SOME VARIATIONAL PRINCIPLES OVER FINITE DIMENTIONAL HILBERT SPACES

Antoine Mhanna

SOME VARIATIONAL PRINCIPLES OVER FINITE DIMENTIONAL HILBERT SPACES.

ANTOINE MHANNA[1]*

Abstract. In this paper a new variational approach concerning functions (continuous) over Hilbert spaces is presented. This will extend the Ky Fan principles (of eigenvalues) to a larger set of functions. Moreover we generalize properties for functions defined over a product of finite dimentional Hilbert spaces and show that the stated conditions are sufficient but not necessary. An obvious generalization of the Courant-Fischer minimax theorem is also given.

1. Introduction and preliminaries

The numerical range of a Hermitian matrix A is the image of the Rayliegh Quotient R_A which is the application :

$$\begin{array}{rcl} R_A : \mathbb{C}^n \backslash \{0\} & \to & \mathbb{R} \\ v & \to & \dfrac{v^* A v}{v^* v}. \end{array}$$

The set of $a \times n$ complex matrices is denoted by $\mathbb{M}_{a,n}(\mathbb{C})$. Let $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$ denote the eigenvalues of an $n \times n$ Hermitian matrix A and let $\sigma_1(A) \geq \cdots \geq \sigma_n(A)$ be the singular values of a matrix $A \in \mathbb{M}_{a,n}(\mathbb{C})$.

Proposition 1.1. [1] *Let A be an $n \times n$ Hermitian matrix then*

$$\begin{aligned} \lambda_1(A) &= \max\{R_A(v)\} \textit{ so } \lambda_1 = R_A(v_1) \textit{ for some } v_1 \in \mathbb{C}^n. \\ &\vdots \\ \lambda_k(A) &= \max\{R_A(v), v \perp v_1, v_2, \cdots, v_{k-1}\} \textit{ and } \\ \lambda_k &= R_A(v_k) \textit{ for some } v_k \perp v_1, v_2, \cdots, v_{k-1}. \end{aligned}$$

The same v_i verifying Proposition 1.1 will verify the following:

Date: 2016.

* Corresponding author.

2010 *Mathematics Subject Classification.* Primary: 15A18; 39B55; 47J20, Secondary: 15A42; 39B62.

Key words and phrases. Variational characterization; Orthogonality; Ky Fan principle; Courant-Fischer theorem.

Proposition 1.2. [1] *Let A be an $n \times n$ Hermitian matrix then*

$$\begin{aligned}
\lambda_n(A) &= \min\{R_A(v)\} \text{ so } \lambda_n = R_A(v_n) \text{ for some } v_n \in \mathbb{C}^n. \\
&\vdots \\
\lambda_k(A) &= \min\{R_A(v), v \perp v_n, v_{n-1}, \cdots, v_{k-1}\} \text{ and} \\
\lambda_k &= R_A(v_k) \text{ for some } v_k \perp v_n, v_{n-1}, \cdots, v_{k-1}.
\end{aligned}$$

2. Main results

Since we will consider $\mathbb{C}$-Hilbert spaces, coefficients of any vector written in a certain basis are in $\mathbb{C}$.

If f is a given function that is defined over a product of vector spaces and a certain domain D, we mean hereafter by $f(sp(\mathfrak{u}_1), \cdots, sp(\mathfrak{u}_n))$ where $\mathfrak{u}_i$ is any set of vectors, the value of $f(x_1, \cdots, x_n)$ for any of the x_i's taken to be in $span(\mathfrak{u}_i)$ and in the domain D.

2.1. Variational characterizations.

Lemma 2.1. *Let $(\mathcal{H}, \|.\|_s)$ be a $\mathbb{K}$-Hilbert space ($\mathbb{K} \equiv \mathbb{R}$ or $\mathbb{C}$) of dimension n where $\|.\|_s$ denotes the norm associated to the scalar product on $\mathcal{H}$. Let f be a continuous function from $\mathcal{H}$ to $\mathbb{R}$. If $r > 0$ is any fixed real number, set:*

$$\begin{aligned}
h_1 &:= \max\{f(x), \|x\|_s = r\} \text{ so } h_1 = f(v_1) \text{ for some } v_1 \in \mathcal{H}. \\
h_2 &:= \max\{f(x), \|x\|_s = r, x \perp v_1\} \text{ so } h_2 = f(v_2) \text{ for some } v_2 \in \mathcal{H}. \\
&\vdots \\
h_n &:= \max\{f(x), \|x\|_s = r, x \perp v_1, v_2, \cdots, v_{n-1}\} \\
&\quad \text{so } h_n = f(v_n) \text{ for some } v_n \perp v_1, \cdots, v_{n-1}.
\end{aligned}$$

and set:

$$\begin{aligned}
q_1 &:= \min\{f(x), \|x\|_s = r\} \text{ so } q_1 = f(w_1) \text{ for some } w_1 \in \mathcal{H}. \\
q_2 &:= \min\{f(x), \|x\|_s = r, x \perp w_1\} \text{ so } q_2 = f(w_2) \text{ for some } w_2 \in \mathcal{H}. \\
&\vdots \\
q_n &:= \min\{f(x), \|x\|_s = r, x \perp w_1, w_2, \cdots, w_{n-1}\} \\
&\quad \text{so } q_n = f(w_n) \text{ for some } w_n \perp w_1, w_2, \cdots, w_{n-1}.
\end{aligned}$$

$\mathcal{A} \equiv (v_1, \cdots, v_n)$ and $\mathcal{B} \equiv (w_1, \cdots, w_n)$ 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 real function f is only defined on $D := B_r$.

