BIT 33 (1993), 687–694.

PERTURBATION ANALYSIS OF THE ORTHOGONAL PROCRUSTES PROBLEM

INGE SÖDERKVIST

Department of Computing Science, University of Umeå, S-90187 Umeå, Sweden.
E-mail: inge@cs.ume.se

Abstract.

Given two arbitrary real matrices A and B of the same size, the orthogonal Procrustes problem is to find an orthogonal matrix M such that the Frobenius norm $\|MA - B\|$ is minimized. This paper treats the common case when the orthogonal matrix M is required to have a positive determinant. The stability of the problem is studied and supremum results for the perturbation bounds are derived.

AMS subject classification: 65K05, 65F35, 65G99.

Keywords: Condition number, Orthogonal Procrustes problem, Perturbation bounds, Singular values.

1. Introduction.

Given any matrices A and $B \in R^{m \times n}$, the orthogonal Procrustes problem is formulated as

$$\min_{M \in \Omega} \|MA - B\|, \tag{1.1}$$

where the norm $\|\cdot\|$ is the matrix *Frobenius norm* and Ω denotes the set of orthogonal m by m matrices. Green [4] solved the problem when A and B both are assumed to have full row rank. The general problem was solved by Schönemann [9] and it has later also been treated by Hanson and Norris [5]. By using the singular value decomposition, $BA^T = U\Sigma_Z V^T$, of the matrix $Z = BA^T$, the solution is given by the *orthogonal polar factor* of Z, i.e.,

$$\hat{M} = UV^T. \tag{1.2}$$

The solution is unique provided the smallest singular value, $\sigma_m(Z)$, of Z is nonzero.

Hanson and Norris [5] and Wahba [13] treat the problem when M is required to have a positive determinant, i.e., the set Ω in (1.1) is replaced by a set Ω_+ defined as

$$\Omega_+ = \{M \in R^{m \times m} \mid M^T M = I;\ \det(M) = +1\}.$$

Received June 1992. Revised December 1992.

We will focus our attention on this problem which arises in many applications where M corresponds to a rotation, for example in kinematics when determining a rotational movement of a body, see e.g., [10]. To express the solution explicitly we have to consider two different cases:

1. $\det(UV^T) = +1$. The solution is given by $\hat{M} = UV^T$ and it is unique if $\sigma_{m-1}(Z) \neq 0$.
2. $\det(UV^T) = -1$. The solution is given by $\hat{M} = U\,\mathrm{diag}(1,\ldots,1,-1)V^T$ and it is unique if $\sigma_{m-1}(Z) \neq \sigma_m(Z)$.

We will study the perturbed problem

$$\min_{M \in \Omega_+} \|M(A + \Delta A) - (B + \Delta B)\|, \tag{1.3}$$

to investigate how the solution $\hat{M} + \Delta M$ depends on the perturbation matrices $\Delta A = \{\delta a_{i,j}\}$ and $\Delta B = \{\delta b_{i,j}\} \in R^{m \times n}$.

When the solution of problem (1.3) is given by the polar decomposition of the matrix $(B + \Delta B)(A + \Delta A)^T$ (case 1 above), it is of course possible to use the well-known bounds for the polar factors, see e.g., [1, 2, 6, 8 and 14]. But by utilizing the special structure of the problem it is possible to derive much sharper bounds.

In section 2, we prove a supremum result for $\|\Delta M\|$, expressed by the residual $\|\hat{M}A - B\|$, and by the amount of perturbations $\|\Delta A\|$ and $\|\Delta B\|$. In section 3 we show how the problem is connected to the *skew symmetric Procrustes problem*. This connection is used on problems with zero residual to derive the same first order bound as in section 2.

2. The perturbation bound.

As shown in section 1, a necessary condition for the unperturbed problem to have a unique solution is that the second smallest singular value of the matrix BA^T is nonzero. Thus, we are only interested in cases where $n \geq m - 1$. However, to simplify notations we will assume $n \geq m$. This is no restriction since we can append a zero column to A and B, if necessary, without changing the problem.

The following theorem gives a supremum result for the error $\|\Delta M\|$.

