BIT 12 (1972), 99–111

PERTURBATION BOUNDS IN CONNECTION WITH SINGULAR VALUE DECOMPOSITION

PER-ÅKE WEDIN

Abstract.

Let A be an $m \times n$-matrix which is slightly perturbed. In this paper we will derive an estimate of how much the invariant subspaces of $A^H A$ and AA^H will then be affected. These bounds have the $\sin\theta$ theorem for Hermitian linear operators in Davis and Kahan [1] as a special case. They are applicable to computational solution of overdetermined systems of linear equations and especially cover the rank deficient case when the matrix is replaced by one of lower rank.

1. Preliminaries.

Let A be an $m \times n$-matrix over the complex field. Then there exists a singular value decomposition of A,

$$A = U \Sigma V^H = U_1 \Sigma_1 V_1^H + U_0 \Sigma_0 V_0^H \tag{1.1}$$

where

$$V_1 = [v_1,\ldots,v_r];\; V_0 = [v_{r+1},\ldots,v_p];\; V = (V_1, V_0)\,,$$

and

$$U_1 = [u_1,\ldots,u_r];\; U_0 = [u_{r+1},\ldots,u_p];\; U = (U_1, U_0);$$

$$\Sigma_1 = \operatorname{diag}(\sigma_1,\ldots,\sigma_r);\; \Sigma_0 = \operatorname{diag}(\sigma_{r+1},\ldots,\sigma_p);\; \Sigma = \operatorname{diag}(\sigma_1,\ldots,\sigma_p)$$

V_1, V_0, V and U_1, U_0, U are assumed to be partial isometries satisfying

$$V^H V = U^H U = I_p;\; V_1^H V_1 = U_1^H U_1 = I_r;\; V_0^H V_0 = U_0^H U_0 = I_{p-r}\,.$$

The rank of A is p and $r \leqq p$.

For the perturbation of A, $B = A+T$, a corresponding singular value decomposition can be made. Take

$$A_j = U_j(A)\,\Sigma_j(A)\,V_j^H(A);\quad B_j = U_j(B)\,\Sigma_j(B)\,V_j^H(B);\qquad j = 0,1\,. \tag{1.2}$$

From (1.1) it is obvious that

$$A+T = A_1+A_0+T = B_1+B_0 = B\,. \tag{1.3}$$

It seems natural to assume that $r = \operatorname{rank}(A_1) = \operatorname{rank}(B_1)$, although

Received November 16, 1971.

this condition is not needed for the generalization of the $\sin\theta$ theorem that we will prove later.

The range of a matrix A is denoted by $R(A)$ and the nullspace of A by $N(A)$. It is an immediate consequence of the decomposition (1.1) and the definition (1.2) that

$$R(A_1) = N(A_1^H)^\perp \quad \text{and} \quad N(A_1) = R(A_1^H)^\perp \tag{1.4}$$

where $\perp$ denotes the orthogonal complement.

$R(A_1)$ and $R(A_0)$ are invariant subspaces of the Hermitian matrix AA^H as are $R(A_1^H)$ and $R(A_0^H)$ of A^HA.

In this paper we are going to estimate the angles between the subspaces in E^m, $R(A_1)$ and $R(B_1)$ as well as between the subspaces in E^n, $R(A_1^H)$ and $R(B_1^H)$.

The orthogonal projection onto a subspace M is denoted by P_M. Angles between subspaces are studied by Davis and Kahan in [1]. The angle between a vector x and a subspace M can be defined by

$$\sin\theta(x,M) = \min_{y\in M} \|x-y\|_2$$

with $\|x\|_2 = 1$.

It follows from the projection theorem that

$$\min_{y\in M} \|x-y\|_2 = \|(I-P_M)x\|_2 .$$

For two subspaces L and M it is natural to define

$$\|\sin\theta(L,M)\| = \|(I-P_M)P_L\| \tag{1.5}$$

for any unitary invariant norm.

Our aim can now be formulated strictly. We want to find good upper bounds for

$$\|\sin\theta(R(B_1),R(A_1))\| \quad \text{and} \quad \|\sin\theta(R(B_1^H),R(A_1^H))\|$$

when we have estimates of $\|T\|$ and the gap between the least singular value of B_1 and the largest singular value of A_0.

