Functional Analysis,
Approximation and
Computation
10 (3) (2018), 27–35

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/faac

Some distances on subspaces

Shanwen Hu[a], Wei Jin[b], Yifeng Xue[c]

[a] *Research Center for Operator Algebras, East China Normal University, Shanghai 200241, P.R. China*
[b] *School of Mathematical Science, East China Normal University, Shanghai 200241, P.R. China*
[c] *School of Mathematical Science and Shanghai Key Laboratory of PMMP, East China Normal University, Shanghai 200241, P.R. China*

Abstract. Let H be an infinite separable complex Hilbert space. Let V be a closed subspace of H with $\dim V = n$ and let $MO_n(V)$ be the set of n–tuples $(e_1, \cdots, e_n)$ such that $e_1, \cdots, e_n$ are mutually orthonormal in V. In this short note, we first show that for an n–tuple of linearly independent unit vectors $(\alpha_1, \cdots, \alpha_n)$ in H, there is an n–tuple of mutually orthonormal vectors $(\gamma_1, \cdots, \gamma_n) = (\alpha_1, \cdots, \alpha_n)A^{-1/2}$ such that

$$\sum_{j=1}^{n} \|\alpha_j - \gamma_j\|^2 = \inf\Big\{\sum_{j=1}^{n} \|\alpha_j - e_j\|^2 \,|\, (e_1, \cdots, e_n) \in MO_n(H)\Big\},$$

where A is the Gram matrix of $(\alpha_1, \cdots, \alpha_n)$; then we define a new distance $d_1(M, N)$ for two subspaces M and N with $\dim M = \dim N < +\infty$ and give a formula to compute $d_2(M, N)$.

1. Introduction

Let H be an infinite separable complex Hilbert space and let $B(H)$ denote the C^*–algebra of all bounded operators on H. For $A \in B(H)$, let $\mathrm{Ker}A$ and $\mathrm{Ran}(A)$ denote the kernel and the range of A, respectively.

Let $S(H)$ denote the set of all closed subspaces in H and

$$S_n(H) = \{V \in S(H) |\ \dim V = n < +\infty\}.$$

Define the gap function $\delta(\cdot, \cdot) \colon S(H) \times S(H) \to \mathbb{R}$ by

$$\delta(S, T) = \begin{cases} 0 & V = \{0\} \\ \sup\{\mathrm{dist}(s, T) |\, s \in S, \|s\| = 1\} & V \neq \{0\} \end{cases},$$

where, $\mathrm{dist}(s, T) = \inf\{\|s - t\| |\, t \in T\}$. Put $\hat{\delta}(S, T) = \max\{\delta(S, T), \delta(T, S)\}$. Then $\hat{\delta}(\cdot, \cdot)$ is a distance on $S(H)$ and $\hat{\delta}(S, T) = \|P_S - P_T\|$, $\forall\, S, T \in S(H)$ (cf. [5]), where P_S (resp. P_T) is the orthogonal projection of H onto S (resp.

2010 *Mathematics Subject Classification*. Primary 47A05; Secondary 15A03.
Keywords. Distance; Subspaces; Trace; Polar decomposition.
Received: 23 December 2017; Accepted: 15 November 2018
Communicated by Dragan S. Djordjević
Research supported by NSFC (No. 11531003) and STCSM (No. 18dz2271000)
Email addresses: swhu@math.ecnu.edu.cn (Shanwen Hu), jin_wei5@163.com (Wei Jin), yfxue@math.ecnu.edu.cn (Yifeng Xue)

T). This distance and the gap function play a very important role in the study of perturbation of operators on H. Please see [5] for details. For $V \in S(H)$ with $\dim V = n$, set

$$MO_n(V) = \{(e_1, \cdots, e_n) | \, e_1, \cdots, e_n \text{ are mutually orthonormal in } V\},$$
$$LI_n(V) = \{(\alpha_1, \cdots, \alpha_n) | \, \alpha_1, \cdots, \alpha_n \text{ are linearly independent in } V \text{ with } \|\alpha_i\| = 1\}.$$

There is another distance on $S_n(H)$ which is a useful tool in Statistics, given by

$$d_0(U, V) = \sqrt{\sum_{i=1}^{n} d^2(e_i, V)}, \quad \forall\, U, V \in S_n(H),$$

where $(e_1, \cdots, e_n) \in MO_n(U)$ (cf. [4], [3], [6]). From [6], $d_0(U, V)$ is independence of the choice of $(e_1, \cdots, e_n)$ and

$$\begin{aligned} d_0(U, V) &= \sqrt{\sum_{i=1}^{n} \|e_i - P_V e_i\|^2} = \sqrt{n - \sum_{i=1}^{n} (P_V P_U e_i, P_U e_i)} \\ &= \sqrt{\mathrm{Tr}(P_U - P_U P_V P_U)}, \end{aligned}$$

where P_U (resp. P_V) is the orthogonal projection of H onto U (resp. V) and $\mathrm{Tr}(\cdot)$ is the canonical trace on the set

$$\mathcal{F}(H) = \{T \in B(H) | \, \dim \mathrm{Ran}(T) < +\infty\}.$$

Note that $<A, B>= \mathrm{Tr}(AB^*)$ defines an inner product on $\mathcal{F}(H)$ of A, B on $\in \mathcal{F}(H)$ and $\|A\|_{\mathrm{Tr}} = \sqrt{\mathrm{Tr}(AA^*)}$ is a norm of A on $\mathcal{F}(H)$.