THEOREM 2.1 *Given a matrix $A \in R^{m \times n}$, $m \leq n$, with singular values $\sigma_1 \geq \ldots \geq \sigma_m$, and a positive number γ, satisfying $\gamma < (\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}$.*

By introducing the set

$$\Gamma = \{B \in R^{m \times n} \mid \min_{M \in \Omega_+} \|MA - B\| \leq \gamma\},$$

for any matrix $B \in \Gamma$ we can define the matrices $\hat{M}$ and $\tilde{M}$ as

$$\hat{M} = \arg\min_{M \in \Omega_+} \|MA - B\|,$$

$$\tilde{M} = \arg\min_{M \in \Omega_+} \|M(A + \Delta A) - (B + \Delta B)\|,$$

where the perturbation matrices $\Delta A \in R^{m\times n}$ *and* $\Delta B \in R^{m\times n}$ *are bounded as*

$$\|\Delta A\| \leq \varepsilon_A < (\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}, \quad \|\Delta B\| \leq \varepsilon_B < (\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}} - \gamma.$$

The perturbation $\Delta M = \tilde{M} - \hat{M}$ *is then bounded as*

$$\sup_{B\in\Gamma} \sup \|\Delta M\| = 2\sqrt{[1 - ((1-\kappa_A^2\varepsilon_A^2)^{\frac{1}{2}}(1-\kappa_B^2\varepsilon_B^2)^{\frac{1}{2}} - \kappa_A\kappa_B\varepsilon_A\varepsilon_B)]}, \tag{2.1}$$

the second supremum taken over $\|\Delta A\| \leq \varepsilon_A$ and $\|\Delta B\| \leq \varepsilon_B$, *and where* κ_A *is* $1/(\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}$ *and* κ_B *is* $1/((\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}} - \gamma)$.

By using the Taylor expansion of (2.1) and by applying the theorem on some interesting special cases we get the following results.

COROLLARY 2.1 *With the same assumption and definitions as in Theorem* 2.1, *define the function* $h(\varepsilon_A, \varepsilon_B, \gamma) = \sup_{B\in\Gamma} \sup \|\Delta M\|$, *the second supremum taken over* $\|\Delta A\| \leq \varepsilon_A$ *and* $\|\Delta B\| \leq \varepsilon_B$. *The function* $h(\varepsilon_A, \varepsilon_B, \gamma)$ *then satisfies*

$$h(\varepsilon_A, \varepsilon_B, \gamma) = \sqrt{2}(\kappa_A\varepsilon_A + \kappa_B\varepsilon_B) + O((\varepsilon_A + \varepsilon_B)^3), \tag{2.2}$$

$$h(\varepsilon_A, 0, \gamma) = 2\sqrt{[1 - (1-\kappa_A^2\varepsilon_A^2)^{\frac{1}{2}}]} = \sqrt{2}\,\kappa_A\varepsilon_A + O(\varepsilon_A^3), \tag{2.3}$$

$$h(0, \varepsilon_B, \gamma) = 2\sqrt{[1 - (1-\kappa_B^2\varepsilon_B^2)^{\frac{1}{2}}]} = \sqrt{2}\,\kappa_B\varepsilon_B + O(\varepsilon_B^3). \tag{2.4}$$

For problems with zero residual ($\gamma = 0$), *we observe that* κ_A *equals* κ_B. *When the perturbation matrices are equal* ($\varepsilon_A = \varepsilon_B = \varepsilon$), *the function* $h(\varepsilon, \varepsilon, 0)$ *is the linear function*

$$h(\varepsilon, \varepsilon, 0) = 2\sqrt{2}\,\kappa_A\varepsilon. \tag{2.5}$$

To prove Theorem 2.1 we need the following lemma.

LEMMA 2.1 *Given the matrices* $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2)$, $B \in R^{2\times 2}$, *and a positive number* γ, *satisfying* $\gamma \leq \|\Sigma\|$. *If* $\|\Sigma - B\| \leq \gamma$ *then*

$$\sigma_1 b_{1,1} + \sigma_2 b_{2,2} \geq \|\Sigma\|(\|\Sigma\| - \gamma). \tag{2.6}$$

PROOF. Take $B = \Sigma + R$, where $\|R\| \leq \gamma$. Then

$$\sigma_1 b_{1,1} + \sigma_2 b_{2,2} = \sigma_1^2 + \sigma_2^2 + \sigma_1 r_{1,1} + \sigma_2 r_{2,2} \geq \|\Sigma\|^2 - \|\Sigma\|\gamma. \quad \blacksquare$$

PROOF OF THEOREM 2.1. We give a strict proof for the case when ε_A is zero, followed by a geometrical arguing to obtain the general result.