We will now define residuals which can be used instead of T. Let $y_1,\ldots,y_r$ be orthonormal vectors spanning the subspace $R(B_1)$. This means that with $Y_1 = [y_1,\ldots,y_r]$,

$$Y_1^HY_1 = I_r \quad \text{and} \quad Y_1Y_1^H = P_{R(B_1)} . \tag{1.6}$$

Analogously define X_1 through

$$X_1^HX_1 = I_r \quad \text{and} \quad X_1X_1^H = P_{R(B_1^H)} . \tag{1.7}$$

Take $D_1 = Y_1^HBX_1$. A convenient but not necessary choice is $X_1 = V_1$

and $Y_1 = U_1$. With this choice we get $D_1 = \Sigma_1(B)$. We can now define the residuals

$$\begin{cases} R_{11} = AX_1 - Y_1 D_1 \\ R_{21} = A^H Y_1 - X_1 D_1^H . \end{cases} \tag{1.8}$$

The connection between R_{11} and T is seen through the rewriting $R_{11} = AX_1 - Y_1 D_1 = (B-T)X_1 - Y_1(Y_1^H B X_1) = -TX_1$. Analogously it can be shown that $R_{21} = -T^H Y_1$.

In the same way partial isometries Y_0 and X_0 corresponding to $R(B_0)$ and $R(B_0^H)$ could be defined.

We have now at hand almost all tools necessary to estimate $\|\sin\theta(R(B_1), R(A_1))\|$ and $\|\sin\theta(R(B_1^H), R(A_1^H))\|$. But that work will wait until paragrah 3 because we will first cite a theorem by Davis and Kahan about the perturbation of Hermitian operators. That theorem also shows how the gap-condition should be imposed on B_1 and A_0. In paragraph 3 we then make a rather natural generalization of Davis' and Kahan's theorem.

2. The $\sin\theta$ theorem for Hermitian matrices.

In this section it is assumed that A and B are Hermitian matrices. It follows that $R(A_1) = R(A_1^H)$ and $R(B_1) = R(B_1^H)$. We also choose $Y_1 = X_1$ and get as a consequence $R_{11} = R_{21} = R_1$. There are several estimates of trigonometric functions of the angles between B_1 and A_1 in [1]. This theorem which we intend to generalize below will be stated. The notations differ slightly from those used in [1], p. 10.

The $\sin\theta$ *theorem.* Assume there is an interval $[\beta, \alpha]$ and a $\delta > 0$ such that the spectrum of A_0 lies entirely in $[\beta, \alpha]$ while that of B_1 lies entirely outside of $(\beta - \delta, \alpha + \delta)$ (or such that the spectrum of B_1 lies entirely in $[\beta, \alpha]$, while that of A_0 lies entirely outside of $(\beta - \delta, \alpha + \delta)$). Then for every unitary invariant norm, $\delta \|\sin\theta(R(B_1), R(A_1))\| \leqq \|R_1\|$.

In the proof of the $\sin\theta$ theorem above a multiple of the identity operator is added to A, translating the spectra of B_0 and A_1 without affecting R. It can accordingly be assumed without loss of generality that $0 \leqq \alpha = -\beta$ in the hypothesis of the theorem. This choice of α and β is subsequently used in [1] (see the proof p. 25). The point is now that we cannot in general make any translation of the singular values of the $m \times n$-matrices A and B. It therefore seems natural to limit our generalization of the $\sin\theta$-theorem to the following formulation of the $\sin\theta$ theorem, which is equivalent for Hermitian matrices.

Adaptable formulation of the $\sin\theta$ *theorem.* Let A and B be Hermitian matrices. Assume there exists an $\alpha \geqq 0$ and a $\delta > 0$ such that

$$\sigma_{\min}(B_1) \geqq \alpha+\delta \quad \text{and} \quad \sigma_{\max}(A_0) \leqq \alpha\,,$$

then for every unitary invariant norm,

$$\delta\|\sin\theta(R(B_1),R(A_1))\| \leqq \|R_1\|\,.$$

3. The $\sin\theta$ theorem for the singular value decomposition.

We are now ready to formulate a generalization of the $\sin\theta$ theorem to $m\times n$-matrices.

