Rev. Téc. Ing., Univ. Zulia
Vol. 14, No. 2, 1991

NAGWA SHERIF
7th Zone, Nasr City-11727
P.O. Box - 4062
Cairo, Egypt

# ON TRANSFORMATIONS AND PERTURBATIONS OF ORTHOGONAL r-FRAMES


## ABSTRACT

A decomposition of $C^n$ into a finite direct sum of orthogonal subspaces can be conveniently represented by its orthogonal projector frame, which is the collection of the corresponding orthogonal projectors. Two such decompositions whose frames are close are known to be linearly homeomorphic and homotopic. In a recent work we compared the resulting geodesic arcs with naturally arising paths resulting from interpolating the balanced transformation, and found them cubically close. In this work we describe an efficient algorithm to compute the balanced transformation.

## RESUMEN

Una descomposición de $C^n$ en una suma directa finita de subespacios ortogonales puede ser representada convenientemente por su cuadro proyector ortogonal, la cual es la colección de los proyectores ortogonales correspondientes. Dos de tales descomposiciones, cuyos cuadros son cerrados, son "homórfico y homotópico". En este trabajo, se decribe un algoritmo eficiente para computar la transformación balanceada.


## 1. INTRODUCTION

Let E, F be two orthogonal r-frames on $C^n$ (i.e., a sequence of r commuting orthogonal projectors which sum to the identity). Assume that in the decomposition of $C^n$ with respect to the frames E, F, the corresponding subspaces are defined as ranges of rectangular matrices, which might as well be assumed to have orthogonal columns. The purpose of this work is to propose an algorithm to compute a particular unitary U which maps E onto F with only matrices of lower order entering the calculations. This unitary U is called the balanced transformation and is optimal in the sense that it deviates minimally from the identity in the Frobenius norm [4].

The computation of U, beside being of interest on its own, will give information about the principal angles. The principal angles have many applications in statistics and numerical analysis. In [1], the statistical models of canonical correlations, factor analysis and stochastic equations are described in terms of the principal angles. Other applications can be found in numerical analysis [16], and theory of approximate least squares [2].

We present the notation and preliminaries in Section 2. We also compare U with another unitary which is geometrically justified as the most natural way to move the frame E onto F. In Section 3 the main result about factorizing U is established. We use this factorization to construct an algorithm to compute U; this is given in Section 4. The algorithm is illustrated by a numerical example. In Section 5 we derive perturbation inequalities for the angles between subspaces and U(F,E)

## 2. NOTATIONS AND PRELIMINARIES

By an orthogonal r-frame E on $C^n$, we mean $E = (E_1,E_2,\dots,E_r)$ where $E_j \in C^{n \times n}$, $1 \le j \le r$, satisfy

$$\begin{cases} \text{(i) } 0 \ne E_j = E_j^2 = E_j^* & 1 \le j \le r, \\ \text{(ii) } \sum_{j=1}^{r} E_j = I \end{cases} \tag{2.1}$$

Clearly equation (2.1) implies that the $E_i$'s are pairwise disjoint, since if

$$x \in \mathcal{R}(E_i),\ \|x\| = \sum_{j=1}^{r} \|E_j x\|^2 \ \text{ so } \sum_{j \ne i} \|E_j x\|^2 = 0 \ .$$

*Key words:* Orthogonal frames, angles between subspaces, matrix inverse square root, perturbations of subspaces.
AMS (MOS) Subject Classificaction. 65F25, 65F30, 15A23, 15A60.

and consequently $E_jx = 0$, $j \neq i$, so $E_jE_i = \delta_{ij}E_i$.

Throughout, r will be fixed and we shall write a frame E to mean an orthogonal r-frame.

Two frames E and F are said to be unitarily similar if there exists a unitary matrix V such that VE = FV, that is, $VE_j = F_jV$, $1 \leq j \leq r$. The unitary similarity orbit of a fixed frame E, denoted by $\mathcal{E}^r(E)$, is the set of frames which are unitarily similar to E, namely

$$\mathcal{E}^r(E) = \{ VEV^*, \text{ V is unitary matrix } \}. \quad (2.2)$$

In [7] the set $\mathcal{E}^r(E)$ is studied where it is shown to be a Riemannian manifold. In fact if F is a close frame to E, then certainly $F \in \mathcal{E}^r(E)$. This will be the case if for example

$$\|E-F\| = \max_{1\leq i\leq r} \| E_i-F_i\| < 1.$$

A particular unitary U which realizes the equivalence of the frames E and F is

$$U = U(F,E) = \left(\sum_{j=1}^{r} F_jE_j\right)\left(\sum_{j=1}^{r} E_jF_jE_j\right)^{-1/2} \quad (2.3)$$

It can be easily checked that U(F,E)E = F U(F,E), so U(F,E) maps th subspace $\mathcal{R}(E_j)$ onto $\mathcal{R}(F_j)$, $1\leq j\leq r$. We also note that $U(F,E)^* = U(E,F)$; for this reason we call it the balanced transformation.

