have the minimum j with $\alpha_{j,i} \neq 0$ suppose $j = 1$ and without loss of generality let it be x_1 and the other one x_i for a certain $i < n$, from $x_1 \perp x_i$ and $\|x_i\|_s = 1$ we obtain

$$\left\{ \begin{array}{l} |\alpha_{1,i}|^2 + \cdots + |\alpha_{n,i}|^2 = 1 \\ |\alpha_{1,1}|^2 + \cdots + |\alpha_{n,1}|^2 = 1 \\ \overline{\alpha_{1,1}}\alpha_{1,i} = -(\overline{\alpha_{2,1}}\alpha_{2,i} + \cdots + \overline{\alpha_{n,1}}\alpha_{n,i}) \end{array} \right.$$

Thus we get:

$$\left\{ \begin{array}{l} |\alpha_{1,1}\alpha_{1,i}|^2 = |(\overline{\alpha_{2,1}}\alpha_{2,i} + \cdots + \overline{\alpha_{n,1}}\alpha_{n,i})|^2 \\ |(\overline{\alpha_{2,1}}\alpha_{2,i} + \cdots + \overline{\alpha_{n,1}}\alpha_{n,i})|^2 + |\alpha_{2,i}|^2|\alpha_{1,1}|^2 \cdots + |\alpha_{n,i}|^2|\alpha_{1,1}|^2 = |\alpha_{1,1}|^2 \\ |(\overline{\alpha_{2,1}}\alpha_{2,i} + \cdots + \overline{\alpha_{n,1}}\alpha_{n,i})|^2 + |\alpha_{1,1}|^2(|\alpha_{2,i}|^2 + |\alpha_{3,i}|^2 + \cdots - 1) = 0 \\ |\alpha_{1,i}|^2 = 0. \end{array} \right.$$

By iterating the same process this time applied to another vector say x_2 instead of x_1 we see that the family of orthonormal vectors we have taken (x_1, \dots, x_k) is of this

form

$$\begin{cases} x_1 = \alpha_{1,1}v_1 + \alpha_{2,1}v_2 + \cdots + \alpha_{n,1}v_n \\ x_2 = \alpha_{2,2}v_2 + \cdots + \alpha_{n,2}v_n \\ \vdots \\ x_{k-1} = \alpha_{k-1,k-1}v_{k-1} + \cdots + \alpha_{n,k-1}v_n \\ x_k = \alpha_{k,k}v_k + \cdots + \alpha_{n,k}v_n \end{cases}$$

where any of the indexed α could be equal zero. Consider $k < n$, by hypothesis we get $f(x_i) \leq f(v_i)$ summing over i we get the required result, if $k = n$ we see that we can have $x_{n-1} = \alpha_{n-1,n-1}v_{n-1} + \alpha_{n-1,n}v_n$ and $x_n = \alpha_{n,n-1}v_{n-1} + \alpha_{n,n}v_n$ again by hypothesis and the case we discussed earlier; that is, when the dimension of E equals 2 we get $f(x_n) + f(x_{n-1}) \leq f(v_{n-1}) + f(v_n)$ and that completes the proof. \square

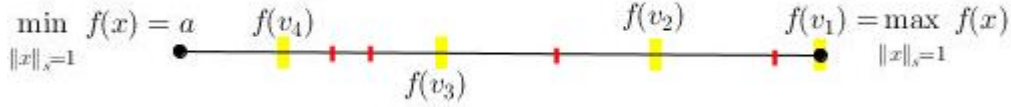


Figure 1: The max condition for a function f over a 4 dimensional space E , where the red marks are the possible $f(x_i), i = 1, \dots, 4$ if (x_1, \dots, x_4) is an orthonormal basis.

Notice that $f(v_n) \leq \dots \leq f(v_1)$ and $f(w_n) \geq \dots \geq f(w_1)$. The proof of Lemma 3.1 suggests that we can relax the max and min conditions, keeping the same terminology, we obtain

Corollary 3.1. *Let $k \leq n, m < z$ and f the function considered defined from B_r to \mathbb{R} .*