Without loss of generality we can assume that $\hat{M} = I$ and $A = \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_m)$, where $\sigma_1 \geq \ldots \geq \sigma_m \geq 0$, see [11] for details.

Let us first consider the 2-dimensional case. The matrix $\tilde{M} = I + \Delta M$ can then be written as $I + \Delta M = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$, and it is easy to see that $\|\Delta M\|$ satisfies

$$\|\Delta M\|^2 = 4 - 4\cos(\phi). \tag{2.7}$$

By computing $F(\phi) = \|(I + \Delta M)\Sigma - (B + \Delta B)\|$ and utilizing the necessary condition for minimum, $\partial F/\partial\phi = 0$, we get the following equation:

$$\tan(\phi) = \frac{\sin(\phi)}{\cos(\phi)} = \frac{r + \sigma_1 \delta b_{2,1} - \sigma_2 \delta b_{1,2}}{q + \sigma_1 \delta b_{1,1} + \sigma_2 \delta b_{2,2}},$$

where $r = \sigma_1 b_{2,1} - \sigma_2 b_{1,2}$ and $q = \sigma_1 b_{1,1} + \sigma_2 b_{2,2}$. Since $\phi = 0$ is assumed to be the solution to the unperturbed problem, it follows that r vanishes.

By defining

$$\alpha = (\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}, \tag{2.8}$$

we get from Lemma 2.1 that q satisfies $q \geq \alpha(\alpha - \gamma)$. Since $\|\Delta B\|$ is less than $\alpha - \gamma$, we conclude that $\tan(\phi)$ is bounded.

It is no restriction to assume that ϕ is nonnegative. Thus, both $\tan(\phi)$ and $\|\Delta M\|$ are increasing functions of ϕ and the worst perturbation ΔB is identified by solving

$$\max_{\|\Delta B\| \leq \varepsilon_B} \frac{\sigma_1 \delta b_{2,1} - \sigma_2 \delta b_{1,2}}{q + \sigma_1 \delta b_{1,1} + \sigma_2 \delta b_{2,2}}. \tag{2.9}$$

The solution to this problem satisfies

$$\Delta B = \begin{bmatrix} -\sigma_1 \sin(\theta) & -\sigma_2 \cos(\theta) \\ \sigma_1 \cos(\theta) & -\sigma_2 \sin(\theta) \end{bmatrix} \varepsilon_B/\alpha. \tag{2.10}$$

To identify the angle θ, corresponding to a maximum, we observe that problem (2.9) becomes

$$\max_{\theta} \frac{\alpha \varepsilon_B \cos(\theta)}{q - \alpha \varepsilon_B \sin(\theta)}, \tag{2.11}$$

which is solved by $\sin(\theta) = \alpha\varepsilon_B/q$. The worst perturbation ΔB in (2.10) is now completely determined. For this worst perturbation, the value of $\tan(\phi)$ is

$$\tan(\phi) = \frac{\alpha \varepsilon_B/q}{(1 - (\alpha \varepsilon_B/q)^2)^{\frac{1}{2}}}. \tag{2.12}$$

Thus, $\tan(\phi)$ and $\|\Delta M\|$ are both maximized for the angle ϕ satisfying $\sin(\phi) = \alpha\varepsilon_B/q$.

To identify $B \in \Gamma$ that minimizes q, and hence maximizes $\sin(\phi)$, we use Lemma 2.1 to get $q_{min} = \alpha(\alpha - \gamma)$. The maximal value of $\sin(\phi)$ is therefore $\varepsilon_B/(\alpha - \gamma)$. By using the definition (2.8) of α, and the expression (2.7) for $\|\Delta M\|$, we conclude

$$\sup_{B \in \Gamma} \sup_{\|\Delta B\| \leq \varepsilon_B} \|\Delta M\| = 2\sqrt{\left\{1 - \left(1 - \frac{\varepsilon_B^2}{((\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}} - \gamma)^2}\right)^{\frac{1}{2}}\right\}}. \tag{2.13}$$

For $m > 2$ we use the real Schur decomposition to rewrite the matrix $I + \Delta M$ as

$$I + \Delta M = W \operatorname{diag}(I_t, \Phi_1, \ldots, \Phi_p) W^T,$$

where W is orthogonal, $t + 2p = m$, and the matrices Φ_i are identified as

$$\Phi_i = \begin{bmatrix} \cos(\phi_i) & -\sin(\phi_i) \\ \sin(\phi_i) & \cos(\phi_i) \end{bmatrix}, \quad i = 1, \ldots p.$$

