Mecánica Computacional Vol XXII
M. B. Rosales, V. H. Cortínez y D. V. Bambill (Editores)
Bahía Blanca, Argentina, Noviembre de 2003

FINDING THE CLOSEST PERSYMMETRIC AND SKEW-SYMMETRIC MATRICES

María G. Eberle[†] **y María C. Maciel**[†]

[†] Departamento de Matemática, Universidad Nacional del Sur,
Avda. Alem 1253, B8000CPB Bahía Blanca, Argentina. TE: 54-(0291)-4595100
Internos 3416/3422. e-mail: geberle@criba.edu.ar-immaciel@criba.edu.ar

Keywords: Symmetric Procrustes problem, persymmetric Procrustes problem, skew-symmetric Procrustes problem, singular value decomposition.

Abstract. *The class of Procrustes problems has many application in the biological, physical and social sciences just as in the investigation of elastic structures. The problem consists of solving a constrained linear least squares problem defined on a set of the space of matrices . The different problems are obtained varying the structure of the matrices belonging to the feasible set. Higham has solved the orthogonal, the symmetric and the positive definite cases. Raydán has studied the rectangular case and minimizing on a feasible set which is an intersection of convex sets of matrices. The method used is based on the alternate projection method. The Toeplitz cases has been analyzed by these authors.*
In this contribution, the theory and algorithm developed by Higham for the symmetric Procrustes problem are extended to the persymmetric and skew-symmetric cases. The singular value decomposition is used to analyze the problems and to characterized their solutions.
Numerical difficulties are discussed and illustrated by examples.

1 INTRODUCTION

Let us consider the class of constrained least squares problem

$$\begin{array}{ll} \min & \|AX - B\| \\ st & X \in \mathcal{P}. \end{array} \tag{1}$$

where $A, B \in \mathrm{I\!R}^{m\times n}$, $\mathcal{P} \subset \mathrm{I\!R}^{n\times n}$ and $\|\bullet\| \equiv \|\bullet\|_F$ denotes the Frobenius norm, defined as

$$\|A\|_F^2 = Tr(A^T A) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2.$$

Different matrix approximation problems belong to the family defined by (1). When $\mathcal{P}$ is the subspace of symmetric matrices, (1) becomes in the symmetric Procrustes problem.

$$\begin{cases} \min & \|AX - B\|_{\mathrm{F}}^2 \\ st & \\ & X = X^T. \end{cases} \tag{2}$$

Higham[1], analyzes this problem by using singular value decomposition, and gives a stable method to compute a solution.

The authors[2] have studied the problem when $\mathcal{P}$ is the subspace of Toeplitz matrices. In this contribution we consider the cases when $\mathcal{P}$ is the set of persymmetric matrices and the set of skew-symmetric matrices. The first problem is established as

$$\begin{cases} \min & \|AX - B\|_{\mathrm{F}}^2 \\ st & \\ & X = EX^T E, \end{cases} \tag{3}$$

where $E = [e_n, \cdots, e_1]$ and e_i is the i^{th} unit vector.

Clearly the skew-symmetric Procrustes problem is

$$\begin{cases} \min & \|AX - B\|_{\mathrm{F}}^2 \\ st & \\ & X = -X^T. \end{cases} \tag{4}$$

The paper is organized as follows. In the next section we present the symmetric problem proposed by Higham. In section 3 we analyze the persymmetric Procrustes problem. Section 4 is devoted to present skew-symmetric case and characterize the set of solutions. In section 5 a framework of the algorithm is presented and some numerical results. This work concludes with final remarks .

2 THE SYMMETRIC PROCRUSTES PROBLEM

The problem (2) with $A,\ B \in \mathrm{I\!R}^{m\times n},\ m > n$, has been studied by Higham[1].

Let us consider the singular value decomposition of A

$$A = P \begin{bmatrix} \sum \\ 0 \end{bmatrix} Q^T,$$

where $P \in \mathrm{I\!R}^{m\times m}$ and $Q \in \mathrm{I\!R}^{n\times n}$ are orthogonal matrices and $\sum = diag(\sigma_1, \cdots, \sigma_n)$, $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$.