The generalized $\sin\theta$ *theorem.* Assume there exists an $\alpha \geqq 0$ and a $\delta > 0$ such that

$$\sigma_{\min}(B_1) \geqq \alpha+\delta \quad \text{and} \quad \sigma_{\max}(A_0) \leqq \alpha\,.$$

Take $\varepsilon = \max(\|R_{11}\|,\|R_{21}\|)$ where R_{11} and R_{21} are defined by (1.8). Then for every unitary invariant norm,

$$(3.1)\qquad \begin{cases} \|\sin\theta(R(B_1),R(A_1))\| \leqq \dfrac{\varepsilon}{\delta} \\[2ex] \|\sin\theta(R(B_1{}^H),R(A_1{}^H))\| \leqq \dfrac{\varepsilon}{\delta}. \end{cases}$$

Proof. We are going to use the same notations as in section 1. From the definition (1.5) it is seen that

$$(3.2)\qquad \|\sin\theta(R(B_1),R(A_1))\| = \|P_{R(A_1)^\perp}P_{R(B_1)}\|$$

and

$$(3.3)\qquad \|\sin\theta(R(B_1{}^H),R(A_1{}^H))\| = \|P_{R(A_1{}^H)^\perp}P_{R(B_1{}^H)}\|\,.$$

It is also known from section 1 that

$$(3.4)\qquad \|R_{11}\| = \|TX_1\| = \|TP_{R(B_1{}^H)}\|$$

and that

$$(3.5)\qquad \|R_{21}\| = \|T^HY_1\| = \|P_{R(B_1)}T\|\,.$$

The pseudo-inverse of a matrix A is denoted by A^+. In the following we will make ardent use of the identities,

$$(3.6)\qquad P_{R(A)} = AA^+ = A^{+H}A^H;\ P_{R(A^H)} = A^+A = A^HA^{+H}\,.$$

The proof now depends on two decompositions,

$$(3.7)\qquad \begin{aligned} P_{R(A_1)^\perp}P_{R(B_1)} &= P_{R(A_1)^\perp}B_1B_1{}^+ = P_{R(A_1)^\perp}(A_1+A_0-B_0-T)B_1{}^+ \\ &= \{-P_{R(A_1)^\perp}TP_{R(B_1{}^H)}+A_0(P_{R(A_1{}^H)^\perp}P_{R(B_1{}^H)})\}B_1{}^+ \end{aligned}$$

and the corresponding identity

$$(3.8) \qquad P_{R(B_1{}^H)}P_{R(A_1{}^H)^\perp} = B_1{}^+\{-P_{R(B_1)}TP_{R(A_1{}^H)^\perp}+P_{R(B_1)}P_{R(A_1)^\perp}A_0\}\,.$$

Take $\mu=\max(\|P_{R(A_1)^\perp}TP_{R(B_1{}^H)}\|,\|P_{R(B_1)}TP_{R(A_1{}^H)^\perp}\|)$. From (3.4) and (3.5) it is seen that

$$(3.9) \qquad \mu \leqq \max(\|R_{11}\|,\|R_{21}\|) = \varepsilon\,.$$

We will temporarily use the notations

$$t_1 = \|\sin\theta(R(B_1),R(A_1))\| = \|P_{R(A_1)^\perp}P_{R(B_1)}\|$$

and

$$t_2 = \|\sin\theta(R(B_1{}^H),R(A_1{}^H))\| = \|P_{R(A_1{}^H)^\perp}P_{R(B_1{}^H)}\|\,.$$

It was assumed that $\sigma_{\max}(A_0)=\|A_0\|_2\leqq\alpha$ and that $1/\|B_1{}^+\|_2=\sigma_{\min}(B_1)\geqq\alpha+\delta$. Now it is known from [3] that for unitary norms and arbitrary matrices C and D such that CD is defined

$$\|CD\| \leqq \|C\|_2\|D\| \leqq \|C\|\cdot\|D\|\,.$$

Then from the decompositions (3.7) and (3.8) we derive the inequalities

$$(3.10) \qquad \begin{cases} t_1 \leqq \dfrac{\mu+\alpha t_2}{\alpha+\delta} \\[2ex] t_2 \leqq \dfrac{\mu+\alpha t_1}{\alpha+\delta}. \end{cases}$$

Assume that t_j is $\max(t_1,t_2)$. (3.10) implies that

$$t_j \leqq \frac{\mu+\alpha t_j}{\alpha+\delta}$$

or simplified that

$$(3.11) \qquad t_j \leqq \mu/\delta\,.$$

Now we take inequality (3.9) into account and insert the notations used originally. Then formula (3.11) implies the inequalities (3.1) which we set out to prove.

If A and B are Hermitian, then the theorem above is identical with the $\sin\theta$ theorem in section 2.