Let the matrices $G = W^T \Sigma$, $C = W^T B$, and $E = W^T \Delta B$ be partitioned in row blocks as

$$G^T = [G_0^T, \ldots, G_p^T], \quad C^T = [C_0^T, \ldots, C_p^T], \quad E^T = [E_0^T, \ldots, E_p^T],$$

$$G_0, C_0, E_0 \in R^{t \times m}, \quad G_i, C_i, E_i \in R^{2 \times m}, \quad i = 1, \ldots p.$$

Each one of the matrices Φ_i is then the solution to the 2-dimensional problem

$$\min_{\Phi_i \in \Omega_+} \|\Phi_i G_i - (C_i + E_i)\|. \tag{2.14}$$

We have already proved that (2.13) holds for 2-dimensional problems. Hence, we apply (2.13) on problem (2.14) to get

$$\|\Delta M\|^2 = \sum_{i=1}^{p} \|\Phi_i - I\|^2 \leq \sum_{i=1}^{p} 4\left(1 - \left(1 - \frac{\varepsilon_i^2}{(\sigma_1(G_i)^2 + \sigma_2(G_i)^{\frac{1}{2}} - \gamma_i)^2}\right)^{\frac{1}{2}}\right), \tag{2.15}$$

where $\varepsilon_i = \|C_i\|$ satisfies $\sum_i^p \varepsilon_i^2 \leq \varepsilon_B^2$ and $\gamma_i = \|G_i - C_i\|$ satisfies $\sum_i^p \gamma_i^2 \leq \gamma^2$. Let $i = k$ be the solution to

$$\min_{1 \leq i \leq p} (\sigma_1^2(G_i) + \sigma_2^2(G_i)).$$

Then $\|\Delta M\|^2$ in (2.15) is maximized when $\varepsilon_k = \varepsilon_B$, $\gamma_k = \gamma$ and $\varepsilon_i = 0$, $\gamma_i = 0$ for $i \neq k$. From the Mirsky Theorem (see e.g., [12] p. 204) we get

$$\sigma_1^2(G_k) + \sigma_2^2(G_k) \geq \sigma_{m-1}^2(G) + \sigma_m^2(G) = \sigma_{m-1}^2 + \sigma_m^2.$$

Hence, we conclude

$$\|\Delta M\| \leq 2\sqrt{\left\{1 - \left(1 - \frac{\varepsilon_B^2}{((\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}} - \gamma)^2}\right)^{\frac{1}{2}}\right\}}. \tag{2.16}$$

The upper bound is attained for

$$B = \Sigma - \operatorname{diag}(0, \ldots, 0, \sigma_{m-1}, \sigma_m)\gamma/(\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}, \tag{2.17}$$

and $\Delta B = \begin{bmatrix} 0 & 0 \\ 0 & \Delta B_{22} \end{bmatrix}$, where $\Delta B_{22} \in R^{2 \times 2}$ is chosen as

$$\Delta B_{22} = \begin{bmatrix} -\kappa_B \varepsilon_B \sigma_{m-1} & -(1 - \kappa_B^2 \varepsilon_B^2)^{\frac{1}{2}} \sigma_m \\ (1 - \kappa_B^2 \varepsilon_B^2)^{\frac{1}{2}} \sigma_{m-1} & -\kappa_B \varepsilon_B \sigma_m \end{bmatrix} \frac{\varepsilon_B}{(\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}}. \tag{2.18}$$

The theorem is now proved for the special case $\varepsilon_A = 0$.

To generalize the result to the case where also a perturbation ΔA is considered, we study the worst 2 by 2 perturbation ΔB, given by (2.10), from a geometrical point of view, see Fig. 1. We observe that θ equals ϕ_B. We also know from the solution of (2.11) and from the minimal value of q, that θ satisfies $\sin(\theta) = \varepsilon_B/(\alpha - \gamma)$. Analogously, the angles ω and ϕ_A corresponding to the worst perturbation ΔA, satisfy $\omega = \phi_A$ and $\sin(\omega) = \varepsilon_A/\alpha$, see Fig. 1.