From the fact $\|QAP\|_{\mathrm{F}} = \|A\|_{\mathrm{F}}$, the problem (2) becomes

$$\left\{ \begin{array}{ll} \min & \| \sum Y - C_1\|_{\mathrm{F}}^2 \\ st & \\ & Y = Y^T, \end{array} \right. \tag{5}$$

where

$$Y = Q^T X Q,$$

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = P^T B Q,$$

$$C_1 \in \mathrm{I\!R}^{n\times n}.$$

The solution of the equivalent problem (5) is

$$(Y_*)_{ij} = \left\{ \begin{array}{ll} \dfrac{\sigma_i (C_1)_{ij} + \sigma_j (C_1)_{ji}}{{\sigma_i}^2 + {\sigma_j}^2} & if \quad {\sigma_i}^2 + {\sigma_j}^2 \neq 0 \\ arbitrary & otherwise, \end{array} \right.$$

and therefore the solution of (2) is given by $X_* = QY_*Q^T$.

Higham[1] characterizes the set of solutions of the problem and gives stability conditions.

3 THE PERSYMMETRIC PROCRUSTES PROBLEM

We remind when a matrix is said to be persymmetric. More details about these matrices can be found in the book of Golub and van Loan[3].

Definition 1 (Persymmetric matrix)
A matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ *is said to be persymmetric if it is symmetric about the NE-SW diagonal,*

$$a_{ij} = a_{n-j+1,\,n-i+1}, \qquad i,j = 1,\cdots,n.$$

The definition means that a persymmetric matrix A can be written as

$$A \;=\; EA^TE \tag{6}$$

where E is the matrix $E = [e_n \cdots e_1]$ and e_i is the i^{th} unit vector.

Let us consider the persymmetric Procrustes problem (3). We begin considering the singular value decomposition of the matrix A,

$$A = P\begin{bmatrix}\Sigma\\0\end{bmatrix}Q^T,$$

where $P \in \mathbb{R}^{m\times m}$, $Q \in \mathbb{R}^{n\times n}$ are orthogonal matrices and the matrix $\Sigma = diag(\sigma_1,\cdots,\sigma_n) \in \mathbb{R}^{n\times n}$ with $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$.

Because of the invariance of the Frobenius norm under orthogonal transformations we can write

$$\begin{aligned}\|AX-B\|^2 &= \left\|P\begin{bmatrix}\Sigma\\0\end{bmatrix}Q^TX-B\right\|^2 = \left\|P^T\left(P\begin{bmatrix}\Sigma\\0\end{bmatrix}Q^TX-B\right)Q\right\|^2\\ &= \left\|\begin{bmatrix}\Sigma\\0\end{bmatrix}Q^TXQ-P^TBQ\right\|^2 = \left\|\begin{bmatrix}\Sigma\\0\end{bmatrix}Y-Z\right\|^2\\ &= \|\Sigma Y-Z_1\|^2 - \|Z_2\|^2,\end{aligned}$$

where

$$\begin{aligned}Y &= Q^TXQ,\\ Z &= \begin{bmatrix}Z_1\\Z_2\end{bmatrix} = P^TBQ,\\ Z_1 &\in \mathbb{R}^{n\times n}.\end{aligned}$$

Observe that, if $X = EX^TE$ and $Y = Q^TXQ$, the matrix Y satisfies

$$Y = UY^TU,$$

where $U = Q^T E Q \in \mathbb{R}^{n\times n}$ is an orthogonal and symmetric matrix. The matrix Y also satisfies $YU = U^T = (YU)^T$.

The objective function can be written in terms of the symmetric matrix $S = YU$, as follows

$$\begin{aligned} \|\Sigma Y - Z_1\|^2 &= \|(\Sigma Y - Z_1)U\|^2 = \|\Sigma YU - Z_1 Y\|^2 \\ &= \|\Sigma S - \bar{Z}_1\|^2. \end{aligned}$$

Therefore the problem (3) becomes

$$\begin{aligned} \min \quad & \|\Sigma S - \bar{Z}_1\| \\ st \quad & S = S^T. \end{aligned} \tag{7}$$

The last problem is the symmetric Procrustes problem. Then, according with Higham[1], we obtain the solution $S^\star$ whose elements are

$$s^\star_{ij} = \begin{cases} \frac{\sigma_i z_{ij} + \sigma_j z_{ji}}{\sigma_i^2 + \sigma_j^2} & if \quad \sigma_i^2 + \sigma_j^2 \neq 0 \\ arbitrary & \quad otherwise, \end{cases}$$

then $Y^\star = S^\star U$ and we can obtain the solution of the problem (3) as $X^\star = QY^\star Q^T$.