- *If* $f(sp(v_m,\cdots,v_z))\geq r^2f(v_z)$ *and* $f\big(\sum_{j=s}^{n}\alpha_jv_j\big)\leq\sum_{j=s}^{n}|\alpha_j|^2f(v_j)$ *(max condition) for all* $s=1,\cdots,n$ *then* $\sum_{i=1}^{k}r^2h_i\geq\sum_{i=1}^{k}f(y_i)$. *In particular if* $r=1$ *then* $\sum_{i=1}^{k}h_i=\max_{B_k}\sum_{i=1}^{k}f(x_i)$.
- *If* $f(sp(w_m,\cdots,w_z))\leq r^2f(w_z)$ *and* $f\big(\sum_{j=s}^{n}\alpha_jw_j\big)\geq\sum_{j=s}^{n}|\alpha_j|^2f(w_j)$ *(min condition) for all* $s=1,\cdots,n$ *then* $\sum_{i=1}^{k}r^2q_i\leq\sum_{i=1}^{k}f(y_i)$. *In particular if* $r=1$ *then* $\sum_{i=1}^{k}q_i=\min_{B_k}\sum_{i=1}^{k}f(x_i)$,

where $B_k=(x_1,\cdots,x_k)$ *denotes an orthonormal basis of dimension* k, $(y_1,\cdots,y_k)$ *is any orthogonal basis of dimension* k *with* $\|x_i\|_s=r$ *for* $1\leq i\leq k$, $\alpha_j\in\mathbb{C}$ *and* $\beta_j\in\mathbb{C}$ *for all* j.

Proof. The case of $k=1$ is obvious, here we assume that $1<k\leq n$, we can write:

$$\begin{cases} x_1=\alpha_{1,1}v_1+\alpha_{2,1}v_2+\cdots+\alpha_{n,1}v_n\\ x_2=\alpha_{1,2}v_1+\alpha_{2,2}v_2+\cdots+\alpha_{n,2}v_n\\ \quad\vdots\\ x_k=\alpha_{1,k}v_1+\alpha_{2,k}v_2+\cdots+\alpha_{n,k}v_n\\ |\alpha_{1,1}|^2+\cdots+|\alpha_{n,1}|^2=r^2\ \ (\|x_1\|_s=r)\\ |\alpha_{1,2}|^2+\cdots+|\alpha_{n,2}|^2=r^2\ \ (\|x_2\|_s=r)\\ \quad\vdots\\ |\alpha_{1,k}|^2+\cdots+|\alpha_{n,k}|^2=r^2\ \ (\|x_k\|_s=r)\\ x_i\perp x_j,\ \text{if } i\neq j.\end{cases}\qquad (A)$$

These k vectors are completed by $n-k$ vectors (of norm r) orthogonal to them and mutually orthogonal; we will impose a supplementary condition over the added $n-k$ vectors as follows; without loss of generality

$$\begin{cases} x_1=\alpha_{1,1}v_1+\alpha_{2,1}v_2+\cdots+\alpha_{n,1}v_n\\ x_2=\alpha_{1,2}v_1+\alpha_{2,2}v_2+\cdots+\alpha_{n,2}v_n\\ \quad\vdots\\ x_k=\alpha_{1,k}v_1+\cdots+\alpha_{n,k}v_n\\ x_{k+1}=\alpha_{1,k+1}v_1+\cdots+\alpha_{n-1,k+1}v_{n-1}\\ \quad\vdots\\ x_n=\alpha_{1,n}v_1+\cdots+\alpha_{k,n}v_k.\end{cases}\qquad (S)$$

It is easily seen that such basis always exists, we denote it by $\mathcal{C}$, the idea is that the change of basis matrix (between basis $\mathcal{C}$ and basis $\mathcal{A}$) is a matrix U satisfying $U^*U = r^2 I_n$ and so applying the max condition we have:

$$f(x_1) + \cdots + f(x_n) \leq r^2(f(v_1) + \cdots + f(v_n))$$

with $f(x_j) \geq r^2 f(v_{n+k-j})$ for all j, $n \geq j > k$. Consequently we obtain:

$$\begin{aligned} f(x_1) + \cdots + f(x_k) &\leq r^2(f(v_1) + f(v_2) + \cdots + f(v_{k-1}) + f(v_n)) \\ &\leq r^2(f(v_1) + f(v_2) + \cdots + f(v_k)). \end{aligned}.$$