As distinguished from Davis and Kahan [1] we work with subspaces $R(A_0)$ and $R(A_1)$ which are orthogonal but not orthogonal complements since $R(A_0)+R(A_1)=R(A)$. Because of this we cannot state a $\sin\theta$ theorem with the residuals corresponding to A_0 and B_0. But it is easy to make a small change of the definition of X_0 and Y_0 to make a similar theorem possible. Let $y_{r+1},\ldots,y_m$ be orthonormal vectors spanning

$R(B_1)^{\perp}$ and $x_{r+1},\ldots,x_n$ be orthonormal and spanning $N(B_1)$. Take $Y_0=[y_{r+1},\ldots,y_m]$, $X_0=[x_{r+1},\ldots,x_n]$ and $D_0=Y_0^H B X_0$. Define

$$R_{10} = AX_0 - Y_0 D_0;\; R_{20} = A^H Y_0 - X_0 D_0^H .$$

As in section 1 it is seen that $R_{10}=-TX_0$ and that $R_{20}=-T^H Y_0$. For unitary invariant norms we get

$$\|R_{10}\| = \|TP_{R(B_1^H)^{\perp}}\|;\quad \|R_{20}\| = \|P_{R(B_1)^{\perp}}T\| .$$

The $\sin\theta$*-theorem with complementary residuals.* Assume there exists an $\alpha \geqq 0$ and a $\delta > 0$ such that

$$\sigma_{\min}(A_1) \geqq \alpha+\delta \quad \text{and} \quad \sigma_{\max}(B_0) \leqq \alpha .$$

Take $\varepsilon=\max(\|R_{10}\|,\|R_{20}\|)$. Then for every unitary invariant norm,

$$(3.12)\qquad \begin{cases} \|\sin\theta(R(A_1),R(B_1))\| \leqq \dfrac{\varepsilon}{\delta} \\[2ex] \|\sin\theta(R(A_1^H),R(B_1^H))\| \leqq \dfrac{\varepsilon}{\delta}. \end{cases}$$

The proof is similar to that already given in this section.

NOTE 1. When A is perturbed, $P_{R(A_0)}$ and $P_{R(A_0^H)}$ are influenced not only by $P_{R(A_1)}$ and $P_{R(A_1^H)}$ but also by $P_{R(A)^{\perp}}$ and $P_{N(A)}$. That is the reason why it is difficult to get estimates of

$$(3.13)\qquad \|\sin\theta(R(B_0),R(A_0))\| \quad \text{and} \quad \|\sin\theta(R(B_0^H),R(A_0^H))\| .$$

If $N(A)=N(B)=\{0\}$ then

$$\begin{aligned} \|\sin\theta(R(A_1^H),R(B_1^H))\| &= \|P_{N(B_1)}P_{N(A_1)^{\perp}}\| = \|P_{N(B_0)^{\perp}}P_{N(A_0)}\| \\ &= \|P_{N(A_0)}P_{N(B_0)^{\perp}}\| = \|\sin\theta(R(B_0^H),R(A_0^H))\| \end{aligned}$$

and we can use (3.12). If $N(A)$ or $N(B)$ is nonempty we need a lower bound of $\sigma_{\min}(B_0)$ in the theorem above to be able to estimate $\|\sin\theta(R(B_0^H),R(A_0^H))\|$. About such problems see the theorem in section 5.

NOTE 2. For unitary invariant norms

$$(3.14)\qquad \|\sin\theta(R(A_1),R(B_1))\| = \|\sin\theta(R(B_1),R(A_1))\|$$

if and only if

$$(3.15)\qquad \operatorname{rank}(A_1) = \operatorname{rank}(B_1) .$$

The proof of this theorem which is given in several papers (see [5]) is based on the following two facts.

1. The singular values less than one are the same for $(I-P_{R(B_1)})P_{R(A_1)}$ and $(I-P_{R(A_1)})P_{R(B_1)}$.
2. A unitary invariant norm of a matrix depends only on the singular values of that matrix.