3.1 Characterization of the solution.

Let us define

$$\mathcal{S} = \{X \in \mathbb{R}^{n\times n}, X = EX^T E, X = arg\, min\|AZ - B\|, \ \forall\, Z \in \mathcal{P}\}.$$

The following lemma gives some properties of the solution set $\mathcal{S}$. They are the corresponding properties for the linear least-squares problems[3] and the symmetric Procrustes problem[1].

Lemma 1 *Let $\mathcal{P}$ and $\mathcal{S}$ defined as above. Then*

(a) $\mathcal{S}$ is a convex set.

(b) $X \in \mathcal{S}$ if and only if $X = EX^T E$ and

$$A^T(XE) + (XE)A^T A = A^T(BE) + (BE)^T A. \tag{8}$$

(c) $\mathcal{S}$ has a unique element X^{MN} of minimal Frobenius norm.

(d) $S = X^{MN}$ *if an only if rank(A) =n.*

The details of Lemma 1 can be found in Eberle[4]. Now we discuss the behavior of the solution when the set of data has been perturbed. It is a theorem similar to theorem 1. given by Higham[1], (page 137).

Theorem 1 *Let* $A,\ B \in \mathbb{R}^{m\times n}$ *be matrices with* $m \geq n$, $rank(A) = n$.
Let X *be a solution of problem (3), and* $\hat{X}$ *the solution of the perturbed problem*

$$\left\{ \begin{array}{ll} \min & \|(A+\Delta A)X - (B+\Delta B)\|^2 \\ st & \\ & X = EX^TE, \end{array} \right. \tag{9}$$

where $\Delta A,\ \Delta B \in \mathbb{R}^{m\times n}$, *such that* $rank(A+\Delta A) = rank(A) = n$. *Let*

$$R = AX - B$$

$$\widehat{R} = (A+\Delta A)\hat{X} - (B+\Delta B),$$

$$\epsilon_A = \frac{\|\Delta A\|_2}{\|A\|_2},$$

$$\kappa = \kappa_2(A),$$

be such that $\kappa\, \epsilon_A < 1$. *Then*

$$\|X - \hat{X}\| \leq \frac{\kappa}{1-\epsilon_A}\left(\epsilon_A\|X\| + \frac{\|\Delta B\|}{\|A\|_2} + \kappa\epsilon_A\frac{\|R\|}{\|A\|_2}\right) + \kappa\epsilon_A\|X\| \tag{10}$$

$$\|R - \widehat{R}\| \leq \epsilon_A\|X\|\ \|A\|_2 + \|\Delta B\| + \kappa\ \epsilon_A\|R\|. \tag{11}$$

Theorem 2 *Let the persymmetric Procrustes problem (3) be, and assume that* $rank(A) = n$. *Let* $X_\star$ *be its solution. Then the residual can be written as:*

$$\rho^2{}_{\mathrm{pp}} = \|AX_\star - B\|^2 = \sum_{j>i} \frac{(\sigma_i(C_1)_{ji} - \sigma_j(C_1)_{ij})^2}{{\sigma_i}^2 + {\sigma_j}^2} + \|C_2\|^2,$$

where

$$C = \left[\begin{array}{c} C_1 \\ C_2 \end{array} \right] = P^T(BE)Q \in \mathbb{R}^{m\times n},$$

$P,\ Q$ *y* $\sum = diag(\sigma_1, \cdots, \sigma_n)$, *are the matrices involved in the singular values decomposition of the matrix* A.

The proof of the previous result can be consulted in the Eberle's thesis[4].

4 THE SKEW-SYMMETRIC PROCRUSTES PROBLEM

This section is devoted to solve the skew-symmetric problem (4). Similarly to the previous cases, we consider the singular value decomposition of the matrix A

$$A = P \begin{bmatrix} \sigma \\ 0 \end{bmatrix} Q^T.$$

Clearly, solve the problem (4) is equivalent to find a solution to the problem

$$\begin{cases} \min & \|\Sigma Y - C_1\|_F^2 \\ st & \\ & Y = -Y^T, \end{cases} \tag{12}$$

where

$$Y = Q^T X Q,$$

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = P^T B Q,$$

$$C_1 \in \mathbb{R}^{n\times n}.$$

Since $Y = Q^T X Q$ we obtain $Y = Q^T X Q = Q^T(-X)Q = -(Q^T X Q) = -(Q^T X Q)^T = -Y^T$.