To prove the minimum characterization we replace v_i by w_i in (A) and (S) to get the system:

$$\begin{cases} x_1 = \alpha_{1,1}w_1 + \alpha_{2,1}w_2 + \cdots + \alpha_{n,1}w_n \\ x_2 = \alpha_{1,2}w_1 + \alpha_{2,2}w_2 + \cdots + \alpha_{n,2}w_n \\ \quad \vdots \\ x_k = \alpha_{1,k}w_1 + \cdots + \alpha_{n,k}w_n & (G) \\ x_{k+1} = \alpha_{1,k+1}w_1 + \cdots + \alpha_{n-1,k+1}w_{n-1} \\ \quad \vdots \\ x_n = \alpha_{1,n}w_1 + \cdots + \alpha_{k,n}w_k \end{cases}$$

and by the min condition it is not difficult to show -like we did previously- that:

$$\begin{aligned} f(x_1) + \cdots + f(x_k) &\geq r^2(f(w_1) + f(w_2) + \cdots + f(w_{k-1}) + f(w_n)) \\ &\geq r^2(f(w_1) + \cdots + f(w_k)). \end{aligned}.$$

Thus we have discussed all possible cases to complete the proof. □

Remark 2.2. Notice that $f(v_n) \leq \cdots \leq f(v_1)$ and $f(w_n) \geq \cdots \geq f(w_1)$. The way we constructed the systems $(S), (G)$ is important and will be used later on (Theorem 2.5).

To clarify things a counter example is easily constructed:

Example 2.3. Let

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

where ϵ is a strictly positive number to be fixed, here the norm $\|.\|_s$ is the euclidien norm, we verify then that $h_1 = \ln(|1 + \epsilon|) = \max\limits_{\|u\|_s = 1} \ln(|x + \epsilon|) = f(v_1)$, 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

$$\begin{aligned} f(v_1) + f(v_2) = h_1 + h_2 &= \ln(\epsilon(1 + \epsilon)) \\ &< f(x_1) + f(x_2) := \ln(\sqrt{2}.\epsilon + 0.5 + \epsilon^2), \end{aligned}$$

whenever $\epsilon > 0$.

Lemma 2.1 and previous statements will entail some well known variational representations concerning the sum moreover the product (in Subsection 2.2) of eigenvalues of matrices. Their are many related and particular results in the mathematical literature that discuss maximum principles see for example [7], [2] and [8] but most of the representations related to matrices (and even operators) were firstly proved by Ky Fan (see [4], [5] and [6]).

Corollary 2.4 (Ky Fan principle). *Let H be an $n\times n$ Hermitian matrix such that $\lambda_1 \geq \cdots, \geq \lambda_n$ are the eigenvalues of H in decreasing order. For any $1 \leq k \leq n$*

$$\sum_{i=1}^{k} \lambda_i(H) = \max_{U^*U=I_k} tr\ (U^*HU) \tag{2.1}$$

$$\sum_{i=1}^{k} \lambda_{n-i+1}(H) = \min_{U^*U=I_k} tr\ (U^*HU) \tag{2.2}$$

Proof. It suffices to notice that $\max_{U^*U=I_k} \mathrm{tr}\ (U^*HU) = \max_{x_i^*x_j=\delta_{ij}} \sum_{i=1}^{k} x_i^*Hx_i$, respectively $\min_{U^*U=I_k} \mathrm{tr}\ (U^*HU) = \min_{x_i^*x_j=\delta_{ij}} \sum_{i=1}^{k} x_i^*Hx_i$, but from Proposition 1.1 respectively Proposition 1.2 if $f(x) = x^*Hx$, $\mathcal{H} \equiv \mathbb{C}^n$ then by applying Lemma 2.1 to f with $r = 1$ we get the required results because upon diagonalizing H we have: for all i, $f\big(\sum_{j=1}^{i} \alpha_j v_j\big) = \sum_{j=1}^{i} |\alpha_j|^2 f(v_j)$ and $f(v_z) \leq f(sp(v_m, \cdots, v_z)) \leq f(v_m)$ when $m < z$ with w_t taken to be equal v_{n-t} for all t. $\square$

2.2. **Generalization to product spaces.** In the previous section we have taken $\mathcal{H}$ to be any Hilbert space, hereafter we consider f a continuous function defined over $\mathcal{H}_{1,j_1} \times \cdots \times \mathcal{H}_{n,j_n}$ for any n, where $\mathcal{H}_{l,j_l}$ denotes a given Hilbert vector-space of dimension j, with scalar product denoted by $\langle\cdot,\cdot\rangle_{\mathcal{H}_{l,j_l}}$ and $\|.\|_{\mathcal{H}_{l,j_l}}$ the associated norm. We will restrict also our real valued functions f to be defined only over the domain $B_{r_1} \times \cdots \times B_{r_n}$, where for any l, r_l is a fixed real number, B_{r_l} stands for the sphere of radius r in $\mathcal{H}_{l,j_l}$.