NOTE 3. For the spectral norm the generalized $\sin\theta$ theorem follows from the $\sin\theta$ theorem for hermitian matrices. To every $m\times n$-matrix C there is defined an $(m+n)\times(m+n)$-matrix

$$\tilde{C} = \begin{pmatrix} 0 & C \\ C^H & 0 \end{pmatrix}.$$

$\tilde{C}$ is hermitian and has the eigenvalues $\pm\sigma_1(C),\ \pm\sigma_2(C),\ldots,\pm\sigma_p(C)$ and $m+n-2p$ eigenvalues equal to zero. It is observed that with the notations from (1.6) and (1.7)

$$Z_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} Y_1 & -Y_1 \\ X_1 & X_1 \end{pmatrix}$$

is a partial isometry corresponding to $R(\tilde{B}_1)$ because $Z_1{}^HZ_1=I$ and $Z_1Z_1{}^H=P_{R(\tilde{B}_1)}$. Take

$$\hat{D}_1 = Z_1{}^H\tilde{B}_1Z_1 \quad\text{and}\quad \hat{R}_1 = \tilde{A}Z_1-Z_1\hat{D}_1\,.$$

From the $\sin\theta$ theorem for hermitian matrices it follows that if

$$\sigma_{\min}(\tilde{B}_1) \geqq \alpha+\delta \quad\text{and}\quad \sigma_{\max}(\tilde{A}_0) \leqq \alpha$$

then

(3.16) $$\delta\cdot\|\sin\theta(R(\tilde{B}_1),R(\tilde{A}_1))\| \leqq \|\hat{R}_1\|\,.$$

But

$$\begin{aligned}\|\sin\theta(R(\tilde{B}_1),R(\tilde{A}_1))\|_2 &= \|(I-P_{R(\tilde{A}_1)})P_{R(\tilde{B}_1)}\|_2 = \\ &= \left\|\begin{pmatrix}(I-P_{R(A_1)})P_{R(B_1)} & 0 \\ 0 & I-P_{R(A_1{}^H)}P_{R(B_1{}^H)}\end{pmatrix}\right\|_2 = \\ &= \max\left(\|\sin\theta(R(B_1),R(A_1))\|_2, \|\sin\theta(R(B_1{}^H),R(A_1{}^H))\|_2\right)\,.\end{aligned}$$

It is further seen that

$$\hat{R}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} R_{11} & R_{11} \\ R_{21} & -R_{21}\end{pmatrix},$$

where R_{11} and R_{21} are given in (1.8), and that the singular values of $\hat{R}_1$ are the singular values of R_{11} and R_{21}. Hence $\|\hat{R}_1\|_2=\max(\|R_{11}\|_2,\|R_{21}\|_2)$.

$\sigma_{\min}(\tilde{B}_1)=\sigma_{\min}(B_1)$ and $\sigma_{\max}(\tilde{A}_0)=\sigma_{\max}(A_0)$ and so (3.16) gives the generalized $\sin\theta$ theorem for the spectral norm.

NOTE 4. The estimates derived from (3.7) and (3.8) can be sharpened if we take

$$\mu_1 = \|P_{R(A_1)^\perp} T P_{R(B_1{}^H)}\| \quad \text{and} \quad \mu_2 = \|P_{R(B_1)} T P_{R(A_1{}^H)^\perp}\| \,.$$

If a counterpart of (3.10) with μ_1 and μ_2 instead of μ is used we get after some work that

$$t_1 \leqq \frac{\mu_1 + \dfrac{\alpha}{\alpha+\delta}\mu_2}{\left(1+\dfrac{\alpha}{\alpha+\delta}\right)} \leqq \frac{\mu}{\delta}$$

(3.17) and

$$t_2 \leqq \frac{\mu_2 + \dfrac{\alpha}{\alpha+\delta}\mu_1}{\left(1+\dfrac{\alpha}{\alpha+\delta}\right)} \leqq \frac{\mu}{\delta}.$$

Here $\mu_1 \leqq \|R_{11}\|$ and $\mu_2 \leqq \|R_{22}\|$. If $\alpha/(\alpha+\delta)$ is small we get from (3.17) that

$$\|\sin\theta(R(B_1), R(A_1))\| \lesssim \frac{\|R_{11}\|}{\delta}$$

and

$$\|\sin\theta(R(B_1{}^H), R(A_1{}^H))\| \lesssim \frac{\|R_{21}\|}{\delta}.$$

4. Applications to the estimation of $\|B_1{}^+ - A_1{}^+\|$.

Throughout section 4 it will be assumed that

$$\text{rank}(A_1) = \text{rank}(B_1) \tag{4.1}$$

so that we can take advantage of (3.14). In this context it is suitable to use that $\|\sin\theta(L,M)\| = \|(I-P_M)P_L\|$ and we can formulate the result in note 2 in the following lemma.

