isid/ms/2014/2
January 23, 2014
http://www.isid.ac.in/~statmath/eprints

Approximation problems in the Riemannian metric on positive definite matrices

Rajendra Bhatia and Tanvi Jain

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

APPROXIMATION PROBLEMS IN THE RIEMANNIAN METRIC ON POSITIVE DEFINITE MATRICES

RAJENDRA BHATIA AND TANVI JAIN

Abstract. There has been considerable work on matrix approximation problems in the space of matrices with Euclidean and unitarily invariant norms. We initiate the study of approximation problems in the space $\mathbb{P}$ of all $n \times n$ positive definite matrices with the Riemannian metric δ_2. Our main theorem reduces the approximation problem in $\mathbb{P}$ to an approximation problem in the space of Hermitian matrices and then to that in $\mathbb{R}^n$. We find best approximants to positive definite matrices from special subsets of $\mathbb{P}$. The corresponding question in Finsler spaces is also addressed.

1. Introduction

Let $\mathbb{M}$ be the space of all $n \times n$ complex matrices. A *matrix approximation problem* consists of finding the best approximant to an element A of $\mathbb{M}$ from a special subset $\mathbb{S}$. For example, $\mathbb{S}$ could consist of all Hermitian, unitary, positive definite, normal, Toeplitz, or circulant matrices, all matrices of rank k, all matrices with a fixed spectrum, etc. The approximation could be sought with respect to the Euclidean norm $\|\cdot\|_2$, or with respect to some other norm. There is a considerable body of work on such approximation problems with unitarily invariant norms. See, e.g., [1], [2]. Some of these problems turn out to be easy, and have elegant solutions. Some others are both hard and intricate. The nature of the set and the norm both play a role in the tractability of the problem. Section IX.7 of [3] provides an introduction to such problems. Several applications are given in [6].

In recent years there has been considerable interest in the metric space $(\mathbb{P}, \delta_2)$ consisting of the space $\mathbb{P}$ of $n \times n$ positive definite matrices

2010 *Mathematics Subject Classification.* Primary 15B48; Secondary 47A58, 52A41, 53B21, 53B40.

Key words and phrases. Matrix approximation problem, positive definite matrix, Riemannian metric, convex set, Finsler metric.

with a natural Riemannian metric δ_2. This is a classical object of differential geometry and has become a major topic in matrix analysis with several applications in diverse areas, such as image processing, radar detection, brain-computer interfacing and machine learning. See [7].

The principal aim of this article is to initiate the study of approximation problems in the space $(\mathbb{P}, \delta_2)$. Just as in the theory of approximation in the space $\mathbb{M}$, we consider special subsets $\mathcal{K}$ in $\mathbb{P}$ and find the best approximant to any element A of $\mathbb{P}$ from the set $\mathcal{K}$. We show that under some mild restrictions on $\mathcal{K}$ (convexity and unitary invariance), the problem can be reduced to an approximation problem in $\mathbb{R}^n$. One of the interesting features of approximation theory in $\mathbb{M}$ has been that very often the same element turns out to be the best approximant in every unitarily invariant norm. Thus for example in every such norm, the best Hermitian approximant to A is its real part $\mathrm{Re}\, A = \frac{1}{2}(A + A^*)$ and the best unitary approximant is the one that occurs in the polar decomposition $A = UP$.

Every unitarily invariant norm $|||\cdot|||$on $\mathbb{M}$ induces a Finsler metric $\delta_{|||\cdot|||}$ on $\mathbb{P}$. The Riemannian metric δ_2 is special among these as it corresponds to the Euclidean norm $\|\cdot\|_2$. It is of some interest to see how the choice of norm affects the approximation problem. We address this question too.

An introduction to the geometry of the space $\mathbb{P}$ can be found in Chapter 6 of [4]. For basic facts of matrix analysis we refer to [3], [4].