Let $m := \min(j_1, \cdots, j_n)$, for $k \leq m$, similarly we set:

$$\begin{aligned}
&\mathfrak{G}_1 := \max\{f(x_1, \cdots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\ \forall l\} \text{ so } \mathfrak{G}_1 = f(z_{1,1}, \cdots, z_{n,1}),\\
&\mathfrak{G}_2 := \max\{f(x_1, \cdots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\ \forall l,\ x_1 \perp z_{1,1}, \cdots, x_n \perp z_{n,1}\}\\
&\qquad \text{so } \mathfrak{G}_2 = f(z_{1,2}, \cdots, z_{n,2}),\\
&\qquad \vdots\\
&\mathfrak{G}_k := \max\{f(x_1, \cdots, x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\ \forall l,\ x_1 \perp z_{1,1}, z_{1,2} \cdots, z_{1,k-1};\\
&\qquad \cdots; x_n \perp z_{n,1}, z_{n,2} \cdots, z_{n,k-1}\} \text{ so } \mathfrak{G}_k = f(z_{1,k}, \cdots, z_{n,k}),
\end{aligned}$$

and:

$$\mathfrak{O}_1 := \min\{f(x_1,\cdots,x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\ \forall l\} \text{ so } \mathfrak{O}_1 = f(c_{1,1},\cdots,c_{n,1}),$$

$$\mathfrak{O}_2 := \min\{f(x_1,\cdots,x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\ \forall l,\ x_1 \perp c_{1,1},\cdots,x_n \perp c_{n,1}\}$$

$$\text{so } \mathfrak{O}_2 = f(c_{1,2},\cdots,c_{n,2}),$$

$$\vdots$$

$$\mathfrak{O}_k := \min\{f(x_1,\cdots,x_n), \|x_l\|_{\mathcal{H}_{l,j_l}} = r_l\ \forall l,\ x_1 \perp c_{1,1},c_{1,2}\cdots,c_{1,k-1};$$

$$\cdots; x_n \perp c_{n,1},c_{n,2}\cdots,c_{n,k-1}\} \text{ so } \mathfrak{O}_k = f(c_{1,k},\cdots,c_{n,k}).$$

For a certain l and k fixed; $\mathfrak{Z}_{l,k} \equiv (z_{l,1},\cdots,z_{l,k})$ and $\mathfrak{D}_{l,k} \equiv (c_{l,1},\cdots,c_{l,k})$ are two orthogonal basis of $\mathcal{H}_{l,j_l}$ each one of dimension k and such that $\|c_{l,s}\|_{\mathcal{H}_{l,j_l}} = \|z_{l,q}\|_{\mathcal{H}_{l,j_l}} = r_l$ for all q and s.

A direct generalization of Lemma 2.1 would be the following:

Theorem 2.5. *Let* $y_{i,1} \leq y_{i,2}$ *for all* $i \leq n$, $\mu = \left\{\max_i y_{i,2} / y_{i,2} \leq m\right\}$, $k \leq m$ *and let* $\mathfrak{X}_{l,k} = (x_{l,1},\cdots,x_{l,k})$ *denote any orthogonal basis of* $\mathcal{H}_{l,j_l}$ *of dimension* k *such that* $\|x_{l,g}\|_{\mathcal{H}_{l,j_l}} = r_l$ *for all* g. *Let* f *be a continuous function from* $D := B_{r_1} \times \cdots \times B_{r_n}$ *into* $\mathbb{R}$.
1) *Suppose we have*

$$f(sp(z_{1,y_{1,1}},\cdots,z_{1,y_{1,2}}),\cdots,sp(z_{n,y_{n,1}},\cdots,z_{n,y_{n,2}})) \geq f(z_{1,\mu},\cdots,z_{n,\mu})$$

then:

- *If* $\displaystyle\sum_{g=1}^{m} f(x_{1,g},\cdots,x_{n,g}) \leq \sum_{i=1}^{m} f(z_{1,i},\cdots,z_{n,i})$ *(sum max condition). Then*

$$\sum_{i=1}^{k} \mathfrak{G}_i = \max_{\mathfrak{X}_{1,k},\cdots,\mathfrak{X}_{n,k}} \sum_{g=1}^{k} f(x_{1,g},\cdots,x_{n,g}).$$

- *If* f *is positive valued and* $\displaystyle\prod_{g=1}^{m} f(x_{1,g},\cdots,x_{n,g}) \leq \prod_{i=1}^{m} f(c_{1,i},\cdots,c_{n,i})$ *(product max condition) then*

$$\prod_{i=1}^{k} \mathfrak{G}_i = \max_{\mathfrak{X}_{1,k},\cdots,\mathfrak{X}_{n,k}} \prod_{g=1}^{k} f(x_{1,g},\cdots,x_{n,g}).$$

2) *Suppose we have*

$$f(sp(c_{1,y_{1,1}},\cdots,c_{1,y_{1,2}}),\cdots,sp(c_{n,y_{n,1}},\cdots,c_{n,y_{n,2}})) \leq f(c_{1,\mu},\cdots,c_{n,\mu})$$

then:

- *If* $\displaystyle\sum_{g=1}^{m} f(x_{1,g},\cdots,x_{n,g}) \geq \sum_{i=1}^{m} f(c_{1,i},\cdots,c_{n,i})$ *(sum min condition). Then*

$$\sum_{i=1}^{k} \mathfrak{O}_i = \min_{\mathfrak{X}_{1,k},\cdots,\mathfrak{X}_{n,k}} \sum_{g=1}^{k} f(x_{1,g},\cdots,x_{n,g}).$$

- *If f is positive valued and* $\prod_{g=1}^{m} f(x_{1,g}, \cdots, x_{n,g}) \geq \prod_{i=1}^{m} f(c_{1,i}, \cdots, c_{n,i})$ *(product min condition) we have*

$$\prod_{i=1}^{k} \mathfrak{D}_i = \min_{\mathfrak{x}_{1,k},\cdots,\mathfrak{x}_{n,k}} \prod_{g=1}^{k} f(x_{1,g}, \cdots, x_{n,g}).$$

Proof. We construct a family of systems (S_i) for all $i \leq n$ like we did for (S) in Lemma 2.1; for each i the m vectors of (S_i) are in $\mathcal{H}_{i,j_i}$, taking any $(x_{1,g}, \cdots, x_{n,g})$ in $(\mathfrak{X}_{1,k}, \cdots, \mathfrak{X}_{n,k})$ and assuming of course that $x_{l,g} \neq x_{l,g'}$ whenever $g \neq g'$, it is not difficult to adopt the proof of Lemma 2.1 to obtain:

$$\sum_{g=1}^{k} f(x_{1,g}, \cdots, x_{n,g}) \leq \sum_{g=1}^{k-1} f(z_{1,g}, \cdots, z_{n,g}) + f(z_{1,m}, \cdots, z_{n,m}) \tag{2.3}$$

$$\leq \sum_{g=1}^{k} f(z_{1,g}, \cdots, z_{n,g}). \tag{2.4}$$

which proves the sum maximum statement. For the sum minimum principle, by the same way we constructed (G) in the proof of Lemma 2.1 we construct n systems denoted by G_i, each G_i has its random initial k mutually orthogonal vectors completed by just $m-k$ vectors (particularly chosen) to form an orthogonal basis of dimension m in $\mathcal{H}_{i,j_i}$, hereupon the proof is straightforward. The product variational principles have also similar arguments, using the same idea with the systems $(S_i)_{i\leq n}$, $(G_i)_{i\leq n}$ and under stated conditions if the sums (for example in (2.3) and (2.4)) are replaced by products we get our desired characterizations. $\square$

Proposition 2.6. *Let $U \in \mathbb{M}_{n,k}$ such that $U^*U = I_k$ and let V be a $k \times k$ unitary matrices, then UV verifies $(UV)^*(UV) = I_k$.*

Corollary 2.7. *Let H be an $n \times n$ P.S.D. matrix (i.e. $\lambda_n \geq 0$) then for all $k \leq n$:*

$$\prod_{i=1}^{k} \lambda_i(H) = \max_{U^*U=I_k} \det(U^*HU) \tag{2.5}$$

$$\prod_{i=1}^{k} \lambda_{n-i+1}(H) = \min_{U^*U=I_k} \det(U^*HU) \tag{2.6}$$

Proof. The proof is a simple application of Theorem 2.5. By Proposition 2.6 we take our V the one that diagonalizes V^*U^*HUV - of course the matrix V is fixed after fixing U and that doesn't interfere with the value of the determinant - but this way we are seeking

$$\max_{x_i^* x_j = \delta_{ij}} \prod_{i=1}^{k} x_i^* H x_i \quad \text{resp.} \quad \min_{x_i^* x_j = \delta_{ij}} \prod_{i=1}^{k} x_i^* H x_i$$

since $\det(V^*U^*HUV) = \det(U^*HU) = \det(H)$ when U, V are unitaries, with $f_H(x) = x^*Hx$, $\mathcal{H}_{1,n} \equiv \mathbb{C}^n$ and $r = 1$ we verify easily that the conditions of

Theorem 2.5 are satisfied, making use of Proposition 1.1 resp. of Proposition 1.2 we get the required characterizations. □

2.3. **Some Extensions.**

Proposition 2.8. *Let k be a fixed integer and let $\alpha_{i,j}$ be any complex numbers with $i \le k$, $j = 1, 2$ such that $\sum_{i=1}^{k} |\alpha_{i,1}|^2 \le 1$ and $\sum_{i=1}^{k} |\alpha_{i,2}|^2 \le 1$ then*

$$\left| \sum_{i=1}^{k} \alpha_{i,1}\alpha_{i,2} \right| \le 1.$$

Proof. This is a direct consequence of the Cauchy-Schwartz inequality but a direct proof goes as follows: By the triangular inequality, it suffices to prove the proposition when all the complex numbers are nonnegative numbers, we proceed by induction when $k = 2$, we have $\alpha_{1,1}^2 + \alpha_{2,1}^2 \le 1$ and $\alpha_{1,2}^2 + \alpha_{2,2}^2 \le 1$, we will prove that