If we want to move the frame E onto F in the most natural and efficient way within the set of r-frames on $C^n$, this will not be achieved by oonsidering the straight line segment. This is because the straight line segment does not lie in $\mathcal{E}^r(E)$, since if

$$t \rightarrow E+t(F-E),\ 0\leq t\leq 1 \text{ lies in } \mathcal{E}^r(E),$$

then

$$F_j(t) = E_j + t(F_j-E_j)(1 \leq j \leq r,\ 0 \leq t \leq 1)$$

is an orthogonal projector. Thus

$$0 = F_j^2(t) - F_j(t) = (t^2-t)(F_j-E_j)^2 \quad 0 \leq t \leq 1,\ 1 \leq j \leq r.$$

This implies that $E_j = F_j$ for all j and hence E = F. However, a locally minimal arc in $\mathcal{E}^r(E)$ which connects E, F will be the geodesic arc $t \rightarrow (F_j(t))_{j=1}^r$, $t \in [0,1]$, F(0) = E, F(1) = F, where $F_j(t)$ is defined by

$$F_j(t) = \exp(tL)\ E_j\ \exp(-tL),\ 1\leq j\leq r. \quad (2.4)$$

Here L is a skew hermitian matrix ($L = -L^*$) and satisfies the matrix equation

$$\exp L - \sum_{j=1}^{r} F_j \exp L\ E_j + \sum_{j=1}^{r} E_j\ L\ E_j = 0 \quad (2.5)$$

It is shown in [7] that the length of the geodesic arc connecting E and F is $\|L\|_F$ ($\|L\|_F = (\mathrm{tr}\ L^*L)^{1/2}$), which justifies calling expL the direct rotation between E and F.

Both unitaries expL and U(F,E) give rise to paths in $\mathcal{E}^r(E)$ connecting E and F. However these paths are in general different [10]. The first unitary has geometric significance. The second unitary U(F,E), is not the most natural way to move the subspace $\mathcal{R}(E_j)$ onto $\mathcal{R}(F_j)$, $1\leq j\leq r$, but still has the advantage that it is expressed algebraically in terms of E and F. Also it is recently shown in [10] that U(F,E) is still close to expL, namely

$$\|U(F,E) - \mathrm{Exp}L\| = O(\|F-E\|^3).$$

So even if one is interested in computing expL via solving (2.5) iteratively, a good initial approximation will be U(F,E).

Let $\mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \dots \oplus \mathcal{L}_r$ and $\mathcal{M}_1 \oplus \mathcal{M}_2 \oplus \dots \oplus \mathcal{M}_r$ be the decomposition of $C^n$ arising from E and F respectively. That is, $\mathcal{L}_j = \mathcal{R}(E_j)$ and $\mathcal{M}_j = \mathcal{R}(F_j)$, $1\leq j\leq r$. There are different ways to identify subspaces of $C^n$. In our case we will define the subspaces using orthonormal matrices. Namely, for $j = 1,2,\dots,r$, let $V_j, W_j \in C^{n \times n_j}$ where

$$\begin{cases} V_j^*V_j = I & V_jV_j^* = E_j \\ W_j^*W_j = I\ , & W_jW_j^* = F_j \end{cases} \quad (2.6)$$

The above identification is unique only to within a post-multiplication by an arbitrary unitary $n_j \times n_j$ matrix.

The balanced transformation U(F,E) can be computed directly using equation (2.3), where the inverse square root of an n x n matrix is to be computed. Such an inverse square root can be computed using, for example, the numerically stable technique suggested in [14]. However, if n is large, the above procedure which is of order $O(n^3)$ will be computationally expensive. The purpose of the next section is to propose a factorization of U(F,E) so that only lower order matrices, $n \times n_j$, and $n_j \times n_j$ will enter the calculations. The saving will be remarkable, when the restriction of U(F,E) to $\mathcal{L}_i$ is required with $n_i \ll n$.

**Remark 2.1.** If the subspaces $\mathcal{L}_j$, $\mathcal{M}_j$ are defined by $A_j$, $B_j$ respectively, $1\leq j\leq r$, then a QR factorization step is needed to get $V_j$, $W_j$.