2. Best Approximants in $(\mathbb{P}, \delta_2)$

Let $\mathbb{H}$ be the set of all $n \times n$ complex Hermitian matrices. In this section we consider $\mathbb{H}$ with the Euclidean norm $\|\cdot\|_2$ defined as

$$\|A\|_2 = (\mathrm{tr}(A^2))^{1/2} = (\sum_{i,j=1}^{n} |a_{ij}|^2)^{1/2},$$

and the space $\mathbb{P}$ with the Riemannian metric δ_2 defined as

$$\delta_2(A, B) = \|\log(A^{-1/2} B A^{-1/2})\|_2$$

An important property of the metric δ_2 that proves to be useful in the study of approximation problems in $\mathbb{P}$ is the exponential metric increasing (EMI) property. See [4], [5].

Theorem 2.1. *(EMI) For any two Hermitian matrices H and K,*

$$\delta_2(\mathrm{e}^H, \mathrm{e}^K) \geq \|H - K\|_2.$$

If H and K commute, then the two sides are equal.

Any two elements A and B in the space $(\mathbb{P}, \delta_2)$ can be joined by a unique geodesic. This is the curve $A\sharp_t B$ $(0 \le t \le 1)$ given by

$$A\sharp_t B = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}.$$

We write $[A, B]$ for the set

$$\{A\sharp_t B : 0 \le t \le 1\}. \tag{2.1}$$

A subset $\mathcal{K}$ of $\mathbb{P}$ is said to be *convex* (*geodesically convex*) if for every pair of elements A, B in $\mathbb{P}$, the geodesic $[A, B]$ lies entirely in $\mathbb{P}$. This notion is different from convexity in the vector spaces. A subset $\mathcal{S}$ of the real vector space $\mathbb{H}$ is called convex if for every pair of elements H, K in $\mathcal{S}$, the line segment $(1-t)H + tK$, $0 \le t \le 1$, lies entirely in $\mathcal{S}$.

It is well-known that if $\mathcal{K}$ is a closed convex subset of the space $(\mathbb{P}, \delta_2)$, then every A in $\mathbb{P}$ has a unique best approximant from $\mathcal{K}$; i.e., there exists a unique element A_0 of $\mathcal{K}$ such that

$$\delta_2(A, A_0) \le \delta_2(A, X)$$

for all $X \in \mathcal{K}$. See [4]. Our aim is to find this best approximant for some special sets $\mathcal{K}$.

We say that a set $\mathcal{K}$ is *unitarily invariant* if for every unitary matrix U, UAU^* is in $\mathcal{K}$ whenever A is. If H is a Hermitian matrix, we denote by $\operatorname{Eig} H$ an n-vector whose components are the n eigenvalues of H. Given a vector x in $\mathbb{R}^n$, we write D_x for the diagonal matrix whose diagonal coincides with x.

It will be convenient to fix some notations. Given a convex, unitarily invariant set $\mathcal{K}$ in $\mathbb{P}$, we associate with it a subset $\mathcal{S}$ of $\mathbb{H}$ defined as

$$\begin{aligned} \mathcal{S} &= \{\log A : A \in \mathcal{K}\} \\ &= \{H \in \mathbb{H} : \mathrm{e}^H \in \mathcal{K}\}. \end{aligned}$$

The subset S of $\mathbb{R}^n$ is defined as

$$S = \{x \in \mathbb{R}^n : x = \operatorname{Eig} H \text{ for some } H \in \mathcal{S}\}.$$

This is the set of all x such that UD_xU^* is in $\mathcal{S}$ for some unitary U.

Our main theorem reduces the problem of finding the best approximant to an element A of $\mathbb{P}$ from the set $\mathcal{K}$ to the problem of approximating $\log A$ from $\mathcal{S}$ and then to approximating $\operatorname{Eig} \log A$ from the set S. As corollaries we consider three special examples where the theorem is applied.