LEMMA. *If and only if* $\text{rank}(A_1)=\text{rank}(B_1)$ *then*

$$\|P_{R(A_1)^\perp} P_{R(B_1)}\| = \|P_{R(B_1)^\perp} P_{R(A_1)}\| \tag{4.2}$$

and

$$\|P_{N(A_1)} P_{N(B_1)^\perp}\| = \|P_{N(B_1)} P_{N(A_1)^\perp}\| \tag{4.3}$$

for unitary invariant norms.

If we want to estimate $\|B_1^+ - A_1^+\|$ using the theorems in [5] or [6] an estimate of $\|B_1 - A_1\|$ is needed. Let us assume that A and B satisfy the conditions of the $\sin\theta$ theorem. Expressed in the notations from that theorem it is then seen that an estimate of $\|B_1 - A_1\|$ involves $1/\delta$ and $\|B-A\|$. In fact we get

$$\begin{aligned} B_1 - A_1 &= P_{R(B_1)}(B-A)P_{R(A_1^H)} + B_1 P_{N(A_1)} - P_{R(B_1)^\perp} A_1 \\ B_1 P_{N(A_1)} &= P_{R(B_1)} T P_{N(A_1)} + P_{R(B_1)} A_0 \\ P_{R(B_1)^\perp} A_1 &= -P_{R(B_1)^\perp} T P_{R(A_1^H)} + B_0 P_{R(A_1^H)} \,. \end{aligned}$$

From these identities and the $\sin\theta$ theorem it is seen that

$$\|B_1 - A_1\| \leqq \|T\| \left(3 + \frac{\|A_0\|}{\delta} + \frac{\|B_0\|}{\delta}\right). \tag{4.4}$$

Even if (4.4) is an overestimate we cannot find a bound which is essentially sharper. In (4.4) we have used that $\|T\|$ is greater than or equal to the number ε in (3.1).

The point is now that we do not need an estimate of $\|B_1 - A_1\|$ to be able to bound $\|B_1^+ - A_1^+\|$. We use the decomposition from [5],

$$B_1^+ - A_1^+ = -A_1^+(B_1 - A_1)B_1^+ - A_1^+ P_{R(B_1)^\perp} + P_{N(A_1)} B_1^+ \tag{4.5}$$

and note that

$$A_1^+(B_1 - A_1)B_1^+ = A_1^+(B-A)B_1^+ = A_1^+ T B_1^+ \,.$$

Hence

$$\|B_1^+ - A_1^+\| \leqq \|T\| \left(\|B_1^+\| \cdot \|A_1^+\| + \frac{\|B_1^+\|}{\delta} + \frac{\|A_1^+\|}{\delta}\right). \tag{4.6}$$

Bounds of $\|B_1^+ - A_1^+\|$ are useful in at least two kinds of problems.

First there is the problem studied in [7] where we want to see how large the perturbation $(A+T)^+b - A^+b$ is in different directions. Evidently $\|B_0^+ - A_0^+\|$ can be much larger than $\|B_1^+ - A_1^+\|$. However that problem is more easily dealt with if a decomposition similar to (4.5) is used for $B^+ - A^+$. If A and B are nonsingular we get $B^{-1} - A^{-1} = -(A_1^+ + A_0^+)T(B_1^+ + B_0^+)$ which in this context is more useful than (4.6).

There are several problems in which we want to determine the pseudoinverse A_1^+ of the matrix A_1, which minimizes $\|A_1 - A\|$ when $\operatorname{rank}(A_1) = r$ (see [3]). In some of these problems the rank of A_1 is not known beforehand but chosen in such a way that $\|A_1 - A\|$ is small but $\sigma_r(A_1)$ is not too small. In general during the computational procedure a matrix $A+T$ is found for which $\|T\|$ is small and $(A+T)^+$ is taken as an approximation of A_1^+. (4.6) can be used directly for this problem. But a sharper estimate of $\|(A+T)^+-A_1^+\|$ can be found if we take advantage of the fact that $B_1=(A+T)=B$. We know that

$$(4.7) \qquad P_{N(A_1)^\perp}P_{N(B_1)} = A_1^+(-T-A_0)P_{N(B_1)} = -A_1^+TP_{N(B_1)}$$

and that

$$(4.8) \qquad P_{R(B_1)^\perp}P_{R(A_1)} = P_{R(B_1)^\perp}(-T-A_0)A_1^+ = -P_{R(B_1)^\perp}TA_1^+ .$$