## 3. A FACTORIZATION OF U(F,E)

The following relations are well known, cf, [3,8], we list them for the sake of completeness. For two frames E and F we define

$$C_j = (F_j + E_j - I)^2 \ , \ 1 \le j \le r \ ; \qquad (3.1)$$

then we have for $1 \le j \le r$,

$$\begin{cases} \text{(i)} \quad 0 \le C_j \le I \\ \text{(ii)} \quad C_jE_j = E_jC_j, \ C_jF_j = F_jC_j \\ \text{(iii)} \ E_jC_jE_j = E_jF_jE_j \ , \ F_jC_jF_j = F_jE_jF_j \ . \end{cases} \qquad (3.2)$$

Equation (2.3) can be equivalently written as

$$\begin{cases} U(F,E) = (\sum_{j=1}^{r} F_jE_j)(\sum_{j=1}^{r} E_jF_jE_j)^{-1/2} \\ \quad = (\sum_{j=1}^{r} F_jE_j)(\sum_{j=1}^{r} E_jC_j^{-1/2}E_j) \\ \quad = \sum_{j=1}^{r} F_jC_j^{-1/2}E_j \ . \end{cases} \qquad (3.3)$$

This follows by direct calculations, using properties of the $C_j$'s listed in (3.2). Further, set

$$T_j = T_j(F_j,E_j) = C_j^{-1/2}(F_j+E_j-I), \quad 1 \le j \le r. \qquad (3.4)$$

and associate $Z_j$ with $T_j$ where

$$Z_j = \frac{1}{2}(I + T_j), \quad 1 \le j \le r. \qquad (3.5)$$

The following theorem records some properties of the $T_j$'s. Also it expresses U(F,E) in terms of the $T_j$'s.

**Theorem 3.1.** Each $T_j$ is a hermitian involutary matrix exchanging $\mathcal{L}_j$ with $\mathcal{M}_j$ and

$$E_jT_jE_j \ge 0, \ F_jT_jF_j \ge 0, \ 1 \le j \le r. \qquad (3.6)$$

The balanced transformation U(F,E) can be expressed in terms of the $T_j$'s as follows

$$U(F,E) = \sum_{j=1}^{r} T_jE_j \ . \qquad (3.7)$$

Further, $Z = (Z_1, \ Z_2, \ \ldots, \ Z_r)$, where the $Z_j$'s are defined by (3.5), satisfies

$$(T(Z,E)T_o)^2 = T(F,E)T_o. \qquad (3.8)$$

Here

$$T(Z,E) = (T_j(Z_j,E_j))_{j=1}^{r} \ , \ T_o = (T_{oj})_{j=1}^{r} \ , \ T_{oj} = 2E_j-I.$$

**Proof.** From equation (3.4), we have by direct calculations using (3.2), $T_j^* = T_j$ and $T_j^2 = I$.

Further $T_jE_j = C_j^{-1/2}F_jE_j = F_jT_j$, since $F_j$ commutes with $C_j$. Hence indeed $T_j$ exchanges $\mathcal{L}_j$ and $\mathcal{M}_j$. To prove (3.6) we note that $F_jT_jF_j \ge 0$ is equivalent to $E_jT_jE_j \ge 0$ since $F_jT_jE_j = T_j(E_jT_jE_j)T_j$, hence we show that $E_jT_jE_j \ge 0$.

$$E_jT_jE_j = C_j^{-1/2} E_jF_jE_j = C_j^{-1/2} E_jC_jE_j = E_jC_j^{1/2}E_j \ge 0,$$

since $C_j \ge 0$. From equation (3.3) we have

$$U(F,E) = \sum_{j=1}^{r} F_jC_j^{-1/2} E_j = \sum_{j=1}^{r} C_j^{-1/2} F_jE_j$$
$$= \sum_{j=1}^{r} C_j^{-1/2} (F_j + E_j - I)E_j = \sum_{j=1}^{r} T_jE_j$$

hence (3.7) follows. Now since $T_j$ is a hermitian involutary matrix, then $Z_j$ is an orthogonal projector. Hence if we define $T_j(Z_j,E_j)$ by equation (3.4), $T_j(Z_j,E_j)$ will be a hermitian involution which exchanges $\mathcal{R}(Z_j)$ with $\mathcal{R}(E_j)$. Further, since

$T_j(Z_j,E_j)E_j = Z_jT_j(Z_j,E_j)$, we have

$$[T_j(Z_j,E_j)(2E_j-I)]^2 = T_j^2(Z_j,E_j)(2Z_j-I)(2E_j-I);$$

but $T_j^2(Z_j, E_j) = I$, hence (3.8) follows.

**Remark 3.1.** The components $Z_j$ of Z are orthoprojectors on subspaces which can be named as the bisector subspaces of $\mathcal{L}_j$ and $\mathcal{M}_j$, this can be seen from equation

(3.8), see also [3] in case of a pair of subspaces. However, in general $Z \notin \mathcal{S}^r(E)$, $r > 2$; in case of a 2-frame $Z = (Z_1, Z_2)$ will be a frame. This follows since

$$Z_1 + Z_2 = \frac{1}{2}(2I + T_1 + T_2), \text{ with } T_2$$

$$= -T_1, \text{ hence indeed } Z_1 + Z_2 = I.$$

We construct an orthonormal basis of the bisector subspace $\mathcal{R}(Z_j)$ in terms of $V_j$ and $W_j$. This construction extends in some sense the calculation of the besector of two unit vectors in the plane.Once this base is established we can compute $T_j$ and consequently U can be computed via equation (3.7).