Theorem 2.2. *Let $\mathcal{K}$ be a closed convex unitarily invariant subset of $\mathbb{P}$. Suppose $\mathcal{S}$ is also convex. Let A be any element of $\mathbb{P}$ and let $\log A = UD_yU^*$ be the spectral decomposition of $\log A$. Let $\Phi(y)$ be the best approximant to y from the set S, and $\tilde{\Phi}(\log A)$ the best approximant to $\log A$ from the set $\mathcal{S}$. Then*

(i) the best approximant to A from $\mathcal{K}$ is $\exp(\tilde{\Phi}(\log A))$,

(ii) $\tilde{\Phi}(\log A) = UD_{\Phi(y)}U^*$.

Remark 2.3. If $\mathcal{K}$ is unitarily invariant, then so is $\mathcal{S}$. However, if $\mathcal{K}$ is convex, then $\mathcal{S}$ need not be convex. This is shown in the example given below. It is easy to see that if $\mathcal{S}$ is a convex unitarily invariant set, then S is convex.

Example 2.4. Let $\mathcal{K}$ be the geodesic from A to B in $\mathbb{P}$, i.e.,

$$\mathcal{K} = \{A\sharp_t B : 0 \leq t \leq 1\}.$$

We show that $\mathcal{S}$ is not always convex.

Suppose $\mathcal{S}$ is convex. Then the line segment $\gamma(t) = (1-t)\log A + t\log B$ lies in $\mathcal{S}$. Thus there exists an injective function θ from $[0,1]$ into itself such that

$$\mathrm{e}^{\gamma(t)} = A\sharp_{\theta(t)}B.$$

Then

$$\det(\mathrm{e}^{\gamma(t)}) = \det(A\sharp_{\theta(t)}B) = \det(A)^{1-\theta(t)}\det(B)^{\theta(t)}.$$

We also have

$$\det(\mathrm{e}^{\gamma(t)}) = \det(A)^{1-t}\det(B)^{t}.$$

This implies that

$$\det(A)^{\theta(t)-t} = \det(B)^{\theta(t)-t}.$$

So, if $\det(A) \neq \det(B)$, then $\theta(t) = t$. In particular, this means that

$$A\sharp_{1/2}B = \exp\Big(\frac{\log A + \log B}{2}\Big).$$

However, it is well-known that this is not always true.

Proof of Theorem 2.2. Since $\mathcal{K}$ is unitarily invariant, and both δ_2 and $\|\cdot\|_2$ are unitarily invariant, we can assume that A is diagonal. For any matrix B, let us denote by $D(B)$ the diagonal part of B. The diagonal matrix $D(B)$ is a convex combination of unitary conjugates of B and $\mathcal{S}$ is a convex unitarily invariant set. Thus if B is an element of $\mathcal{S}$, then so is $D(B)$.

Since $\log A$ is diagonal, we have

$$\|\log A - D(\tilde{\Phi}(\log A))\|_2^2 \leq \|\log A - \tilde{\Phi}(\log A)\|_2^2.$$

But $\tilde{\Phi}(\log A)$ is the best approximant to $\log A$ from $\mathcal{S}$. Hence $\tilde{\Phi}(\log A) = D(\tilde{\Phi}(\log A))$. In other words $\tilde{\Phi}(\log A)$ is diagonal.

Now let X be any element of $\mathcal{K}$. Then by Theorem 2.1

$$\delta_2(A, X) \geq \|\log A - \log X\|_2.$$

Since $\log X \in \mathcal{S}$, this gives

$$\delta_2(A, X) \geq \|\log A - \tilde{\Phi}(\log A)\|_2.$$

As seen above $\tilde{\Phi}(\log A)$ commutes with $\log A$. So, again by Theorem 2.1, we get

$$\delta_2(A, X) \geq \delta_2(A, \exp(\tilde{\Phi}(\log A))).$$

This proves (i).