Let $\hat{\xi} = (\xi_1, \cdots . \xi_n)$ and $\hat{\gamma} = (\gamma_1, \cdots, \gamma_n)$ be two n–tuples of vectors in H. Define the distance between $\hat{\xi}$ and $\hat{\gamma}$ by

$$d(\hat{\xi}, \hat{\gamma}) = \sqrt{\sum_{j=1}^{n} \|\xi_j - \gamma_j\|^2}.$$

We now define two distances on $S_n(H)$ by

$$d_1(M, N) = \inf\{d(\hat{\xi}, \hat{\gamma}) | \, \hat{\xi} \in MO_n(M), \ \hat{\gamma} \in MO_n(N)\},$$
$$d_2(M, N) = \inf\{d(\hat{\xi}, \hat{\gamma}) | \, \hat{\xi} \in LI_n(M), \ \gamma \in LI_n(N)\}.$$

2. Distance to $MO_n(H)$

Let $\hat{\alpha} = (\alpha_1, \cdots, \alpha_n) \in LI_n(H)$. In this section, we will compute the distance

$$d(\hat{\alpha}, MO_n(H)) = \inf\{d(\hat{\alpha}, \hat{\xi}) | \, \hat{\xi} \in MO_n(H)\}.$$

Since $\hat{\alpha} \in LI_n(H)$, the Gram matrix $A = G(\hat{\alpha}) = ((\alpha_j, \alpha_i))_{n \times n}$ is definite–positive. By using Gram Schmidt process to $(\alpha_1, \cdots, \alpha_n)$, we get an n–tuple of orthonormal vectors $(\beta_1, \cdots, \beta_n)$. Then there is a upper triangular and non–singular matrix X of order n such that $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)X$. Thus, $A = G(\hat{\alpha}) = (XX^*)^{-1}$.

Now let $X^* = U(XX^*)^{-1/2}$ be the polar decomposition of X^*. Then

$$(\beta_1, \cdots, \beta_n)U = (\alpha_1, \cdots, \alpha_n)(XX^*)^{-1/2} = (\alpha_1, \cdots, \alpha_n)A^{-1/2}. \tag{1}$$

Put $(\gamma_1, \cdots, \gamma_n) = (\alpha_1, \cdots, \alpha_n)A^{-1/2}$. Then $(\gamma_1, \cdots, \gamma_n) \in MO_n(H)$ by (1).

Define orthogonal projections on H by $P_i x = (x, \alpha_i)\alpha_i$, $\forall\, x \in H$, $i = 1, \cdots, n$. Set $P = \sum\limits_{i=1}^{n} P_i$. Then P is a bounded linear operator of finite rank. Moreover,

$$P(\alpha_1, \cdots, \alpha_n) = (P\alpha_1, \cdots, P\alpha_n) = (\alpha_1, \cdots, \alpha_n)A. \tag{2}$$

Let I be the identity operator on H and I_n be the identity matrix of order n. Then by (2), we have for any $\lambda > 0$,

$$P(\lambda I + P^2)^{-1}(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A(\lambda I_n + A^2)^{-1}. \tag{3}$$

Since the range of P is closed in H, it follows from [5, Proposition 3.1.7] that P has the Moore–Penrose inverse $P^\dagger$, that is, $P^\dagger$ is a bounded linear and self–adjoint operator which satisfies following conditions:

$$PP^\dagger P = P,\ P^\dagger PP^\dagger = P^\dagger,\ P^\dagger P = PP^\dagger.$$

Then by [5, Corollary 4.2.9] and (3), we have

$$P^\dagger(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A^{-1}. \tag{4}$$

Sine $t^{1/2}$ can be approximated unifirmly by a sequence of polynomials on $[0, n]$, we have by (2) and (4),

$$\begin{aligned} P^{1/2}(\alpha_1, \cdots, \alpha_n) &= (\alpha_1, \cdots, \alpha_n)A^{1/2}, \\ (P^\dagger)^{1/2}(\alpha_1, \cdots, \alpha_n) &= (\alpha_1, \cdots, \alpha_n)A^{-1/2} = (\gamma_1, \cdots, \gamma_n). \end{aligned}$$

Therefore, $P^{1/2}(\gamma_1, \cdots, \gamma_n) = (\alpha_1, \cdots, \alpha_n)$. So we have

Lemma 2.1. *Let $(\alpha_1, \cdots, \alpha_n) \in LI_n(H)$ and let P, A be as above. Then there exists $(\gamma_1, \cdots, \gamma_n) \in MO_n(H)$ such that*

$$P(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A,\ P^{1/2}(\gamma_1, \cdots, \gamma_n) = (\alpha_1, \cdots, \alpha_n),$$