$$M := \sqrt{1 - \alpha_{2,1}^2}\sqrt{1 - \alpha_{2,2}^2} + \alpha_{2,1}\alpha_{2,2} \le 1,$$

and consequently we will have $\alpha_{1,1}\alpha_{1,2} + \alpha_{2,2}\alpha_{2,1} \le 1$, but if $w := \alpha_{2,2} - \alpha_{2,1}$ we get:

$$M \le 1 \iff (1 - \alpha_{2,1}^2)(1 - \alpha_{2,2}^2) \le (1 - \alpha_{2,1}\alpha_{2,2})^2 \tag{2.7}$$
$$\iff (1 - \alpha_{2,1})(1 + \alpha_{2,2})(1 - \alpha_{2,2})(1 + \alpha_{2,1}) \le (1 - \alpha_{2,1}\alpha_{2,2})^2 \tag{2.8}$$
$$\iff (1 - \alpha_{2,1}\alpha_{2,2} - w)(1 - \alpha_{2,1}\alpha_{2,2} + w) \le (1 - \alpha_{2,1}\alpha_{2,2})^2 \tag{2.9}$$
$$\iff (1 - \alpha_{2,1}\alpha_{2,2})^2 - w^2 \le (1 - \alpha_{2,2}\alpha_{2,1})^2. \tag{2.10}$$

Suppose the result true for $k = n$ let us prove it for $k = n + 1$, we will take $\sum_{i=1}^{n+1} |\alpha_{i,1}|^2 \le 1$ and $\sum_{i=1}^{n+1} |\alpha_{i,2}|^2 \le 1$ and set without loss of generality $s_1^2 = \alpha_{n,1}^2 + \alpha_{n+1,1}^2$ and $s_2^2 = \alpha_{n,2}^2 + \alpha_{n+1,2}^2$. By the induction hypothesis $\sum_{i=1}^{n-1} \alpha_{i,1}\alpha_{i,2} + s_1 s_2 \le 1$ but then it is easy to verify that

$$\sum_{i=1}^{n+1} \alpha_{i,1}\alpha_{i,2} \le \sum_{i=1}^{n-1} \alpha_{i,1}\alpha_{i,2} + s_1 s_2 \le 1,$$

thus $\left| \sum_{i=1}^{k} \alpha_{i,1}\alpha_{i,2} \right| \le \sum_{i=1}^{k} |\alpha_{i,1}\alpha_{i,2}| \le 1$, which is the desired result. □

Lemma 2.9. *Let k, h be two fixed integers and let $\alpha_{i,j}$ be any complex numbers with $i \le k$, $j \le h$ such that $\sum_{i=1}^{k} |\alpha_{i,j}|^2 \le 1$ for all j, then*

$$\left| \sum_{i=1}^{k} \alpha_{i,1} \cdots \alpha_{i,h} \right| \le 1.$$

Proof. The proof will follow from Proposition 2.8, by noticing that

$$\left|\sum_{i=1}^{k}\alpha_{i,1}\cdots\alpha_{i,h}\right| \le \sum_{i=1}^{k}|\alpha_{i,1}\cdots\alpha_{i,h}| \le \sum_{i=1}^{k}|\alpha_{i,1}\alpha_{i,2}| \le 1.$$

□

Example 2.10. Let $(\mathcal{H}_{1,m},\|.\|_1)$ respectively $(\mathcal{H}_{2,g},\|.\|_2)$ be two $\mathbb{C}$-Hilbert spaces of dimension m respectively of dimension g, with $m \le g$. Suppose that $P := (p_1,\cdots,p_m)$ is an orthonormal basis of $\mathcal{H}_{1,m}$ and $R := (e_1,\cdots,e_g)$ is an orthonormal basis of $\mathcal{H}_{2,g}$. For all $j \le m$, $l \neq l'$, if we have $x_{1,j} = \sum_{i=1}^{m}\alpha_{1,i,j}p_i$, $x_{2,j} = \sum_{i=1}^{g}\alpha_{2,i,j}e_i$ such that $\|x_{1,j}\|_1 = \|x_{2,j}\|_2 = 1$, $x_{1,l} \perp x_{1,l'}$ and $x_{2,l} \perp x_{2,l'}$ then the following holds

$$\left|\sum_{j=1}^{m}\alpha_{1,i,j}\alpha_{2,i,j}\right| \le 1,$$

for all i, $i \le m$, this is true because we can verify (by completing the set of vectors in each space into an orthonormal basis and associating to it the unitary change of basis matrix) that $\sum_{j=1}^{m}|\alpha_{h,i,j}|^2 \le 1$ for $h = 1,2$ and all i.

A triangular inequality can easily generalize the result if we consider three or more Hilbert spaces, for example letting m denote the least dimension of n Hilbert vector-spaces, for any i, $i \le m$ we can write:

$$\left|\sum_{j=1}^{m}\alpha_{1,i,j}\cdots\alpha_{n,i,j}\right| \le 1,$$

where $\{(\alpha_{s,1,j},\cdots,\alpha_{s,q,j}), j \le m\}$ are the coefficients of the m mutually orthogonal unit vectors $(x_{s,j})_{j\le m}$ written in any orthonormal basis of $\mathcal{H}_{s,q}$ (for some q), we leave the details to the reader.