- *If $f(\text{span}(v_z, \dots, v_m)) \leq r f(v_m)$ and*

$$f(\alpha_{n-1}v_{n-1} + \alpha_n v_n) + f(\beta_{n-1}v_{n-1} + \beta_n v_n) \leq r(f(v_{n-1}) + f(v_n))$$

then $\sum_{i=1}^k r h_i = \max_{B_k} \sum_{i=1}^k f(x_i)$. In addition if f is positive valued we have also

$$\prod_{i=1}^k r h_i = \max_{B_k} \prod_{i=1}^k f(x_i).$$

- If $f(\text{span}(v_z, \dots, v_m)) \leq f(v_m)$ and

$$f(\alpha_{n-1}v_{n-1} + \alpha_nv_n) + f(\beta_{n-1}v_{n-1} + \beta_nv_n) \leq f(v_{n-1}) + f(v_n)$$

then $\sum_{i=1}^k h_i = \max_{B_k} \sum_{i=1}^k f(x_i)$. In addition if f is positive valued we have also

$$\prod_{i=1}^k h_i = \max_{B_k} \prod_{i=1}^k f(x_i).$$

- If $f(\text{span}(w_z, \dots, w_m)) \geq rf(w_z)$ and

$$f(\alpha_{n-1}w_{n-1} + \alpha_nw_n) + f(\beta_{n-1}w_{n-1} + \beta_nw_n) \geq r(f(w_{n-1}) + f(w_n))$$

then $\sum_{i=1}^k rq_i = \min_{B_k} \sum_{i=1}^k f(x_i)$. In addition if f is positive valued we have also

$$\prod_{i=1}^k rq_i = \min_{B_k} \prod_{i=1}^k f(x_i).$$

- If $f(\text{span}(w_z, \dots, w_m)) \geq f(w_z)$ and

$$f(\alpha_{n-1}w_{n-1} + \alpha_nw_n) + f(\beta_{n-1}w_{n-1} + \beta_nw_n) \geq f(w_{n-1}) + f(w_n)$$

then $\sum_{i=1}^k q_i = \min_{B_k} \sum_{i=1}^k f(x_i)$. In addition if f is positive valued then we have also

$$\prod_{i=1}^k q_i = \min_{B_k} \prod_{i=1}^k f(x_i).$$

where $B_k = (x_1, \dots, x_k)$ denotes an orthogonal basis of dimension k such that $\|x_i\|_s = r$ for $1 \leq i \leq k$ and $\alpha_i \in \mathbb{C}$, $\beta_i \in \mathbb{C}$ for $i = n-1, n$.

Here is an example showing that the conditions in Lemma 3.1 are necessary.

Example 3.1. Let

$$\begin{aligned} f : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ u = (x, y) &\rightarrow \ln(|x + \epsilon|). \end{aligned}$$

where ϵ is a strictly positive number to be fixed, we verify then that $h_1 = \ln(|1 + \epsilon|)$
 $= \max_{\|x\|_s=1} \ln(|x + \epsilon|) = f(v_1)$, by taking $v_1 = (1, 0)$, $v_2 = (0, \pm 1)$, $x_1 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ and
 $x_2 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ we have $v_1 \perp v_2$, $x_1 \perp x_2$ but

$$f(v_1) + f(v_2) = \ln(\epsilon(1 + \epsilon)) = h_1 + h_2 < f(x_1) + f(x_2) = \ln(\sqrt{2} + 0.5 + \epsilon^2)$$

whenever $\epsilon < 0.5 + \sqrt{2}$.

Lemma 3.1 and the previous results will entail important variational representations concerning the sum and product of eigenvalues and singular values:

Corollary 3.2 (Ky-Fan principle). *Let H be an $n \times n$ Hermitian matrix such that $\lambda_1 \geq \dots \geq \lambda_n$ are the eigenvalues of H in decreasing order. For any $1 \leq k \leq n$*

$$\sum_{i=1}^k \lambda_i(H) = \max_{U^*U=I_k} \operatorname{tr}(U^*HU) \quad (2)$$

$$\sum_{i=1}^k \lambda_{n-i+1}(H) = \min_{U^*U=I_k} \operatorname{tr}(U^*HU) \quad (3)$$