4.1 Solving the problem

In order to find a minimizer of the functional $\|\Sigma Y - C_1\|_F^2$ with the condition that Y be skew-symmetric, let us consider

$$\begin{aligned} \|\Sigma Y - C_1\|_F^2 &= Tr((\Sigma Y - C_1)^T(\Sigma Y - C_1)) \\ &= Tr(Y^T(\sigma^T\Sigma)Y) - 2Tr((\Sigma Y)^T C_1) + Tr(C_1^T C_1) \\ &= \sum_{i=1}^n \left[\sum_{j=1}^n {\sigma_j}^2 (Y^T)_{ij} Y_{ji}\right] - 2\sum_{i=1}^n \left[\sum_{j=1}^n Y_{ij}\sigma_j (C_1)_{ji}\right] + \sum_{i=1}^n ((C_1)_{ii})^2 \end{aligned} \tag{13}$$

The assumptions $(Y^T)_{ik} = -(Y)_{ik}$, and $Y_{ki} = (Y^T)_{ik} = -(Y)_{ik}$, allow that problem (4) be written as follows

$$\begin{aligned} \|\Sigma Y - C_1\|_{\mathrm{F}}^2 &= \sum_{i=1}^{n}\left[\sum_{j=1}^{n}{\sigma_j}^2 (Y_{ij})^2\right] - 2\sum_{i=1}^{n}\left[\sum_{j=1}^{n} Y_{ij}\sigma_j (C_1)_{ji}\right] + \sum_{i=1}^{n}((C_1)_{ii})^2 \\ &= F(Y_{ij}), \qquad 1 \le i,j \le n. \end{aligned}$$

Computing the derivative with respect to each Y_{ij}, we obtain

$$\frac{\partial F}{\partial Y_{ij}} = 2({\sigma_i}^2 + {\sigma_j}^2)Y_{ij} + 2\sigma_j(C_1)ij - 2\sigma_i(C_1)ji.$$

From the first order necessary condition we obtain

$$Y_{ij} = \frac{\sigma_i(C_1)_{ij} - \sigma_j(C_1)_{ji}}{{\sigma_i}^2 + {\sigma_j}^2}.$$

Because of

$$\frac{\partial^2 F}{\partial Y_{ij}^2} = 2({\sigma_i}^2 + {\sigma_j}^2) > 0,$$

the matrix Y is a minimizer of F .

Therefore a solution of the problem (12) will be

$$(Y_*)_{ij} = \begin{cases} \dfrac{\sigma_i(C_1)_{ij} - \sigma_j(C_1)_{ji}}{{\sigma_i}^2 + {\sigma_j}^2} & if \quad {\sigma_i}^2 + {\sigma_j}^2 \neq 0 \\ arbitrary & otherwise, \end{cases} \tag{14}$$

and a solution of (4) is $X_* = QY_*Q^T$.

4.2 Characterization of the solution.

Let us define

$$\mathcal{S} = \{X \in \mathbb{R}^{n\times n}, X = -X^T, X = arg\ min\|AZ - B\|, \ \ for\ all\ Z = -Z^T\}.$$

The following lemma gives the properties of the solution set $\mathcal{S}$. They are the corresponding properties for the linear least-squares problems[3], the symmetric Procrustes problem[1] and the persymmetric Procrustes problem[4].

Lemma 2 *Let $\mathcal{P}$ and $\mathcal{S}$ defined as above. Then*

(a) $X \in \mathcal{S}$ if and only if $X = -X^T$ and

$$A^T(XE) + (XE)A^T A \;=\; A^T(BE) - (BE)^T A. \tag{15}$$

(b) $\mathcal{S}$ is a convex set.

(c) $\mathcal{S}$ has a unique element X^{MN} of minimal Frobenius norm.

(d) $S = X^{MN}$ if an only if rank(A) =n.