**Theorem 3.2.** Let $\{V_j\}_{j=1}^r$ and $\{W_j\}_{j=1}^r$ be as defined in (2.6).

(i) There exists an orthonormal matrix $X_j$, $1 \le j \le r$, such that $\mathcal{R}(X_j) = \mathcal{M}_j$, and $X_j$ is the closest orthonormal basis to $V_j$.

(ii) Set $Y_j = W_j^* V_j$, then $X_j$ in part (i) can be expressed as follows:

$$X_j = W_j Y_j (Y_j^* Y_j)^{-1/2}, \quad 1 \le j \le r.$$

(iii) If $G_j = X_j + V_j$, then $N_j = G_j (G_j^* G_j)^{-1/2}$, $1 \le j \le r$, is an orthonormal basis of the bisector subspace $\mathcal{R}(Z_j)$.

**Proof.** Define $H_j$, $1 \le j \le r$, by

$$H_j = W_j^* \, U(F,E) V_j.$$

Upon using equations (2.3) and (2.6) we have $H_j^* H_j = H_j H_j^* = I$. Let

$$X_j = W_j H_j,$$

so $X_j^* X_j = I$, and $X_j X_j^* = W_j W_j^* = F_j$; and indeed $X_j$ is a basis for $\mathcal{M}_j = \mathcal{R}(F_j)$ which is closest to $V_j$ (because $X_j = U(F,E)V_j$). To prove (ii), we have

$$H_j = W_j^* \, U(F,E) V_j = W_j^* T_j V_j,$$

as follows from the factorization of U(F,E) in equation (3.7). Hence

$$H_j = W_j^* (E_j + F_j - I) C_j^{-1/2} V_j$$

$$= W_j^* (V_j V_j^* + W_j W_j^* - I) C_j^{-1/2} V_j$$

$$= W_j^* C_j^{-1/2} V_j .$$

But

$$C_j V_j = (I - E_j - F_j + E_j F_j + F_j E_j) V_j$$

$$= (I - V_j V_j^* - W_j W_j^* + V_j V_j^* W_j W_j^* + W_j W_j^* V_j V_j^*) V_j$$

$$= V_j (W_j^* V_j)^* (W_j^* V_j).$$

Hence if we set

$$Y_j = W_j^* V_j, \; L_j = Y_j^* Y_j$$

we get

$$C_j V_j = V_j L_j$$

$$C_j^2 V_j = C_j V_j L_j = V_j L_j^2$$

Inductively, $C_j^m V_j = V_j L_j^m$ for any positive integer m. Hence $f(C_j)V_j = V_j f(L)$ for any continuous function on [0,1], so it is true for the inverse square root function, that is,

$$C_j^{-1/2} V_j = V_j L_j^{-1/2}.$$

But

$$H_j = W_j^* C_j^{-1/2} V_j$$

$$= W_j^* V_j L_j^{-1/2} = Y_j (Y_j^* Y_j)^{-1/2}.$$

Thus

$$X_j = W_j Y_j (Y_j^* Y_j)^{-1/2}, \quad 1 \le j \le r.$$

Next we use $X_j$ to establish basis for $\mathcal{R}(Z_j)$. We set

$$G_j = X_j + V_j.$$

Now $G_j$ is a basis for $\mathcal{R}(Z_j)$; this is because $G_j = X_j + V_j = T_j V_j + V_j = 2Z_j V_j$, hence $\mathcal{R}(G_j) = \mathcal{R}(Z_j)$. An orthonormal basis $N_j$ for $\mathcal{R}(G_j)$ can be established as

$$N_j = G_j(G_j^* G_j)^{-1/2}, \quad 1 \le j \le r.$$

That is, $N_j$ is the unitary polar factor in the polar decomposition of $G_j$. So indeed $N_j N_j^* = Z_j$ and the proof is complete.

**Remark 3.2.** We note that all inverse square root operations involve matrices of lower order $n_j \times n_j$. Further, if it is only required to compute $U(F,E)E_k$, then factorization (3.7) reduces the problem to computing $T_k E_k$ only. These two points illustrate the advantage of using (3.7) to compute U(F,E) rather than the direct formula (3.3).

**Remark 3.3.** In some statistical applications, one is interested in bases for the bisector subspace. For example, in factor analysis, the choice of coordinate system plays a prominent role. Here one is interested in referring a set of observations to especially chosen reference axes defined in some Euclidean space. In particular instances it is desired to define a coordinate system located "mid way between" two other coordinate systems [12]. In these instances we can apply Theorem 3.2 (iii) to find such a coordinate system.

The case r = 2 is particularly important. In this case it can be shown [8] that expL = U(F,E). Further, one can check that indeed U(F,E) satisfies the equation

$$U^2(F,E) = (2F_1 - I)(2E_1 - I).$$

The above equation suggests that compute U(F,E), one has to compute the principal square root of $(2F_1-I)(2E_1-I)$. The procedure suggested by Theorem 3.2 will be computationally efficient, since from equation (3.7) we have $U(F,E) = T_1E_1 + T_2(I-E)$. But $T_2 = -T_1$ hence $U(F,E) = T_1(2E_1-I)$. So we need only to compute

$$Z_1 = N_1 N_1^*.$$

We end this section by pointing out that such decompositions of $C^n$ arise when $\{\mathcal{L}_j\}$ and $\{\mathcal{M}_j\}$ are reducing subspaces of two nearby operators. Such case was studied in [5] for the case of 2-frames and for the case of r-frames in [9].