To prove (ii), we can again assume that $\log A = D_y$. Since $\Phi(y)$ is the best approximant to y from S, for all diagonal matrices $D \in \mathcal{S}$

$$\|\log A - D_{\Phi(y)}\|_2 \leq \|\log A - D\|_2,$$

and since $\tilde{\Phi}(\log A)$ must be a diagonal matrix, $\tilde{\Phi}(\log A) = D_{\Phi(y)}$. ∎

Corollary 2.5. *Let*

$$\mathcal{K} = \{X \in \mathbb{P} : \det(X) = 1\}.$$

and let A be any element of $\mathbb{P}$. Then the best approximant to A from the set $\mathcal{K}$ is $A_0 = A/(\det(A))^{1/n}$.

Proof. The set $\mathcal{K}$ is clearly closed and unitarily invariant. Using the relation

$$\det(X \sharp_t Y) = (\det X)^{1-t} (\det Y)^t, \quad 0 \leq t \leq 1,$$

we see that $\mathcal{K}$ is convex. In this case

$$\mathcal{S} = \{H \in \mathbb{H} : \operatorname{tr}(H) = 0\}.$$

The best approximant to any element K of $\mathbb{H}$ from the set $\mathcal{S}$ is

$$\tilde{\Phi}(K) = K - \frac{\operatorname{tr} K}{n} I.$$

So, by Theorem 2.2, the best approximant to A from $\mathcal{K}$ is

$$\begin{aligned}\exp(\tilde{\Phi}(\log A)) &= \exp(\log A - \frac{\operatorname{tr}(\log A)}{n} I) \\ &= A/(\det A)^{1/n}.\end{aligned}$$

■

The form of A_0 in Corollary 2.5 is one that could be guessed from the description of $\mathcal{K}$. This is less so in the next case we consider.

We denote by $\|A\|$, the operator norm of A, i.e.,

$$\|A\| = \sup\{\|Ax\| : \|x\| = 1\} = s_1(A),$$

the maximum singular value of A. Recall that every Hermitian matrix H has a Jordan decomposition $H = H_+ - H_-$ in which $H_\pm$ are positive semidefinite. These are called the positive and negative parts of H, respectively.

Corollary 2.6. *Let*

$$\mathcal{K} = \{X \in \mathbb{P} : \|X\| \leq 1\},$$

and let A be any element of $\mathbb{P}$. Then the best approximant to A from $\mathcal{K}$ is

$$A_0 = \exp(-(\log A)_-).$$

Proof. The set $\mathcal{K}$ is clearly closed and unitarily invariant. Its convexity follows from the relation

$$\|X \natural_t Y\| \leq \|X\|^{1-t}\|Y\|^t, \quad 0 \leq t \leq 1.$$

See [4]. This set consists of all positive definite matrices X with their maximum eigenvalue $\lambda_1^{\downarrow}(X) \leq 1$. Hence $\mathcal{S}$ consists of all Hermitian matrices whose maximum eigenvalue is nonpositive. In other words $\mathcal{S}$ is the set of all negative semidefinite matrices. By Theorem IX.7.3 in [3] the best approximant to any element K of $\mathbb{H}$ from the set $\mathcal{S}$ is $-K_-$. Hence by Theorem 2.2 the best approximant to A from $\mathcal{K}$ is $\exp(-(\log A)_-)$. ■

We next consider the set

$$\mathcal{K} = \{X \in \mathbb{P} : \| \wedge^k X\| \leq 1\}, \tag{2.2}$$

$1 \leq k \leq n$. The special cases $k = 1$ and n are the ones considered in Corollary 2.6 and 2.5, respectively. The general case turns out to be more intricate. For handling this we reduce the problem to an ordinary convex program in $\mathbb{R}^n$ and then use the Karush-Kuhn-Tucker (KKT) optimization theorem. See [8].