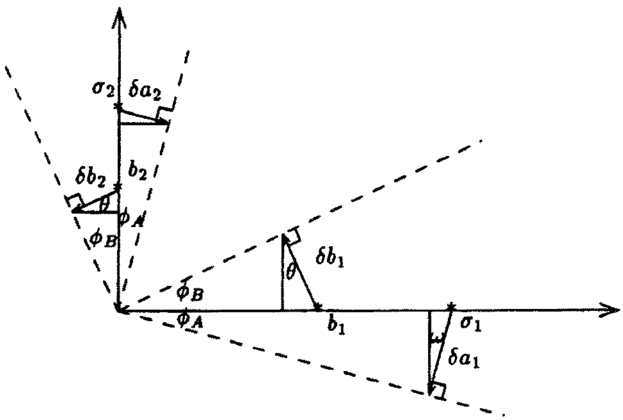

Fig. 1. A 2-D illustration of how the worst perturbations $\Delta B = [\delta b_1, \delta b_2]$ and $\Delta A = [\delta a_1, \delta a_2]$, affect the rotation matrix by making the rotation angle $\phi_A + \phi_B$ as large as possible.

The rotation angle, ϕ, corresponding to the perturbed rotation matrix equals $\phi_B + \phi_A$. Hence,

$$\cos(\phi) = \cos(\phi_B + \phi_A) = \cos(\phi_B)\cos(\phi_A) - \sin(\phi_B)\sin(\phi_A)$$

$$= (1 - \varepsilon_A^2/\alpha^2)^{\frac{1}{2}}(1 - \varepsilon_B^2/(\alpha - \gamma)^2)^{\frac{1}{2}} - \varepsilon_A\varepsilon_B/\alpha(\alpha - \gamma). \tag{2.19}$$

As in the case for $\varepsilon_A = 0$, we generalize the results to arbitrary dimension by replacing σ_1 and σ_2 with σ_{m-1} and σ_m, in the definition (2.8) of α. By doing so and inserting the expression (2.19) into the equation (2.7) we get the general result (2.1). ■

Let us now give some comments about Theorem 2.1. First, the sensitivity of the problem is determined by the condition numbers $\kappa_A = 1/(\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}$ and $\kappa_B = 1/((\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}} - \gamma)$. Hence, if the two smallest singular values of the matrix A are small, or if the residual is large, the problem is ill-conditioned.

Second, the worst matrix $B \in \Gamma$ defined in (2.17) has singular values that satisfy

$$1/(\sigma_{m-1}^2(B) + \sigma_m^2(B))^{\frac{1}{2}} = 1/((\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}} - \gamma) = \kappa_B.$$

Using this fact, the symmetry in the problem, with respect to A and B, becomes more evident.

Third, if we assume $\sigma_m > 0$ and restrict the residual, γ, and the allowed amounts of perturbations, ε_A and ε_B, to satisfy $\gamma < \sigma_m$, $\varepsilon_A < \sigma_m$, and $\varepsilon_B < \sigma_m - \gamma$, then the theorem also holds for the general orthogonal Procrustes problem (1.1) provided that the unperturbed solution has positive determinant. This is true because these restrictions imply that $\text{sign}(\det((B + \Delta B)(A + \Delta A)^T))$ equals $\text{sign}(\det(BA^T))$, i.e., the solution $\tilde{M}$ to the perturbed general problem has positive determinant.

Finally, we compare our first order bound (2.2) with the bound obtained by applying the results derived in [1] and [8] for the polar decomposition. The result for the orthogonal polar factor $M + \Delta M$ of a matrix $Z + \Delta Z$ is

$$\|\Delta M\| \leq \frac{2\|\Delta Z\|}{\sigma_{m-1}(Z) + \sigma_m(Z)} + O(\|\Delta Z\|^2). \tag{2.20}$$

In our application Z corresponds to BA^T and ΔZ to $B\Delta A^T + \Delta BA^T + \Delta B\Delta A^T$. Using the expression (2.17) for the matrix B that makes the problem most ill-conditioned we get the relations

$$\|\Delta Z\| \leq \sigma_1(\varepsilon_A + \varepsilon_B) + O((\varepsilon_A + \varepsilon_B)^2) \tag{2.21}$$

and

$$\sigma_{m-1}(Z) + \sigma_m(Z) = (\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}((\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}} - \gamma) = 1/(\kappa_A\kappa_B). \tag{2.22}$$

These relations inserted into (2.20) give

$$\|\Delta M\| \leq \sigma_1\kappa_A\kappa_B(\varepsilon_A + \varepsilon_B). \tag{2.23}$$

Thus, our result (2.2) is much sharper if $\sigma_1 \gg (\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}$ or if the residual is large.