Lemma 2.2. *Let P and A be as above. Then*

(1) $\mathrm{Tr}(P^{1/2}) = \mathrm{Tr}_n(A^{1/2})$;
(2) $\sigma(P)\backslash\{0\} = \sigma(A)$, *where $\sigma(P)$ and $\sigma(A)$ are the spectrum of P and A, respectively.*

Proof. (1) Set $M = \mathrm{span}\{\alpha_1, \cdots, \alpha_n\}$. Clearly, $M = \mathrm{span}\{\gamma_1, \cdots, \gamma_n\} = \mathrm{Ran}(P^{1/2})$. Let $\{\eta_i\}_{i=1}^\infty$ be the orthonormal basis for $M^\perp$. Then $P^{1/2}\eta_i = 0$ for $i \geq 1$. Thus,

$$\mathrm{Tr}(P^{1/2}) = \sum_{j=1}^{n}(P^{1/2}\gamma_j, \gamma_j) + \sum_{j=1}^{\infty}(P^{1/2}\eta_j, \eta_j) = \sum_{j=1}^{n}(\alpha_j, \gamma_j).$$

Put $A^{1/2} = (c_{ij})_{n\times n}$. Since $(\alpha_1, \cdots, \alpha_n) = (\gamma_1, \cdots, \gamma_n)A^{1/2}$, it follows that $(\alpha_j, \gamma_j) = c_{jj}$, $j = 1, \cdots, n$ and consequently,

$$\mathrm{Tr}(P^{1/2}) = \sum_{j=1}^{n}(\alpha_j, \gamma_j) = \sum_{j=1}^{n} c_{jj} = \mathrm{Tr}_n(A^{1/2}).$$

(2) Let $\lambda \notin \sigma(P)$. Then there is $C \in B(H)$ such that $C(P - \lambda I) = (P - \lambda I)C = I$. Thus, from $P(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A$, we get that

$$(\alpha_1, \cdots, \alpha_n) = C(\alpha_1, \cdots, \alpha_n)(A - \lambda I_n). \tag{5}$$

Let $\begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix} \in \mathbb{C}^n$ such that $(A - \lambda I_n)\begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix} = 0$. Then by (5),

$$\mu_1\alpha_1 + \cdots + \mu_n\alpha_n = 0. \tag{6}$$

Since $\alpha_1, \cdots, \alpha_n$ are linearly independent, it follows from (6) that $\mu_j = 0$, $j = 1, \cdots, n$ and so that $\mathrm{Ker}(A - \lambda I_n) = 0$. This means that $\sigma(A) \subset \sigma(P)\backslash\{0\}$.

On the other hand, let $\lambda \in \sigma(P)\backslash\{0\}$. Choose $\xi \in H\backslash\{0\}$ such that $(P - \lambda I)\xi = 0$. Note that $\xi \in \mathrm{Ran}(T) \subset \mathrm{span}\{\alpha_1, \cdots, \alpha_n\}$. So there are $k_1, \cdots, k_n \in \mathbb{C}$ such that

$$\xi = (\alpha_1, \cdots, \alpha_n)\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}, \quad |k_1|^2 + \cdots + |k_n|^2 \neq 0.$$

Then from

$$(P - \lambda I)(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)(A - \lambda I_n)$$

we get that $(A - \lambda I_n)\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = 0$. Thus $\lambda \in \sigma(A)$ and hence $\sigma(A) \subset \sigma(P)\backslash\{0\}$. □

Lemma 2.3. *Let D, $W \in B(H)$ satisfy conditions: $D \geq 0$ and $\|W\| \leq 1$. Then* $|\mathrm{Tr}(DW)| \leq \mathrm{Tr}(D)$.

Proof. Since $< X, Y >= \mathrm{Tr}(XY^*)$, $\forall\, X, Y \in \mathcal{F}(H)$ defines an inner product on $\mathcal{F}(H)$, we have

$$|\mathrm{Tr}(DW)|^2 = |\mathrm{Tr}(D^{1/2}(W^*D^{1/2})^*)|^2 \leq \mathrm{Tr}(D)\mathrm{Tr}(D^{1/2}WW^*D^{1/2}). \tag{7}$$

Since $WW^* \leq \|WW^*\|I \leq I$, we get that $\mathrm{Tr}(D^{1/2}WW^*D^{1/2}) \leq \mathrm{Tr}(D)$ and hence $|\mathrm{Tr}(DW)| \leq \mathrm{Tr}(D)$. □

Theorem 2.4. *Let $(\alpha_1, \cdots, \alpha_n)$, $(\gamma_1, \cdots, \gamma_n)$ and P, A be as above. Then*

$$\inf\{\sum_{j=1}^{n} \|\alpha_j - \xi_j\|^2 \,|\, (\xi_1, \cdots, \xi_n) \in MO_n(H)\} = \sum_{j=1}^{n} \|\alpha_j - \gamma_j\|^2 = 2n - 2\mathrm{Tr}_n(A^{1/2}).$$