As noticed in Theorem 2.5 the order at which we write the elements of the orthogonal basis is important, we associate to every basis $(x_1,\cdots,x_k)$ the group of permutation $\mathfrak{S}_k$.

Definition 2.11. Given a certain basis $A := (x_1,\cdots,x_k)$ of dimension k, the no fix permutation basis of A is $A' := (x_{\mathfrak{d}(1)},\cdots,x_{\mathfrak{d}(k)})$ where $\mathfrak{d}$ is a permutation of the k elements with no fixed point.

Keeping the terminology used in this section we can now state our main Lemma:

Lemma 2.12. *Let n be a fixed integer and let f_i be nonnegative real numbers for $i \le m$, ordered in decreasing order with $f_i = 0$ when $i > m$. Let f be the function*

$$\begin{aligned} f: B_{1_1}\times\cdots\times B_{1_n} &\to \mathbb{R}^+ \\ (x_1,\cdots,x_n) &\to \sum_i f_i\left|\langle x_1,s_{1,i}\rangle_{\mathcal{H}_{1,j_1}}\right|\left|\langle x_2,s_{2,i}\rangle_{\mathcal{H}_{2,j_2}}\right|\cdots\left|\langle x_n,s_{n,i}\rangle_{\mathcal{H}_{n,j_n}}\right|, \end{aligned}$$

where $(s_{l,1},\cdots,s_{l,k})$ *is an orthonormal basis of dimension* k *in* $\mathcal{H}_{l,j_l}$*, then* $\mathfrak{G}_i = f_i$*,* $\mathfrak{O}_i = 0$ *for all* i *and the function* f *verifies the following statements of Theorem 2.5:*

- $\sum_{g=1}^{m} f(x_{1,g},\cdots,x_{n,g}) \le \sum_{i=1}^{m} f(z_{1,i},\cdots,z_{n,i})$
- *For all* $k \le m$, $\sum_{i=1}^{k} \mathfrak{O}_i = \min_{\mathfrak{X}_{1,k},\cdots,\mathfrak{X}_{n,k}} \sum_{g=1}^{k} f(x_{1,g},\cdots,x_{n,g})$.

Proof. By writing any unit vector $x_l \in \mathcal{H}_{l,j_l}$ in terms of the corresponding $S_{l,j_l} := (s_{l,1},\cdots,s_{l,j_l})$ basis and by using Lemma 2.9 we get the required result. Note in this case that the basis $\mathfrak{Z}_{l,k} \equiv (z_{l,1},\cdots,z_{l,k})$ can be taken to be the arbitrarily chosen $S_{l,k} = (s_{l,1},\cdots,s_{l,k})$ and $\mathfrak{D}_{l,k} \equiv (c_{l,1},\cdots,c_{l,k})$ are also the $S_{l,k}$ except one, say $S_{n,k}$ which will be replaced by $S'_{n,k}$ (the no fix permutation basis). □

One can ask if the conditions of Theorem 2.5 are necessary and the answer is no as the next example shows:

Example 2.13. Let $A \in \mathbb{M}_{a,n}(\mathbb{C})$ fixed, $m := \min(a,n)$, the norm $\|.\|_s$ is the one associated to the usual scalar product on $\mathbb{C}^k$ for some k, if $A = W\Sigma V$ is the singular value decomposition, $W = [t_1 \cdots t_a]$, $V^* = [s_1 \cdots s_n]$ and Σ is the $a \times n$ matrix: $\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & 0 \\ 0 & \cdots & \sigma_{\mathfrak{r}} & \\ \hline & 0 & & 0 \end{array}\right]$ of rank $\mathfrak{r}$, then it can be verified that

$$A = \sum_{i=1}^{m} \sigma_i(A) t_i s_i^*, \tag{2.11}$$

(this is known as the *dyadic* decomposition of A, see [3] for details) and so:

$$x^*Ay = x^*W\Sigma V y = \sum_{i=1}^{m} \sigma_i(x^*t_i)(s_i^*y).$$

Let us define the continuous function f_A as:

$$\begin{array}{rcl} f_A:\ D & \to & \mathbb{R} \\ (x,y) & \to & |x^*Ay| = \left|\sum_{i=1}^{m} \sigma_i(x^*t_i)(s_i^*y)\right|, \end{array}$$

where $D := \left\{(x,y) \in \mathbb{C}^a \times \mathbb{C}^n; \|x\|_s = \|y\|_s = 1\right\}$.