By an ordinary convex program (P) we mean a problem of the following form.

$$\begin{array}{lll} \text{Minimize} & f_0(x) & \text{over } \mathbb{R}^n \text{ subject to the constraints} \\ & f_i(x) \le 0, \quad 1 \le i \le m. & \end{array}$$

where each $f_j(x)$ is a real-valued convex function.

The sets $\mathcal{S}$ and S associated with $\mathcal{K}$ are given by

$$\mathcal{S} = \{H \in \mathbb{H} : \sum_{i=1}^{k} \lambda_i^{\downarrow}(H) \le 0\}$$

and

$$S = \{x \in \mathbb{R}^n : \sum_{i=1}^{k} x_i^{\downarrow} \le 0\}. \tag{2.3}$$

It can be verified that $\mathcal{K}$ and $\mathcal{S}$ are closed, convex and unitarily invariant. Hence we need to find the best approximant to a given vector y in $\mathbb{R}^n$ from S. Since S is unitarily invariant and

$$x^{\downarrow} - y^{\downarrow} \prec x - y$$

for all x, y in $\mathbb{R}^n$, we can assume that $y \in \mathbb{R}^{n\downarrow}$, i.e., $y_1 \ge \cdots \ge y_n$, and $S \subseteq \mathbb{R}^{n\downarrow}$. Thus we can interpret this approximation problem as the following ordinary convex program.

$$\begin{array}{lll} (P_0) & \min & \sum_{i=1}^{n} (x_i - y_i)^2 \quad \text{subject to} \\ & & f_j(x) \le 0, \qquad 1 \le j \le n, \end{array}$$

where

$$f_1(x) = \sum_{j=1}^{k} x_i$$

and

$$f_{j+1}(x) = x_{j+1} - x_j, \; 1 \le j \le n-1.$$

We solve this ordinary convex program by using a special case of KKT optimization theorem. For the convenience of the reader we state the theorem.

Theorem 2.7. *Let (P) be an ordinary convex program. Suppose $(\lambda_1, \ldots, \lambda_m)$ is a nonnegative vector in $\mathbb{R}^m$ such that the infimum of the function*

$$h = f_0 + \lambda_1 f_1 + \cdots + \lambda_m f_m$$

is finite and equals $\min f_0$*. Let* D *be the set of points where* h *attains its infimum over* $\mathbb{R}^n$*. Then a vector* $\overline{x}$ *in* D *is an optimal vector for* (P) *if for all* $1 \leq i \leq m$*,* $f_i(\overline{x}) = 0$ *whenever* $\lambda_i > 0$*, and* $f_i(\overline{x}) \leq 0$ *otherwise.*

Since the Euclidean norm is strictly convex, we obtain a unique optimal vector for (P_0).

The best approximant to any element in S is obviously itself, so we assume that $y \notin S$,i.e., $y_1 \geq y_2 \geq \cdots \geq y_n$ and $y_1 + \cdots + y_k > 0$. Let $\overline{y}_k = \frac{y_1+\cdots+y_k}{k}$. The following two cases arise while solving the problem (P_0).

Case 1. $y_k - y_{k+1} \geq \overline{y}_k$.
Case 2. $y_k - y_{k+1} < \overline{y}_k$.

In Case 1, the vector $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_n)$ given by

$$\begin{array}{rcll} \tilde{y}_i & = & y_i - \overline{y}_k & 1 \leq i \leq k \\ \tilde{y}_j & = & y_j & k+1 \leq j \leq n. \end{array}$$

satisfies the conditions of the KKT optimization theorem and hence is the unique optimal solution for (P_0).