Proof. Let M and $\{\eta_j\}_{j=1}^{\infty}$ be as in the proof of Lemma 2.2. For any $(\xi_1, \cdots, \xi_n) \in MO_n(H)$, set $N = \mathrm{span}\{\xi_1, \cdots, \xi_n\}$. Let $\{\xi'_j\}_{j=1}^{\infty}$ be an orthonormal basis for $N^{\perp}$. Then there is a unitary operator U on H such that $U\gamma_i = \xi_i$, $i = 1, \cdots, n$ and $U\eta_j = \xi'_j$, $j \geq n + 1$. Thus,

$$\begin{aligned}\sum_{j=1}^{n} \|\alpha_j - \xi_j\|^2 &= 2n - \sum_{j=1}^{n} [(\alpha_j, \xi_j) + (\xi_j, \alpha_j)] \\ &= 2n - \sum_{j=1}^{n} ([U^*P^{1/2} + P^{1/2}U]\gamma_j, \gamma_j) - \sum_{j=1+n}^{\infty} ([U^*P^{1/2} + P^{1/2}U]\eta_j, \eta_j) \\ &= 2n - \mathrm{Tr}(U^*P^{1/2} + P^{1/2}U).\end{aligned}$$

Note that

$$\mathrm{Tr}(U^*P^{1/2} + P^{1/2}U) = 2\mathrm{Tr}(P^{1/2}\mathrm{Re}(U)) \leq 2\mathrm{Tr}(P^{1/2}) = 2\mathrm{Tr}_n(A^{1/2})$$

by Lemma 2.2 for $\|\mathrm{Re}(U)\| = \frac{1}{2}\|U^* + U\| \leq 1$. Thus,

$$\inf\{\sum_{j=1}^{n} \|\alpha_j - \xi_j\|^2 \,|\, (\xi_1, \cdots, \xi_n) \in MO_n(H)\} \geq 2n - 2\mathrm{Tr}_n(A^{1/2}).$$

Now using the relation $(\alpha_1, \cdots, \alpha_n) = (\gamma_1, \cdots, \gamma_n)A^{1/2}$, we get that

$$\sum_{j=1}^{n} \|\alpha_j - \gamma_j\|^2 = 2n - \sum_{j=1}^{n} [(\alpha_j, \gamma_j) + (\gamma_j, \alpha_j)] = 2n - 2\mathrm{Tr}_n(A^{1/2}).$$

The statement is proved. □

Remark 2.5. Let $\hat{\alpha} = (\alpha_1, \cdots, \alpha_n)$ be an n–tuple of linearly independent unit vectors in H and let $A = ((\alpha_j, \alpha_i))_{n\times n}$ be the Gram matrix. By applying Gram–Schmidt process to $\alpha_1, \cdots, \alpha_n$, we get an n–tuple $\hat{\beta} = (\beta_1, \cdots, \beta_n)$ of orthonormal vectors. Put $\hat{\gamma} = (\gamma_1, \cdots, \gamma_n) = (\alpha_1, \cdots, \alpha_n)A^{-1/2} \in MO_n(H)$. Theorem 2.4 says

$$d(\hat{\alpha}, \hat{\gamma}) = d(\hat{\alpha}, MO_n(H)) \le d(\hat{\alpha}, \hat{\beta}),$$

that is, $\hat{\gamma}$ has least square property.

For example. Let $H = L^2([0,1])$. Set $\alpha_1(t) = 1$, $\alpha_2(t) = \sqrt{3}\,t$, $\forall\, t \in [0,1]$. Then

$$A = \begin{pmatrix} 1 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 1 \end{pmatrix},\ \beta_1(t) = 1,\ \beta_2(t) = \sqrt{12}(t - \frac{1}{2}),\ \forall\, t \in [0,1]$$

and $A^{1/2} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$, $A^{-1/2} = \begin{pmatrix} \sqrt{3} & -1 \\ -1 & \sqrt{3} \end{pmatrix}$. So from $(\gamma_1, \gamma_2) = (\alpha_1, \alpha_2)A^{-1/2}$, we get that $\gamma_1(t) = \sqrt{3}(1-t)$, $\gamma_2(t) = 3t - 1$, $\forall\, t \in [0,1]$. Consequently,

$$\sum_{j=1}^{2} \|\alpha_j - \gamma_j\|^2 = 2(2 - \sqrt{3}),\ \sum_{j=1}^{2} \|\alpha_j - \beta_j\|^2 = 1,\ \sum_{j=1}^{2} \|\alpha_j - \beta_{3-j}\|^2 = 4 - \sqrt{3}.$$

In [2, Corollary 3.10], authors proved that if $(\alpha_1, \cdots, \alpha_n)$ is an n–tuple of linear independent unit vectors in H with $|(\alpha_i, \alpha_j)| < \frac{\epsilon}{2(n-1)}$, $1 \le i < j \le n$, where $n \ge 2$ and $\epsilon \in (0,1)$. Then is an n–tuple of mutually orthogonal unit vectors $(\gamma_1, \cdots, \gamma_n)$ in H such that $\|\alpha_i - \gamma_i\| < \epsilon$.