Proof
(a) Suppose that $X \in \mathcal{S}$, clearly $X = -X^T$. In order to prove (15), let us consider X a solution of (4), then $X = QYQ^T$ where Y is obtained as (14). Then

$$\begin{aligned}
A^TAX + XA^TA &= A^TA(QYQ^T) + (QYQ^T)A^TA \\
&= (P\Sigma Q^T)^T(P\Sigma Q^T)QYQ^T + QYQ^T(P\Sigma Q^T)^T(P\Sigma Q^T) \\
&= Q\Sigma P^T P\Sigma Q^T QYQ^T + QYQ^T Q\Sigma P^T P\Sigma Q^T \\
&= Q\Sigma^2 YQ^T + QY\Sigma^2 Q^T = Q(\Sigma^2 Y + Y\Sigma^2)Q^T.
\end{aligned}$$

Considering the ij-element of the previous matrix:

$$\begin{aligned}
(A^TAX + XA^TA)_{ij} &= (Q(\Sigma^2Y + Y\Sigma^2)Q^T)_{ij} \\
&= \sum_{k=1}^{n} Q_{ik}[(\Sigma^2Y + Y\Sigma^2)Q^T(\Sigma^2Y + Y\Sigma^2)Q^T]_{kj} \\
&= \sum_{k=1}^{n} Q_{ik}\left[\sum_{l=1}^{n}(\Sigma^2Y + Y\Sigma^2)_{kl}(Q^T)_{lj}\right] \\
&= \sum_{k=1}^{n} Q_{ik}\left[\sum_{l=1}^{n}[(\Sigma^2Y)_{kl} + (Y\Sigma^2)_{kl}](Q^T)_{lj}\right] \\
&= \sum_{k=1}^{n} Q_{ik}\left[\sum_{l=1}^{n}({\sigma_k}^2Y_{kl} + Y_{kl}{\sigma_l}^2)(Q^T)_{lj}\right] \\
&= \sum_{k=1}^{n} Q_{ik}\left[\sum_{l=1}^{n}[Y_{kl}({\sigma_k}^2 + {\sigma_l}^2)](Q^T)_{lj}\right] \\
&= \sum_{k=1}^{n} Q_{ik}\left[\sum_{l=1}^{n}\left(\frac{\sigma_i(C_1)_{ij} - \sigma_j(C_1)_{ji}}{{\sigma_i}^2 + {\sigma_j}^2}\right)({\sigma_k}^2 + {\sigma_l}^2)(Q^T)_{lj}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} ((C_1)_{ij} - \sigma_j (C_1)_{ji})(Q^T)_{lj} \right] \\
&= \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} [\sigma_k (P^T BQ)_{kl} - \sigma_l (P^T BQ)_{lk}](Q^T)_{lj} \right] \\
&= \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_k (P^T BQ)_{kl} (Q^T)_{lj} \right] - \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_l (P^T BQ)_{lk} (Q^T)_{lj} \right]. \qquad (16)
\end{aligned}
$$

The first sum of the right hand side of (16) is

$$
\begin{aligned}
\sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_k (P^T BQ)_{kl} (Q^T)_{lj} \right] &= \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_k \sum_{s=1}^{n} (P^T)_{ks} (BQ)_s l (Q^T)_{lj} \right] \\
&= \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_k \sum_{s=1}^{n} (P^T)_{ks} (\sum_{r=1}^{n} B_{sr} Q_{rl})(Q^T)_{lj} \right] \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{s=1}^{n} Q_{ik} \sigma_k (P^T)_{ks} (\sum_{r=1}^{n} B_{sr} Q_{rl} (Q^T)_{lj}) \\
&= \sum_{s=1}^{n} \left(\sum_{k=1}^{n} Q_{ik} \sigma_k (P^T)_{ks} \right) \sum_{r=1}^{n} B_{sr} \sum_{l=1}^{n} Q_{rl} (Q^T)_{lj}) \\
&= \sum_{s=1}^{n} (A^T)_{is} B_{sj} = (A^T B)_{ij}.
\end{aligned}
$$

Similarly, the second term of the right hand side of (16):

$$
\begin{aligned}
\sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_l (P^T BQ)_{lk} (Q^T)_{lj} \right] &= \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_l \sum_{s=1}^{n} (P^T)_{ls} (BQ)_{sk} (Q^T)_{lj} \right] \\
&= \sum_{k=1}^{n} Q_{ik} \left[\sum_{l=1}^{n} \sigma_l \sum_{s=1}^{n} (P^T)_{ls} (\sum_{r=1}^{n} B_{sr} Q_r k)(Q^T)_{lj} \right] \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} \sum_{s=1}^{n} (P^T)_{ls} \sigma_l (Q^T)_{lj} (\sum_{r=1}^{n} B_{sr} Q_{rk} Q_{ik}) \\
&= \sum_{s=1}^{n} \left(\sum_{l=1}^{n} (P^T)_{ls} \sigma_l (Q^T)_{lj} \right) \sum_{r=1}^{n} B_s r \sum_{k=1}^{n} Q_{ik} (Q^T)_{kr}) \\
&= \sum_{s=1}^{n} A_{sj} (B^T)_{is} = (B^T A)_{ij}.
\end{aligned}
$$