We next consider Case 2. For $1 \leq m \leq n-k$, let

$$\begin{aligned} \mu_m &= \frac{(m+1)(y_1 + \cdots + y_{k-1}) + y_k + \cdots + y_{k+m}}{(m+1)k - m}, \\ \nu_m &= \frac{m(y_1 + \cdots + y_{k-1} - m(k-1)y_k) + k(y_{k+1} + \cdots + y_{k+m})}{(m+1)k - m} \\ z_m &= y_k - \mu_m + \nu_m \end{aligned}$$

$1 \leq m \leq n-k$. It can be verified that $\mu_1 > 0$, $\nu_1 > 0$, $\mu_1 > \nu_1$ and $y_{k+1} > z_1$. Let

$$p = \max\{m : 1 \leq m \leq n-k, \mu_m > 0, \nu_m > 0, \mu_m > \nu_m \text{ and } y_{k+m} > z_m\}.$$

Let $\hat{y} = (\hat{y}_1, \ldots, \hat{y}_n)$ be the vector given by

$$\begin{array}{rcll} \hat{y}_i & = & y_i - \mu_p & 1 \leq i \leq k-1 \\ \hat{y}_k & = & z_p & \\ \hat{y}_i & = & \hat{y}_k & k+1 \leq i \leq k+p \\ \hat{y}_i & = & y_i & i > k+p. \end{array}$$

In this case $\hat{y}$ turns out to be the unique optimal solution of (P_0).

Thus the best approximant $\Phi(y)$ to y from S is given by

$$\Phi(y) = \begin{cases} y & \text{if } y \in S \\ \tilde{y} & \text{if } \ y_k - y_{k+1} \geq \overline{y}_k \\ \hat{y} & \text{if } \ y_k - y_{k+1} < \overline{y}_k. \end{cases} \tag{2.4}$$

Corollary 2.8. *Let*

$$\mathcal{K} = \{X \in \mathbb{P} : \| \wedge^k X\| \leq 1\}.$$

Let A be any element of $\mathbb{P}$ and let $\log A = UD_yU^*$ be the spectral decomposition of $\log A$. Then the best approximant to A from $\mathcal{K}$ is

$$A_0 = U \exp(D_{\Phi(y)})U^*,$$

where $\Phi(y)$ is given by (2.4).

3. Best Approximants in $(\mathbb{P}, \delta_{|||\cdot|||})$

In this section we indicate how results of Section 2 may be extended to Finsler metrics on $\mathbb{P}$ arising from unitarily invariant norms on $\mathbb{M}$.

Recall that a norm $|||\cdot|||$ on $\mathbb{M}$ is said to be *unitarily invariant* if $|||UAV||| = |||A|||$ for all A and unitary U, V. Such a norm arises from a symmetric gauge function $|||\cdot|||$ on $\mathbb{R}^n$. It gives rise to a Finsler metric $\delta_{|||\cdot|||}$ on $\mathbb{P}$ defined as

$$\delta_{|||\cdot|||}(A, B) = \left|\left|\left|\log(A^{-1/2}BA^{-1/2})\right|\right|\right|.$$

It has been shown in [5] that in the metric space $(\mathbb{P}, \delta_{|||\cdot|||})$ the curve $[A, B]$ defined in (2.1) is a geodesic joining A and B. (This geodesic is unique if geodesics in $|||\cdot|||$ are unique.) We say that a set $\mathcal{K}$ in $(\mathbb{P}, \delta_{|||\cdot|||})$ is convex if it is convex in $(\mathbb{P}, \delta_2)$.

Remark 3.1. The crucial EMI property is also valid in $(\mathbb{P}, \delta_{|||\cdot|||})$; we have

$$\delta_{|||\cdot|||}(\mathrm{e}^H, \mathrm{e}^K) \geq |||H - K|||,$$

for all Hermitian matrices H, K and the two sides are equal if H and K commute. See [5].

Theorem 2.2 remains true: the approximation problem in $(\mathbb{P}, \delta_{|||\cdot|||})$ reduces to the approximation problem in $(\mathbb{H}, |||\cdot|||)$ and then to that in $(\mathbb{R}^n, |||\cdot|||)$. The best approximant is unique whenever the norm is strictly convex.