From (4.5) it follows that

$$\|B_1^+-A_1^+\| \leqq \|A_1^+\|\cdot\|T\|\cdot\|B_1^+\|+\|A_1^+\|\cdot\|P_{R(A_1)}P_{R(B_1)^\perp}\| + \|B_1^+\|\cdot\|P_{N(A_1)}P_{N(B_1)^\perp}\|$$

and since the lemma implies that

$$\|P_{N(A_1)}P_{N(B_1)^\perp}\| = \|P_{N(A_1)^\perp}P_{N(B_1)}\|$$

we can use (4.7) and (4.8) to get the estimate

$$(4.9) \qquad \|B_1^+-A_1^+\| \leqq 3\|T\|\cdot\|A_1^+\|\max(\|A_1^+\|,\|B_1^+\|) .$$

The constant 3 in (4.9) can be changed to $(\sqrt{5}+1)/2$ for the spectral norm and to $\sqrt{2}$ for the Euclidean matrix norm. The technique necessary for that improvement is given in [6], section 6.

5. An extension of the original $\sin\theta$ theorem.

A_j and B_j, $j=0$, 1 are still defined as in section 1. We return to the definition of X_0 and Y_0 as partial isometries corresponding to $R(B_0)$ and $R(B_0^H)$ that is

$$(5.1) \qquad \begin{aligned} P_{R(B_0^H)} &= X_0X_0^H; \quad I = X_0^HX_0 \\ P_{R(B_0)} &= Y_0Y_0^H; \quad I = Y_0^HY_0 \end{aligned}$$

and take $D_0=Y_0^HBX_0$. As before we define the residuals

$$(5.2) \qquad R_{01} = AX_0-Y_0D_0; \quad R_{02} = A^HY_0-X_0D_0^H$$

and

$$(5.3) \qquad \varepsilon = \max(\|R_{01}\|,\|R_{02}\|) .$$

Theorem (extension of the original $\sin\theta$*-theorem).*

Assume there is an interval $[\beta,\alpha]$ and a $\delta>0$ such that the singular values of B_0 lie entirely in $[\beta,\alpha]$ while the singular values of A_1 lie entirely outside of $(\beta-\delta,\alpha+\delta)$ (or such that the singular values of A_1 lie entirely in $[\beta,\alpha]$ while those of B_0 lie entirely outside of $(\beta-\delta,\alpha+\delta)$). Then for every unitary invariant norm

$$\max\left(\|\sin\theta(R(B_0),R(A_0))\|,\|\sin\theta(R(B_0^H),R(A_0^H))\|\right) \leqq \frac{\varepsilon}{\delta}\cdot k \tag{5.4}$$

with

$k = \sqrt{2}$ for the spectral norm and the Euclidean matrix norm
$k \leqq 2$ for all unitary invariant norms.

PROOF. As was mentioned in section 2 we cannot make any translation of the singular values of A and B. Instead we decompose

$$A_1 = A_2+A_3 \tag{5.5}$$

with $A_2^HA_3=0$ and $A_3A_2^H=0$. Then A_2 and A_3 have the same singular values as A_1. We assume that all singular values of A_1 which are less than or equal to $\beta-\delta$ belong to A_2 and those which are greater than or equal to $\alpha+\delta$ belong to A_3.

By definition

$$\|\sin\theta(R(B_0),R(A_0))\| = \|(I-P_{R(A_0)})P_{R(B_0)}\| . \tag{5.6}$$

We make the decomposition

$$\begin{aligned}(I-P_{R(A_0)})P_{R(B_0)} &= (P_{R(A)^\perp}+P_{R(A_2)}+P_{R(A_3)})P_{R(B_0)} = \\ &= (P_{R(A)^\perp}+P_{R(A_2)})P_{R(B_0)}+P_{R(A_3)}P_{R(B_0)} .\end{aligned} \tag{5.7}$$