The above result can be improved as the following:

Corollary 2.6. *Let $\hat{\alpha} = (\alpha_1, \cdots, \alpha_n)$ be an n–tuple of linear independent unit vectors in H with $|(\alpha_i, \alpha_j)| < \frac{\epsilon}{n-1}$, $1 \le i < j \le n$ for $\epsilon \in (0,1)$. Then is an n–tuple of mutually orthogonal unit vectors $\hat{\gamma} = (\gamma_1, \cdots, \gamma_n)$ in H such that $\|\alpha_i - \gamma_i\| < \epsilon$, $i = 1, \cdots, n$ and $d(\hat{\alpha}, \hat{\gamma}) = d(\hat{\alpha}, MO_n(H))$.*

Proof. Let A and P be as in Theorem 2.4. By Lemma 2.1, there is $\hat{\gamma} = (\gamma_1, \cdots, \gamma_n) \in MO_n(H)$ such that $P^{1/2}(\gamma_1, \cdots, \gamma_n) = (\alpha_1, \cdots, \alpha_n)$. Thus, $d(\hat{\alpha}, \hat{\gamma}) = d(\hat{\alpha}, MO_n(H))$ by Theorem 2.4 and $\|\alpha_i - \gamma_i\| \le \|I - P^{1/2}\|$, $i = 1, \cdots, n$.

According to Gerŝgorin' theorem, the spectrum of $A = ((\alpha_j, \alpha_i))_{n\times n}$ is contained in $\{\lambda \in \mathbb{R} \,|\, |1 - \lambda| \le \epsilon\}$ for $|(\alpha_i, \alpha_j)| < \frac{\epsilon}{n-1}$, $1 \le i < j \le n$. By Lemma 2.2 (2), $\sigma(P)\backslash\{0\} = \sigma(A) \subset (1-\epsilon, 1+\epsilon)$. Thus,

$$\|\alpha_i - \gamma_i\| \le \|I - P^{1/2}\| \le 1 - \sqrt{1-\epsilon} < \epsilon, \quad i = 1, \cdots, n.$$

□

3. The distance between $MO_n(M)$ and $MO_n(N)$

In this section, we will compute the distance $d_1(M, N)$ for $M, N \in S_n(H)$.

Let $U(H)$ denote the set of all unitary operators in $B(H)$. For a closed subspace K of H, let P_K denote the orthogonal projection of H onto K.

Theorem 3.1. *Let M, $N \in S_n(H)$. Then*

$$\begin{aligned} d_1(M, N) &= \inf\{ \sqrt{\mathrm{Tr}((P_M - P_N U)^*(P_M - P_N U))} \,|\, U \in U(H) \text{ with } UM = N\} \\ &= \sqrt{2\mathrm{Tr}(P_M - (P_M P_N P_M)^{1/2})}. \end{aligned}$$

Proof. Let $(e_1, \cdots, e_n, e_{n+1}, \cdots)$ be an orthonormal basis for H with $(e_1, \cdots, e_n) \in MO_n(M)$. Let $U \in U(H)$ with $UM = N$. Put $f_n = Ue_n$, $n = 1, 2, \cdots$. Then $(f_1, \cdots, f_n, \cdots)$ be an orthonormal basis for H with $(f_1, \cdots, f_n) \in MO_n(N)$. Thus,

$$\begin{aligned}\mathrm{Tr}((P_M - P_N U)^*(P_M - P_N U)) &= \sum_{i=1}^{n}((P_M - P_N U)^*(P_M - P_N U)e_i, e_i) \\ &\quad + \sum_{i=n+1}^{\infty}((P_M - P_N U)^*(P_M - P_N U)e_i, e_i) \\ &= \sum_{i=1}^{n}[(e_i, e_i) - (e_i, f_i) - (f_i, f_i) + (f_i, f_i)] \\ &= \sum_{i=1}^{n}\|e_i - f_i\|^2\end{aligned}$$

and hence

$$d_1(M, N) \leq \inf\{\sqrt{\mathrm{Tr}((P_M - P_N U)^*(P_M - P_N U))}\,|\, U \in U(H) \text{ with } UM = N\}.$$

On the other hand, let $(e_1, \cdots, e_n, e_{n+1}, \cdots)$ (resp. $(f_1, \cdots, f_n, f_{n+1}, \cdots)$) be an orthonormal basis for H (resp. N) with $(e_1, \cdots, e_n) \in MO_n(M)$ (resp. $(f_1, \cdots, f_n) \in MO_n(N)$). Define a unitary operator W on H such that $We_n = f_n$, $n = 1, 2, \cdots$. Then $WM = N$ and

$$\begin{aligned}\sum_{i=1}^{n}\|e_i - f_i\|^2 &= \sum_{i=1}^{\infty}((P_M - P_N W)e_i, (P_M - P_N W)e_i) \\ &= \mathrm{Tr}((P_M - P_N W)^*(P_M - P_N W))\end{aligned}$$

for $P_M e_i = P_N f_i = 0$ when $i > n$. So we have

$$d_1(M, N) \leq \inf\{\sqrt{\mathrm{Tr}((P_M - P_N U)^*(P_M - P_N U))}\,|\, U \in U(H) \text{ with } UM = N\}.$$