And if H is P.S.D that is λ_n is nonnegative we have

$$\prod_{i=1}^k \lambda_i(H) = \max_{U^*U=I_k} \det(U^*HU) \quad (4)$$

$$\prod_{i=1}^k \lambda_{n-i+1}(H) = \min_{U^*U=I_k} \det(U^*HU). \quad (5)$$

Proof. It suffices to notice that $\max_{U^*U=I_k} \operatorname{tr}(U^*HU) = \max_{x_i^*x_j=\delta_{ij}} \sum_{i=1}^k x_i^*Hx_i$ applying Proposition 2.1 and Lemma 3.1 to the function $f(x) = x^*Hx$ completes the proof, since $f(\operatorname{span}(v_z, \dots, v_m)) \leq f(v_m)$ when $m < z$ (see Proposition 1.1) and for all i we have $f(\alpha_i v_i + \alpha_{i+1} v_{i+1}) = |\alpha_i|^2 f(v_i) + |\alpha_{i+1}|^2 f(v_{i+1})$. For the product assertion if H is P.S.D we have $\max_{U^*U=I_k} \det(U^*HU) = \max_{x_i^*x_j=\delta_{ij}} (x_i^*Hx_j)$ but since U^*HU is Hermitian we can suppose that it is a diagonal matrix we apply again Proposition 2.1 and Lemma 3.1 to the positive valued function $f(x) := x^*Hx$ to complete the proof. The minimum versions are similarly proven. \square

Definition 3.1. *A matrix $P \in \mathbb{M}_{m,n}$ is said to be a rank k partial isometry or a k partial isometry if $\sigma_1(P) = \dots = \sigma_k(P) = 1$ and all other singular values are equal to zero.*

Let us denote the set of the rank k partial isometries in $\mathbb{M}_{m,n}$ by $L_{m,n}^k$.

Corollary 3.3. *Let $A \in \mathbb{M}_{m,n}$ have singular values $\sigma_1(A) \geq \dots \geq \sigma_q(A) \geq 0$, where*

$q = \min\{m, n\}$. For each $k = 1, \dots, q$ we have :

$$\sum_{i=1}^k \sigma_i(A) = \max \{ |tr X^* A Y| : X \in \mathbb{M}_{m,k}, Y \in \mathbb{M}_{n,k}, X^* X = I = Y^* Y \} \quad (6)$$

$$= \max \{ Re(tr X^* A Y) : X \in \mathbb{M}_{m,k}, Y \in \mathbb{M}_{n,k}, X^* X = I = Y^* Y \} \quad (7)$$

$$= \max \{ |tr A P| : P \in L_{n,m}^k \} \quad (8)$$

Proof. (6) and (7) representations will follow from Theorem 2.1 and a direct generalization of Lemma 3.1. Let us show first that they are equivalent to the last one. If $X \in \mathbb{M}_{m,k}$ and $Y \in \mathbb{M}_{n,k}$ satisfy $Y^* Y = I = X^* X$, $tr X^* A Y = tr A Y X^*$ and $C := Y X^* \in \mathbb{M}_{n,m}$ has $C^* C = X Y^* Y X^* = X X^*$. Since the k largest singular values of $X X^*$ are the same as those of $X^* X = I \in \mathbb{M}_k$, we conclude that $C = Y X^*$ is a rank k partial isometry. Conversely, if $C \in \mathbb{M}_{n,m}$ is a given rank k partial isometry, then the singular value decomposition of C is

$$C = W \Sigma V = \begin{pmatrix} W_{n,k} & * \\ & \end{pmatrix} \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_{m,k} \\ * \end{pmatrix} = W_{n,k} V_{k,m} = W_{n,k} (V_{k,m}^*)^*$$

and we verify that $W_{n,k}^* W_{n,k} = I_k = V_{k,m}^* V_{k,m}$. Thus the asserted variational formulas are equivalent. For (6) and (7), recall from Theorem 2.1 that

$$\sigma_k(A) = \max_{\substack{x \in \text{span}\{t_1, \dots, t_{k-1}\}^\perp \\ y \in \text{span}\{s_1, \dots, s_{k-1}\}^\perp}} \frac{x^* A y}{\|x\| \|y\|}$$

for all $k \leq \min(m, n)$, where we had $x^* A y = x^* W \Sigma V y = \sum_{i=1}^n \sigma_i(x^* t_i)(s_i^* y)$. This summation makes sense since $\sigma_i = 0$ if $i > \min(m, n)$. If we consider adopting the proof of the max condition in Lemma 3.1 to $f(x, y) = |x^* A y|$, the max condition would be slightly different, if $f(x)$ is the function defined from $B_1 \times B_1$ to \mathbb{R} by $f(x, y) = |x^* A y|$ then we have from the expression of $f(x)$ that whenever $z < l$ and $p < j$ $f(\text{span}(t_z, \dots, t_l), \text{span}(s_p, \dots, s_j)) \leq f(t_u, s_u)$ where $u := \min(z, p)$, we only need to show that