## 4. ALGORITHM

Let, in the decomposition of $\mathbb{C}^n$, the subspaces be defined by rectangular matrices $(V_j)_{j=1}^r$, $(W_j)_{j=1}^r$, which we assume orthonormal; cf.(2.6). In the sequel we shall need to compute the polar decomposition of a rectangular matrix. There are different techniques to achieve this [6,13]. One approach is based on the use of SVD of the given matrix. Let $A \in C^{k\times\ell}$, $k \ge \ell$ be a full rank matrix, consider SVD of A

$$A = Q \Sigma P^*, \quad \Sigma = \begin{bmatrix} D \\ 0 \end{bmatrix} \tag{4.1}$$

where $D = \text{diag}\ \{S_1(A), S_2(A), \ldots, S_\ell(A)\}$, $S_1(A) \ge S_2(A) \ge \ldots \ge S_\ell(A) \ge 0$.

Here $\{S_i(A)\}_{i=1}^{\ell}$ are called the singular values of A,P and Q are unitaries. If we partition Q as $Q = [Q_1, Q_2]$ where $Q_1 \in C^{k\times\ell}$, then in the polar decomposition of A, A = BH, the unitary polar factor B is $B = Q_1P^*$. Note that $B = A(A^*A)^{-1/2}$. In [6] another approach was proposed to construct the polar factor of a square matrix by applying the iteration

$$B_0 = A$$

$$B_{r+1} = \frac{1}{2}(B_r)(B_r^{*-1}). \tag{4.2}$$

Then $B_r \to B$ quadratically. If the matrix A is not square a QR factorization step is needed and then we apply (4.2) to R. The latter approach does not give information about singular values.

**Remark 4.1.** The algorithm to be described will enable us also to compute the angles between subspaces $\mathcal{L}_j$, $\mathcal{M}_j$. Each pair of subspaces $\mathcal{L}_j$, $\mathcal{M}_j$ is characterized in terms of certain angles called principal angles. These constitute the spectrum of a hermitian positive definite matrix $\Theta_j$. In fact

$$C_j = \cos^2\Theta_j, \quad I - C_j = \ell.c.^2(\Theta_j). \tag{4.3}$$

The spectrum of $C_j^{1/2}$ is the same as the set of singular values of $W_j^*V_j$, namely $\{(\cos\Theta_{jk})_{k=1}^{n_j}\}$, where $(\Theta_{jk})_{k=1}^{n_j}$ are the principal angles between $\mathcal{L}_j$, $\mathcal{M}_j$. Also the spectrum $\{\sin\,\theta_{jk}\}_{k=1}^{n_j}$ of $\ell.c.\Theta_j$ is the same as that of $F_j-E_j$, and is the same as the set of singular values of $W_j^{\perp *}V_j$ (here $W_j^{\perp}$ is orthonormal bases of $\mathcal{M}_j^{\perp}$ which can be obtained from $W_i$ $i \ne j$) that is. For the proof of these facts we refer to [5,15,17].

We now summarize the computational procedure to compute U, as well as other relevant quantities such as the principal angles or bases for bisector subspaces, when the orthonormal matrices $(V_j)_{j=1}^r$; $(W_j)_{j=1}^r$ are given.

Step 1. For j = 1,...,r do Step 2 to Step 6

Step 2 Set $Y_j = W_j^*V_j$

Step 3 Find SVD of $Y_j$, set $B_j = Y_j(Y_j^*Y_j)^{-1/2}$.

Step 4 Set $X_j = W_jB_j$, $G_j = X_j + V_j$.

Step 5 Compute $N_j = G_j(G_j^* G_j)^{-1/2}$

Step 6 Set $Z_j = N_j N_j^*$, $T_j = 2Z_j - I$, $E_j = V_j V_j^*$.

Step 7 Set $U = \sum_{j=1}^{r} T_j E_j$.

In applying the previous algorithm, the angles between subspaces can be computed, if required, in Step 2 as pointed out in Remark 4.1. In Step 5, we can compute any orthonormal bases for $R(G_j)$, for example a QR factorization step will be enough, however $N_j$ is the optimal one [13]. Finally the inverse square root encountered may also be computed as in [14].

We illustrate the previous algorithm by the following numerical example.