Let $P_N P_M = V(P_M P_N P_M)^{1/2}$ be the polar decomposition of $P_N P_M$ on H, where $\mathrm{Ran}(V) = \mathrm{Ran}(P_N P_M)$ and $\mathrm{Ran}(V^*) = \mathrm{Ran}(P_M P_N)$ (cf. [5, Proposition 1.5.16]). Since $\dim M = \dim N = n$, $\dim(\mathrm{Ran}(V)) = \dim(\mathrm{Ran}(V^*))$ and

$$M = \mathrm{Ran}(V^*) \oplus [(M \cap (\mathrm{Ran}(V^*))^{\perp}], \quad N = \mathrm{Ran}(V) \oplus [N \cap (\mathrm{Ran}(V))^{\perp}],$$

we have $\dim[(M \cap (\mathrm{Ran}(V^*))^{\perp}] = \dim[(N \cap (\mathrm{Ran}(V))^{\perp}]$. Consequently, there is a unitary operator W on H such that $P_N P_M = W(P_M P_N P_M)^{1/2}$ and $WM = N$. By using the identity $P_N P_M = W(P_M P_N P_M)^{1/2}$ with $W \in U(H)$ and $WM = N$, we get that

$$\mathrm{Tr}((P_M - P_N W)^*(P_M - P_N W)) = 2\,\mathrm{Tr}(P_M - (P_M P_N P_M)^{1/2}) \tag{8}$$

and for any $U \in U(H)$ with $UM = N$,

$$\mathrm{Tr}((P_M - P_N U)^*(P_M - P_N U)) = 2\,\mathrm{Tr}(P_M - \frac{1}{2}(W^*U + U^*W)(P_M P_N P_M)^{1/2}). \tag{9}$$

Note that $\|\frac{1}{2}(W^*U + U^*W)\| \leq 1$ and $\frac{1}{2}(W^*U + U^*W)$ is self–adjoint. It follows from Lemma 2.3 that

$$\mathrm{Tr}(\frac{1}{2}(W^*U + U^*W)(P_M P_N P_M)^{1/2}) \leq \mathrm{Tr}((P_M P_N P_M)^{1/2}). \tag{10}$$

Finally, combining (8), (9) with (10), we obtain that

$$\begin{aligned}&\inf\{\sqrt{\mathrm{Tr}((P_M - P_N U)^*(P_M - P_N U))}\,|\, U \in U(H) \text{ with } UM = N\} \\ &\qquad = \sqrt{2\mathrm{Tr}(P_M - (P_M P_N P_M)^{1/2})}.\end{aligned}$$

□

Corollary 3.2. *$d_1(\cdot,\cdot)$ is a metric on $S_n(H)$.*

Proof. Clearly, $d_1(M,N)=0$ iff $M=N$ and $d_1(M,N)=d_1(N,M)$ for $M,N\in S_n(H)$. Now let $M,N,O\in S_n(H)$. Then for any $U,V\in U(H)$ with $UM=N$, $VN=O$, we have by Theorem 3.1,

$$\begin{aligned} d_1(M,O) &\le \mathrm{Tr}^{1/2}(P_M-P_O(VU))^*(P_M-P_O(VU)) = \|P_M-P_O(VU)\|_{\mathrm{Tr}} \\ &= \|(P_N-P_OV)U+P_M-P_NU\|_{\mathrm{Tr}} \le \|P_N-P_OV\|_{\mathrm{Tr}}+\|P_M-P_NU\|_{\mathrm{Tr}} \end{aligned}$$

Thus, by using Theorem 3.1 again, we get that

$$d_1(M,N)\le d_1(N,O)+d_1(M,N).$$

□

Corollary 3.3. *Let $M,N\in S_n(H)$. Then*

$$\frac{1}{\sqrt{2}}\delta(M,N)\le d_0(M,N)\le d_1(M,N)\le \sqrt{2}\,d_0(M,N)\le \sqrt{2n}\,\delta(M,N).$$

Proof. Since $P_MP_NP_M\le I_H$, it follows that $P_MP_NP_M\le (P_MP_NP_M)^{1/2}$. So

$$\mathrm{Tr}(P_M-P_MP_NP_M)\ge \mathrm{Tr}(P_M-(P_MP_NP_M)^{1/2})$$

and hence $d_0(M,N)\ge \dfrac{1}{\sqrt{2}}d_1(M,N)$ by Theorem 3.1.

Note that $k=\dim\mathrm{Ran}(P_MP_NP_M)\le n$. $P_MP_NP_M$ has eigenvalues $\lambda_1,\cdots,\lambda_t$ with $t\le k$ and $\lambda_j>0$, $j=1,\cdots,t$. From

$$\lambda_1^{1/2}+\cdots+\lambda_t^{1/2}\le \sqrt{t}(\lambda_1+\cdots+\lambda_t^{1/2})$$

we get that $\mathrm{Tr}((P_MP_NP_M)^{1/2})\le \sqrt{n}\sqrt{\mathrm{Tr}(P_MP_NP_M)}$. Thus

$$\begin{aligned} \mathrm{Tr}(P_M-(P_MP_NP_M)^{1/2}) &\ge \sqrt{n}(\sqrt{n}-\sqrt{\mathrm{Tr}(P_MP_NP_M)}) \\ &= \sqrt{n}\frac{n-\mathrm{Tr}(P_MP_NP_M)}{\sqrt{n}+\sqrt{\mathrm{Tr}(P_MP_NP_M)}} \\ &\ge \frac{1}{2}\mathrm{Tr}(P_M-P_MP_NP_M) \end{aligned}$$

and furthermore, $d_1(M,N)\ge d_0(M,N)$.