Since $|x^*Ay| \le \sum_{i=1}^{m} \sigma_i \left|(x^*t_i)\right| \left|(s_i^*y)\right|$, by Lemma 2.12 for each k, $k = 1,\cdots,m$ we have:

$$\mathfrak{G}_1 = \sigma_1(A) = \max \frac{|x^*Ay|}{\|x\|\|y\|} \quad \textit{and} \quad \mathfrak{G}_k = \sigma_k(A) = \max_{\substack{x\in \mathrm{sp}\{t_1,\cdots,t_{k-1}\}^\perp \\ y\in \mathrm{sp}\{s_1,\cdots,s_{k-1}\}^\perp}} \frac{|x^*Ay|}{\|x\|\|y\|}; \text{ while}$$

it is easy to exhibit two vectors (x, y) such that

$$f_A(x, y) = f_A(sp(t_{y_{1,1}}, \cdots, t_{y_{1,2}}), sp(s_{y_{2,1}}, \cdots, s_{y_{2,2}})) \leq f_A(t_\mu, s_\mu),$$

it is well known -see [8]- that for all $k \leq m$:

$$\bullet \sum_{i=1}^{k} \mathfrak{G}_i = \max\left\{|trX^*AY| : X \in \mathbb{M}_{a,k}, Y \in \mathbb{M}_{n,k}, X^*X = I_k = Y^*Y\right\}$$

$$= \max_{B_k, C_k} \sum_{i=1}^{k} f_A(x_i, y_i) = \sum_{i=1}^{k} \sigma_i(A),$$

where x_i is a column of X, y_i a column of Y, $B_k = (x_1, \cdots, x_k)$ respectively $C_k = (y_1, \cdots, y_k)$ are any two orthonormal basis of dimension k in $\mathbb{C}^a$ respectively in $\mathbb{C}^n$.

2.4. **Courant-Fischer Theorem.** The notations here are those introduced in Subsection 2.2.

Theorem 2.14. *Let f be a continuous function from $D := B_{r_1} \times \cdots \times B_{r_n}$ into $\mathbb{R}$. Let $y_{i,1} \leq y_{i,2}$ for all $i \leq n$, $\mu = \left\{\max_i y_{i,2} / y_{i,2} \leq m\right\}$ and $k \leq m$.*
If we have:

$$[1] \quad f(sp(z_{1,y_{1,1}}, \cdots, z_{1,y_{1,2}}), \cdots, sp(z_{n,y_{n,1}}, \cdots, z_{n,y_{n,2}})) \geq f(z_{1,\mu}, \cdots, z_{n,\mu})$$

then:

$$\mathfrak{G}_k = \max_{\substack{dim(E_{l,k})=k,\\ \forall l, l \leq n.}} \min_{\substack{x_l \in (E_{l,k} \cap D_l),\\ \forall l, l \leq n.}} f(x_1, \cdots, x_n),$$

and if we have:

$$[2] \quad f(sp(c_{1,y_{1,1}}, \cdots, c_{1,y_{1,2}}), \cdots, sp(c_{n,y_{n,1}}, \cdots, c_{n,y_{n,2}})) \leq f(c_{1,\mu}, \cdots, c_{n,\mu})$$

then:

$$\mathfrak{Q}_k = \min_{\substack{dim(E_{l,k})=k,\\ \forall l, l \leq n.}} \max_{\substack{x_l \in (E_{l,k} \cap D_l),\\ \forall l, l \leq n.}} f(x_1, \cdots, x_n)$$

where $D_l = \{x \in \mathcal{H}_{l,j_l} / \|x\|_{\mathcal{H}_{l,j_l}} = r_l\}$, for all $l \leq n$.

Proof. For any fixed k and for all l, we have $E_{l,k} \cap sp(z_{l,k}, \cdots, z_{l,j_l}) \neq \phi$ respectively $E_{l,k} \cap sp(c_{l,k}, \cdots, c_{l,j_l}) \neq \phi$, which implies from [1] respectively [2] the two required characterizations. □

If A is any $h \times h$ Hermitian matrix, the case $n = 1$, $r = 1$, $\mathcal{H}_{1,h} \equiv \mathbb{C}^h$ and $f = R_A$ in the previous theorem gives the well known Courant-Fischer theorem.

Acknowledgment. I want to thank Prof. Roger Horn and Prof. Ajit Iqbal Singh for their helpful comments.

References

[1] F. Zhang, *Matrix Theory Basic Results and Techniques,* 2^{nd} *Edition*, Sec. 8.3, (Universitext) Springer, 2013.
[2] L. Mirsky, *Maximum principles in matrix theory*, Proc. Glasgow Math. Assoc. **4**, (1958), pp. 34-37.
[3] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*, Part I: Lectures 4-5, SIAM, null edition, 1997.
[4] K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations I.*, Proc. Nat. Acad. Sci. (U.S.A.) **35**, (1949), pp. 652-655.
[5] K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations II.*, Proc. Proc. Nat. Acad. Sci. (U.S.A.) **36**, (1950), pp. 31-35.
[6] K. Fan, *Maximum propreties and inequalities for the eigenvalues of completely continuous operators*, Proc. Nat. Acad. Sci. (U.S.A.) **37**, (1951), pp. 760-766.
[7] M. Marcus and B. N. Moyls, *On the maximum principle of Ky Fan*, Canad. J. Math. **9**, (1957), pp. 313-320.
[8] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Sec. 3.4, Cambridge University Press, New York, 1991.

[1] Fakra, Beirut, Lebanon.
E-mail address: tmhanat@yahoo.com