**Example 4.1.** Consider the following subspaces in $\mathbb{R}^4$ determined by

$$V_1 = e_1 \quad V_2 = [e_2, e_3] \quad V_3 = e_4$$

and

$$W_1 = \begin{bmatrix} -0.5 \\ 0.5 \\ -0.5 \\ 0.5 \end{bmatrix} \quad W_2 = \begin{bmatrix} -0.5 & 0.5 \\ -0.5 & -0.5 \\ 0.5 & -0.5 \\ -0.5 & -0.5 \end{bmatrix} \quad W_3 = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ -0.5 \end{bmatrix}$$

Applying the previous algorithm, the principal angles between $\mathcal{R}(V_j)$ and $\mathcal{R}(W_j)$, $1 \le j \le 3$ are

$$\left\{\frac{\pi}{3}\right\}, \left\{\frac{\pi}{4}, \frac{\pi}{4}\right\}, \left\{\frac{\pi}{3}\right\}$$

The balanced transformations is

$$U(F,E) = \begin{bmatrix} 0.50000 & 0.00000 & -0.70711 & 0.50000 \\ -0.50000 & 0.70711 & 0.00000 & -0.50000 \\ 0.50000 & 0.00000 & 0.70711 & -0.50000 \\ 0.50000 & 0.70711 & 0.00000 & 0.50000 \end{bmatrix}$$

We remark that iteration (4.2) can be applied in Step 2 instead of the SVD if the principal angles are not requiered.

## 5. A PERTURBATION INEQUALITY

Let E, F be the frames associated with the decompositon of $C^n$ into $\mathcal{L}_1 \oplus \dots \oplus \mathcal{L}_r$ and $\mathcal{M}_1 \oplus \dots \oplus \mathcal{M}_r$. Suppose these subspaces are perturbed so that we have $\tilde{\mathcal{L}}_1 \oplus \tilde{\mathcal{L}}_2 + \dots \oplus \tilde{\mathcal{L}}_r$ and $\tilde{\mathcal{M}}_1 \oplus \dots \oplus \tilde{\mathcal{M}}_r$. As before suppose $\tilde{E}$, $\tilde{F}$ are defined by orthonormal matrices $\left\{\tilde{V}_j\right\}_{j=1}^{r}$ and $\left\{\tilde{W}_j\right\}_{j=1}^{r}$ as in (2.6).

We set for $1 \le j \le r$

$$\begin{cases} C(\Theta_j) = \operatorname{diag}\left\{\sigma_{j1},\dots,\sigma_{j,n_j}\right\}, \ \sigma_{j1} \ge \sigma_{j2} \ge \dots \ge \sigma_{j,n_j} \ge 0 \\ S(\Theta_j) = \operatorname{diag}\left\{\mu_{j1},\dots,\mu_{j,n_j}\right\}, \ \mu_{j1} \ge \mu_{j2} \ge \dots \ge \mu_{j,n_j} \ge 0 \end{cases} \tag{5.1}$$

$$\begin{cases} (\mu_{jk})_{k=1}^{n_j} = (\sin \theta_{jk})_{k=1}^{n_j}, \\ (\sigma_{jk})_{k=1}^{n_j} = (\cos \theta_{jk})_{k=1}^{n_j}, \end{cases} \tag{5.2}$$

where $(\theta_{jk})_{k=1}^{n_j}$ are the principal angles between $\mathcal{L}_j$, $\mathcal{M}_j$. Also we set

$$\begin{cases} C(\tilde{\Theta}_j) = \operatorname{diag}\left\{\tilde{\sigma}_{j1},\dots,\tilde{\sigma}_{j,n_j}\right\}, \ \tilde{\sigma}_{j1} \ge \tilde{\sigma}_{j2} \ge \dots \ge \tilde{\sigma}_{j,n_j} \ge 0 \\ S(\tilde{\Theta}_j) = \operatorname{diag}\left\{\tilde{\mu}_{j1},\dots,\tilde{\mu}_{j,n_j}\right\}, \ \tilde{\mu}_{j1} \ge \tilde{\mu}_{j2} \ge \dots \ge \tilde{\mu}_{j,n_j} \ge 0. \end{cases} \tag{5.3}$$

A relation similar to (5.2) holds where $\left\{\tilde{\theta}_{jk}\right\}_{k=1}^{n_j}$ are the principal angles between $\tilde{\mathcal{L}}_j$, $\tilde{\mathcal{M}}_j$. The purpose of this section is to derive some perturbation inequalities for $C(\Theta_j)$ and $S(\Theta_j)$ and U(F,E) in terms of the perturbations in $V_j$ and $W_j$.

The perturbation bounds in this section will be cast in terms of unitarily invariant norms. A unitarily invariant norm on $C^{m \times n}$ is a matrix norm with the additional property that for $A \in C^{m \times n}$

$$\|PAQ\| = \|A\|,$$

if P,Q are unitaries. We shall be dealing with matrices of varying dimensions, hence we shall consider a family of unitarily invariant norms defined on

$$\bigcup_{m\ n=1}^{\infty} C^{m \times n}.$$

We refer to [15] for details about unitarily invariant norms. In particular $\|.\|_2$ wil denote the spectral norm.