By the definition of $d_0(M,N)$ and $\delta(M,N)$, we have $d_0(M,N)\le \sqrt{n}\,\delta(M,N)$. Now choose $(e_1,\cdots,e_n)\in MO_n(M)$ and $(f_1,\cdots,f_n)\in MO_n(N)$ such that

$$d_1(M,N)=\sqrt{\sum_{i=1}^n\|e_i-f_i\|^2}.$$

For any $\xi\in M$ with $\|\xi\|=1$, we can find $\lambda_1,\cdots,\lambda_n\in\mathbb{C}$ such that $\xi=\sum_{i=1}^n\xi_ie_i$. So $\sum_{i=1}^n|\lambda_i|^2=1$ and

$$d(\xi,N)\le \|\sum_{i=1}^n\lambda_ie_i-\sum_{i=1}^n\lambda_if_i\|\le \sqrt{\sum_{i=1}^n\|e_i-f_i\|^2}=d_1(M,N).$$

Finally, $\delta(M,N)\le d_1(M,N)\le \sqrt{2}\,d_0(M,N)$. □

4. The distance between $LI_n(M)$ and $LI_n(N)$

In this section, we will compute the distance $d_2(M,N)$ for $M,N \in S_n(H)$.

Lemma 4.1. *Let $V \in S_n(H)$. Then for any $\epsilon > 0$ and any $\alpha \in V$ with $\|\alpha\| = 1$, there are linearly independent vectors $\alpha_1, \cdots, \alpha_n$ in V with $\alpha_1 = \alpha$ and $\|\alpha_i\| = 1$, $i = 2, \cdots, n$ such that $\|\alpha_i - \alpha\| < \epsilon$, $i = 2, \cdots, n$.*

Proof. Choose $\beta_2, \cdots, \beta_n$ in V such that $\|\beta_i\| = \epsilon$, $i = 2, \cdots, n$ and $\{\alpha, \beta_2, \cdots, \beta_n\}$ is mutually orthogonal. Put

$$\alpha_1 = \alpha, \quad \alpha_i = (\alpha + \beta_i)/\|\alpha + \beta_i\|, \quad i = 2, \cdots, n.$$

Then $\alpha_1, \cdots, \alpha_n$ are linearly independent and

$$\begin{aligned}\|\alpha_i - \alpha\|^2 &= \frac{\|(1 - \|\alpha + \beta_i\|)\alpha + \beta_i\|^2}{\|\alpha + \beta_i\|^2} = 2\,\frac{1 - \sqrt{1+\epsilon^2} + \epsilon^2}{1 + \epsilon^2} \\ &= \frac{2\epsilon^2}{\sqrt{1+\epsilon^2}(1 + \sqrt{1+\epsilon^2})} < \epsilon^2, \quad i \geq 2\end{aligned}$$

and hence $\|\alpha_i - \alpha\| < \epsilon$, $i = 2, \cdots, n$. □

Let V_1, V_2 be two closed subspaces in H with $V_1 \cap V_2 = \{0\}$. Then by Definition 9.4 and Lemma 9.5 in [1], we have

$$\sup\{|(x,y)| \mid x \in V_1, y \in V_2, \|x\| = \|y\| = 1\} = \|P_{V_1}P_{V_2}\|.$$

Thus we have

Lemma 4.2. *Let V_1, V_2 be two closed subspaces in H with $V_1 \cap V_2 = \{0\}$. Then*

$$\inf\{\|x - y\| \mid x \in V_1, y \in V_2, \|x\| = \|y\| = 1\} = \sqrt{2 - 2\|P_{V_1}P_{V_2}\|}.$$

Theorem 4.3. *Let $M,N \in S_n(H)$. Then*

$$d_2(M,N) = \begin{cases} \sqrt{2n}\sqrt{1 - \|P_M P_N\|} & M \cap N = \{0\} \\ 0 & M \cap N \neq \{0\} \end{cases}.$$

Proof. We first suppose that $M \cap N = \{0\}$.

Let $\Xi = (\xi_1, \cdots, \xi_n) \in M$ and $\Gamma = (\gamma_1, \cdots, \gamma_n) \in N$. Then by Lemma 4.2,

$$\begin{aligned} d(\Xi, \Gamma) &= \sqrt{\sum_{j=1}^{n} \|\xi_j - \gamma_j\|^2} \\ &\geq \sqrt{n}\min\{\|x - y\| \mid x \in M, y \in N, \|x\| = \|y\| = 1\} \\ &= \sqrt{2n}\sqrt{1 - \|P_M P_N\|} \end{aligned}$$

and so that $d_2(M,N) \geq \sqrt{2n}\sqrt{1 - \|P_M P_N\|}$.