$$f(ct_i + et_{i+1}, ds_i + as_{i+1}) + f(wt_i + rt_{i+1}, ls_i + os_{i+1}) \leq f(t_i, s_i) + f(t_{i+1}, s_{i+1})$$

for any complex numbers c, e, d, r, o, w, a, l and all $i, i = 1, \dots, \min(m, n)$. Now fix $x_1 = ct_i + et_{i+1}$, $y_1 = ds_i + as_{i+1}$, $x_2 = wt_i + rt_{i+1}$ and $y_2 = ls_i + os_{i+1}$, $y_2 \perp y_1$

and $x_2 \perp x_1$, $\|x_1\|_s = \|y_1\|_s = \|y_2\|_s = \|x_2\|_s = 1$, similarly to Lemma 3.1 from the orthogonality-norm conditions

$$\begin{cases} |d|^2 = |o|^2 \\ |l|^2 = |a|^2 \\ |c|^2 = |r|^2 \\ |w|^2 = |e|^2 \end{cases}$$

and by the Cauchy-Schwartz inequality i will assume that all our complex scalars are real 'positive' because $f(ct_i + et_{i+1}, ds_i + rs_{i+1}) + f(wt_i + rt_{i+1}, ls_i + os_{i+1}) \leq f(|c|t_i + |e|t_{i+1}, |d|s_i + |r|s_{i+1}) + f(|w|t_i + |r|t_{i+1}, |l|s_i + |o|s_{i+1})$ and the upper bound is well in the range of the function f which is restricted to $B_1 \times B_1$. The assumption tells us that $c = r, a = l, o = d, w = e$, compute to get

$$f(x_1, y_1) + f(x_2, y_2) = (\sigma_i + \sigma_{i+1})(cd + ea) \leq f(t_i, s_i) + f(t_{i+1}, s_{i+1}) = \sigma_i + \sigma_{i+1}$$

if and only if $cd + ea \leq 1$, but $c^2 + e^2 = d^2 + a^2 = 1$ thus if $m := a - c$

$$\sqrt{1 - e^2}\sqrt{1 - a^2} + ea \leq 1 \iff (1 - c^2)(1 - a^2) \leq (1 - ac)^2 \quad (9)$$

$$\iff (1 - c)(1 + a)(1 - a)(1 + c) \leq (1 - ca)^2 \quad (10)$$

$$\iff (1 - ca - m)(1 - ca + m) \leq (1 - ca)^2 \quad (11)$$

$$\iff (1 - ca)^2 - m^2 \leq (1 - ca)^2 \quad (12)$$

By using Theorem 2.1 and applying Lemma 3.1 the proof is completed. \square

Since $f(x, y)$ is positive valued, a direct consequence of this characterization is that for any $k \leq \min(m, n)$, $\prod_{i=1}^k \sigma_i = \max_{B_k, C_k} \prod_{i=1}^k f(x_i, y_i)$, where $B_k = (x_1, \dots, x_k)$ and $C_k = (y_1, \dots, y_k)$ are two orthonormal basis of \mathbb{C}^k . We can do a similar generalization to the min condition for the function $f(x, y)$ as the following corollary shows:

Corollary 3.4. For $k \leq \min(m, n)$, $A \in \mathbb{M}_{m,n}(\mathbb{C})$ fixed, define f_A as:

$$\begin{aligned} f_A : \{ (x, y) \in \mathbb{C}^m \times \mathbb{C}^n; \|x\|_s = \|y\|_s = 1 \} &\rightarrow \mathbb{R} \\ (x, y) &\rightarrow |x^* Ay| \end{aligned}$$

and

$$\mathfrak{q}_1 := \min\{f_A(x, y), \|x\|_s = \|y\|_s = 1\} \text{ so } \mathfrak{q}_1 = f_A(d_1, e_1) \text{ for some } (d_1, e_1) \in \mathbb{C}^m \times \mathbb{C}^n.$$

$$\mathfrak{q}_2 := \min\{f_A(x, y), \|x\|_s = \|y\|_s = 1, x \perp d_1, y \perp e_1\}$$

$$\text{so } \mathfrak{q}_2 = f_A(d_2, e_2) \text{ for some } (d_2, e_2) \in \mathbb{C}^m \times \mathbb{C}^n.$$