Conversely, if $X = -X^T$ and $A^TAX + XA^TA = A^TB - B^TA$, let us define

$$F(X) = \|AX - B\|_F^2 = Tr((AX - B)^T(AX - B))$$

and prove that $\forall E \in \mathbb{R}^{n\times n}$, with $E = -E^T$ the inequality:

$$F(X + E) \geq F(X).$$

holds.
Since

$$\begin{aligned} F(X+E) - F(X) &= Tr((A(X+E) - B)^T(A(X+E) - B)) \\ &\quad - Tr((AX - B)^T(AX - B)). \end{aligned} \tag{17}$$

From the fact $(X + E)^T = -(X + E)$,

$$Tr(((X + E)^TA^T - B^T)(A(X + E) - B))) = -Tr(((X + E)A^T + B^T)(A(X + E) - B))$$

then

$$\begin{aligned} F(X+E) &= -Tr(XA^TAX + XA^TAE - XA^TB + EA^TAX + EA^TAE \\ &\quad -EA^TB + B^TAX + B^TAE - B^TB) \\ &= -Tr(XA^TAX) - Tr(XA^TAE) + Tr(XA^TB) - Tr(EA^TAX) \\ &\quad -Tr(EA^TAE) + Tr(EA^TB) - Tr(B^TAX) - Tr(B^TAE) + Tr(B^TB) \\ &= \Big(-Tr(XA^TAE) - Tr(EA^TAX) - Tr(EA^TAE) + Tr(EA^TB) - Tr(B^TAE)\Big) \\ &\quad - \Big(Tr(XA^TAX - XA^TB - B^TAX - B^TB)\Big) \\ &= -Tr(XA^TAE) - Tr(A^TAXE) + Tr(A^TBE) - Tr(B^TAE) \\ &\quad -Tr(EA^TAE) - Tr(XA^TAX - XA^TB - B^TAX - B^TB). \end{aligned}$$

The second term of the left hand side of (17) is

$$\begin{aligned} F(X) &= Tr((AX - B)^T(AX - B)) = Tr((-XA^T - B^T)(AX - B)) \\ &= -Tr((XA^T + B^T)(AX - B)) = -Tr(XA^TAX - XA^TB + B^TAX - B^TB) \end{aligned}$$

then the difference (17) becomes

$$\begin{aligned} F(X+E)-F(X) &= -Tr(XA^TAE)-Tr(A^TAXE)+Tr(A^TBE)-Tr(B^TAE) \\ &\quad -Tr(EA^TAE) \\ &= -Tr[(XA^TA+A^TAX-A^TBE+B^TA)E] \\ &\quad -Tr(-E^TA^TAE) \end{aligned}$$

$$\begin{aligned} &= -Tr[((XA^TA+A^TAX)-(A^TBE-B^TA))E]+Tr((AE)^T(AE)) \\ &= -Tr[0.E]+\|AE\|_{\mathtt{F}}^2 \\ &= \|AE\|_{\mathtt{F}}^2 > 0, \end{aligned}$$

which allows us to guarantee that X minimizes F and hence $X \in \mathcal{S}$.

(b) $\mathcal{S}$ is a convex set because if $X_1,\ X_2 \in \mathcal{S},\ \ \mu \in \mathbb{R}$, then $(\mu X_1 + (1-\mu)X_2) \in \mathcal{S}$:

$$A^TA((\mu X_1+(1-\mu)X_2))+((\mu X_1+(1-\mu)X_2))A^TA = A^TB-B^TA$$

$$\begin{aligned} A^TA((\mu X_1)+(1-\mu)X_2) &= A^TAX_1\mu + A^TAX_2(1-\mu) \\ &= A^TAX_1\mu + A^TAX_2 - A^TAX_2\mu \\ ((\mu X_1)+(1-\mu)X_2)A^TA &= X_1\mu A^TA + (1-\mu)X_2A^TA \\ &= X_1\mu A^TA + X_2A^TA - \mu X_2A^TA. \end{aligned}$$

Adding

$$A^TA(\mu X_1+(1-\mu)X_2)+(\mu X_1+(1-\mu)X_2)A^TA = A^TB-B^TA$$

(c) If in the definition of the elements Y_{ij} the arbitrary values are defined equal to zero, the minimum Frobenius norm element is obtained.