3. The connection with the skew symmetric Procrustes problem.

Every matrix $(\hat{M} + \Delta M) \in \Omega_+$ can be represented by a skew symmetric matrix S as $(\hat{M} + \Delta M) = \exp(S) = e^S$, (see [3] p. 287). Using this representation, problem (1.3) becomes

$$\min_{S=-S^T} \|e^S(A + \Delta A) - (B + \Delta B)\|. \tag{3.1}$$

If we assume $\hat{M} = I$, $A = \Sigma$, and that the residual to the unperturbed problem is zero, i.e., $B = \Sigma$, the first order approximation of problem (3.1) is the linear *skew symmetric Procrustes problem*

$$\min_{S=-S^T} \|S(\Sigma + \Delta A) - (\Delta B - \Delta A)\|. \tag{3.2}$$

This problem is closely related to the *symmetric Procrustes problem* treated by Higham [7]. According to the theory by Higham for the symmetric case, we can show that the solution to (3.2) satisfies

$$\sup_{\Delta A,\Delta B} \|S\| = \frac{\sqrt{2}(\|\Delta A\| + \|\Delta B\|)}{(\tilde{\sigma}_{m-1}^2 + \tilde{\sigma}_m^2)^{\frac{1}{2}}}$$

$$(3.3) \qquad = \frac{\sqrt{2}(\|\Delta A\| + \|\Delta B\|)}{(\sigma_{m-1}^2 + \sigma_m^2)^{\frac{1}{2}}} + O((\|\Delta A\| + \|\Delta B\|)^2),$$

where $\tilde{\sigma}_{m-1}$ and $\tilde{\sigma}_m$ are the two smallest singular values of the matrix $\Sigma + \Delta A$. Since $\Delta M = S + O(\|S\|^2)$, we note that (3.3) is the same first order result for $\|\Delta M\|$ as we get by inserting $\gamma = 0$ into (2.2).

Acknowledgements.

I would like to thank Prof. Per-Åke Wedin for supporting me in writing this paper. Thanks also to the referees for their valuable comments on earlier versions of the paper and for encouraging me to extend Theorem 2.1 to include problems with nonzero residual.

REFERENCES

1. A. Barrlund. *Perturbation bounds on the polar decomposition.* BIT, 30:101–113, 1989.
2. C.-H. Chen and J.-G. Sun. *Perturbation bounds for the polar factors.* Journal of Computational Mathematics, 7:1–6, 1989.
3. F. R. Gantmacher. *The Theory of Matrices, Vol. I.* Chelsea Publishing Company, New York, 1977. Translation of the Russian-Language Book Teoriya Matrits.
4. B. F. Green. *The orthogonal approximation of an oblique structure in factor analysis.* Psychometrika, 17:429–440, 1952.
5. R. J. Hanson and M. J. Norris. *Analysis of measurements based on the singular value decomposition,* SIAM Journal of Scientific and Statistical Computing, 2:363–373, 1981.
6. N. J. Higham. *Computing the polar decomposition – with applications.* SIAM Journal of Scientific and Statistical Computing, 7:1160–1174, 1986.
7. N. J. Higham. *The symmetric Procrustes problem.* BIT, 28:133–143, 1988.
8. C. Kenney and A. J. Laub. *Polar decomposition and matrix sign function condition estimates.* SIAM Journal of Scientific and Statistical Computing, 12:488–504, 1991.
9. P. H. Schönmann. *A generalized solution of the orthogonal Procrustes problem.* Psychometrika, 31:1–10, 1966.
10. I. Söderkvist and P.-Å. Wedin. *Determining the movements of the skeleton using well-configured markers.* To appear in Journal of Biomechanics.
11. I. Söderkvist and P-Å. Wedin. *A practical condition number for a configuration of points defining a rigid body movement.* Technical Report UMINF-91.29, Institute of Information Processing, University of Umeå, 1991.
12. G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, Inc., London, 1990.
13. G. Wahba. *Problem* 65-1*: A least squares estimate of satellite attitude.* SIAM Review, 7:409, 1965. Solutions in 8:384–386, 1966.
14. Per-Åke Wedin. *Some thought on perturbation bounds for the polar decomposition and other perturbation problems in linear algebra. A first sketch.* Technical Report UMINF-91.30, Institute of Information Processing University of Umeå, 1991.