An analogous decomposition can be made for $(I-P_{R(A_0^H)})P_{R(B_0^H)}$. We will now prove that

$$\|(P_{R(A)^\perp}+P_{R(A_2)})P_{R(B_0)}\| \leqq \frac{\varepsilon}{\delta} \tag{5.8}$$

and that

$$\|P_{R(A_3)}P_{R(B_0)}\| \leqq \frac{\varepsilon}{\delta}. \tag{5.9}$$

Let us for a moment assume that proof clear. Because $P_{R(A)^\perp}+P_{R(A_2)}$ is orthogonal to $P_{R(A_3)}$ it follows that

$$\|(I-P_{R(A_0)})P_{R(B_0)}\|_2 = \sup_{\|x\|=1}\|(I-P_{R(A_0)})P_{R(B_0)}x\|_2 \leqq \sqrt{2}\,\frac{\varepsilon}{\delta}$$

and that

$$\|(I-P_{R(A_0)})P_{R(B_0)}\|_E^2 = \|(P_{R(A)^\perp}+P_{R(A_2)})P_{R(B_0)}\|_E^2+\|P_{R(A_3)}P_{R(B_0)}\|_E^2 \leqq 2\frac{\varepsilon^2}{\delta^2},$$

$$\|(I-P_{R(A_0)})P_{R(B_0)}\|_E \leqq \sqrt{2}\,\frac{\varepsilon}{\delta}.$$

The inequality $\|(I-P_{R(A_0)})P_{R(B_0)}\| \leqq 2(\varepsilon/\delta)$ follows for all unitary invariant norms from the triangular inequality and (5.8) and (5.9).

We now turn to the proof of (5.9) which we rewrite as in (3.8):

$$\begin{aligned}(5.10)\qquad P_{R(B_0)}P_{R(A_3)} &= P_{R(B_0)}AA_3^+ = P_{R(B_0)}(B-T)A_3^+ \\ &= B_0(P_{R(B_0^H)}P_{R(A_3^H)})A_3^+ - P_{R(B_0)}TP_{R(A_3^H)}A_3^+ .\end{aligned}$$

A corresponding decomposition can be made in E^n of $P_{R(B_0^H)}P_{R(A_3^H)}$. It is also observed that $\|A_3^+\| \leqq 1/(\alpha+\delta)$ and that $\|B_0\| \leqq \alpha$. The proof of the $\sin\theta$ theorem in section 3 can then be copied to give (5.9) and at the same time

$$\|P_{R(A_3^H)}P_{R(B_0^H)}\| \leqq \frac{\varepsilon}{\delta}.$$

The proof of (5.8) depends on the identity,

$$\begin{aligned}(5.11)\qquad (P_{R(A)^\perp}+P_{R(A_2)})P_{R(B_0)} &= (P_{R(A)^\perp}+P_{R(A_2)})BB_0^+ \\ &= (P_{R(A)^\perp}+P_{R(A_2)})(A+T)B_0^+ \\ &= \{A_2(P_{R(A_2^H)}+P_{R(A^H)^\perp})P_{R(B_0^H)} + (P_{R(A)^\perp}+P_{R(A_2)})TP_{R(B_0^H)}\}B_0^+ .\end{aligned}$$

This decomposition is similar to (3.7). Since $\|B_0^+\| \leqq 1/\beta$ and $\|A_2\| \leqq \beta-\delta$ we can use the same technique of proof as in section 3 to get (5.8). That (5.8) and (5.9) give (5.4) has already been shown and hence the theorem is proved.

NOTE. It is not likely that the constant k in (5.4) can be chosen smaller than $\sqrt{2}$. Taking Hermitian matrices A and B, then B_0 might have eigenvalues both in $[-\alpha,-\beta]$ and $[\beta,\alpha]$. On the other hand it seems probable that the theorem can be sharpened so that $k=\sqrt{2}$ can be chosen for all unitary invariant norms.

REFERENCES

1. Ch. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation.* III, SIAM J. Numer. Anal. 7 (1970), 1–46.
2. T. Kato, *Perturbation theory for linear operators*, Springer, Berlin and New York, 1966.
3. L. Mirsky, *Symmetric gauge functions and unitarily invariant norms*, Quart, J, Math. Oxford, Ser. 2, 11 (1960), 50–59.
4. A. Ruhe, *Perturbation bounds for means of eigenvalues and invariant subspaces*, BIT 10 (1970), 343–354.
5. P.–Å. Wedin, *Perturbation theory for pseudo-inverses*, to appear.

6. P.–Å. Wedin, *On pseudoinverses of perturbed matrices*, Lund Un. Comp. Sc. Tech. Rep. 1969.
7. M. K. Gavurin, *Ill–conditioned systems of linear algebraic equations*, Zh. Vychisl. Mat. i Mat. Fiz., 2 (1962), 387–397.

DEPARTMENT OF COMPUTER SCIENCES
LUND UNIVERSITY
SÖLVEGATAN 14A
S-22362 LUND, SWEDEN