The best approximants obtained for the examples in Corollaries 2.5 and 2.6 for the Euclidean norm work for all unitarily invariant norms. The reason being that in both cases, $\tilde{\Phi}(\log A)$ is the best approximant to $\log A$ from $\mathcal{S}$ independent of the norm.

For Corollary 2.6, this follows from Theorem IX.7.3 of [3]. For the convenience of the reader we briefly sketch the proof for Corollary 2.5.

Let $\mathcal{K}$ be the set given in Corollary 2.6. Then the set S associated to it is

$$S = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 0\}.$$

Let $\overline{y}$ be the n-vector with all its components equal to $\frac{\sum_{i=1}^{n} y_i}{n}$, and let J be the $n \times n$ matrix with all its entries $1/n$. The matrix J is doubly stochastic. By using the relation

$$\overline{y} = J(y - z),$$

and Theorem II.1.9 of [3], it can be verified that $\Phi(y) = y - \overline{y}$ is the best approximant to y from S. This proves that the best approximant to any element from the set $\mathcal{K}$ of Corollary 2.6 is the same for every unitarily invariant norm.

However this does not happen in case of Corollary 2.8, i.e., in this case the best approximant to A from $\mathcal{K}$ is dependent on the choice of the norm.

Let $\|\cdot\|$ denote the symmetric gauge function on $\mathbb{R}^n$ defined by

$$\|(x_1, \ldots, x_n)\| = \max_{1 \le i \le n} |x_i|.$$

This induces the operator norm on $\mathbb{M}$ and the metric $\delta_{\|\cdot\|}$ on $\mathbb{P}$. In the following example we show that the best approximants from the set $\mathcal{K}$ described in (2.2) in the two spaces $(\mathbb{P}, \delta_2)$ and $(\mathbb{P}, \delta_{\|\cdot\|})$ are not always the same.

Example 3.2. Let $\mathcal{K}$ be the subset of 4×4 positive definite matrices defined as

$$\mathcal{K} = \{X \in \mathbb{P} : \| \wedge^2 X\| \le 1\}.$$

The set S associated with $\mathcal{K}$ is

$$S = \{x \in \mathbb{R}^4 : x_1^{\downarrow} + x_2^{\downarrow} \le 0\}.$$

Consider the vector $y = (5, 1, 1, -1)$. Let $\tilde{y} = (2, -2, -2, -2)$ and $\hat{y} = (1, -1, -1, -1)$. Then $\tilde{y}$ is a best approximant to y from S in $(\mathbb{R}^n, \|\cdot\|)$. By (2.4), $\hat{y}$ is the unique best approximant to y from S

in the Euclidean norm; but it is not a best approximant in $\|\cdot\|$ since $\|y-\hat{y}\| > \|y-\tilde{y}\|$.

References

1. T. Ando, *Approximation in trace norm by positive semidefinite matrices*, Linear Algebra Appl. **71** (1985), 15-21.
2. T. Ando, T. Sekiguchi and T. Suzuki, *Approximation by positive operators*, Math. Z. **131** (1973), 273-282.
3. R. Bhatia, *Matrix Analysis,* Springer, 1997.
4. R. Bhatia, *Positive Definite Matrices,* Princeton University Press, 2007.
5. R. Bhatia, *On the exponential metric increasing property,* Linear Algebra Appl. **375** (2003) 211-220.
6. N. J. Higham, *Matrix nearness problems and applications,* in M. J. C. Grover and S. Barnett eds.,Applications of Matrix Theory, Oxford University Press, 1989.
7. F. Nielsen and R. Bhatia, eds., *Matrix Information Geometry,* Springer, 2013.
8. R. T. Rockafellar, *Convex Analysis,* Princeton University Press, 1970.

Indian Statistical Institute, New Delhi, India
E-mail address: rbh@isid.ac.in

Indian Statistical Institute, New Delhi, India
E-mail address: tanvi@isid.ac.in