The following theorem is well known [12].

**Theorem 5.1.** Let $\rho_1 \geq \rho_2 \geq \ldots \geq \rho_p$ and $\sigma_1 \geq \ldots \geq \sigma_p$ be the singular values of the matrices A, B, then

$$\|\text{diag}\,(\rho_1 - \sigma_1, \ldots, \rho_p - \sigma_p)\| \leq \|A-B\|$$

in any unitarily invariant norm.

Let $C(\Theta_j)$, $S(\Theta_j)$, $C(\tilde{\Theta}_j)$, $S(\tilde{\Theta}_j)$ be defined as in (5.1) and (5.3), then we now prove the perturbation inequalities

**Theorem 5.2**

(i) $\|E_j - \tilde{E}_j\| \leq 2 \min (\|V_j - \tilde{V}_j\|, \|V_j^{\perp} - \tilde{V}_j^{\perp}\|$

(ii) $\|C(\Theta_j)-C(\tilde{\Theta}_j)\| \leq \|V_j - \tilde{V}_j\|, \|W_j - \tilde{W}_j\|$

(iii) $\|S(\Theta_j)- S(\tilde{\Theta}_j)\| \leq \|V_j - \tilde{V}_j\| + \|W_j^{\perp} - \tilde{W}_j^{\perp}\|$

in any unitarily invariant norm.

**Proof.**

$$\|E_j - \tilde{E}_j\| = \|V_jV_j^* - \tilde{V}_j\tilde{V}_j^*\| \leq \|V_j + \tilde{V}_j\|\|V_j - \tilde{V}_j\|$$

$$\leq 2\,\|V_j - \tilde{V}_j\|$$

$$\|E_j-\tilde{E}_j\| = \|(I-E_j) - (I-\tilde{E}_j)\| = \|V_j^{\perp}\, V_j^{\perp *} - \tilde{V}_j^{\perp}\, \tilde{V}_j^{\perp *}\|$$

$$\leq 2\,\|V_j^{\perp} - \tilde{V}_j^{\perp}\|.$$

Hence (a) follows. A similar inequality holds for

$$\|F_j - \tilde{F}_j\|.$$

For part (ii), we have

$$\|W_j^*\, V_j - \tilde{W}_j^*\, \tilde{V}_j\| \leq \|\, W_j^* \quad V_j - \tilde{W}_j^*\, V_j\| +$$

$$\|\tilde{W}_j\, V_j - \tilde{W}_j^*\, \tilde{V}_j\| \leq \|W_j - \tilde{W}_j\| + \|V_j - \tilde{V}_j\|.$$

However the singular values of $W_j^*\, V_j$ are precisely the diagonal elements of $C(\Theta_j)$ (Remark 4.1), similarly for $\tilde{W}_j^*\, \tilde{V}_j$. Now we apply Theorem 5.1 to get

$$\|C(\Theta_j)-C(\tilde{\Theta}_j)\| \leq \|W_j^*\, \tilde{V}_j-\tilde{W}_j^*\, \tilde{V}_j\| \leq \|W_j - \tilde{W}_j\| + \|V_j - \tilde{V}_j\|.$$

Similarly we can prove (iii).

**Remark 5.1.** The constant in the inequality in part (i) is reduced to 1 in case of the spectral norm while it is $\sqrt{2}$ in the Frobenius norm. This is because $V_j^*\, \tilde{V}_j^{\perp}$ has the same singular values as $V_j^{\perp *}\, \tilde{V}_j$. Namely, we have

$$\|E_j - \tilde{E}_j\| = \|V_jV_j^* - \tilde{V}_j\tilde{V}_j\|$$

$$= \left\| \begin{bmatrix} V_j^* \\ V_j^{\perp *} \end{bmatrix} [V_jV_j^* - \tilde{V}_j\tilde{V}_j^*]\,[\tilde{V}_j\tilde{V}_j^{\perp}\,]\right\|$$

$$= \left\| \begin{bmatrix} 0 & V_j^*\,\tilde{V}_j^{\perp} \\ -V_j^{\perp *}\,\tilde{V}_j & 0 \end{bmatrix} \right\|.$$

In particular

$$\|E_j - \tilde{E}_j\|_2 = \|V_j^{\perp *}\, \tilde{V}_j\|_2 = \|V_j^{\perp *}\,(\tilde{V}_j-V_j)\|_2$$

$$\leq \|V_j-\tilde{V}_j\|_2.$$

Similarly

$$\|E_j - \tilde{E}_j\|_F \leq \sqrt{2}\,\|\, V_j - \tilde{V}_j\|_F.$$

Finally we present a perturbation inequality for U(F,E). For that we need the following theorem which is also of interest.

**Theorem 5.3.** Let K, $\tilde{K}$ be skew hermitian matrices, then

$$\|e^K - e^{\tilde{K}}\| \leq \|K - \tilde{K}\|$$

in any unitarily invariant norm.