\vdots

$$\mathfrak{q}_k := \min\{f(x, y), \|x\|_s = \|y\|_s = 1, x \perp d_1, \dots, d_{k-1}, y \perp e_1, \dots, e_{k-1}\}$$

$$\text{so } \mathfrak{q}_k = f(d_k, e_k) \text{ for some } (d_k, e_k) \in \mathbb{C}^m \times \mathbb{C}^n.$$

then we have

$$\sum_{i=1}^k \mathfrak{q}_i = \min \{|\text{tr} X^* A Y| : X \in \mathbb{M}_{m,k}, Y \in \mathbb{M}_{n,k}, X^* X = I = Y^* Y\} \quad (13)$$

$$= \min_{B_k, C_k} \sum_{i=1}^k f(x_i, y_i) := \mathfrak{w} = 0 \quad (14)$$

where x_i is a column of X , y_i a column of Y , $B_k = (x_1, \dots, x_k)$ and $C_k = (y_1, \dots, y_k)$

Proof. Since $f_A(x, y)$ is positive, if we showed that for all $k \leq \min(m, n)$ we have

$$\sum_{i=1}^k \mathfrak{q}_i = 0, \text{ then } \sum_{\mathfrak{B}_k, \mathfrak{C}_k, i=1}^k f(x_i, y_i) \text{ is necessarily nothing but } \mathfrak{w} \text{ where } \mathfrak{B}_k \equiv (d_1, \dots, d_k)$$

and $\mathfrak{C}_k \equiv (e_1, \dots, e_k)$. Recall from Theorem 2.1 that

$$f_A = |x^* A y| = |x^* W \Sigma V y| = \left| \sum_{i=1}^n \sigma_i(x^* t_i)(s_i^* y) \right|$$

We seek the minimum: \mathfrak{q}_1 of f_A over $B_1 \times B_1$, let us take the two orthonormal basis of

dimension k , $T_k = (t_1, \dots, t_k)$ and $S_k = (s_1, \dots, s_k)$ where (t_1, \dots, t_k) are the columns

of W and (s_1, \dots, s_k) are the columns of V^* , set $S_{\mathfrak{d}(k)}$ to be the orthonormal basis

$(s_{\mathfrak{d}(1)}, \dots, s_{\mathfrak{d}(k)}) = (\mathfrak{a}_1, \dots, \mathfrak{a}_k)$ where \mathfrak{d} is a permutation with no fixed point then by

a direct computation we can take explicitly \mathfrak{B}_k to be the basis (t_1, \dots, t_k) , \mathfrak{C}_k to be

$S_{\mathfrak{d}(k)}$ which is previously defined and get

$$\mathfrak{q}_1 = f_A(t_1, \mathfrak{a}_1) = 0 \quad (15)$$

$$\mathfrak{q}_2 = f_A(t_2, \mathfrak{a}_2) = 0 \quad (16)$$

$$\vdots \quad \quad \quad \vdots \quad (17)$$

$$\mathfrak{q}_k = f_A(t_k, \mathfrak{a}_k) = 0 \quad (18)$$

Another way to see the result is to apply the generalization form of Corollary 3.1; that is, when the function is defined on a product space $\mathbb{C}^m \times \mathbb{C}^n$; we need to show, first if $z < l$, $p < j$ and $o = \max(l, j)$ that

$$f_A(\text{span}(d_z, \dots, d_l), \text{span}(e_p, \dots, e_j)) \geq f(d_o, e_o) = 0$$

which is true, second if we consider without loss of generality that $m \leq n$ we should have for any $\alpha_i \in \mathbb{C}$, $\beta_i \in \mathbb{C}$ with $m - 1 \leq i \leq n$:

$$f_A(\alpha_{m-1}d_{m-1} + \alpha_m d_m, \sum_{i=m-1}^n \beta_i e_i) \geq f_A(d_{m-1}, e_{m-1}) + f_A(d_m, e_m) := 0 + 0 = 0$$

which is also true and that completes the proof. □

References

- [1] F.Zhang, *Matrix Theory Basic Results and Techniques, 2nd Edition*
(Universitext) Springer, 2013.
- [2] R. A. Horn and C. R Johnson, *Topics in Matrix Analysis*, Cambridge University
Press, New York, 1991.