Choose $x_0 \in M$ and $y_0 \in N$ with $\|x_0\| = \|y_0\| = 1$ such that

$$\|x_0 - y_0\| = \min\{\|x - y\| \mid x \in M, y \in N, \|x\| = \|y\| = 1\}.$$

For any $\epsilon > 0$, by Lemma 4.1, there are unit vectors $x_1, \cdots, x_{n-1}$ in M and unit vectors $y_1, \cdots, y_{n-1}$ in N such that $\{x_0, x_1, \cdots, x_{n-1}\}$ and $\{y_0, y_1, \cdots, y_{n-1}\}$ are linearly independent, respectively, and $\|x_i - x_0\| < \epsilon$, $\|y_i - y_0\| < \epsilon$, $i = 1, \cdots, n-1$. Therefore

$$d_2^2(M,N) \leq \sum_{j=0}^{n-1} \|x_j - y_j\|^2 < \|x_0 - y_0\|^2 + (n-1)(2\epsilon + \|x_0 - y_0\|)^2.$$

Let $\epsilon \to 0^+$ and then we have, by Lemma 4.2,

$$d_2(M,N) \le \sqrt{n}\|x_0 - y_0\| = \sqrt{2n}\sqrt{1 - \|P_M P_N\|}.$$

Now assume that $M \cap N \ne \{0\}$. Take $\alpha \in M \cap N$ with $\|\alpha\| = 1$. Then for any $\epsilon > 0$, there are unit vectors $\alpha_1, \cdots, \alpha_{n-1} \in M$ and $\beta_1, \cdots, \beta_{n-1} \in N$ such that $\{\alpha, \alpha_1, \cdots, \alpha_{n-1}\}$ and $\{\alpha, \beta_1, \cdots, \beta_{n-1}\}$ are linearly independent, respectively, and $\|\alpha_i - \alpha\| < \epsilon$, $\|\beta_i - \alpha\| < \epsilon$, $i = 1, \cdots, n-1$, by Lemma 4.2. Thus,

$$d_2^2(M,N) \le \sum_{j=1}^{n-1} \|\alpha_j - \beta_j\|^2 \le \sum_{j=1}^{n-1} (\|\alpha_j - \alpha\| + \|\beta_j - \alpha\|)^2 \le 4(n-1)\epsilon^2$$

and consequently, $d_2(M,N) = 0$. □

Remark 4.4. Clearly, $d_1(M,N) \ge d_2(M,N)$ for any $M, N \in S_n(H)$. But we can claim that there is no any constant $c > 0$ such that

$$d_1(M,N) \le c\, d_2(M,N), \quad M, N \in S_n(H) \text{ with } M \cap N = \{0\}.$$

In fact, let $e_1, \cdots, e_{2n-2}$ be mutually orthonormal vectors in H ($n \ge 2$) and put $V = \text{span}\{e_1, \cdots, e_{2n-2}\}$. For any integer $k \ge 1$, choose $f_k, g_k \in V^\perp$ such that f_k, g_k are linearly independent and

$$\|f_k\| = \|g_k\| = 1,\ (f_k, g_k) = (g_k, f_k),\ \|f_k - g_k\| < \frac{1}{k}.$$

Set

$$M_k = \text{span}\{e_1, \cdots, e_{n-1}, f_k\} \text{ and } N_k = \text{span}\{e_n, \cdots, e_{2n-2}, g_k\},\ k \ge 1.$$

Then $M_k, N_k \in S_n(H)$, $M_k \cap N_k = \{0\}$ and

$$P_{M_k} P_{N_k} x = (x, g_k)(g_k, f_k) f_k,\ P_{M_k} P_{N_k} P_{M_k} x = (x, f_k)|(f_k, g_k)|^2 f_k,\ \forall\, x \in H.$$

Thus $(P_{M_k} P_{N_k} P_{M_k})^{1/2} x = (x, f_k)|(f_k, g_k)| f_k$, $\forall\, x \in H$ and

$$\|P_{M_k} P_{N_k}\| = |(f_k, g_k)| = \frac{1}{2}(2 - \frac{1}{k^2})$$

and hence $d_1(M_k, N_k) \to \sqrt{2(n-1)}$ and $d_2(M_k, N_k) \to 0$ as $k \to \infty$.

References

[1] F. Deutsch, Best Approximation in Inner Product Spaces, CMS Books in Math. 7, Springer–Verlag, Now York, Berlin, Heidelberg, 2001.

[2] S. Hu and Y. Xue, Completeness of n–tuple of projections in C^*–algebras, J. Operator Theory, **72**(1) (2014), 49–70.

[3] X. Sun and Q. Cheng, On subspace distance, in Lecture Notes in Computer Science, 4142, 81–89, Springer Berlin Heidelberg, 2006.

[4] L. Wang, X. Wang and J. Feng, Subspace distance analysis with application to adaptive Bayesian algorithm for face recognition, Pattern Recognition **39** (3) (2006), 456-464.

[5] Y. Xue, Stable Perturbations of Operators and Related Topics, World Scientific, 2012.

[6] G. Zuccon, L.A. Azzopardi and C.J. van Rijsbergen, Semantic Spaces: Measuring the distance between different subspaces, 225–236, P. Bruza, D. Sofge, W. Lawless, K. van Rijsbergen, M. Klusch (Eds), Quantum Interaction, Lecture Notes in Artificial Intelligence, 5494, Springer-Verlag, Berlin, 2009.