**Proof.** The proof is based on the following identity

$$\frac{d}{dt}\, e^{(1-t)K}\, e^{t\tilde{K}} = -K\, e^{(1-t)K}\, e^{t\tilde{K}} + e^{(1-t)K}\, e^{t\tilde{K}}\, \tilde{K}\,.$$

This identity is introduced and used in [18]. Hence upon integration

$$e^{K} - e^{\tilde{K}} = \int_0^1 (-K\, e^{(1-t)K}\, e^{t\tilde{K}} + e^{(1-t)K}\, e^{t\tilde{K}}\, \tilde{K})dt$$

$$\| e^{K} - e^{\tilde{K}} \| \leq \int_0^1 \| e^{(1-t)K} \|_2 \| K - \tilde{K} \| \, \| e^{t\tilde{K}} \|_2 dt.$$

However $\| e^{(1-t)K} \|_2 = \| e^{t\tilde{K}} \|_2 = 1 (e^{(1-t)K}, e^{t\tilde{K}}$ being unitaries) hence

$$\| e^{K} - e^{\tilde{K}} \| \leq \| K - \tilde{K} \|.$$

In [8], U(F,E) was locally characterized, and it was shown that if K = log U(F,E), then K is the unique solution of the operator equation

$$\exp K - \sum_{j=1}^{r} F_j \exp KE_j - \sum_{j=1}^{r} \sinh K = 0$$

The above theorem shows that

$$\| U(\tilde{F}, \tilde{E}) - U(F,E) \| \leq \| K - \tilde{K} \|.$$

In case of 2-frame with $\tilde{E} = E$, let $\tilde{\theta}$ be the angel matrix between $\mathcal{M}_1$, $\tilde{\mathcal{M}}_1$; it is the same as the of $\mathcal{M}_2$, $\tilde{\mathcal{M}}_2$. Hence

$$\| U(\tilde{F},\tilde{E}) - U(F,E) \| = \| (U(\tilde{F},F) - I)\, U(F,E) \|$$
$$= \| U(\tilde{F},F) - I \| \leq \| \hat{\theta} \|.$$

The last inequality follows from Theorem 5.3.

Finally we remark that all the results in this work are still valid if we have orthogonal r-frames on a Hilbert space.

## REFERENCES


[1] COHEN, C.: An investigation of the geometry of subspaces for some multivariate statistical models, Ph.D. Thesis, Dept. of Indust. Eng., University of Illianois, Urban I 11(1969).

[2] DAHLQUIST, G.; SJÖBERG, B. and SVENSSON, P.: Comparison of averages with the method of least squares, Math. Comp. 22:833-845 (1968).

[3] DAVIS, C.: Separation of two linear subspace, Acta. Sci. Math. (Szeged) 19: 172-187 (1958).

[4] DAVIS, C.: The rotation of eigenvectors by a perturbation, J. Math. Anal. Appl. 6: 139-173(1963).

[5] DAVIS, C.; KAHAN, W.: Rotation of eigenvectors by a perturbation III, SIAM J. Number. Anal. 7:1-46(1970).

[6] HIGHAM, N.J.: Computing the polar decomposition with applications, SIAM J. Numer. Anal. 7:1-46(1970).

[7] KOVARIK, Z.V.: Manifods of frames of projectors, Linerar Algebra Appl. 31: 151-158(1980).

[8] KOVARIK, Z.V. and SHERIF, N.: Characterization of similarities between n-frames of projectors, Linerar Algebra Appl. 57:57-69(1984).

[9] KOVARIK, Z.V. and SHERIF, N.: Perturbation of invariant subspaces, Linear Algebra Appl. 64:93-113(1985).

[10] KOVARIK, Z.V. and SHERIF, N.: Goedesics and near-geodesics in the manifolds of projector frames, Linear Algebra Appl., 99:259-277(1988).

[11] McCAMMON, R.B.: Half rotation in N-dimensional Eudlidean space. Comm. Assoc. Comput. Mach. 9:688-689(1966).

[12] MIRSKY, L.: Symmetric gauge functions and unitarily invariant norms, Quart. J. Math. Oxford 11:50-59(1960).

[13] PHILIPPE, B.: An algorithm to improve nearly orthonormal set of vectors on a vector processor. SIAM J. Alg. Disc. Meth. 8(3):396-403(1987).

[14] SHERIF, N.: On the computation of a matrix inverse square root, Computing, to appear.

[15] STEWART, G.W.: On the perturbation of pseudo-inverse, projections and linear least square problems SIAM Rev. 19:634-662(1977).

[16] VARAH, J.M.: Computing invariant subspaces of a general matrix with the eigensystem is poorly conditioned, Math. Comp. 24:137-149(1970).

[17] WEDIN, P.A.: Pertubation theory for pseudo-inverses, BIT 13:217-232(1973).

[18] WERMUTH, E.M.E.: Two remarks on matrix exponentials, Linear Algebra Appl. 117:127-132(1989).


Recibido el 04 de Mayo de 1990