(d) If $rank(A)=n$, the singular values are non zero and the unique solution is X^{MN}.

5 NUMERICAL EXPERIENCE

It is clear that the computational procedure which allows to find the solutions for both problems (3) and (4) is simple. The strategies have been implemented in Fortran, by

using FORTRAN POWER STATION (1993) in an environment PC with a processor Pentium(R2) Intelmmx(TM) Technology, 31.0 MB de RAM. and 32 bits of virtual memory. The singular value decomposition of A was computed by using a variation of the Golub and Reinch[5] algorithm. The LINPACK subroutines have been used to make it.
Given the persymmetric Procrustes problem (3) we consider several sizes of the matrices A and B.

Example 5.1 Let A and B be as follows:

$$A = \begin{bmatrix} 5 & 3 & 2 \\ 1 & 2 & 4 \\ 6 & 0 & 3 \\ -1 & 2 & -3 \end{bmatrix}, \qquad B = \begin{bmatrix} 15 & 10 & -3 \\ 1 & 5 & 3 \\ 15 & 6 & -3 \\ 2 & 3 & -2 \end{bmatrix}.$$

The solution of the persymmetric Procrustes problem is

$$X_\star = \begin{bmatrix} -.9896 & 0.9203 & 2.9339 \\ 0.0315 & 1.8791 & 0.9203 \\ .9838 & 0.0315 & -0.9896 \end{bmatrix}.$$

Example 5.2 Now let A and B be as follows:

$$A = \begin{bmatrix} -1 & 2 & 3 & -2 \\ 1 & 4 & 5 & -1 \\ 0 & 2 & 3 & 7 \\ -1 & 2 & -1 & 0 \\ 3 & 4 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 12 & 0 & 1 & 3 \\ 1 & -1 & 2 & -2 \\ 1 & 3 & 2 & 1 \\ 0 & 2 & 1 & 10 \\ 1 & 1 & 2 & 3 \\ -4 & 3 & -1 & 2 \\ 10 & 9 & 0 & 1 \end{bmatrix}.$$

The solution matrix is

$$X_\star = \begin{bmatrix} -6.2156 & 6.0507 & 6.6790 & -9.6773 \\ 7.9559 & -1.4131 & 4.7940 & 6.6790 \\ -2.6760 & 4.9199 & -1.4131 & 6.0507 \\ 6.4650 & -2.6760 & 7.9559 & -6.2156 \end{bmatrix}.$$

For the same choice of matrices we find the solution of (4).
With the matrices of example 5.1 the solution is

$$X^* = \begin{bmatrix} 2.9339 & -0.9203 & .9896 \\ 0.9203 & 1.8791 & -0.0315 \\ -0.9896 & 0.0315 & .9838 \end{bmatrix}.$$

The solution of (4) by using the matrices of example 5.2 is

$$X_\star = \begin{bmatrix} -9.6773 & -6.6790 & -6.0507 & -6.2156 \\ 6.6790 & 4.7940 & 1.4131 & 7.9559 \\ 6.0507 & -1.4131 & 4.9199 & 2.6760 \\ 6.2156 & -7.9559 & -2.6760 & 6.4650 \end{bmatrix}.$$

More examples of larger sizes can be formed in Eberle[4].

REFERENCES

[1] N.J. HIGHAM. The symmetric Procrustes problem. *BIT*, 28:133–143, 1988.

[2] M.G. EBERLE and M.C. MACIEL Finding the closest Toeplitz matrix. Accepted for publication in *Computational and Applied Mathematics*, 2003.

[3] G.H. GOLUB and Ch.F. VAN LOAN. *Matrix Computations.* Johns Hopkins University Press, Baltimore, Maryland, 1989.

[4] M.G. EBERLE. Métodos de proyecciones alternas en problemas de optimización persimétricos. Master's thesis, Dept. de Matemática, Universidad Nacional del Sur, Bahía Blanca, Argentina, 1999.

[5] G.H. GOLUB and C. REINSCH. Singular value decomposition and least squares problems. *Numerische Mathematik*, 14:403–420, 1970.